Kenn Gerdes *Editor*

Prokaryotic Toxin-Antitoxins

Springer

Prokaryotic Toxin-Antitoxins

Kenn Gerdes
Editor

Prokaryotic Toxin-Antitoxins

Springer

Editor
Kenn Gerdes
Centre for Bacterial Cell Biology
Institute for Cell and Molecular Bioscience
Newcastle University
Newcastle upon Tyne
UK

ISBN 978-3-642-33252-4 ISBN 978-3-642-33253-1 (eBook)
DOI 10.1007/978-3-642-33253-1
Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2012950379

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

This monograph on Prokaryotic Toxin–Antitoxins (TAs), the first of its kind, is meant to give an overview of a complex and rapidly developing scientific field. It will hopefully be useful not only as an introduction to the topic for newcomers to the field but also as a tool for established researchers to gain a comprehensive overview of the field. I've made an attempt to cover almost all aspects of Toxin–Antitoxins by inviting leading research groups to summarize and discuss the present stage of their respective field, resulting in a collection of 18 excellent Chapters that summarize the latest progress in the TA field. Initially, it seemed a daunting task to coordinate this monograph but almost all researchers that I approached with an invitation responded positively—and they were all able to produce their chapters with excellence—I'm very happy and grateful to the Authors that the progress with the monograph went so smoothly. I'd also like to apologize to the research groups that were not invited to contribute due to space constraints.

The monograph makes an attempt to give a timely and systematic overview of the TA field, which is a very complex one. The complexity originates from several facts: (i) the bewildering number of TA loci in single organisms; (ii) the very mild or almost absent phenotypes of single TA gene deletions; (iii) different research labs using different strains. These facts may have (mis)led different research groups to interpret similar observations differently thus adding to the complexity within the field. From the outside scientist, this complexity may look like confusion and, to my view, is one of the best arguments to produce a serene and unbiased monograph that enables the readers to make a judgement themselves. The Introduction (Chap. 1) tries to give an organized definition, overview, and discussion of this multifaceted problem before the individual TA gene families are presented and discussed.

2nd, August 2012 Kenn Gerdes

Contents

Chapter 1
Introduction

Kenn Gerdes

Abstract In the last decade, the field of "Prokaryotic Toxin—Antitoxins" has developed into an exciting branch of molecular microbiology, with many basic discoveries that have increased our understanding of these genes and that may lead to important breakthroughs in medical microbiology and biotechnology.

1.1 The Mystifying Numbers of Prokaryotic Toxin-Antitoxin Loci

In the last decade, the field of "Prokaryotic Toxin—Antitoxins" has developed into an exciting branch of molecular microbiology, with many basic discoveries that have increased our understanding of these genes and that may lead to important breakthroughs in medical microbiology and biotechnology. The first TA loci were discovered in the 1980s (Ogura and Hiraga 1983; Gerdes et al. 1985) and hints of the molecular mechanisms behind the ability of these genes to increase plasmid maintenance (Gerdes et al. 1986; Hiraga et al. 1986) instigated further detailed molecular analyses to obtain a deeper understanding of how these peculiar genes are regulated (Van Melderen et al. 1994; Gerdes et al. 1997). However, the vast expansion in the number of sequenced prokaryotic genomes in the last decade has led to an even greater expansion of the number of known genes and gene families. Unsurprisingly, this has led to a parallel large, confusing expansion of the number of putative TA genes in prokaryotic genomes. For microbiologists outside

K. Gerdes (✉)
Centre for Bacterial Cell Biology, Institute for Cell and Molecular Biosciences, Newcastle University, Richardson Road, Newcastle Upon Tyne, NE2 4AX, UK
e-mail: kenn.gerdes@ncl.ac.uk

K. Gerdes (ed.), *Prokaryotic Toxin-Antitoxins*,
DOI: 10.1007/978-3-642-33253-1_1,

the TA field, this increase may perhaps seed the feeling that almost any gene pair in which one partner without the concomitant presence of the other is toxic to the host cell can be classified as a TA pair. If so, the term "TA pair" would not be very useful. It is an important aim of this first monograph on prokaryotic toxin-antitoxins to give a systematic overview of this exciting and rapidly developing field. Several reviews in the literature give an introduction to TA loci (Buts et al. 2005; Gerdes et al. 2005; Condon 2006; Gerdes and Wagner 2007; Yamaguchi et al. 2011; Yamaguchi and Inouye 2011; Hayes and Van Melderen 2011; Arcus et al. 2011; Blower et al. 2011b) and Chap. 14 herein describes the TA database (http://bioinfo-mml.sjtu.edu.cn/TADB/) that gives access to an instant, automatic annotation of TA loci in many different prokaryotic organisms (Shao et al. 2011).

1.2 The Three Major Classes of TA Loci

Based on the nature and mechanism of action of a given antitoxin, Finbarr Hayes suggested to divide TA loci into different types (Hayes 2003) and his notation has been adopted by the literature. In **type I** TA loci, toxin expression is controlled by antisense RNAs that inhibit the translation of a toxin encoding mRNA (Gerdes and Wagner 2007). Paradigm and novel type I TA loci are treated in Chaps. 2 (Weaver lab) and 3 (Fozo lab), respectively. In **type II** loci, the antitoxin is a protein that combines with and neutralizes the toxin. A major part of this monograph is devoted to the description of type II loci because of the large number of evolutionary independent type II gene families that are known (Chaps. 4–14 and 16–18). As described in Chap. 15, **type III** loci were discovered only recently by George Salmond's group (Fineran et al. 2009). Here, the antitoxin is a small RNA that combines with and neutralizes the toxin by direct protein—RNA contact. Interestingly, type III toxins are RNases that are structurally similar to MazF, a canonical type II toxin (Blower et al. 2011a). Thus, type III loci may have evolved from an ancestral *mazEF* locus by replacing the antitoxin with a small RNA that binds to and neutralizes the toxin.

1.3 Definition of TA Pairs

To create a general definition of TA pairs is not straightforward. For instance, in some cases overproduction of one protein of a pair of proteins that function together in an essential cellular process is detrimental and inhibits cell growth. One such example was described previously (Allen and Kornberg 1991). DnaB and DnaC proteins interact to fulfill their functions in DNA replication. Overproduction of DnaC is toxic because it inhibits the essential helicase activity of DnaB. By contrast, the simultaneous overproduction of DnaC and DnaB relieved DnaC's inhibition of replication and therefore also abrogated the toxicity of DnaC.

Obviously, we do not call DnaBC a TA pair because toxin activity of a given TA pair should be the major biological function. So, is it possible to generate a useful, operational definition of what should be called a TA locus? I believe that by setting up a number of common properties derived from what is known about gene pairs already accepted to constitute TA pairs, it will in many cases be possible to determine if a given gene pair under scrutiny should be called a TA pair. However, there will certainly also be cases in which this will not be possible. To avoid confusion in the field, we should be restrictive and use such terms as "putative TA pair" or "potential TA pair" if the picture is not entirely clear. This leaves us with the easier task to list the properties of "canonical" and "bona fide" TA pairs.

1.3.1 Canonical TA Pairs

In particular for type II TA loci, it is possible to list a number of common traits that apply to most of the members belonging to the gene families *ccdAB*, *phd/doc*, *relBE* (including *higBA* and *mqsrA* families), *mazEF*, *vapBC*, *parDE* and probably others.

1. Two adjacent genes, one encoding a stable toxin and the other encoding a relatively unstable antitoxin.
2. Auto-repression of TA operon transcription by the antitoxin and co-repression of transcription by the toxin.
3. Control of TA operon transcription by conditional cooperativity (derepression by excess toxin).
4. Broad phylogeny based on toxin sequence similarity; members present on both chromosomes and mobile genetic elements.

These are the most common characteristics of *bona fide* TA gene families but there are, of course, TAs that do not fulfil all criteria. For example, the *ccdAB* family has a relatively narrow phylogeny (Chap. 4) and yet constitutes one of the best-studied TA models. In one case, the activity of a MazF homologue is controlled by a general transcription factor encoded by a gene located distal to the *mazF* gene (Nariya and Inouye 2008). Type II TA loci that encode three components have been described and may be more widespread than previously thought (Smith and Rawlings 1997; Hallez et al. 2010; Bordes et al. 2011).

1.4 Model Organisms to Study TA Loci

The large number of TA loci in many obligatory or facultative free-living organisms makes their genetic analysis difficult because full penetrance of a phenotype will often require multiple deletions (Maisonneuve et al. 2011). It is therefore important to focus research on selected model organisms. The use of

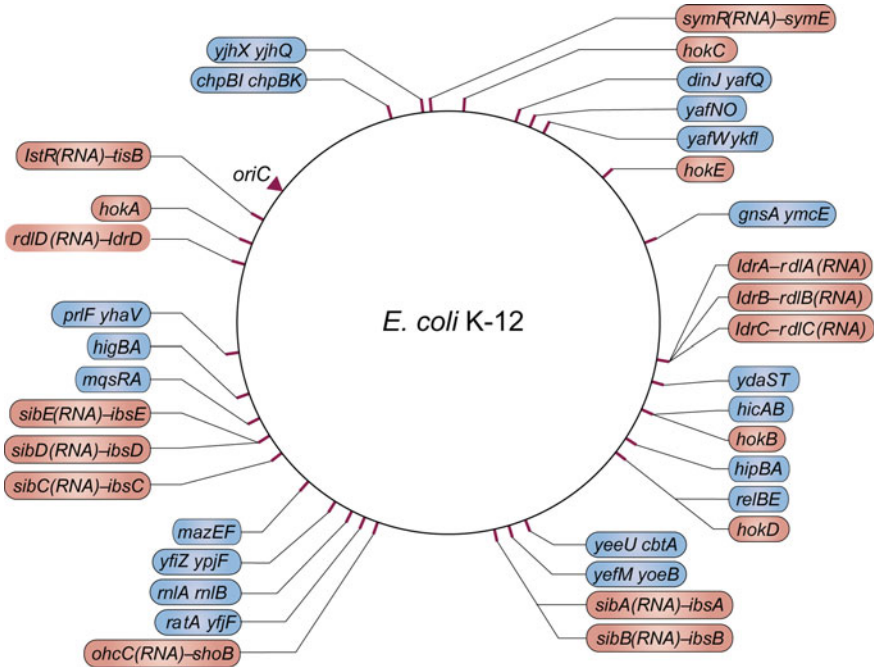


Fig. 1.1 Localization of type I (*red*) and type II (*blue*) toxin-antitoxin loci of *E. coli* K-12. Toxin-antitoxin activity of the gene pairs *yjhX yjhQ* and *gnsA ymcE* has not been confirmed. Adapted from Yamaguchi and Inouye (2011)

E. coli K-12 to study TAs has been fruitful. A recent review lists a host of 36 TA loci in *E. coli* K-12, divided on 17 type I and 19 type II loci (Fig. 1.1). As seen from the figure, the TAs are more or less randomly localized on the chromosome, except for the loci of two type I families (*ldr* and *sib*/*ibs*) that are clustered as repetitive elements.

E. coli is fast and convenient to work with but important analyses of TA loci in other model organisms are under way in many laboratories. For pathogenic organisms, the analysis of TA loci from the highly persistent human pathogen *Mycobacterium tuberculosis* will be particularly exciting to follow. As described in Chap. 8 (Arcus lab) and Chap. 17 (Geneveaux lab), it has already been shown that TA loci are activated during certain conditions in *M. tuberculosis* (Ramage et al. 2009; Bordes et al. 2011; Keren et al. 2011). Chap. 18 (Espinosa lab) shows that *Streptococcus pneumoniae* is also developing into an exciting model organism with hitherto many uncharacterized TAs. Other important persistent pathogens are *Pseudomonas aeruginosa* and *Salmonella enterica*, and we will hopefully soon know more about the biological roles played by TA loci in these organisms.

1.5 Biological Roles of TA Loci

As described in Chap. 16 (Diaz-Orejas lab), both type I and type II TA loci were discovered due to their capability to increase plasmid maintenance while Chap. 15 (Salmond lab) describes that type III loci were discovered due to their ability to aborting phage infection (Fineran et al. 2009). Some type I and II loci can exclude bacteriophages (Pecota and Wood 1996; Hazan and Engelberg-Kulka 2004) and it would not be surprising if some type III loci can stabilize plasmids. These phenotypes can be seen as logical consequences of the properties of the TA gene-encoded components: the antitoxins are almost always more unstable than the corresponding toxin. Thus, in plasmid-free cells or after phage infection, the unstable antitoxin will decay, leading to activation of the toxin and cell stasis or killing. In turn cell stasis or killing leads to increased plasmid maintenance or reduced phage spreading, respectively. However, the penetrance of these phenotypes is highly variable. For example, the *ccd* locus stabilizes the F plasmid 2–5 fold, a factor that should be compared to the 1,000 to 10,000 fold stabilization obtained by the *par* (*sop*) locus of F (Ogura and Hiraga 1983; Boe et al. 1987; Jensen et al. 1995). Similarly, *parD* (*kis/kid*) of R1/R100, *phd/doc* of P1 and *relBE* of P307 also have very modest effects on plasmid stabilization (Bravo et al. 1987; Lehnherr et al. 1993; Grønlund and Gerdes1999) and the *relBE* locus of the *E. coli* chromosome stabilizes plasmids as well as *relBE* from plasmid P307 (Grønlund and Gerdes 1999). In this connection it should be added that *parD (kis/kid)* is activated if the plasmid copy-number is very low and increase plasmid maintenance by a "rescue" mechanism (Pimentel et al. 2005; Lopez-Villarejo et al. 2012). It has not been investigated if this plasmid rescue mechanism is genral. In case of *parDE* from the broad-host-range plasmid RK2/RP4, plasmid stabilization is highly efficient in some hosts while inefficient in others (Sia et al. 1995). These observations certainly manifest that TA loci function in plasmid stabilization. Similarly, TA loci can stabilize chromosomal segments and integrative conjugative elements (Szekeres et al. 2007; Wozniak and Waldor 2009). Clearly, TA loci can increase the genes encoded both by chromosomes and mobile genetic elements, but it remains uncertain whether gene stabilization is the main biological function of TAs. Such a question is difficult to answer but evidence described in Chaps. 5 (Gerdes lab), 11 (Lewis lab), 12 (Meinhart lab), and 17 (Geneveaux lab) suggest that chromosomal TA loci function in persistence and/or stress responses (Christensen et al. 2001; Keren et al. 2004, 2011; Balaban et al. 2004; Shah et al. 2006; Bordes et al. 2011; Mutschler et al. 2011). Chap. 13 (Garrett lab) shows that type II TA loci are highly abundant in archaea, thus implicating a broad biological role to be played by these elements. Chapter 13 also for the first time presents a thorough bioinformatic analysis of TA loci in archaea.

Regarding the mysteriously and perplexing number of functions ascribed to TA loci, I can't help to emphasising the remarkable observation (described in Chap. 5) that progressive deletion of ten type II loci of *E. coli* K-12 gradually reduced the level of persistence (Maisonneuve et al. 2011). This is the first time that a common

phenotype has been observed for chromosome-encoded TA loci belonging to different gene families and, to my view, underscores that persistence is a major function of type II loci. It will be interesting to learn if the persistence of important pathogenic bacteria depends on TA loci.

Acknowledgments I thank Manuel Espinosa Padron for comments to the manuscript. This work was supported by the Welcome Trust and the European Research Council (ERC).

References

Allen, G. C, Jr, & Kornberg, A. (1991). Fine balance in the regulation of DnaB helicase by DnaC protein in replication in *Escherichia coli*. *Journal of Biological Chemistry, 266*, 22096–22101.

Arcus, V. L., McKenzie, J. L., Robson, J., & Cook, G. M. (2011). The PIN-domain ribonucleases and the prokaryotic VapBC toxin-antitoxin array. *Protein Engineering Design and Selection, 24*, 33–40.

Balaban, N. Q., Merrin, J., Chait, R., Kowalik, L., & Leibler, S. (2004). Bacterial persistence as a phenotypic switch. *Science, 305*, 1622–1625.

Blower, T. R., Pei, X. Y., Short, F. L., Fineran, P. C., Humphreys, D. P., Luisi, B. F., et al. (2011a). A processed noncoding RNA regulates an altruistic bacterial antiviral system. *Nature Structural and Molecular Biology, 18*, 185–190.

Blower, T. R., Salmond, G. P., & Luisi, B. F. (2011b). Balancing at survival's edge: The structure and adaptive benefits of prokaryotic toxin-antitoxin partners. *Current Opinion in Structural Biology, 21*, 109–118.

Boe, L., Gerdes, K., & Molin, S. (1987). Effects of genes exerting growth inhibition and plasmid stability on plasmid maintenance. *Journal of Bacteriology, 169*, 4646–4650.

Bordes, P., Cirinesi, A. M., Ummels, R., Sala, A., Sakr, S., Bitter, W., et al. (2011). SecB-like chaperone controls a toxin-antitoxin stress-responsive system in Mycobacterium tuberculosis. *Proceedings of the National Academy of Sciences USA, 108*, 8438–8443.

Bravo, A., Detorrontegui, G., & Diaz, R. (1987). Identification of components of a new stability system of plasmid R1, pard, that is close to the origin of replication of this plasmid. *Molecular and General Genetics, 210*, 101–110.

Buts, L., Lah, J., Dao-Thi, M. H., Wyns, L., & Loris, R. (2005). Toxin-antitoxin modules as bacterial metabolic stress managers. *Trends in Biochemical Sciences, 30*, 672–679.

Christensen, S. K., Mikkelsen, M., Pedersen, K., & Gerdes, K. (2001). RelE, a global inhibitor of translation, is activated during nutritional stress. *Proceedings of the National Academy Sciences USA, 98*, 14328–14333.

Condon, C. (2006). Shutdown decay of mRNA. *Molecular Microbiology, 61*, 573–583.

Fineran, P. C., Blower, T. R., Foulds, I. J., Humphreys, D. P., Lilley, K. S., & Salmond, G. P. (2009). The phage abortive infection system, ToxIN, functions as a protein-RNA toxin-antitoxin pair. *Proceedings of the National Academy Sciences USA, 106*, 894–899.

Gerdes, K., & Wagner, E. G. H. (2007). RNA antitoxins. *Current Opinion in Microbiology, 10*, 117–124.

Gerdes, K., Larsen, J. E., & Molin, S. (1985). Stable inheritance of plasmid R1 requires two different loci. *Journal of Bacteriology, 161*, 292–298.

Gerdes, K., Rasmussen, P. B., & Molin, S. (1986). Unique type of plasmid maintenance function—Postsegregational killing of plasmid-free cells. *Proceedings of the National academy of Sciences of the USA, 83*, 3116–3120.

Gerdes, K., Gultyaev, A. P., Franch, T., Pedersen, K., & Mikkelsen, N. D. (1997). Antisense RNA-regulated programmed cell death. *Annual Review of Genetics, 31*, 1–31.

Gerdes, K., Christensen, S. K., & Lobner-Olesen, A. (2005). Prokaryotic toxin-antitoxin stress response loci. *Nature Reviews Microbiology, 3*, 371–382.

Grønlund, H., & Gerdes, K. (1999). Toxin-antitoxin systems homologous with relBE of *Escherichia coli* plasmid P307 are ubiquitous in prokaryotes. *Journal of Molecular Biology, 285*, 1401–1415.

Hallez, R., Geeraerts, D., Sterckx, Y., Mine, N., Loris, R., & Van, M. L. (2010). New toxins homologous to ParE belonging to three-component toxin-antitoxin systems in *Escherichia coli* O157:H7. *Molecular Microbiology, 76*, 719–732.

Hayes, F. (2003). Toxins-antitoxins: Plasmid maintenance, programmed cell death, and cell cycle arrest. *Science, 301*, 1496–1499.

Hayes, F., & Van Melderen, L. (2011). Toxins-antitoxins: Diversity, evolution and function. *Critical Reviews in Biochemistry and Molecular Biology, 46*, 386–408.

Hazan, R., & Engelberg-Kulka, H. (2004). *Escherichia coli* mazEF-mediated cell death as a defense mechanism that inhibits the spread of phage P1. *Molecular Genetics and Genomics, 272*, 227–234.

Hiraga, S., Jaffe, A., Ogura, T., Mori, H., & Takahashi, H. (1986). F-Plasmid ccd mechanism in escherichia-coli. *Journal of Bacteriology, 166*, 100–104.

Jensen, R. B., Grohmann, E., Schwab, H., Diazorejas, R., & Gerdes, K. (1995). Comparison of ccd of F, parde of Rp4, and pard of R1 using a novel conditional replication control-system of plasmid R1. *Molecular Microbiology, 17*, 211–220.

Keren, I., Shah, D., Spoering, A., Kaldalu, N., & Lewis, K. (2004). Specialized persister cells and the mechanism of multidrug tolerance in *Escherichia coli*. *Journal of Bacteriology, 186*, 8172–8180.

Keren, I., Minami, S., Rubin, E., & Lewis, K. (2011). Characterization and transcriptome analysis of mycobacterium tuberculosis persisters. *MBio, 2*, e00100–e00111.

Lehnherr, H., Maguin, E., Jafri, S., & Yarmolinsky, M. B. (1993). Plasmid addiction genes of bacteriophage-P1-Doc, which causes cell-death on curing of prophage, and Phd, which prevents host death when prophage is retained. *Journal of Molecular Biology, 233*, 414–428.

Lopez-Villarejo, J., ago-Navarro, E., Hernandez-Arriaga, A. M., & az-Orejas, R. (2012). Kis antitoxin couples plasmid R1 replication and parD (kis, kid) maintenance modules. *Plasmid, 67*, 118–127.

Maisonneuve, E., Shakespeare, L. J., Jorgensen, M. G., & Gerdes, K. (2011). Bacterial persistence by RNA endonucleases. *Proceedings of the National Academy Sciences USA, 108*, 13206–13211.

Mutschler, H., Gebhardt, M., Shoeman, R. L., & Meinhart, A. (2011). A novel mechanism of programmed cell death in bacteria by toxin-antitoxin systems corrupts peptidoglycan synthesis. *PLoS Biology, 9*, e1001033.

Nariya, H., & Inouye, M. (2008). MazF, an mRNA interferase, mediates programmed cell death during multicellular *Myxococcus* development. *Cell, 132*, 55–66.

Ogura, T., & Hiraga, S. (1983). Mini-F plasmid genes that couple host-cell division to plasmid proliferation. *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences, 80*, 4784–4788.

Pecota, D. C., & Wood, T. K. (1996). Exclusion of T4 phage by the hok/sok killer locus from plasmid R1. *Journal of Bacteriology, 178*, 2044–2050.

Pimentel, B., Madine, M. A., & de la Cueva-Mendez, G. (2005). Kid cleaves specific mRNAs at UUACU sites to rescue the copy number of plasmid R1. *EMBO Journal, 24*, 3459–3469.

Ramage, H. R., Connolly, L. E., & Cox, J. S. (2009). Comprehensive functional analysis of mycobacterium tuberculosis toxin-antitoxin systems: implications for pathogenesis, stress responses, and evolution. *PLoS Genetics, 5*, e1000767.

Shah,D., Zhang,Z.G., Khodursky,A., Kaldalu,N., Kurg,K., Lewis,K. (2006) Persisters: a distinct physiological state of E-coli. *Bmc Microbiology* 6.

Shao, Y., Harrison, E. M., Bi, D., Tai, C., He, X., Ou, H. Y., et al. (2011). TADB: A web-based resource for type 2 toxin-antitoxin loci in bacteria and archaea. *Nucleic Acids Research, 39*, D606–D611.

Sia, E. A., Roberts, R. C., Easter, C., Helinski, D. R., & Figurski, D. H. (1995). Different relative importances of the par operons and the effect of conjugal transfer on the maintenance of intact promiscuous plasmid RK2. *Journal of Bacteriology, 177*, 2789–2797.

Smith, A. S., & Rawlings, D. E. (1997). The poison-antidote stability system of the broad-host-range Thiobacillus ferrooxidans plasmid pTF-FC2. *Molecular Microbiology, 26*, 961–970.

Szekeres, S., Dauti, M., Wilde, C., Mazel, D., & Rowe-Magnus, D. A. (2007). Chromosomal toxin-antitoxin loci can diminish large-scale genome reductions in the absence of selection. *Molecular Microbiology, 63*, 1588–1605.

Van Melderen, L., Bernard, P., & Couturier, M. (1994). Lon-dependent proteolysis of ccdA is the key control for activation of ccdB in plasmid-free segregant bacteria. *Molecular Microbiology, 11*, 1151–1157.

Wozniak, R. A., & Waldor, M. K. (2009). A toxin-antitoxin system promotes the maintenance of an integrative conjugative element. *PLoS Genetics, 5*, e1000439.

Yamaguchi, Y., & Inouye, M. (2011). Regulation of growth and death in *Escherichia coli* by toxin-antitoxin systems. *Nature Reviews Microbiology, 9*, 779–790.

Yamaguchi, Y., Park, J. H., & Inouye, M. (2011). Toxin-antitoxin systems in bacteria and archaea. *Annual Review of Genetics, 45*, 61–79.

Chapter 2
Type I Toxin-Antitoxin Loci: *hok/sok* and *fst*

Keith Weaver

Abstract The first type I TA locus to be discovered was *hok*/*sok* of plasmid R1 due to its ability to stabilize heterologous replicons in *Escherichia coli*. The unraveling of the mechanism of plasmid stabilization yielded a surprise: *hok*/*sok* increases plasmid maintenance by killing of plasmid-free cells, a mechanism known as post-segregational killing (PSK). The *hok*/*sok* locus codes for two RNAs: Sok antisense RNA that inhibits translation of the toxin-encoding *hok* mRNA. The system is regulated such that *hok* mRNA is translated in plasmid-free cells only. Initially, it was difficult to understand how a *cis*-acting antisense (Sok-RNA) that is fully complementary to its target RNA (*hok* mRNA) apparently could regulate translation of the mRNA in a reversible manner. The elucidation of the underlying mechanism behind the paradox revealed a unique type of gene regulation that involves mRNA processing and refolding. The *par* locus of plasmid pAD1 from *Enterococcus faecalis* also stabilizes plasmids by PSK and also encodes two RNAs that interact. However, the mechanism by which the mRNA (RNA I) avoids irreversible inactivation by the antisense RNA (RNA II) is entirely different from that of *hok*/*sok*. Surprisingly, the chromosome of *E. coli* K-12 encodes type I TA loci called *ldr* that appears to share properties with both *hok*/*sok* and *par* of pAD1. This chapter reviews the *hok*/*sok*, *par*, and *ldr* loci.

K. Weaver (✉)
Division of Basic Biomedical Sciences, Sanford School of Medicine,
University of South Dakota, Vermillion, 57069, USA
e-mail: kweaver@usd.edu

K. Gerdes (ed.), *Prokaryotic Toxin-Antitoxins*,
DOI: 10.1007/978-3-642-33253-1_2,

2.1 Introduction

Type I toxin-antitoxin systems deploy an RNA antitoxin that interferes, either directly or indirectly, with toxin translation. The first type I TA system described was *hok/sok* of plasmid R1, and it remains the best characterized. The *par* determinant of *Enterococcus faecalis* plasmid pAD1 was the first type I TA system characterized in Gram-positive bacteria and has many intriguing similarities and differences with the *hok/sok* system. Both *hok/sok* and *par* function to stabilize their respective plasmids by a mechanism known as post-segregational killing (PSK). Both *hok/sok* and *par* also have multiple chromosomal homologs whose functions remain uncertain. The more recently described *ldrD* locus, present on the *Escherichia coli* K-12 chromosome, surprisingly encodes a *par*-like toxin regulated by a *hok/sok*-like mechanism (Kawano et al. 2002). This chapter describes these three systems, comparing their key features.

2.2 The *hok/sok* Paradigm

The *hok/sok* TA/PSK system was first described more than 25 years ago as a locus that stabilized *E. coli* plasmid R1 greater than 100-fold (Gerdes et al. 1985). Ensuing detailed work to characterize the system has established many of the principles that guided the study of more recently described type I TA loci and RNA-regulated systems in general. These include (1) differential stabilities of toxin and antitoxin RNAs; (2) sequence motifs involved in rapid interactions between regulatory and target RNAs; (3) RNA-mediated inhibition of translation and targeted destruction of mRNA; (4) importance of intramolecular stable and metastable structures in controlling the timing of mRNA translation; (5) accumulation of a pool of inactive toxin mRNA that can later be activated upon plasmid loss; (6) the role of membrane active peptide toxins; (7) the presence of chromosomal homologs of plasmid-encoded PSK systems. These basic principles of *hok/sok* function will be described below. For more details see a recent review (Gerdes and Wagner 2007).

The principal components of the *hok/sok* system are *hok* mRNA, which encodes two genes, *hok* (host killing) and *mok* (modulation of killing), and the Sok (suppression of killing) regulatory RNA. As shown in Fig. 2.1a, *mok* overlaps with *hok* extensively and is required for its translation. Sok-RNA is transcribed from the opposite strand of that of *hok* mRNA and is complementary to the *mok* translation initiation region. Sok RNA is produced from a strong promoter while the *hok* RNA promoter is weak (Franch et al. 1997), so Sok RNA is in molar excess over *hok* RNA in the presence of the plasmid (Gerdes et al. 1990b). However, Sok RNA has a half-life of about 30 s while the *hok* RNA half-life is approximately 20 min. Therefore, if a cell fails to inherit a plasmid copy at cell division, Sok RNA is modified by poly(A) polymerase and rapidly depleted by RNase E (Dam

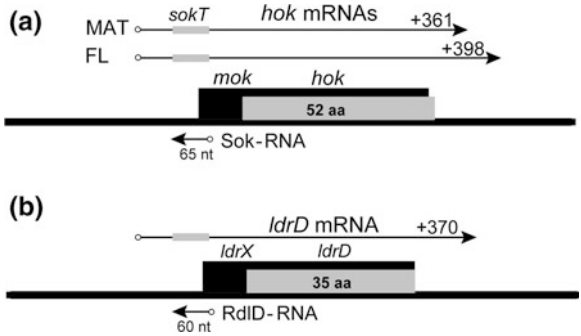


Fig. 2.1 Genetic organization of the *hok/sok* and *ldrD/rdlD* loci. **a** *hok/sok* of plasmid R1. Organization of *hok* and *mok* genes are shown as *gray* and *black boxes*, respectively, on the DNA represented by a *black line*. *hok* and *sok* transcripts are shown above and below the line with transcription start sites represented by a *small circle* and their length in nucleotides (*nt*) indicated. FL and MAT (mature) represent the full length, inactive and processed, active *hok* mRNAs, respectively. *sokT* is the Sok-RNA target in *hok* mRNA, indicated by a *small gray box* on the *hok* transcripts. The *mok* reading frame is required for *hok* translation but the *mok* gene product has no known function. The *mok* reading frame is out-of-frame with *hok* and terminates 1 nt downstream of *hok*. **b** *ldrD* locus of *E. coli* K-12. Genes and transcripts are indicated as in (**a**). *ldrX* denotes an open reading frame that starts with the rare start-codon UUG and overlaps with *ldrD*. *ldrX* and *ldrD* are in-frame and they therefore share the same termination codon. *ldrX* may facilitate *ldrD* translation in the same manner as *mok* stimulates *hok* translation. Reprinted by permission from Gerdes and Wagner 2007

Mikkelsen and Gerdes 1997) freeing *hok* RNA for translation of the Hok toxin resulting in cell death.

Sok inhibits *hok* translation indirectly by interfering with ribosome binding at the *mok* Shine-Dalgarno (SD) sequence (Thisted and Gerdes 1992). In addition, the Sok RNA-*hok* mRNA complex is efficiently cleaved by RNase III, making Sok-mediated repression of *hok* translation irreversible (Franch et al. 1999b; Gerdes et al. 1992). This presents a problem for *hok/sok* functionality since there is more than enough Sok in plasmid-containing cells to eliminate all of the *hok* RNA, thereby leaving none to be translated upon plasmid loss. This problem is solved by the sequential formation of alternative conformations of the *hok* mRNA (Fig. 2.2). Initially, during transcription, the nascent transcript adopts a metastable structure that sequesters both the *mok* and *hok* SDs and the preferred Sok recognition structure, preventing premature translation or destruction of the transcript (Møller-Jensen et al. 2001). Once transcription is complete, the full length *hok* transcript adopts a conformation with several important features: (1) the 5′ and 3′ ends form a blunt end pairing interaction which contributes to the stability of the transcript (Franch and Gerdes 1996), (2) the SD sequences for both *mok* and *hok* are sequestered in stem-loop structures, (3) the preferred Sok binding site is buried within the structure of the *hok* RNA (Gultyaev et al. 1997, 2000). This inactive *hok* RNA isomer is neither translated nor destroyed by Sok binding and so can accumulate as a pool in plasmid-containing cells. Finally, the full length *hok* transcript is slowly processed from the 3′ end by RNase II and PNPase (Franch

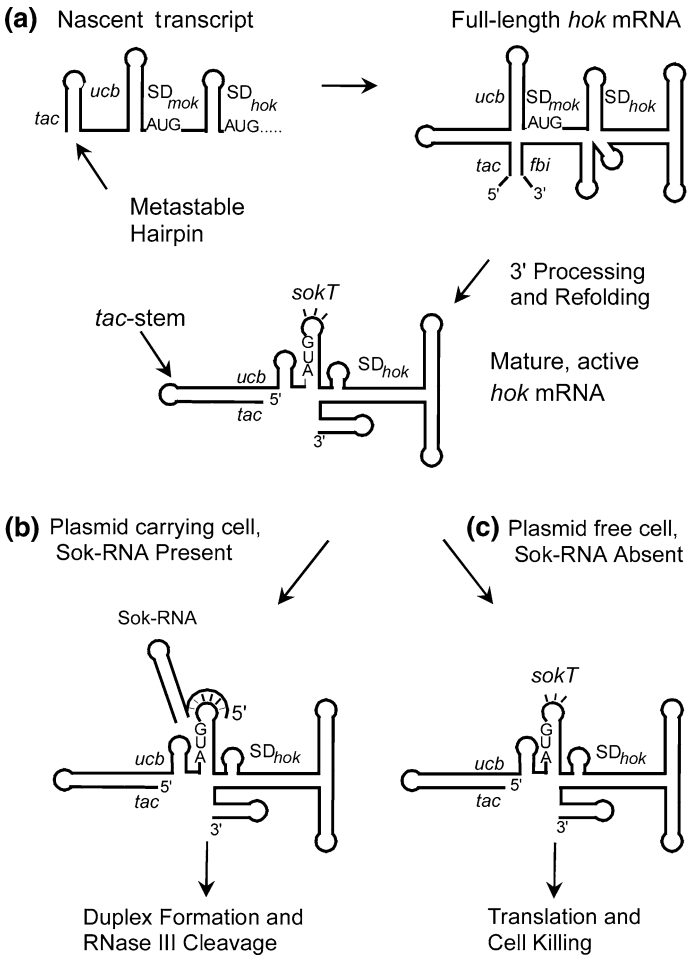


Fig. 2.2 The *hok* mRNA folding pathway (**a**) and the fates of activated *hok* mRNA (**b**, **c**). *tac* denotes the translational activator sequence that sequentially base pairs with three different sequence elements within *hok* mRNA: in the metastable hairpin; with the upstream complementary box (*ucb*) and with the fold-back inhibition (*fbi*) element. The *ucb* element is complementary to both *tac* and SD of *mok*. The everted bases of the U-turn in refolded *hok* mRNA are indicated with *dashes*. Reprinted by permission from Gerdes and Wagner 2007

et al. 1997). Removal of 39 nucleotides at the 3′-end of *hok* mRNA triggers refolding of the RNA resulting in exposure of the *mok* SD and the formation of a stem-loop structure (SokT') that is the preferred recognition site for Sok RNA. SokT' contains a U-turn motif which everts three bases in the loop (Franch et al. 1999a), increasing the rate of Sok binding by 10-fold and allowing rapid inhibition of translation with subsequent RNase III cleavage. As long as the plasmid is retained, Sok continually targets any processed *hok* RNA for destruction. But if the

plasmid is lost, Sok RNA is rapidly depleted and newly processed *hok* RNA is free to bind ribosomes for translation.

The Hok toxin is a 52 amino acid peptide with a predicted transmembrane domain at its N-terminus and a long C-terminal tail. Expression of Hok in plasmid-free segregants leads to the formation of "ghost cells", showing a condensation of cell material at the poles and a translucent cell center (Gerdes et al. 1986a, b). Induced overexpression of Hok leads to rapid cell death, loss of membrane potential, and arrest of respiration. These effects precede the cell morphology changes suggesting that they are the initial effects of the toxin (Gerdes et al. 1986a). Fractionation experiments suggest that Hok is membrane localized and topological assays with the Hok homolog, Gef, indicate that the C-terminal tail protrudes into the periplasm (Poulsen et al. 1991). Hok is the founding member of a family of toxic peptides that include both plasmid-encoded and chromosomal members (Gerdes et al. 1990a). Alignments reveal conservation in the putative membrane spanning domain, around two conserved cysteine residues, and at an AYE motif near the C-terminus. Mutagenesis experiments demonstrated the importance of the AYE motif (Gerdes et al. 1990a) and Gef was found to form disulfide bonds in vivo (Poulsen et al. 1991). However, mutation of the Cys residues had only a minor effect on toxicity. Two other Hok homologs, SrnB and PndA, were shown to increase membrane permeability and complement λ holins, suggesting a primary effect on membrane integrity (Ito and Ohnishi 1983; Sakikawa et al. 1989). Hok is not toxic to cells when added externally unless it is forced into cells by electroporation (Pecota et al. 2003) suggesting that it has a target within the cytoplasmic membrane. The identity of this target is yet to be defined.

The Hok-like toxins are commonly found in *hok/sok*-like loci complete with *mok*-like leader regions, Sok-like regulatory RNAs, and the intramolecular regulatory elements necessary for *hok* mRNA conformational changes (Gerdes et al. 1990b; Gultyaev et al. 1997). Therefore, it is likely that plasmid-encoded homologs perform the same function as *hok/sok*, and indeed this has been demonstrated for the F-encoded *srnB* and R483-encoded *pnd* loci (Nielsen et al. 1991). The function of the chromosomal loci is unclear. The five *hok/sok* homologs on the *E. coli* K-12 chromosome appear to have been inactivated by insertion elements, point mutations, or genetic rearrangements, but the possibility for reactivation by unknown signals has not been ruled out (Pedersen and Gerdes 1999). *hok/sok* loci have also been identified on the chromosomes of other enteric bacteria (Faridani et al. 2006), but their functions are similarly obscure.

2.3 par_{pAD1} and Its Relatives

The pAD1 *par* determinant was originally identified as a locus required for maximal stability of the plasmid's basic replicon (Weaver et al. 1993). The first indication that *par* might be a TA system came from the investigation of a

serendipitously isolated pAD1 mini-plasmid that triggered host cell death when induced with cAD1, a peptide pheromone usually required for induction of plasmid conjugation functions (Weaver and Clewell 1989). Later work showed that this phenomenon resulted from the fortuitous fusion of a pheromone-inducible promoter to the toxin-encoding mRNA of the *par* locus, RNA I (Weaver et al. 1996; Weaver and Tritle 1994). Sequence and RNA analysis identified a short transcript convergently transcribed and partially complementary to RNA I (Weaver and Tritle 1994), designated RNA II. It was later demonstrated that RNA II was capable of counteracting the toxic effects of RNA I both *in cis* and *in trans*, confirming its role as the antitoxin of the system (Greenfield and Weaver 2000; Weaver et al. 1996). Toxicity was shown to be due to a 33 amino acid open reading frame designated Fst for *faecalis* stabilizing toxin (Greenfield et al. 2000). It was further demonstrated that the *par* locus, contained on a fragment of 457 nt, stabilized heterologous plasmids at the expense of host cell growth, confirming its role as a PSK system (Weaver 1995; Weaver et al. 1996, 1998). More recently, multiple *par* homologs, sharing both toxin homology and similarity in genetic organization, have been identified on the plasmids and chromosomes of many Gram-positive bacteria (Fozo et al. 2010; Kwong et al. 2010; Weaver et al. 2009). It is presumed that the plasmid-encoded *par* homologs perform a function similar to that of pAD1 *par*, but the function of the chromosomal homologs is as yet unknown. For consistency in nomenclature, we have recommended using either the name of the mobile element or the chromosomal locus designation in subscript with the *par* component (Weaver et al. 2009) and will use that convention in this chapter.

2.3.1 The Genetic Organization of par_{pAD1} and the Interaction of Its RNAs

The genetic organization of par_{pAD1} and the structure of its transcripts are shown in Fig. 2.3 and 2.4. The *par* RNAs are convergently transcribed and share a bidirectional intrinsic terminator. The complementary terminator loop provides one region of complementarity at which the two RNAs interact. The RNAs are also transcribed across a pair of direct repeats, DRa and DRb, which provide a second region of complementarity between RNA $\mathrm{I}_{\mathrm{pAD1}}$ and RNA $\mathrm{II}_{\mathrm{pAD1}}$. Interaction at both the terminator loop and the direct repeats is essential for proper regulation of $\mathrm{Fst}_{\mathrm{pAD1}}$ translation, but the function of these interactions differs. As in *hok/*sok, the interaction between the par_{pAD1} RNAs is initiated at a U-turn motif, in this case in the terminator loop of RNA $\mathrm{I}_{\mathrm{pAD1}}$ (Greenfield et al. 2001). Mutations in the terminator loop reduce the rate of interaction of the RNAs in vitro (Greenfield et al. 2001) and abrogate RNA $\mathrm{II}_{\mathrm{pAD1}}$-mediated protection in vivo (Greenfield and Weaver 2000), suggesting that the rate of interaction is important for translational suppression. Following the initial reversible interaction between the terminator loops, binding is rapidly extended to the DRa and DRb repeats

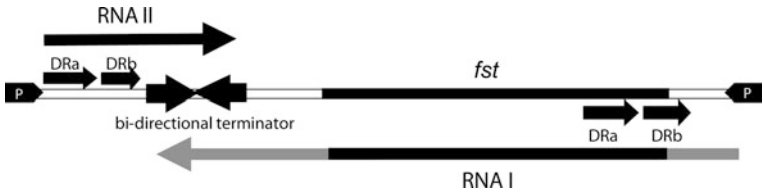


Fig. 2.3 Genetic organization of the pAD1 *par* locus. Converging promoters (*black arrowheads* labeled P) transcribe the toxin-encoding RNA I (*shaded arrow below line*) and the antitoxin RNA II (*dark arrow above line*) toward a bi-directional intrinsic transcriptional terminator (*converging arrows on line*). The RNAs are transcribed across direct repeats (*small arrows* labeled DRa and DRb *above* and *below line*) at which interaction occurs, suppressing translation of the Fst_{pAD1} coding sequence (*black box*)

sequestering the initiation codon, interfering with ribosome binding, and inhibiting translation of the toxic peptide, Fst_{pAD1} (Greenfield et al. 2000, 2001).

The par_{pAD1} locus, therefore, has features of both *cis*- and *trans*-RNA regulated systems (Brantl 2009), unlike *hok/sok* which is a classic *cis*-regulated system. Like in *cis*-regulated systems, the genes for RNA I_{pAD1} and RNA II_{pAD1} overlap providing critical regions of complementarity required for interaction. However, overlap occurs at the 3′ ends of the genes rather than the 5′ ends as in most *cis*-regulated systems. Like *trans*-regulated systems, interaction between the par_{pAD1} RNAs occurs at dispersed regions of complementarity in which different interaction sites play different roles.

2.3.2 Critical Roles of RNA I_{pAD1} Intramolecular Structures in par_{pAD1} Regulation

Like the *hok* mRNA, intramolecular structures of RNA I_{pAD1} affect ribosome access to the SD sequence and RNA stability. Two RNA I intramolecular structures, 5′-SL and 5′-UH (boxed and labeled in Fig. 2.4), impact par_{pAD1} function. 5′-SL is a stem-loop that sequesters the Fst_{pAD1} SD, suppressing translation (Greenfield et al. 2000; Shokeen et al. 2008). Translational suppression is not complete since low levels of translation can be observed in vitro and wild-type RNA I_{pAD1} is toxic in vivo in the absence of RNA II_{pAD1}. However, mutations destabilizing the 5′-SL cannot be established in cells expressing RNA II_{pAD1} (Shokeen et al. 2008) in spite of the fact that it is capable of binding to and suppressing translation from such mutants in vitro (Greenfield et al. 2000). This discrepancy between in vivo and in vitro results may relate to the timing of RNA I_{pAD1}'s interaction with its two competing partners, ribosomes and RNA II_{pAD1}. Because the interaction between the RNAs is initiated at the terminator loop, the ribosome binding site of RNA I_{pAD1} is transcribed and available for ribosome binding before RNA II_{pAD1} can initiate binding. The 5′-SL is postulated to temporarily inhibit ribosome binding until the terminator loop can be transcribed. It is

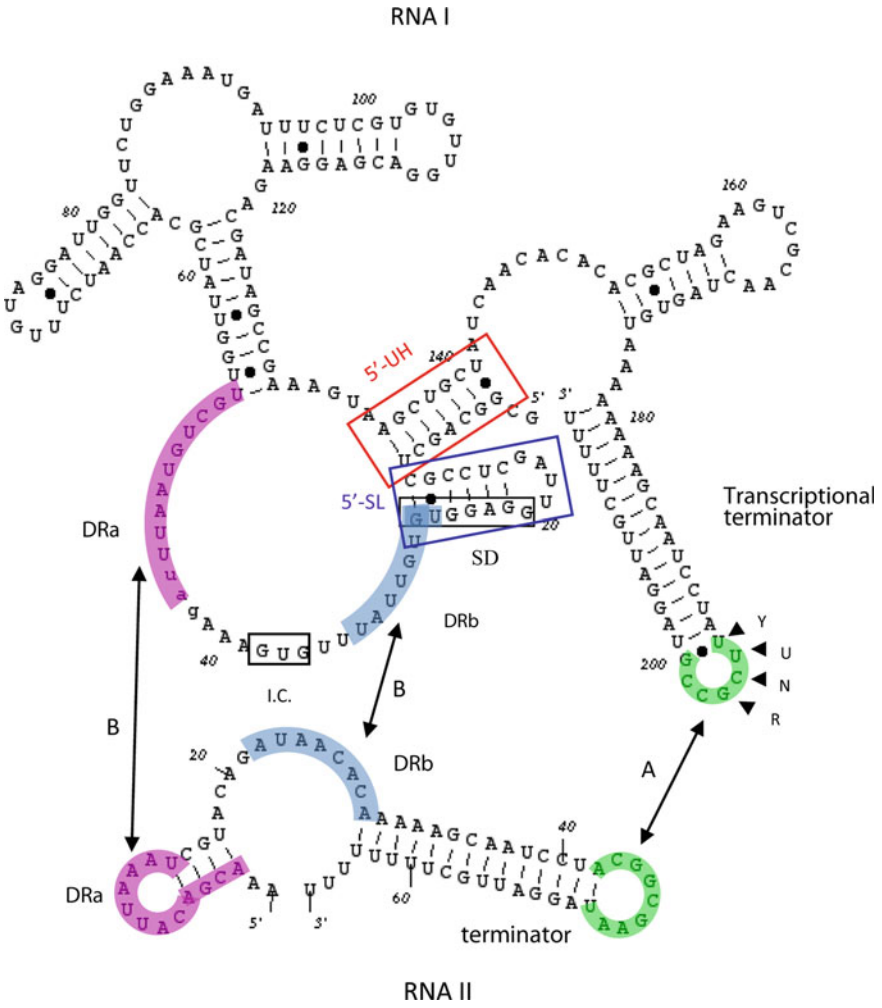


Fig. 2.4 Secondary structures and interaction elements of RNA I and II encoded by *par* of pAD1. The specific regions of interaction between the RNAs are *shaded different colors* and labeled. Interaction is initiated at the U- turn motif (labeled YUNR) present in the loop of the terminator of RNA I (*green shaded*). This interaction is indicated by the *arrow* labeled A. The interaction then extends to the direct repeat sequences DRa (*pink shaded*) and DRb (*blue shaded*). This interaction is indicated by *arrows* labeled B and is responsible for preventing translation of Fst_{pAD1}, since the initiation codon (I.C.) and the ribosome binding site (SD) are sequestered by the interacting RNAs. The two structures, 5′-SL (*blue box*) and 5′-UH (*red box*) are responsible for preventing premature translation of Fst_{pAD1} and RNA I_{pAD1} stability, respectively. This figure has been modified from one appearing in Shokeen et al 2008

also possible that RNA I_{pAD1} is processed to remove the 5′-SL in vivo before it can be translated, but no such processing product has been observed in spite of multiple attempts.

The 5′-UH is an "upstream helix" composed of the extreme 5′ end of the RNA I_{pAD1} transcript and a complementary sequence further downstream that folds back to interact with it. This helix sequesters the 5′ nucleotides from digestion by cellular RNases and is at least partially responsible for the greater stability of RNA I_{pAD1} relative to RNA II_{pAD1} (Shokeen et al. 2009). Mutations in the 5′-UH result in a >fourfold drop in RNA I_{pAD1} half-life from >40 min to around 9 min; the half-life of free RNA II_{pAD1} is approximately 4 min. Mutation of the 5′-UH makes RNA I_{pAD1} more susceptible to RNases J1 and J2, which have 5′-3′-exonuclease activity (Mathy et al. 2007). Whether these are the primary RNases responsible for degradation of RNA I_{pAD1} is not clear. It is also possible that other features of RNA I_{pAD1}, e.g., its relatively inaccessible 3′ end and its compact structure, may also contribute to its stability.

2.3.3 Interaction of par_{pAD1} RNAs Facilitates the Accumulation of a Stable Complex

As in the *hok/sok* system, a sufficient pool of the toxin message must accumulate to allow translation after the plasmid has been lost. Unlike the *hok/sok* system, RNA I_{pAD1} does not appear to adopt alternate structures that prevent regulatory RNA interaction and consequent degradation. Rather, interaction of the par_{pAD1} RNAs leads to stabilization of the RNAs and accumulation of the RNA I_{pAD1}-RNA II_{pAD1} complex. In the presence of RNA I_{pAD1}, RNA II_{pAD1} basal levels increase more than twofold and half-life increases from 4 to 16 min (Weaver et al. 2004). Similarly, the basal level and stability of the RNA I 5′-UH mutant were increased more than twofold in the presence of RNA II (Shokeen et al. 2009). These results suggest that formation of the RNA I_{pAD1}-RNA II_{pAD1} complex protects both RNAs from degradation by cellular RNases. While most regulatory RNAs appear to destabilize their targets, target stabilization is not without precedent (Opdyke et al. 2004). These results led to the following model for regulation of par_{pAD1} function. Following transcription of RNA I_{pAD1}, the 5′-SL prevents ribosome binding until interaction with RNA II_{pAD1} can occur. The translationally inactive complex then accumulates as a pool in the cells with RNA I_{pAD1} to RNA II_{pAD1} ratios maintained at about 1:1.1 (Weaver et al. 2004). It is possible that the discontinuous nature of the interacting sites in the complex prevents efficient degradation by RNase III which requires at least two helical turns of double stranded RNA for binding and activity (Robertson 1982). The lower stability of RNA II_{pAD1} suggests that it is preferentially removed from the complex and degraded by means that are yet to be described. This removal must be active, since in vitro results suggest that the complex does not spontaneously dissociate (Weaver et al. 2004) and could involve RNA helicase and/or targeted RNase action. If plasmid remains in the cell, sufficient RNA II_{pAD1} is produced to replace that removed from the complex. If the plasmid is lost, degraded RNA II_{pAD1} cannot be replaced, the Fst_{pAD1} ribosome

Fig. 2.5 The Fst_{pAD1} toxin. The sequence of the Fst_{pAD1} toxin is shown in *black* using standard single letter amino acid abbreviations. The putative transmembrane domain is *boxed* in *yellow*. All mutations disrupting the toxicity of the peptide are shown *below* the sequence in *green* while those not affecting toxicity are shown *above* the sequence in *red*. Mutations introducing termination codons are represented by (*). This figure is reproduced by permission from Weaver et al. 2009

binding site becomes accessible, either through the processing of the 5′-SL or by the utilization of a ribosomal standby site (de Smit and van Duin 2003), and sufficient toxin is produced to kill the cell.

2.3.4 The Fst_{pAD1} Toxin is a Small, Probably Membrane Localized, Peptide

Like Hok toxin, the Fst_{pAD1} toxin is a small peptide with a predicted transmembrane domain (Fig. 2.5). Beyond that, the two peptides show no significant homology and Fst_{pAD1} lacks cysteine residues and the C-terminal AYE motif important for Hok function. In addition to the transmembrane domain, the 33 amino acid Fst_{pAD1} peptide has a charged N-terminus and a highly charged C-terminal tail. Alanine scanning mutagenesis supplemented with select conservative and non-conservative amino acid changes revealed that the putative transmembrane domain was most important for toxin function (Weaver et al. 2009). Alanine substitutions in many of the amino acids in this region lost toxicity, while substitutions with bulky hydrophobic amino acids leucine and valine were tolerated, consistent with a role in membrane transit. The centrally located P11 residue appears especially important for function since substitution with four different amino acids, including alanine and acidic and basic amino acids, eliminated toxicity. At the N-terminus, substitutions of the two charged amino acids, K2 and D3, with either acidic or basic amino acids retained toxicity. However, alanine substitutions were nontoxic suggesting that polarity at the N-terminus is important for function. Substitution of L4 with either alanine or a charged amino acid eliminated toxicity. In contrast, the charged C-terminal tail appears to contribute little to toxin function. In most cases, alanine substitutions and substitutions reversing amino acid charge had no effect. Indeed, a nonsense mutation at D25

resulting in truncation of the C-terminus was still toxic. A nonsense mutation in the adjacent L24 was nontoxic. It is important to note that the mechanism of testing toxicity could not distinguish degrees of toxicity, so it was not possible to test if toxic mutations might have had reduced toxicity.

An atomic resolution structure of Fst_{pAD1} has been determined in the membrane mimetic dodecylphosphocholine (DPC) by NMR spectroscopy (Göbl et al. 2010). These results indicated that Fst forms a transmembrane α-helix with the first two and the last seven amino acids protruding. The charged C-terminal seven amino acids are disordered and were predicted to extend from the cytoplasmic side of the membrane. These authors suggested that the primary function of membrane insertion was to facilitate interactions with a specific target rather than being directed against the membrane itself. They also predicted that the disordered C-terminus might become structured upon recognition of the target, but this conclusion conflicts with mutagenic studies indicating that the last eight amino acids are not essential for toxicity (Weaver et al. 2009).

2.3.5 *Overexpression of Fst_{pAD1} Affects Nucleoid Structure, Segregation, and Cell Division*

Fst_{pAD1} is toxic to *E. faecalis* (Patel and Weaver 2006; Weaver et al. 2003), *S. aureus* (Weaver et al. 2009), and *B. subtilis* (Patel and Weaver 2006) when overexpressed from the native RNA I transcript. Toxicity can also be observed in *E. coli* if the 5′-SL structure is disrupted (Shokeen et al. 2008). In all four species, the primary effect of toxin overexpression is condensation of the nucleoid. In *E. coli* and *B. subtilis*, this results in elongation of cells, perhaps because the collapsed nucleoid interferes with formation of the division septum at the cell center. In *S. aureus*, the division septum forms and invaginates but the nucleoid is frequently trapped at the convergence point and completion of cell division is inhibited. *E. faecalis* cells initially elongate, then produce misplaced division septae, and finally mis-segregate the nucleoid, producing cells containing little or no DNA. The different effects of Fst may reflect differences in the control of cell division in the different species. In *E. coli* and *B. subtilis*, nucleoid occlusion (NO) systems (Wu and Errington 2012) apparently prevent the formation of division septae over the condensed chromosome. In both *S. aureus* and *E. faecalis* NO appears to be ineffective in stopping invagination of the cell wall or Fst abrogates its function. In *S. aureus*, new cell wall growth occurs only at the septum (Pinho and Errington 2003), so the presence of a condensed nucleoid effectively blocks both division and growth. In the chaining ovococci, however, cell wall growth occurs both longitudinally and septally (Morlot et al. 2003), allowing elongation of Fst-exposed cells with the nucleoid trapped at the division site. In at least some cells, the partition apparatus mobilizes the condensed chromosome, but only into one of the daughter cells.

2.3.6 Fst_{pAD1} is Active at the Membrane but Its Specific Target is Unknown

The putative transmembrane domain of Fst_{pAD1} and its importance to toxin function suggest that the peptide is membrane localized. However, exposure to Fst_{pAD1}, unlike Hok (Gerdes et al. 1986b), does not result in the leakage of cell contents and the formation of "ghost cells." An increase in cell permeability is observed following Fst_{pAD1} overexpression but only after the appearance of cell growth and division anomalies, suggesting that membrane defects may be a secondary effect (Patel and Weaver 2006; Weaver et al. 2003). Nisin and Fst have synergistic effects suggesting that they have different but complementary targets (Weaver et al. 2003). Nisin is a pore forming lantibiotic that docks on lipid II and also affects peptidoglycan synthesis (Wiedemann et al. 2001). Unlike nisin but like Hok, synthetic Fst_{pAD1} has no effect on cell growth when added externally (Weaver et al. 2003), suggesting either that it is modified in some way within the cell or targets a component present only on the inner surface of the membrane or in the cytoplasm. Recent microarray data indicate that exposure to Fst_{pAD1} results in induction of a variety of energy-dependent membrane transporters; interference with this induction by RNA polymerase mutation or interference of ABC transporter activity with reserpine leads to Fst resistance (Brinkman and Weaver, unpublished). It is possible that hyperactivity of energy-utilizing membrane transporters depletes the cells of energy thereby leading to the observed toxic effect.

2.3.7 par_{pAD1} Homologs are Widespread in Gram-Positive Organisms

Work by several groups has revealed that Fst_{pAD1} belongs to a large family of RNA-regulated peptide toxins that also includes Hok (Fozo et al. 2008, 2010; Kwong et al. 2010; Weaver et al. 2009). These peptides are smaller than 60 amino acids, hydrophobic, and predicted to contain an α-helical transmembrane domain. Indeed, many of the smaller peptides may consist solely of the transmembrane helix. Most are toxic when overexpressed in their native host (Fozo et al. 2008). An exhaustive bioinformatic search across 774 bacterial genomes identified hundreds of these peptides in the γ-proteobacteria and Firmicutes that were divided into eight families (Fozo et al. 2010). Fst_{pAD1} is the founding member of the Fst/Ldr family of peptide toxins, which in this analysis consisted of 161 members. (The same study identified 59 Hok homologs in γ-proteobacteria.) In addition, Kwong et al. reported the identification of more than 200 Fst-related peptides in a diversity of Gram-positive bacteria (Kwong et al. 2010). While there is likely significant overlap between these two lists, it seems apparent that Fst peptides are ubiquitous in Gram-positive bacteria and the related Ldr peptides are prevalent in the γ-proteobacteria. In the Gram-positive bacteria, examination of the DNA

sequences surrounding the Fst peptides revealed the existence of all of the elements originally defined in the par_{pAD1} locus, including the convergent promoters for RNA I and RNA II transcripts, a bi-directional intrinsic terminator, the DRa and DRb interacting sequences, and sequences providing the 5′-SL and 5′-UH of RNA I, suggesting that they may be regulated in a similar manner to par_{pAD1} (Kwong et al. 2010; Weaver et al. 2009). U-turn motifs were not always present in the terminator loop, however, suggesting that some features of the interaction pathway might differ in individual systems. In addition, while the general features of the Fst-encoding *par* loci are conserved, their sequences are not, particularly in the DRa and DRb regions predicted to be involved in RNA–RNA interaction. This feature would allow related *par* systems present on different plasmids to operate in the same cell without interfering with one another. Furthermore, a number of *par*-homologs are chromosomally encoded. While many of these are associated with integrated mobile genetic elements or their remnants, some are not. For example, Fst_{EF0409} is located between genes that appear to be associated with mannitol transport and metabolism and is present in all sequenced *E. faecalis* but not *E. faecium* strains. Recent work in our laboratory indicates that it neither interferes with nor is essential for pAD1 *par* function (Weaver, unpublished results).

The function of the chromosomally located *par* loci, like most other chromosomal TA systems, is unknown. Some undoubtedly are remnants of integrated mobile genetic elements or serve to stabilize certain regions of the chromosome. Interestingly, five of the *par* homologs not associated with MGE are intimately linked to genes involved in carbohydrate metabolism (Kwong et al. 2010; Weaver et al. 2009). This includes par_{EF0409} located between genes for mannitol PTS components in *E. faecalis*, $par_{SSP0870}$ located between genes for 6-phosphoglucono-lactonase and an aldehyde dehydrogenase in *Staphylococcus saprophyticus*, $par_{LSEI2682}$ situated between genes for mannose-6-P isomerase and a two-component signal transduction system in *Lactobacillus caseii*, a locus in *S. aureus* MRSA252 located between genes encoding a putative ABC transporter and glycerate kinase, and a *Listeria monocytogenes* locus downstream of a gene encoding a glycosyl hydrolase. The locations of these *par* homologs along with the association of Fst effects on ABC transporters are suggestive of a role in fine-tuning carbohydrate metabolism. This possibility is under active investigation.

2.4 The *ldr/rdl* Locus of *E. coli* K-12: An Fst-Like Peptide Regulated in a *hok/sok*-Like Manner

The long direct repeat (LDR) sequences are four approximately 500 base pair sequences identified by sequencing of the *E. coli* K-12 genome. Three are tandem repeats located at 27.4 min on the chromosomal map in the largest intergenic region of the genome. The fourth, LDR-D, is located at 79.7 min, symmetrically opposite to the other three on the chromosomal map. All four loci encode a small

open reading frame of about 35 amino acids, the LDR-D version of which has been analyzed in detail (Kawano et al. 2002). Overexpression of *ldrD* results in rapid loss of viability and condensation of the nucleoid with subsequent elongation of cells. A small RNA, *rdlD*, is transcribed antisense to the upstream untranslated sequence of *ldrD* and suppresses its translation. As expected of a TA locus, the *ldrD* mRNA is more stable than the *rdlD* antisense RNA. Interestingly, the *ldr* loci also encode a *mok*-like open reading frame that could be regulated by *rdl* RNA and could affect *ldr* translation in a manner similar to the *hok/sok* locus (Fig. 2.1b), although the possibility of *hok*-like conformational changes in the *ldr* transcript have not been investigated. LDR-like sequences are conserved in several *E. coli* strains and *Salmonella* species but their function is unknown. Deletion of the sequences from the genome leads to no obvious phenotypic abnormalities and cloned sequences are unable to stabilize plasmid vectors. Microarray analysis suggests that overexpression of LdrD may affect purine metabolism, but the natural conditions that would lead to such expression are unknown.

More recently, a link was established between the LdrD and Fst toxins by exhaustive PSI-BLAST (Fozo et al. 2010). A superfamily signature consisting of a transmembrane helix followed by a highly conserved tryptophan with a C-terminal tail of charged amino acids was suggested, though the Fst_{pAD1} prototype lacks the conserved tryptophan. The possible relationship between these peptides is further strengthened by their strikingly similar effects upon overexpression in *E. coli*. How these apparently related peptides came to reside in such disparate hosts is a mystery. Although Fst homologs are frequently present on mobile genetic elements, phylogenetic analysis showing coherence between the phylogeny of Fst/Ldr peptides and their hosts of origin suggests that their distribution is not due to recent horizontal gene transfer. The significant differences in the mechanisms of regulation of Ldr and Fst translation also argue for a rather distant relationship. If the LdrD regulatory mechanism turns out to be highly similar to the *hok/sok* system, it will be interesting to speculate on the evolutionary path that fused an Fst-like peptide to a *hok/sok*-like TA locus.

2.5 Conclusion

To date, the *hok/sok* and *par* systems remain the best studied type I TA systems. While detailed analysis has identified many similarities between the two systems, there are also significant differences. For example, while both systems utilize regulatory RNAs to control the translation of their respective toxins and U-turn motifs are critical for interaction timing, the Sok RNA is a classic *cis*-acting antisense RNA while RNA II_{pAD1} interacts with its targets via dispersed regions of complementarity in a manner more similar to *trans*-acting RNA regulators. Both the *hok* RNA and RNA I_{pAD1} adopt compact secondary structures that stabilize the RNAs and suppresses premature translation initiation. However, while the *hok* RNA structure prevents Sok binding in order to allow a stable pool of the toxin

message to accumulate, RNA I_{pAD1} interaction with RNA II_{pAD1} is apparently immediate and it is the complex that accumulates. Both Hok and Fst_{pAD1} are membrane active peptide toxins that must be produced internally to exert their effects. But while Hok expression leads to the production of ghost cells, Fst_{pAD1} (and LdrD) causes nucleoid condensation and division inhibition. It seems likely that *hok/sok* and *par* evolved independently and found partially convergent means to solve similar problems. The surprising hybrid nature of the *ldr/rdl* systems certainly deserves more attention from an evolutionary perspective.

While the *hok/sok* and *par* systems are relatively well-defined, important characteristics remain to be determined. The most pressing issue regarding RNA-mediated regulation is the mechanism of the release of RNA II_{pAD1}-mediated repression of RNA I_{pAD1} translation. What enzymes and what processes are required to remove RNA II_{pAD1} from the RNA complex? Is the 5′-SL of RNA I_{pAD1} removed prior to translation? The mechanism of action of the toxins from all three systems is still shrouded in mystery. What are the targets on the cytoplasmic membrane to which these peptides bind? How do they insert into the membrane? How do Fst_{pAD1} and LdrD affect nucleoid structure? Why does Fst_{pAD1} have different effects on cell division in different hosts? Examination of this last question may provide insights into the regulation of cell division and chromosomal partition in cocci.

It is now clear that RNA-regulated peptide toxins are ubiquitous on bacterial chromosomes. But, as with most of their type II cousins, the function of these systems is mostly unknown. The conservation of the critical elements of *par* loci suggests that there is selective pressure for their maintenance. Their association with genes involved in carbohydrate metabolism is suggestive that they may be involved in regulating this process, but much work remains to determine what role they play.

References

Brantl, S. (2009). Bacterial chromosome-encoded small regulatory RNAs. *Future Microbiol, 4*(1), 85–103. doi:102217/174609134185.

Dam Mikkelsen, N., & Gerdes, K. (1997). Sok antisense RNA from plasmid R1 is functionally inactivated by RNase E and polyadenylated by poly(A) polymerase I. *Molecular Microbiology, 26*(2), 311–320. doi:101046/j1365-295819975751936x.

de Smit, M. H., & van Duin, J. (2003). Translational standby sites: How ribosomes may deal with the rapid folding kinetics of mRNA. *Journal of Molecular Biology, 331*(4), 737–743. doi:101016/s0022-2836(03)00809-x.

Faridani, O. R., Nikravesh, A., Pandey, D. P., Gerdes, K., & Good, L. (2006). Competitive inhibition of natural antisense Sok-RNA interactions activates Hok-mediated cell killing in *Escherichia coli*. *Nucleic Acids Research, 34*(20), 5915–5922. doi:101093/nar/gkl750.

Fozo, E. M., Hemm, M. R., & Storz, G. (2008). Small toxic proteins and the antisense RNAs that repress them. *Microbiology and Molecular Biology Reviews, 72*(4), 579–589. doi:101128/mmbr00025-08.

Fozo, E. M., Makarova, K. S., Shabalina, S. A., Yutin, N., Koonin, E. V., & Storz, G. (2010). Abundance of type I toxin-antitoxin systems in bacteria: Searches for new candidates and discovery of novel families. *Nucleic Acids Research, 38*(11), 3743–3759. doi:101093/nar/gkq054.

Franch, T., & Gerdes, K. (1996). Programmed cell death in bacteria: Translational repression by mRNA end-pairing. *Molecular Microbiology, 21*(5), 1049–1060. doi:101046/j1365-29581996771431x.

Franch, T., Gultyaev, A. P., & Gerdes, K. (1997). Programmed cell death by hok/sok of plasmid R1: Processing at the hok mRNA 3′-end triggers structural rearrangements that allow translation and antisense RNA binding. *Journal of Molecular Biology, 273*(1), 38–51. doi:101006/jmbi19971294.

Franch, T., Petersen, M., Wagner, E. G. H., Jacobsen, J. P., & Gerdes, K. (1999a). Antisense RNA regulation in prokaryotes: rapid RNA/RNA interaction facilitated by a general U-turn loop structure. *Journal of Molecular Biology, 294*(5), 1115–1125. doi:101006/jmbi19993306.

Franch, T., Thisted, T., & Gerdes, K. (1999b). Ribonuclease III Processing of coaxially stacked RNA helices. *Journal of Biological Chemistry, 274*(37), 26572–26578. doi:101074/jbc2743726572.

Gerdes, K., & Wagner, E. G. H. (2007). RNA antitoxins. *Current Opinion in Microbiology, 10*(2), 117–124. doi:101016/jmib200703003.

Gerdes, K., Larsen, J. E., & Molin, S. (1985). Stable inheritance of plasmid R1 requires two different loci. *Journal of Bacteriology, 161*(1), 292–298.

Gerdes, K., Bech, F. W., Jorgensen, S. T., Lobner-Olesen, A., Rasmussen, P. B., Atlung, T., Boe, L., Karlstrom, O., Molin, S., & von Meyenburg, K. (1986a). Mechanism of postsegregational killing by the hok gene product of the parB system of plasmid R1 and its homology with the relF gene product of the E. coli relB operon. EMBO Journal, *5,* 2023–2029.

Gerdes, K., Rasmussen, P. B., & Molin, S. (1986b). Unique type of plasmid maintenance function: postsegregational killing of plasmid-free cells. *Proceedings of the National Academy of Sciences of the United States of America, 83*(10), 3116–3120.

Gerdes, K., Poulsen, L. K., Thisted, T., Nielsen, A. K., Martinussen, J., & Andreasen, P. H. (1990a). The hok killer gene family in gram-negative bacteria. *New Biologist, 2*(11), 946–956.

Gerdes, K., Thisted, T., & Martinussen, J. (1990b). Mechanism of post-segregational killing by the hoklsok system of plasmid R1: sok antisense RNA regulates formation of a hok mRNA species correlated with killing of plasmid-free cells. *Molecular Microbiology, 4*(11), 1807–1818. doi:101111/j1365-29581990tb02029x.

Gerdes, K., Nielsen, A., Thorsted, P., & Wagner, E. G. H. (1992). Mechanism of killer gene activation. Antisense RNA-dependent RNase III cleavage ensures rapid turn-over of the stable Hok, SrnB and PndA effector messenger RNAs. *Journal of Molecular Biology, 226*(3), 637–649. doi:10.1016/0022-2836(92)90621-p.

Göbl, C., Kosol, S., Stockner, T., Rückert, H. M., & Zangger, K. (2010). Solution structure and membrane binding of the toxin Fst of the par addiction module. *Biochemistry, 49*(31), 6567–6575. doi:101021/bi1005128.

Greenfield, T. J., & Weaver, K. E. (2000). Antisense RNA regulation of the pAD1 par post-segregational killing system requires interaction at the 5′ and 3′ ends of the RNAs. *Molecular Microbiology, 37*(3), 661–670. doi:101046/j1365-2958200002034x.

Greenfield, T. J., Ehli, E., Kirshenmann, T., Franch, T., Gerdes, K., & Weaver, K. E. (2000). The antisense RNA of the par locus of pAD1 regulates the expression of a 33-amino-acid toxic peptide by an unusual mechanism. *Molecular Microbiology, 37*(3), 652–660. doi:101046/j1365-2958200002035x.

Greenfield, T. J., Franch, T., Gerdes, K., & Weaver, K. E. (2001). Antisense RNA regulation of the par post-segregational killing system: Structural analysis and mechanism of binding of the antisense RNA, RNAII and its target, RNAI. *Molecular Microbiology, 42*(2), 527–537. doi:101046/j1365-2958200102663x.

Gultyaev, A. P., Franch, T., & Gerdes, K. (1997). Programmed cell death by hok/sok of plasmid R1: Coupled nucleotide covariations reveal a phylogenetically conserved folding pathway in the hok family of mRNAs. *Journal of Molecular Biology, 273*(1), 26–37. doi:101006/jmbi19971295.

Gultyaev, A. P., Franch, T., & Gerdes, K. (2000). Coupled nucleotide covariations reveal dynamic RNA interaction patterns. *RNA, 6*(11), 1483–1491.

Ito, R., & Ohnishi, Y. (1983). The roles of RNA polymerase and RNAase I in stable RNA degradation in *Escherichia coli* carrying the srnB+ gene. *Biochimica et Biophysica Acta, 739*(1), 27–34. doi:101016/0167-4781(83)90040-4.

Kawano, M., Oshima, T., Kasai, H., & Mori, H. (2002). Molecular characterization of long direct repeat (LDR) sequences expressing a stable mRNA encoding for a 35-amino-acid cell-killing peptide and a cis-encoded small antisense RNA in *Escherichia coli*. *Molecular Microbiology, 45*(2), 333–349. doi:101046/j1365-2958200203042x.

Kwong, S. M., Jensen, S. O., & Firth, N. (2010). Prevalence of Fst-like toxin-antitoxin systems. *Microbiology, 156*(4), 975–977. doi:101099/mic0038323-0.

Mathy, N., Bénard, L., Pellegrini, O., Daou, R., Wen, T., & Condon, C. (2007). 5′-to-3′ exoribonuclease activity in bacteria: Role of RNase J1 in rRNA maturation and 5′ stability of mRNA. *Cell, 129*(4), 681–692. doi:101016/jcell200702051.

Møller-Jensen, J., Franch, T., & Gerdes, K. (2001). Temporal translational control by a metastable RNA structure. *Journal of Biological Chemistry, 276*(38), 35707–35713. doi:101074/jbcM105347200.

Morlot, C., Zapun, A., Dideberg, O., & Vernet, T. (2003). Growth and division of *Streptococcus pneumoniae*: Localization of the high molecular weight penicillin-binding proteins during the cell cycle. *Molecular Microbiology, 50*(3), 845–855. doi:101046/j1365-2958200303767x.

Nielsen, A. K., Thorsted, P., Thisted, T., Wagner, E. G. H., & Gerdes, K. (1991). The rifampicin-inducible genes srn6 from F and pnd from R483 are regulated by antisense RNAs and mediate plasmid maintenance by kiiling of plasmid-free segregants. *Molecular Microbiology, 5*(8), 1961–1973. doi:101111/j1365-29581991tb00818x.

Opdyke, J. A., Kang, J.-G., & Storz, G. (2004). GadY, a small-RNA regulator of acid response genes in *Escherichia coli*. *Journal of Bacteriology, 186*(20), 6698–6705. doi:101128/jb186206698-67052004.

Patel, S., & Weaver, K. E. (2006). Addiction toxin Fst has unique effects on chromosome segregation and cell division in *Enterococcus faecalis* and *Bacillus subtilis*. *Journal of Bacteriology, 188*(15), 5374–5384. doi:101128/jb00513-06.

Pecota, D. C., Osapay, G., Selsted, M. E., & Wood, T. K. (2003). Antimicrobial properties of the *Escherichia coli* R1 plasmid host killing peptide. *Journal of Biotechnology, 100*(1), 1–12. doi:101016/s0168-1656(02)00240-7.

Pedersen, K., & Gerdes, K. (1999). Multiple hok genes on the chromosome of *Escherichia coli*. *Molecular Microbiology, 32*(5), 1090–1102. doi:101046/j1365-2958199901431x.

Pinho, M. G., & Errington, J. (2003). Dispersed mode of *Staphylococcus aureus* cell wall synthesis in the absence of the division machinery. *Molecular Microbiology, 50*(3), 871–881. doi:101046/j1365-2958200303719x.

Poulsen, L. K., Refn, A., Molin, S., & Andersson, P. (1991). Topographic analysis of the toxic Gef protein from *Escherichia coli*. *Molecular Microbiology, 5*(7), 1627–1637. doi:101111/j1365-29581991tb01910x.

Robertson, H. D. (1982). *Escherichia coli* ribonuclease III cleavage sites. *Cell, 30*(3), 669–672. doi:101016/0092-8674(82)90270-7.

Sakikawa, T., Akimoto, S., & Ohnishi, Y. (1989). The pnd gene in *E. coli* plasmid R16: Nucleotide sequence and gene expression leading to cell Mg2+ release and stable RNA degradation. *Biochimica et Biophysica Acta, 1007*(2), 158–166. doi:10.1016/0167-4781(89)90034-1.

Shokeen, S., Patel, S., Greenfield, T. J., Brinkman, C., & Weaver, K. E. (2008). Translational regulation by an intramolecular stem-loop is required for intermolecular RNA regulation of the par addiction module. *Journal of Bacteriology, 190*(18), 6076–6083. doi:101128/jb00660-08.

Shokeen, S., Greenfield, T. J., Ehli, E. A., Rasmussen, J., Perrault, B. E., & Weaver, K. E. (2009). An intramolecular upstream helix ensures the stability of a toxin-encoding RNA in *Enterococcus faecalis*. *Journal of Bacteriology, 191*(5), 1528–1536. doi:101128/jb01316-08.

Thisted, T., & Gerdes, K. (1992). Mechanism of post-segregational killing by the hok/sok system of plasmid R1: Sok antisense RNA regulates hok gene expression indirectly through the

overlapping mok gene. *Journal of Molecular Biology, 223*(1), 41–54. doi:101016/0022-2836(92)90714-u.

Weaver, K. E. (1995). *Enterococcus faecalis* plasmid pAD1 replication and maintenance. *Developments in Biological Standardization, 85*, 89–98.

Weaver, K. E., & Clewell, D. B. (1989). Construction of *Enterococcus faecalis* pAD1 miniplasmids: Identification of a minimal pheromone response regulatory region and evaluation of a novel pheromone-dependent growth inhibition. *Plasmid, 22*(2), 106–119. doi:101016/0147-619x(89)90020-6.

Weaver, K. E., & Tritle, D. J. (1994). Identification and characterization of an *Enterococcus faecalis* plasmid pAD1-encoded stability determinant which produces two small RNA molecules necessary for its function. *Plasmid, 32*(2), 168–181. doi:101006/plas19941053.

Weaver, K. E., Clewell, D. B., & An, F. (1993). Identification, characterization, and nucleotide sequence of a region of *Enterococcus faecalis* pheromone-responsive plasmid pAD1 capable of autonomous replication. *Journal of Bacteriology, 175*(7), 1900–1909.

Weaver, K. E., Jensen, K. D., Colwell, A., & Sriram, S. (1996). Functional analysis of the *Enterococcus faecalis* plasmid pAD1-encoded stability determinant par. *Molecular Microbiology, 20*(1), 53–63. doi:101111/j1365-29581996tb02488x.

Weaver, K. E., Walz, K. D., & Heine, M. S. (1998). Isolation of a derivative of *Escherichia coli–Enterococcus faecalis* shuttle vector pAM401 temperature sensitive for maintenance in *E. faecalis* and its use in evaluating the mechanism of pAD1par-dependent plasmid stabilization. *Plasmid, 40*(3), 225–232. doi:101006/plas19981368.

Weaver, K. E., Weaver, D. M., Wells, C. L., Waters, C. M., Gardner, M. E., & Ehli, E. A. (2003). *Enterococcus faecalis* plasmid pAD1-encoded Fst toxin affects membrane permeability and alters cellular responses to lantibiotics. *Journal of Bacteriology, 185*(7), 2169–2177. doi:101128/jb18572169-21772003.

Weaver, K. E., Ehli, E. A., Nelson, J. S., & Patel, S. (2004). Antisense RNA regulation by stable complex formation in the *Enterococcus faecalis* plasmid pAD1 par addiction system. *Journal of Bacteriology, 186*(19), 6400–6408. doi:101128/jb186196400-64082004.

Weaver, K. E., Reddy, S. G., Brinkman, C. L., Patel, S., Bayles, K. W., & Endres, J. L. (2009). Identification and characterization of a family of toxin-antitoxin systems related to the *Enterococcus faecalis* plasmid pAD1 par addiction module. *Microbiology, 155*(9), 2930–2940. doi:101099/mic0030932-0.

Wiedemann, I., Breukink, E., van Kraaij, C., Kuipers, O. P., Bierbaum, G., de Kruijff, B., et al. (2001). Specific binding of nisin to the peptidoglycan precursor lipid II combines pore formation and inhibition of cell wall biosynthesis for potent antibiotic activity. *Journal of Biological Chemistry, 276*(3), 1772–1779. doi:101074/jbcM006770200.

Wu, L. J., & Errington, J. (2012). Nucleoid occlusion and bacterial cell division. *Nature Reviews Microbiology, 10*(1), 8–12.

Chapter 3
Novel Type I Toxin-Antitoxins Loci

Elizabeth Fozo

Abstract Although plasmid encoded type I TA loci have been known for many years, recent discoveries have identified novel, chromosomally encoded type I loci. Many novel type I families have now been identified but their exact biological roles are not known. Here, I describe their discovery, their distribution, and the possible functions for these unusual gene pairs.

3.1 Introduction

Type I TA loci consist of two genes: one encoding a small, hydrophobic potentially toxic protein (the toxin) often 60 amino acids or less; and a second encoding a small RNA (the antitoxin) encoded on the opposite strand of DNA to the protein gene. These sRNA antitoxins act by base pairing to complementary sequences within the toxin mRNA. This leads to the formation of a double-stranded RNA, which can repress translation of the toxin mRNA and/or destabilize the toxin mRNA, leading to decreased levels of toxic protein. Type I loci, including *hok*/*sok* of R1 and *fst* of pAD1, were initially described on plasmids where they serve a role in plasmid maintenance (see Chap. 2). For clarity, I will refer to these as plasmid-based loci. Homologs of *hok*/*sok* and *fst* were later found encoded within bacterial chromosomes (Faridani et al. 2006; Pedersen and Gerdes 1999; Weaver et al. 2009).

Excitingly, in 2002, a new type I locus was described. This locus, *ldr-rdl*, had no discernable homology to any plasmid sequence and was found encoded within the *Escherichia coli* K-12 chromosome and related bacteria (Kawano et al. 2002).

E. Fozo (✉)
Department of Microbiology, University of Tennessee, Knoxville, TN 37996, USA
e-mail: efozo@utk.edu

K. Gerdes (ed.), *Prokaryotic Toxin-Antitoxins*,
DOI: 10.1007/978-3-642-33253-1_3,

These chromosomal loci, with no known plasmid-encoded homologs, are herein referred to as the novel type I loci. Additional novel type I loci have since been discovered either by experimental approaches to identify sRNAs, experimental serendipity, or by bioinformatics approaches.

3.2 Features of Novel Type I Loci

The novel loci have many interesting features, some shared with the plasmid-based loci, and others unique to those encoded on chromosomes. The main feature of all type I loci is that the toxic protein is small (60 amino acids or less) and hydrophobic. Essentially, the proteins appear to be no more than a transmembrane domain, with either a short N-terminus or C-terminus tail. For those with C-termini tails, the C-terminus is typically rich in polar or aromatic residues (Fozo et al. 2010).

The novel loci are encoded distant from their flanking genes, often by as many as 200 nucleotides (nt) or more. This observation at first suggests that these are mobile genetic elements that can be acquired by horizontal gene transfer. However, for those examined, there is no significant difference between the G + C content of the locus and the surrounding chromosomal content. Also, there is no indication of insertion or repetitive elements reminiscent of horizontal gene transfer.

The novel toxins possess rather long 5′ or 3′ untranslated regions (UTRs). The UTR in many cases is actually longer than the coding sequence of the toxin gene. Usually, the RNA antitoxin bind to these long UTRs. The UTRs are folded into stable stem-loop structures that, in the cases of the 5′ UTRs, sequester the ribosome binding site and start codon and thereby inhibit translation. Thus, the UTRs can repress toxin translation, independently of the action of the antitoxin RNA.

Type I antitoxins function similarly to the well-studied chromosomally encoded sRNAs. However, the base pairing potential of an antitoxin for its target is much more extensive than that of a conventional sRNA. Typical sRNAs pair over 6–12 nt and often require the RNA chaperone Hfq to stabilize the interaction (Waters and Storz 2009). The antitoxins have a much greater region of complementarity to their targets, from 18 to 60 nt or more. This extensive complementarity is likely why the known antitoxins do not rely on the bacterial protein Hfq to facilitate their interactions with their targets.

Another interesting feature of some of these chromosomal systems (like Ldr-Rdl, Ibs-Sib, and Zor-Orz) is that they are tandemly duplicated within the same intergenic region. Strains may also contain multiple copies of the same locus scattered about the genome. Why these loci are duplicated is not known; in some cases, the duplicated copies are practically identical in sequence, whereas for others, the loci are rather divergent in sequence. Whether these divergent homologs have unique cellular functions remains an unexplored area of research.

Table 3.1 Novel Type I Loci

Traditional	Arrangement[a]	Sequence[b]	Organism[c]
Lrd-Rdl		MTFAELGMAF WHDLAAPVIA GILASMINVNW LNKRK	*E. coli*
TxpA-RatA		MSTYESLMVM IGFANLIGGI MTWVISLLTL LFMLRKKDTH PIYITVKEKC LHEDPPIKG	*B. subtilis*
Ibs-Sib		MMRLVIILIV LLLISFSAY	*E. coli*
Brg-SR4		MTVYELSMIM INFGGLILNT VLLIFNIMMI VTSSQKKK	*B. subtilis*
YhzE-AsYhzE		MSGGYSNGFA LLVVLFILLI IVGAAYIY	*B. subtilis*
YonT-AsYonT		MLEKMGIVVA FLISLTVLTI NSLTIVEKVR NLKNGTSKKK KRIRKRLRPK RQRQRIRR	*B. subtilis*
Nontraditional			
TisB-IstR-1		MNLVDIAILI LKLIVAALQL LDAVLKYLK	*E. coli*
ShoB-OhsC		MTDCRYLIKR VIKIIIAVLQ LILLFL	*E. coli*
Zor-Orz		MDSLTQKLTV LIAVLELLVA LLRLIDLLK	*E. coli*

[a] The black arrow indicates the toxin mRNA with the square representative of the open reading frame within the mRNA; the red arrow is indicative of the antitoxin
[b] A representative protein sequence is listed for each toxin, including LdrD, IbsC, ZorO
[c] The organism listed is where the family was first described

The organization of the novel type I loci can be varied. For example, some antitoxins are encoded antisense to the long 5′ or 3′ UTR or directly antisense to the coding region of the toxin mRNA (Table 3.1). In a few cases, the antitoxin is encoded divergently to the toxin. Such an arrangement has not been reported for plasmid-based type I toxins. In cases where the sRNA is not encoded directly antisense to the toxin mRNA, one may ask, "how can there be pairing and regulation?" In these cases, the antitoxin RNA has 18–21 nt of perfect complementarity to the toxin mRNA, allowing for regulation. A more detailed discussion of these unusual loci will follow below.

3.3 Classification of Novel Loci Based on Gene Arrangement

Based upon the genetic arrangement of the novel loci, herein I will classify the novel loci into two main categories: conventional novel loci and the unconventional novel loci. The conventional loci are arranged such that the antitoxin is

encoded opposite to the 5′ or 3′ UTR or the coding region of the toxin (Table 3.1). This is the most common organizational structure and is also seen with the plasmid-encoded loci. These toxin families also tend to be widespread in nature. The unconventional loci were discovered recently. Here, the antitoxin is encoded divergent from the toxin, but they do share some extensive sequence complementarity. Fewer of these loci have been identified and their taxonomic distribution is more limited.

3.4 Conventional Loci

3.4.1 Ldr-Rdl Family: The First Family to be Described

The Ldr-Rdl family was the first novel type I locus discovered (Kawano et al. 2002). Originally, three Long Direct Repeat (LDR) sequences were noted within the *E. coli* strain MG1655. These repeats, LDR-A, LDR-B, and LDR-C, are approximately 530 nt in length, and are found in tandem to each other on the *E. coli* chromosome. In their analysis, Kawano et al. detected a fourth repeat (LDR-D) sequence, encoded distal from the previous three (Kawano et al. 2002). They noted two distinct transcripts encoded opposite of each other within the LDR-D repeat. One of these transcripts contained an open reading frame, named *ldrD*, that encoded a highly hydrophobic, 35 amino acid peptide. The *ldrD* mRNA was shown to have a long 5′ UTR (180 nt) and encoded antisense to the UTR is *rdlD*. RdlD, the second transcript, was shown to be a sRNA. The same genetic arrangement was noted in the other three LDR repeats. Given the similarities to the plasmid-encoded type I loci, the authors hypothesized that this was a toxin–antitoxin locus, and designed a "rescue" experiment to prove this (Fig. 3.1). Overexpression of LdrD was toxic to *E. coli*, but co-expression of the sRNA RdlD prevented this toxicity (Kawano et al. 2002). A final regulatory consequence noted was that RdlD had a much shorter half-life than the *ldrD* mRNA, similar to what is reported for the *hok* toxin-encoding mRNA and Sok antisense sRNA, where *hok* mRNA has a much longer half-life than Sok-RNA (Gerdes et al. 1990). It is probable that the instability of RdlD is important for allowing translation of the LdrD mRNA during specific growth conditions.

RdlD can potentially base pair to the 5′ UTR of *ldrD*, approximately 45–100 nt upstream of the start codon of *ldrD*. How then could RdlD prevent expression of *ldrD*? Scrutiny of the DNA sequence revealed a potential small open reading frame, denoted *ldrX* that is entirely located within the pairing region RdlD (Gerdes and Wagner 2007). The authors hypothesized that translation of *ldrX* is necessary for translation of *ldrD*. This situation appears to be analogous to the *hok* mRNA; here the 5′-end of the small open reading frame *mok,* which is required for translation of *hok*, overlaps with the Sok pairing region (Thisted and Gerdes 1992).

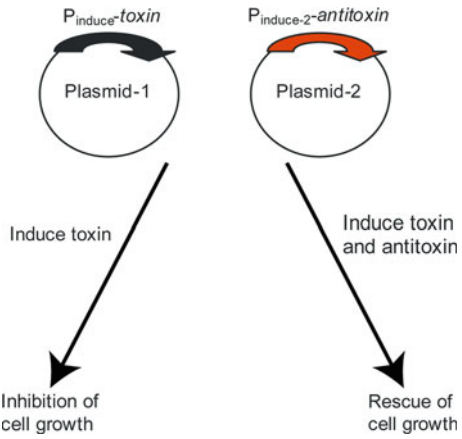


Fig. 3.1 Typical rescue experiment to confirm the presence of an antitoxin gene. The toxin gene is cloned behind an inducible promoter on a plasmid. The antitoxin is cloned behind a different inducible promoter on a compatible plasmid. Induction of the toxin gene alone results in inhibition of cell growth due to cell stasis or death. Co-expression of an antitoxin prevents the inhibition

There have been attempts to understand how overproduction of LdrD leads to cell death. Quite soon after overexpression of LdrD, DNA condensation is observed; this feature is not shared by other toxins, and indicates that LdrD may have unique targets in the cell (Kawano et al. 2002; Fontaine, unpublished observation). Although highly hydrophobic, overproduction of LdrD caused significantly slower dissipation of proton motive force when compared to other type I toxins (Fontaine and Fozo, unpublished observation). Combined, this suggests that the membrane is not the primary target of LdrD.

The biological function of the Ldr family is still unknown. Deletion approaches have not yielded strong phenotypes; however, only a limited number of growth conditions have been tested (Hobbs et al. 2010; Kawano et al. 2002). Rather little is known about transcriptional control of this family; further work in this area could give clues to function as well.

The Ldr toxin family is much broader that initially anticipated. Using a bioinformatics approach to identify all possible Ldr homologs, it was discovered that the Ldr and the Fst family (Fst was originally described as a plasmid-based toxin, see Chap. 2) are actually part of the same "superfamily" (Fozo et al. 2010). This finding was especially interesting when considering the effects of LdrD and Fst-induced toxicity: both cause DNA condensation when overproduced. Again, this feature is not found in other type I loci, suggesting unique biochemical properties of these proteins.

3.4.2 TxpA-RatA: Gram Positives Get into the Act

The first novel chromosomal locus found in a Gram-positive bacteria was TxpA-RatA, discovered in an sRNA screen in *Bacillus subtilis*. An RNA encoded by the intergenic region of *ybqM* (*txpA*)-*ybqN* was detected by microarray analysis. Further work confirmed that this sRNA (RatA) is encoded convergently to *txpA*

and that the 3′ ends of the antisense and mRNA overlap by approximately 75 nt (Silvaggi et al. 2005; Table 3.1). This is very similar to the *fst* locus of plasmid pAD1 of *Enterococcus faecalis* (Weaver et al. 1996). Given the similar arrangement to *fst*, and that *txpA* encodes a 59 amino acid hydrophobic protein, the authors examined whether *txpA-ratA* constitues a bona fide type I toxin–antitoxin locus (Silvaggi et al. 2005). Overproduction of TxpA caused cell death in *B. subtilis*, whereas co-expression of RatA prevented cell death. Furthermore, a strain deleted for RatA has increased levels of *txpA* mRNA compared to the parental strain. This suggests that binding of RatA to the toxin mRNA leads to degradation of the toxin-encoding mRNA; hence, the increase in *txpA* levels in a *ratA* deletion strain.

Functional analysis has been accomplished with a *ratA* deletion strain (Silvaggi et al. 2005). The authors noted lysis of *ΔratA* on agar plates after several days of growth. This phenotype could be complemented by providing RatA in *trans*. Additionally, suppressor colonies arose that did not lyse. Sequencing of these strains revealed mutations in *txpA*, including premature stop codons and missense mutations. These results confirm the regulatory role of RatA and demonstrate the inherent toxicity of *txpA* overexpression.

How TxpA overproduction causes cell death is not clear. It is highly hydrophobic, and may potentially integrate into the membrane. This, along with the observed lysis phenotype, suggests that TxpA forms pores in the cell membrane. Surprisingly, two separate studies confirmed that TxpA overproduction in *E. coli* has no impact on cell growth, suggesting that the target of TxpA is missing in *E. coli*, or that TxpA does not reach toxic levels in *E. coli* (Fozo et al. 2010; Silvaggi et al. 2005).

The number of TxpA homologs across a wide array of species has grown considerably through novel bioinformatic analyses. In a search to discover new toxin–antitoxin loci, the gene *ef3263* from *E. faecalis* V583 was identified as a putative toxic protein (Fozo et al. 2010). This protein was then used as "bait" in a modified PSI-BLAST search. In this approach, the top hits of the search were then used as "baits" to pull out all possible homologs to EF3263. This repetitive search continued until no further homologs were obtained. Through this approach, it was shown that EF3263 is a distant homolog of TxpA. The searches revealed that TxpA is rather well represented across bacterial species and in fact there are six family members within the chromosome *E. faecalis* V583 alone (Table 3.1). Of note, the sequence of the family members is incredibly diverse, even within the same strain (Table 3.2).

It is important to understand the biological function of TxpA. The locus is encoded within *skin*, a large phage-like element within the *B. subtilis* chromosome (Silvaggi et al. 2005). During sporulation, the *skin* element is excised from the mother spore, but remains present in the forespore, and will be present in future germinating cells. One hypothesis is that *txpA-ratA* functions similarly to plasmid addiction loci; it kills off the mother cell, while allowing the forespore to survive. Studies examining whether or not a *txpA-ratA* locus deletion strain is deficient in sporulation and/or causes the death of the mother cell are needed to confirm this hypothesis.

Table 3.2 Chromosomal TxpA homologs of *E. faecalis*[a]

Protein	Organism	Sequence
TxpA	*B. subtilis*	MSTYESLMVM IGFANLIGGIMTWVISLLTL LFMLRKKDTH PIYITVKEKC LHEDPPIKG
EF3263	*E. faecalis*	MTVFEALMLA IAFATLIVKI SNKNDKK
EF0723	*E. faecalis*	MFLSVEAALG LMGFATVVTI IFVILALVLD NKNNRS
EF3247	*E. faecalis*	MCREVMAVAI QSEREVMRMH VFPKFTERRG LLSAYETIQT ILGFGMFTIA LIALIVKLLK NDKKK
EF3249	*E. faecalis*	MNVSTKIYER RGLLSIAEAL ALMISFGSFI ATLIFGILKV VKEDKKK
EF3087	*E. faecalis*	MYPMHFNERS IFLSIEAALE LMISFAAFVA LLIFGILEAT KNNKK
EF3088	*E. faecalis*	MYDGGSELER QVMPMCIRCT LTKGAFFLSI EATLELMISF ATLVALLIFG ILEATKNDKK

[a] There is also a homolog on plasmid pTEF1

What about the function of TxpA in species beyond *B. subtilis*? The sequences of TxpA homologs are quite diverse, varying greatly in length and amino acid content. Additionally, species like *E. faecalis* do not form spores, nor do they have a *skin* element within their chromosome. The sequences of the six TxpA members of *E. faecalis* are very different from one another (Table 3.1). Overproduction of these proteins in *E. coli* and *E. faecalis* has shown that only one is toxic to either of these bacteria (Fozo et al. 2010; Miracle and Fozo, unpublished observations). Given the lack of toxicity and the diversity in sequences, it is possible that these homologs have evolved separate functions. Thus, although *txpA* in *B. subtilis* may have a role in spore formation, the role(s) in other species is a mystery.

3.4.3 Ibs-Sib: The Smallest Toxins

Serendipity has also played a hand in the discovery of novel type I loci. Upon completion of the *E. coli* K-12 genome, a series of four repeat sequences were denoted. The sequence repeat was referred to as the QUAD repeat, and analysis noted that within each repeat was strong promoter elements, but no open reading frame following the promoter sequence was observed (Rudd 1999). Thus, the QUADs were hypothesized to encode stable RNAs, and several papers did report detection of the putative RNAs (Argaman et al. 2001; Hershberg et al. 2003; Rivas et al. 2001; Wassarman et al. 2001). Following revision of the genomic sequence, a fifth repeat was found, and the name was changed to the SIB (*s*hort *i*ntergenic a*b*undant sequences) repeat (Fozo et al. 2008b). The repeats are found in *E. coli* strains and related species, and are often repeated multiple times in the genome, sometimes in tandem to each other.

Overexpression of the Sibs gave an unexpected phenotype, and sequence gazing revealed that, encoded on the opposite strand relative to each *sib* was a small open reading frame (Table 3.1). These reading frames encode 18–19 amino acid proteins that were highly toxic to *E. coli* when overexpressed. However, expression of the Sib RNAs could prevent this toxicity (Fozo et al. 2008b). These small proteins were consequently named Ibs (*i*nhibition *b*rings *s*tasis).

The Sib antitoxins, unlike the other type I antitoxins described to date, completely overlap the coding sequence of the Ibs mRNA (Table 3.1). Interestingly, two forms of the Sib RNAs can be detected: a full-length of approximately 140 nt, and a shorter 110 nt form, due to differences in the 3′ end of the RNA. The shorter form mapped such that it ends in the predicted ribosome binding site for the cognate toxin. Processing of the *ibs*-Sib RNA complex likely leads to a shorter Sib form. An *rnc* deletion strain, which lacks RNase III, does have higher levels of the full-length Sib RNAs than a wild-type strain, suggesting that *ibs* mRNAs and Sib antisense RNAs form duplexes that are cleaved by RNase III (Fozo, unpublished results).

Given that the Sib RNAs are very similar to each other, there could be cross talk in their regulation of the Ibs. Initial experiments showed that this was not the case. For example, SibC could prevent IbsC-mediated toxicity, but not IbsE-toxicity (Fozo et al. 2008b). An additional study mapped the region of specificity to the most variable region of Sib sequence (Han et al. 2010). They showed that two regions of the Sib RNA make contact the *ibs* target initially, similar to what was reported with the interaction between RNAII and RNAI encoded by *fst* (Greenfield et al. 2000; Greenfield and Weaver 2000). These two initial contact domains (TRD1 and TRD2) were critical for the specificity observed for a Sib sRNA (Han et al. 2010).

Overproduction of the Ibs toxins leads to a rapid depolarization of the cell membrane and increased expression of the *psp* operon (Fozo et al. 2008). The *psp* operon is induced in response to stresses that impact membrane integrity and/or proton motive force. This induction correlates well with the observed membrane damage induced by the Ibs proteins.

The Ibs proteins are incredibly small and hydrophobic, yet rather toxic upon overproduction. A mutagenesis study was established to determine what residues were critical for Ibs toxicity (Mok et al. 2010). Incredibly, multiple single amino acid substitutions could be made, and toxicity was still maintained. Residues within the putative transmembrane domain, however, were critical for toxicity. These residues may be important for membrane localization or could play a role in protein–protein interactions.

What is the biological function for the Ibs? This question still remains unanswered. However, some clues were obtained using a strain in which the promoter of *sibC* was mutated, while leaving *ibsC* intact, and leading to increased *ibsC* mRNA levels. Although no major growth effects were observed, there was elevated expression of the *psp* operon, suggesting that the cell membrane integrity was compromised (Fozo et al. 2008b). Indeed, a greater portion of the mutant strain (7.5 % of the population) had higher levels of proton motive force

dissipation as compared to the wild-type strain (1 % of the population; Fozo and Fontaine, unpublished observations). Although membrane damage was seen only in a subset of the population, it does suggest that elevated levels of IbsC can have detrimental effects, and that IbsC expression could be uneven within a population of cells. Further experiments examining sensitivity of the strain to physiological conditions known to induce the *psp* response (ethanol, heat, etc.) did not show any significant effects/differences compared to a wild-type strain.

One major difficulty in elucidating the true function of this locus is the inability to detect the *ibs* mRNA and protein *E. coli*. Detection of *ibs* transcripts via northern was possible only upon deletion of the cognate *sib* promoter (Fozo et al. 2008b). Mapping of the transcription start site of *ibs* was possible only by the use of multicopy plasmids (Han et al. 2010). The promoter identified is neither very strong nor was there evidence for binding sites of known transcriptional regulators.

3.4.4 BsrG-Sr4: Limited to Spβ Prophage

Screens to identify sRNAs, along with sequence analysis, led to the recent discovery of another novel type I system in *B. subtilis*, the BsrG-Sr4 locus (Jahn et al. 2012). Reports had indicated that there was an sRNA (Sr4) encoded by the intergenic region of *bsrG-yokL* of *B. subtilis* (Irnov et al. 2010; Saito et al. 2009). Jahn et al. demonstrated that the two RNAs (*bsrG* and Sr4) converge at their 3′ ends, and that they overlap by approximately 120 nt (Jahn et al. 2012). Similar to other type I TA loci encoding 3′ overlapping toxin mRNAs and antitoxin RNAs, RNA pairing appears to induce degradation of the target RNA. The *bsrG* gene encodes a small protein of 38 amino acids. Overproduction of the small protein was highly toxic to *Bacillus*, but co-expression of the SR4 RNA could alleviate this toxicity (Jahn et al. 2012).

Interestingly, the level of the toxin-encoding mRNA drops dramatically in response to heat shock. Careful analysis revealed that this is not owing to repression of transcription, but rather to RNA instability that is independent of SR4 (Jahn et al. 2012). How this phenomenon is related to function is not clear.

Functional studies of BsrG produced results similar to what was seen with TxpA. A strain deleted for SR4 had much higher levels of *bsrG* mRNA. The strain also produced cell lysis on agar plates, as was seen with TxpA (Jahn et al. 2012; Silvaggi et al. 2005). This lysis occurred more rapidly than in the case TxpA, suggesting that either the total levels of BsrG were higher than TxpA or that BsrG itself is more toxic. Suppressors did arise easily in this strain, and many were mapped to mutations in the coding sequence itself—including frame shifts and premature stop codons.

What is the biological function for BsrG? The toxin is encoded by the SPβ prophage element of *B. subtilis*. Homologs are found only within the prophage, which is present in very few species. Given its location in the genome, the locus may serve to maintain the prophage within the population; however, further experiments are needed.

3.4.5 The YhzE Family: A Nontoxic Toxin?

YhzE, a 28 amino acid hydrophobic protein from *B. subtilis* ssp. *subtilis* str. 168, was identified in a computational search as a putative type I toxin (Fozo et al. 2010). Although *yhzE* was annotated, there was a clear, duplicated unannotated homolog encoded in tandem. To distinguish between these genes, the previously annotated gene is now referred to as *yhzE*-1 and the newly identified gene as *yhzE*-2 (Fozo et al. 2010). Many homologs were found across various species of Firmicutes and *B. subtilis* 168 has eight separate members of this family. A highly expressed sRNA encoded convergently to the 3′ end of *yhzE-1* was detected by Northern analysis, suggesting that expression of *yhzE-1* may be regulated by this RNA. However, overexpression of YhzE-1 in *E. coli* did not impact growth, similar to what was reported for TxpA and BsrG, nor was toxicity observed in *B. subtilis* (Jahn et al. 2012; Fozo et al. 2010; Silvaggi et al. 2005).

3.4.6 YonT: Spβ Prophage Déjà Vu

YonT was identified in the same bioinformatic search as YhzE (Fozo et al. 2010). Like *bsrG*, *yonT* is encoded by Sp*β* prophage. Northern analysis detected an approximately 120 nt sRNA convergently transcribed relative to the toxin-encoding gene (overlapping 3′ ends). Overproduction of YonT was toxic in *E. coli*, something rather unusual for the previously described type I toxins from *Bacillus*.

So what is the function of YonT? This locus has not yet been characterized phenotypically or biochemically. Given its location within SP*β*, it may serve to maintain the prophage within the population and it is possible that the *bsrG-sr*4 and *yonT* loci both contribute to the maintenance of the prophage.

3.5 The Unconventional Loci

For unconventional type I loci, the toxins are encoded divergently from their antisense RNAs. Thus, the base pairing potential between these gene pairs is much more limited compared to conventional type I loci. These families are much more limited in their distribution and are found mainly in *E. coli*, *Shigella*, and *Salmonella* species.

3.5.1 TisB-IstR-1: SOS-Induced Toxin

The TisB-IstR-1 pair was discovered in a screen to identify novel small regulatory RNAs in *E. coli* K-12 (Vogel et al. 2004; Argaman et al. 2001). Initially, the authors detected an sRNA (IstR-1) divergent from the uncharacterized operon of

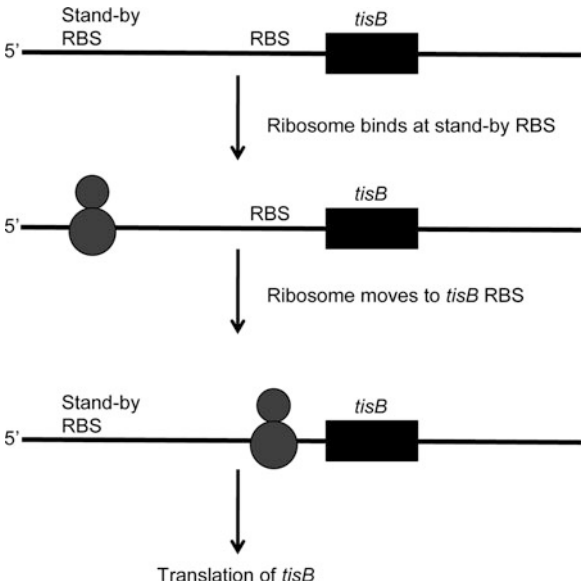


Fig. 3.2 Model describing regulation of translation of *tisB*. Indicated are the "stand-by" ribosome binding site of *tisA* and the true ribosome binding site for *tisB* (see text and Darfeuille et al. 2007 for details)

ysdAB, now known as *tisAB* (Vogel et al. 2004). The *tis* mRNA is composed of two overlapping genes, *tisA* and *tisB*, that encode short proteins. Overexpression of *tisAB* is very toxic to *E. coli*, and this toxicity is shown to be due to *tisB*, and not *tisA* (Vogel et al. 2004). In fact, *tisA* appears not to be translated, and it may function solely to insure translation of *tisB* (Darfeuille et al. 2007). Co-expression of the sRNA IstR-1 can repress TisB toxicity.

Control of *tisB* expression has been thoroughly investigated. Accessibility of the *tisB* ribosome binding site (RBS) is rather limited owing to the secondary structure of the mRNA as well as binding of the antitoxin IstR-1 RNA. However, toe printing experiments revealed that the ribosome could bind far upstream of the *tisB* RBS, and this region mapped to the *tisA* RBS (Fig. 3.2; Darfeuille et al. 2007). Further characterization showed that the *tisA* RBS serves as a "stand-by" ribosome site; the ribosome is unable to bind at the "real" site due to obstruction of the site and instead binds an upstream "stand-by" site (Unoson and Wagner 2007). The ribosome is essentially in a holding position until the obstruction over the true RBS is relieved (in the case of *tisB*, the folded mRNA structure breathes), allowing the ribosome to slide and begin translation at the correct site.

So what is the function of TisB? When overproduced, TisB localizes to the inner membrane of *E. coli*, leading to membrane damage (Unoson and Wagner 2008). Consistently, overproduction of TisB induces transcription of genes that respond to membrane damage (Fozo et al. 2008b). TisB overproduction also leads to reduced replication, transcription, and translation rates but these are probably indirect effects (Unoson and Wagner 2008). Taken together, membrane damage, along with the effects on macromolecule biosynthesis explain how TisB overproduction could cause cell death.

However, what happens when TisB is expressed at endogenous levels, and not from a multicopy plasmid? Transcription of *tisB* is induced by DNA damage, and thus studies have focused on whether TisB plays a role in the SOS response (Vogel et al. 2004). One study examined competition between a wild-type strain and one in which *istR-1* was deleted. This mutant strain has elevated levels of *tisB* compared to the wild-type strain. In response to long-term growth and repeated exposure to the DNA damaging agent mitomycin C, the mutant strain was eventually outcompeted by the wild type (Unoson and Wagner 2008). A study published by a second group reported that overproduction of TisAB could lead to inhibition of an SOS response (Weel-Sneve et al. 2008). Together, these studies suggest that the *tis-istR-1* locus is important for proper responses to DNA damage.

A final study has linked *tisB* expression to persistence (see Chap. 11). The majority of cells treated with the antibiotic ciprofloxacin, which causes DNA damage, are killed; however, a subpopulation known as persisters can survive this treatment. Deletion of *tisB* led to a dramatic reduction in the number of persister cells formed upon ciprofloxacin treatment. How TisB can lead to the formation of persister cells is not clear, but perhaps TisB expression causes slow cell growth, allowing them to survive, and recover from DNA damage.

Taken together from multiple studies, it appears that induction of *tisB* in response to SOS damage is an important component for cell fitness. However, these studies also show that *tisB* levels must be tightly regulated since its expression could be detrimental.

3.5.2 ShoB-OhsC: Another Case of Tight Translation Control

Similar to the TisB-IstR study, a cloning-based screen used to identify novel sRNAs in *E. coli* was instrumental in discovering yet another type I locus, *shoB-ohsC* (Kawano et al. 2005). These authors detected two RNA divergent transcripts encoded by the intergenic region of *yfhL-acpS*. That share a 19 nt region of sequence complementarity (Fig. 3.3). One of the genes, named *shoB*, encodes a putative 26 amino acid hydrophobic peptide while the second gene, *ohsC*, encodes an sRNA. Similarity searches show that this locus is limited to *E. coli* and *Shigella* species, and is thus even more limited in its distribution than the *tisB-istR*-1 locus (Fozo et al. 2008b, 2010; Kawano et al. 2005).

Overproduction of ShoB was very toxic to *E. coli*, but this toxicity could be repressed by overexpression of OhsC-RNA. Additionally, ShoB overproduction led to a rapid decrease in proton motive force, suggesting that cell death was likely due to membrane damage (Fozo et al. 2008b).

The endogenous expression of *shoB* mRNA and its antitoxin OhsC are interesting in that their expression patterns are reciprocal; in cells grown in minimal medium, *shoB* mRNA is readily detected in exponential phase, but not stationary phase. In contrast, the OhsC-RNA is readily detected in cells grown to stationary phase (Kawano et al. 2005). Translational control of ShoB may be related to that

Fig. 3.3 Genetic organization of the divergently transcribed *shoB—ohsC* genes of *E. coli* K-12 MG1655. The *shoB* mRNA is shown as a *gray arrow* pointing *left-ward* while the *ohsC* antisense gene is shown as a *gray arrow* pointing *right-ward*. The *shoB* reading frame is indicated by the *black arrow* and the 19-nucleotide regions of complementarity between *shoB* mRNA and OhsC antisense RNA are indicated by the *white boxes*

of *tisB* (reviewed in Fozo et al. 2008a). Like in the case of *tisB*, there are multiple 5′ ends for the *shoB* transcript. Analysis of translational reporter fusions showed that the full-length *shoB* mRNA is not translated, whereas a 5′-trancated mRNA was, albeit at a low level. Translation of *shoB* was increased in an *ohsC* deletion strain, confirming that OhsC functions to repress *shoB* (Fozo et al. 2008b). But even without OhsC present, translation of the toxin was still quite low. Thus, overall translation of ShoB is highly repressed, and this is independent of the antitoxin. One clue as to how this repression occurs is due to the predicted secondary structure of the *shoB* mRNA: the mRNA is tightly folded so that the ribosome binding site and start codon are sequestered by stable stem-loop structures (Fozo et al. 2008a). The biological function of ShoB is not yet known.

3.5.3 Zor-Orz: The Newest Members

In a search for new type I loci, two hydrophobic, duplicated annotated proteins within the *E. coli* O15:H7 EDL933 genome, *z3289* and *z3290*, were identified (Fozo et al. 2010). These proteins are 29 amino acids in length, encoded in tandem and differ by a single amino acid change. Homologs are found within pathogenic *E. coli* and commensal strains, and *Shigella* species. In some cases, only one of the genes is present. However, neither gene is present in laboratory strains such as *E. coli* K-12 MG1655. Northern analyses confirmed the presence of two sRNAs, encoded antisense and divergent of the annotated proteins (Table 3.1).

Overproduction of either protein was toxic to MG1655. To determine whether the putative antitoxins OrzO (originally denoted as sRNA-1, divergent to *z*3289) and OrzP (originally denoted as sRNA-2, divergent to *z*3290), are indeed antitoxins, a rescue experiment was performed (Wen and Fozo, unpublished data). The sRNAs could indeed repress toxicity associated with overproduction of *zorO* (*z3289*) and *zorP* (*z3290*) (Fig. 3.4).

As with the multiple *ibs-sib* loci, the two *zor-orz* loci are highly similar. This similarity raises the question: do the antitoxins cross-regulate expression of the noncognate toxin? Preliminary experiments indicate that this not the case; only OrzO can repress *zorO* expression (Fig. 3.4). Mutagenesis experiments have narrowed the region required for regulatory specificity and current work is in progress to unravel the exact details for this specificity.

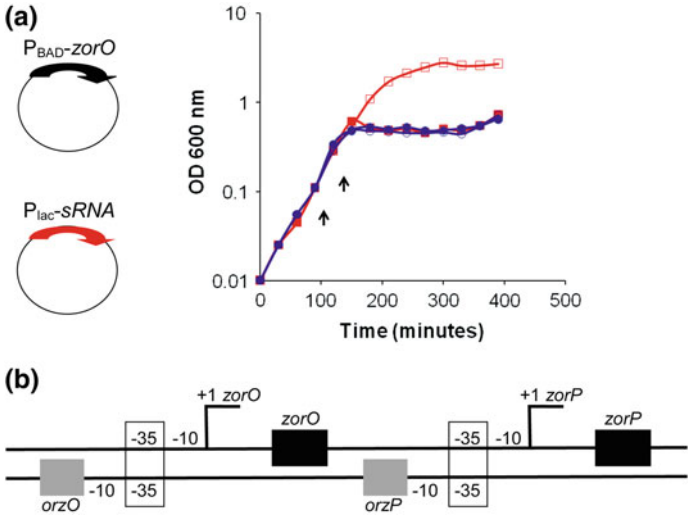


Fig. 3.4 Zors are type I toxins with unusual arrangements. **a** Repression of ZorO-induced toxicity. One plasmid contained *zorO* cloned behind the arabinose-inducible P_{BAD} and the other plasmid contained either *orzO* (in *red*) or *orzP* (in *blue*) cloned behind an IPTG-inducible P_{lac} promoter in *E. coli* MG1655. Addition of 0.2 % arabinose or 1 mM IPTG is indicated by the *arrows*. The closed symbols indicate no IPTG added; the open symbols indicate the addition of IPTG. **b** Schematic of the *zor-orz* locus of *E. coli* O157:H7 EDL933. The toxin open reading frames are indicated by the *black boxes*; the *arrow* indicates the start of toxin transcription. The genes encoding the antitoxins are in *gray*. The –10 and –35 indicate the positioning of the predicted sigma-70 promoter elements. Note that the –35 for *zorO* and *orzO* overlap, as does the –35 for *zorP* and *orzP*

Remarkably, the antitoxins share putative overlapping -35 promoter elements with the cognate toxin –35 promoter element (Fig. 3.4). One could imagine competition for RNA polymerase binding between the divergent but adjacent promoters (Fozo et al. 2010).

3.6 The Big Question: What are Type I TA Loci Doing in Bacteria?

Elucidating the biological role of these genes has been challenging. Unlike the toxins of type II loci, there is no similarity to known enzymatic domains or well-characterized proteins to give guidance for biochemical studies. Type I toxins essentially resemble a transmembrane domain. The speculation has long been that they simply form pores in the membranes of cells.

If they simply form pores in membranes, than why are there so many families and so many duplicated families within one genome? For example, *E. coli* O157:H7 EDL933 has 6 Ibs-Sib, 2 Zor-Orz, 4 Ldr-Rdl, 1 Tis-IstR, and 1 ShoB-

OhsC. The limited data generated show the different toxin families are expressed under different conditions in the cell. To elucidate their function, a variety of approaches must be used. There has been recent success in using large-scale approaches to identify phenotypes of small RNAs and small proteins; similar approaches may be useful for phenotypic characterization of the type I loci. More in depth gene expression studies may also provide useful; they have been extremely beneficial for elucidating potential functions for TisB. Biochemical approaches as well must be considered. Traditionally, work with small proteins has proven challenging, but there has been some recent success using epitope tagged proteins to identify both localization and putative interacting partners (Hemm et al. 2008; Ramamurthi et al. 2006, 2009).

3.7 Finding More

Novel chromosome-encoded type I toxins has been discovered via by identification of sRNAs, by serendipity or through a bioinformatic approach. Why were they missed in the first place? An open reading frame is typically not annotated unless it is larger than 50 codons, thus leaving many small genes unnoticed. Additionally, the type I toxins identified have been found only in γ Proteobacteria and Firmicutes. Does this mean that type I TA loci are limited to these groups? Likely not: given their small sizes and divergent sequences, bioinformatic searches are somewhat limited in predicting homologs. Relatively, few changes in sequence can cause such search algorithms to fail, thus limiting the number of putative homologs that are identified. For example, both BLAST and PSI-BLAST using the Ibs as bait failed to identify the Ibs homologs in *Helicobacter pylori* (Fozo et al. 2010, Sharma et al. 2010). Additionally, traditional BLAST searches were unable to identify that *ef3263* is a member of the *txpA* family of *B. subtilis*.

Identification of the unconventional chromosomal loci is difficult by using bioinformatic tools. Attempts to design algorithms to find similar loci are not trivial, because the identification of limited stretches of similarities is challenging. The ShoB-OhsC locus was identified in a search for sRNAs and then it was noted that there was sequence complementarity between the divergent genes. For the *zor-orz* pairs, the proteins were identified as having properties common to type I toxins, but the Orz sRNAs were identified only through using RNA folding algorithms of the region.

So, do type I loci exist in other species? It is my belief that new families in additional bacterial species do exist and eventually will be identified. For example, RNA sequencing data are now revealing many overlapping, small transcripts in a variety of bacteria and some of these are predicted to be type I loci (Steglich et al. 2008). Our knowledge regarding the novel type I toxin-antitoxin loci is still in its infancy. However, given the recent advances in identification and functional characterization, it is likely that the function of many known, and yet to be discovered loci, will be revealed in the upcoming years.

References

Argaman, L., Hershberg, R., Vogel, J., Bejerano, G., Wagner, E. G., Margalit, H., et al. (2001). Novel small RNA-encoding genes in the intergenic regions of *Escherichia coli*. *Current Biology, 11*, 941–950.

Darfeuille, F., Unoson, C., Vogel, J., & Wagner, E. G. (2007). An antisense RNA inhibits translation by competing with standby ribosomes. *Molecular Cell, 26*, 381–392.

Faridani, O. R., Nikravesh, A., Pandey, D. P., Gerdes, K., & Good, L. (2006). Competitive inhibition of natural antisense Sok-RNA interactions activates Hok-mediated cell killing in *Escherichia coli*. *Nucleic Acids Research, 34*, 5915–5922.

Fozo, E. M., Hemm, M. R., & Storz, G. (2008a). Small toxic proteins and the antisense RNAs that repress them. *Microbiology and Molecular Biology Reviews, 72*, 579–589.

Fozo, E. M., Kawano, M., Fontaine, F., Kaya, Y., Mendieta, K. S., Jones, K. L., et al. (2008b). Repression of small toxic protein synthesis by the sib and OhsC small RNAs. *Molecular Microbiology, 70*, 1076–1093.

Fozo, E. M., Makarova, K. S., Shabalina, S. A., Yutin, N., Koonin, E. V., & Storz, G. (2010). Abundance of type I toxin-antitoxin systems in bacteria: searches for new candidates and discovery of novel families. *Nucleic Acids Research, 38*, 3743–3759.

Gerdes, K., Thisted, T., & Martinussen, J. (1990). Mechanism of post-segregational killing by the *hok/sok* system of plasmid R1: Sok antisense RNA regulates formation of a *hok* mRNA species correlated with killing of plasmid-free cells. *Molecular Microbiology, 4*, 1807–1818.

Gerdes, K., & Wagner, E. G. (2007). RNA antitoxins. *Curr Opin Microbiol, 10*, 117–124.

Greenfield, T. J., Ehli, E., Kirshenmann, T., Franch, T., Gerdes, K., & Weaver, K. E. (2000). The antisense RNA of the *par* locus of pAD1 regulates the expression of a 33-amino-acid toxic peptide by an unusual mechanism. *Molecular Microbiology, 37*, 652–660.

Greenfield, T. J., & Weaver, K. E. (2000). Antisense RNA regulation of the pAD1 *par* post-segregational killing system requires interaction at the 5′ and 3′ ends of the RNAs. *Molecular Microbiology, 37*, 661–670.

Han, K., Kim, K. S., Bak, G., Park, H., & Lee, Y. (2010). Recognition and discrimination of target mRNAs by Sib RNAs, a *cis*-encoded sRNA family. *Nucleic Acids Research, 38*, 5851–5866.

Hemm, M. R., Paul, B. J., Schneider, T. D., Storz, G., & Rudd, K. E. (2008). Small membrane proteins found by comparative genomics and ribosome binding site models. *Molecular Microbiology, 70*, 1487–1501.

Hershberg, R., Altuvia, S., & Margalit, H. (2003). A survey of small RNA-encoding genes in *Escherichia coli*. *Nucleic Acids Research, 31*, 1813–1820.

Hobbs, E. C., Astarita, J. L., & Storz, G. (2010). Small RNAs and small proteins involved in resistance to cell envelope stress and acid shock in *Escherichia coli*: Analysis of a bar-coded mutant collection. *Journal of Bacteriology, 192*, 59–67.

Irnov, I., Sharma, C. M., Vogel, J., & Winkler, W. C. (2010). Identification of regulatory RNAs in *Bacillus subtilis*. *Nucleic Acids Research, 38*, 6637–6651.

Jahn, N., Preis, H., Wiedemann, C., & Brantl, S. (2012). BsrG/SR4 from *Bacillus subtilis*–the first temperature-dependent type I toxin-antitoxin system. *Molecular Microbiology, 83*, 579–598.

Kawano, M., Oshima, T., Kasai, H., & Mori, H. (2002). Molecular characterization of long direct repeat (LDR) sequences expressing a stable mRNA encoding for a 35-amino-acid cell-killing peptide and a cis-encoded small antisense RNA in *Escherichia coli*. *Molecular Microbiology, 45*, 333–349.

Kawano, M., Reynolds, A. A., Miranda-Rios, J., & Storz, G. (2005). Detection of 5′- and 3′-UTR-derived small RNAs and *cis*-encoded antisense RNAs in *Escherichia coli*. *Nucleic Acids Research, 33*, 1040–1050.

Mok, W. W., Patel, N. H., & Li, Y. (2010). Decoding toxicity: Deducing the sequence requirements of IbsC, a type I toxin in *Escherichia coli*. *Journal of Biological Chemistry, 285*, 41627–41636.

Pedersen, K., & Gerdes, K. (1999). Multiple *hok* genes on the chromosome of *Escherichia coli*. *Molecular Microbiology, 32*, 1090–1102.

Ramamurthi, K. S., Clapham, K. R., & Losick, R. (2006). Peptide anchoring spore coat assembly to the outer forespore membrane in *Bacillus subtilis*. *Molecular Microbiology, 62*, 1547–1557.

Ramamurthi, K. S., Lecuyer, S., Stone, H. A., & Losick, R. (2009). Geometric cue for protein localization in a bacterium. *Science, 323*, 1354–1357.

Rivas, E. R., Klein, J., Jones, T. A., & Eddy, S. R. (2001). Computational identification of noncoding RNAs in *E. coli* by comparative genomics. *Current Biology, 11*, 1369–1373.

Rudd, K. E. (1999). Novel intergenic repeats of *Escherichia coli* K-12. *Research in Microbiology, 150*, 653–664.

Saito, S., Kakeshita, H., & Nakamura, K. (2009). Novel small RNA-encoding genes in the intergenic regions of *Bacillus subtilis*. *Gene, 428*, 2–8.

Sharma, C. M., Hoffmann, S., Darfeuille, F., Reignier, J., Findeiss, S., Sittka, A., et al. (2010). The primary transcriptome of the major human pathogen *Helicobacter pylori*. *Nature, 464*, 250–255.

Silvaggi, J. M., Perkins, J. B., & Losick, R. (2005). Small untranslated RNA antitoxin in *Bacillus subtilis*. *Journal of Bacteriology, 187*, 6641–6650.

Steglich, C., Futschik, M. E., Lindell, D., Voss, B., Chisholm, S. W., & Hess, W. R. (2008). The challenge of regulation in a minimal photoautotroph: Non-coding RNAs in *Prochlorococcus*. *PLoS Genetics, 4*, e1000173.

Thisted, T., & Gerdes, K. (1992). Mechanism of post-segregational killing by the *hok/sok* system of plasmid R1. Sok antisense RNA regulates *hok* gene expression indirectly through the overlapping *mok* gene. *Journal of Molecular Biology, 223*, 41–54.

Unoson, C., & Wagner, E. G. (2007). Dealing with stable structures at ribosome binding sites: Bacterial translation and ribosome standby. *RNA Biology, 4*, 113–117.

Unoson, C., & Wagner, E. G. (2008). A small SOS-induced toxin is targeted against the inner membrane in *Escherichia coli*. *Molecular Microbiology, 70*, 258–270.

Vogel, J., Argaman, L., Wagner, E. G., & Altuvia, S. (2004). The small RNA IstR inhibits synthesis of an SOS-induced toxic peptide. *Current Biology, 14*, 2271–2276.

Wassarman, K. M., Repoila, F., Rosenow, C., Storz, G., & Gottesman, S. (2001). Identification of novel small RNAs using comparative genomics and microarrays. *Genes and Development, 15*, 1637–1651.

Waters, L. S., & Storz, G. (2009). Regulatory RNAs in bacteria. *Cell, 136*, 615–628.

Weaver, K. E., Jensen, K. D., Colwell, A., & Sriram, S. I. (1996). Functional analysis of the *Enterococcus faecalis* plasmid pAD1-encoded stability determinant *par*. *Molecular Microbiology, 20*, 53–63.

Weaver, K. E., Reddy, S. G., Brinkman, C. L., Patel, S., Bayles, K. W., & Endres, J. L. (2009). Identification and characterization of a family of toxin-antitoxin systems related to the *Enterococcus faecalis* plasmid pAD1 *par* addiction module. *Microbiology, 155*, 2930–2940.

Weel-Sneve, R., Bjoras, M., & Kristiansen, K. I. (2008). Overexpression of the LexA-regulated *tisAB* RNA in E. coli inhibits SOS functions; implications for regulation of the SOS response. *Nucleic Acids Research, 36*, 6249–6259.

Chapter 4
Type II Toxin-Antitoxin Loci: The *ccdAB* and *parDE* Families

Marie Deghorain, Nathalie Goeders, Thomas Jové and Laurence Van Melderen

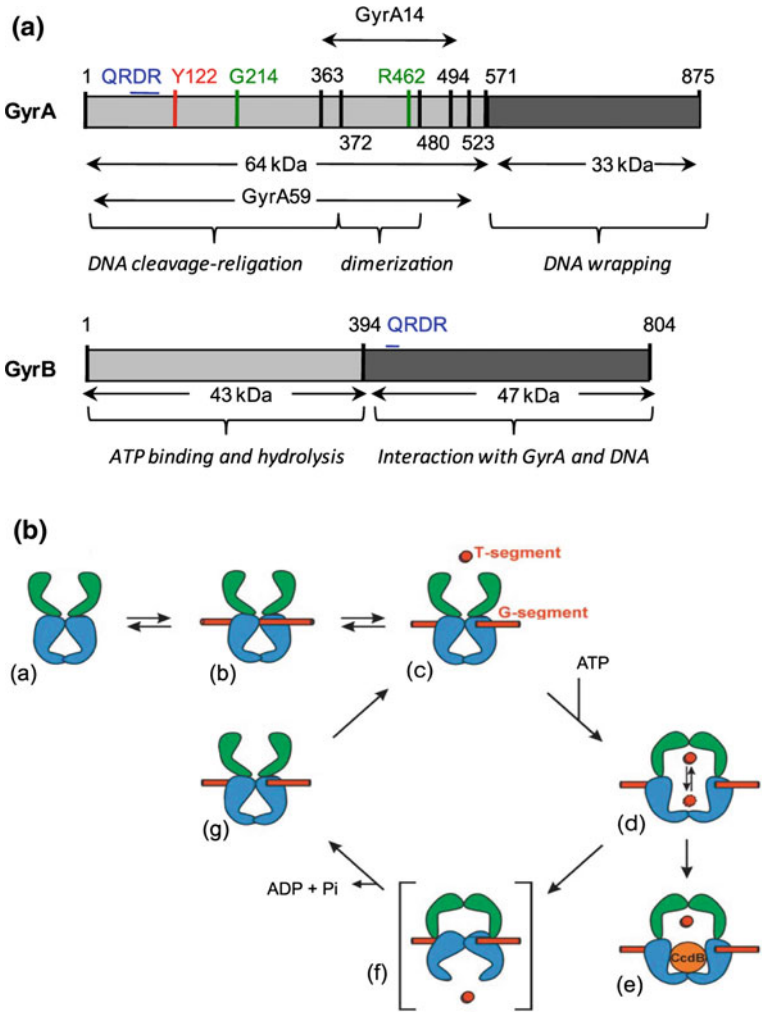
Abstract The ccd_F and $parDE_{RK2}$ systems were originally identified on the *E. coli* F plasmid and on the RK2 broad-host-range plasmid, respectively. Early reports indicated that these two loci promote plasmid maintenance in growing population by killing daughter bacteria that did not inherit a plasmid copy at cell division by a mechanism termed 'post-segregational killing'. The toxins encoded by these two loci target DNA-gyrase, an essential topoisomerase II, although the molecular details appear to be different. Interestingly, $CcdB_F$ and $ParE_{RK2}$ do not share a common ancestor, and therefore appear to have evolved independently. CcdB and ParE protein families share structural similarity with toxins belonging to the MazF and RelE toxin families, respectively, members of which inhibit translation by different mechanisms. In addition to focus on the molecular mechanism of poisoning of DNA-gyrase by $CcdB_F$ and $ParE_{RK2}$ as well as structural data, we describe here the different systems encoding CcdB-like and ParE-like toxins and highlight the specific features of those systems in terms of antitoxin structure and properties, autoregulation, biological functions, and evolution.

4.1 Introduction

The ccd_F locus is encoded by the *Escherichia coli* F plasmid and was the first type II system to be identified (Ogura and Hiraga 1983). The $parDE_{RK2}$ locus of broad-host-range plasmid RK2 was identified a few years later (Roberts et al. 1990). The

M. Deghorain · N. Goeders · T. Jové · L. Van Melderen (✉)
Faculté des Sciences, Laboratoire de génétique et physiologie bactérienne, IBMM, Université Libre de Bruxelles, 12 rue des Professeurs Brachet et Jeener, 6041 Gosselies, Belgique
e-mail: lvmelder@ulb.ac.be

K. Gerdes (ed.), *Prokaryotic Toxin-Antitoxins*,
DOI: 10.1007/978-3-642-33253-1_4,



Ccd_F and $ParE_{RK2}$ toxins represent so far the only examples of type II toxins targeting DNA-gyrase. In the canonical systems, they are linked the $CcdA_F$ and $ParD_{RK2}$ antitoxins, respectively. More recently, homologs of $ParE_{RK2}$ have been found associated with different antitoxins, highlighting the 'mix and match' between toxin and antitoxins super-families (Leplae et al. 2011; Hayes and Van Melderen 2011). In this chapter, we will therefore deal with the toxins and the antitoxins separately.

◀ **Fig. 4.1** *E. coli* DNA-gyrase, domains and catalytic cycle. **a** Domain organization of GyrA and GyrB subunits. GyrA and GyrB subunits are organized in structural domains. The N-terminal domain (NTD) of GyrA (59-64 kDa) includes the DNA cleavage religation (residues 1–363) and the dimerization (residues 372–480) regions. It contains the catalytic tyrosine 122 (Y122), and encompasses the GyrA59 region (residues 2–523). The GyrA C-terminal domain (CTD, 33 kDa) is a DNA-binding domain responsible for DNA wrapping. The CTD of GyrB (43 kDa) interacts with DNA and GyrA subunits. The GyrB NTD (47 kDa) is involved in ATP binding and hydrolysis. Resistance to quinolones is conferred by mutations in QRDR (Quinolones Resistance Determining Region, *blue bars*) in GyrA (residues 84–106) or in GyrB (residues 426–447), while resistance to $CcdB_F$ is conferred by mutations of the G214 and R462 residues in GyrA. The minimal region for $CcdB_F$ binding covers residues 363–494 (GyrA14). For both GyrA and GyrB subunits, NTD and CTD are indicated in *light gray* and in *dark gray*, respectively. Numbers indicate amino acids. Adapted from (Van Melderen, 2002). **b** Molecular model for the catalytic cycle of DNA-gyrase and its poisoning by $CcdB_F$ GyrA is shown in *blue*, GyrB in *green*, DNA in *red* and $CcdB_F$ in *orange*. (**a**) DNA-gyrase is in its resting state in the absence of DNA. The GyrB clamp is open. (**b**) A first segment of DNA (the G segment) can bind reversibly to the resting state of the enzyme. (**c**) The G segment is cleaved and both parts remain covalently attached to the active site tyrosine residues Tyr122 of GyrA. This is the so-called cleavable complex. A second segment of DNA (the T segment) can enter the GyrB clamp. (**d**) Upon ATP binding, the GyrB clamp closes. Closing of the GyrB clamp pushes the conformational equilibrium of the GyrA dimer to its open conformation, and the T segment is allowed to cross the G gate, thus adding two turns of negative supercoil. (**e**) The $CcdB_F$ toxin enters the cleavable complex before the T-segment passage. After binding, the G gate cannot close due to sterical hindrance and the cleavable complex is trapped. (**f**) In absence of $CcdB_F$, the primary exit gate of GyrA transiently opens and closes to release the T segment. (**g**) After or during release of the T segment, ATP is hydrolyzed and the enzyme returns to its resting state. At this point, another round of supercoiling can start, or alternatively the G segment can be religated and the supercoiled DNA released from the enzyme. Figure and legend taken from (Dao-Thi et al. 2005)

4.2 The CcdB and ParE Toxins Inhibit DNA-gyrase, a Bacterial Type II Topoisomerase

4.2.1 DNA-gyrase Structure and Catalytic Cycle

The DNA-gyrase (EC 5.99.1.3) is one of the four topoisomerases found in bacteria. Topoisomerases are essential enzymes that regulate DNA topology and prevent chromosome entanglements (Collin et al. 2011). Topoisomerases catalyze DNA supercoiling and relaxation and introduce and remove catenanes and knots. These reactions require the passage of a DNA segment across a transient single-strand or double-strand DNA break. Topoisomerase classification relies on the type of breaks these enzymes introduce into their DNA substrates, type I introducing single-strand breaks (Topo I and Topo III) and type II double-strand breaks (DSB) (DNA-gyrase and Topo IV). Unlike other bacterial topoisomerase, DNA gyrase introduces negative supercoils in an ATP-dependent reaction (Gellert et al. 1976). As a consequence, upon DNA-gyrase inhibition, relaxation of circular DNA molecules occurs. The negative supercoiling activity of DNA-gyrase is essential during replication and transcription processes to counteract the positive

supercoiling resulting from DNA unwinding ahead of the polymerases and thereby prevents 'topological stress' that would lead to inhibition of these essential processes (Nollmann et al. 2007). Bacterial promoters are differentially sensitive to the chromosome supercoiling status, which reflects an indirect function of DNA-gyrase in regulation of gene expression (Travers and Muskhelishvili 2005). A minor contribution of DNA-gyrase to post-replication chromosome decatenation has also been shown (Zechiedrich and Cozzarelli 1995).

DNA-gyrase of *E. coli* (~350 kDa) is composed of two subunits, GyrA and GyrB, which form a 'heart-shaped' $GyrA_2$-$GyrB_2$ hetero-tetramer. The GyrA subunit (875 aa, 97 kDa) is responsible for DNA binding, double-strand DNA-cleavage and religation reactions while the GyrB subunit (804 aa, 90 kDa) provides energy for the reaction by ATP hydrolysis (Fig. 4.1 a). The amino-terminal domain (GyrA59, residues 2–523) of the GyrA subunit has been crystallized as a dimer (Fig. 4.1a, Loris et al. 1999; Morais Cabral et al. 1997). It displays a large central cavity closed at the 'top' by the 'DNA gate' formed by the proximal amino-terminal subdomain. This subdomain contains the catalytic residue tyrosine 122 (Fig. 4.1a), which covalently binds to the 5'-end of the cleaved DNA. At the 'bottom' of the Gyr59 dimer, the GyrA dimerization domain constitutes the 'exit gate' that closes the $GyrA_2$ cavity. During its catalytic cycle, the DNA-gyrase transiently cleaves DNA, passes another DNA segment through the DSB, religates the DSB and releases the transferred DNA through the 'exit gate' (Fig. 4.1b) (Roca and Wang 1994). The relaxation process occurs by the reverse mechanism without ATP hydrolysis.

4.2.2 Toxins Inhibiting DNA-gyrase

DNA-gyrase is essential and constitutes a pivotal target for antibacterial molecules [for a recent review, see (Collin et al. 2011)]. Numerous DNA-gyrase inhibitors have been described so far, with the quinolone antibiotics being the most famous. Quinolone resistant strains isolated from clinical settings harbor mutations in a restricted region of the *gyrA* gene located close to the catalytic tyrosine 122 and known as QRDR (quinolone resistance determining region; Fig. 4.1a). Quinolones trap the DNA*DNA-gyrase complex in an open conformation and inhibit the religation step (Drlica et al. 2009). These ternary DNA*DNA-gyrase*quinolone complexes constitute a barrier for replication fork progression and generate DSBs by an as yet unknown mechanism. DSBs trigger the SOS response (Box 1) and accumulation of DSBs leads to chromosome fragmentation and eventually cell death. In addition, reactive oxygen species (ROS) are produced and induce lethal damages in lipids, proteins, and DNA (Dwyer et al. 2007). In addition to quinolones, the well-characterized type II $CcdB_F$ and $ParE_{RK2}$ toxins also target DNA-gyrase.

4.2.3 The CcdB Toxin: Structure and DNA-gyrase Inhibition Mechanism

The $CcdB_F$ monomer is a 101 aa protein (11.7 kDa) (Bex et al. 1983) consisting of an amino-terminal β-sheet of five anti-parallel strands followed by a 13 residues long α-helix plus a three-strand sheet ('wing sheet') inserted between the third and the fourth β-sheet strands (Loris et al. 1999) (Fig. 4.2a, b). The active $CcdB_F$ toxin is a homo-dimer and constitutes a globular protein with a hydrophobic core comprised between the helix and the β-sheets. The three last carboxy-terminal aa of $CcdB_F$ (Trp99-Gly100-Ile101) are involved in $CcdB_F$ toxicity (Bahassi et al. 1995) and extend out of the compact structure (Fig. 4.2b). The 'wing' sheets are extruded out of the dimeric structure (Fig. 4.2a and b) and are involved in the binding of the $CcdA_F$ antitoxin.

To determine the cellular target of the $CcdB_F$ toxin, the group of Martine Couturier isolated $CcdB_F$-resistant *E. coli* mutants and mapped the mutation to the *gyrA* gene (*gyrA462*), showing that $CcdB_F$ targets the DNA-gyrase. Using in vitro DNA-gyrase assays, it was shown that $CcdB_F$ inhibits the supercoiling activity of DNA-gyrase and binds to the free GyrA form (Maki et al. 1996, 1992). In addition to this inhibitory activity, $CcdB_F$ converts DNA-gyrase into a poison that generate DSBs in substrate DNA, indicating that, like quinolones, $CcdB_F$ poisons a post-cleavage intermediate of DNA-gyrase (Bernard et al. 1993). However, quinolones interact differently with the DNA-gyrase as the GyrA462 mutant is not resistant to quinolones (neither to $ParE_{RK2}$, see below) and vice versa, mutations in the QRDR do not confer resistance to $CcdB_F$ (Bernard and Couturier 1992; Critchlow et al. 1997). The arginine at position 462, located within the GyrA dimerization domain (Fig. 4.1a), is the key residue for $CcdB_F$-mediated inhibition of DNA-gyrase as it directly interacts with the tryptophan 99 of $CcdB_F$ (Fig. 4.2b) (Dao-Thi et al. 2005, 2004) and overexpression of a GyrA fragment (GyrA14, Fig. 4.1a) protects *E. coli* from $CcdB_F$-mediated killing (Dao-Thi et al. 2005). The arginine 462 residue is accessible to $CcdB_F$ when DNA-gyrase is cycling on DNA, at the DNA passage step. In addition, several cycles of ATP hydrolysis and presumably strand passage events are required for efficient $CcdB_F$ poisoning (Scheirer and Higgins 1997), suggesting that the step targeted by $CcdB_F$ in the catalytic cycle is short lived.

Therefore, the model proposed for DNA-gyrase poisoning by $CcdB_F$ is that it targets the open conformational state, prior to DNA passage and inhibits the religation step (Fig. 4.1b). Formation of CcdB*DNA*DNA-gyrase ternary complexes leads to cell death as explained above. Quinolones stabilize an earlier intermediate of the catalytic cycle but lead to a comparable phenotype [SOS induction and cell death (Smith and Maxwell 2006)].

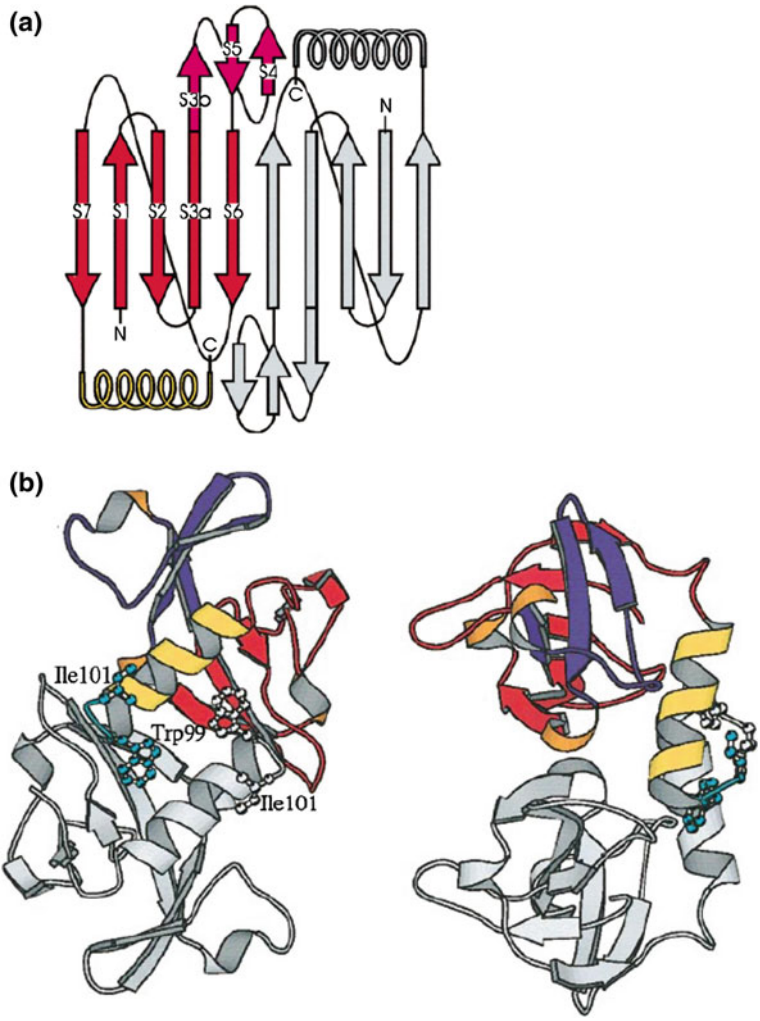


Fig. 4.2 Three-dimensional structure of the F plasmid encoded $CcdB_F$ toxin. **a** Topology diagram of the $CcdB_F$ dimer. For one monomer, secondary structural elements are labeled and color coded. The main five-stranded antiparallel β-sheet is shown in *red*, the smaller three-stranded wing sheet in *purple*. The C-terminal α-helix is colored *yellow*. The second monomer is shown in *gray*. **b** Ribbon representation of the $CcdB_F$ dimer in two different orientations, roughly 90 ° apart. The same color code as used in **a**. depicts secondary structure elements of one monomer. The other monomer is shown in *gray*. The three isolated turns of α310-helix are colored *ochre*. The three C-terminal residues (Trp99, Gly100 and Ile101), mutations of which suppress the killer phenotype, are colored *light blue* and are shown as ball-and-stick model. Figure and legend are adapted from (Loris et al. 1999)

4.2.4 The ParE Toxin: Structure and DNA-gyrase Inhibition Mechanism

The *parDE*$_{RK2}$ system was identified as a locus promoting stabilization of the broad-host-range RK2 plasmid (Roberts et al. 1994). It is composed of the ParE$_{RK2}$ toxin (103 aa, 12 kDa) and the ParD$_{RK2}$ antitoxin (Johnson et al. 1996). Using a replication thermosensitive mutant plasmid, the group of Helinski showed that the loss of the plasmid at nonpermissive temperature inhibited cell growth and induced cell filamentation. Later on, they showed that ParE$_{RK2}$ inhibited DNA-gyrase-mediated supercoiling and induced DSBs in vitro (Jiang et al. 2002). Further evidences supporting that DNA-gyrase is the ParE target were obtained from the ParE homologs in *E. coli* O157:H7 (Hallez et al. 2010) and in *Vibrio cholerae* (Yuan et al. 2011). These ParE$_{RK2}$ homologs induced the SOS system, co-localized with the nucleoid (*E. coli* O157:H7 homolog) and induced DSBs in an in vitro DNA-gyrase assay (*V. cholerae* homologs) (Yuan et al. 2010). However, the molecular mechanism for ParE$_{RK2}$-mediated poisoning appears to be quite distinct to that of CcdB$_F$, e.g., mutations in GyrA that confer resistance to quinolones or CcdB$_F$ do not confer resistance to ParE$_{RK2}$ and ParE$_{RK2}$ does not inhibit DNA-gyrase mediated relaxation (Yuan et al. 2010). In addition, ParE$_{RK2}$ and CcdB$_F$ do not share similar fold and belong to two unrelated super-families (Leplae et al. 2011), supporting the idea that they might act on DNA-gyrase by different molecular mechanisms. The three-dimensional structure of ParE$_{RK2}$ was solved recently and it consists in a dimer in which each monomer contains two α-helices followed by a three-stranded antiparallel β-sheet, the last being crucial for ParE toxicity (Dalton and Crosson 2010).

4.2.5 Properties of the Antitoxins Associated with CcdB and ParE Toxins

Toxins belonging to a given super-family can be associated to antitoxins that belong to different super-families (Leplae et al. 2011). The CcdB$_F$-like toxins constitute an exception to this observation, as these toxins are exclusively associated to CcdA$_F$ antitoxins. For the ParE$_{RK2}$-like toxins, the situation is different and these toxins are predicted to be associated to at least four different super-families (see below).

Common features are found in the different super-families of antitoxins described so far. Most of them are composed of two structurally and functionally distinct domains, i.e., a N-terminal region that contains a DNA-binding domain and a C-terminal domain that is the toxin-binding domain. Some exceptions are now described; one example will be detailed below in which the antitoxin is devoid of any known DNA-binding domain. In addition, in the cases that have been studied, antitoxins are less stable than toxins. This differential stability is responsible for toxin activation.

4.2.6 The CcdA Antitoxin

In addition to the well-characterized plasmidic ccd_F system, chromosomal *ccd* systems from *E. coli* O157:H7 and *Erwinia chrysanthemi* as well a *ccd* system from *Vibrio fischeri* super integron were also experimentally validated (Wilbaux et al. 2007; Saavedra De Bast et al. 2008; Rowe-Magnus et al. 2003). As mentioned above, these CcdB toxins are all associated to antitoxins of the CcdA superfamily.

Resolution of the $CcdA_F$ structure by NMR revealed that $CcdA_F$ is a dimer in solution and when bound to DNA. Dimerization occurs via its N-terminal region folded as a ribbon-helix-helix (RHH) DNA-binding domain and leads to formation of an anti-parallel β-sheet. As for other RHH regulators, insertion of this positively charged β-sheet into the major groove of DNA leads to the protein-DNA complex (Madl et al. 2006). The C-terminal domain of $CcdA_F$ interacts with the CcdB toxin. A truncated form of $CcdA_F$ (CcdA41) lacking the 31 N-terminal amino acids is able to protect cells from the toxic activity of $CcdB_F$ (Bernard and Couturier 1991; Salmon et al. 1994). These data show that the two domains act independently. In contrast to the N-terminal domain, the C-terminal region is intrinsically disordered and becomes structured upon binding to $CcdB_F$ (Madl et al. 2006). Due to its unstructured nature, the C-terminal region of $CcdA_F$ plays a key role in essential processes. First, it is likely that this disordered domain is responsible for the short half-life of the antitoxin, and thus is implicated in the ccd_F system activation. Second, it is essential for the ability of the antitoxin to dissociate the $CcdB_F$*DNA-gyrase complex (rejuvenation). Finally, it allows $CcdA_F$ to interact with two different binding sites on $CcdB_F$ with distinct affinities, and thus participates to the fine tuning of autoregulation (De Jonge et al. 2009).

4.2.7 CcdA is Degraded by the ATP-Dependent Lon Protease

The metabolic stability of two antitoxins belonging to the CcdA super-family, the F plasmid $CcdA_F$ and the *E. coli* O157:H7 $CcdA_{O157}$, was assessed in vivo. Both antitoxins are degraded by the ATP-dependent Lon protease (Van Melderen et al. 1994, 1996; Wilbaux et al. 2007). Their half-lives were estimated at 30 min in the absence of $CcdB_F$. $CcdA_F$ is stabilized up to 60 min in the presence of CcdB (Van Melderen et al. 1994). This was confirmed for $CcdA_F$ by in vitro experiments. $CcdA_F$ proteolytic degradation is probably favored by its intrinsically disordered C-terminal domain, which exposes hydrophobic residues recognized by Lon, whereas folding of CcdA consequently to CcdB binding probably masks or renders these residues inaccessible (Van Melderen et al. 1996).

4.2.8 $CcdA_F$ Rejuvenates the $CcdB_F$*DNA-gyrase Complex

$CcdA_F$ has the ability to reverse the toxic activity of $CcdB_F$ by dissociating the $CcdB_F$*DNA-gyrase complex. This property relies on the flexibility of the C-terminal region and its ability to cover two partially overlapping binding sites on the $CcdB_F$ dimers (Kampranis et al. 1999; Dao-Thi et al. 2002; De Jonge et al. 2009). $CcdA_F$ binds first a high affinity site on $CcdB_F$ dimers bound to DNA-gyrase. This interaction induces a conformational change in the toxin dimer thereby disrupting the $CcdB_F$*DNA-gyrase complex. This allows then $CcdA_F$ to zip around the toxin once it is freed and leads to the formation of the $CcdA_F$*$CcdB_F$ complex in which both affinity sites are occupied. Rejuvenation and toxin neutralization are possible even at low antitoxin concentrations owing to the toxin–antitoxin high affinity (Dao-Thi et al. 2002; De Jonge et al. 2009).

4.2.9 Antitoxins Associated with ParE Toxins

ParE toxins that were experimentally validated are associated with four different super-families of antitoxins, the ParD, PasA, Phd, and RelB super-families (Leplae et al. 2011; Hallez et al. 2010; Fiebig et al. 2010; Yuan et al. 2011). In the canonical $parDE_{RK2}$ system, the $ParE_{RK2}$ toxin is associated to the $ParD_{RK2}$ antitoxin, belonging to the ParD super-family. This super-family has a small number of representative sequences (Leplae et al. 2011). The $ParD_{RK2}$ antitoxin is composed of an amino-terminal RHH-fold involved in DNA binding and dimerization (Oberer et al. 2007). The C-terminal domain is intrinsically disordered as for the $CcdA_F$ antitoxin (Oberer et al. 2007).

Recently, three systems containing a ParE-like toxin were validated in *E. coli* O157:H7 (Hallez et al. 2010). Two of the systems show a particular genetic organization and are composed of three components, one of them being a transcriptional regulator (PaaR). The antitoxins (PaaA) belong to the RelB super-family but are devoid of any known-DNA binding domain (Leplae et al. 2011). It was shown that the PaaA2 antitoxin is unusually unstable as compared to other antitoxins such as $CcdA_F$. The half-life of PaaA2 is estimated at 6 min in a wild-type strain and the ClpAP, ClpXP, and Lon proteases are involved in its degradation (Hallez et al. 2010). In the third system, the antitoxin ($ParD_{EDL933}$) belongs to the PasA super-family, which is unrelated to the ParD super-family.

In *Caulobacter crescentus*, four *parDE* systems have been recently described (Fiebig et al. 2010). Among these, two antitoxins belong to the ParD super-family, whereas the two other toxins were found to be associated to different super-families of antitoxins (Leplae et al. 2011). The antitoxin associated to ParE3 does not share any similarities with ParD as reported by Fiebig et al. and belongs to a novel antitoxin family (Leplae et al. 2011; our unpublished data). The ParE1 toxin is associated to a PasA antitoxin (Leplae et al. 2011; our unpublished data). Interestingly, this antitoxin is stably folded in solution (Dalton and Crosson 2010).

Furthermore, the toxin annotated as $RelE_2$ by Fiebig et al. (2010) was shown experimentally to induce the SOS response in *E. coli*, suggesting that this toxin is also a ParE toxin (Leplae et al. 2011). The associated antitoxin does not show any similarities neither with ParD nor other antitoxins and represents a novel family of antitoxins (Leplae et al. 2011; our unpublished data). Thus, in *C. crescentus*, five toxins homologous to $ParE_{RK2}$ have been validated and show association with different antitoxin super-families.

Additional examples of non-ParD antitoxins associated with ParE toxins were also found in *Vibrio cholerae* using a bioinformatics analysis (Leplae et al. 2011; our unpublished data). Antitoxins annotated as $ParD_1$ and $ParD_3$ were similar to Phd, while the annotated $ParD_2$ antitoxin did not share any similarities with known antitoxins.

4.3 Autoregulation of the *ccdAB* and *parDE* operons

A classical feature of type II TA systems is their organization in operon, as well as the ability of the TA complex to negatively regulate its own expression at the level of transcription. As discussed above, in general, the antitoxin carries a DNA-binding domain and is responsible for DNA binding. The toxin usually acts as a co-repressor.

4.3.1 Autoregulation of the ccd *Operon*

The antitoxin:toxin ratio controls transcription of the *ccdAB* operon. Depending on the ratios (1:1 or 1:2), $CcdA_F$ and $CcdB_F$ form hetero complexes of different stoechiometry that rely on the two CcdA binding sites present on $CcdB_F$ dimers (Dao-Thi et al. 2002; De Jonge et al. 2009). Transcriptional repression is maximal in the presence of the hetero-tetrameric $(CcdA_F)_2$-$(CcdB_F)_2$ complex. Three operators are described in the promoter region of the ccd_F operon, a main 6-bp palindromic sequence (5'-GTATAC-3') and two lower affinity binding sites (5'-TATA-3'). Dnase I footprinting showed that $CcdA_F$ alone gives no clear protection patterns, whereas the $(CcdA_F)_2$-$(CcdB_F)_2$ complexes form cooperative chains that cover up to 120 bp in the operator/promoter region. These alternating $CcdA_F$-$CcdB_F$ chains are formed through the interactions of the two C-termini of $CcdA_F$ in the dimer with a different $CcdB_F$ dimer (De Jonge et al. 2009). The authors proposed that $CcdA_F$ alone is not able to fully repress the operon expression owing to a weak $CcdA_F$-operator complex with high off and on rates (Dao-Thi et al. 2002). When $CcdB_F$ is in excess, for instance in case of excessive antitoxin degradation, a hexameric $(CcdA_F)_2$-$(CcdB_F)_4$ complex is formed which is unable to interact with DNA (Afif et al. 2001). In this condition, if transcription and translation are functional, synthesis of $CcdA_F$ and $CcdB_F$ will occur and re-established a proper $CcdA_F$:$CcdB_F$ ratio. This phenomenon, co-repression at a

1:1 ratio and de-repression at higher toxin levels, was termed conditional cooperatively (Overgaard et al. 2008) and is common in type II TA loci (De Jonge et al. 2009). Similarly to the ccd_F system, the chromosomally encoded ccd_{O157} system is regulated by $CcdA_{O157}$ and $CcdB_{O157}$. However, the activity of the ccd_{O157} promoter is 5-fold lower as compared to the plasmidic one and repression is less effective (Wilbaux et al. 2007).

4.3.2 *Autoregulation of the* parE-*Encoding Operons*

Roberts et al. showed that $ParD_{RK2}$ alone is sufficient to repress its promoter, in contrast to most antitoxins (Roberts et al. 1993). $ParD_{RK2}$ binds a 33-bp site containing a direct and an inverted repeat in the promoter region of the $parDE_{RK2}$ system. DNase I footprints showed that $ParD_{RK2}$ protects a 48-bp region near the promoter (Roberts et al. 1993). Similarly, transcription reporter fusions showed that the antitoxins associated to $ParE_1$ and $ParE_2$ of *V. cholerae* were sufficient to repress transcription from their respective promoters (Yuan et al. 2011). Regulation of the three-component *paaR2–paaA2–parE2* system of *E. coli* O157:H7 appears to be different. β-galactosidase assays showed that autoregulation is dependent on the PaaA2–ParE2 TA complex and on the predicted PaaR2 transcriptional regulator. Molecular details are not elucidated yet, however, complete repression requires the three components (Hallez et al. 2010).

4.4 Biological Functions of the *ccd* and *parE*-Containing Systems

4.4.1 *Plasmid-Encoded* ccd_F *and* $parDE_{RK2}$ *Systems: The Post-Segregational Killing Phenomenon*

As mentioned above, ccd_F was discovered on the F plasmid and $parDE_{RK2}$ on RK2. Despite their large sizes (95 and 60 kb, respectively) and their low copy number (from 1 copy per chromosomal equivalent for F to 5–8 copies for RK2), these plasmids are maintained in growing bacterial populations with a remarkable stability. It is believed that the ccd_F and $parDE_{RK2}$ loci contribute to plasmid maintenance, in parallel with active segregation (*par*) and multimer resolution systems (Austin et al. 1981). TA loci act by a mechanism called post-segregational killing (PSK) (Gerdes et al. 1986b; Jaffe et al. 1985) or addiction (Yarmolinsky 1995) (Fig. 4.3a). The molecular basis of PSK relies on the differential stability of the toxin and antitoxin (Tsuchimoto et al. 1992; Van Melderen et al. 1994). When a plasmid is not transmitted to a daughter cell, the antitoxin is rapidly degraded and the antitoxin pool is not replenished. The toxin is therefore released from its

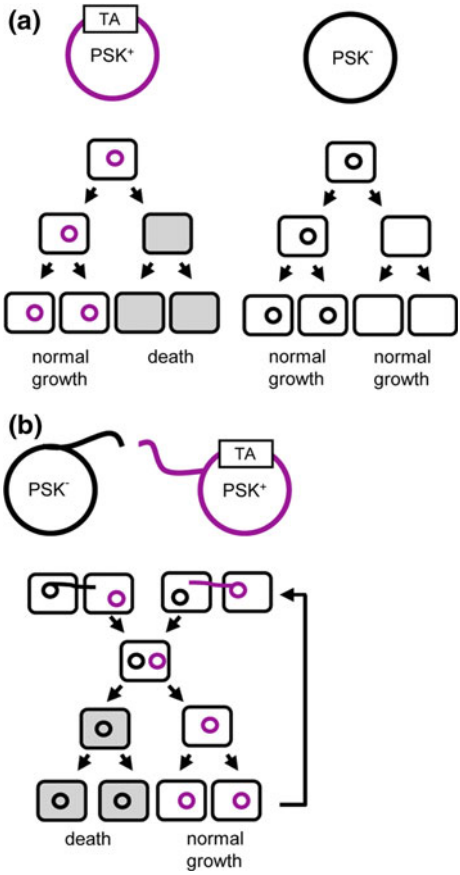


Fig. 4.3 Advantage conferred by plasmid-encoded TA systems. **a** Vertical transmission. TA systems increase plasmid prevalence in growing bacterial populations by post-segregational killing (PSK). PSK^+ plasmid is shown in *purple*, *left panel*. Daughter bacteria that inherit a plasmid copy at cell division grow normally. If daughter bacteria do not inherit a plasmid copy, degradation of the labile antitoxin proteins by the host ATP-dependent proteases will liberate the stable toxin. This will lead to the selective killing of the plasmid-free bacteria (*in gray*). When considering only vertical transmission, TA systems increase the prevalence of the plasmid in the population as compared with plasmids devoid of TA systems (PSK^- plasmid in *black*, *right panel*). **b** Horizontal transmission. Plasmid–plasmid competition. The PSK^+ plasmid (*in purple*) and the PSK^- plasmid (*in black*) belong to the same incompatibility group and are conjugative. Under conditions in which conjugation occurs, conjugants containing both plasmids are generated. Because the two plasmids are incompatible, they cannot be maintained in the same bacteria. The 'loss' of the PSK^+ plasmid will lead to the killing of bacteria containing the PSK^- plasmid through the PSK mechanism (*in gray*), thereby outcompeting the PSK2 plasmid. On the contrary, the loss of the PSK^- plasmid will be without any deleterious effect on the PSK^+ plasmid. Through multiple events of conjugation, the fitness of the PSK^+ plasmid will be increased (*arrow*). Figure and legend taken from (Van Melderen and Saavedra De Bast 2009)

inhibition by the antitoxin, and may eventually kill the plasmid-free cell. TA loci involved in PSK have been described for a number of plasmids, including virulence plasmids, in a wide range of bacterial species, e.g., *hok*/*sok* and *kis-kid* on the R1 plasmid (Gerdes et al. 1986a, 1986b; Munoz-Gomez et al. 2005), *phd-doc* on the P1 bacteriophage (Lehnherr et al. 1993) and *mvpA-mvpT* (or vapBC) on a *Shigella flexneri* virulence plasmid (Sayeed et al. 2000).

On the F plasmid, the ccd_F system coexists with two additional systems belonging to the *hok*/*sok* family of type I TA systems [*flm* (Loh et al. 1988) and *srn* (Golub and Panzer 1988)]. These three systems work independently and differ in their PSK efficiency, the ccd_F system being considered as the poorest killer (Jensen and Gerdes 1995, Jensen et al. 1995). Interestingly, a consequence of the CcdB toxin activity is the activation of the SOS system (see above). Since the SOS response involves mutagenesis, it was proposed that a consequence of the CcdB toxicity might be the generation of genetic diversity in a small fraction of the bacterial population, which might represent an advantage for adaptation under specific conditions (Couturier et al. 1998). In addition, one cannot exclude that the ccd_F system might also play a physiological role in plasmid-containing cells under specific conditions that compromise the $CcdA_F$:$CcdB_F$ ratio. It was reported that the Ccd_F proteins production increases in plasmid-containing cells during stationary phase compared to exponentially growing cells. Interestingly, this was correlated to an increase of the mutation rate, suggesting a possible $CcdB_F$-mediated mutagenesis in resting cells (Aguirre-Ramirez et al. 2006).

Killing of plasmid-free daughter cell increases the vertical stability of plasmids by increasing the prevalence of cells carrying the plasmid in a population (Jaffe et al. 1985; Gerdes et al. 1986a). However, the presence of a type II TA system does not increase the effective frequency of vertical inheritance and the stability hypothesis fails to explain the apparent success of type II systems on plasmid (Cooper and Heinemann 2000). The 'competition' model proposed by Cooper and Heinemann (2000) provides an attractive hypothesis for the selection and maintenance of plasmid-encoded PSK systems during evolution (Fig. 4.3b). In a competition experiment, they showed that a conjugative plasmid carrying a PSK system (PSK^+) is able to outcompete a conjugative plasmid of the same incompatibility group but devoid of a PSK system (PSK^-). Therefore, by eliminating competitor plasmids in the bacterial progeny, the presence of a PSK system on a plasmid increases its relative fitness in the population. Indeed, when a cell inherits two incompatible PSK^+ and PSK^- plasmids, the loss of the PSK^+ plasmid leads to killing the cells containing the PSK^- plasmid through the PSK mechanism while the loss of the PSK^- plasmid has no deleterious effect on the PSK^+ plasmid.

4.4.2 Chromosomally Encoded ccd *and* parDE *Loci*

Bacterial genomes sequencing revealed the presence of numerous and various TA genes in bacterial chromosomes, raising the important question of their biological

function(s) (Pandey and Gerdes 2005; Leplae et al. 2011). Several chromosomally encoded systems related to the ccd_F system or the $parDE_{RK2}$ systems have been studied. They originate from *E. coli* O157:H7 (ccd_{O157}; Wilbaux et al. 2007; Mine et al. 2009; Hallez et al. 2010), *E. chrysanthemi* (ccd_{Ech}; Saavedra De Bast et al. 2008), *C. crescentus* (*parE*-containing systems; Fiebig et al. 2010) and *V. cholerae* (*parE*-conaining systems; Yuan et al. 2010, 2011). In contrast to plasmid-encoded TA loci that play a pivotal role in plasmid maintenance, a common and unitary function has not been assigned to the chromosomally encoded *ccd* and *parDE* systems. Their physiological roles, if any, are diverse, and seem dependent of several factors such as host species properties, genomic location, age, or any other factor that may influence the evolution of TA systems (Van Melderen and Saavedra De Bast 2009).

4.4.2.1 The Anti-Addiction Model

The anti-addiction model relies on interactions between plasmid- and chromosomally encoded systems (Saavedra De Bast et al. 2008). Following this model, chromosomally encoded TA systems might play a protective role against PSK-mediated by homologous plasmid-encoded systems as long as their antitoxins are able to counteract the toxic activity of plasmid-encoded systems. This type of interactions was evidenced for the ccd_{Ech} located in the *E. chrysanthemi* chromosome and the ccd_F system of the F plasmid (Saavedra De Bast et al. 2008). The $CcdA_{Ech}$ is able to counteract the toxic activity of the $CcdB_F$ toxin, and the ccd_{Ech} system is able to protect the cells against ccd_F-mediated PSK. In addition, it was found by competition experiments that a gain of fitness is conferred by ccd_{Ech} under PSK conditions. Therefore, by acting as anti-addiction systems, chromosomally encoded TA modules might serve as defense mechanisms against invading DNA (plasmid or phages), similarly to restriction modification and CRISPR systems (Barrangou et al. 2007; Kobayashi 2001; Tock and Dryden 2005).

The positive effect on fitness by the anti-addiction systems may have ensured their conservation in bacterial chromosomes, up to a certain point of evolution. This phenomenon might also have favored the selection of plasmid-encoded systems variants, which are no longer antagonized by anti-addiction modules. Harmonious coexistence of chromosomally and plasmid-encoded *ccd* modules has been observed in *E. coli* O157:H7 (Wilbaux et al. 2007). In this case, the ccd_{O157} system encoded on the chromosome is not able to counteract the toxic activity of the closely related ccd_F system found in the pO157 natural virulence plasmid, even in overproduction conditions (Wilbaux et al. 2007). The chromosomally encoded ccd_{O157} system appears thus to have lost its anti-addiction properties and might be devoid of any current physiological function. Consistently, decay of the $ccdB_{0157}$ gene was observed in many natural isolates of *E. coli* (Mine et al. 2009).

4.4.2.2 The Genome Stabilization Model

Another possible role for chromosomally encoded TA systems might be the stabilization of large regions and the prevention against large-scale genome reductions. TA systems found in bacterial chromosomes are often embedded within prophages or other mobile genetic elements such as superintegrons (SIs). As observed for plasmids, the addiction properties of TA modules can act against the loss of chromosomal regions (Rowe-Magnus et al. 2003; Szekeres et al. 2007). Hence, the combination of $parDE_1$ and $relBE_1$ TA systems from the *Vibrio vulnificus* SI was shown to reduce large DNA deletions in different genomic contexts (Szekeres et al. 2007). Similarly, the 3 ParE-containing systems encoded within the SI of the *Vibrio cholerae* chromosome II might act to maintain the integrity of the SI, and therefore to prevent the loss of the chromosome II. However, the 3 ParE-containing systems contributed to DNA degradation occurring in cells lacking chromosome II, but not to the other phenotype traits observed in these cells (for instance, growth defect or morphological aberrations) (Yuan et al. 2010, 2011). It is therefore possible that these systems function as a part of a more complex process for chromosome II stabilization, due to their ability to mediate a PSK-like phenomenon as described for plasmid maintenance.

4.4.2.3 Chromosomally Encoded *ccd* and *parDE* Loci of *C. crescentus*

ParE-containing systems have also been identified in the *C. crescentus* chromosome (Fiebig et al. 2010; Leplae et al. 2011; see above). These TA systems are not essential for *C. crescentus* viability in the laboratory conditions that were tested (Fiebig et al. 2010). All the ParE-containing systems are functionally independent, as crosstalk was not observed between toxins and antitoxins belonging to different systems. In addition, they appear differentially regulated by distinct environmental conditions suggesting that they might be part of the host regulatory networks and that their regulation might be more complex than that of their counterpart on plasmids. (Fiebig et al. 2010). Therefore, the genetic context in which TA modules are inserted might influence their physiological function by affecting their expression levels. In the same line of idea, systems composed of three components, such as the *paaR2-paaA2-parE2* system of *E. coli* O157:H7 (see above) in which the transcriptional regulation is carried out by the PaaR2, might constitute other examples of integration.

4.4.2.4 The Selfish Hypothesis

TA modules may be considered as selfish entities (Van Melderen and Saavedra De Bast 2009). Indeed, toxin and antitoxin genes are linked together by a strong interdependence. They move within genomes by horizontal transfer and are able to maintain themselves in bacterial population at the expense of their host cell, at least when they are located on plasmid. Hence, one cannot exclude that TA

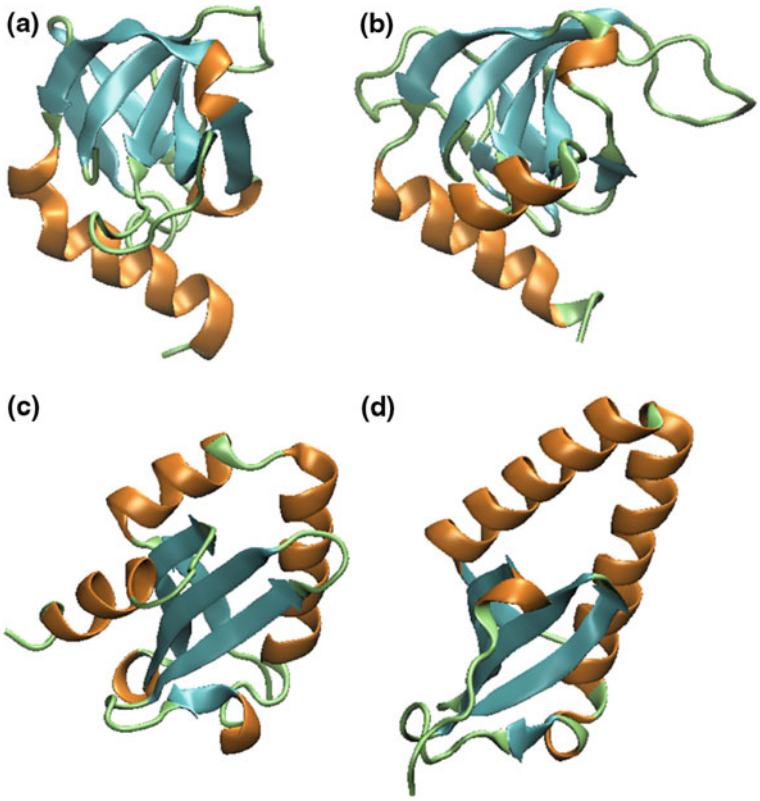


Fig. 4.4 Structures of toxins of the MazF/Kid/CcdB super-family and the RelE/ParE super-family $CcdB_F$ monomer (**a**; PDB ID: 4VUB) and Kid monomer (**b**; PDB ID: 1M1F) belong to the same super-family. RelE (**c**; PDB ID: 3KXE) and ParE (**d**; PDB ID: 2KC9) share a similar fold and belong to the RelE/ParE super-family

systems might act for stabilization of the DNA molecules that carry them, such as plasmids, mobile genetic elements (SI, phages,..), or chromosomes. The physiological functions described above might thus be viewed as a consequence of TA systems addiction properties.

4.5 Evolution of $CcdB_F$ and $ParE_{RK2}$ Toxins

TA systems were initially classified in families based on sequence similarities between the toxins (Jorgensen et al. 2009; Pandey and Gerdes 2005). Structural and bioinformatics data showed that some toxin families sharing very low or even undetectable primary sequence similarities and/or targeting different cellular processes are phylogenetically related and are grouped in super-families (Leplae et al. 2011). Resolution of the Kid toxins structure for instance unraveled its

similarity with CcdB despite low sequence similarity and led to the definition of the MazF/CcdB super-family (Hargreaves et al. 2002; Kamada et al. 2003). Likewise, bioinformatics approaches showed that ParE and RelE toxins are distantly related (Anantharaman and Aravind 2003; Guglielmini et al. 2008). Resolution of their structures later on confirmed this as they share a common fold (Dalton and Crosson 2010). These two super-families are, however, unrelated to the nine other toxin super-families as defined in Leplae et al. 2011.

The CcdB/MazF and ParE/RelE super-families are quite abundant in bacterial genomes and appear to have a broad phyletic distribution (Leplae et al. 2011). In addition, these super-families appear to be chromosomally encoded preferentially; however, there is most likely a strong bias since much more genome sequences are available than plasmidic ones (Leplae et al. 2011).

4.5.1 The CcdB/MazF Super-Family

The first toxin structure to be resolved was that of $CcdB_F$ (Loris et al. 1999). Subsequent resolution of the structures of MazF and Kid (Kamada et al. 2003; Hargreaves et al. 2002) unraveled their structural similarity with $CcdB_F$ despite a very low sequence similarity (Fig. 4.4a and b). Interestingly, these two toxins, although sharing 3-D structures, target different physiological processes. While $CcdB_F$ targets DNA-gyrase, MazF and Kid are endoribonucleases. Furthermore, $CcdB_F$ lacks residues implicated in mRNA cleavage and, reciprocally, the residues of the C-terminal helix important for $CcdB_F$ activity are missing in Kid (Hargreaves et al. 2002). ToxN which belongs to a type III system, also shares the CcdB/MazF/Kid-fold (Blower et al. 2011). As Kid, ToxN has an RNase activity but different residues are implicated in their toxic activity and they have different specificities (Diago-Navarro et al. 2009; Kamphuis et al. 2006; Blower et al. 2009). Overlay of their structures showed that the catalytic residues of both toxins are located in the same pocket, which leads to the hypothesis that the difference in residues probably accounts for their different specificities (Blower et al. 2011).

4.5.2 The ParE/RelE Super-Family

RelE and $ParE_{RK2}$ also target different processes despite their similar fold (Fig. 4.4c and d), as RelE cleaves mRNAs in a translation-dependent manner, whereas $ParE_{RK2}$ toxin targets DNA-gyrase. Similarly to the MazF/CcdB super-family, the different activities appear to lie in different regions of the toxins as $ParE_1$ of *C. crescentus* and $ParE_{RK2}$ lack the catalytic residues important for RelE toxic activity (Dalton and Crosson 2010; Neubauer et al. 2009).

4.6 Applications Based on the $CcdB_F$ Toxin

The $CcdB_F$ protein was used to construct positive-selection cloning vectors relying on the specific disruption of the toxic *ccdB*$_F$ gene (Bernard et al. 1994; Gabant et al. 1998). This allows the recovery of only recombinant colonies and significantly improves the efficiency of the cloning reaction. These cloning vectors can be amplified and prepared in a GyrA462 mutant, since the gyrA462 mutation confers resistance to $CcdB_F$ (Bernard and Couturier 1992; Bernard et al. 1994). This $CcdB_F$-based positive selection was used by Invitrogen™ (Carlsbad, CA) for developing the Gateway® technology. This method involves a cloning technique based on the lambda site-specific recombination system (Papagiannis et al. 2007) and consists in a two-step procedure. In a first step, a PCR product is directly cloned into an entry vector, and is transferred in a second step to a destination vector provided with *ccdB*. This results in the replacement of the toxic *ccdB*$_F$ gene by the DNA fragment of interest and the positive selection of recombinant clones.

The discovery of $CcdB_F$ and its interaction with DNA-gyrase also opened up the possibility to generate novel inhibitors of bacterial topoisomerase. Bacterial resistance to antibiotics is continuously emerging and constitutes a critical problem for public health that needs to be overcome. Recently, a synthetic peptide derived from CcdB and able to inhibit simultaneously DNA-gyrase and topoisomerase IV was obtained, providing a potent new class of antibacterial agents (Trovatti et al. 2008).

Box 4.1: The SOS response

The SOS response is a regulatory network triggered in response to DNA damages that otherwise would compromise DNA replication and ultimately lead to cell death (Erill et al. 2007). A large panel of exogenous and endogenous events causes DNA damages, such as oxidative stress, UV radiation, low pH, antibiotics, or toxins from TA systems such as CcdB and ParE. Horizontal gene transfer events might constitute another source of SOS-inducing ssDNA (e.g., conjugation or IS transposition) (Baharoglu et al. 2010; Lane et al. 1994). DNA damages induce replication fork stalling which in turns generate DSBs. The RecBCD and RecFOR complexes initiate DSBs repair and produce single-strand DNA (ssDNA). This ssDNA constitutes the induction signal for the SOS response. The RecA recombinase binds to ssDNA and forms nucleofilaments that activate its co-protease activity (RecA*; Cox 2007). RecA* promotes self-cleavage of the LexA transcriptional repressor, thereby inducing the expression of about 40 genes in *E. coli* (referred as the "SOS regulon"). Expression of these genes promote three essential roles for the SOS response: (1) transient inhibition of cell division and DNA repair; (2) genome evolution through the expression of error-prone DNA polymerases (Galhardo et al. 2009) and mobilization of mobile genetic elements such as prophages, ICE, transposons, or integron gene cassettes (Rokney et al. 2008; Beaber et al. 2004; Kuan and Tessman 1992; Guerin et al. 2009) and; (3) dormancy of a small subpopulation (persisters cells). Interestingly, this latter

mechanism appears to be mediated by type I SOS-induced TA systems (Dorr et al. 2010; Wang and Wood 2011).

Acknowledgments Work in the laboratory of LVM is supported by the Fonds de la Recherche (FRSM-3.4530.04), Fondation Van Buuren and Fonds Jean Brachet. The authors report no conflicts of interest.

References

Afif, H., Allali, N., Couturier, M., & Van Melderen, L. (2001). The ratio between CcdA and CcdB modulates the transcriptional repression of the ccd poison-antidote system. *Molecular Microbiology, 41*, 73–82.

Aguirre-Ramirez, M., Ramirez-Santos, J., Van Melderen, L., & Gomez-Eichelmann, M. C. (2006). Expression of the F plasmid ccd toxin-antitoxin system in *Escherichia coli* cells under nutritional stress. *Canadian Journal of Microbiology, 52*, 24–30.

Anantharaman, V., & Aravind, L. (2003). New connections in the prokaryotic toxin-antitoxin network: Relationship with the eukaryotic nonsense-mediated RNA decay system. *Genome Biology, 4*, R81.

Austin, S., Ziese, M., & Sternberg, N. (1981). A novel role for site-specific recombination in maintenance of bacterial replicons. *Cell, 25*, 729–736.

Baharoglu, Z., Bikard, D., & Mazel, D. (2010). Conjugative DNA transfer induces the bacterial SOS response and promotes antibiotic resistance development through integron activation. *PLoS Genetics, 6*, e1001165.

Bahassi, E. M., Salmon, M. A., Van Melderen, L., Bernard, P., & Couturier, M. (1995). F plasmid CcdB killer protein: ccdB gene mutants coding for non-cytotoxic proteins which retain their regulatory functions. *Molecular Microbiology, 15*, 1031–1037.

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., et al. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. *Science, 315*, 1709–1712.

Beaber, J. W., Hochhut, B., & Waldor, M. K. (2004). SOS response promotes horizontal dissemination of antibiotic resistance genes. *Nature, 427*, 72–74.

Bernard, P., & Couturier, M. (1991). The 41 carboxy-terminal residues of the miniF plasmid CcdA protein are sufficient to antagonize the killer activity of the CcdB protein. *Molecular and General Genetics, 226*, 297–304.

Bernard, P., & Couturier, M. (1992). Cell killing by the F plasmid CcdB protein involves poisoning of DNA-topoisomerase II complexes. *Journal of Molecular Biology, 226*, 735–745.

Bernard, P., Kezdy, K. E., Van Melderen, L., Steyaert, J., Wyns, L., Pato, M. L., et al. (1993). The F plasmid CcdB protein induces efficient ATP-dependent DNA cleavage by gyrase. *Journal of Molecular Biology, 234*, 534–541.

Bernard, P., Gabant, P., Bahassi, E. M., & Couturier, M. (1994). Positive-selection vectors using the F plasmid ccdB killer gene. *Gene, 148*, 71–74.

Bex, F., Karoui, H., Rokeach, L., Dreze, P., Garcia, L., & Couturier, M. (1983). Mini-F encoded proteins: Identification of a new 10.5 kilodalton species. *EMBO Journal, 2*, 1853–1861.

Blower, T. R., Fineran, P. C., Johnson, M. J., Toth, I. K., Humphreys, D. P., & Salmond, G. P. (2009). Mutagenesis and functional characterization of the RNA and protein components of the toxIN abortive infection and toxin-antitoxin locus of Erwinia. *Journal of Bacteriology, 191*, 6029–6039.

Blower, T. R., Pei, X. Y., Short, F. L., Fineran, P. C., Humphreys, D. P., Luisi, B. F., et al. (2011). A processed noncoding RNA regulates an altruistic bacterial antiviral system. *Nature Structural and Molecular Biology, 18*, 185–190.

Collin, F., Karkare, S., & Maxwell, A. (2011). Exploiting bacterial DNA gyrase as a drug target: Current state and perspectives. *Applied Microbiology and Biotechnology, 92*, 479–497.

Cooper, T. F., & Heinemann, J. A. (2000). Postsegregational killing does not increase plasmid stability but acts to mediate the exclusion of competing plasmids. *Proceedings of National Academy Science USA, 97*, 12643–12648.

Couturier, M., Bahassi, E. M., & Van Melderen, L. (1998). Bacterial death by DNA gyrase poisoning. *Trends in Microbiology, 6*, 269–275.

Cox, M. M. (2007). Regulation of bacterial RecA protein function. *Critical Reviews in Biochemistry and Molecular Biology, 42*, 41–63.

Critchlow, S. E., O'Dea, M. H., Howells, A. J., Couturier, M., Gellert, M., & Maxwell, A. (1997). The interaction of the F plasmid killer protein, CcdB, with DNA gyrase: Induction of DNA cleavage and blocking of transcription. *Journal of Molecular Biology, 273*, 826–839.

Dalton, K. M., & Crosson, S. (2010). A conserved mode of protein recognition and binding in a ParD-ParE toxin-antitoxin complex. *Biochemistry, 49*, 2205–2215.

Dao-Thi, M. H., Charlier, D., Loris, R., Maes, D., Messens, J., Wyns, L., et al. (2002). Intricate interactions within the ccd plasmid addiction system. *Journal of Biological Chemistry, 277*, 3733–3742.

Dao-Thi, M. H., Van Melderen, L., De Genst, E., Buts, L., Ranquin, A., Wyns, L., et al. (2004). Crystallization of CcdB in complex with a GyrA fragment. *Acta Crystallographica. Section D, Biological Crystallography, 60*, 1132–1134.

Dao-Thi, M. H., Van Melderen, L., De Genst, E., Afif, H., Buts, L., Wyns, L., et al. (2005). Molecular basis of gyrase poisoning by the addiction toxin CcdB. *Journal of Molecular Biology, 348*, 1091–1102.

De Jonge, N., Garcia-Pino, A., Buts, L., Haesaerts, S., Charlier, D., Zangger, K., et al. (2009). Rejuvenation of CcdB-poisoned gyrase by an intrinsically disordered protein domain. *Molecular Cell, 35*, 154–163.

Diago-Navarro, E., Kamphuis, M. B., Boelens, R., Barendregt, A., Heck, A. J., van den Heuvel, R. H., et al. (2009). A mutagenic analysis of the RNase mechanism of the bacterial kid toxin by mass spectrometry. *FEBS Journal, 276*, 4973–4986.

Dorr, T., Vulic, M., & Lewis, K. (2010). Ciprofloxacin causes persister formation by inducing the TisB toxin in *Escherichia coli*. *PLoS Biology, 8*, e1000317.

Drlica, K., Hiasa, H., Kerns, R., Malik, M., Mustaev, A., & Zhao, X. (2009). Quinolones: Action and resistance updated. *Current Topics in Medicinal Chemistry, 9*, 981–998.

Dwyer, D. J., Kohanski, M. A., Hayete, B., & Collins, J. J. (2007). Gyrase inhibitors induce an oxidative damage cellular death pathway in *Escherichia coli*. *Molecular Systems Biology, 3*, 91.

Erill, I., Campoy, S., & Barbe, J. (2007). Aeons of distress: An evolutionary perspective on the bacterial SOS response. *FEMS Microbiology Reviews, 31*, 637–656.

Fiebig, A., Castro Rojas, C. M., Siegal-Gaskins, D., & Crosson, S. (2010). Interaction specificity, toxicity and regulation of a paralogous set of ParE/RelE-family toxin-antitoxin systems. *Molecular Microbiology, 77*, 236–251.

Gabant, P., Szpirer, C. Y., Couturier, M., Faelen, M. (1998). Direct selection cloning vectors adapted to the genetic analysis of gram-negative bacteria and their plasmids. *Gene, 207*, 87–92

Galhardo, R. S., Do, R., Yamada, M., Friedberg, E. C., Hastings, P. J., Nohmi, T., et al. (2009). DinB upregulation is the sole role of the SOS response in stress-induced mutagenesis in *Escherichia coli*. *Genetics, 182*, 55–68.

Gellert, M., Mizuuchi, K., O'Dea, M. H., & Nash, H. A. (1976). DNA gyrase: An enzyme that introduces superhelical turns into DNA. *Proceedings of National Academy Science USA, 73*, 3872–3876.

Gerdes, K., Bech, F. W., Jorgensen, S. T., Lobner-Olesen, A., Rasmussen, P. B., Atlung, T., et al. (1986a). Mechanism of postsegregational killing by the hok gene product of the parB system of plasmid R1 and its homology with the relF gene product of the *E. coli* relB operon. *EMBO Journal, 5*, 2023–2029.

Gerdes, K., Rasmussen, P. B., & Molin, S. (1986b). Unique type of plasmid maintenance function: Postsegregational killing of plasmid-free cells. *Proceedings of National Academy Science USA, 83*, 3116–3120.

Golub, E. I., & Panzer, H. A. (1988). The F factor of *Escherichia coli* carries a locus of stable plasmid inheritance stm, similar to the parB locus of plasmid RI. *Molecular and General Genetics, 214*, 353–357.

Guerin, E., Cambray, G., Sanchez-Alberola, N., Campoy, S., Erill, I., Da Re, S., et al. (2009). The SOS response controls integron recombination. *Science, 324*, 1034.

Guglielmini, J., Szpirer, C., & Milinkovitch, M. C. (2008). Automated discovery and phylogenetic analysis of new toxin-antitoxin systems. *BMC Microbiology, 8*, 104.

Hallez, R., Geeraerts, D., Sterckx, Y., Mine, N., Loris, R., & Van Melderen, L. (2010). New toxins homologous to ParE belonging to three-component toxin-antitoxin systems in *Escherichia coli* O157:H7. *Molecular Microbiology, 76*, 719–732.

Hargreaves, D., Santos-Sierra, S., Giraldo, R., Sabariegos-Jareno, R., de la Cueva-Mendez, G., Boelens, R., et al. (2002). Structural and functional analysis of the kid toxin protein from *E. coli* plasmid R1. *Structure, 10*, 1425–1433.

Hayes, F., & Van Melderen, L. (2011). Toxins-antitoxins: Diversity, evolution and function. *Critical Reviews in Biochemistry and Molecular Biology, 46*, 386–408.

Jaffe, A., Ogura, T., & Hiraga, S. (1985). Effects of the ccd function of the F plasmid on bacterial growth. *Journal of Bacteriology, 163*, 841–849.

Jensen, R. B., & Gerdes, K. (1995). Programmed cell death in bacteria: Proteic plasmid stabilization systems. *Molecular Microbiology, 17*, 205–210.

Jensen, R. B., Grohmann, E., Schwab, H., Diaz-Orejas, R., & Gerdes, K. (1995). Comparison of ccd of F, parDE of RP4, and parD of R1 using a novel conditional replication control system of plasmid R1. *Molecular Microbiology, 17*, 211–220.

Jiang, Y., Pogliano, J., Helinski, D. R., & Konieczny, I. (2002). ParE toxin encoded by the broad-host-range plasmid RK2 is an inhibitor of *Escherichia coli* gyrase. *Molecular Microbiology, 44*, 971–979.

Johnson, E. P., Strom, A. R., & Helinski, D. R. (1996). Plasmid RK2 toxin protein ParE: Purification and interaction with the ParD antitoxin protein. *Journal of Bacteriology, 178*, 1420–1429.

Jorgensen, M. G., Pandey, D. P., Jaskolska, M., & Gerdes, K. (2009). HicA of *Escherichia coli* defines a novel family of translation-independent mRNA interferases in bacteria and archaea. *Journal of Bacteriology, 191*, 1191–1199.

Kamada, K., Hanaoka, F., & Burley, S. K. (2003). Crystal structure of the MazE/MazF complex: Molecular bases of antidote-toxin recognition. *Molecular Cell, 11*, 875–884.

Kamphuis, M. B., Bonvin, A. M., Monti, M. C., Lemonnier, M., Munoz-Gomez, A., van den Heuvel, R. H., et al. (2006). Model for RNA binding and the catalytic site of the RNase Kid of the bacterial parD toxin-antitoxin system. *Journal of Molecular Biology, 357*, 115–126.

Kampranis, S. C., Howells, A. J., & Maxwell, A. (1999). The interaction of DNA gyrase with the bacterial toxin CcdB: Evidence for the existence of two gyrase-CcdB complexes. *Journal of Molecular Biology, 293*, 733–744.

Kobayashi, I. (2001). Behavior of restriction-modification systems as selfish mobile elements and their impact on genome evolution. *Nucleic Acids Research, 29*, 3742–3756.

Kuan, C. T., & Tessman, I. (1992). Further evidence that transposition of Tn5 in *Escherichia coli* is strongly enhanced by constitutively activated RecA proteins. *Journal of Bacteriology, 174*, 6872–6877.

Lane, D., Cavaille, J., & Chandler, M. (1994). Induction of the SOS response by IS1 transposase. *Journal of Molecular Biology, 242*, 339–350.

Lehnherr, H., Maguin, E., Jafri, S., & Yarmolinsky, M. B. (1993). Plasmid addiction genes of bacteriophage P1: Doc, which causes cell death on curing of prophage, and phd, which prevents host death when prophage is retained. *Journal of Molecular Biology, 233*, 414–428.

Leplae, R., Geeraerts, D., Hallez, R., Guglielmini, J., Dreze, P., & Van Melderen, L. (2011). Diversity of bacterial type II toxin-antitoxin systems: A comprehensive search and functional analysis of novel families. *Nucleic Acids Research, 39*, 5513–5525.

Loh, S. M., Cram, D. S., & Skurray, R. A. (1988). Nucleotide sequence and transcriptional analysis of a third function (Flm) involved in F-plasmid maintenance. *Gene, 66*, 259–268.

Loris, R., Dao-Thi, M. H., Bahassi, E. M., Van Melderen, L., Poortmans, F., Liddington, R., et al. (1999). Crystal structure of CcdB, a topoisomerase poison from *E. coli*. *Journal of Molecular Biology, 285*, 1667–1677.

Madl, T., Van Melderen, L., Mine, N., Respondek, M., Oberer, M., Keller, W., et al. (2006). Structural basis for nucleic acid and toxin recognition of the bacterial antitoxin CcdA. *Journal of Molecular Biology, 364*, 170–185.

Maki, S., Takiguchi, S., Miki, T., & Horiuchi, T. (1992). Modulation of DNA supercoiling activity of *Escherichia coli* DNA gyrase by F plasmid proteins. Antagonistic actions of LetA (CcdA) and LetD (CcdB) proteins. *Journal of Biological Chemistry, 267*, 12244–12251.

Maki, S., Takiguchi, S., Horiuchi, T., Sekimizu, K., & Miki, T. (1996). Partner switching mechanisms in inactivation and rejuvenation of *Escherichia coli* DNA gyrase by F plasmid proteins LetD (CcdB) and LetA (CcdA). *Journal of Molecular Biology, 256*, 473–482.

Mine, N., Guglielmini, J., Wilbaux, M., & Van Melderen, L. (2009). The decay of the chromosomally encoded ccdO157 toxin-antitoxin system in the *Escherichia coli* species. *Genetics, 181*, 1557–1566.

Morais Cabral, J. H., Jackson, A. P., Smith, C. V., Shikotra, N., Maxwell, A., & Liddington, R. C. (1997). Crystal structure of the breakage-reunion domain of DNA gyrase. *Nature, 388*, 903–906.

Munoz-Gomez, A. J., Lemonnier, M., Santos-Sierra, S., Berzal-Herranz, A., & Diaz-Orejas, R. (2005). RNase/anti-RNase activities of the bacterial parD toxin-antitoxin system. *Journal of Bacteriology, 187*, 3151–3157.

Neubauer, C., Gao, Y. G., Andersen, K. R., Dunham, C. M., Kelley, A. C., Hentschel, J., et al. (2009). The structural basis for mRNA recognition and cleavage by the ribosome-dependent endonuclease RelE. *Cell, 139*, 1084–1095.

Nollmann, M., Crisona, N. J., & Arimondo, P. B. (2007). Thirty years of *Escherichia coli* DNA gyrase: From in vivo function to single-molecule mechanism. *Biochimie, 89*, 490–499.

Oberer, M., Zangger, K., Gruber, K., & Keller, W. (2007). The solution structure of ParD, the antidote of the ParDE toxin antitoxin module, provides the structural basis for DNA and toxin binding. *Protein Science, 16*, 1676–1688.

Ogura, T., & Hiraga, S. (1983). Mini-F plasmid genes that couple host cell division to plasmid proliferation. *Proceedings of National Academy Science USA, 80*, 4784–4788.

Overgaard, M., Borch, J., Jørgensen, M. G., & Gerdes, K. (2008). Messenger RNA interferase RelE controls relBE transcription by conditional cooperativity. *Molecular Microbiology, 69*, 841–857.

Pandey, D. P., & Gerdes, K. (2005). Toxin-antitoxin loci are highly abundant in free-living but lost from host-associated prokaryotes. *Nucleic Acids Research, 33*, 966–976.

Papagiannis, C. V., Sam, M. D., Abbani, M. A., Yoo, D., Cascio, D., Clubb, R. T., et al. (2007). Fis targets assembly of the Xis nucleoprotein filament to promote excisive recombination by phage lambda. *Journal of Molecular Biology, 367*, 328–343.

Roberts, R. C., Burioni, R., & Helinski, D. R. (1990). Genetic characterization of the stabilizing functions of a region of broad-host-range plasmid RK2. *Journal of Bacteriology, 172*, 6204–6216.

Roberts, R. C., Spangler, C., & Helinski, D. R. (1993). Characteristics and significance of DNA binding activity of plasmid stabilization protein ParD from the broad host-range plasmid RK2. *Journal of Biological Chemistry, 268*, 27109–27117.

Roberts, R. C., Strom, A. R., & Helinski, D. R. (1994). The parDE operon of the broad-host-range plasmid RK2 specifies growth inhibition associated with plasmid loss. *Journal of Molecular Biology, 237*, 35–51.

Roca, J., & Wang, J. C. (1994). DNA transport by a type II DNA topoisomerase: Evidence in favor of a two-gate mechanism. *Cell, 77*, 609–616.

Rokney, A., Kobiler, O., Amir, A., Court, D. L., Stavans, J., Adhya, S., et al. (2008). Host responses influence on the induction of lambda prophage. *Molecular Microbiology, 68*, 29–36.

Rowe-Magnus, D. A., Guerout, A. M., Biskri, L., Bouige, P., & Mazel, D. (2003). Comparative analysis of superintegrons: Engineering extensive genetic diversity in the *Vibrionaceae*. *Genome Research, 13*, 428–442.

Saavedra De Bast, M., Mine, N., & Van Melderen, L. (2008). Chromosomal toxin-antitoxin systems may act as antiaddiction modules. *Journal of Bacteriology, 190*, 4603–4609.

Salmon, M. A., Van Melderen, L., Bernard, P., & Couturier, M. (1994). The antidote and autoregulatory functions of the F plasmid CcdA protein: A genetic and biochemical survey. *Molecular and General Genetics, 244*, 530–538.

Sayeed, S., Reaves, L., Radnedge, L., & Austin, S. (2000). The stability region of the large virulence plasmid of *Shigella flexneri* encodes an efficient postsegregational killing system. *Journal of Bacteriology, 182*, 2416–2421.

Scheirer, K. E., & Higgins, N. P. (1997). The DNA cleavage reaction of DNA gyrase. Comparison of stable ternary complexes formed with enoxacin and CcdB protein. *Journal of Biological Chemistry, 272*, 27202–27209.

Smith, A. B., & Maxwell, A. (2006). A strand-passage conformation of DNA gyrase is required to allow the bacterial toxin, CcdB, to access its binding site. *Nucleic Acids Research, 34*, 4667–4676.

Szekeres, S., Dauti, M., Wilde, C., Mazel, D., & Rowe-Magnus, D. A. (2007). Chromosomal toxin-antitoxin loci can diminish large-scale genome reductions in the absence of selection. *Molecular Microbiology, 63*, 1588–1605.

Tock, M. R., & Dryden, D. T. (2005). The biology of restriction and anti-restriction. *Current Opinion in Microbiology, 8*, 466–472.

Travers, A., & Muskhelishvili, G. (2005). DNA supercoiling—a global transcriptional regulator for enterobacterial growth? *Nature Reviews Microbiology, 3*, 157–169.

Trovatti, E., Cotrim, C. A., Garrido, S. S., Barros, R. S., & Marchetto, R. (2008). Peptides based on CcdB protein as novel inhibitors of bacterial topoisomerases. *Bioorganic and Medicinal Chemistry Letters, 18*, 6161–6164.

Tsuchimoto, S., Nishimura, Y., & Ohtsubo, E. (1992). The stable maintenance system pem of plasmid R100: Degradation of PemI protein may allow PemK protein to inhibit cell growth. *Journal of Bacteriology, 174*, 4205–4211.

Van Melderen, L. (2002). Molecular interactions of the CcdB poison with its bacterial target, the DNA gyrase. *International Journal of Medical Microbiology, 291*, 537–544.

Van Melderen, L., & Saavedra De Bast, M. (2009). Bacterial toxin-antitoxin systems: More than selfish entities? *PLoS Genetics, 5*, e1000437.

Van Melderen, L., Bernard, P., & Couturier, M. (1994). Lon-dependent proteolysis of CcdA is the key control for activation of CcdB in plasmid-free segregant bacteria. *Molecular Microbiology, 11*, 1151–1157.

Van Melderen, L., Thi, M. H., Lecchi, P., Gottesman, S., Couturier, M., & Maurizi, M. R. (1996). ATP-dependent degradation of CcdA by Lon protease. Effects of secondary structure and heterologous subunit interactions. *Journal of Biological Chemistry, 271*, 27730–27738.

Wang, X., & Wood, T. K. (2011). Toxin-antitoxin systems influence biofilm and persister cell formation and the general stress response. *Applied and Environment Microbiology, 77*, 5577–5583.

Wilbaux, M., Mine, N., Guerout, A. M., Mazel, D., & Van Melderen, L. (2007). Functional interactions between coexisting toxin-antitoxin systems of the ccd family in *Escherichia coli* O157:H7. *Journal of Bacteriology, 189*, 2712–2719.

Yarmolinsky, M. B. (1995). Programmed cell death in bacterial populations. *Science, 267*, 836–837.

Yuan, J., Sterckx, Y., Mitchenall, L. A., Maxwell, A., Loris, R., & Waldor, M. K. (2010). Vibrio cholerae ParE2 poisons DNA gyrase via a mechanism distinct from other gyrase inhibitors. *Journal of Biological Chemistry, 285*, 40397–40408.

Yuan, J., Yamaichi, Y., & Waldor, M. K. (2011). The three vibrio cholerae chromosome II-encoded ParE toxins degrade chromosome I following loss of chromosome II. *Journal of Bacteriology, 193*, 611–619.

Zechiedrich, E. L., & Cozzarelli, N. R. (1995). Roles of topoisomerase IV and DNA gyrase in DNA unlinking during replication in *Escherichia coli*. *Genes and Development, 9*, 2859–2869.

Chapter 5
Type II Toxin-Antitoxin Loci: The *relBE* Family

Kenn Gerdes

Abstract *relBE* of *Escherichia coli* K-12 is a paradigm TA locus. Here, I describe the discovery of *relBE* and review the genetic, physiological, biochemical and structural analyses that have led to important insights into TA biology. Five *relBE* homologues of K-12 are also described while the 6th (*mqsRA*) is treated in a separate Chapter. Finally, a discussion of the possible biological functions of *relBE* and other type II TA loci sets the stage for future work on prokaryotic type II TA loci.

5.1 Introduction

The *relE* gene of *Escherichia coli* K-12 was recognized as a "toxin" encoded by a TA locus in 1998 (Gotfredsen and Gerdes 1998). Soon thereafter, the first *relBE* homologs were identified (Grønlund and Gerdes 1999) and now we know that there are at least seven *relBE*-like loci in K-12 (Christensen-Dalsgaard et al. 2010). Many of these loci have similar properties to *relBE* but there are significant exceptions. For example, the *mqsRA* locus has properties so far not seen with *relBE* loci or with other TA loci. The *mqsRA* locus is therefore described in Chap. 6. This chapter describes *relBE* and the other five *relBE* homologous loci of *E. coli* K-12. It should be noted that *relBE* is now one of the best-described TA loci in terms of regulation of activity and transcription, structural insights into the TA complex, the free toxin and the free antitoxin. In particular, the target of RelE, the 70S ribosome, has been solved in two different states bound with RelE.

K. Gerdes (✉)
Institute for Cell and Molecular Biosciences, Centre for Bacterial Cell Biology, Newcastle University, Newcastle upon Tyne, NE2 4AX, UK
e-mail: kenn.gerdes@ncl.ac.uk

K. Gerdes (ed.), *Prokaryotic Toxin-Antitoxins*,
DOI: 10.1007/978-3-642-33253-1_5,

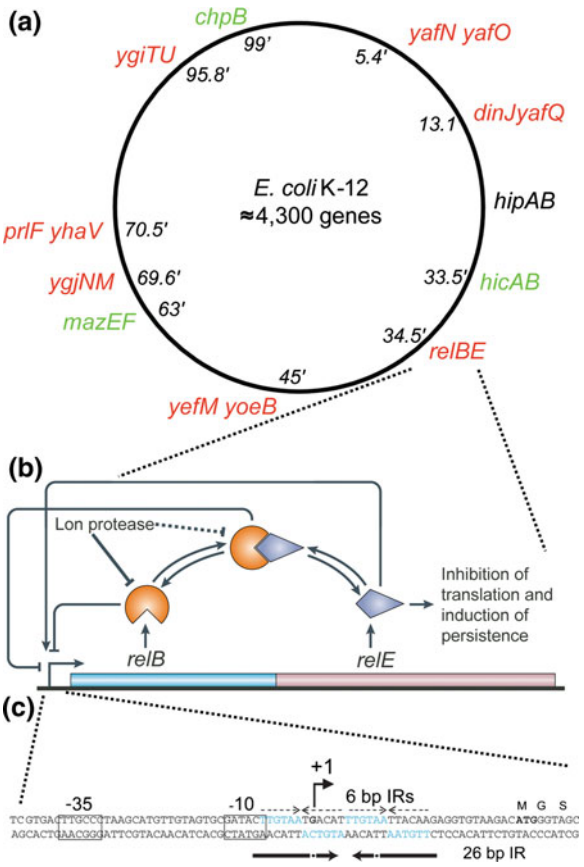


Fig. 5.1 TA loci of *E. coli* K-12 and a *close-up* on *relBE*. **a** Schematic of the *E. coli* K-12 chromosome showing the location of type II TA loci. **b** *Blow up* schematic of the *relBE* locus showing the regulatory connections that are described in the text. **c** DNA sequence of the *relBE* promoter region showing *relO*, a 26 bp inverted repeat (IR) (opposing *arrows below* the sequence) whose legs each contain smaller inverted repeats shown *above* the sequence. The two small IRs both bind RelB dimers. +1 denotes the transcriptional start site. Promoter sequences are also shown. **c** was adapted from (Overgaard et al. 2009). Note that *ygiTU* is now called *mqsRA* and *ygjNM* is called *higBA*

5.2 Discovery of *relB* and *relE* of *E. coli* K-12

Originally, the *relB* gene of the *relBEF* operon (Fig. 5.1a, b) was identified by single-point mutations that conferred the so-called "delayed-relaxed response". Wild-type *E. coli* cells respond to amino acid (aa) starvation by synthesizing (p)ppGpp that elicits the stringent response (Potrykus and Cashel 2008). The "alarmone" (p)ppGpp is synthesized by RelA [(p)ppGpp synthetase I], which is activated by the presence of uncharged tRNA at the ribosomal A-site. In turn, (p)ppGpp triggers profound physiological changes that help the cells cope with nutritional stress (Magnusson et al. 2005). Most importantly, (p)ppGpp shuts down rRNA transcription and stimulates transcription of biosynthetic operons. Cells defective in synthesizing (p)ppGpp during aa starvation (because they carry a mutation in *relA*) are said to be "relaxed" because their synthesis of stable RNA (rRNA + tRNA) continues during aa starvation. In the delayed-relaxed response, synthesis of stable RNA halts for about approx. 10 after onset of starvation and then continues—hence the term "delayed-relaxed" (Lavalle 1965; Bech et al. 1985). Certain single amino acid changes confer the delayed-relaxed

response. In addition, these *relB* mutants have a severe growth lag in re-feeding experiments that hinted at the activation of a translational inhibitor in the mutant strains. We now know that this translational inhibitor is RelE and that the delayed-relaxed response can be explained by hyperactivation of RelE (Christensen and Gerdes 2004). Hyperactivation of RelE occurs because point mutations in *relB* (e.g. *relB101*) that confer the delayed-relaxed response destabilizes RelB (see below). Thus, at the onset of aa starvation, mutant RelB is degraded very rapidly by Lon (Christensen and Gerdes 2004). This, in turn, leads to hyperactivation of RelE, which rapidly inhibits translation and thereby reduces the drain on the pools of charged tRNAs. The resulting recharging of tRNA terminates RelA activation which consequently leads to a normalization of the (p)ppGpp level and resumption of stable RNA synthesis. The delay in the triggering of the stringent response can thus be explained as the time it takes Lon to degrade the mutant RelB and activate RelE (Christensen and Gerdes 2004).

The *relE* gene is located immediately downstream of *relB* and upstream of *relF* (Fig. 5.1a). The *relF* gene is a *hok* homolog that we called *hokD* (Pedersen and Gerdes 1999) (see Chap. 2). The analysis of the *relBE* locus showed that *relE* encodes a very efficient inhibitor of cell growth and that inhibition of cell growth was due to inhibition of translation (Gotfredsen and Gerdes 1998; Pedersen et al. 2002). It was also obvious that *relBE* had the genetic organization of a type II TA locus and, consistent with this notion, induction of *relB* transcription counteracted the inhibitory effect of RelE (Gotfredsen and Gerdes 1998).

5.3 RelE's Mechanism of Action

The discovery that RelE inhibits translation was interesting and we decided to pursue the molecular mechanism. Induction of plasmid-encoded *relE* inhibited translation but not transcription or replication (Christensen et al. 2001). To elucidate the molecular mechanism we turned to biochemistry. Indeed, the initial experiments showed that purified RelE inhibited translation in an S30 extract and that purified RelB counteracted RelE (Pedersen et al. 2002). However, S30 extracts are crude mixtures and we turned to a fully reconstituted in vitro translation system developed in Måns Ehrenberg's laboratory.

5.3.1 RelE is a Ribosome-Dependent mRNase

Remarkably, addition of substoichiometric amounts of RelE led to cleavage of mRNA positioned at the ribosomal A-site, between the 2nd and 3rd nucleotide (Pedersen et al. 2003). Although the cleavage rate was low, there is no doubt that RelE functions as an enzyme because one RelE molecule was able to cleave multiple A-site codons. RelE can cleave all codons but has codon preferences; codons with a Guanine base at the 3rd position is cleaved much faster in vitro than

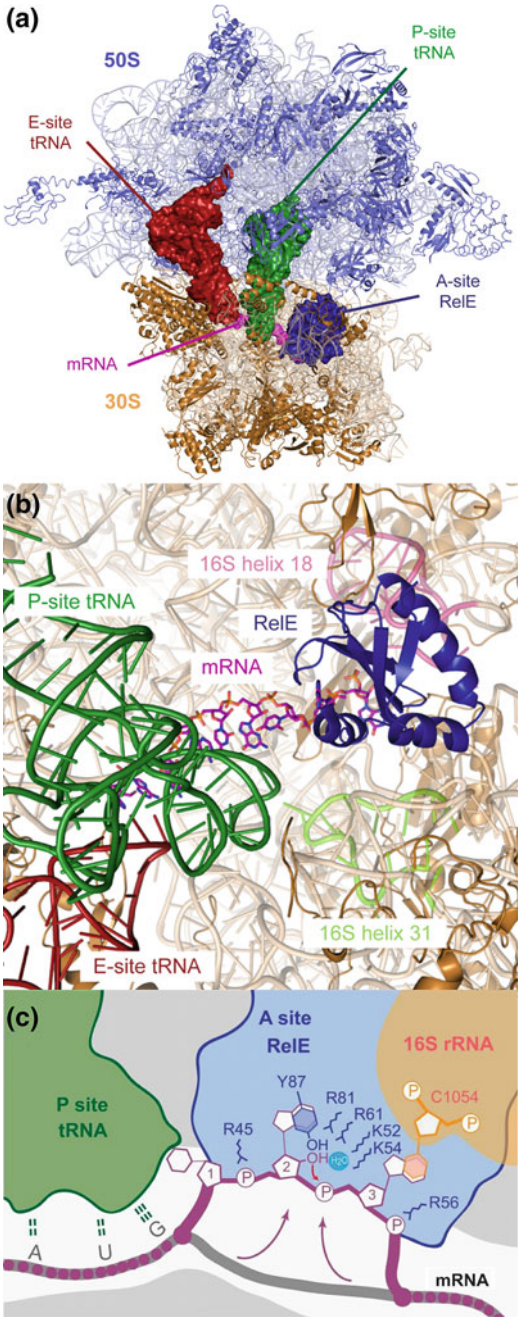

(a)
50S
P-site tRNA
E-site tRNA
A-site RelE
mRNA
30S
(b)
16S helix 18
P-site tRNA
RelE
mRNA
16S helix 31
E-site tRNA
(c)
A site RelE
16S rRNA
P site tRNA
Y87
R81
R61
K52
K54
R45
C1054
R56
mRNA

Fig. 5.2 The co-crystal structure of the 70S ribosome of Thermus thermophilus and RelE. **a** *Top* view of the 70S ribosome with the 50S (*blue*) and 30S (*wheat*) subunits surrounding RelE (A site, *blue*), $tRNA^{fMet}$ (P site, *green*), a noncognate tRNAfMet (E site, *red*), and mRNA (*magenta*). **a** and **b** are based on the precleavage structure. In this crystal form, the ribosomes contain a non-cognate $tRNA^{fMet}$ bound to the AAA codon in the E site, $tRNA^{fMet}$ bound to the AUG codon in the P site, and UAG in the A site, a stop codon which has previously been shown to elicit efficient cleavage by RelE on the *E. coli* ribosome in vitro (Pedersen et al. 2003) **b** Close-up of the A and P sites of the 30S subunit viewed from the interface to the 50S. RelE (*blue cartoon*) spans the 16S rRNA from the head (helix 31 region, *green*) to the body (helix 18, *pink*). The mRNA is shown in *purple sticks*, and the P and E site tRNAs coloured as in **a**. **c** Overview of the RelE cleavage mechanism. RelE (*blue*) occupies the A site where it pulls the mRNA (*purple*) into its active site (*arrows*). Here, the second nucleotide (''2'') stacks with Y87 (*blue*), while the third nucleotide (''3'') stacks with C1054 (*orange*). Relevant basic side chains and a water molecule in the active site are shown. The *red arrow* indicates the nucleophilic attack of the 20-O^- of the phosphate bond between the 2nd and 3rd nucleotide of the codon loaded at the A-site. Adapted from (Neubauer et al. 2009)

codons with G at the 2nd position. RelE activity was entirely dependent on the presence of ribosomes; in their absence, RelE did not cleave naked mRNAs. Consistently, RelE does not cleave mRNA outside coding regions of mRNAs in vivo (Christensen and Gerdes 2003).

RelE from *E. coli* is structurally related to extracellular microbial RNases that cleave RNA specifically after guanosine residues, independent of the ribosome (Neubauer et al. 2009; Li et al. 2009). However, RelE lacks a conserved aa generating the canonical enzymatic triad, thus explaining why RelE itself has no enzymatic activity in vitro (Li et al. 2009). Later structural analyses performed in collaboration with the laboratories of Venki Ramakrishnan and Ditlev Brodersen confirmed that, indeed, RelE binds to the ribosomal A-site and cleaves mRNA between the 2nd and 3rd nucleotide (Fig. 5.2a) (Neubauer et al. 2009). Binding of RelE to the ribosomal A-site only slightly changes RelE's conformation; thus, the ribosome does not activate RelE simply by changing its structure. Rather, RelE induces a change in the conformation of mRNA positioned at the A-site that promotes cleavage, a change that also requires that the mRNA interacts with a highly conserved residue in 16S rRNA helix 34, as visualized in Fig. 5.2b. The mechanism of RelE cleavage is shown schematically in Fig. 5.2c. This picture may at least in part explain why the ribosome is required for RelE to cleave mRNA. RelE is also structurally related to RegB of coliphage T4 that cleaves mRNA site specifically to activate translation (Odaert et al. 2007). The structures of two RelE homologs from archaea further strengthen the conclusion that RelE is related to microbial RNases (Takagi et al. 2005; Francuski and Saenger 2009).

RelE cleaves mRNA within reading frames in vivo and mutations that change the reading frame consistently change the cleavage pattern (Christensen and Gerdes 2003). The product of RelE cleavage is a ribosome stalled on an mRNA that lacks native stop-codon, a so-called "non-stop" mRNA. Obviously, this is detrimental because the ribosome is trapped and, furthermore, acts as a roadblock for ribosomes that may be engaged in translation of the upstream part of the mRNA. However, bacteria can rescue such trapped ribosomes by at least two mechanisms. The first and well-described mechanism depends on trans-translation

by tmRNA that is loaded into the A-site of ribosomes stalled at a non-stop mRNA (Ivanova et al. 2005; Keiler 2008). We tested the inference that overproduction of tmRNA and the tmRNA stabilizing protein (SrmB) would counteract RelE activity; indeed such overproduction reduced the growth-inhibitory effect of RelE overproduction. Conversely, deletion of *ssrA* encoding tmRNA increased the cell's sensitivity to RelE overproduction (Christensen and Gerdes 2003). Interestingly, overproduction of tmRNA also counteracts ribosome-independent such as MazF (Christensen et al. 2003), indicating that MazF also generates stalled ribosomes at non-stop mRNAs. The second ribosome rescue mechanism, which probably functions as a backup of the trans-translation reaction, is provided by ArfA, a hydrolase that can release nascent, incomplete peptides by hydrolysis of peptidyl-tRNA (Chadani et al. 2010). Interestingly, ArfA is only expressed when trans-translation is inactive (Chadani et al. 2010; Garza-Sanchez et al. 2011).

5.3.2 *RelE Cleaves A-site Codons of Eukaryotic and Mitochondrial Ribosomes*

Using a very sensitive Basic Local Alignment Search Tool procedure, we discovered numerous *relBE* loci in Gram-positive and archaeal organisms (Grønlund and Gerdes 1999). Strikingly, ectopic expression of archaeal RelE homologs in *E. coli* yielded mRNA cleavage patterns very similar to that of *E. coli* RelE, indicating that the mechanism of mRNA cleavage is conserved across the prokaryotic domains.

RelE homologs have not been identified in eukaryotes. Interestingly, however, RelE cleaves mRNA positioned at the A-site of eukaryotic ribosomes and has been used as a tool for "printing" the mRNA location on such ribosomes (Andreev et al. 2008). Consistently, RelE production induces apoptosis of lung cancer and osteosarcoma cell lines (Yamamoto et al. 2002; Shi et al. 2008) and severely inhibits the growth of yeast cells (Kristoffersen et al. 2000). These effects of RelE expression could be counteracted by RelB. Finally, engineering RelE to be imported into mammalian mitochondria revealed that RelE also cleaves mRNA codons positioned at A-site of mitochondrial ribosomes (Temperley et al. 2010). This latter property of RelE was exploited to show that mitoribosomes stalled at AGA and AGG arginine codons (mitochondria lack the corresponding tRNAs) located at the termini of reading frames devoid of conventional termination codons backtrack one base. The one base backtracking leads to the positioning of a conventional UAG stop codon in the A-site and thereby allows for regular termination of translation.

5.4 Regulation of *relBE* Transcription and RelE Activity

This section describes the complex regulation of *relBE* transcription and of RelE activity. Initially, *relBE* seemed to be a typical *E. coli* operon under simple negative autogenous control. However, recently we described a novel regulatory element

called "conditional cooperativity" that controls *relBE* transcription. Interestingly, although conditional cooperativity is a common property of the evolutionary independent TA gene families, the underlying molecular mechanisms are different (conf. Chaps. 9 and 16).

5.4.1 The RelBE Complex Autoregulates relBE Transcription

The employment of conventional *lacZ* fusion technology showed that the unrepressed *relBE* promoter (P_{relBE}) is very strong, with an activity comparable to that of the λP_L promoter and that RelB alone represses *relBE* transcription (Gotfredsen and Gerdes 1998). The presence of RelE enhances RelB's repression of the P_{relBE}::*lacZ* transcriptional fusion such that the fully repressed promoter has a $\approx$800-fold lower activity than that of the unrepressed promoter (Gotfredsen and Gerdes 1998; Overgaard et al. 2009). These results indicate that the main regulator of *relBE* transcription is the RelBE complex.

RelB is a dimer that interacts with operator DNA via its N-terminal Ribbon-Helix-Helix (RHH) domain (Li et al. 2008; Overgaard et al. 2009) and with RelE via its C-terminal domain (Cherny et al. 2007; Li et al. 2009). Biochemical analyses showed that the operator in P_{relBE} consists of two juxtaposed inverted repeats (called *relO*) to which two RelB dimers bind (Fig. 5.1c) (Overgaard et al. 2008, 2009; Li et al. 2008). In an excellent paper, Mitsuhiko Ikura and coworkers reported that the repressor complex had a $RelB_2 \cdot RelB_2$ stoichiometry (Li et al. 2008) whereas we reported a $RelB_2 \cdot$ RelE stoichiometry (Overgaard et al. 2008). The reason for this discrepancy is not known. Both groups agree that RelE enhances the binding of RelB to *relO* by strongly increasing cooperativity of the binding reaction. Foot printing analysis showed that only RelB contacts operator DNA and that RelE's stimulating effect on DNA binding is indirect (Li et al. 2008; Overgaard et al. 2009). These observations support that $RelB_2 \cdot$ RelE represses transcription of the *relBE* operon and is consistent with the in vivo data on regulation. Both groups also agree that two RelB dimers bind two juxtaposed inverted repeats in the promoter region and that the two operators function independently of each other.

5.4.2 Lon Protease Degrades RelB

Due to the delayed-relaxed response elicited by mutations in *relB*, we analyzed *relBE* transcription during aa starvation (Christensen et al. 2001). We observed that aa starvation elicited a dramatic 30-fold increase in the rate of *relBE* transcription. This increase correlated with a decrease in the cellular level of RelB, consistent with autogenous transcriptional control. It also indicated that RelB was degraded during aa starvation. Indeed, deletion of *lon* encoding the common stress protease Lon severely reduced the response and simultaneously prevented decay of

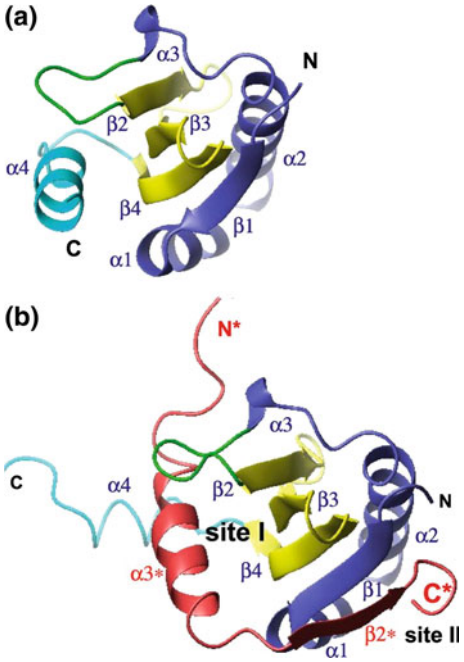


Fig. 5.3 RelB triggers structural rearrangements in RelE. **a** Structure of free RelE that exhibits a tertiary fold similar to microbial RNases and RegB of T4. **b** Structure of RelE in complex with the C-terminal domain of RelB. Upon interaction with RelE the unstructured tail of RelB folds into an α-helix (α3*) and a β sheet (β2*). Simultaneously, helix α3* of RelB displaces α4 of RelE such that it swings out from the surface of the central β-sheet. Adapted from (Li et al. 2009)

RelB (Christensen et al. 2001). Since *clp* and other protease-deficient alleles had no such effect, we concluded that RelB is degraded by Lon. Indeed, later biochemical analyses confirmed this notion (Overgaard et al. 2008, 2009). Interestingly, all other of the 11 known type II antitoxins of *E. coli* that have been tested are also degraded by Lon (Christensen et al. 2003; Jørgensen et al. 2009; Christensen-Dalsgaard et al. 2010; Wang et al. 2011), raising the possibility that these TA loci are controlled by a common cellular mechanism.

5.4.3 RelB Triggers Structural Rearrangements in RelE

RelB and RelE forms a complex in which RelE is enzymatically inactive (Pedersen et al. 2002, 2003; Overgaard et al. 2008). RelB was estimated to be present in 550–1100 molecules per cell while RelE was approximately 10-fold less abundant (Overgaard et al. 2009). This relatively modest excess of RelB can keep RelE inactivated only if RelB and RelE form a very tight complex. Using Surface Plasmon Resonance, we estimated an apparent dissociation constant $K_d = k_{off}/k_{on} = (1.66 \times 10^{-4}\ s^{-1})/(5.01 \times 10^5\ M^{-1}\ s^{-1}) = 0.33$ nM, very similar to the strong binding affinities reported for other TA pairs (Dao-Thi et al. 2005; Khoo et al. 2007; Overgaard et al. 2009).

The tertiary structure of free RelE of *E. coli* and RelE in complex with the C-terminal domain of RelB has been solved. RelE is a monomer with a compact structure consisting of a four-stranded antiparallel β sheet flanked by three α

helices (Fig. 5.3a) (Neubauer et al. 2009; Li et al. 2009). The tertiary fold of RelE is similar to that of microbial RNases and the RegB nuclease of T4 (Odaert et al. 2007; Neubauer et al. 2009; Li et al. 2009).

The free C-terminal domain of RelB is unstructured (Cherny et al. 2007). However, upon binding to RelE, the C-terminal domain of RelB adopts an $\alpha\beta$ configuration that forms a concave interface that interacts with a convex interface of RelE (Fig. 5.3b). This "embracement" of RelE by RelB explains the very strong interaction between the molecules (Overgaard et al. 2009). Remarkably, in the RelBE complex, helix α3* of RelB displaces α4 of RelE such that α4 swings out from the surface of the central β-sheet of RelE. This RelB-mediated structural displacement destroys the catalytic active site of RelE (Li et al. 2009).

5.4.4 RelE in The RelBE Complex Can be Activated

The strong RelB–RelE interaction raises an important question: Can RelE, which is in complex with RelB, be released from RelB such that it becomes active? If yes, how does it occur? The answer to the first question is: yes, RelE in complex with RelB can be activated. This conclusion comes from an experiment in which *relBE* transcription was driven by a synthetic P_{lac} promoter (Christensen et al. 2001). First, transcription of *relBE* was induced to express RelB and RelE. Since RelB in the induction period was expressed at a higher level than RelE, cellular translation was not inhibited. Then, at time zero, IPTG was removed and aa starvation induced. This treatment resulted in a severe inhibition of translation that depended on the presence of RelE. Thus, RelE that was in complex with RelB can indeed be released and thereby activated. It is not known if RelE activation is passive or active. The k_{off} rate determined in vitro corresponds to a half-life of the RelBE complex of ≈ 70 min. Intuitively, such a long half-life supports that an active mechanism exists. Since Lon degrades RelB, it is tempting to speculate that Lon can degrade both free RelB and RelB bound to RelE. The RelB C-terminus that interacts with the presumed low-affinity site of RelE in the $RelB_2 \cdot RelE$ complex might breathe (see below) and thereby allow Lon to gain access to the complex and thereby release RelE by degrading RelB.

5.4.5 relBE Transcription is Regulated by Conditional Cooperativity

Although the RelB level decreased immediately during aa starvation of wild type cells, the initial decrease was followed by recovery, with a new steady-state level of ≈ 50 % of that of the pre-starvation level, thus raising the possibility of additional elements controlling *relBE* expression (Christensen et al. 2001; Overgaard et al. 2008). Using a non-toxic variant of RelE that interacts normally with RelB, we found that overexpression of RelE leads to a dramatic increase in *relBE*

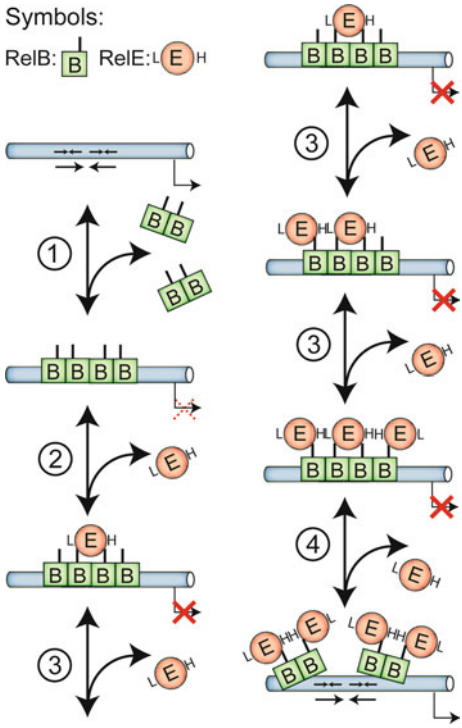


Fig. 5.4 Model explaining the mechanism of conditional cooperativity controlling *relBE* operon transcription. The *relBE* operator DNA (*blue tube*) contains a large 26 bp inverted repeat (opposing *large arrows*) of which each leg contains inverted repeats of 6 bp to which a RelB dimer binds (opposing *small arrows*) RelB dimers are shown as *green squares*. The C-termini of RelB are shown as *black bars* sticking out from the DNA binding domain of RelB. Broken arrow pointing right-wards indicates the transcriptional start site. **1** Two $RelB_2$ dimers bind with no or weak cooperativity to the adjacent 6 bp inverted repeats and partially repress transcription from the promoter (*dashed red cross*). **2** Addition of a low amount of RelE (*red circles*) leads to formation of a $RelB_2 \cdot RelE$ complex that binds with high cooperativity to operator DNA because it bridges two RelB dimers. L and H denote hypothetical low- and high-affinity RelB binding sites of RelE. **3** Additional RelE fills up the additional RelE-binding sites of the RelB · RelE operator complex. **4** When the amount of RelE exceeds that of RelB, free RelE displaces or breaks the interaction between a RelB C-terminus and the low-affinity binding site of RelE. This breakage destabilizes the operator complex and thereby activates *relBE* transcription

transcription (Overgaard et al. 2008). Biochemical analyses revealed that, surprisingly, the RelB/RelE ratio controls whether RelBE complexes bind to *relO*: when the RelB/RelE ratio was larger than one, then a $RelB_2 \cdot RelE$ complex bound strongly and cooperatively to *relO*. However, when the RelB/RelE ratio was equal to or less than one then the complex did not bind to *relO*. Thus, the RelB/RelE ratio controls complex binding to *relO* and thereby the transcription-rate of *relBE* in vivo. Our biochemical data also indicated a model for how RelE could do this (Overgaard et al. 2008). The model is visualized in Fig. 5.4. The model proposes

that RelB and RelE can form two complexes that have different RelB/RelE stoichiometry and therefore different binding affinities for *relO*. Whereas a $RelB_2 \cdot RelE$ complex binds strongly and cooperatively to *relO*, a $RelB_2 \cdot RelE_2$ complex has much lower affinity for the operator. The model further assumes that one RelE molecule has two binding sites for RelB, a low-affinity site and a high-affinity site. Thus, in the $RelB_2 \cdot RelE$ complex, one of the RelB monomers interacts with the high-affinity site of RelE whereas the other monomer interacts with the low-affinity site. When RelB > RelE, all complexes will be of the form $RelB_2 \cdot RelE$ that binds *relO* tightly and cooperatively. However, if for some reason the RelB/RelE ratio shifts (e.g., due to RelB degradation during aa starvation) such that RelE > RelB then the high-affinity RelB binding-sites of RelE will out compete the low-affinity RelB–RelE interaction and generate a $RelB_2 \cdot RelE_2$ complex that does not bind to *relO*. Thus, an excess of RelE will by this mechanism stimulate *relBE* transcription and restore the RelB/RelE ratio such that RelE activity will be quenched. As described in Chap. 9, later structural analyses revealed that transcription the *phd/doc* operon of P1 is regulated by exactly this mechanism (Garcia-Pino et al. 2010) whereas conditional cooperativity in *vapBC* operon control is regulated by a different mechanism (Winther and Gerdes 2012) (see also Chap. 8). Due to the ubiquitous presence of conditional cooperativity in evolutionary unrelated TA gene systems, we believe that this peculiar type of transcriptional control plays an important role in TA biology. One possibility is that it secures that fortuitous excess of toxin is rapidly sequestered by thus to prevent inadvertent inhibition of cell growth, as suggested previously (Afif et al. 2001). However, as discussed below, the phenomenon could also play a regulatory role in the generation or resuscitation of persister cells. It should be noted that these possibilities are not mutually exclusive. Interestingly, the three *relBE* loci of *Mycobacterium tuberculosis* are also regulated by conditional cooperativity (Yang et al. 2010).

5.5 RelE Homologs of *E. coli* K-12

TA loci encoding RelE homologs have also been analyzed quite extensively. The chromosome of *E. coli* encodes in total seven *relE* homologous genes all of which have an adjacent upstream or downstream antitoxin gene. The location of these 7 loci on the physical map of *E. coli* chromosome is shown in Fig. 5.1a. These seven *relE* genes are quite diverse, in some cases so diverse that conventional BLAST analyses will not identify the homologs. In such cases, the relatedness of a given toxin to RelE was confirmed by similarities in secondary and/or tertiary structures. For example, the unusual TA locus *mqsRA* encodes MqsR that exhibits an RNase fold similar to that of RelE (Brown et al. 2009). The *mqsRA* locus is described separately in Chap. 6 while *relBE* loci of pathogens are described in Chap. 8, Chap. 17, and Chap. 18 The following describes the *relBE* homologs of *E. coli* K-12, with emphasis on properties that differ from those of *relBE*.

5.5.1 *yefM yoeB*

The *yefM yoeB* locus was originally identified by sequence similarity with the *axe* (antitoxin)–*txe* (toxin) locus of an *Enterococcus faecium* multidrug-resistant plasmid (Grady and Hayes 2003). Similar to components of other addiction modules, the YoeB toxin is a basic protein (10 kDa), whereas the YefM antitoxin is acidic (11 kDa) and structurally unfolded (Cherny and Gazit 2004). The *yefM yoeB* cassette functions as a TA module (Grady and Hayes 2003; Cherny and Gazit 2004). We showed that similar to RelE, YoeB toxin induces cleavage of translated mRNA (Christensen et al. 2004). Moreover, overexpression of Lon specifically activated YoeB-dependent mRNA cleavage, indicating that YefM antitoxin is more sensitive to Lon cleavage than the other antitoxins of *E. coli* (Christensen et al. 2004). This inference is supported by the observation that inhibition of translation by VapC specifically activates YoeB (Winther and Gerdes 2009).

YoeB toxin has an RNase fold similar to that of RelE and furthermore, can cleave RNA in vitro independent of the ribosome (Kamada and Hanaoka 2005). Consistently, YoeB has a complete RNase fold containing all aa required for catalytic activity. Thus, this difference raised the possibility that YoeB could cleave mRNA independent of the ribosome in vivo, similar to what has been found for MazF and HicA (Zhang et al. 2003; Jørgensen et al. 2009). However, a detailed analysis of the mRNA cleavage patterns mediated by YoeB overexpression indicates that YoeB selectively cleaves mRNA codons positioned at the ribosomal A-site, thus supporting our initial interpretation (Christensen-Dalsgaard and Gerdes 2008). A more recent study showed that YoeB preferentially cleaves mRNAs at their second codon in a ribosome-dependent manner (Zhang and Inouye 2009). In most model mRNAs, we have observed that RelE and its homologs preferentially cleave mRNA at the second codon (often AAA) and at the termination codon. The most likely explanation of this effect is that the A-site is open for a longer time span at these codons than at elongation codons. The "dwelling" at these two codons is due to the fact that translation initiation and termination are considerably slower processes than elongation. In turn, the dwelling of the ribosome allows for more time at which the RNase can gain access to the A-site codon.

Comparison of the crystal structures of free YoeB and YoeB in complex with YefM revealed that the C-terminal tail of YoeB undergoes a conformational change when YefM binds to it (Kamada and Hanaoka 2005). In the free YoeB structure, the three C-terminal aa residues fold into a canonical conformation seen in the catalytic site of microbial RNases such as RNase Sa. In the TA complex structure, a change in the position of YoeB Tyr82 toward YoeB His83 results in the movement of His83 and other residues in YoeB such that the RNases becomes catalytically inactive (Kamada and Hanaoka 2005). Thus, the mechanisms by which the YoeB and RelB antitoxins inhibit their cognate RNases are very different at the molecular level. The functional significance, if any, of this difference is unknown.

5.5.2 *dinJ yafQ*

The *dinJ YafQ* locus was recognized as a putative *relBE* homologous locus by BLAST searches (Gerdes 2000). The *dinJ yafQ* operon promoter contains a LexA binding site, raising the possibility that the promoter is induced during the SOS response (Lewis et al. 1994; Fernandez De Henestrosa et al. 2000) and LexA as well as the DinJ-YafQ complex binds to the promoter region (Prysak et al. 2009). Consistently, *dinJ yafQ* transcription increased modestly after mitomycin C treatment of the cells (Prysak et al. 2009). Ectopic production of YafQ inhibited translation and this inhibition was counteracted by DinJ, thereby establishing *dinJ yafQ* as a *bona fide* TA locus (Motiejunaite et al. 2007). Interestingly, ectopic production of YafQ inhibited translation only 40 %, thus raising the possibility that YafQ may have target specificity. Using a different expression system, we found that indeed YafQ inhibited translation by $\approx$ 50 % only, very different from other mRNases expressed in the same way (Gerdes lab, unpublished). Purified YafQ exhibited robust ribonuclease activity in vitro that was independent of the ribosomes (Prysak et al. 2009). However, in vivo, YafQ cleaved mRNAs preferentially at AAA (lysine) codons (Prysak et al. 2009). This cleavage was dependent on the reading frame in the mRNA, indicating that like RelE and YoeB, YafQ cleaves mRNA located at the ribosomal A-site. The possible target selectivity by YafQ warrants further investigation.

5.5.3 *yafNO*

The *yafNO* genes are located downstream of *dinB* encoding DNA polymerase IV that is involved in error-prone DNA repair and is part of the SOS regulon (Wagner et al. 1999; Van Dyk et al. 2001). By sequence similarity, *yafN* was recognized as a *relB* homolog, raising the possibility that the neighboring gene (*yafO*) could be a toxin (Gerdes 2000). Later work showed that indeed expression of YafO is toxic and that YafN counteracts the toxicity of YafO (Brown and Shaw 2003; Singletary et al. 2009). Transcription of *yafNO* was proposed to be regulated by the SOS response (Singletary et al. 2009). We showed recently that *yafNO* has its own promoter located between the *dinB* and *yafN* genes. The *yafNO* promoter is autoregulated by YafN and the approx. four fold increase in *yafNO* transcription during the SOS response is solely due to the increase in *dinB* transcription (Christensen-Dalsgaard et al. 2010). This, of course, does not preclude that *yafNO* plays an as yet unidentified role during the SOS response, even though Rosenberg and colleagues showed that the locus plays no role in stress-induced mutagenesis (Singletary et al. 2009). We have also shown that ectopic expression of YafO inhibits translation very efficiently, by ribosome-dependent cleavage of mRNA. Similar to all other antitoxins of *E. coli* that have been tested, Lon degrades the YafN antitoxin (Christensen-Dalsgaard et al. 2010).

5.5.4 higBA *(Formerly* ygjNM*)*

The online program RASTA-Bacteria (**R**apid **A**utomated **S**can for **T**oxins and **A**ntitoxins in **Bacteria**) suggested that the the *ygjNM* locus might constitute a TA pair (Sevin and Barloy-Hubler 2007). Indeed, *ygjN* turned out to encode a ribosome-dependent mRNase while *ygjM* encodes an antitoxin that autoregulates transcription of the *ygjNM* locus. Since the gene order is reversed as compared to that of *relBE*, we renamed the locus *higBA*, consistent with previously accepted nomenclature (Tian et al. 1996; Christensen-Dalsgaard and Gerdes 2006). Similar to all other type II antitoxins of *E. coli* that have been tested, Lon degrades the HigA antitoxin (Christensen-Dalsgaard et al. 2010).

5.5.5 prlF yhaV

Bioinformatics analysis suggested that the *prlF yhaV* might be a TA pair (Anantharaman and Aravind 2003; Coles et al. 2005) and this was confirmed experimentally (Schmidt et al. 2007). Curiously, mutations in *prlF* suppresses the lethality associated with a toxic LamB-LacZ hybrid protein that jams the SecYEG transport system (Kiino and Silhavy 1984; Kiino et al. 1990). These mutations also cause hyperactivation of Lon (Snyder and Silhavy 1992). The molecular mechanisms behind these mysterious phenotypes are not known but it is possible that the *prlF* mutations leads to activation of YhaV mRNase and that this activation leads to a reduction of the mRNA that encodes the toxic fusion protein. PrlF and YhaV forms a highly stable complex in vitro, probably with a $PrlF_2 \cdot YhaV_4$ stoichiometry (Schmidt et al. 2007). PrlF contains a swapped-hairpin (AbrB-like) DNA-binding domain that mediates binding of the TA complex to operators in the promoter region and thereby mediates autoregulation of TA transcription (Schmidt et al. 2007). YhaV has ribonuclease activity in vitro that is independent of the ribosomes. The in vivo activity of YhaV mRNase has not been investigated.

5.6 Phylogeny of the RelE Superfamily

As with all other TA families, RelE toxins are better conserved than RelB antitoxins. In fact, RelB antitoxins are so diverse that it is not useful to divide the homologs into subfamilies. Rather, a *relB* gene should be defined as a gene adjacent to and upstream of a *relE* gene and encoding an inhibitor of the cognate RelE toxin. Similarly, *higA* and *mqsA* genes should be defined as being adjacent to and downstream of a *relE* homolog (called *higB* or *mqsR*), the difference between a *higA* and *mqsA* genes being that the former have their DNA binding domain at the N-terminus while the latter has it at the C-terminus. Even *relE* genes are highly

diverse (Gerdes 2000; Pandey and Gerdes 2005; Jørgensen et al. 2009) and it is often impossible to determine whether an RNase belongs to the RelE family without having access to the tertiary structure—that is—comparison of the tertiary structures may be the only way to finally determine whether a given protein belongs to the RelE superfamily (Takagi et al. 2005). Thus, although MqsR and YoeB have RNase activity, only the tertiary structures settled the issue whether these RNases were related to RelE (Kamada and Hanaoka 2005; Brown et al. 2009). Recently, a novel branch of the RelE superfamily was discovered and analyzed using the model system *brnTA* from *Brucella abortus*, a facultative intracellular pathogen (Heaton et al. 2012). The toxin gene, *brnT*, is located upstream of *brnA* that encodes an antitoxin with its DNA-binding domain at its C-terminus. Thus, *brnTA* subfamily resembles most closely the *mqsRA* subfamily in genetic organization and gene regulation. Interestingly, *brnTA* transcription was strongly induced by low pH and oxidative stress, stressors that *B. abortus* encounters during infection (Heaton et al. 2012). *B. abortus* is known to cause recurrent infections and it is possible that its 4 TA loci contribute to its persistence and thereby to relapses of infection.

ParE toxins (Chaps. 4 and 16) encoded by type II *parDE* loci inhibits DNA gyrase, similar to CcdB (Jiang et al., 2002). We observed that *parE* and *relE* homologs share sequence similarity (Pandey and Gerdes 2005). In fact, deep-branching homologs of RelE and ParE cannot be discriminated from each other by sequence similarity alone and ParE and RelE thus belong to the same superfamily of toxins. This observation also suggested that RelE and ParE share a common ancestral origin. Indeed, molecular homology modeling has shown that the tertiary structure of ParE is related to those of the RelE family members whose structures are known (Barbosa et al. 2010). Curiously, CcdB and MazF have similar tertiary folds belongs to the same superfamily (Kamada et al. 2003). Thus, two toxin superfamilies have both evolved to target the same two cellular processes, suggesting that, as also argued further below, the two superfamilies may have related cellular functions.

5.7 Biological Functions of the *relBE* Family

During the years, many functions have been assigned to TA loci (Magnuson 2007). It has also been proposed that they have no biological function (Van Melderen 2010). However, there is little doubt that several different functions can be ascribed to TA loci. Different functions are in some cases not mutually exclusive and it is therefore a possibility that a given TA locus may actually accomplish several functions simultaneously. However, it is possible that TA loci evolved to perform a common biological function, that is, that they have a "main" function. It is even possible that this main function has yet to be revealed. In the following, I present and discuss my own view on this issue as it evolved.

5.7.1 Plasmid and Chromosome-Segment Stabilization

Both type I and II TA loci were identified by their ability to stabilize plasmids (Ogura and Hiraga 1983; Gerdes et al. 1985). *ccdAB* was the first type II locus that was discovered due to its ability to stabilize plasmids in growing bacterial cultures. Plasmid stabilization by *ccdAB* was shown to be due to killing of new-born plasmid free cells (Hiraga et al. 1986; Jensen et al. 1995). Killing of plasmid-free segregants is also the mechanism by which the type I *hok*/*sok* locus of plasmid R1 stabilizes plasmids (Gerdes et al. 1986b). CcdB is a DNA gyrase inhibitor (Chap. 4). Thus, ectopic production of CcdB inhibits DNA replication and thereby, cell division. This inhibition of cell division is seen microscopically as the formation of long filamentous, non-dividing cells that stay viable during a considerable time span. So, although cell death eventually ensues, it is possible that the cell killing seen with CcdB overproduction is an overproduction artifact. This interpretation is supported by the observation that CcdA can fully resuscitate CcdB-inhibited cells even several hours after *ccdB* induction (Jørgensen and Gerdes, unpublished observation). The important point that I wish to make here is that the initial observations with plasmid stabilization by TA loci were interpreted by us and others to mean that, similar to type I loci (Gerdes et al. 1986a), type II loci encode toxins that kill the cells.

This view was dramatically changed when we discovered that cells inhibited by RelE or MazF could be fully resuscitated even three hours after ectopic production of the toxins (Pedersen et al. 2002). This robust cell viability assay told us that the biological function of RelE and MazF probably did not involve cell killing. Many type II loci are known to stabilize plasmids; however, the folds-of-stabilization are in most cases modest, much weaker than the effect of plasmid-encoded partitioning loci (Boe et al. 1987; Jensen et al. 1995; Grønlund and Gerdes 1999). Although type II loci that stabilize plasmids efficiently do exist, it is a rare observation (Roberts and Helinski 1992). A further indication that the primary biological function of type II TA is not plasmid stabilization comes from the observation that a plasmid-encoded *relBE* homolog stabilizes a test-plasmid to the same modest degree as the native chromosome-encoded *relBE* locus of *E. coli* K-12 (Grønlund and Gerdes 1999). One way to interpret these results is that the plasmid stabilization mediated by type II TA loci is an effect that unavoidably arises from the fact that the antitoxins are metabolically unstable while the toxins are more stable. In this view, the degree of stabilization is determined by intrinsic parameters of the components of the systems, such as antitoxin half-life, TA complex stability, and total amount of TA complex expressed in the cell.

Vibrio cholerae has 13 TA loci in its super-integron, all located adjacent to *attC* sites. Nine of these loci belong to the *relBE*/*higBA* superfamily. Due to the adjacent *attC* sites, these TA loci are considered *bona fide* integron cassettes that can undergo lateral transfer via the common integron mechanism (Pandey and Gerdes 2005). Integron elements are known to undergo frequent recombination events and it has been observed that the TA loci stabilize integron elements (Szekeres et al. 2007; Wozniak and Waldor 2009; Yuan et al. 2011). It is thus

possible that a major role of these particular TA loci is to stabilize other genetic elements within the super-integron. A test of this proposal is warranted but probably requires the construction of a *V. cholerae* strain devoid of most or all of its 13 TA loci.

5.7.2 Phage Defence

It has been proposed that TA loci may contain phage attacks by mediating abortive infection. The model is that as an infecting phage inhibits host cell translation the toxins will be activated due to decay of the antitoxins. In turn, phage translation and spreading will also be inhibited and the culture survives at the expense of a few infected cells. This phenomenon may be particularly relevant in the case of type III TA loci (Chap. 15). Briefly, the type III locus *toxIN* efficiently protects *Erwinia carotovora* from several different phages by providing the host cell with an abortive infection mechanism (Fineran et al. 2009). The locus encodes a small RNA (ToxI-RNA) that binds directly to the ToxN toxin. Interestingly, the crystal structure revealed that ToxN is a MazF homologue (Blower et al. 2011). Type I and II TA loci have also been suggested to protect cell cultures against spreading of phage, although the effects were rather modest (Pecota and Wood 1996; Hazan and Engelberg-Kulka 2004).

5.7.3 TA loci as Stress-Response-Elements

We proposed previously that TA loci may function as stress-response elements that are activated during nutritional starvation, in particular aa starvation (Christensen et al. 2001). This suggestion originated from the clear observation that, during aa starvation, *relBE* of K-12 reduces the global translation-rate from 10 to 5 % of the pre-starvation level. This may be a significant drop and may reduce the level of translational errors by reducing the drain on charged tRNAs (Sørensen 2001). Moreover, a *vapBC* locus of Sulfolobus is required for cell survival during heat shock (Cooper et al. 2009). Interestingly, the metabolic stability of a HigA antitoxin is controlled be a SecB homologous chaperone, further implicating TA loci in stress responses (Bordes et al. 2011) (see Chap. 17).

5.7.4 Persistence and TA Loci

Almost all bacteria produce rare cells that are tolerant to antibiotics. The phenomenon, called *bacterial persistence*, was discovered by Joseph Bigger who explored how cells of *Staphylococcus aureus* were killed by penicillin. Bigger

found that even though penicillin killed the vast majority of the cells in growing bacterial cultures there were often surviving bacteria that he called "persisters" (Bigger 1944). It is now known that persisters are cells that escape killing by antibiotics because they are in a state of slow growth or dormancy (Balaban et al. 2004). Almost all bacteria that have been tested express the persistence phenotype (Lewis 2010). In particular, many pathogenic bacteria, such as *Streptococcus pyogenes*, *Pseudomonas aeruginosa*, and *Salmonella enterica*, generate persistent or recurrent infections and the major human pathogen *M. tuberculosis* can persist for decades in an individual before an infection breaks out. Thus, the understanding of the molecular mechanism(s) behind persistence is important scientifically and medically.

The first evidence that bacterial persistence has a genetic basis came from the isolation of mutations in *E. coli* that increased the levels of persisters. The mutations were mapped to the *hipA* locus (high persistence) (Moyed and Bertrand 1983). HipA is a toxin encoded by the type II *hipAB* TA locus that also encodes HipB antitoxin (Chap. 11). Interestingly, direct connections between the persistence phenomenon and TA loci were indicated independently by several other studies. Two reports showed that persister cells contained elevated levels of TA mRNAs (Keren et al. 2004; Shah et al. 2006). Moreover, as described above, we showed that ectopic production of type II toxins induces a persistence-like state from which the cells could be resuscitated (Pedersen et al. 2002). This applies not only to RelE and MazF, but also to other type II toxins (Christensen-Dalsgaard and Gerdes 2006; 2010). Thus, TA loci were promising candidates of genes responsible for persistence and we decided to investigate the phenomenon further

It was known that, in *E. coli* K-12, deletion of a single TA locus did not affect the level of persister cells (Keren et al. 2004; Shah et al. 2006). To analyze this systematically, we generated strains carrying single deletions of the 10 TA loci that encode mRNases (Maisonneuve et al. 2011). None of these strains exhibited a reduction of persister cell formation (Fig. 5.5a). However, deletion of all 10 TA loci in the same strain reduced persistence 100–200-fold (Fig. 5.5b). This reduction was not due to deletion of one or two specific loci. Rather, the progressive deletion of more and more TA loci resulted in a gradual decrease in persistence (Fig. 5.5c). Curiously, deletion of the 4 first TA loci did not affect the persistence level significantly (Fig. 5.5c). However, a similar pattern was observed when we deleted the TA loci in reversed order: in this case, deletion of the first three loci did not affect the persistence level (Maisonneuve et al. 2011). All type II antitoxins of *E. coli* K-12 that have been tested are degraded by Lon. Consistently, a *lon* deletion strain had a highly reduced level of persistence (Maisonneuve et al. 2011). Deletion of other ATP-dependent proteases, such as *clpP* or *hslUV* had no effect on persistence. These results are consistent with a model in which persistence is induced by stochastic activation of Lon, followed by activation of one or more TA loci in a low fraction of the cells of a growing population ($\approx 10^{-4}$/cell/generation). Thus, for the first time we have been able to assign a common phenotype to TA loci belonging to evolutionary independent gene families. It should be noted that these results do not exclude that TA loci evolved to do something different and that

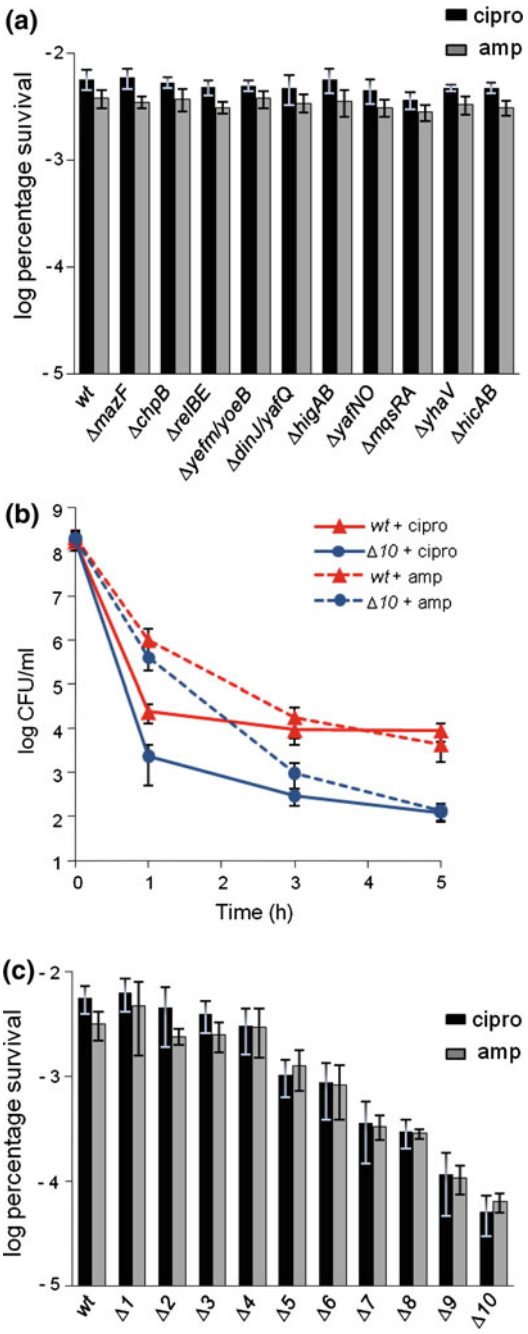


Fig. 5.5 TA-encoded mRNases are required for persister cell formation. **a** Cells of MG1655 (*wt*) and isogenic deletion strains (*mazF*, *chpB*, *relBE*, *yefMyoeB*, *higBA*, *dinJ yafQ*, *yafNO*, *mqsRA*, *yhaV*, and *hicAB*) were exposed to 1 μg/mL ciprofloxacin (*black bars*) or 100 μg/mL ampicillin (*gray bars*) in exponential growth phase. The percentage of survival after 5 h of antibiotic treatment was compared with that of the wt strain (*log scale*). amp, ampicillin; cipro, ciprofloxacin. **b** Exponentially growing cultures of MG1655 (wt) and MG1655 *Δ10TA* strains were exposed to 1 μg/mL ciprofloxacin or 100 μg/mL ampicillin. **c** Exponentially growing cells of MG1655 carrying increasing numbers of TA locus deletions were exposed to 1 μg/mL ciprofloxacin (*black bars*) or 100 μg/mL ampicillin (*gray bars*). Percentage of survival after 5 h of antibiotic treatment was compared with the wt strain (*log scale*). Error bars indicate the Standard deviation. For genotypes of the individual TA locus deletion strains, see (Maisonneuve et al. 2011)

the effect seen on persistence reflects a reduced viability of the cells that differentiate into persistence. Future work may answer this question.

Acknowledgments This work was supported by the Wellcome Trust. I thank members of the Gerdes group and of the Centre for Bacterial Cell Biology for friendly and stimulating discussions.

References

Afif, H., Allali, N., Couturier, M., & Van Melderen, L. (2001). The ratio between CcdA and CcdB modulates the transcriptional repression of the ccd poison-antidote system. *Molecular Microbiology, 41*, 73–82.

Anantharaman, V., & Aravind, L. (2003). New connections in the prokaryotic toxin–antitoxin network: Relationship with the eukaryotic nonsense-mediated RNA decay system. *Genome Biology, 4*, R81.

Andreev, D., Hauryliuk, V., Terenin, I., Dmitriev, S., Ehrenberg, M., & Shatsky, I. (2008). The bacterial toxin RelE induces specific mRNA cleavage in the A site of the eukaryote ribosome. *RNA, 14*, 233–239.

Balaban, N. Q., Merrin, J., Chait, R., Kowalik, L., & Leibler, S. (2004). Bacterial persistence as a phenotypic switch. *Science, 305*, 1622–1625.

Barbosa, L. C., Garrido, S. S., Garcia, A., Delfino, D. B., & Marchetto, R. (2010). Function inferences from a molecular structural model of bacterial ParE toxin. *Bioinformation, 4*, 438–440.

Bech, F. W., Jorgensen, S. T., Diderichsen, B., & Karlstrom, O. H. (1985). Sequence of the Relb Transcription Unit from *Escherichia-Coli* and Identification of the RelB Gene. *EMBO Journal, 4*, 1059–1066.

Bigger, J. W. (1944). Treatment of staphylococcal infections with penicillin by intermittent sterilisation. *Lancet, ii*, 497–500.

Blower, T. R., Pei, X. Y., Short, F. L., Fineran, P. C., Humphreys, D. P., Luisi, B. F., et al. (2011). A processed noncoding RNA regulates an altruistic bacterial antiviral system. *Nature Structural and Molecular Biology, 18*, 185–190.

Boe, L., Gerdes, K., & Molin, S. (1987). Effects of genes exerting growth inhibition and plasmid stability on plasmid maintenance. *Journal of Bacteriology, 169*, 4646–4650.

Bordes, P., Cirinesi, A. M., Ummels, R., Sala, A., Sakr, S., Bitter, W., et al. (2011). SecB-like chaperone controls a toxin–antitoxin stress-responsive system in mycobacterium tuberculosis. *Proceedings of the National Academy of Sciences of the U S A, 108*, 8438–8443.

Brown, J. M., & Shaw, K. J. (2003). A novel family of *Escherichia coli* toxin–antitoxin gene pairs. *Journal of Bacteriology, 185*, 6600–6608.

Brown, B. L., Grigoriu, S., Kim, Y., Arruda, J. M., Davenport, A., Wood, T. K., et al. (2009). Three dimensional structure of the MqsR:MqsA complex: A novel TA pair comprised of a toxin homologous to RelE and an antitoxin with unique properties. *PLoS Pathogens, 5*, e1000706.

Chadani, Y., Ono, K., Ozawa, S., Takahashi, Y., Takai, K., Nanamiya, H., et al. (2010). Ribosome rescue by Escherichia coli ArfA (YhdL) in the absence of trans-translation system. *Molecular Microbiology, 78*, 796–808.

Cherny, I., & Gazit, E. (2004). The YefM antitoxin defines a family of natively unfolded proteins—implications as a novel antibacterial target. *Journal of Biological Chemistry, 279*, 8252–8261.

Cherny, I., Overgaard, M., Borch, J., Bram, Y., Gerdes, K., & Gazit, E. (2007). Structural and thermodynamic characterization of the *Escherichia coli* RelBE toxin–antitoxin system: Indication for a functional role of differential stability. *Biochemistry, 46*, 12152–12163.

Christensen, S. K., & Gerdes, K. (2003). RelE toxins from bacteria and Archaea cleave mRNAs on translating ribosomes, which are rescued by tmRNA. *Molecular Microbiology, 48*, 1389–1400.

Christensen, S. K., & Gerdes, K. (2004). Delayed-relaxed response explained by hyperactivation of RelE. *Molecular Microbiology, 53*, 587–597.

Christensen, S. K., Mikkelsen, M., Pedersen, K., & Gerdes, K. (2001). RelE, a global inhibitor of translation, is activated during nutritional stress. *Proceedings of the National Academy of Sciences of the U S A, 98*, 14328–14333.

Christensen, S. K., Maenhaut-Michel, G., Mine, N., Gottesman, S., Gerdes, K., & Van Melderen, L. (2004). Overproduction of the lon protease triggers inhibition of translation in Escherichia coli: Involvement of the yefM-yoeB toxin–antitoxin system. *Molecular Microbiology, 51*, 1705–1717.

Christensen, S. K., Pedersen, K., Hansen, F. G., & Gerdes, K. (2003). Toxin–antitoxin loci as stress-response-elements: ChpAK/MazF and ChpBK cleave translated RNAs and are counteracted by tmRNA. *Journal of Molecular Biology, 332*, 809–819.

Christensen-Dalsgaard, M., & Gerdes, K. (2006). Two higBA loci in the Vibrio cholerae superintegron encode mRNA cleaving enzymes and can stabilize plasmids. *Molecular Microbiology, 62*, 397–411.

Christensen-Dalsgaard, M., & Gerdes, K. (2008). Translation affects YoeB and MazF messenger RNA interferase activities by different mechanisms. *Nucleic Acids Research, 36*, 6472–6481.

Christensen-Dalsgaard, M., Jørgensen, M. G., & Gerdes, K. (2010). Three new RelE-homologous mRNA interferases of Escherichia coli differentially induced by environmental stresses. *Molecular Microbiology, 75*, 333–348.

Coles, M., Djuranovic, S., Soding, J., Frickey, T., Koretke, K., Truffault, V., et al. (2005). AbrB-like transcription factors assume a swapped hairpin fold that is evolutionarily related to double-psi beta barrels. *Structure, 13*, 919–928.

Cooper, C. R., Daugherty, A. J., Tachdjian, S., Blum, P. H., & Kelly, R. M. (2009). Role of vapBC toxin–antitoxin loci in the thermal stress response of Sulfolobus solfataricus. *Biochemical Society Transactions, 37*, 123–126.

Dao-Thi, M. H., Van Melderen, L., De Genst, E., Afif, H., Buts, L., Wyns, L., et al. (2005). Molecular basis of gyrase poisoning by the addiction toxin CcdB. *Journal of Molecular Biology, 348*, 1091–1102.

De Fernandez-Henestrosa, A. R., Ogi, T., Aoyagi, S., Chafin, D., Hayes, J. J., Ohmori, H., et al. (2000). Identification of additional genes belonging to the LexA regulon in Escherichia coli. *Molecular Microbiology, 35*, 1560–1572.

Fineran, P. C., Blower, T. R., Foulds, I. J., Humphreys, D. P., Lilley, K. S., & Salmond, G. P. (2009). The phage abortive infection system, ToxIN, functions as a protein-RNA toxin–antitoxin pair. *Proceedings of the National Academy of Sciences of the U.S.A., 106*, 894–899.

Francuski, D., & Saenger, W. (2009). Crystal structure of the antitoxin–toxin protein complex RelB-RelE from Methancaldococcus jannaschii. *Journal of Molecular Biology, 393*, 898–908.

Garcia-Pino, A., Balasubramanian, S., Wyns, L., Gazit, E., De Greve, H., Magnuson, R. D., et al. (2010). Allostery and intrinsic disorder mediate transcription regulation by conditional cooperativity. *Cell, 142*, 101–111.

Garza-Sanchez, F., Schaub, R. E., Janssen, B. D., & Hayes, C. S. (2011). tmRNA regulates synthesis of the ArfA ribosome rescue factor. *Molecular Microbiology, 80*, 1204–1219.

Gerdes, K. (2000). Toxin–antitoxin modules may regulate synthesis of macromolecules during nutritional stress. *Journal of Bacteriology, 182*, 561–572.

Gerdes, K., Bech, F. W., Jorgensen, S. T., Lobnerolesen, A., Rasmussen, P. B., Atlung, T., et al. (1986a). Mechanism of Postsegregational killing by the hok gene-product of the parb system of plasmid R1 and its homology with the Relf gene-product of the Escherichia-Coli Relb operon. *EMBO Journal, 5*, 2023–2029.

Gerdes, K., Larsen, J. E., & Molin, S. (1985). Stable inheritance of plasmid R1 requires two different loci. *Journal of Bacteriology, 161*, 292–298.

Gerdes, K., Rasmussen, P. B., & Molin, S. (1986b). Unique type of plasmid maintenance function—postsegregational killing of plasmid-free cells. *Proceedings of the National Academy of Sciences of the U.S.A., 83*, 3116–3120.

Gotfredsen, M., & Gerdes, K. (1998). The Escherichia coli relBE genes belong to a new toxin–antitoxin gene family. *Molecular Microbiology, 29*, 1065–1076.

Grady, R., & Hayes, F. (2003). Axe-Txe, a broad-spectrum proteic toxin–antitoxin system specified by a multidrug-resistant, clinical isolate of Enterococcus faecium. *Molecular Microbiology, 47*, 1419–1432.

Grønlund, H., & Gerdes, K. (1999). Toxin–antitoxin systems homologous with relBE of Escherichia coli plasmid P307 are ubiquitous in prokaryotes. *Journal of Molecular Biology, 285*, 1401–1415.

Hazan, R., & Engelberg-Kulka, H. (2004). Escherichia coli mazEF-mediated cell death as a defence mechanism that inhibits the spread of phage P1. *Molecular Genetics and Genomics, 272*, 227–234.

Heaton, B. E., Herrou, J., Blackwell, A. E., Wysocki, V. H., & Crosson, S. (2012). Molecular structure and function of the novel BrnT/BrnA toxin–antitoxin system of brucella abortus. *Journal of Biological Chemistry, 287*, 12098–12110.

Hiraga, S., Jaffe, A., Ogura, T., Mori, H., & Takahashi, H. (1986). F-Plasmid Ccd mechanism in Escherichia-Coli. *Journal of Bacteriology, 166*, 100–104.

Ivanova, N., Pavlov, M. Y., & Ehrenberg, M. (2005). tmRNA-induced release of messenger RNA from stalled ribosomes. *Journal of Molecular Biology, 350*, 897–905.

Jensen, R. B., Grohmann, E., Schwab, H., Diazorejas, R., & Gerdes, K. (1995). Comparison of Ccd of F, Parde of Rp4, and Pard of R1 using a novel conditional replication control-system of plasmid R1. *Molecular Microbiology, 17*, 211–220.

Jiang, Y., Pogliano, J., Helinski, D. R., & Konieczny, I. (2002). ParE toxin encoded by the broad-host-range plasmid RK2 is an inhibitor of Escherichia coli gyrase. *Molecular Microbiology, 44*, 971–979.

Jørgensen, M. G., Pandey, D. P., Jaskolska, M., & Gerdes, K. (2009). HicA of Escherichia coli defines a novel family of translation-independent mRNA interferases in bacteria and archaea. *Journal of Bacteriology, 191*, 1191–1199.

Kamada, K., & Hanaoka, F. (2005). Conformational change in the catalytic site of the ribonuclease YoeB toxin by YefM antitoxin. *Molecular Cell, 19*, 497–509.

Kamada, K., Hanaoka, F., & Burley, S. K. (2003). Crystal structure of the MazE/MazF complex: Molecular bases of antidote-toxin recognition. *Molecular Cell, 11*, 875–884.

Keiler, K. C. (2008). Biology of trans-translation. *Annual Review of Microbiology, 62*, 133–151.

Keren, I., Shah, D., Spoering, A., Kaldalu, N., & Lewis, K. (2004). Specialized persister cells and the mechanism of multidrug tolerance in Escherichia coli. *Journal of Bacteriology, 186*, 8172–8180.

Khoo, S. K., Loll, B., Chan, W. T., Shoeman, R. L., Ngoo, L., Yeo, C. C., et al. (2007). Molecular and structural characterization of the PezAT chromosomal toxin–antitoxin system of the human pathogen Streptococcus pneumoniae. *Journal of Biological Chemistry, 282*, 19606–19618.

Kiino, D. R., Phillips, G. J., & Silhavy, T. J. (1990). Increased expression of the bifunctional protein PrlF suppresses overproduction lethality associated with exported beta-galactosidase hybrid proteins in Escherichia coli. *Journal of Bacteriology, 172*, 185–192.

Kiino, D. R., & Silhavy, T. J. (1984). Mutation prlF1 relieves the lethality associated with export of beta-galactosidase hybrid proteins in Escherichia coli. *Journal of Bacteriology, 158*, 878–883.

Kristoffersen, P., Jensen, G. B., Gerdes, K., & Piskur, J. (2000). Bacterial toxin–antitoxin gene system as containment control in yeast cells. *Applied and Environmental Microbiology, 66*, 5524–5526.

Lavalle, R. (1965). New mutants for regulation of RNA synthesis. *Bull Soc Chim Biol (Paris), 47*, 1567–1570.

Lewis, K. (2010). Persister cells. *Annual Review of Microbiology, 64*, 357–372.

Lewis, L. K., Harlow, G. R., Gregg-Jolly, L. A., & Mount, D. W. (1994). Identification of high affinity binding sites for LexA which define new DNA damage-inducible genes in Escherichia coli. *Journal of Molecular Biology, 241*, 507–523.

Li, G. Y., Zhang, Y., Inouye, M., & Ikura, M. (2008). Structural mechanism of transcriptional autorepression of the Escherichia coli RelB/RelE antitoxin/toxin module. *Journal of Molecular Biology, 380*, 107–119.

Li, G. Y., Zhang, Y., Inouye, M., & Ikura, M. (2009). Inhibitory mechanism of Escherichia coli RelE-RelB toxin–antitoxin module involves a helix displacement near an mRNA interferase active site. *Journal of Biological Chemistry, 284*, 14628–14636.

Magnuson, R. D. (2007). Hypothetical functions of toxin–antitoxin systems. *Journal of Bacteriology, 189*, 6089–6092.

Magnusson, L. U., Farewell, A., & Nyström, T. (2005). ppGpp a global regulator in Escherichia coli. *Trends in Microbiology, 13*, 236–242.

Maisonneuve, E., Shakespeare, L. J., Jorgensen, M. G., & Gerdes, K. (2011). Bacterial persistence by RNA endonucleases. *Proceedings of the National Academy of Sciences of the U S A, 108*, 13206–13211.

Motiejunaite, R., Armalyte, J., Markuckas, A., & Suziedeliene, E. (2007). Escherichia coli dinJ-yafQ genes act as a toxin–antitoxin module. *FEMS Microbiology Letters, 268*, 112–119.

Moyed, H. S., & Bertrand, K. P. (1983). hipA, a newly recognized gene of Escherichia coli K-12 that affects frequency of persistence after inhibition of murein synthesis. *Journal of Bacteriology, 155*, 768–775.

Neubauer, C., Gao, Y. G., Andersen, K. R., Dunham, C. M., Kelley, A. C., Hentschel, J., et al. (2009). The structural basis for mRNA recognition and cleavage by the ribosome-dependent endonuclease RelE. *Cell, 139*, 1084–1095.

Odaert, B., Saida, F., Aliprandi, P., Durand, S., Crechet, J. B., Guerois, R., et al. (2007). Structural and functional studies of RegB, a new member of a family of sequence-specific ribonucleases involved in mRNA inactivation on the ribosome. *Journal of Biological Chemistry, 282*, 2019–2028.

Ogura, T., & Hiraga, S. (1983). Mini-F plasmid genes that couple host-cell division to plasmid proliferation. *Proceedings of the National Academy of Sciences of the U.S.A., 80*, 4784–4788.

Overgaard, M., Borch, J., & Gerdes, K. (2009). RelB and RelE of Escherichia coli form a tight complex that represses transcription via the ribbon-helix-helix motif in RelB. *Journal of Molecular Biology, 394*, 183–196.

Overgaard, M., Borch, J., Jorgensen, M. G., & Gerdes, K. (2008). Messenger RNA interferase RelE controls relBE transcription by conditional cooperativity. *Molecular Microbiology, 69*, 841–857.

Pandey, D. P., & Gerdes, K. (2005). Toxin–antitoxin loci are highly abundant in free-living but lost from host-associated prokaryotes. *Nucleic Acids Research, 33*, 966–976.

Pecota, D. C., & Wood, T. K. (1996). Exclusion of T4 phage by the hok/sok killer locus from plasmid R1. *Journal of Bacteriology, 178*, 2044–2050.

Pedersen, K., Christensen, S. K., & Gerdes, K. (2002). Rapid induction and reversal of a bacteriostatic condition by controlled expression of toxins and antitoxins. *Molecular Microbiology, 45*, 501–510.

Pedersen, K., & Gerdes, K. (1999). Multiple hok genes on the chromosome of Escherichia coli. *Molecular Microbiology, 32*, 1090–1102.

Pedersen, K., Zavialov, A. V., Pavlov, M. Y., Elf, J., Gerdes, K., & Ehrenberg, M. (2003). The bacterial toxin RelE displays codon-specific cleavage of mRNAs in the ribosomal A site. *Cell, 112*, 131–140.

Potrykus, K., & Cashel, M. (2008). (p)ppGpp still magical? *Annual Review of Microbiology, 62*, 35–51.

Prysak, M. H., Mozdzierz, C. J., Cook, A. M., Zhu, L., Zhang, Y., Inouye, M., et al. (2009). Bacterial toxin YafQ is an endoribonuclease that associates with the ribosome and blocks translation elongation through sequence-specific and frame-dependent mRNA cleavage. *Molecular Microbiology, 71*, 1071–1087.

Roberts, R. C., & Helinski, D. R. (1992). Definition of a minimal plasmid stabilization system from the broad-host-range plasmid Rk2. *Journal of Bacteriology, 174*, 8119–8132.

Schmidt, O., Schuenemann, V. J., Hand, N. J., Silhavy, T. J., Martin, J., Lupas, A. N., et al. (2007). prIF and yhaV encode a new toxin–antitoxin system in Escherichia coli. *Journal of Molecular Biology, 372*, 894–905.

Sevin, E. W., & Barloy-Hubler, F. (2007). RASTA-Bacteria: A web-based tool for identifying toxin–antitoxin loci in prokaryotes. *Genome Biology, 8*, R155.

Shah, D., Zhang, Z. G., Khodursky, A., Kaldalu, N., Kurg, K., & Lewis, K. (2006). Persisters: A distinct physiological state of E-coli. *Bmc Microbiology, 6*, 53.

Shi, Y. L., Bao, L., Shang, Z. L., & Yao, S. X. (2008). RelE toxin protein of mycobacterium tuberculosis induces growth inhibition of lung cancer A-549 cell. *Sichuan Da Xue Xue Bao Yi Xue Ban, 39*, 368–372.

Singletary, L. A., Gibson, J. L., Tanner, E. J., McKenzie, G. J., Lee, P. L., Gonzalez, C., et al. (2009). An SOS-regulated type 2 toxin–antitoxin system. *Journal of Bacteriology, 191*, 7456–7465.

Snyder, W. B., & Silhavy, T. J. (1992). Enhanced export of beta-galactosidase fusion proteins in prlF mutants is Lon dependent. *Journal of Bacteriology, 174*, 5661–5668.

Sørensen, M. A. (2001). Charging levels of four tRNA species in Escherichia coli Rel(+) and Rel(-) strains during amino acid starvation: A simple model for the effect of ppGpp on translational accuracy. *Journal of Molecular Biology, 307*, 785–798.

Szekeres, S., Dauti, M., Wilde, C., Mazel, D., & Rowe-Magnus, D. A. (2007). Chromosomal toxin—antitoxin loci can diminish large-scale genome reductions in the absence of selection. *Molecular Microbiology, 63*, 1588–1605.

Takagi, H., Kakuta, Y., Okada, T., Yao, M., Tanaka, I., & Kimura, M. (2005). Crystal structure of archaeal toxin–antitoxin RelE-RelB complex with implications for toxin activity and antitoxin effects. *Nature Structural and Molecular Biology, 12*, 327–331.

Temperley, R., Richter, R., Dennerlein, S., Lightowlers, R. N., & Chrzanowska-Lightowlers, Z. M. (2010). Hungry codons promote frameshifting in human mitochondrial ribosomes. *Science, 327*, 301.

Tian, Q. B., Hayashi, T., Murata, T., & Terawaki, Y. (1996). Gene product identification and promoter analysis of hig locus of plasmid Rts1. *Biochemical and Biophysical Research Communications, 225*, 679–684.

Van Dyk, T. K., DeRose, E. J., & Gonye, G. E. (2001). LuxArray, a high-density, genomewide transcription analysis of Escherichia coli using bioluminescent reporter strains. *Journal of Bacteriology, 183*, 5496–5505.

Van Melderen, L. (2010). Toxin–antitoxin systems: Why so many, what for? *Current Opinion in Microbiology, 13*, 781–785.

Wagner, J., Gruz, P., Kim, S. R., Yamada, M., Matsui, K., Fuchs, R. P., et al. (1999). The dinB gene encodes a novel E. coli DNA polymerase, DNA pol IV, involved in mutagenesis. *Molecular Cell, 4*, 281–286.

Wang, X., Kim, Y., Hong, S. H., Ma, Q., Brown, B. L., Pu, M., et al. (2011). Antitoxin MqsA helps mediate the bacterial general stress response. *Nature Chemical Biology, 7*, 359–366.

Winther, K. S., & Gerdes, K. (2009). Ectopic production of VapCs from Enterobacteria inhibits translation and trans-activates YoeB mRNA interferase. *Molecular Microbiology, 72*, 918–930.

Winther, K. S., & Gerdes, K. (2012). Regulation of Enteric vapBC Transcription: Induction by VapC Toxin Dimer-Breaking. *Nucleic Acids Research, 40*, 4347–4357.

Wozniak, R. A., & Waldor, M. K. (2009). A toxin–antitoxin system promotes the maintenance of an integrative conjugative element. *PLoS Genetics, 5*, e1000439.

Yamamoto, T. A. M., Gerdes, K., & Tunnacliffe, A. (2002). Bacterial toxin RelE induces apoptosis in human cells. *FEBS Letters, 519*, 191–194.

Yang, M., Gao, C., Wang, Y., Zhang, H., & He, Z. G. (2010). Characterization of the interaction and cross-regulation of three mycobacterium tuberculosis RelBE modules. *PLoS ONE, 5*, e10672.

Yuan, J., Yamaichi, Y., & Waldor, M. K. (2011). The three vibrio cholerae chromosome II-encoded ParE toxins degrade chromosome I following loss of chromosome II. *Journal of Bacteriology, 193*, 611–619.

Zhang, Y., & Inouye, M. (2009). The inhibitory mechanism of protein synthesis by YoeB, an Escherichia coli toxin. *Journal of Biological Chemistry, 284*, 6627–6638.

Zhang, Y. L., Zhang, J. J., Hoeflich, K. P., Ikura, M., Qing, G. L., & Inouye, M. (2003). MazF cleaves cellular mRNAs specifically at ACA to block protein synthesis in Escherichia coli. *Molecular Cell, 12*, 913–923.

Chapter 6
Type II Toxin-Antitoxin Loci: The Unusual *mqsRA* Locus

Niilo Kaldalu, Villu Kasari, Gemma Atkinson and Tanel Tenson

Abstract The *mqsRA* locus of *Escherichia coli* K-12 codes for a translation-independent GCU site-specific endoribonuclease MqsR and an antitoxin MqsA, which has an additional function as a transcriptional regulator of other genes. Besides binding to the regulatory region of its own promoter, the antitoxin MqsA binds to the promoter regions of *cspD*, *rpoS*, *spy*, and *mcbR*. By repressing these target genes, MqsRA participates in regulation of the general stress response and biofilm formation. Structural analyses have shown that MqsR belongs to the RelE superfamily and MqsR is thus the first known ribosome-independent mRNase belonging to this toxin family.

6.1 Introduction: Discovery of *mqsRA* of *Escherichia coli* K-12

The *mqsRA* locus (genes b3022 and b3021) was previously called *ygiUT*. Transcription of *mqsR* was found to be increased eightfold in biofilms (Ren et al. 2004). On the basis of this observation and the widespread conservation of b3022 homologs in proteobacteria, researchers in the lab of Thomas Wood tested the effect of this gene on biofilm formation and found that it is necessary for cross-species quorum-sensing autoinducer 2 (AI-2)-mediated biofilm growth. Hence, they renamed the gene *mqsR*—the motility quorum-sensing regulator (Gonzalez Barrios et al. 2006). The effect of MqsRA on biofilm growth has recently found a mechanistic explanation (Wang et al. 2011) and is discussed in detail below.

N. Kaldalu · V. Kasari · G. Atkinson · T. Tenson (✉)
Institute of Technology, University of Tartu, Nooruse 1, Tartu, 50411, Estonia
e-mail: Tanel.Tenson@ut.ee

K. Gerdes (ed.), *Prokaryotic Toxin-Antitoxins*,
DOI: 10.1007/978-3-642-33253-1_6,

Meanwhile, *mqsR* was shown to be significantly overexpressed in a dormant subpopulation of *E. coli* enriched for persisters—temporarily non-growing bacteria that are not killed by bactericidal antibiotics [for review see: (Lewis 2010), and Chap. 11]. In a subpopulation of dormant bacteria, identified by the lack of de novo synthesis of ribosomal RNA, *mqsR* was the most highly upregulated gene (Shah et al. 2006). Along with the genes of other TA loci, *mqsR* and *mqsA* were also found to be overexpressed in the persister-enriched subpopulation of ampicillin-refractory cells, isolated via ampicillin-induced lysis of sensitive bacteria (Keren et al. 2004b). From their TA-like gene organization, growth-arresting effect of B3022 and resemblance of MqsA to transcriptional repressors, *mqsR* and *mqsA* were predicted to encode a new TA system (Shah et al. 2006). Independently, *mqsRA*-homologous loci were predicted to constitute a new TA family (Makarova et al. 2009). This prediction was experimentally confirmed by four research groups (Brown et al. 2009; Yamaguchi et al. 2009; Christensen-Dalsgaard et al. 2010; Kasari et al. 2010; Kim et al. 2010) and MqsR was shown to be a GCU-specific endoribonuclease (Yamaguchi et al. 2009; Christensen-Dalsgaard et al. 2010). This chapter gives an overview of the function and structure of MqsRA, with additional discussion of MqsA's auxiliary role in transcriptional regulation.

6.2 Phylogenetic Distribution of *mqsRA*

The MqsR family comprises two phylogenetically distinct groups (Makarova et al. 2009) of which Group 1 is common in gamma and beta-proteobacteria, and can also be found in alpha and delta-proteobacteria, in chlorobi, and in acidobacteria (*Acidobacter*) (Kasari et al. 2010) while the more divergent Group 2 is present in chlorobi and firmicutes, and also in some phages and prophages (Figs. 6.1 and 6.2) (Makarova et al. 2009). However, only one of the toxin active site residues (residue Y81 of *E. coli* MqsR), and none of the antitoxin-interacting residues are conserved in Group 2, raising the possibility that Group 2 is functionally distinct from Group 1 (Fig. 6.2).

Most MqsR-encoding genomes contain a single *mqsR* gene, the exceptions being *Thermoanaerobacter pseudethanolicus,* which codes for two, and *Beggiatoa sp*. and *Geobacter uraniireducens,* which code for three MqsR proteins (Makarova et al. 2009) (Fig. 6.1). As has been observed for other toxin-antitoxin systems (Makarova et al. 2009), the presence of *mqsR* genes differs among closely related bacteria. For example, only about one-half of the sequenced *E. coli* genomes contain this gene (Kasari et al. 2010).

In the TA units, the *mqsR* toxin is closely followed (only a couple of nucleotides in between or overlapping) by an antitoxin coding gene. Two distinct helix-turn-helix (HTH) families of antitoxins are genetically linked to the MqsR toxins (Makarova et al. 2009). The proteins of the first subfamily, represented by MqsA, contain an additional N-terminal Zn finger domain, and are linked to the Group 1

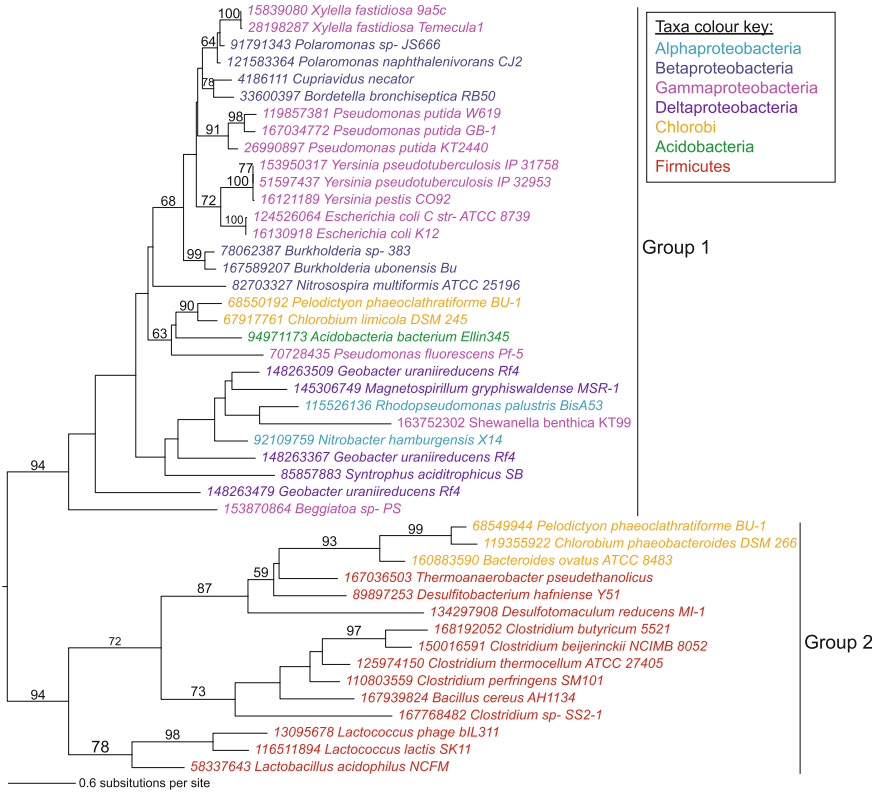


Fig. 6.1 Phylogenetic tree of MqsR homologs belonging to Groups 1 and 2. The phylogeny was constructed from a RaxML (Stamatakis 2006) maximum likelihood analysis of an alignment of MqsR homologs (Makarova et al. 2009) with gap-rich regions excluded (data set dimensions: 96 positions from 45 sequences). A second copy of highly divergent homolog from *Beggiatoa sp.* PS was removed from the dataset before phylogenetic analysis. Branch lengths are proportional to the number of amino acid substitutions (see scale bar in *lower left*) and numbers on branches show percentage bootstrap support values that are greater than 50 %. Taxa are coloured according to the taxonomy in the *inset box*

toxins. The second subfamily (typified by *Lactococcus* phage bIL311 protein Orf21) is linked to Group 2 toxins and is characterized by the fusion of an N-terminal HTH domain and a domain of the GepA (genetic element protein A) family. The latter family together with its Group 2 MqsR counterpart is present mostly in phages and prophages of Firmicutes (Makarova et al. 2009). It is not completely unexpected that MqsR has switched its antitoxin partner in different evolutionary lineages, given that such partner shifts have occurred at least once in most toxin-antitoxin families (Arbing et al. 2010; Leplae et al. 2011).

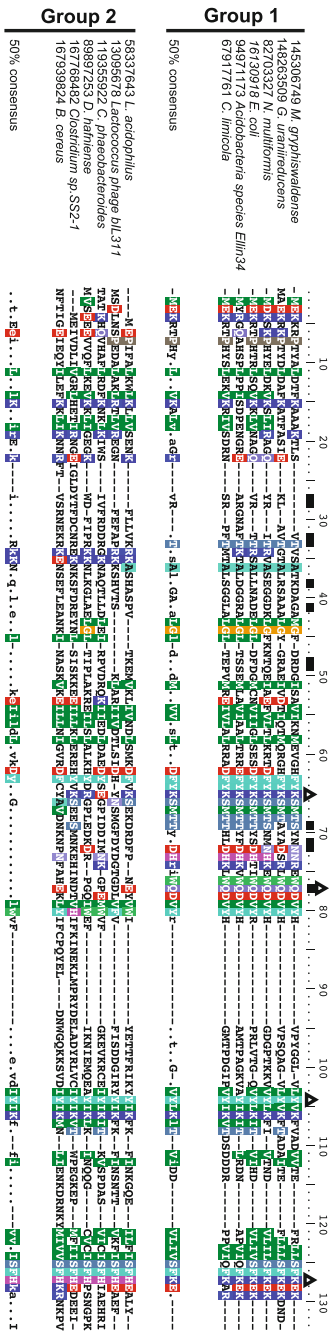


Fig. 6.2 MqsR group1 and 2 alignment. Representative sequences from alignment of MqsR group1 and 2 (Makarova et al. 2009) are shown along with the 50 % consensus sequences for each group, as calculated from the full alignment. *Triangles above the alignment* show the location of residues important for toxicity of *E. coli* MqsR and *squares* show MqsA-interacting residues

6.3 MqsR's Mechanism of Action

MqsR is a 98-residue protein with molecular weight of 11.231 kDa and MqsA is a 131-residue protein of 14.702 kDa. As expected for a toxin-antitoxin system, production of MqsR stops growth of *E. coli,* and production of MqsA can reverse the growth arrest (Brown et al. 2009; Yamaguchi et al. 2009; Christensen-Dalsgaard et al. 2010; Kasari et al. 2010). MqsA is able to resuscitate bacteria after several hours of MqsR expression, showing that MqsR does not kill them (Christensen-Dalsgaard et al. 2010; Kasari et al. 2010). Expression of MqsR causes elongation of bacterial cells (Kasari et al. 2010) and rapid shutdown of protein synthesis (Yamaguchi et al. 2009; Christensen-Dalsgaard et al. 2010; Kasari et al. 2010). Secondary structure predictions suggest that MqsR belongs to the RelE superfamily of RNases and, indeed, it was shown to encode an endoribonuclease (Yamaguchi et al. 2009; Christensen-Dalsgaard et al. 2010). MqsR cleaves model mRNAs in vivo (Yamaguchi et al. 2009; Christensen-Dalsgaard et al. 2010) and in vitro (Yamaguchi et al. 2009). Cleavage by MqsR is translation independent as demonstrated by in vitro experiments (Yamaguchi et al. 2009) and by comparative degradation of a normal mRNA and its start codon-deficient mutant version in living cells (Christensen-Dalsgaard et al. 2010). Thus, MqsR is the first known toxin of the RelE superfamily that cleaves mRNA independent of the ribosome. The endoribonuclease activity of the purified His-tagged MqsR was completely inhibited by the purified MqsA-His (Yamaguchi et al. 2009). Primer extension analysis of the in vivo and in vitro cleaved RNA indicated that MqsR cleaves RNA before or after the G residue in GCU sequences. Cleavage of some GCA/G/C sites could also be detected in vivo (Yamaguchi et al. 2009; Christensen-Dalsgaard et al. 2010). Alanine-scanning mutagenesis of evolutionarily and structurally conserved residues showed that K56, Q68, Y81, and K96 play a role in the MqsR-mediated toxicity and probably comprise the active site of the toxin (Fig. 6.3) (Brown et al. 2009).

6.4 Tertiary Structures of MqsR and MqsA

The structures of *E. coli* MqsR and MqsA have been solved by X-ray crystallography (Brown et al. 2009) (Fig. 6.3). MqsR is a small, globular protein, consisting of a central six-stranded β-sheet (b1-b3-b4-b5-b6-b2) and three adjacent α-helices. The MqsR fold is most similar to the bacterial toxins YoeB and RelE, both of which are ribonucleases and adopt a microbial RNase fold, the RelE-like fold.

MqsA has several specific features compared to the other antitoxins. First, the MqsA protein is larger than MqsR; the only other TA locus that encodes an antitoxin that is larger than its cognate toxin is *hicAB* (Jorgensen et al. 2009). Typically, most free antitoxins have at least one very flexible or unstructured domain that changes its conformation or acquires structure only upon binding to

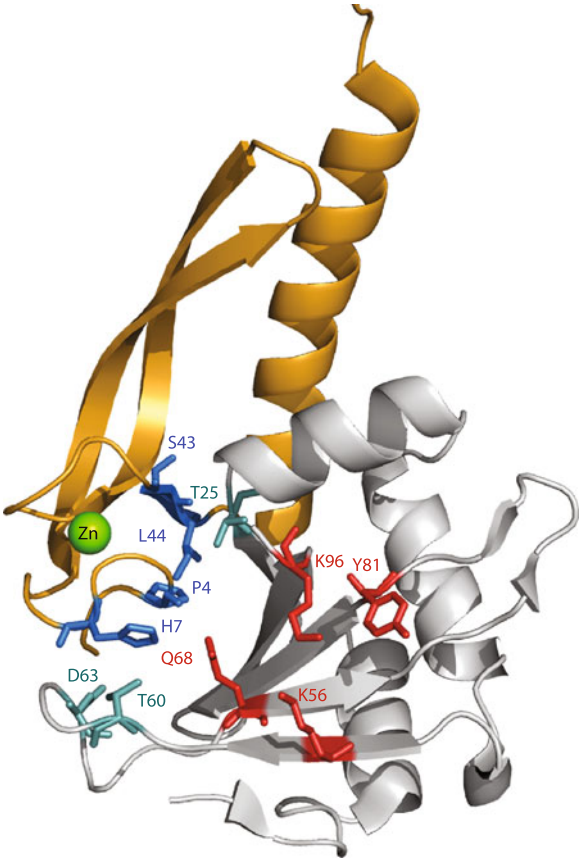


Fig. 6.3 The three-dimensional structure of the *E. coli* MqsR:MqsA-N complex. The N-terminal domain of the antitoxin MqsA is shown in *orange* and the toxin MqsR is in *gray* (PDB ID 3HI2; Brown et al. 2009). Those residues important for toxicity of MqsR are shown as *red sticks*, with other sticks indicating the most strongly conserved sites of MqsR that are involved in interactions with MqsA. The *green sphere* shows the location of the zinc ion

the toxin (Yamaguchi et al. 2011). MqsA, however, is structured throughout and does not seem to change conformation upon toxin binding. The MqsA monomer adopts a novel, elongated fold, composed of two structurally distinct domains (the HTH domain and the zinc-binding domain) connected by a flexible linker. The HTH domain is classified as a member of the XRE family of helix-turn-helix (XRE-HTH) proteins. The zinc-binding domain contains a long five-turn α-helix, a twisted β-sheet, and loops that connect the secondary structure elements and coordinate a zinc ion. The zinc ion serves a structural and not catalytic role, and is coordinated by four cysteines. In spite of the low sequence identity (12 % identity, 18 % similarity among different species) in the N-terminal domains of MqsA proteins, these cysteines are perfectly conserved, suggesting that all MqsA proteins adopt a similar fold (Brown et al. 2009).

Like all known TA systems, both the MqsR:MqsA complex and MqsA alone regulate their own transcription by binding to DNA. MqsA binds DNA predominantly via its C-terminal HTH domain, with direct binding of conserved Asn and Arg residues in the recognition helix to the DNA major groove (Brown et al.

2011). Binding of MqsA induces a large bend in the DNA accompanied by extensive domain rearrangements in the protein. The N-terminal domains of MqsA rotate and twist via a two-residue hinge, changing the protein from a highly extended conformation into a narrow, elongated DNA "clamp." The residues of the N-terminal domains contact the phosphoribose backbone of DNA and thereby enhance binding (Brown et al. 2011).

Unlike the toxin-inhibiting moieties of many other antitoxins, the inhibitory N-terminal domain of MqsA does not neutralize the toxicity of its cognate toxin through direct binding to the active site (Brown et al. 2009). Similarly to the toxins and antitoxins of HipB-HipA, Phd-Doc, and ε-ζ TA systems, binding of MqsA renders MqsR ineffective by inducing substantial conformational changes (Yamaguchi et al. 2011). In the complex of MqsR:MqsA bound to DNA, the toxin active sites of the two MqsR monomers face one another and define a narrow (13–15 Å) channel that makes the active sites inaccessible to mRNA (Brown et al. 2011).

6.5 Autoregulation of *mqsRA* Transcription

Similar to other known type II TA genes, *mqsR* and *mqsA* are co-transcribed and transcriptionally autoregulated (Figure 6.4). However, the order of the genes in the two-gene operon is unusual for TA units: the first gene (*mqsR*) encodes for the toxin and the second gene (*mqsA*) for the antitoxin. Only two other groups of TA loci (*hicAB* and *higBA*) have such gene order (Christensen-Dalsgaard and Gerdes 2006; Jorgensen et al. 2009). Transcription from the P_{MqsRA} promoter starts 18 bp upstream of the start codon of *mqsR* (Christensen-Dalsgaard et al. 2010; Kasari et al. 2010). Two transcripts, which have identical 5′-ends but different 3′-ends are transcribed from the promoter (Kasari et al. 2010). An alternative transcription start site 109 bp upstream of the *mqsR* start codon was mapped using a primer that binds very close to the actual 5′ end of the mRNA (Yamaguchi et al. 2009) and is probably incorrect. Unlike the *higBA* locus of plasmid Rts1, which has an extra *higA* promoter within the *higB* gene (Tian et al. 1996), no separate promoter was found in front of *mqsR* (Yamaguchi et al. 2009; Christensen-Dalsgaard et al. 2010; Kasari et al. 2010).

As is usual for TA systems, MqsA represses transcription of *mqsRA* (Figure 6.4). Ectopically expressed MqsA represses the P_{MqsRA}::*lacZ* fusion about 20-fold (Christensen-Dalsgaard et al. 2010; Kasari et al. 2010). Both MqsA and the MqsR:MqsA complex bind to two palindromic sequences (5′-*ACCT*TTT*AGGT*), which flank the putative −35 box of the promoter. Gel shift assays show that MqsA binds DNA with higher affinity when in complex with MqsR (Brown et al. 2009; Yamaguchi et al. 2009).

Transcription of *mqsRA* is activated in a number of conditions: in biofilms (Ren et al. 2004), in non-dividing persister cells (Keren et al. 2004b; Shah et al. 2006), in response to oxidative stress (Kim et al. 2010), and in response to activation of the HipA toxin that leads to the development of persistence (Kasari et al. 2010).

One explanation for this activation is the decreased level of the transcriptional repressor MqsA due to degradation. Degradation of MqsA by Lon protease has been confirmed under oxidative stress (Wang et al. 2011). Proteolysis of MqsA also results in a relative increase of free MqsR and changes the toxin/antitoxin ratio, which can lead to destabilization of the ternary toxin:antitoxin:promoter DNA complexes. Such a mechanism of induction has been described for *relBE* of *E. coli*, which is transcriptionally activated when the [RelE]:[RelB] ratio is shifted toward high [RelE] (Overgaard et al. 2008) (for details see Chap. 5) and *vapBC* of *Salmonella enterica* (Winther and Gerdes 2012). Derepression by an excess of the toxin has been proposed to be common to TA loci (Overgaard et al. 2008) and appears to be the case for *mqsRA* since ectopic production of MqsR enhances transcription of the chromosomal *mqsRA* operon (Kim et al. 2010). The putative mechanism of such derepression is formation of TA complexes of different stoichiometry and DNA-binding capacity. It is common that toxins and antitoxins work in multimeric complexes that may have a different stoichiometry (Zhang et al. 2003; Overgaard et al. 2008; Arbing et al. 2010). Therefore, it is not surprising that different multimeric complexes of MqsR:MqsA have been purified and described. Both a dimer of dimers, composed of two copies of MqsR and two copies of MqsA (Brown et al., 2009), and a larger complex with the molecular mass of 90 kDa, which likely consists of two MqsR dimers and one MqsA dimer (Yamaguchi et al. 2009) have been reported. The former is similar to the DNA-binding complex of RelE:RelB (Overgaard et al. 2008) and the latter to the purified MazF:MazE complex (Zhang et al. 2003). When separately expressed and purified, MqsR and MqsA both exist as homodimers (Yamaguchi et al. 2009).

Unexpectedly, ectopic production of MqsA also strongly stimulates transcription of chromosomal *mqsR* (Kim et al. 2010). This may be explained by an outcompeting of the MqsR:MqsA complex from the *mqsRA* operator sequence by the overexpressed MqsA, which is intrinsically a weaker repressor.

6.6 Regulation by MqsRA

The MqsRA system has been shown to regulate biofilm formation, curli and cellulose production, stress resistance, motility, and persister formation (Gonzalez Barrios et al. 2006; Kim et al. 2010; Wang et al. 2011; Hong et al. 2012). Surprisingly, deletion of *mqsR* caused transcriptional repression of 239 genes and induction of 76 genes more than twofold in the log-phase cultures of *E. coli* (Kim et al. 2010). Some of these massive changes may be a result of the increase in stability of particular transcripts due to the lack of cleavage by MqsR. However, many effects are due to modulation of stability and DNA-binding capacity of MqsA. MqsA has been shown to bind regulatory sequences of other genes than *mqsRA* and directly regulate their transcription.

6.6.1 Auxiliary Transcriptional Regulation

It is generally thought that antitoxins are specialized proteins that lack any other function except control over their cognate toxins. However, antitoxins as DNA-binding proteins can potentially bind to multiple sites in a genome and directly regulate transcription of many genes. The intrinsic instability of these proteins may enable a coordinated physiological response triggered by the proteolytic machinery.

MqsA can bind specifically not only to its own promoter but also to other promoters in the genome such as the promoters for *mcbR* and *spy* (Brown et al. 2009). This surprising observation demonstrates a somewhat relaxed binding stringency, since these two promoters do not contain canonical binding sites for MqsA. The crystal structure shows that each MqsA monomer binds DNA via residues Asn97 and Arg101, which make base-specific interactions with eight nucleotides including four from one strand (5′-TAAC) and four from the anti-parallel strand (5′-AGGT). The intervening nucleotides of the palindrome seem to have no effect on binding affinity (Brown et al. 2011). The *E. coli* genome contains tens of such palindromes, most of them within coding regions but some in the intergenic and putative promoter regions (Wang et al. 2011). A number of studies from the lab of Thomas Wood have demonstrated that binding of MqsA and the MqsR:MqsA complex to at least some of these sites has a regulatory effect on bacterial physiology (Kim et al. 2009, 2010; Wang et al. 2011; Hong et al. 2012).

Potential binding sites of the MqsR:MqsA complex in biofilm cells have been identified experimentally using a nickel-enrichment DNA micro-array. Among other sites, MqsA:MqsR binds to the promoter regions of *cspD*, *mcbR*, *spy* and *bssR*. A whole-transcriptome analysis revealed that deletion of MqsR from the chromosome and overproduction of MqsR and MqsA induce and repress a large number of genes (Kim et al. 2010). Among those were many stress response-related genes, including *rpoS* that encodes the σ-factor S. σ^S is a key regulator of stress response and subject to complex, multi-level regulation. Another gene affected by MqsA and MqsR is *cspD*, which encodes a stress-induced DNA replication inhibitor. As shown by a gel-shift assay, the MqsR:MqsA complex binds specifically to the *cspD* promoter. In response to oxidative stress, transcription of *cspD* was induced 5 ± 1-fold in wild-type bacteria but did not change in the Δ*mqsRA* mutant (Kim et al. 2010).

The promoter region of *rpoS* contains an MqsA binding site (5′-*ACCT*TGC*AGGT*) verified by gel shift assay. Upon mutating the palindrome to 5′-*ACCT*TGCTCAC, binding of MqsA to P_{rpoS} was abolished. To confirm that MqsA regulates *rpoS* expression through the binding to this palindrome, it was shown that P_{rpoS} with the mutated binding site is no more repressed by the antitoxin. Via repression of *rpoS*, MqsA regulates synthesis of 3,5-cyclic diguanylic acid (c-di-GMP), stress resistance, motility, and biofilm formation (Wang et al. 2011).

The physiological relevance of auxiliary transcriptional regulation by MqsRA is unclear, since only some *E. coli* strains encode this TA system. Such moonlighting

regulation may simply be a side effect that is tolerated by the cell, rather than a functional mechanism. Alternatively, it may have evolved through positive selection and have an adaptive value in certain conditions. At the same time, in some other environments, an absence of the *mqsRA* locus or this auxiliary transcriptional control may give a selective advantage to bacteria. This would create a bet hedging-like situation whereas the *mqsRA*$^+$ and *mqsRA*$^-$ strains are both preserved in microbiota and their relative frequency is determined by the environmental conditions. Transcriptional repression as a secondary function is in fact not limited to the MqsRA TA system. Another antitoxin, DinJ of the YafQ-DinJ TA pair also decreases the levels of σ^S with similar physiological consequences. DinJ binds specifically to the LexA palindrome in front of the *cspE* gene and represses its transcription; cold shock protein CspE, in turn, enhances translation of the *rpoS* mRNA. Inactivation of CspE abolishes the effect of DinJ on RpoS (Hu et al. 2012).

6.6.2 Effects on Biofilm Formation

Enhancement of biofilm formation was the first phenotype described for *mqsR* (Gonzalez Barrios et al. 2006). According to present knowledge, this is brought about by transcriptional regulation of *rpoS* by MqsA (Wang et al. 2011) (Fig. 6.4). *MqsR* functions as a stabilizer of MqsA and a facilitator of its binding to DNA (Brown et al. 2009). The various results that have been obtained from testing the effects of *mqsRA* on biofilm formation (Kasari et al. 2010; Kim et al. 2010) can find their explanation within this scheme.

In the absence of stress, MqsA represses *rpoS*. In response to stress, MqsA is degraded by Lon, increased levels of RpoS activate expression of diguanylate cyclases and increase cellular concentration of c-di-GMP. Another consequence of the increase in RpoS is upregulation of CsgD, a positive regulator of synthesis of the biofilm components cellulose and curli. At the same time, decrease in expression of flagellar regulators FlhDC leads to the loss of motility and initiation of biofilm formation (Wang et al. 2011).

6.6.3 MqsRA and Formation of Persister Cells

TA loci may be primarily responsible for persister formation because deletion of ten type II TA loci of *E. coli* K-12 reduced persistence significantly (Maisonneuve et al. 2011). In a recent study, a mutant of *mqsR* with increased toxicity [comparable to the high persistence mutant *hipA*7 (Moyed and Bertrand 1983)] was found to have increased persister frequency and was used to investigate the transcriptome of persister cells (Hong et al. 2012). The authors noticed the *rpoS*-dependent genes of stress resistance (acid resistance, osmotic resistance and multidrug resistance) were repressed in the high toxicity mutant strain compared to

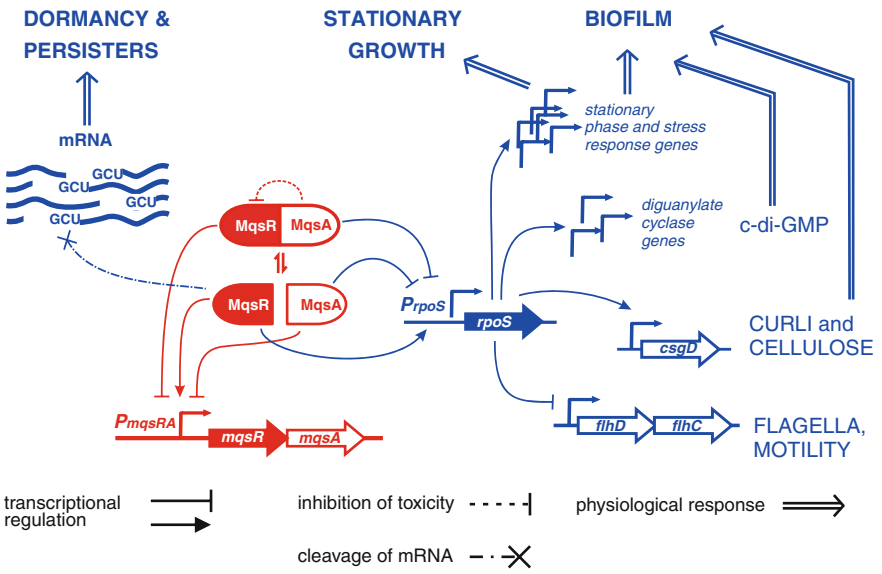


Fig. 6.4 Autoregulation and auxiliary regulation by MqsRA. Autoregulatory interactions are in *red* and all other regulatory interactions are in *blue*. MqsR (toxin) and MqsA (antitoxin), which are encoded by co-transcribed genes, form a tight complex and MqsA inhibits MqsR through direct protein–protein interaction. MqsA, both alone and in complex with MqsR, binds to the regulatory regions of its own promoter (P_{mqsRA}) and the promoter of *rpoS* (P_{rpoS}) to repress transcription. Free MqsR in excess disrupts this DNA–protein interaction and induces transcriptional derepression. MqsA is subject to stress-inducible proteolytic degradation. Proteolysis of MqsA leads to derepression of *rpoS*; promotes biofilm formation through increase of c-di-GMP, curli, and cellulose production; and upregulates stationary phase-specific genes. Alternatively, activation of MqsR causes fragmentation of mRNA and leads cells into dormancy. Upon stress and limitation of resources, bacteria are faced to the choice of developmental pathway and become either growing stationary phase cells or dormant persisters

the wild-type strain. On the basis of this observation, they conceived a model of choice between two developmental pathways—one that enables growth in stressful conditions and depends on *rpoS*, and the other one that leads to dormancy and persister formation and depends on toxin-antitoxin systems. In accordance with this hypothesis, deletion of *rpoS* as well as stress resistance genes *gadB*, *gadX*, *mdtF* and *osmY* increased persistence dramatically (Hong et al. 2012). The testing conditions used in this study might have been critical for the outcome. Cultures were grown to a turbidity of 0.5 at 600 nm and adjusted to a turbidity of 1 at the beginning of antibiotic treatment. As a result, the cultures were approaching stationary phase and were in a stage when persister frequency of wild-type *E. coli* starts to increase rapidly (Keren et al. 2004a) and TA systems become activated (Sangurdekar et al. 2006). At this growth phase, a bulk of the persisters is most probably formed de novo, while in younger batch cultures, persisters are cells of the overnight culture used for inoculation that did not resume growth after transfer

to the fresh medium (Keren et al. 2004a; Joers et al. 2010). For the *hipBA* overactive mutant it has been shown that the persister frequency depends heavily on growth conditions (Luidalepp et al. 2011). These intriguing results certainly inspire further studies to establish the mechanisms of persistence.

Acknowledgments Our research on the TA systems is supported by European Regional Development Fund through the Centre of Excellence in Chemical Biology and by grant 8822 from Estonian Science Foundation.

References

Arbing, M. A., Handelman, S. K., et al. (2010). Crystal structures of Phd-Doc, HigA, and YeeU establish multiple evolutionary links between microbial growth-regulating toxin-antitoxin systems. *Structure, 18*(8), 996–1010.

Brown, B. L., Grigoriu, S., et al. (2009). Three dimensional structure of the MqsR:MqsA complex: A novel TA pair comprised of a toxin homologous to RelE and an antitoxin with unique properties. *PLoS Pathogens, 5*(12), e1000706.

Brown, B. L., Wood, T. K., et al. (2011). Structure of the *Escherichia coli* antitoxin MqsA (YgiT/b3021) bound to its gene promoter reveals extensive domain rearrangements and the specificity of transcriptional regulation. *The Journal of biological chemistry, 286*(3), 2285–2296.

Christensen-Dalsgaard, M., & Gerdes, K. (2006). Two higBA loci in the *Vibrio cholerae* superintegron encode mRNA cleaving enzymes and can stabilize plasmids. *Molecular Microbiology, 62*(2), 397–411.

Christensen-Dalsgaard, M., Jorgensen, M. G., et al. (2010). Three new RelE-homologous mRNA interferases of *Escherichia coli* differentially induced by environmental stresses. *Molecular Microbiology, 75*(2), 333–348.

Gonzalez Barrios, A. F., Zuo, R., et al. (2006). Autoinducer 2 controls biofilm formation in *Escherichia coli* through a novel motility quorum-sensing regulator (MqsR, B3022). *Journal of Bacteriology, 188*(1), 305–316.

Hong, S. H., Wang, X., et al. (2012). Bacterial persistence increases as environmental fitness decreases. *Microbial Biotechnology, 5*(4), 509–522.

Hu, Y., Benedik, M. J., et al. (2012). Antitoxin DinJ influences the general stress response through transcript stabilizer CspE. *Environmental Microbiology, 14*(3), 669–679.

Joers, A., Kaldalu, N., et al. (2010). The frequency of persisters in *Escherichia coli* reflects the kinetics of awakening from dormancy. *Journal of Bacteriology, 192*(13), 3379–3384.

Jorgensen, M. G., Pandey, D. P., et al. (2009). HicA of *Escherichia coli* defines a novel family of translation-independent mRNA interferases in bacteria and archaea. *Journal of Bacteriology, 191*(4), 1191–1199.

Kasari, V., Kurg, K., et al. (2010). The *Escherichia coli* mqsR and ygiT genes encode a new toxin-antitoxin pair. *Journal of Bacteriology, 192*(11), 2908–2919.

Keren, I., Kaldalu, N., et al. (2004a). Persister cells and tolerance to antimicrobials. *FEMS Microbiology Letters, 230*(1), 13–18.

Keren, I., Shah, D., et al. (2004b). Specialized persister cells and the mechanism of multidrug tolerance in *Escherichia coli*. *Journal of Bacteriology, 186*(24), 8172–8180.

Kim, Y., Wang, X., et al. (2009). Toxin-antitoxin systems in *Escherichia coli* influence biofilm formation through YjgK (TabA) and fimbriae. *Journal of Bacteriology, 191*(4), 1258–1267.

Kim, Y., Wang, X., et al. (2010). *Escherichia coli* toxin/antitoxin pair MqsR/MqsA regulate toxin CspD. *Environmental Microbiology, 12*(5), 1105–1121.

Leplae, R., Geeraerts, D., et al. (2011). Diversity of bacterial type II toxin-antitoxin systems: A comprehensive search and functional analysis of novel families. *Nucleic Acids Research, 39*(13), 5513–5525.

Lewis, K. (2010). Persister cells. *Annual Review of Microbiology, 64*, 357–372.

Luidalepp, H., Joers, A., et al. (2011). Age of inoculum strongly influences persister frequency and can mask effects of mutations implicated in altered persistence. *Journal of Bacteriology, 193*(14), 3598–3605.

Maisonneuve, E., Shakespeare, L. J., et al. (2011). Bacterial persistence by RNA endonucleases. *Proceedings of the National Academy of Sciences of the United States of America, 108*(32), 13206–13211.

Makarova, K. S., Wolf, Y. I., et al. (2009). Comprehensive comparative-genomic analysis of type 2 toxin-antitoxin systems and related mobile stress response systems in prokaryotes. *Biology Direct, 4*, 19.

Moyed, H. S., & Bertrand, K. P. (1983). hipA, a newly recognized gene of *Escherichia coli* K-12 that affects frequency of persistence after inhibition of murein synthesis. *Journal of Bacteriology, 155*(2), 768–775.

Overgaard, M., Borch, J., et al. (2008). Messenger RNA interferase RelE controls relBE transcription by conditional cooperativity. *Molecular Microbiology, 69*(4), 841–857.

Ren, D., Bedzyk, L. A., et al. (2004). Gene expression in *Escherichia coli* biofilms. *Applied Microbiology and Biotechnology, 64*(4), 515–524.

Sangurdekar, D. P., Srienc, F., et al. (2006). A classification based framework for quantitative description of large-scale microarray data. *Genome Biology, 7*(4), R32.

Shah, D., Zhang, Z., et al. (2006). Persisters: A distinct physiological state of *E. coli*. *BMC Microbiology, 6*, 53.

Stamatakis, A. (2006). RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics, 22*(21), 2688–2690.

Tian, Q. B., Hayashi, T., et al. (1996). Gene product identification and promoter analysis of hig locus of plasmid Rts1. *Biochemical and Biophysical Research Communications, 225*(2), 679–684.

Wang, X., Kim, Y., et al. (2011). Antitoxin MqsA helps mediate the bacterial general stress response. *Nature Chemical Biology, 7*(6), 359–366.

Winther, K. S., & Gerdes, K. (2012). Regulation of enteric vapBC transcription: Induction by VapC toxin dimer-breaking. *Nucleic Acids Research, 40*(10), 4347–4357.

Yamaguchi, Y., Park, J. H., et al. (2009). MqsR, a crucial regulator for quorum sensing and biofilm formation, is a GCU-specific mRNA interferase in *Escherichia coli*. *Journal of Biological Chemistry, 284*(42), 28746–28753.

Yamaguchi, Y., Park, J. H., et al. (2011). Toxin-antitoxin systems in bacteria and archaea. *Annual Review of Genetics, 45*, 61–79.

Zhang, J., Zhang, Y., et al. (2003). Characterization of the interactions within the mazEF addiction module of *Escherichia coli*. *The Journal of biological chemistry, 278*(34), 32300–32306.

Chapter 7
Type II Toxin-Antitoxin Loci: The *mazEF* Family

Yoshihiro Yamaguchi and Masayori Inouye

Abstract The *mazEF* locus is one of the most extensively characterized toxin-antitoxin (TA) systems. MazF is an endoribonuclease that cleaves RNA at a specific sequence. MazF is conserved in most bacterial and some archaeal species. Since the discovery of MazF in *Escherichia coli*, a number of MazF homologues and other mRNA interferases with different mRNA cleavage specificities have been elucidated. Here we describe their unique biochemical features, the regulatory mechanisms of the MazF activity and presumed physiological roles of MazF in the cells.

7.1 Introduction

Among the cellular targets for the toxins encoded by TA loci, mRNA is the most common (Yamaguchi et al. 2011). Out of 34 TA systems in *E. coli* K-12 identified so far, 10 have been identified to regulate mRNA stability (Yamaguchi and Inouye 2009). Since all these 10 toxins cleave cellular mRNAs, they are termed "mRNA interferases" (Inouye 2006; Yamaguchi and Inouye 2009). Based on how these mRNA interferases cleave mRNAs, they are classified into two types; ribosome-independent mRNA interferases, which can cleave mRNAs in the absence of ribosomes, and ribosome-dependent mRNA interferases, which cleave mRNA only when they associate with ribosomes (Yamaguchi and Inouye 2009). Ribosome-independent mRNA interferases are sequence-specific endoribonucleases

Y. Yamaguchi · M. Inouye (✉)
Department of Biochemistry, Robert Wood Johnson Medical School,
Center for Advance Biotechnology and Medicine,
Piscataway, NJ 08854, USA
e-mail: Inouye@cabm.rutgers.edu

K. Gerdes (ed.), *Prokaryotic Toxin-Antitoxins*,
DOI: 10.1007/978-3-642-33253-1_7,

cleaving mRNA at specific RNA sequences (Yamaguchi and Inouye 2009). The first such enzyme is MazF (MazF-ec) from *E. coli* that cleaves mRNA at ACA sequences to effectively inhibit protein synthesis (Zhang et al. 2003b).

MazE and MazF ("*ma-ze*" means "What is it?" in Hebrew) were found during the determination of the sequence of the *relA* gene encoding 3′,5′-bispyrophosphate (ppGpp) synthase in *E. coli* (Fig. 7.1) (Metzger et al. 1988). The *mazE-mazF* system, which consists of two genes, *mazE* and *mazF*, which are overlapping by one base pair is located downstream of the *relA* gene (Fig. 7.1). Sequence analysis revealed that MazF-ec is homologous to the *pemK* genes on plasmid pR100 (Masuda et al. 1993). In this report, they termed two chromosomal homologs to *pemK* as *chpA* (now termed *mazF*) and *chpB* (now termed *chpBK*), respectively (Masuda et al. 1993).

We became interested quite accidentally in the *mazEF* system during the characterization of GTP-binding proteins or GTPases in *E. coli,* which were assumed to play crucial roles in signal transduction. The first such protein was Era, encoded by the *era* gene (Ahnn et al. 1986; March et al. 1988). Era was found to be an essential GTPase that specifically bound GTP or GDP to hydrolyze GTP to GDP (Ahnn et al. 1986; March et al. 1988). GTP-binding proteins are known to change their conformation from an active (GTP-bound) to an inactive (GDP-bound) form. GTP-binding proteins interact with GTPase activating proteins or guanine nucleotide releasing proteins to modulate the GTPase activity, and with several downstream target proteins for signal transduction. Indeed, we identified a protein called MazG, which interacted with Era (Zhang and Inouye 2002). Subsequently, we characterized MazG as a dual functional enzyme, having both nucleoside triphosphate pyrophosphohydrolase and pyrophosphatase activities (Zhang and Inouye 2002). The *mazG* gene is located downstream of *mazEF* operon (Fig. 7.1). This finding led us to wonder what the function of MazF is, and why it is toxic to the cells. We were clearly lured by the name of MazE, "what is it".

Our study revealed that MazF-ec was an ACA-sequence-specific endoribonuclease that when induced cleaved almost all cellular mRNAs, completely inhibiting cell growth (Zhang et al. 2003b). *mazEF* is one of the best characterized TA systems (Aizenman et al. 1996; Kamada et al. 2003; Li et al. 2006; Metzger et al. 1988; Zhang et al. 2003b, 2005a). It has been shown that in the cells two MazF dimers form a stable heterohexamer with a MazE dimer so that the MazF toxicity is completely suppressed (Kamada et al. 2003, Zhang et al. 2003a, b). However, while the MazF dimer by itself is stable without forming a complex with MazE dimers, the MazE dimers in the absence of MazF are highly sensitive to ATP-dependent serine proteases, ClpAP (Aizenman et al. 1996) and Lon (Christensen et al. 2003) (Fig. 7.1). Under stress conditions, these proteases are induced, which then preferentially eliminate the MazE dimers in the cells to activate the MazF mRNA interferase activity. This results in digesting almost all cellular mRNAs leading cell growth arrest (Zhang et al. 2003b).

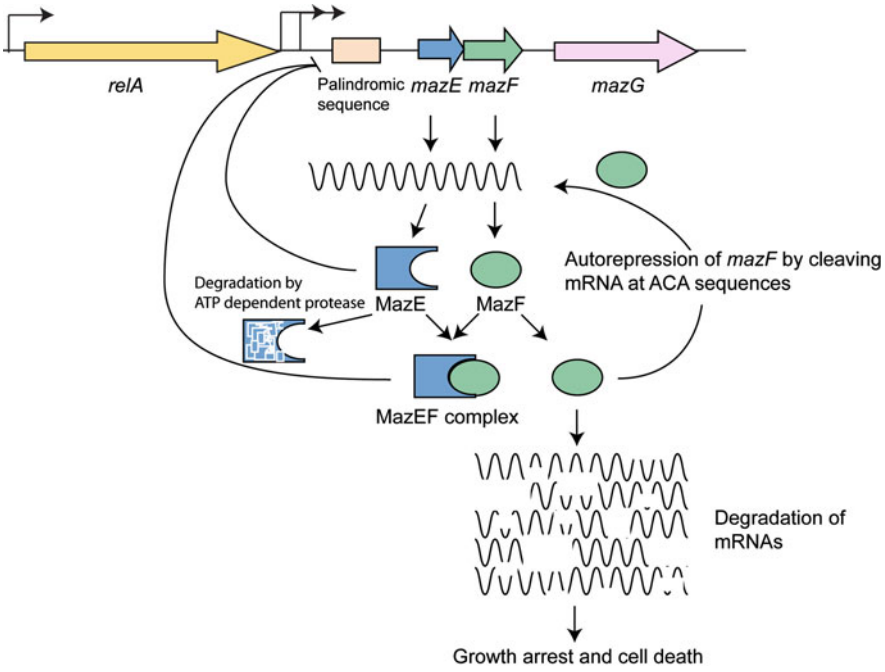


Fig. 7.1 Regulatory mechanisms of the *mazE-mazF* system in *E. coli*. Both *mazE* and *mazF* genes overlapping by one base pair are co-transcribed from the same promoter. Although a same number of MazE and MazF molecules are potentially produced from the transcript, only one MazE dimer is required to neutralize two MazF dimers by forming a heterohexamer complex, which is assumed to secure complete suppression of the MazF toxicity under normal growth conditions. However, under stress conditions, MazE is subjected to be degraded by an ATP-dependent protease because of its highly fragile structure in comparison with the MazF structure. The MazE-MazF complex functions as a repressor for the *mazE-mazF* operon, so that the MazE degradation results in derepression of the transcription of the *mazE-mazF* operon. MazE dimer, by itself, also functions as a weak repressor of the *mazE-mazF* operon. MazE and MazF expression is also negatively regulated at the level of their transcripts, as the *mazF* mRNA contains nine ACA sequences, while the *mazE* transcript contains only two ACA sequences. This two-tier autoregulatory system at the levels of transcription and translation is the key feature for the *mazE-mazF* regulation

To date, a large number of MazF homologs have been identified from bacteria to archaea, and some bacterial species have more than one MazF homologs (Gerdes et al. 2005; Shao et al. 2011). MazF homologs identified so far specifically cleave mRNAs at either three-, five-, or seven-base specific sequences (Yamaguchi et al. 2011).

7.2 MazF in *E. coli* (MazF-ec)

7.2.1 Determination of the Target of MazF-ec

To identify the cellular target of MazF-ec, a cell-free system prepared from *E. coli* BW25113 cells carrying pBAD-*mazF*-ec (inducible in the presence of arabinose) permeabilized by toluene treatment was used (Zhang et al. 2003b). It was found that ATP-dependent [^{35}S] methionine incorporation was completely inhibited when cells were preincubated for 10 min in the presence of arabinose before toluene treatment. However, the incorporation of [α-^{32}P] dTTP and [α-^{32}P] UTP was not significantly affected under similar conditions, demonstrating that MazF-ec inhibits protein synthesis, but not DNA replication or RNA synthesis (Zhang et al. 2003b).

On the basis of these observations, Northern blot analysis of total cellular RNA extracted at different time intervals after arabinose induction of MazF-ec was carried out (Zhang et al. 2003b). For the two genes tested, the *ompA* mRNA was observed only at the zero time point, and the *secE* mRNA was detected only up to 2.5 min after *mazF*-ec induction, indicating that MazF-ec was an endoribonuclease targeting cellular mRNAs. From the analysis of the polysome pattern of *E. coli* BW25113 cells carrying the pBAD-*mazF* plasmid, it was found that MazF disrupted polysomes, either by inhibiting translation initiation and/or by degrading mRNAs (Zhang et al. 2003b). Furthermore, purified MazF-ec inhibited protein synthesis in an *E. coli* cell-free RNA/protein synthesis system (Zhang et al. 2003b).

To further characterize the specificity of the MazF-ec endoribonuclease activity, primer extension was carried out to determine the precise cleavage site sequences. By comparing the cleavage sites in the *mazG* and *yeeW* mRNAs, it became evident that an ACA sequence was indeed present in all of the cleavage sites (Zhang et al. 2003b). The mRNAs were cleaved between the first A and the second C residues in the ACA sequences, suggesting that MazF was an endoribonuclease that specifically recognizes the ACA sequence.

A 30-base RNA sequence from the *mazG* mRNA, containing an ACA sequence in the middle was synthesized to further characterize RNA cleavage by MazF-ec (Zhang et al. 2003b). The synthetic RNA was completely cleaved, yielding shorter distinct fragments. When the RNA substrate was annealed with an antisense DNA, the RNA cleavage was inhibited. A similar result was also obtained from an antisense RNA, indicating that MazF-ec cannot cleave the ACA sequence in RNA/DNA or RNA/RNA duplexes. It was also found that MazF-ec could not cleave single-stranded DNA with the same 30-base sequence as the RNA substrate, indicating that MazF-ec is an ACA-specific endoribonuclease for single-stranded RNA (Zhang et al. 2003b).

7.2.2 Mechanism of RNA Cleavage by MazF

Using a synthetic 15-base RNA, it was shown that the hydrolysis reaction mediated by MazF could occur at either the 5′ or 3′ phosphodiester linkage of the first A residue in the ACA sequence, yielding a free 5′-OH group on the 3′ cleavage product, and thus a 3′-phosphate (or a 2′, 3′-cyclic phosphate) on the 5′ cleavage product (Zhang et al. 2005a). These data were further substantiated by MALDI-mass spectrometric analysis of the MazF-digested 15-mer RNA. MALDI-mass spectrometric analysis after MazF digestion of shorter (13-, 11-, and 7-base) RNAs revealed that (i) every substrate was cleaved at the 5′-end to the first A residue of the ACA sequence, (ii) the resulting 5′ products had almost exclusively a 2′,3′-cyclic phosphate at their 3′-end, and (iii) all of the resulting 3′ products contained a free 5′-OH group. These results further confirmed that MazF was an endoribonuclease that cleaves a phosphodiester linkage at the 5′ side of a phosphodiester bond. Using a 12-base RNA, it was shown that there was no specific requirement for the residue directly upstream of the ACA sequence for the MazF cleavage. All of the 5′ cleavage products contained a 2′,3′-cyclic phosphate at their 3′ end, indicating that the 2′-OH group at the cleavage site played an essential role for the hydrolysis reaction as in the case of ribonuclease A. This requirement was further confirmed using RNAs in which the 2′-OH group was modified with a methyl group at different sites (Zhang et al. 2003b).

Although it seems to be evident that the MazF-ec recognition sequence is ACA on the basis of the results described above, a report has been published claiming that the MazF-ec recognition triplet is NAC (N is preferentially A and U) rather than ACA (Munoz-Gomez et al. 2004). In this report, the authors used three short RNAs synthesized by T7 RNA polymerase, corresponding to the HIV-1 TAR RNA, the CopT RNA of plasmid R1 and its antisense, and the CopA RNA. From the analysis of MazF-cleaved products by primer extension experiments, the authors concluded that MazF cleaved these RNAs between N and A of the NAC sequences. The authors speculated that the difference between their results and the results from our laboratory may be due to His-tag, as they used non-His-tagged MazF-ec. However, this seems highly unlikely, since identical results were observed in in vitro (with use of His-tagged MazF-ec) and in vivo (with non-His-tagged MazF-ec) in our experiments (Zhang et al. 2003b, 2005a). It appears that these authors were detecting minor cleavage sites with these three small RNAs having very limited sequences in vitro. To support this notion, it should be noted that there is only one ACA site in the CopT RNA and no other ACA sequences in other RNAs. Most notably, the ACA sequence in the CopT RNA was indeed cleaved with the least amount of MazF-ec used in their experiment. At this concentration, none of the other cleavage sites were detected.

The fact that MazF-ec is unable to effectively cleave triplet sequences other than ACA was further shown with use of penta-oligonucleotides, the U1A2C3A4U5, in which all possible bases at each position were substituted to other bases (Zhang et al. 2005a). Any base substitutions in the ACA sequence (A2, C3, or A4) completely

blocked RNA cleavage by MazF-ec, whereas any base changes at residues U1 and U5 did not interfere with the RNA cleavage. These results clearly demonstrate that the MazF-ec recognition sequence is ACA, not UAC.

It was also reported that MazF-ec inhibits translation by a ribosome- and codon-dependent mechanism in a very similar manner to RelE (Pedersen et al. 2003; Christensen et al. 2003), while another study showed that MazF-ec cleaves mRNAs in a totally codon-independent manner, since MazF-ec cleaves mRNAs equally well at the same site both in vivo and in vitro (without the addition of ribosomes) (Zhang et al. 2003b). The codon-independent mRNA cleavage by MazF-ec was further substantiated by using small synthetic RNAs, as short as a seven-base RNA, which were able to serve as MazF-ec substrates (Zhang et al. 2005a). In the work by Christensen et al. (2003), they apparently missed a distinct cleavage band of the *lpp* mRNA by MazF-ec, which exists very close to the 5′-end. On the basis of the sequencing ladder provided in that report, the cleavage site can be identified as ACA located in the 5′-untranslated region.

The sequence analysis of the *E. coli* genome reveals that *E. coli* genes contain on average 11 ACA sequences per gene (Zhang et al. 2003b). Interestingly, the 51 genes in the *E. coli* genome contain no ACA sequences. It remains to be determined whether these ACA-less genes play important physiological roles in cells producing MazF-ec under stress conditions. Notably, among the ACA-less genes, there are two cold-shock genes, *cspB* and *cspG*, which are important for cold adaptation (Xia et al. 2001) suggesting that these RNA chaperones may also play a role in the cells whose growth is inhibited as a result of MazF induction. It is intriguing to note that the *mazF-ec* mRNA itself contains as many as nine ACA sequences in its rather short ORF (only 333 bases long), among which four ACA sequences are clustered in the middle of the ORF, while the *mazE* ORF (249 bases long) contains only 2 ACA sequences. As expected, the *mazF* mRNA is susceptible to MazF-ec in vitro (Zhang et al. 2003b), indicating that *mazF-ec* expression is negatively autoregulated by its own gene product (Fig. 7.1). It is notable that when the ACA sequences in the *mazF-ec* ORF were replaced with MazF-resistant sequences (without changing the amino acid sequence of MazF), the MazF-ec toxicity increased significantly (Suzuki and Inouye, unpublished data). It should be noted that although MazF induction completely blocks cell growth as most of cellular RNA are digested by MazF, the cells are still fully metabolically active since the MazF-induced cells are able to efficiently produce a protein from an ACA-less mRNA. This system is called the Single-Protein Production system (Suzuki et al. 2005) as discussed in a later section.

7.2.3 *MazE-ec and MazF-ec Structure and Function in* E. coli

The crystal structure of the MazF-ec:MazE-ec complex has been determined, revealing that a symmetric central MazE-ec dimer uses two unstructured C-terminal extensions, each of which binds to a dimer of MazF-ec, forming a stable 2:4

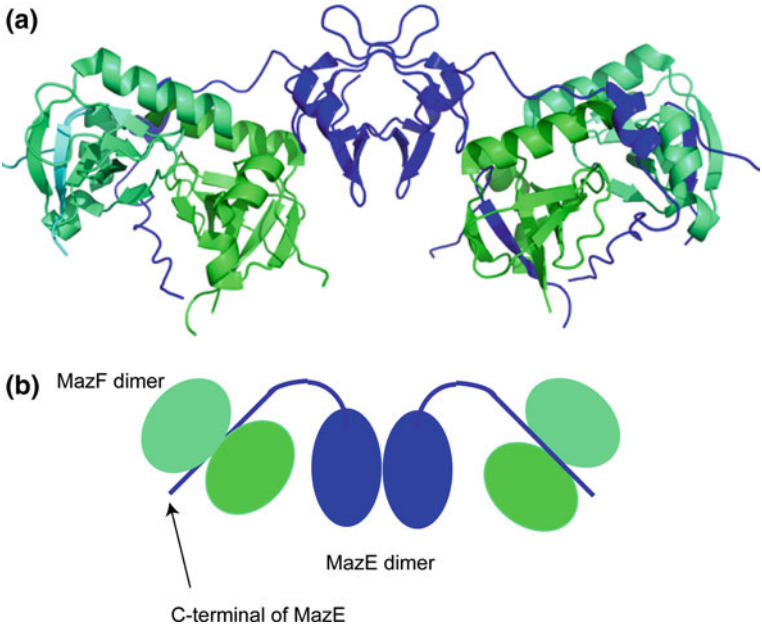


Fig. 7.2 The crystal structure of the MazE-MazF complex. **a** Crystal structure of the MazE-MazF complex. The central MazE dimer is represented by *blue* and the MazF dimer by *green* and lyme. **b** Schematic representation of the structure of the MazE-MazF complex

heterohexamer complex ($MazF_2$-$MazE_2$-$MazF_2$) (Fig. 7.2) (Kamada et al. 2003). This 1-to-2 ratio of MazE-ec to MazF-ec in the complex is worth noting as it is a highly effective manner for toxin neutralization. Interestingly, the unstructured C-terminal region of MazE-ec is highly negatively charged and covers the entire cleft of one of the symmetric concave interfaces created between two MazF-ec molecules in a MazF-ec dimer to inhibit the MazF-ec activity. The highly acidic or negatively charged C-terminal MazE-ec extension is considered to mimic RNA substrates binding to the MazF-ec active center (Zhang et al. 2003b, 2005a). NMR structural studies show that the MazF-ec homodimer containing two symmetric RNA binding sites is capable of interacting with a 24-residue MazE-ec C-terminal peptide (Li et al. 2006). There is negative cooperativity between these two sites, so that a single MazE-ec molecule occupying one of the identical binding sites on a MazF-ec dimer affects the conformation of both sites at the same time to efficiently inhibit the MazF-ec activity (Li et al. 2006). At present, however, it is not known how the *E. coli* MazF-ec dimer specifically recognizes and cleaves the ACA sequence.

7.3 MazF Homologs in Bacteria and Archaea

As mentioned earlier, there are two chromosomal *pemK* homologs in *E. coli* K-12, MazF-ec and ChpBK (Masuda et al. 1993). Since the discovery of MazF-ec, many other MazF homologs and other mRNA interferases with different mRNA cleavage specificities have been found (Yamaguchi and Inouye 2009; Yamaguchi et al. 2011). *Mycobacterium tuberculosis* has at least seven MazF homologs (MazF-mt1 to -mt7), four of which (MazF-mt1, -mt3, -mt4, and -mt6) cause cell growth arrest when induced in *E. coli* (Zhu et al. 2006). It has also been demonstrated that a MazF homolog from *Myxococcus xanthus* (MazF-mx) is essential for the programmed cell death of this developmental bacterium during the fruiting body formation (Nariya and Inouye 2008). Recently, it was shown that MazF-hw from a halophilic archaeon from the Sinai Peninsula, *Haloquadratum walsbyi*, is a seven-base cutter cleaving RNA at UUACUCA sequences (Yamaguchi et al. 2012).

7.3.1 ChpBK in E. coli

The ChpBK TA system consisting of *chpBI* and *chpBK* encode antitoxin ChpBI and toxin ChpBK, respectively. ChpBK consists of 116 amino acid residues, and its sequence shows 35 % identity and 52 % similarity to *E. coli* MazF (Zhang et al. 2005b). Interestingly when ChpBK expression is induced, the total protein synthesis is reduced only by approximately 60 %, and a low level of cellular protein synthesis is observed even at 60 min after ChpBK induction (Zhang et al. 2005b). In contrast, induction of MazF-ec under similar conditions almost complete inhibition of protein synthesis in vivo is observed (Zhang et al. 2003b), indicating that ChpBK is less detrimental to the cells than MazF-ec, although the induction of both ChpBK and MazF results in cell growth inhibition.

The ChpBK-mediated cleavage sites in the *era* and *mazG* mRNAs in vivo were identified as XACY, where the cleavage occurs either before or after the A residue (Zhang et al. 2005b). C-residues were notably absent from the X and Y positions in these in vivo cleavage sites. The same cleavage sites were identified by in vitro experiments with purified ChpBK. To further pinpoint the cleavage specificity of ChpBK, a U1A2C3A4 sequence in the *mazG* mRNA was mutated to all possible bases at each position (Zhang et al. 2005b). Any base changes of the AC sequence (A2 and C3) completely blocked the cleavage by ChpBK, while the U1 position could be substituted with any other bases, although U or C was preferred over A or G. When A4 was replaced with a U-residue, the RNA was effectively cleaved, while its substitution with a G-residue significantly reduced RNA cleavage. On the other hand, the A4 substitution with C completely inhibited the RNA cleavage. Together, ChpBK is a sequence-specific endoribonuclease that cleaves mRNAs at ACY sequences (where Y is G, A, or U, but not C). Furthermore, using a synthetic 11-base RNA, ChpBK was found to cleave the substrate, yielding a 2′,3′-cyclic

phosphate at the 3′ end of the 5′ product and a 5′-OH group at the 5′ end of the 3′ product, in a manner identical to that demonstrated for MazF-ec (Zhang et al. 2005b).

It is interesting to note that complete cleavage of the *mazG* mRNA by MazF-ec was observed in 5 min after induction, as judged from the density of the band of the full-length mRNA, while the cleavage by ChpBK was much slower and a substantial amount of the full-length mRNA was still visible even at 90 min after induction (Zhang et al. 2005b). This again supports the notion that ChpBK is intrinsically a less potent toxin than MazF-ec. Interestingly, the pI value of ChpBK is 5.2, while the pI values of MazF-ec and PemK are basic at 8.3 and 11.0, respectively. The weaker toxic effect of ChpBK may be at least partially attributed to its acidic pI value (the basic property of MazF may help the ionic interaction with acidic RNA substrates) and important for its physiological role in the cells, since ChpBK may arrest cell growth in a less damaging manner than MazF-ec.

7.3.2 PemK in E. coli

The *pemI-pemK* TA system and the *kis-kid* TA system are involved in the stable maintenance of two closely related incFII low copy plasmids, plasmid R100 (Tsuchimoto et al. 1988, 1992) and plasmid R1 (Bravo et al. 1987; Ruiz-Echevarria et al. 1991), respectively. These two systems turned out to be identical (Engelberg-Kulka and Glaser 1999). Kid (PemK) has been shown to inhibit ColE1 replication, acting on the initiation of DNA synthesis, but does not inhibit P4 DNA replication in vitro (Ruiz-Echevarria et al. 1995). Toxin Kid (PemK) and antitoxin Kis (PemI) have been shown to function not only in bacteria but also efficiently in a wide range of eukaryotes. Kid inhibits cell proliferation in yeast, *Xenopus laevis*, and human cells, and the inhibition is released by Kis (de la Cueva-Mendez et al. 2003). These results suggest that there is a common target for Kid in both prokaryotes and eukaryotes.

A primer extension assay performed on the *mazG* and *era* mRNA partially digested by PemK showed that PemK was a ribonuclease that cleaved primarily at the 5′ or 3′ side of the A-residue in the UAH sequence (where H is C, A, or U), with one exception that the cleavage was observed between U and G residues in UGC (Zhang et al. 2004). The sequence UAC appeared in 11 out of the 18 cleavage sites determined. This specificity was further confirmed with a synthetic 30-base RNA containing one UAC sequence, where the RNA was cleaved equally at either the 5′ or 3′ side of the A-residue in the UAC sequence. However, in the primer extension experiment using the full-length *mazG* mRNA in vivo, this UAC sequence was cleaved only at the 5′ side of the A-residue. The reason for this difference between in vitro and in vivo is not clear at present. The PemK cleavage of the 30-base RNA substrate is blocked by either antisense RNA or DNA, indicating that PemK is a ribosome-independent endoribonuclease that cleaves single-stranded RNA at UAC sequences.

Based on the fact that Kid (PemK) inhibits ColE1 replication at the initiation stage in vitro, but has no significant effect on the P4 DNA replication, DnaB was proposed as the target for the inhibitory action of Kid (Ruiz-Echevarria et al. 1995). However, the ColE1 replication is known to be initiated by RNA II and inhibited by RNA I (Cesareni et al. 1991; Davison 1984), whereas the P4 DNA replication is mainly regulated by protein (Briani et al. 2001), ribonucleases are expected to play a role in the control of the ColE1 plasmid replication (Jung and Lee 1995). Indeed, RNA II contains several UAC sequences, importantly two of which exist in the loop regions of the first and second stem-loop structures (Tomizawa 1984). Therefore, the observed inhibition of ColE1 DNA replication by Kid is probably due to the degradation of RNA II by the Kid endoribonuclease activity. Furthermore, the fact that Kid inhibits the growth of various eukaryotic cells (de la Cueva-Mendez et al. 2003) can be readily explained by its mRNA interferase activity against cellular mRNAs.

7.3.3 MazF Homologs from M. tuberculosis

M. tuberculosis contains seven MazF homologs, and at least four of them show toxicity, as they cause cell growth arrest when induced in *E. coli* (Zhu et al. 2006). These MazF homologs have been shown to be sequence-specific mRNA interferases as described below (Zhu et al. 2006, 2008).

7.3.3.1 MazF-mt1

When expression of MazF-mt1 was induced in *E. coli* in vivo, a specific cleavage of the *era* mRNA between a U- and an A-residue was detected by primer extension in a time-dependent manner (Zhu et al. 2006). The identical cleavage site was detected in vitro using the *era* mRNA synthesized by T7 RNA polymerase and purified MazF-mt1. In addition to the site reported for the in vivo cleavage (CU^ACC; ^ represents cleavage site), a weak cleavage site (UU^ACA) was also detected in vitro. For further biochemical characterization, a synthetic 15-base RNA was used, which was cleaved between U and A in the UAC sequence by purified His_6MazF-mt1 (Zhu et al. 2006). When the DNA and RNA substrates were mixed and treated with the enzyme, only the RNA substrate was cleaved, indicating that MazF-mt1 was an endoribonuclease. To further test the cleavage specificity, five 11-base RNA substrates were synthesized, one with the UACA sequence in the center (5′-AUAUACAUAUG-3′) and the others in which every position in UACA was substituted with a G residue. The first U and the second A residues could not be replaced with G. The third C residue also appears to be important, as the cleavage between U and A was significantly reduced when this C was replaced with G. On the other hand, the fourth A residue was replaceable with G, demonstrating that MazF-mt1 was an endoribonuclease that preferentially cleaves UAC sequences (61).

7.3.3.2 MazF-mt3

In the earlier work with MazF-ec (Zhang et al. 2003b), ChpBK (Zhang et al. 2005b), PemK (Zhang et al. 2004) and MazF-mt1, and MazF-mt6 (Zhu et al. 2006), in vivo and in vivo primer extension was sufficient to determine their cleavage site specificities, since all these mRNA interferases recognize only three-base sequences in mRNAs. However, the cleavage site specificity increase from three to four bases or longer, the determination of the specific cleavage site sequences become more difficult. This was the case for MazF-mt3 when in vivo primer extension experiments were used. To circumvent this problem, a general method was developed using the 3.5-kb MS2 phage RNA as a RNA substrate (Zhu et al. 2008). In this method, the use of CspA, an RNA chaperone is essential as CspA unwinds all the secondary structures in the MS2 RNA, which are resistant to mRNA interferase cleavage. Using this method, the MS2 RNA was indeed cleaved at several sites. In order to cover the entire MS2 RNA sequence, a total of 22 primers were designed for the primer extension experiments. In most cases the addition of CspA, significantly enhanced RNA cleavage. As a result, eight cleavage sites were identified, whose consensus sequence was found to be CUCCU. MazF-mt3 cleaves between the U-residue in the second position and the C-residue in the third position (CU^CCU) (Zhu et al. 2008). In addition to the eight CUCCU sites, four other cleavage sites were observed, whose consensus sequence was UUCCU. MazF-mt3 cleaves these sites between the U-residue in the second position and the C-residue in the third position (UU^CCU).

It was speculated that the pentad sequences identified as the cleavage sites for MazF-mt3 (CU^CCU and UU^CCU) might be involved in the stringent control of the synthesis of specific proteins under particular conditions, either by protecting mRNAs from being cleaved (by having no pentad sequences or by containing them much less frequently) or by being more sensitive to the MazF-mt3 endoribonuclease activity (by having many of the pentad sequences in an mRNA) (Zhu et al. 2008). Indeed, it was found that there are certain genes which have the MazF-mt3 pentad recognition sequences at a much lower frequency than expected. Out of 10 genes with the least number of pentad sequences, four were identified as members of the PPE gene family, which has been proposed to play a role in the immunopathogenicity of *M. tuberculosis* (Brennan and Delogu 2002). Three of these four genes, PPE54, PPE5553, and PPE56, are in a single locus on the chromosome of *M. tuberculosis,* although they do not appear to be part of a single operon. All four PPE genes encode members of the major polymorphic tandem repeat subfamily of PPE proteins (PPE_MPTR) (Gey van Pittius et al. 2006). The under-representation of MazF-mt3 pentad recognition sequences in specific members of the PE and PPE family suggests that the mRNAs for these genes are relatively resistant to MazF-mt3 compared with most of other mRNAs encoded from the *M. tuberculosis* genome.

7.3.3.3 MazF-mt6

MazF-mt6 preferentially cleaves mRNA in U-rich regions and after a U-residue. (U/C)U^ (A/U)C(U/C) has been identified as a consensus cleavage sequence for MazF-mt6. Cleavage also occurs after G or A-residues in some cases, although all cleavage sites contain UU, UC, or CU pairs (Zhu et al. 2006).

7.3.3.4 MazF-mt7

Specific cleavage sites for MazF-mt7 were also identified using the MS2 RNA-CspA system (Zhu et al. 2008). Most of the cleavage sites contain the sequence UCGCU, where MazF-mt7 cleaves between the first U and the second C residues (U^CGCU). Some cleavage sites have a one-base difference from the UCGCU consensus and a few have a two-base difference. All these cleavage sites, however, share the central G-residue and most of them also have a following C-residue. MazF-mt7 is thus an mRNA interferase that also recognizes a pentad sequence; however, it appears to be less stringent than MazF-mt3.

Similar bioinformatics analysis to that described for the MazF-mt3 recognition sequence was carried out to search for *M. tuberculosis* mRNAs that are either resistant or hyper-susceptible to cleavage by MazF-mt7 (Zhu et al. 2008). Of the top 10 predicted MazF-mt7-resistant genes, four belong to the PE or PPE family. Three of the 10 genes are PPE_MPTR and one is a PE_PGRS family member. Interestingly, both PPE55 and PPE56 are resistant, not only to MazF-mt7, but also to MazF-mt3. PPE55 is known to be highly immunogenic, and is expressed during incipient, subclinical *M. tuberculosis* infection. PPE56 is highly homologous (67 %) to PPE55 (Singh et al. 2005). Based on these data, it has been proposed that the expression of different PE and PPE family proteins may be regulated at certain stages of infection by mRNA interferases during *M. tuberculosis* pathogenesis (Zhu et al. 2006, 2008).

The function of most genes in the PE/PPE family is not well understood at present; however, a number of these proteins appear to localize to the mycobacterial cell wall or cell surface, and several have been shown to provoke an immune response in vivo or play a role in host–pathogen interactions (Basu et al. 2007; Brennan and Delogu 2002; Denny and Smith 2004; Mishra et al. 2008). In addition, members of these families have been shown to be differentially expressed in different tissues or at distinct times during growth in vitro (Dheenadhayalan et al. 2006). Thus, to the extent that MazF-mt3 and MazF-mt7 can affect the production of specific PE or PPE family proteins in vivo, these mRNA interferases could play a role in the pathogenesis of tuberculosis. Members of the PPE_MPTR superfamily are present in multiple copies only in *M. tuberculosis* and its close relatives, and are totally absent in the fast-growing nonpathogenic mycobacteria, including *M. smegmatis*. This finding suggests that these genes may have been positively selected during the evolution of these slow-growing pathogens, raising an interesting possibility that the target specificity of these mRNA interferases might have

played a role in the evolutionary process to expand the PE/PPE family in the slow-growing pathogenic mycobacteria (Zhu et al. 2008).

Interestingly, all these toxin genes appear to be co-transcribed with an overlapping upstream gene, as in the case of the *E. coli mazE-mazF* operon. However, none of these upstream gene products are homologous to MazE. Therefore, it remains to be determined whether these are antitoxins against their cognate toxins. The answer to this question may provide important insights into the role of *M. tuberculosis* MazF homologs in the pathogenicity of this human pathogen.

7.3.4 *MazF-mx in* Myxococcus xanthus

Surprisingly, not all *mazF* genes are co-transcribed with a cognate antitoxin gene. In some bacteria, the gene encoding a MazF homolog exists as a single gene without an upstream (or downstream) antitoxin gene. It has been speculated that the expression of these solitary toxin genes has to be tightly regulated in the cells and/or their cognate antitoxin gene(s) must exist somewhere else on the genome. It was shown that *M. xanthus* MazF (MazF-mx) is encoded by a monocistronic operon without any cognate antitoxin gene (Nariya and Inouye 2008). Genomic analysis on the *M. xanthus* genomic database (TIGR) suggests that there is only a single MazF homolog (MazF-mx) with no identifiable MazE homolog. MazF-mx has 24 % identity and 58 % similarity to *E. coli* MazF.

Using a yeast two-hybrid screen with MazF-mx as a bait and an *M. xanthus* genomic library, MrpC was found as a candidate of antitoxin for MazF-mx. Interestingly, MrpC is a member of the CRP transcription regulator family and its gene is located 4.44 Mbp downstream of the *mazF-mx* gene on the chromosome. Importantly, the *mrpC* gene is essential for *M. xanthus* development (Sun and Shi 2001) and is a key early developmental transcription activator of the gene encoding FruA, another essential developmental regulator (Ueki and Inouye 2003). The MrpC and MazF interaction was also demonstrated by pull-down assays (Nariya and Inouye 2008). Severe cell toxicity by MazF-mx was observed in the *ΔmrpC* strain when induced in *M. xanthus* (Nariya and Inouye 2008). In particular, MazF-mx expression is toxic in the absence of MrpC expression, confirming the notion that MrpC functions as an antitoxin to MazF-mx.

Interestingly, however, MazF-mx expression did not exhibit strong cellular toxicity in *E. coli*, suggesting that MazF-mx may cleave mRNAs at more specific sites than *E. coli* MazF-ec. When MS2 phage RNA was incubated with purified MazF-mx, it was cleaved into two major bands with many minor bands, suggesting that MS2 RNA may contain only one preferential cleavage site for MazF-mx, which was determined to be GAGU^UGCA. Preincubation of MazF-mx with MrpC resulted in a complete inhibition of the cleavage. When higher concentrations of MazF-mx were used, new cleavage sites appeared at AUGU^CAGG, ACGU^AAUA, and ACGU^AAAG, with several other minor cleavage sites. Interestingly, all cleavage sites are located in predicted single-stranded regions,

consistent with the previous finding that MazF-ec cleaves only single-stranded RNAs (Zhang et al. 2003b, 2005a). From the alignment of all cleavage sites identified, the preferred cleavage sequence for MazF-mx was identified to be GU^UGC, in which the first G-residue can be replaced with an A-residue.

7.3.5 MazF in Gram-Positive Bacteria

Bacillus subtilis also contains a MazF homolog, MazF-bs (EndoA) (Pellegrini et al. 2005). Using BLAST search, it was found that MazF-bs has significant identity (62.8–94 %) and similarity (79.3–96.6 %) to MazF homologs from various Gram-positive bacteria such as *Staphylococcus aureus* (Fig. 7.3). However, MazF-bs has low identity (19.2–37.3 %) and similarity (38.3–55.9 %) to MazF homologs from Gram-negative bacteria. Using the BLASTCLUST program, it was also found that MazF-bs is widely prevailing in Gram-positive bacteria; a total of 102 MazF homologs from 99 strains (61 species), of which 97 strains are Gram-positive bacteria (Park et al. 2011). This result showed that MazF-bs homologs mainly exist in Gram-positive bacteria and are widely distributed within Gram-positive bacterial species. As the TA systems are proposed to have invaded into the bacterial genomes using the horizontal gene transfer mechanism (Van Melderen and Saavedra De Bast 2009), MazF-bs homologs might have been preferentially spread by the same mechanism within Gram-positive bacterial species. Supporting this notion, even if *S. aureus* and *Halothermothrix orenii* are located far from *B. subtilis* in the phylogenetic tree, both MazF-sa and MazF–ho have significantly high homology with MazF-bs (Park et al. 2011).

Although there is a very high structural similarity between MazF-bs and MazF–ec (Pellegrini et al. 2005), MazE-ec could not neutralize the MazF-bs toxicity, suggesting that each antitoxin has been co-evolved with its cognate toxin. *S. aureus* contains one MazF homolog, MazF-sa, which was shown to cleave mRNA at XUUX (X is A, C, or G) (Fu et al. 2007, 2009). However, MazF-sa was found to have more of a stringent cleavage specificity, cleaving mRNAs at UACAU sequences (Zhu et al. 2009). Although it was earlier reported that MazF-bs cleaves RNA at UAC (Pellegrini et al. 2005), the consensus cleavage sequence of MazF-bs has been recently determined to be U^ACAU with use of MS2 RNA as substrate (Park et al. 2011). Since MazF-bs has significantly high homology with other MazF homologs from Gram-positive bacteria, these MazF-bs homologs may have the same cleavage specificity cleaving RNA at UACAU.

By a computational analysis of the number of the pentad sequence in every ORF on the *B. subtilis* genome, 10 of the top 20 genes containing the pentad sequence are involved in secondary metabolite biosynthesis, transport, and catabolism according to the COG database (Park et al. 2011). In the *S. aureus* genome, the genes which contain significantly abundant MazF-sa cleavage sites have been identified to be related to pathogenic factors (Zhu et al. 2009). On the other hand, none of the top 10 UACAU-rich genes in *S. aureus* exist in the

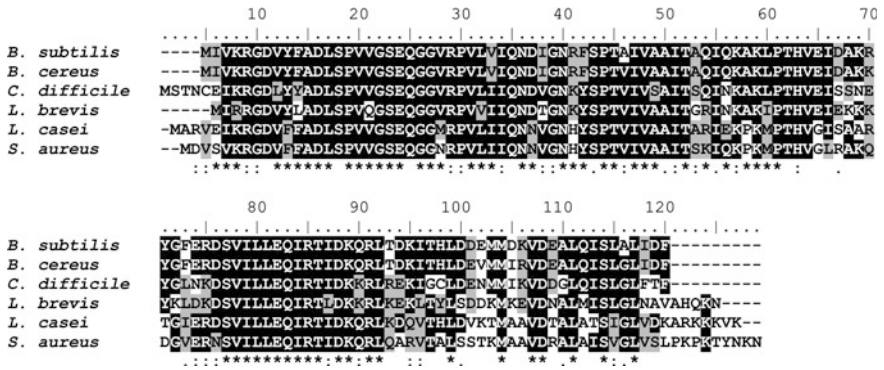


Fig. 7.3 Sequence alignments of MazF homologs from Gram-positive bacteria. Alignments of *B. subtilis* MazF (MazF-bs) with those of *Bacillus cereus* (REF), *Clostridium difficile* (REF), *Lactobacillus brevis* (REF), *Lactobacillus casei* (REF), and *Staphylococcus aureus* are shown. Identical residues are shown in *black*, homologous residues in *gray*

B. subtilis genome. The reason for this observation is not well understood at present, but the difference may have something to do with the pathogenicity of these bacteria.

7.3.6 *MazF-hw from the Archaeon* Haloquadratum walsbyi

To date, the longest recognition sequence for MazF homologs consists of seven bases, which was found for a MazF homolog, MazF-hw, from *H. walsbyi* (Walsby 1980). This MazF homolog cleaves RNAs at a specific seven-base sequence (UU^ACUCA) (Yamaguchi et al. 2012). Such a highly sequence-specific mRNA interferase may be called as a RNA restriction enzyme as they cleave RNA into large specific fragments like DNA restriction enzymes. While DNA restriction enzymes require a palindromic sequence at the cleavage sites, RNA restriction enzymes do not. Therefore, for six-base sequences, there are only 64 different sequences for DNA restriction enzymes, while there are 4,096 possible sequences for RNA restriction enzymes or mRNA interferases.

Interestingly, out of all 2,610 ORFs on the *H. walsbyi* genome, only 183 contain at least one UUACUCA sequence. Of these 183 genes, one gene, *boa,* encoding a transcriptional regulator for rhodopsin, contains three of the seven-base sequence, suggesting that the light-driven proton pump is effectively regulated at the level of mRNA by an mRNA interferase (Yamaguchi et al. 2012). It has been shown that *E. coli* MazF-ec causes complete cell growth arrest, however, the cells retain full metabolic activities (Zhang et al. 2003b). This physiological quasi-dormant state caused by MazF induction is considered to play an important role in cellular survival under various stress conditions in nature. Notably, the MazF-hw activity was strongly inhibited in the presence of 10 mM $MgCl_2$ or 400 mM NaCl, while

H. walsbyi requires 3 M NaCl for growth and is resistant to 2 M $MgCl_2$ (Stoeckenius 1981; Walsby 1980), suggesting that MazF-hw is likely not to be active under normal growth conditions in nature, provided that the salt concentration in the cell is also high. Although the actual intra cellular salt concentration in *H. walsbyi* has not been estimated, a number of enzymes from this organism have been shown to require high salt concentrations for their activities (Lanyi 1974; Schwartz 1979), suggesting that the intra cellular salt concentration is quite high.

H. walsbyi lives by floating on the surface of saturated salt water (Walsby 1980), thus effectively utilizing sun light for the production of ATP. However, it is assumed that upon hypo-osmotic stress in nature such as rain fall or influx of water from a river lowering the specific gravity of saturated salt water, the cells can no more float on the surface, resulting in reduction of ATP production. This eventually causes the decrease of the cellular salt concentration to activate the activation of MazF-hw, which then degrades the mRNA for the transcription activator for the rhodopsin gene. Notably, most of the genes containing two of the seven-base sequence are associated with ATP consumption including ABC-type transporters, cAMP-dependent carbon starvation protein A, sulfatase, and FAD-linked oxidase. These genes may be regulated under the hypo-osmotic stresses, as the mRNAs from these genes are likely more susceptible to MazF-hw than other cellular mRNAs. If this is the case, protein-mediated mRNA interference is used to silence specific gene expression in this archaeon.

7.4 Physiological Function of MazF

Since TA systems are thought to play an important part in cell physiology in *E. coli,* their induction has been examined under several stress conditions (Gerdes et al. 2005). MazF expression is induced under a number of different stresses such as nutrient starvation, the addition of antibiotics, DNA damage, and heat (Aizenman et al. 1996; Hazan et al. 2004; Sat et al. 2001); in addition, stress-induced MazF is suggested to cause cell growth arrest and eventually programmed cell death in *E. coli* MC4100 (Aizenman et al. 1996; Belitsky et al. 2011; Hazan et al. 2004; Kolodkin-Gal and Engelberg-Kulka 2008; Kolodkin-Gal et al. 2007; Kolodkin-Gal et al. 2008; Sat et al. 2001). Although the mechanism is still unclear, it was also reported that MazF induces an exclusive group of proteins, some of which are called survival proteins since the deletion of any one of the genes for these proteins increases stress-induced cell death in *E. coli* MC4100 (Amitai et al. 2009). Other proteins induced by MazF induction are called death proteins, since the rate of cell death caused by antibiotics is substantially reduced by any deletions of the death protein genes (Amitai et al. 2009). Notably, however, the effect of the deletion of the genes for survival proteins on enhancing cell death induced by antibiotics is quite low compared with the wild type *E. coli* MC4100. More recently, it was shown that MazF cleaves some specific mRNA at ACA sequences

at closely upstream of the start codon in *E. coli* (Vesper et al. 2011). These authors also showed that upon MazF induction, 16S rRNA in 30S ribosomes is cleaved. However, it has been shown that the genes without ACA sequences can be efficiently translated after induction of MazF, and thus the ribosomes after induction of MazF seem to be fully functional (Mao et al. 2009; Suzuki et al. 2005). In fact, it has been demonstrated that both 23S and 16S rRNAs are stable after MazF induction (Zhang et al. 2003b). It is possible that the cell death caused by antibiotics (spectinomycin or nalidixic acid) shown by Vesper et al. (2011) may be mediated by other genes in addition to the *mazF* gene, since after the addition of antibiotics, many genes are shown to be up or down regulated (Hong et al. 2009; Kohanski et al. 2008).

In the case of *M. xanthus*, MazF-mx has an essential role in fruiting body formation, during which 80 % of the cell population has to lyse or go through obligatory altruistic cell death, which requires MazF-mx induction (Nariya and Inouye 2008). On the other hand, obligatory cell death and lysis are not observed for *E. coli,* when cells are starved of nutrients, indicating that MazF-ec and MazF-mx play different roles in the cells.

7.4.1 *Roles of the* mazE-mazF *TA System in DNA Damage-dependent Replication Arrest*

Thymine starvation resulting in cell death is known as "thymineless death" (TLD) (Ahmad et al. 1998). Despite its discovery more than 50 years ago, the mechanism(s) for TLD remains elusive. Recently, it has been suggested that the *mazE-mazF* TA system is the sole element responsible for TLD, and that thymine depletion causes MazF induction leading to cell death in *E. coli* MC4100 (Engelberg-Kulka et al. 2004; Sat et al. 2003). However, more recently this view was disputed, demonstrating that *mazE-mazF* as well as *relB-relE* cannot account solely for TLD (Fonville et al. 2010). These authors showed that RecA and approximately 40 other proteins up regulated during SOS response play essential roles in TLD. They further showed that TLD also occurs by blocking of cell division caused by accumulation of single-stranded DNA (the signal of SOS response).

Recently, it was reported that there are two cell-death pathways, one by *recA*- and *lexA*-mediated cell death and the other by *mazEmazF*-mediated cell death in *E. coli* after DNA damage caused by norfloxacin or nalidixic acid (Erental et al. 2012). In this report, the authors showed that the *mazEmazF* system represses the cell death mediated by *recA* and *lexA* by inhibiting the transcription of *recA* after the addition of nalidixic acid. This observation apparently conflicts with the results by Fonville et al. (2010) as described above and also with an early work, in which RecA protein (then described as protein X) is induced upon the addition of nalidixic acid (Inouye 1969). Further research is needed for more detail study on RecA production under stress conditions and its roles in cell physiology.

7.4.2 *Roles of the* mazE-mazF *TA system in DNA Damage-independent Replication Arrest*

Hydroxyurea (HU) is known to cause DNA damage independently of the arrest of DNA replication. HU inhibits class I ribonucleoreductase, leading to deoxyribonucleotide triphosphate depletion in the cells to inhibit DNA replication (Foti et al. 2005). After the addition of HU, iron uptake systems and the expression of toxins (MazF and RelE) are activated, resulting in the production of prematurely terminated protein fragments and also stimulating envelope stress responses in *E. coli* (Godoy et al. 2006). Incorporated ferrous ion induces the formation of hydroxy radicals *via* the Fenton reaction, causing cell death (Imlay et al. 1988; Imlay and Linn 1986).

Deletion of *mazE-mazF* can increase resistance to HU (Davies et al. 2009). Conversely, the Δ*relB-relE* strain is indistinguishable from the parental strain with respect to survival in the presence of HU. Interestingly, however, when both *mazE-mazF* and *relB-relE* are deleted, HU resistance substantially increases to a level far above that observed in a Δ*mazE-mazF* strain alone. Deletion of *mazE-mazF* causes only a small decrease in hydroxyl radical formation, while deletion of *relB-relE* has no detectable effect. However, the Δ*mazE-mazF*Δ*relB-relE* strain exhibits higher reduction in hydroxyl radical levels, consistent with the observed increase in survival. Notably, microarray results indicated that HU treatment seems to cause toxin activation mainly at a posttranscriptional level (Davies et al. 2009). In this analysis, not only *mazE-mazF* and *relB-relE* but also other TA systems such as *mqsRmqsA* and *higBhigA* were induced after HU treatment (Davies et al. 2009), suggesting that these toxins may have roles in HU-induced cell death.

It has been reported that extracellular death factor (EDF) consisting of Asn–Asn-Trp-Asn-Asn is involved in *mazE-mazF*-mediated cell death in *E. coli* MC4100 (Kolodkin-Gal and Engelberg-Kulka 2008; Kolodkin-Gal et al. 2007). In these reports, the addition of bacteriostatic antibiotics (rifampicin and chloramphenicol) was shown to lead to *mazE-mazF*-mediated cell death in MC4100 at a high cell density, but not at a low cell density (Kolodkin-Gal et al. 2007). Recently, it was reported that EDF binds directly to MazF, leading to release MazF from the MazE-MazF complex (Belitsky et al. 2011). At present it is not known how the pentapeptide, Asn–Asn-Trp-Asn-Asn, is generated, since *E. coli* has no proteins containing this sequence, although it was suggested that NNWDN peptide may be quarried from *Zwf* protein, glucose-6-phosphate 1-dehydrogenase, which may be subsequently amidated to generate NNWNN (Kolodkin-Gal et al. 2007).

Cell death mediated by *mazE-mazF* seems to be different from HU-mediated cell death, since HU-mediated cell death is independent of the culture density. Furthermore, HU-mediated cell death occurs in both MC4100 and MG1655 strains with similar growth inhibition kinetics, while MC1655 but not MC4100 is resistant to *mazE-mazF*-mediated cell death (Fonville et al. 2010). The mechanisms by which HU-mediated deoxynucleoside triphosphates depletion activates the MazF and RelE toxins are not yet clearly understood, however, it is apparent that HU-mediated cell death is at least partially dependent on TA systems.

HU-mediated activation of MazF and RelE likely leads to the production of truncated proteins and misfolded proteins, which may result in protein carbonylation (a marker of oxidative stress) (Dukan et al. 2000) and the formation of reactive oxygen species as is the case when cells are treated with aminoglycosides (Kohanski et al. 2007). Consistent with this, MazF overexpression has been shown to cause protein carbonylation (Kolodkin-Gal et al. 2008).

7.4.3 Stringent Response Under Amino Acid Starvation

At least 10 TA systems including *mazE-mazF* in *E. coli* have been shown to be induced by amino acid starvation (Christensen and Gerdes 2003; Christensen-Dalsgaard et al. 2010; Jorgensen et al. 2009). The *mazE-mazF* system is related to the stringent response, since the *mazE-mazF* operon is located downstream of *relA*, the gene for ppGpp synthase (Aizenman et al. 1996) (Fig. 7.1). Indeed, it has been demonstrated that amino acid starvation leads to activation of bacterial toxins RelE and MazF, resulting in mRNA cleavage to inhibit translation (Christensen et al. 2001; Gerdes et al. 2005). However, the mRNA cleavage mediated by ribosomes stalling under various conditions occurs independently of bacterial toxins (Hayes 2003; Sunohara et al. 2004). It was demonstrated that mRNA cleavage occurs through serine starvation in a five-TA deletion mutant (Δ*mazEmazF*, Δ*chpBIK*, Δ*relBrelE*, Δ*yafMyoeB* and Δ*yafQdinJ*) as well as wild type (Li et al. 2008). These results indicate that mRNA cleavage during amino acid starvation may be primarily mediated by stalled ribosomes in a toxin-independent manner. It has been suggested that the induction of RelE and other toxins occurs only in the presence of high concentrations of a serine analog, serine hydroxamine (Li et al. 2008). Therefore, the toxin-mediated mRNA cleavage may require other TA systems and become significant only when cells are exposed to a severe starvation condition.

7.4.4 Network of TA Systems

One intriguing possibility is that there is a network(s) of TA systems, whose regulation is coordinated under various stress conditions. In *M. tuberculosis,* there are around 80 TA systems, of which 27 are VapB-VapC systems (Ramage et al. 2009). Interestingly, a VapB antitoxin can interact with noncognate toxins, MazF toxins in *M. tuberculosis* (Zhu et al. 2010). Such a heterogeneous TA interaction is interesting as two or more different TA systems in *E. coli* may be induced at the same time (Kohanski et al. 2007). It is also possible that the activation of one toxin may result in the activation of other toxins in order to survive under harsh stress conditions in nature or to achieve effective cell death.

7.5 Application of the *mazE-mazF* System as a Tool in Molecular Biology and Biotechnology

Since toxins from bacterial TA systems target to a wide variety of cellular functions including DNA replication, mRNA stability, ribosome assembly, translation, cell division and cell wall synthesis, toxins by themselves or their modified forms may be utilized as a tool for molecular biology, recombinant DNA, the development of antibiotics, and even for therapy and prevention for human diseases. We describe a few examples for the application of toxins for biotechnology and gene therapy.

7.5.1 Single-Protein Production System Using MazF

Although MazF from *E. coli* cleaves single-stranded RNAs at ACA sequences, cellular tRNAs appear to be protected from cleavage because of their extensive secondary structure, and furthermore rRNAs are also resistant to MazF as they are extensively protected by ribosomal proteins (Zhang et al. 2003b). Therefore, MazF expression results in nearly complete degradation of mRNA, leading to severe reduction of protein synthesis in conjunction with growth arrest (Zhang et al. 2003b). However, RNA and DNA synthesis are not inhibited by MazF induction, indicating that metabolic activities for ATP production and nucleotide biosynthesis are retained in cells overproducing MazF. These cells seem to be "undead" even if they are not able to grow. In order to better define this "undead" state, one can test the ability of the cells overproducing MazF for protein synthesis by inducing an mRNA that is devoid of ACA sequences, and examine whether the cells are still able to produce a protein from the ACA-less (thus resistant to MazF) mRNA.

Indeed, it was found that MazF-induced cells carrying an ACA-less mRNA were capable to efficiently produce a protein from the ACA-less mRNA. This converts *E. coli* cells into a bioreactor producing only a single protein of interest, establishing the single-protein production (SPP) system (Suzuki et al. 2005). To apply this system for protein production, an mRNA for a protein of interest has to be first engineered to eliminate all ACA sequences in the mRNA by single-base substitutions without altering the amino acid sequence of the protein encoded. This ACA-less mRNA thus becomes resistant to MazF, while almost all other cellular mRNAs are sensitive to MazF cleavage, enabling the MazF-induced cells to produce the protein only from the engineered ACA-less mRNA, establishing the SPP system (Fig. 7.4) (Suzuki et al. 2005). Using this system, high levels of production of various target proteins have been achieved in the total absence of cell growth (Mao et al. 2009, 2011; McVey et al. 2006). It has been demonstrated that SPP protein production is maintained at a high level at least for 96 h in the virtual absence of background protein synthesis (Suzuki et al. 2005).

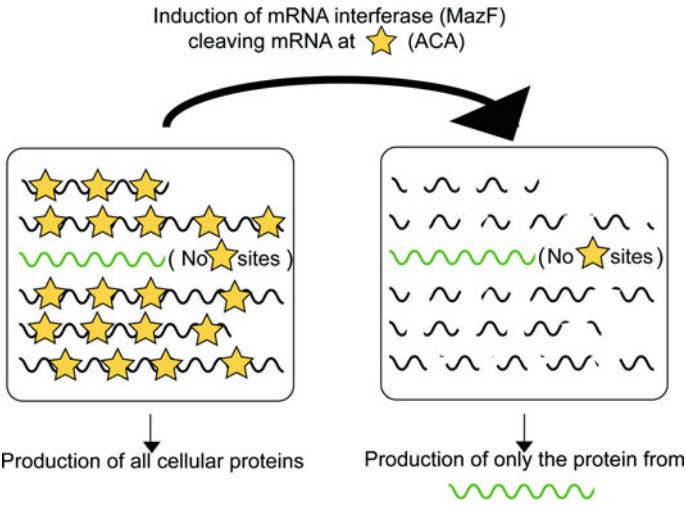


Fig. 7.4 The mechanism of the single-protein production (*SPP*) system in *E. coli.* Since almost all genes (in *black*) contain ACA sequences (*orange stars*), only target protein (encoded on ACA-less gene in *green*) is produced after MazF induction in *E. coli*

In addition to the usefulness of the SPP system for the production bacterial, yeast, and human proteins, the technology was also effective for overexpression of integral inner membrane proteins whose natural levels of expression in the cells are usually very low (Mao et al. 2011). Notably, the SPP system is very useful for nuclear magnetic resonance (NMR) protein structure work, as one can obtain very high signal to noise ratios when new protein synthesis is monitored by isotopic labeling (note that isotopes are not being incorporated into any other cellular proteins except for the target protein). This isotope-labeling technology may enable NMR structural and functional studies of membrane proteins without purification (Mao et al. 2009) as well as a protein of interest in living cells. Other important features of the SPP system for protein production should also be noted: First, the culture for the SPP system can be substantially condensed for protein production without affecting yield of the protein, since cells are not growing. This in turn results in remarkable cost saving (Mao et al. 2009). In particular, when highly expensive isotopes and amino acid analogs are used for protein production, the condensed SPP system (cSPP) is likely to be the method of choice. Second, toxic amino acid analogs such as selenomethionine, fluorophenylalanine can be efficiently incorporated into a protein of interest, since these toxic amino acids are not being incorporated into any other cellular proteins but into the target protein (Suzuki et al. 2006). This together with the cSPP method may provide significant advantage over conventional methods for incorporation of highly toxic amino acid analogs into a protein.

7.5.2 The Use of MazF for Gene Therapy

MazF has been shown to be useful to regulate cell growth and death in eukaryotic cells, as its induction in mammalian cells also causes striking degradation of cellular mRNAs to inhibit protein synthesis leading to NBK/BIK and BAK-dependent apoptotic cell death (Shimazu et al. 2007). This result raises an intriguing possibility that MazF or other mRNA interferases from bacteria may be used as a tool for gene therapy against cancer, AIDS or other human diseases. Indeed, it has been recently demonstrated that $CD4^+$ cells carrying the *E. coli mazF* gene under the Tar promoter from HIV-1 acquires resistance against HIV-1 infection (Chono et al. 2011a). In the HIV-1 life cycle immediately after HIV-1 infection, Tat (transactivator of transcription), an early regulatory protein encoded by the HIV-1 genome, is produced, which subsequently binds to the TAR (transactivation response) sequence to induce the transcription of the HIV-1 genome leading to the expression of other HIV-1 proteins (Berkhout et al. 1989). Therefore, for prevention of HIV-1 infection, it may be a best strategy to preferentially destroy the HIV-1 transcript upon HIV-1 infection.

Recently, a novel approach was reported for possible treatment and prevention of HIV-1 and hepatitis C virus (HCV) infection by using MazE-MazF fusion protein (Shapira et al. 2012) (Fig. 7.5). The proteases encoded by a number of RNA viruses such as human immunodeficiency virus (HIV-1) and HCV play an essential role in viral infection, as they are required for processing of viral encoded polyproteins (Carrere-Kremer et al. 2004; Freed 1998). Thus, HIV-1 and HCV proteases have been considered to be ideal drug targets (Foote et al. 2011, Ghosh et al. 2011; McHutchison and Patel 2002). However, the major problem to use these proteases as a drug target is that viruses readily develop resistance to newly developed drugs, resulting in vicious cycles for drug development (Brik and Wong 2003; Shafer et al. 2007). Therefore, instead of using these proteases as drug targets, one may utilize these RNA viral proteases positively by activating the MazF-ec activity (Shapira et al. 2012; Park et al. 2012). For this purpose, the C-terminal unstructured peptide from MazE, which is known to inhibit the MazF activity was fused with a linker having a specific protease cleavage site(s) for viral proteases. Therefore, only when cells are infected with HIV-1 or HCV, viral proteases activate the latent toxin, the MazE-MazF fusion protein. This will then eliminate all cellular mRNAs leading to Bak-dependent apoptotic cell death (Shimazu et al. 2007). It is remained to be examined if the mRNA interferase thus induced are able to effectively disintegrate viral RNA in the infected cells.

7.5.3 Other Applications of MazF for Biotechnology

Two other applications of MazF have been reported. First, the *mazF* gene can be used as a selective marker for deletion of specific genes in *B. subtilis* (Zhang et al. 2006). Second, MazF and its homologs can be used for mRNA interference

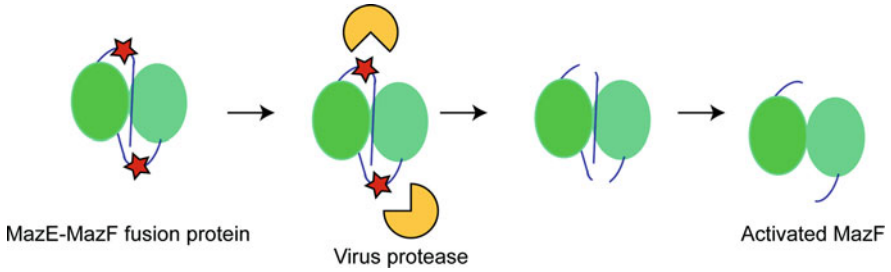


Fig. 7.5 Schematic representation of the activation of the MazE-MazF fusion protein by a virus protease. A MazE-MazF fusion protein is shown in *green* and *blue*, where the MazF molecule is shown by a *green oval shape* and the C-terminal fragment of MazE by a *blue line*. The *red star* indicates the cleavage site for a specific protease. A specific protease indicated by a *yellow shape* cleaves the fusion protein at the star position releasing the MazE fragment from the N-terminus end of MazF to activate the MazF dimer

(protein-mediated mRNA interference) to silence expression of a specific gene or a group of specific genes (Yamaguchi et al. 2012). This protein-base mRNA interference may be as effective as RNA-based mRNA interference mediated by siRNA and antisense RNA. To date, a number of MazF homologs are found in bacteria and archaea, having a wide range of cleavage specificities from three to seven bases. As described earlier, pathogenic bacteria such as *M. tuberculosis* (Zhu et al. 2006, 2008) and *S. aureus* (Zhu et al. 2009) contain mRNA interferases that recognize specific five-base RNA sequences, which are either overpresented or underpresented in genes associated with their pathogenicity. In *M. xanthus*, a five-base sequence-specific mRNA interferase may be involved in programmed cell death during fruiting body formation (Nariya and Inouye 2008). These five-base specific mRNA interferases may be sufficient to silence a group of genes sharing cellular functions such as pathogenecity and cellar development, but not specific enough to target only a single gene or only a few limited target genes. To achieve protein-mediated mRNA interference comparable to RNA-mediated mRNA interference, much higher sequence-specific mRNA interferases are required. As MazF-hw from an archaeon has been identified as a seven-base cutter, it raises an intriguing possibility that there may be many other MazF homologs for other seven base or longer RNA sequences, so that only a specific mRNA or a specific group of cellular mRNAs can be silenced using these MazF homologs.

In *E. coli*, out of 233 ORFs containing the seven-base sequence for MazF-hw, only four genes (*lolD, rplC, rpmD* and *rpoB*) are essential for cell growth (Yamaguchi et al. 2012). Upon MazF-hw induction, mRNAs from all four of these genes were degraded so that cell growth is arrested. However, when the MazF-hw target sequences, UUACUCA in the four essential genes were replaced to other MazF-hw resistant sequences without the change of the amino acid sequences, the cell growth was no more sensitive to MazF-hw induction. However, cell growth again became sensitive to MazF-hw, when any one of the four genes regains a UUACUCA sequence. These data demonstrate that specific genes can be silenced

by a sequence-specific mRNA interferase (Yamaguchi et al. 2012). Although the seven-base cutter, MazF-hw, is not specific enough to regulate expression of specific human genes, it could be possible to engineer MazF-hw to be more specific for longer sequences targeting only a few human genes. Notably, in addition to antisense RNA and siRNA technology, protein-mediated RNA interference with the use of mRNA interferases will open up a new avenue to interfere with the expression of a specific mRNA or a specific group of mRNAs, preventing viral infection and harmful gene expression from bacteria to human.

7.6 Concluding Remarks

Among the TA systems, a number of toxins have been identified as sequence-specific endoribonucleases or mRNA interferases (Yamaguchi and Inouye 2011, Yamaguchi et al. 2011). MazF is an extensively characterized ribosome-independent mRNA interferase. MazF in *E. coli* may cleave almost all mRNAs when induced in *E. coli*, as MazF-ec cleaves RNA at ACA sequences. However, some mRNAs in *E. coli* are considered to be resistant to MazF-ec, as they either do not contain any ACA sequences or the ACA sequences are protected from cleavage by the secondary structures in the mRNA (Amitai et al. 2009). It should be noted that there are a number of strategies for cells to take under stress conditions to regulate biosynthetic pathways. Among them, mRNAs seem to be the least deleterious target for the cells, comparing to DNA synthesis, protein synthesis, and other biosynthetic pathways, since mRNAs can be easily replenished as compared to DNA, proteins and other cellular components. Therefore, the regulation of mRNA stability by mRNA interferases under stress conditions seems to be the most sensible strategy for cell growth control.

Another interesting aspect is the recent finding of a highly sequence-specific mRNA interferase from an archaeon cleaving a specific 7-base sequence, possibly existing once in every 16,384 bases (Yamaguchi et al. 2012), suggesting that the use of highly sequence-specific mRNA interferases may be used for regulation of specific genes in nature. This novel regulatory mechanism in mRNA interference may be as important in nature as mRNA interference mediated by RNA such as antisense RNA and siRNA. The protein-based mRNA interference by mRNA interferases also opens new avenues in biotechnology. The SPP system (Mao et al. 2009; Suzuki et al. 2005) has been shown to be highly useful in NMR protein structure biology and may be used for other areas of biotechnology such as high-throughput drug screening. Since MazF-ec showed toxicity in mammalian cells (Shimazu et al. 2007), the *mazF* gene may be used for gene therapy against cancer, AIDS and other human diseases (Chono et al. 2011a, b). Furthermore, the development of novel antibiotics inhibiting the TA complex formation in the cells is very attractive, as such antibiotics would have two synergistic effects: one to release free toxin in the cells and the other to derepress the TA operon expression resulting in more toxin production, since the TA complex functions as a repressor for its own operon.

It is not clear at present why the cleavage sites for MazF homologs consist of only odd base numbers (3, 5, or 7 bases). There might be an intrinsic restriction in the dimer structure of MazF, which only allows expansion of the cleavage site specificity by two bases at a time. It remains to be determined if this odd number rule for RNA recognition is prevailing in all other MazF homologs, whose RNA cleave site specificity has yet to be determined.

Protein-based mRNA interference is an exciting new area of research in basic science as well as medicine and biotechnology. It should also be noted that there may be many unknown mRNA interferases in nature, which are not homologous to MazF. We will for sure see many more exciting and innovating discoveries in near future.

Acknowledgments We would like to dedicate this chapter for late Dr. Tsutomu Shimazu, who demonstrated for the first time the MazF toxicity in mammalian cells (Shimazu et al. 2007). We also thank to Dr. Vikas Nanda, Mr. Jung-Ho Park, and Ms. Sehrish Ajimal for their critical reading of this chapter. This work was supported by an NIH grant, 1RO1GM081567 and 3RO1GM081567-02S1.

References

Ahmad, S. I., Kirk, S. H., & Eisenstark, A. (1998). Thymine metabolism and thymineless death in prokaryotes and eukaryotes. *Annual Review of Microbiology, 52*, 591–625.

Ahnn, J., March, P. E., Takiff, H. E., & Inouye, M. (1986). A GTP-binding protein of *Escherichia coli* has homology to yeast RAS proteins. *Proceedings of the National academy of Sciences of the United States of America, 83*, 8849–8853.

Aizenman, E., Engelberg-Kulka, H., & Glaser, G. (1996). An *Escherichia coli* chromosomal "addiction module" regulated by guanosine 3′,5′-bispyrophosphate: a model for programmed bacterial cell death. *Proceedings of the National academy of Sciences of the United States of America, 93*, 6059–6063.

Amitai, S., Kolodkin-Gal, I., Hananya-Meltabashi, M., Sacher, A., & Engelberg-Kulka, H. (2009). *Escherichia coli* MazF leads to the simultaneous selective synthesis of both "death proteins" and "survival proteins". *PLoS Genetics, 5*, e1000390.

Basu, S., Pathak, S. K., Banerjee, A., Pathak, S., Bhattacharyya, A., Yang, Z., et al. (2007). Execution of macrophage apoptosis by PE_PGRS33 of *Mycobacterium tuberculosis* is mediated by Toll-like receptor 2-dependent release of tumor necrosis factor-alpha. *Journal of Biological Chemistry, 282*, 1039–1050.

Belitsky, M., Avshalom, H., Erental, A., Yelin, I., Kumar, S., London, N., et al. (2011). The *Escherichia coli* extracellular death factor (EDF) induces the endoribonucleolytic activities of the toxins MazF and ChpBK. *Molecular Cell, 41*, 625–635.

Berkhout, B., Silverman, R. H., & Jeang, K. T. (1989). Tat trans-activates the human immunodeficiency virus through a nascent RNA target. *Cell, 59*, 273–282.

Bravo, A., de Torrontegui, G., & Diaz, R. (1987). Identification of components of a new stability system of plasmid R1, ParD, that is close to the origin of replication of this plasmid. *Molecular and General Genetics, 210*, 101–110.

Brennan, M. J., & Delogu, G. (2002). The PE multigene family: a 'molecular mantra' for mycobacteria. *Trends in Microbiology, 10*, 246–249.

Briani, F., Deho, G., Forti, F., & Ghisotti, D. (2001). The plasmid status of satellite bacteriophage P4. *Plasmid, 45*, 1–17.

Brik, A., & Wong, C. H. (2003). HIV-1 protease: mechanism and drug discovery. *Organic & Biomolecular Chemistry, 1*, 5–14.

Carrere-Kremer, S., Montpellier, C., Lorenzo, L., Brulin, B., Cocquerel, L., Belouzard, S., et al. (2004). Regulation of hepatitis C virus polyprotein processing by signal peptidase involves structural determinants at the p7 sequence junctions. *The Journal of biological chemistry, 279*, 41384–41392.

Cesareni, G., Helmer-Citterich, M., & Castagnoli, L. (1991). Control of ColE1 plasmid replication by antisense RNA. *Trends in Genetics, 7*, 230–235.

Chono, H., Matsumoto, K., Tsuda, H., Saito, N., Lee, K., Kim, S., et al. (2011a). Acquisition of HIV-1 resistance in T lymphocytes using an ACA-specific *E. coli* mRNA interferase. *Human Gene Therapy, 22*, 35–43.

Chono, H., Saito, N., Tsuda, H., Shibata, H., Ageyama, N., Terao, K., et al. (2011b). In vivo safety and persistence of endoribonuclease gene-transduced CD4+ T cells in cynomolgus macaques for HIV-1 gene therapy model. *PLoS ONE, 6*, e23585.

Christensen, S. K., & Gerdes, K. (2003). RelE toxins from bacteria and archaea cleave mRNAs on translating ribosomes, which are rescued by tmRNA. *Molecular Microbiology, 48*, 1389–1400.

Christensen, S. K., Mikkelsen, M., Pedersen, K., & Gerdes, K. (2001). RelE, a global inhibitor of translation, is activated during nutritional stress. *Proceedings of the National academy of Sciences of the United States of America, 98*, 14328–14333.

Christensen, S. K., Pedersen, K., Hansen, F. G., & Gerdes, K. (2003). Toxin-antitoxin loci as stress-response-elements: ChpAK/MazF and ChpBK cleave translated RNAs and are counteracted by tmRNA. *Journal of Molecular Biology, 332*, 809–819.

Christensen-Dalsgaard, M., Jorgensen, M. G., & Gerdes, K. (2010). Three new RelE-homologous mRNA interferases of *Escherichia coli* differentially induced by environmental stresses. *Molecular Microbiology, 75*, 333–348.

Davies, B. W., Kohanski, M. A., Simmons, L. A., Winkler, J. A., Collins, J. J., & Walker, G. C. (2009). Hydroxyurea induces hydroxyl radical-mediated cell death in *Escherichia coli*. *Molecular Cell, 36*, 845–860.

Davison, J. (1984). Mechanism of control of DNA replication and incompatibility in ColE1-type plasmids–a review. *Gene, 28*, 1–15.

de la Cueva-Mendez, G., Mills, A. D., Clay-Farrace, L., Diaz-Orejas, R., & Laskey, R. A. (2003). Regulatable killing of eukaryotic cells by the prokaryotic proteins Kid and Kis. *EMBO Journal, 22*, 246–251.

Denny, P. W., & Smith, D. F. (2004). Rafts and sphingolipid biosynthesis in the kinetoplastid parasitic protozoa. *Molecular Microbiology, 53*, 725–733.

Dheenadhayalan, V., Delogu, G., Sanguinetti, M., Fadda, G., & Brennan, M. J. (2006). Variable expression patterns of Mycobacterium tuberculosis PE_PGRS genes: evidence that PE_PGRS16 and PE_PGRS26 are inversely regulated in vivo. *Journal of Bacteriology, 188*, 3721–3725.

Dukan, S., Farewell, A., Ballesteros, M., Taddei, F., Radman, M., & Nystrom, T. (2000). Protein oxidation in response to increased transcriptional or translational errors. *Proceedings of the National academy of Sciences of the United States of America, 97*, 5746–5749.

Engelberg-Kulka, H., & Glaser, G. (1999). Addiction modules and programmed cell death and antideath in bacterial cultures. *Annual Review of Microbiology, 53*, 43–70.

Engelberg-Kulka, H., Sat, B., Reches, M., Amitai, S., & Hazan, R. (2004). Bacterial programmed cell death systems as targets for antibiotics. *Trends in Microbiology, 12*, 66–71.

Erental, A., Sharon, I., & Engelberg-Kulka, H. (2012). Two programmed cell death systems in *Escherichia coli*: An apoptotic-like death is inhibited by the mazEF-mediated death pathway. *PLoS Biology, 10*, e1001281.

Fonville, N. C., Bates, D., Hastings, P. J., Hanawalt, P. C., & Rosenberg, S. M. (2010). Role of RecA and the SOS response in thymineless death in *Escherichia coli*. *PLoS Genetics, 6*, e1000865.

Foote, B. S., Spooner, L. M., & Belliveau, P. P. (2011). Boceprevir: a protease inhibitor for the treatment of chronic hepatitis C. *Annals of Pharmacotherapy, 45*, 1085–1093.

Foti, J. J., Schienda, J., Sutera, V. A., Jr, & Lovett, S. T. (2005). A bacterial G protein-mediated response to replication arrest. *Molecular Cell, 17*, 549–560.

Freed, E. O. (1998). HIV-1 gag proteins: diverse functions in the virus life cycle. *Virology, 251*, 1–15.

Fu, Z., Donegan, N. P., Memmi, G., & Cheung, A. L. (2007). Characterization of MazFSa, an endoribonuclease from *Staphylococcus aureus*. *Journal of Bacteriology, 189*, 8871–8879.

Fu, Z., Tamber, S., Memmi, G., Donegan, N. P., & Cheung, A. L. (2009). Overexpression of MazFsa in *Staphylococcus aureus* induces bacteriostasis by selectively targeting mRNAs for cleavage. *Journal of Bacteriology, 191*, 2051–2059.

Gerdes, K., Christensen, S. K., & Lobner-Olesen, A. (2005). Prokaryotic toxin-antitoxin stress response loci. *Nature Reviews Microbiology, 3*, 371–382.

Gey van Pittius, N. C., Sampson, S. L., Lee, H., Kim, Y., van Helden, P. D., & Warren, R. M. (2006). Evolution and expansion of the *Mycobacterium tuberculosis* PE and PPE multigene families and their association with the duplication of the ESAT-6 (esx) gene cluster regions. *BMC Evolutionary Biology, 6*, 95.

Ghosh, A. K., Chapsal, B. D., Parham, G. L., Steffey, M., Agniswamy, J., Wang, Y. F., et al. (2011). Design of HIV-1 protease inhibitors with C3-substituted hexahydrocyclopentafuranyl Urethanes as P2-ligands: synthesis, biological evaluation, and protein-ligand X-ray crystal structure. *Journal of Medicinal Chemistry, 54*, 5890–5901.

Godoy, V. G., Jarosz, D. F., Walker, F. L., Simmons, L. A., & Walker, G. C. (2006). Y-family DNA polymerases respond to DNA damage-independent inhibition of replication fork progression. *EMBO Journal, 25*, 868–879.

Hayes, F. (2003). Toxins-antitoxins: plasmid maintenance, programmed cell death, and cell cycle arrest. *Science, 301*, 1496–1499.

Hazan, R., Sat, B., & Engelberg-Kulka, H. (2004). *Escherichia coli mazEF*-mediated cell death is triggered by various stressful conditions. *Journal of Bacteriology, 186*, 3663–3669.

Hong, J., Ahn, J. M., Kim, B. C., & Gu, M. B. (2009). Construction of a functional network for common DNA damage responses in *Escherichia coli*. *Genomics, 93*, 514–524.

Imlay, J. A., Chin, S. M., & Linn, S. (1988). Toxic DNA damage by hydrogen peroxide through the Fenton reaction in vivo and in vitro. *Science, 240*, 640–642.

Imlay, J. A., & Linn, S. (1986). Bimodal pattern of killing of DNA-repair-defective or anoxically grown *Escherichia coli* by hydrogen peroxide. *Journal of Bacteriology, 166*, 519–527.

Inouye, M. (1969). Unlinking of cell division from deoxyribonucleic acid replication in a temperature-sensitive deoxyribonucleic acid synthesis mutant of *Escherichia coli*. *Journal of Bacteriology, 99*, 842–850.

Inouye, M. (2006). The discovery of mRNA interferases: implication in bacterial physiology and application to biotechnology. *Journal of Cellular Physiology, 209*, 670–676.

Jorgensen, M. G., Pandey, D. P., Jaskolska, M., & Gerdes, K. (2009). HicA of Escherichia coli defines a novel family of translation-independent mRNA interferases in bacteria and archaea. *Journal of Bacteriology, 191*, 1191–1199.

Jung, Y. H., & Lee, Y. (1995). RNases in ColE1 DNA metabolism. *Molecular Biology Reports, 22*, 195–200.

Kamada, K., Hanaoka, F., & Burley, S. K. (2003). Crystal structure of the MazE/MazF complex: molecular bases of antidote-toxin recognition. *Molecular Cell, 11*, 875–884.

Kohanski, M. A., Dwyer, D. J., Hayete, B., Lawrence, C. A., & Collins, J. J. (2007). A common mechanism of cellular death induced by bactericidal antibiotics. *Cell, 130*, 797–810.

Kohanski, M. A., Dwyer, D. J., Wierzbowski, J., Cottarel, G., & Collins, J. J. (2008). Mistranslation of membrane proteins and two-component system activation trigger antibiotic-mediated cell death. *Cell, 135*, 679–690.

Kolodkin-Gal, I., & Engelberg-Kulka, H. (2008). The extracellular death factor: physiological and genetic factors influencing its production and response in *Escherichia coli*. *Journal of Bacteriology, 190*, 3169–3175.

Kolodkin-Gal, I., Hazan, R., Gaathon, A., Carmeli, S., & Engelberg-Kulka, H. (2007). A linear pentapeptide is a quorum-sensing factor required for *mazEF*-mediated cell death in *Escherichia coli*. *Science, 318*, 652–655.

Kolodkin-Gal, I., Sat, B., Keshet, A., & Engelberg-Kulka, H. (2008). The communication factor EDF and the toxin-antitoxin module *mazEF* determine the mode of action of antibiotics. *PLoS Biology, 6*, e319.

Lanyi, J. K. (1974). Salt-dependent properties of proteins from extremely halophilic bacteria. *Bacteriol Reviews, 38*, 272–290.

Li, G. Y., Zhang, Y., Chan, M. C., Mal, T. K., Hoeflich, K. P., Inouye, M., et al. (2006). Characterization of dual substrate binding sites in the homodimeric structure of *Escherichia coli* mRNA interferase MazF. *Journal of Molecular Biology, 357*, 139–150.

Li, X., Yagi, M., Morita, T., & Aiba, H. (2008). Cleavage of mRNAs and role of tmRNA system under amino acid starvation in *Escherichia coli*. *Molecular Microbiology, 68*, 462–473.

Mao, L., Inoue, K., Tao, Y., Montelione, G. T., McDermott, A. E., & Inouye, M. (2011). Suppression of phospholipid biosynthesis by cerulenin in the condensed single-protein-production (cSPP) system. *Journal of Biomolecular NMR, 49*, 131–137.

Mao, L., Tang, Y., Vaiphei, S. T., Shimazu, T., Kim, S. G., Mani, R., et al. (2009). Production of membrane proteins for NMR studies using the condensed single protein (cSPP) production system. *Journal of Structural and Functional Genomics, 10*, 281–289.

March, P. E., Lerner, C. G., Ahnn, J., Cui, X., & Inouye, M. (1988). The *Escherichia coli* Ras-like protein (Era) has GTPase activity and is essential for cell growth. *Oncogene, 2*, 539–544.

Masuda, Y., Miyakawa, K., Nishimura, Y., & Ohtsubo, E. (1993). *chpA* and *chpB*, *Escherichia coli* chromosomal homologs of the pem locus responsible for stable maintenance of plasmid R100. *Journal of Bacteriology, 175*, 6850–6856.

McHutchison, J. G., & Patel, K. (2002). Future therapy of hepatitis C. *Hepatology, 36*, S245–S252.

McVey, C. E., Amblar, M., Barbas, A., Cairrao, F., Coelho, R., Romao, C., et al. (2006). Expression, purification, crystallization and preliminary diffraction data characterization of *Escherichia coli* ribonuclease II (RNase II). *Acta Crystallographica, Section F: Structural Biology and Crystallization Communications, 62*, 684–687.

Metzger, S., Dror, I. B., Aizenman, E., Schreiber, G., Toone, M., Friesen, J. D., et al. (1988). The nucleotide sequence and characterization of the relA gene of *Escherichia coli*. *Journal of Biological Chemistry, 263*, 15699–15704.

Mishra, K. C., de Chastellier, C., Narayana, Y., Bifani, P., Brown, A. K., Besra, G. S., et al. (2008). Functional role of the PE domain and immunogenicity of the *Mycobacterium tuberculosis* triacylglycerol hydrolase LipY. *Infection and Immunity, 76*, 127–140.

Munoz-Gomez, A. J., Santos-Sierra, S., Berzal-Herranz, A., Lemonnier, M., & Diaz-Orejas, R. (2004). Insights into the specificity of RNA cleavage by the *Escherichia coli* MazF toxin. *FEBS Letters, 567*, 316–320.

Nariya, H., & Inouye, M. (2008). MazF, an mRNA interferase, mediates programmed cell death during multicellular *Myxococcus* development. *Cell, 132*, 55–66.

Park, J. H., Yamaguchi, Y., & Inouye, M. (2011). Bacillus subtilis MazF-bs (EndoA) is a UACAU-specific mRNA interferase. FEBS Lett.

Park, J. H., Yamaguchi, Y., & Inouye, M. (2012). Intramolecular regulation of the sequence-specific mRNA interferase activity of MazF Fused to a MazE Fragment with a Linker Cleavable by Specific Proteases. *Applied and Environmental Microbiology, 78*, 3794–3799.

Pedersen, K., Zavialov, A. V., Pavlov, M. Y., Elf, J., Gerdes, K., & Ehrenberg, M. (2003). The bacterial toxin RelE displays codon-specific cleavage of mRNAs in the ribosomal A site. *Cell, 112*, 131–140.

Pellegrini, O., Mathy, N., Gogos, A., Shapiro, L., & Condon, C. (2005). The *Bacillus subtilis* ydcDE operon encodes an endoribonuclease of the MazF/PemK family and its inhibitor. *Molecular Microbiology, 56*, 1139–1148.

Ramage, H. R., Connolly, L. E., & Cox, J. S. (2009). Comprehensive functional analysis of *Mycobacterium tuberculosis* toxin-antitoxin systems: implications for pathogenesis, stress responses, and evolution. *PLoS Genetics, 5*, e1000767.

Ruiz-Echevarria, M. J., Berzal-Herranz, A., Gerdes, K., & Diaz-Orejas, R. (1991). The kis and kid genes of the parD maintenance system of plasmid R1 form an operon that is autoregulated at the level of transcription by the co-ordinated action of the Kis and Kid proteins. *Molecular Microbiology, 5*, 2685–2693.

Ruiz-Echevarria, M. J., Gimenez-Gallego, G., Sabariegos-Jareno, R., & Diaz-Orejas, R. (1995). Kid, a small protein of the parD stability system of plasmid R1, is an inhibitor of DNA replication acting at the initiation of DNA synthesis. *Journal of Molecular Biology, 247*, 568–577.

Sat, B., Hazan, R., Fisher, T., Khaner, H., Glaser, G., & Engelberg-Kulka, H. (2001). Programmed cell death in *Escherichia coli*: some antibiotics can trigger *mazEF* lethality. *Journal of Bacteriology, 183*, 2041–2045.

Sat, B., Reches, M., & Engelberg-Kulka, H. (2003). The *Escherichia coli mazEF* suicide module mediates thymineless death. *Journal of Bacteriology, 185*, 1803–1807.

Schwartz, W. (1979). Microbial life in extreme environments. XII und 465 S., 15 Abb., 36 Tab., 1 Taf. (farbig). In D. J. KUSHNER (Ed.), *Zeitschrift für allgemeine Mikrobiologie* 19, (pp. 447–447). London: Academic Press. $ 19.60.

Shafer, R. W., Rhee, S. Y., Pillay, D., Miller, V., Sandstrom, P., Schapiro, J. M., et al. (2007). HIV-1 protease and reverse transcriptase mutations for drug resistance surveillance. *Aids, 21*, 215–223.

Shao, Y., Harrison, E. M., Bi, D., Tai, C., He, X., Ou, H. Y., et al. (2011). TADB: A web-based resource for type 2 toxin-antitoxin loci in bacteria and archaea. *Nucleic Acids Research, 39*, D606–D611.

Shapira, A., Shapira, S., Gal-Tanamy, M., Zemel, R., Tur-Kaspa, R., & Benhar, I. (2012). Removal of hepatitis C virus-infected cells by a zymogenized bacterial toxin. *PLoS ONE, 7*, e32320.

Shimazu, T., Degenhardt, K., Nur, E. K. A., Zhang, J., Yoshida, T., Zhang, Y., et al. (2007). NBK/BIK antagonizes MCL-1 and BCL-XL and activates BAK-mediated apoptosis in response to protein synthesis inhibition. *Genes & Development, 21*, 929–941.

Singh, K. K., Dong, Y., Patibandla, S. A., McMurray, D. N., Arora, V. K., & Laal, S. (2005). Immunogenicity of the *Mycobacterium tuberculosis* PPE55 (Rv3347c) protein during incipient and clinical tuberculosis. *Infection and Immunity, 73*, 5004–5014.

Stoeckenius, W. (1981). Walsby's square bacterium: fine structure of an orthogonal procaryote. *Journal of Bacteriology, 148*, 352–360.

Sun, H., & Shi, W. (2001). Genetic studies of mrp, a locus essential for cellular aggregation and sporulation of *Myxococcus xanthus*. *Journal of Bacteriology, 183*, 4786–4795.

Sunohara, T., Jojima, K., Tagami, H., Inada, T., & Aiba, H. (2004a). Ribosome stalling during translation elongation induces cleavage of mRNA being translated in *Escherichia coli*. *Journal of Biological Chemistry, 279*, 15368–15375.

Sunohara, T., Jojima, K., Yamamoto, Y., Inada, T., & Aiba, H. (2004b). Nascent-peptide-mediated ribosome stalling at a stop codon induces mRNA cleavage resulting in nonstop mRNA that is recognized by tmRNA. *RNA, 10*, 378–386.

Suzuki, M., Roy, R., Zheng, H., Woychik, N., & Inouye, M. (2006). Bacterial bioreactors for high yield production of recombinant protein. *Journal of Biological Chemistry, 281*, 37559–37565.

Suzuki, M., Zhang, J., Liu, M., Woychik, N. A., & Inouye, M. (2005). Single protein production in living cells facilitated by an mRNA interferase. *Molecular Cell, 18*, 253–261.

Tomizawa, J. (1984). Control of ColE1 plasmid replication: the process of binding of RNA I to the primer transcript. *Cell, 38*, 861–870.

Tsuchimoto, S., Nishimura, Y., & Ohtsubo, E. (1992). The stable maintenance system pem of plasmid R100: degradation of PemI protein may allow PemK protein to inhibit cell growth. *Journal of Bacteriology, 174*, 4205–4211.

Tsuchimoto, S., Ohtsubo, H., & Ohtsubo, E. (1988). Two genes, *pemK* and *pemI*, responsible for stable maintenance of resistance plasmid R100. *Journal of Bacteriology, 170*, 1461–1466.

Ueki, T., & Inouye, S. (2003). Identification of an activator protein required for the induction of fruA, a gene essential for fruiting body development in *Myxococcus xanthus*. *Proceedings of the National academy of Sciences of the United States of America, 100*, 8782–8787.

Van Melderen, L., & Saavedra De Bast, M. (2009). Bacterial toxin-antitoxin systems: more than selfish entities? *PLoS Genetics, 5*, e1000437.

Vesper, O., Amitai, S., Belitsky, M., Byrgazov, K., Kaberdina, A. C., Engelberg-Kulka, H., et al. (2011). Selective translation of leaderless mRNAs by specialized ribosomes generated by MazF in *Escherichia coli*. *Cell, 147*, 147–157.

Walsby, A. E. (1980). A square bacterium. *Nature, 283*, 69–71.

Xia, B., Ke, H., & Inouye, M. (2001). Acquirement of cold sensitivity by quadruple deletion of the cspA family and its suppression by PNPase S1 domain in *Escherichia coli*. *Molecular Microbiology, 40*, 179–188.

Yamaguchi, Y., & Inouye, M. (2009). mRNA interferases, sequence-specific endoribonucleases from the toxin-antitoxin systems. *Progress of Molecular Biology Translation Science, 85*, 467–500.

Yamaguchi, Y., & Inouye, M. (2011). Regulation of growth and death in *Escherichia coli* by toxin-antitoxin systems. *Nature Reviews Microbiology, 9*, 779–790.

Yamaguchi, Y., Nariya, H., Park, J. H., & Inouye, M. (2012). Inhibition of specific gene expressions by protein-mediated mRNA interference. *National Communications, 3*, 607.

Yamaguchi, Y., Park, J. H., & Inouye, M. (2011). Toxin-antitoxin systems in bacteria and archaea. *Annual Review of Genetics, 45*, 61–79.

Zhang, J., & Inouye, M. (2002). MazG, a nucleoside triphosphate pyrophosphohydrolase, interacts with Era, an essential GTPase in *Escherichia coli*. *Journal of Bacteriology, 184*, 5323–5329.

Zhang, J., Zhang, Y., & Inouye, M. (2003a). Characterization of the interactions within the *mazEF* addiction module of *Escherichia coli*. *Journal of Biological Chemistry, 278*, 32300–32306.

Zhang, J., Zhang, Y., Zhu, L., Suzuki, M., & Inouye, M. (2004). Interference of mRNA function by sequence-specific endoribonuclease PemK. *Journal of Biological Chemistry, 279*, 20678–20684.

Zhang, Y., Zhang, J., Hara, H., Kato, I., & Inouye, M. (2005a). Insights into the mRNA cleavage mechanism by MazF, an mRNA interferase. *Journal of Biological Chemistry, 280*, 3143–3150.

Zhang, Y., Zhang, J., Hoeflich, K. P., Ikura, M., Qing, G., & Inouye, M. (2003b). MazF cleaves cellular mRNAs specifically at ACA to block protein synthesis in *Escherichia coli*. *Molecular Cell, 12*, 913–923.

Zhang, Y., Zhu, L., Zhang, J., & Inouye, M. (2005b). Characterization of ChpBK, an mRNA interferase from *Escherichia coli*. *Journal of Biological Chemistry, 280*, 26080–26088.

Zhu, L., Inoue, K., Yoshizumi, S., Kobayashi, H., Zhang, Y., Ouyang, M., et al. (2009). *Staphylococcus aureus* MazF specifically cleaves a pentad sequence, UACAU, which is unusually abundant in the mRNA for pathogenic adhesive factor SraP. *Journal of Bacteriology, 191*, 3248–3255.

Zhu, L., Phadtare, S., Nariya, H., Ouyang, M., Husson, R. N., & Inouye, M. (2008). The mRNA interferases, MazF-mt3 and MazF-mt7 from *Mycobacterium tuberculosis* target unique pentad sequences in single-stranded RNA. *Molecular Microbiology, 69*, 559–569.

Zhu, L., Zhang, Y., Teh, J. S., Zhang, J., Connell, N., Rubin, H., et al. (2006). Characterization of mRNA interferases from *Mycobacterium tuberculosis*. *Journal of Biological Chemistry, 281*, 18638–18643.

Zhu, L., Sharp, J. D., Kobayashi, H., Woychik, N. A., & Inouye, M. (2010). Noncognate Mycobacterium tuberculosis toxin-antitoxins can physically and functionally interact. *The Journal of Biological Chemistry, 285*, 39732–39738.

Zhang, X. Z., Yan, X., Cui, Z. L., Hong, Q., & Li, S. P. (2006). mazF, a novel counter-selectable marker for unmarked chromosomal manipulation in Bacillus subtilis. *Nucleic Acids Research, 34*, e71.

Chapter 8
Type II Toxin-Antitoxins: Structural and Functional Aspects of Type II Loci in Mycobacteria

Vickery L. Arcus and Gregory M. Cook

Abstract Several prokaryotes have genomes that are densely populated with toxin–antitoxin (TA) loci. For example, the aquatic bloom-forming cyanobacterium *Microcystis aeruginosa* habours more than 110 TA loci, and the nematode symbiont *Photorhabdus luminescens* has at least 58 TAs encoded in its genome. Amongst this group of unrelated bacteria lies *Mycobacterium tuberculosis* whose genome encodes more than 68 TAs. *M. tuberculosis* causes the devastating human disease tuberculosis (TB) and is, therefore, the subject of intense research. The presence of so many TAs begs the question as to their evolutionary origin, contemporary biological functions and their potential as therapeutic targets in this important human pathogen. This chapter reviews the distribution of TAs in the mycobacterial genus and the experimental results that provide clues to TA function in *M. tuberculosis*.

8.1 Introduction

The mycobacteria provide a unique view of the acquisition, evolution and contemporary biological roles for toxin–antitoxin (TA) loci in prokaryotes. This genus is the subject of great interest as it covers a number of important human and mammalian pathogens along with many environmental organisms. For example, a

V. L. Arcus (✉)
Department of Biological Sciences, University of Waikato, Private Bag 3105, Hamilton 3240, New Zealand
e-mail: varcus@waikato.ac.nz

G. M. Cook
Department of Microbiology and Immunology, Otago School of Medical Sciences, University of Otago, 56 Dunedin 9054, New Zealand
e-mail: gregory.cook@otago.ac.nz

K. Gerdes (ed.), *Prokaryotic Toxin-Antitoxins*,
DOI: 10.1007/978-3-642-33253-1_8,

number of the environmental mycobacteria are studied for their potential application in bioremediation due to their ability to metabolise a wide range of xenobiotic compounds.

Mycobacterium tuberculosis, the causative agent of tuberculosis (TB), claims the lives of approximately 1.7 million people per annum. The evolutionary success of *M. tuberculosis* is borne out in the fact that an estimated one-third of the world's population carries the organism and around 9–10 million new cases of TB are reported each year. One of the main problems in the treatment of TB infection is the capacity of *M. tuberculosis* to enter a persistent state that is less susceptible to antibiotics, dictating long treatment regimes. The global TB burden has recently been compounded by the increase in multidrug resistant (MDR) and extensively drug resistant (XDR) strains of *M. tuberculosis* and the increased susceptibility of HIV-infected individuals (Caminero et al. 2010).

M. tuberculosis has a genome that is densely populated with TA loci (Arcus et al. 2005; Gerdes et al. 2005) (Fig. 8.1 and Table 8.1). There is some debate about the precise numbers of TA loci that reside in the genome—the laboratory pathogenic strain *M. tuberculosis* H37Rv has at least 65 TA loci spread around the genome (Shao et al. 2011). Some authors have predicted a total of up to 88 TA loci for *M. tuberculosis* H37Rv (Ramage et al. 2009). By and large, these loci are conserved across the "*M. tuberculosis* complex"—the collective noun for a group of mycobacteria that includes *Mycobacterium bovis* (the cause of TB in cattle and humans), *Mycobacterium africanum* (a human pathogen) and *Mycobacterium microti* (a pathogen of voles). The TA loci of the *M. tuberculosis* complex are dominated by the *vapBC* family of TA operons ($\sim$45 *vapBCs* in each genome), accompanied by smaller numbers of TAs from the *mazEF* and *relBE* families. The available experimental evidence for mycobacteria indicates that the VapC, MazF and RelE toxins all target mRNA transcripts for degradation (and/or binding) and collectively, these operons represent a significant arsenal of post-transcriptional regulators encoded in the *M. tuberculosis* genome. In other enteric bacteria such as *Salmonella* and *Shigella*, the VapCs are site-specific RNases that cleave initiator tRNA in the anticodon loop (Winther and Gerdes 2011).

The fact that *M. tuberculosis* is responsible for more human deaths per annum when compared to any other infectious agent has seen recent, very significant scientific effort invested in deciphering the biology, pathology and immunology of this organism. The presence of many TA loci in the genome of *M. tuberculosis* is enigmatic and places this organism amongst the top 10 prokaryotes in terms of numbers of chromosomal TA loci (Arcus et al. 2011). This in turn begs the question, what are the biological roles of the TA loci in *M. tuberculosis*? And could the role of these autotoxins be exploited to facilitate more effective treatments for this devastating disease?

TAs have been linked to growth arrest, persistence, antibiotic resistance and the formation of 'non-culturable' cells in a range of bacteria (Gerdes et al. 2005; Lewis 2010; Schumacher et al. 2009; Shleeva et al. 2004). *M. tuberculosis* is particularly successful as a pathogen precisely because of its ability to enter a persistent state where it can survive for decades inside granulomas undergoing

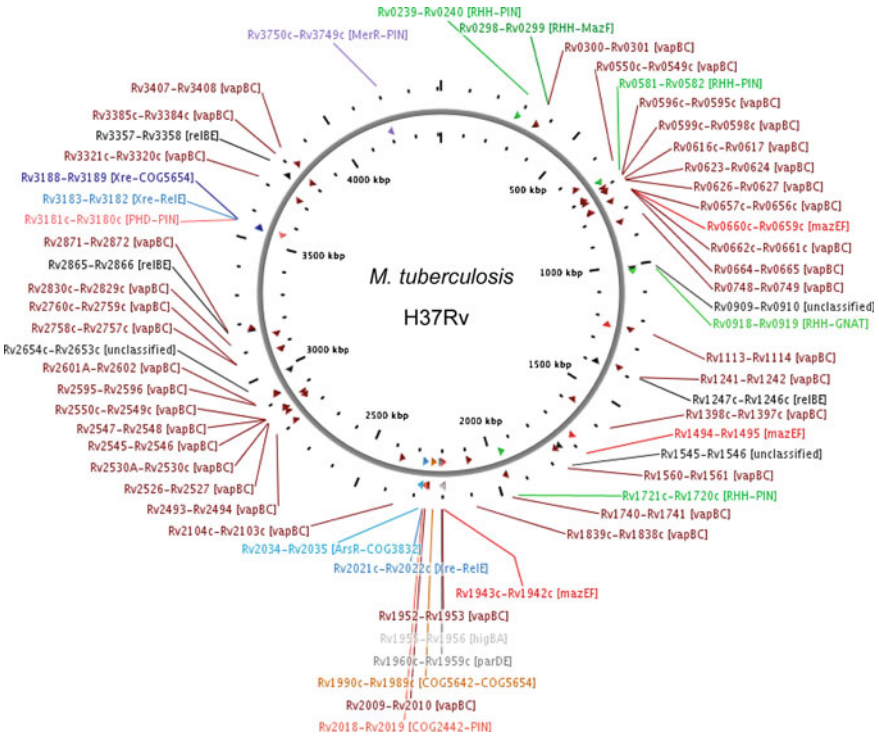


Fig. 8.1 The distribution of TA operons in the genome of *M. tuberculosis* H37Rv. This genome map is taken directly from that produced at the TADB (Shao et al. 2011). TA loci are labelled with their ORF numbers and coloured according to the TA family. A total of 62 TAs are identified in this figure compared to 67 TAs listed in Table 8.2 and 88 TAs identified by Ramage et al. (2009)

virtually no growth and maintaining very low metabolic rates. This 'dormant' lifestyle not only evades the immune response, but also necessitates very long courses of antimicrobial therapy (up to 12 months) to clear any *M. tuberculosis* infection. It is thus tempting to make the connection between large numbers of chromosomal TAs in *M. tuberculosis* and its 'dormant' or persistent non-replicating lifestyle. This observation remains a coincidence for *M. tuberculosis* and there is as yet, no experimental evidence to link TAs with dormancy and persistence in this pathogen. However, significant experimental progress is being made in this area and the link between mRNA interferase TAs and persistence has recently been made in *Escherichia coli* (Maisonneuve et al. 2011).

In terms of TA loci, the mycobacterial species that lie outside the *M. tuberculosis* complex provide a stark contrast. They have very few TA loci encoded in their genomes (Table 8.1). For example, *Mycobacterium smegmatis* has three TA loci (one each of *VapBC*, *Phd–Doc* and *MazEF*) (Frampton et al. 2011) and the causative agent of leprosy, *Mycobacterium leprae*, has none. It can be argued that *M. leprae* is

Table 8.1 Predicted type II TA loci in different mycobacterial species

Organism	Description	Predicted number TA loci
Mycobacterium tuberculosis H37Rv	Pathogen, causes tuberculosis	65
Mycobacterium tuberculosis H37Ra	Non-pathogenic tuberculosis lab strain	66
Mycobacterium leprae	Pathogen, causes leprosy	0
Mycobacterium ulcerans	Pathogen, causes skin lesions	2
Mycobacterium smegmatis mc^2155	Model organism found in soil	3
Mycobacterium marinum	Opportunistic pathogen, free living	1
Mycobacterium gilvum	Environmental, PAH[a] degrading	10
Mycobacterium sp. KMS	Environmental, PAH[a] degrading	15
Mycobacterium vanbaalenii	Environmental, PAH[a] degrading	14

Data taken from Makarova et al. (2009), Sevin and Barloy-Hubler (2007), Shao et al. (2011)
[a] *PAH* polycyclic aromatic hydrocarbon

a special case as its genome has undergone significant decay during recent evolution, shedding nearly half of all its genes (Cole et al. 2001). Indeed, five TA operons can be identified amongst its many pseudogenes (Ramage et al. 2009).

A group of fast-growing environmental organisms that have been studied for their potential application in bioremediation also have small numbers of TA operons encoded in their genomes (both *Mycobacterium gilvum* and *Mycobacterium sp.* KMS have 6 TAs) along with a few additional TAs encoded on plasmids (Table 8.1). Another environmental species, *Mycobacterium vanbaalenii* has 14 TAs encoded in its 6.5 Mb genome. Thus, the mycobacteria can be loosely divided into two groups—the fast-growing environmental organisms with few TAs and the pathogenic slow-growing organisms with abundant TAs. The exception to this generalisation being *M. leprae*.

This chapter reviews the experimental evidence that has been reported for TAs in mycobacteria in an effort to define the biological roles for TAs in this genus. The target for the majority of research has been *M. tuberculosis* due to the importance of this organism in human health. To reflect this, the chapter places some emphasis on the possible biological roles for TAs in the pathology of *M. tuberculosis*.

8.2 Genome-Wide Surveys of Type II TA Loci in Mycobacteria

There have been a number of bioinformatics efforts directed at making a comprehensive survey of type II TA systems across prokaryotes (Makarova et al. 2009; Pandey and Gerdes 2005; Sevin and Barloy-Hubler 2007; Shao et al. 2011). As the search algorithms improve, more and more potential TAs have been predicted. This bioinformatics effort has led to a severe bottleneck as there are now vastly more predicted TA operons than can reasonably be experimentally investigated.

Despite this, several research groups have focussed on the mycobacteria and *M. tuberculosis* in particular, using medium- and high-throughput experimental approaches to investigate the TAs in this genus.

Inouye and colleagues characterised the *mazEF* TAs from *M. tuberculosis* (Zhu et al. 2006). They identified seven *mazEF* operons in the chromosome of *M. tuberculosis* (Table 8.2) and their experimental approach involved expressing the toxic component, MazF, in *E. coli*, and using bacterial growth as the readout for toxicity. They were able to show that of the seven identified *M. tuberculosis mazEF* loci, four exhibited toxicity when expressed in *E. coli*. Importantly, they were able to demonstrate that the toxicity was a result of sequence-specific RNAse activity by MazF (a similar biochemical activity when compared to MazF from *E. coli*) and thus these seven TAs were potentially mRNA interferases in *M. tuberculosis*. They subsequently demonstrated the sequence specificity for two of these MazF toxins (MazF-mt3 and MazF-mt7) and linked this biochemical activity to possible roles in pathogenicity for *M. tuberculosis* (Zhu et al. 2008). A layer of complexity was added by the observation that there is cross-talk between the MazEF and VapBC toxins and antitoxins in *M. tuberculosis* (Zhu et al. 2010). In at least one case, a MazE antitoxin is able to interact with either MazF or one of two VapC toxins abrogating their activity. Further, an MazF (Rv1102c) has been shown to cause growth arrest and an increased number of persister cells in *M. smegmatis* (Han et al. 2010) and the MazF-like protein Rv1991c and its upstream antitoxin have been independently biochemically characterised (Zhao and Zhang 2008). One MazF toxin has been suggested to interact with DNA topoisomerase I in *M. tuberculosis* (Huang and He 2010).

Gupta cloned 38 *M. tuberculosis* toxins and their cognate antitoxins and tested their toxicity in *E. coli* under regulated overexpression (Gupta 2009). Gupta found that 10 of these toxins were bacteriostatic or bacteriocidal when expressed in *E. coli* and that these effects could be rescued by coexpression of the cognate antitoxin.

Similarly, Ramage and Cox used extensive bioinformatics to identify 88 TA operons in the chromosome of *M. tuberculosis*. They then cloned and expressed 65 of these TAs using the related organism *M. smegmatis* as a host and showed that 30 of the toxins inhibited the growth of *M. smegmatis* cells (Ramage et al. 2009). They demonstrated archetypal TA behaviour by coexpression of the toxin and antitoxin components and showed rescue of the toxic phenotype. These investigators described the association of many TAs in *M. tuberculosis* with known genomic islands (implying a possible evolutionary origin via horizontally transferred DNA) and in a small number of cases demonstrated up-regulation of the TA operon under conditions of stress; thus, making the link between TAs and the stress response. This is particularly pertinent in the case of *M. tuberculosis* as it is able to combat the immune system by a well-studied stress response that induces a near-dormant lifestyle; thus, evading both the immune system's defenses and antibiotic treatment. Ramage and Cox demonstrated RNase activity for two *M. tuberculosis* VapC proteins in vitro and translation inhibition for four VapC proteins in vivo. They were able to rule out cross-talk amongst four related VapBC

Table 8.2 Type II TA loci in *Mycobacterium tuberculosis* H37Rv

ORF	Description
Rv0064-0065[a]	*vapBC*
Rv0229c[a, b]	*vapBC*
Rv0239-0240	*vapBC*
Rv0277A-0277c[a]	*vapBC*
Rv0298-0299	new
Rv0300-0301	*vapBC*
Rv0456A-0456B[a]	*mazEF*
Rv0549-0550	*vapBC*
Rv0581-0582	*vapBC*
Rv0595c-0596c	*vapBC*
Rv0598c-0599c	*vapBC*
Rv0608-0609	*vapBC*
Rv0616-0617[a]	*vapBC*
Rv0623-0624	*vapBC*
Rv0626-0627	*vapBC*
Rv0656-0657	*vapBC*
Rv0659c-0660c	*mazEF*
Rv0661c-0662c	*vapBC*
Rv0664-0665	*vapBC*
Rv0748-0749	*vapBC*
Rv0909-0910	new
Rv0959A-0960[a]	*vapBC*
Rv1102c-1103c[a]	*mazEF*
Rv1114-1113	*vapBC*
Rv1241-1242	*vapBC*
Rv1246c-1247c	*relBE*
Rv1397c-1398c	*vapBC*
Rv1494-1495	*mazEF*
Rv1560-1561	*vapBC*
Rv1720c-1721c	*vapBC*
Rv1740-1741	*vapBC*
Rv1838c-1839c	*vapBC*
Rv1942c-1943c	*mazEF*
Rv1952-1953	*vapBC*
Rv1955-1956	*higBA*
Rv1959-1960	*parDE*
Rv1962c-1962A[a]	*vapBC*
Rv1982c-1982A[a]	*vapBC*
Rv1991A-1991c	*mazEF*
Rv2009-2010	*vapBC*
Rv2018-2019	*vapBC*
Rv2021c-2022c	*RelBE*
Rv2063A-2063	*mazEF*
Rv2103c-2104c	*vapBC*
Rv2142A-2142c[a]	*parDE*

(continued)

Table 8.2 (continued)

ORF	Description
Rv2274A-2274c[a]	*mazEF*
Rv2493-2494	*vapBC*
Rv2526-2527	*vapBC*
Rv2530A-2530	*vapBC*
Rv2545-2546	*vapBC*
Rv2547-2548	*vapBC*
Rv2549c-2550c	*vapBC*
Rv2595-Rv2596	*vapBC*
Rv2601-2602	*vapBC*
Rv2653c-2654c	new
Rv2757c-2758c	*vapBC*
Rv2759c-2760c	*vapBC*
Rv2801A-2801c[a]	*mazEF*
Rv2829c-2830c	*vapBC*
Rv2862A-Rv2863[a]	*vapBC*
Rv2865-2866	*relBE*
Rv2871-2872	*vapBC*
Rv3320c-3321c	*vapBC*
Rv3357-3358	*relBE*
Rv3384c-3385c	*vapBC*
Rv3407-3408	*vapBC*
Rv3697c-3697A[a]	*vapBC*

Data taken from Makarova et al. 2009; Ramage et al. 2009; Sevin and Barloy-Hubler 2007; Shao et al. 2011)

[a] Absent in Fig. 8.1

[b] This ORF has toxin and antitoxin as a single gene fusion

TAs and confirmed that in these cases, each VapB antitoxin was a specific inhibitor of the cognate VapC toxin (Ramage et al. 2009).

Mizrahi and colleagues investigated the role of ten *VapBC* TAs from *M. tuberculosis* by first testing the toxicity of each VapC in *M. smegmatis*, then demonstrating rescue by the cognate antitoxin, and finally by confirming VapC toxicity under induced expression in *M. tuberculosis* (Ahidjo et al. 2011). They found that four of the 10 VapC toxins had an effect on growth in *M. smegmatis* and these four were also toxic to *M. tuberculosis*. Like Ramage and Cox, these investigators also showed that two of the mycobacterial VapC toxins (Rv0065 and Rv0617) were sequence specific RNases, although they did not elucidate the precise sequences targeted by these two VapC enzymes.

These genome-wide approaches to testing TA behaviour under conditional expression systems in *E. coli*, *M. smegmatis* and *M. tuberculosis* presented both concordant and discordant results. In some cases, the reported results disagreed even when the same expression host was used. For example, the VapC proteins Rv0595c, Rv2549c and Rv3320c were shown to be toxic in one study (Ahidjo et al. 2011) and nontoxic in another when expressed in *M. smegmatis* (Ramage

et al. 2009). However, these discordant results are most probably a function of the peculiarities of different expression systems rather than genuine conflict. The TAs are notorious for inconsistent expression due to their toxicity and variable solubility. Many researchers have reported spontaneous mutations on plasmids presumably to escape toxicity, as well as a failure to express or insoluble expression, under a range of conditions (Ahidjo et al. 2011; Mckenzie et al. 2012a; Robson et al. 2009).

8.3 *vapBC* Loci in Mycobacteria

There are at least 45 *vapBC* TA loci encoded in each of the genomes of organisms in the *M. tuberculosis* complex (Table 8.2). All other species of mycobacteria have between 1 and 6 *vapBC* TA modules. The null hypothesis is that the *vapBC* operons are recent genomic invaders of the *M. tuberculosis* genome and are selfish modules with no biological significance. However, there is preliminary evidence to suggest that this is not the case. Several *vapBC* loci have been shown to be necessary for the transition from slow to fast growth in *M. tuberculosis* (Beste et al. 2009) and their absence from the *M. leprae* genome suggests that they can be easily lost. This implies that the *vapBC* operons along with other TAs have an important role to play in the biology of the *M. tuberculosis* complex group of organisms.

The *vapBC* TA loci are defined by their *vapC* component belonging to the PIN-domain family of proteins. The PIN-domains are a large protein family with representatives in all three major branches of life. They were originally named for their sequence similarity to the N-terminal domain of an annotated **Pi**lT protein (**Pi**lT **N**-terminal domain), although this historical annotation stems from a domain fusion between a PIN-domain and a PilT ATPase domain. A functional link that connects the PIN-domains with type IV pili (PilT), has not been demonstrated.

The PFam database (Finn et al. 2010) lists 6,253 proteins belonging to the PIN-domain family (PF01850) from 1,341 different species including eukaryotes, eubacteria and archaea. Remarkably, 1,473 PIN-domain sequences are found in the mycobacteria and yet their biological functions in these bacteria are not well understood or well studied.

12 years ago, Clissold and Ponting used bioinformatics to predict that the PIN-domain proteins were Mg^{2+}-dependent RNAses, and suggested that the conserved, active site residues were similar in architecture to phage T4 RNAse H and the Flap endonucleases (Clissold and Ponting 2000). They used PSI-BLAST to detect remote sequence homologues and the PIN-domain sequence signature is now more formally encoded in a Hidden Markov Model (found at PFam, Fig. 8.2). The defining feature of the PIN-domain protein family is three well-conserved acidic residues at approximate positions 4, 40 and 93 in the $\sim$130 amino acid sequence. The fourth acidic residue is less well conserved at $\sim$112. Apart from these three or four acidic residues, there is poor sequence conservation across the family

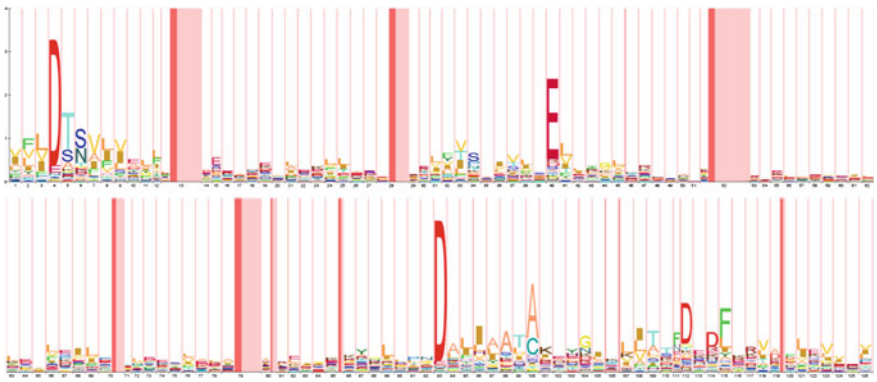


Fig. 8.2 A visual depiction of the hidden Markov model that defines the PIN-domain family of proteins. The height of each letter is proportional to the amount of "information" it provides about the respective position in the PIN-domain family. The width of each column is also an indication of the importance of this position in defining the family. Sequence regions that contain insertions (and therefore almost no information about the family) are shown as *dark* and *light pink columns*. The width of these columns (*dark* + *light pink*) represents the expected length of the insertion and the width of the *dark pink column* is the probability that at least one amino acid is inserted at this point in the sequence. Taken from PFam (Finn et al. 2010) using HMM Logo (Schuster-Böckler et al. 2004)

(Fig. 8.2); however, this is offset by the conservation of three-dimensional structure seen in 11 PIN-domain structures in the Protein Data Bank. The three-dimensional structures cluster the conserved acidic residues together in a putative active site (Arcus et al. 2004; Bunker et al. 2008; Levin et al. 2004).

Subsequent bioinformatics analysis of the genomic context for many PIN-domains in prokaryotes showed that in many cases the gene preceding each PIN-domain had features that resembled known transcription factors and this gene formed an operon with the PIN-domain (Fig. 8.3). This led to the prediction that these operons were TA loci (Anantharaman and Aravind 2003; Arcus et al. 2005; Pandey and Gerdes 2005).

8.4 VapC Structure and Biochemistry

The PIN-domain fold is described in SCOP (Andreeva et al. 2008) as a three-layer $\alpha/\beta/\alpha$ sandwich containing a five-stranded parallel β-sheet in the centre of the structure (Fig. 8.4a). The topology of the protein traces repeating units of β-strand, helix, in a series of right-handed turns. The strand order of the parallel β-sheet is 32145 and, in combination with the right-handed spiral topology, places helices 1–4 above the β-sheet and helices 5–7 below it. This fold brings together three or four conserved acidic residues to form an active site which binds Mg^{2+} or Mn^{2+} ions (Fig. 8.4b) (Arcus et al. 2004; Bunker et al. 2008). A polar residue at $i + 1$ or $i + 2$ following the first acidic residue appears to play a structural role, often

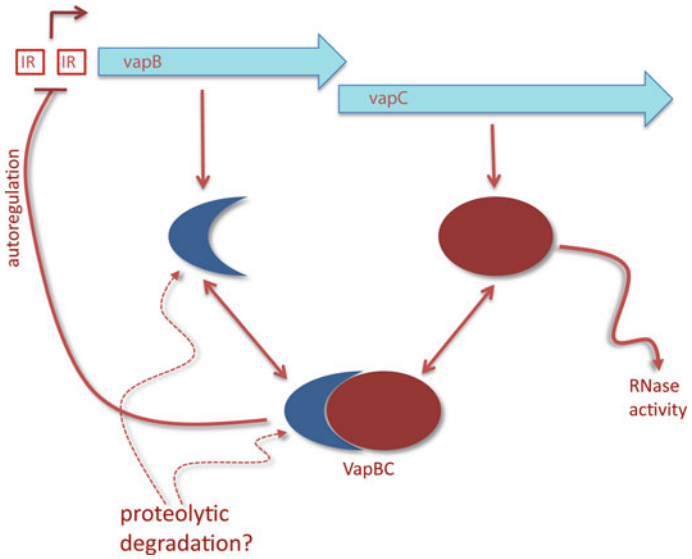


Fig. 8.3 Organisation of a typical *vapBC* TA locus. The *vapB* and *vapC* genes are encoded on different reading frames and these overlap in an operon (shown as *cyan arrows*). The *vapB* and *vapC* genes encode two proteins that form a tight complex. This complex binds to inverted repeats (IR) in the promoter region leading to autoregulation of operon transcription. Activation of VapC RNAse activity is probably a result of proteolytic degradation of the more labile VapB antitoxin

forming a hydrogen bond with its neighbouring acidic residue and presumably positioning it appropriately for catalysis (Fig. 8.4b). All of the prokaryotic PIN-domain structures to date are dimers and dimerisation configures the active sites in a groove along the long axis of the structure (Fig. 8.4c).

The structures for two VapBC TA complexes from *M. tuberculosis* have been solved by Miallau and colleagues (Miallau et al. 2009) (Fig. 8.5a and b) and these elegantly illustrate the tight binding between VapB and VapC and a tetramer of VapBC heterodimers. This arrangement is similar to the structure of FitAB which is a VapBC homologue from *Neisseria gonorrhoeae* (Mattison et al. 2006). The structure of the FitAB hetero-octomer bound to DNA (Fig. 8.5c) beautifully illustrates the autoregulation by *vapBC* loci whereby the TA complex binds to inverted repeats found in the promoter region of the operon (Fig. 8.5c). The tight binding of the antitoxins also indicates the location of the active site groove that would accommodate single-stranded RNA and cleave this substrate. Indeed, based on their structures, Miallau et al. (2009) proposed a catalytic mechanism for *M. tuberculosis* VapC involving two metal ions similar to that of the endo- and exo-nuclease FEN-1, a member of the FLAP nuclease superfamily. This is controversial as comparisons of substrate interactions for the PIN- and NYN- (Nedd4-BP1, YacP nuclease) domains with the FLAP nuclease domains suggests that PIN- and NYN-domains are more likely to coordinate a single metal for catalysis

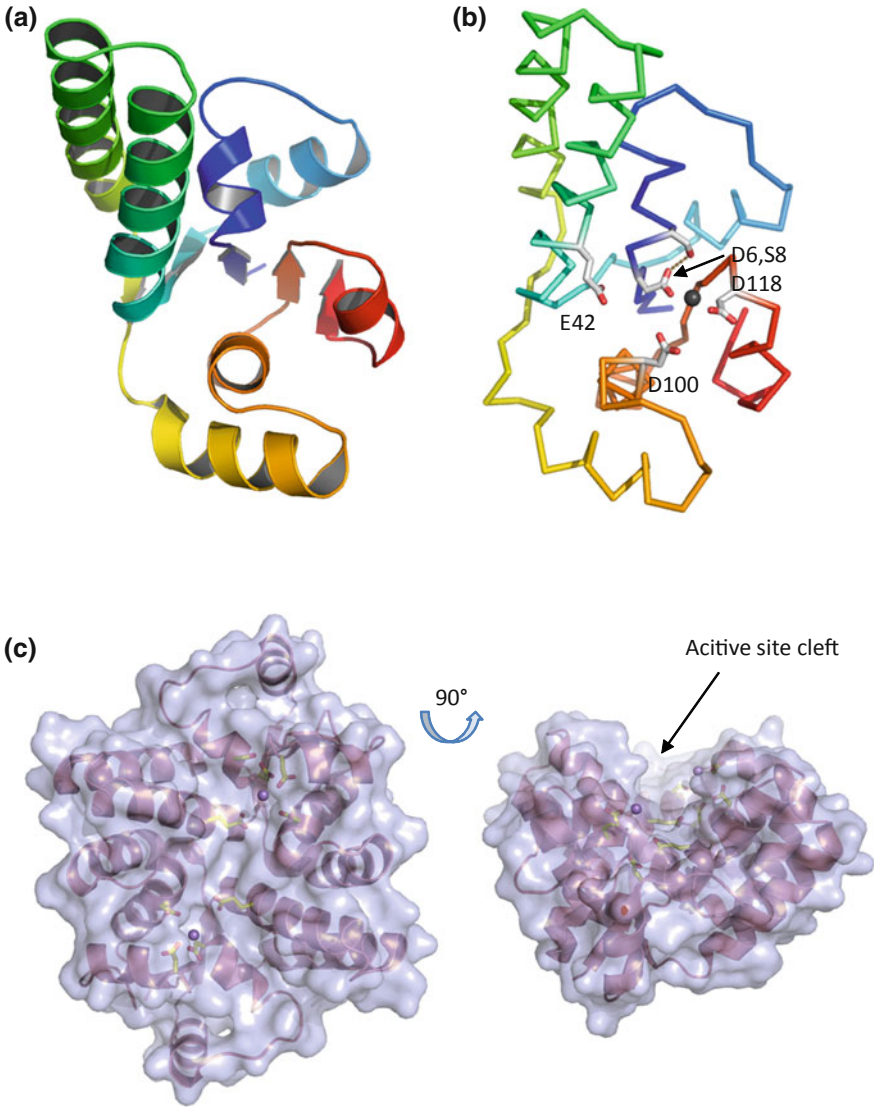


Fig. 8.4 The structure of the PIN-domains. **a** A cartoon representation of a PIN-domain structure coloured *blue-to-red* from the N-terminus to the C-terminus. The five-stranded parallel β-sheet points out of the plane and has strand order 32145 from left to right. The third β-strand is yellow and partially obscured by α-helices 3 and 4. **b** The same structure as a Cα trace showing the four conserved acidic residues, the conserved polar residue and the active site Mn^{2+} ion (in *black*). **c** A PIN-domain dimer as a partially transparent surface, showing the active site channel and two Mn^{2+} ions (in *purple*). Conserved residues are *yellow* and visible beneath the surface. A 90° rotation of the dimer (*at right*) clearly shows the active site cleft. Coordinates were taken from 2FE1 (Bunker et al. 2008) and images were drawn with Pymol (http://www.pymol.org/)

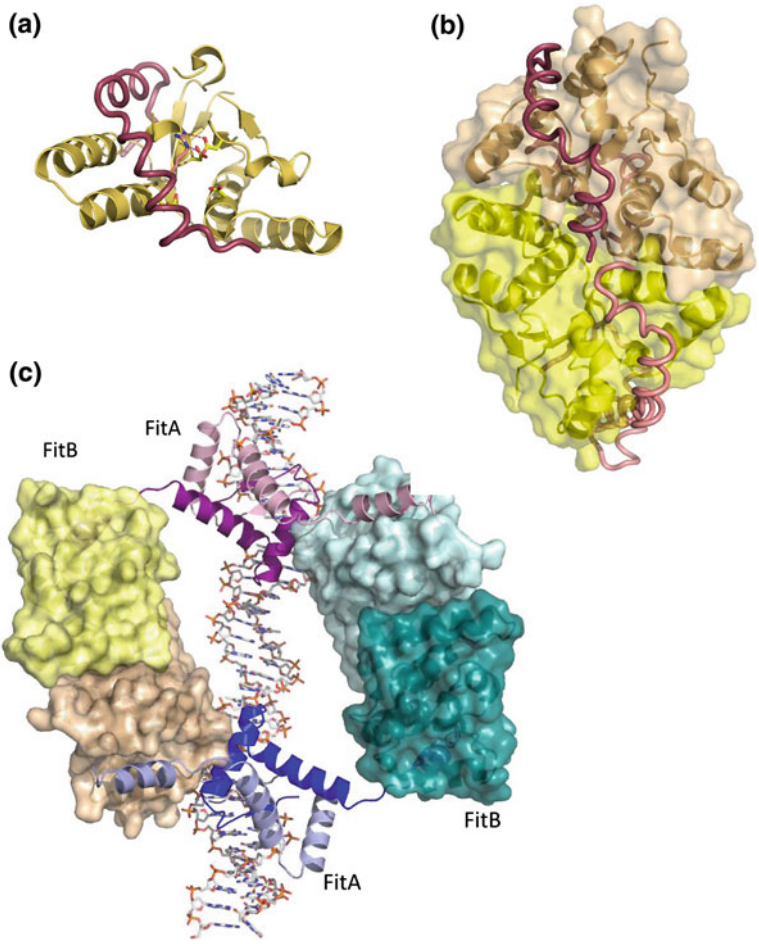


Fig. 8.5 Structures of VapBC protein complexes. **a** The structure of VapC5 (Rv0627) is coloured *orange* and depicted as a cartoon with the four conserved acidic residues in *yellow* as *sticks*. A peptide derived from VapB5 (Rv0626) is bound to VapC5 and is coloured *red*. A key residue, Arg75, is shown forming a salt bridge with one of the conserved acidic residues from VapC5. **b** The structure of VapBC (Rv0300-Rv0301) from *M. tuberculosis*. The VapC dimer is shown as a transparent surface in *yellow* and *brown* with secondary structure elements shown beneath the surface. The C-termini of the two VapB antitoxins are depicted as *red ribbons*. This VapBC dimer is part of a larger hetero-octomeric $(VapBC)_4$ protein complex. **c** The structure of a VapBC homologue, FitAB from *Neisseria gonorrhoeae* (Mattison et al. 2006) bound to promoter DNA. At the *top* of the figure FitA forms a dimer (*two chains coloured pink and purple*). This FitA dimer binds to one half of the DNA inverted repeat. A FitA dimer is also seen at the *bottom* of the figure (*two chains coloured blue and light blue*) binding to the second half of the DNA inverted repeat. Each FitA monomer binds to a FitB monomer—the two FitB dimers lie to the *left* and *right* of the DNA (FitB chains are coloured clockwise from *top right—light green, dark green, tan, yellow*). Double-stranded DNA lies behind the hetero-octomeric FitAB structure

(Anantharaman and Aravind 2006). A crystal structure of VapC from the thermophilic archaea *Pyrobaculum aerophilum* contains just a single Mn^{2+} ion in the active site (Bunker et al. 2008).

Recently, we have shown that the VapC toxins cleave single-stranded RNA in a sequence-specific manner and that the $VapC_{Rv0065}$ and $VapC_{Rv0617}$ from *M. tuberculosis* target the same RNA sequence (Mckenzie et al. 2012a). The products of RNA cleavage are a 5′ RNA fragment that contains a 3′ OH moiety and a 3′ RNA fragment carrying the 5′ phosphate group. This adds further weight to a single-metal catalytic mechanism for VapC RNA cleavage (Mckenzie et al. 2012a). Woychik and colleagues also report sequence-specific RNAse activity on mRNA substrates for $VapC_{Rv0595c}$ from *M. tuberculosis* (Sharp et al. 2012). However, they observe that this catalytic activity is relatively weak and instead ascribe the observed biological functions (translation inhibition and growth arrest) to stable, sequence-specific RNA binding (Sharp et al. 2012).

8.5 The Role of VapBC and Other TAs in *M. smegmatis*

M. smegmatis has been used widely as a model species for studying the mycobacteria genus. Like *M. tuberculosis*, it undergoes an oxygen-limited induced non-replicative state. It is fast growing, non-pathogenic, transformable and provides a powerful model for studying mycobacterial physiology. Unlike *M. tuberculosis*, *M. smegmatis* has a small number of TA operons encoded in its genome: single copies of *phd–doc, mazEF* and *vapBC*. Intriguingly, *the phd–doc* and *vapBC* TAs flank an operon that is duplicated in the genome suggesting that the origin of these TAs in *M. smegmatis* may be in the protection of horizontally transferred DNA.

M. smegmatis VapBC has all the hallmarks of a *bona fide* TA system (see Fig. 8.3): *vapB* and *vapC* form an operon with overlapping open reading frames; they form a benign VapBC protein complex upon expression; VapBC binds to inverted-repeat DNA sequences in the promoter region leading to autoregulation of transcription; inhibition of translation and growth regulation via VapC RNAse activity; and this effect is offset by coexpression of VapB with VapC (Robson et al. 2009). In *M. smegmatis*, conditional expression of VapC is bacteriostatic and not bactericidal pointing towards a role in persistence and growth arrest. More recently, we have shown that the contemporary function for the *vapBC* operon in *M. smegmatis* is as a post-transcriptional regulator of sugar uptake and metabolism (Mckenzie et al. 2012b). Sequence-specific RNAse activity by VapC targets mRNA transcripts that encode sugar transporters and catabolic genes (cleaving mRNA at AUAU and/or AUAA in vitro and in vivo), down-regulating these transcripts as a function of the anabolic rate (growth rate). The *M. smegmatis vapBC* deletion mutant consumes glycerol faster than the glycerol consumption rate required for growth and the molar growth yield on glycerol is lower in the mutant suggesting that growth is uncoupled (Mckenzie et al. 2012b). In addition, a triple deletion *M. smegmatis* mutant (Δ*vapBC*, *mazEF, phd/doc*) was unable to

survive over longer periods of time in rich media due to possible lethal imbalances between catabolic and anabolic rates and/or an increased susceptibility to oxidative stress (Frampton et al. 2011). This effect was not evident in any of the single or double deletion mutants.

In contrast to *M. smegmatis*, *M. tuberculosis* harbours 45 *vapBC* TA systems and has very few sugar transporters (Titgemeyer et al. 2007). No glycerol-specific uptake systems have been identified in *M. tuberculosis* and it has been suggested that facilitated diffusion of glycerol might be sufficient given the slow growth rate (Niederweis 2008). This suggests that *M. tuberculosis* has no mechanism for regulating sugar metabolism at the level of transport, and therefore tighter control of catabolic pathways might be required. Recent work has shown that *M. tuberculosis* does not exhibit substrate preferences i.e. they are able to co-metabolize several substrates simultaneously (de Carvalho et al. 2010). This is achieved by a compartmentalised scheme for metabolism and therefore the need for more than one VapBC module may be anticipated. The genetic requirement in *M. tuberculosis* for two putative VapC proteins (Rv2103c and Rv2494) for strain fitness at slow growth rate and an antitoxin gene (Rv0596c) for fitness at fast growth rate (Beste et al. 2009) lend support to this argument and show obvious routes to testing VapC-mediated regulation and its biological consequences in *M. tuberculosis*.

Very recently, we have shown that $VapC_{Rv0065}$ and $VapC_{Rv0617}$ target the same RNA sequence (Mckenzie et al. 2012a). These two proteins are only distantly related sharing just 22 % amino acid sequence identity. Intuitively, one might expect that distantly related VapCs would target unique subsets of mRNA transcripts as a means to regulate multiple independent metabolic pathways. However, the identical target sequence for these two VapCs may point towards a role for VapBC in global control of mRNA in the *M. tuberculosis* cell under conditions where VapC is mobilised. This mobilisation may be due to the up-regulation of cellular proteases under specific conditions that would readily degrade VapB and liberate VapC inside the cell. Global control of translation by VapC has been observed in the unrelated Enteric bacteria. VapC from *Shigella flexneri* 2a virulence plasmid pMYSH6000 and *Salmonella enterica* have been shown to inhibit translation via cleavage of initiator tRNA (Winther and Gerdes 2011).

8.6 *relBE* and *higBA* Loci in Mycobacteria

Korch et al. (2009) have studied the three RelBE TAs in *M. tuberculosis* (Table 8.2). In a similar approach to previous studies, they demonstrated toxicity for the three RelE proteins by growth inhibition in *M. smegmatis* following induced expression and the rescue of this effect via coexpression of the cognate antitoxin. The physical interaction between toxin and antitoxin was demonstrated along with autoregulation by the protein–protein complex. In an interesting twist to archetypal TA behaviour, two of the RelB antitoxins acted as transcriptional activators when present alone (Korch et al. 2009). The toxin genes were shown to

be expressed in the late stages of macrophage infection by *M. tuberculosis* pointing towards a role in pathogenesis for these TAs.

Further evidence for *M. tuberculosis* TA cross-talk was demonstrated by Yang et al. (2010) who showed cross-talk between the Rel toxin and antitoxin proteins. Two of the three RelB antitoxins showed interactions with each of the three RelE-like toxins, whereas the third antitoxin was specific for its cognate RelE toxin. Further, in two cases, mixed RelBE TA protein complexes were able to bind to the promoter region of the non-cognate TA showing additional cross-talk at the level of autoregulation (Yang et al. 2010).

The role of the *M. tuberculosis* RelBE TAs in persistence and resistance to antibiotic treatment was investigated by Boshoff and colleagues (Singh et al. 2010). They found that all three toxins were up-regulated in vivo but that individual deletion strains were not compromised during murine infection. The RelE toxins appeared to have an effect on increased tolerance to various antibiotics in vitro, however, this effect was not duplicated in vivo.

The *M. tuberculosis higBA* TA operon includes a third, downstream gene Rv1957, which encodes a SecB-like chaperone and is required for proper folding and function of the HigA antitoxin (Bordes et al. 2011). In addition, the autoregulatory role of the HigA antitoxin has recently been reported (Fivian-Hughes and Davis 2010). The toxicity of HigB was demonstrated indirectly insofar as the *higA* antitoxin gene could not be deleted in *M. tuberculosis*.

8.7 Discussion

Since their discovery 25 years ago (Gerdes et al. 1986; Jaffé et al. 1985), the TA operons have been the subject of a great deal of research effort, and several possible biological roles for these ubiquitous prokaryotic genes have been postulated (Hayes 2003; Gerdes et al. 2005; Kolodkin-Gal et al. 2007; Szekeres et al. 2007). The proposed biological roles range from selfish genetic elements to programed cell death, to roles in growth regulation, persistence and antibiotic resistance (Magnuson 2007; Van Melderen and Saavedra De Bast 2009).

The first TAs were identified on plasmids and implicated in plasmid maintenance acting as suicide elements in daughter cells that do not inherit the plasmid upon cell division (Jaffe et al. 1985; Gerdes et al. 1986). This was evidence that TAs act as selfish modules. Other studies have shown that TAs placed on plasmids outcompete the same plasmid without this operon over hundreds of generations (Cooper and Heinemann 2000). Thus, it is highly likely that TAs move around amongst the horizontal gene pool and confer fitness on the mobile genetic elements which carry them. The evolutionary origin of chromosomal copies of TAs, as mobile selfish elements, is also supported by their association in genomes with large numbers of transposable elements. For example, the organism with the most number of TAs, *Microcystis aeroginosa* has 11.8 % of its genome composed of insertion sequences and inverted-repeat transposable elements (Kaneko et al.

2007). TAs in the genome of *M. tuberculosis* have also been associated with mobile elements (Arcus et al. 2005; Ramage et al. 2009). This is further reinforced by the observation that the closest homologues of many of the *M. tuberculosis* TAs are found in unrelated soil organisms and do not conform to the phylogeny of the mycobacteria. Mycobacterial TAs are often found associated with virulence factors, transposases and repetitive elements, suggesting a role in the maintenance of virulence factors (Arcus et al. 2005).

Although it seems likely that the evolutionary origin of the TAs is amongst the horizontal gene pool and that they are able to invade plasmids and genomes and behave as selfish elements, it is also likely that they have been co-opted over time for functions which benefit the host organism. There are several lines of evidence for this co-option hypothesis.

First, TA systems can be lost from a genome and so their selfish behaviour does not permanently fix them in a parent genome. In the cases of the obligate intracellular pathogens, *M. leprae* and *Treponema pallidum,* there are no TAs and the TA loci have most probably been lost via reductive evolution. For *M. leprae*, there are five *pseudo*-TA loci containing stop codons and this attests to the loss of these elements over time. In contrast, their facultative intracellular relatives, *M. tuberculosis*, *M. bovis* and *Treponema denticola* respectively, have large numbers of TA loci. Genome decay along with concomitant loss of TAs can also be seen when comparing the genomes of *Rickettsia conorii* (15 TAs) and *Rickettsia prowzekii* (0 TAs). Genome decay is thought to be an important evolutionary process in the context of intracellular pathogens whose evolution has driven these organisms into more and more specialised intracellular niches (Andersson and Andersson 1999).

Second, a number of TA knockout strains have deleterious phenotypes. A *HipBA* deletion mutant showed a reduction in *E. coli* persister cells (a subpopulation of cells with low metabolic rates) and by inference, a reduction in antibiotic tolerance (Lewis 2007; Schumacher et al. 2009). A *Myxococcus mazF* deletion mutant was unable to make the transition from vegetative growth to fruiting body formation and MazF was shown to be required for programed cell death during *Myxococcus* development (Nariya and Inouye 2008). These observations must be tempered—the *Myxococcus* results have recently been challenged (Lee et al. 2012) and many TA knockout strains show no discernable phenotype under experimental conditions (Tsilibaris et al. 2007).

Thus, it appears likely that the chromosomally encoded mycobacterial TAs have their evolutionary origins amongst mobile genetic elements, but have then been co-opted for contemporary functional roles in many organisms.

Experimental evidence for the contemporary functional roles of the mycobacterial TAs is emerging. Several *M. tuberculosis* TAs are implicated in growth regulation. A total of 16 of the 45 *vapBC* operons in *M. tuberculosis* are either required or positively selected under changes in growth rate (Beste et al. 2009). Six *M. tuberculosis* TAs are up-regulated in response to the antibiotic vancomycin (Provvedi et al. 2009) and there are links between TA activity and other stressors. For example, the *M. smegmatis* triple TA knockout mutant is more susceptible to oxidative stress when compared to wild-type (Frampton et al. 2011).

There is a wide range of indirect evidence that the TAs in mycobacteria are functional. Many of the 65 toxins show toxicity upon induced expression in various hosts and in nearly all cases, this is offset by their cognate antitoxin. However, direct evidence for the role of TAs in mycobacterial survival and physiology is scarce. The *vapBC* operon in *M. smegmatis* is responsible for coupling metabolism with growth rate and the Δ*vapBC* mutant indiscriminately consumes glycerol and inefficiently converts this carbon to biomass. The challenge for TA biologists is to devise strategies to investigate the biology of the ∼88 TAs in *M. tuberculosis* where there may be cross-talk, selfish origins and subtle phenotypes in this most difficult of organisms to study. Overcoming this challenge will reap huge rewards as TB is the most significant infectious organism on Earth.

Acknowledgments We thank Valerie Mizrahi for helpful discussions. We gratefully acknowledge funding from the Health Research Council of New Zealand and the Marsden Fund that has supported research in the Arcus and Cook laboratories.

References

Ahidjo, B. A., Kuhnert, D., Mckenzie, J. L., et al. (2011). VapC toxins from *Mycobacterium tuberculosis* are ribonucleases that differentially inhibit growth and are neutralized by cognate VapB antitoxins. *PLoS ONE, 6*, e21738. doi:10.1371/journal.pone.0021738.

Anantharaman, V., & Aravind, L. (2003). New connections in the prokaryotic toxin–antitoxin network: Relationship with the eukaryotic nonsense-mediated RNA decay system. *Genome Biology, 4*, R81. doi:10.1186/gb-2003-4-12-r81.

Anantharaman, V., & Aravind, L. (2006). The NYN domains: Novel predicted RNAses with a PIN domain-like fold. *RNA Biology*, *3*, 18–27.

Andersson, J. O., & Andersson, S. G. (1999). Insights into the evolutionary process of genome degradation. *Current Opinion in Genetics & Development, 9*, 664–671.

Andreeva, A., Howorth, D., Chandonia, J.-M., et al. (2008). Data growth and its impact on the SCOP database: New developments. *Nucleic Acids Research, 36*, D419–D425. doi:10.1093/nar/gkm993.

Arcus, V. L., Bäckbro, K., Roos, A., et al. (2004). Distant structural homology leads to the functional characterization of an archaeal PIN domain as an exonuclease. *Journal of Biological Chemistry, 279*, 16471–16478. doi:10.1074/jbc.M313833200.

Arcus, V. L., Rainey, P. B., & Turner, S. J. (2005). The PIN-domain toxin–antitoxin array in mycobacteria. *Trends in Microbiology, 13*, 360–365. doi:10.1016/j.tim.2005.06.008.

Arcus, V. L., Mckenzie, J. L., Robson, J., & Cook, G. M. (2011). The PIN-domain ribonucleases and the prokaryotic VapBC toxin–antitoxin array. *Protein Engineering Design Select, 24*, 33–40. doi:10.1093/protein/gzq081.

Beste, D., Espasa, M., Bonde, B., & Kierzek, A. (2009). The genetic requirements for fast and slow growth in Mycobacteria. *PLoS ONE, 4*, e5349. doi:10.1371/journal.pone.0005349.

Bordes, P., Cirinesi, A.-M., Ummels, R., et al. (2011). SecB-like chaperone controls a toxin–antitoxin stress-responsive system in *Mycobacterium tuberculosis*. *Proceedings of the National academy of Sciences of the United States of America, 108*, 8438–8443. doi:10.1073/pnas.1101189108.

Bunker, R. D., Mckenzie, J. L., Baker, E. N., & Arcus, V. L. (2008). Crystal structure of PAE0151 from *Pyrobaculum aerophilum*, a PIN-domain (VapC) protein from a toxin–antitoxin operon. *Proteins, 72*, 510–518. doi:10.1002/prot.22048.

Caminero, J. A., Sotgiu, G., Zumla, A., & Migliori, G. B. (2010). Best drug treatment for multidrug-resistant and extensively drug-resistant tuberculosis. *The Lancet Infectious Diseases, 10*, 621–629. doi:10.1016/S1473-3099(10)70139-0.

Clissold, P., & Ponting, C. (2000). PIN domains in nonsense-mediated mRNA decay and RNAi. *Current Biology, 10*, R888–R890.

Cole, S. T., Eiglmeier, K., Parkhill, J., et al. (2001). Massive gene decay in the leprosy bacillus. *Nature, 409*, 1007–1011. doi:10.1038/35059006.

Cooper, T. F., & Heinemann, J. A. (2000). Postsegregational killing does not increase plasmid stability but acts to mediate the exclusion of competing plasmids. *Proceedings of the National Academy of Sciences* USA, *97*, 12643–12648. doi:10.1073/pnas.220077897.

de Carvalho, L. P. S., Fischer, S. M., Marrero, J., et al. (2010). Metabolomics of *Mycobacterium tuberculosis* reveals compartmentalized co-catabolism of carbon substrates. *Chemistry & Biology, 17*, 1122–1131. doi:10.1016/j.chembiol.2010.08.009.

Finn, R. D., Mistry, J., Tate, J., et al. (2010). The Pfam protein families database. *Nucleic Acids Research, 38*, D211–D222. doi:10.1093/nar/gkp985.

Fivian-Hughes, A. S., & Davis, E. O. (2010). Analyzing the regulatory role of the HigA antitoxin within *Mycobacterium tuberculosis*. *Journal of Bacteriology, 192*, 4348–4356. doi:10.1128/JB.00454-10.

Frampton, R. A., Aggio, R. B. M., Villas-Bôas, S. G., et al. (2011). Toxin–antitoxin systems of *Mycobacterium smegmatis* are essential for cell survival. *Journal of Biological Chemistry, 287*, 5340–5356. doi:10.1074/jbc.M111.286856.

Gerdes, K., Rasmussen, P. B., & Molin, S. (1986). Unique type of plasmid maintenance function: Postsegregational killing of plasmid-free cells. *Proceedings of the National academy of Sciences of the United States of America, 83*, 3116–3120.

Gerdes, K., Christensen, S. K., & Løbner-Olesen, A. (2005). Prokaryotic toxin–antitoxin stress response loci. *Nature Reviews Microbiology, 3*, 371–382. doi:10.1038/nrmicro1147.

Gupta, A. (2009). Killing activity and rescue function of genome-wide toxin–antitoxin loci of *Mycobacterium tuberculosis*. *FEMS Microbiology Letters, 290*, 45–53. doi:10.1111/j.1574-6968.2008.01400.x.

Han, J.-S., Lee, J. J., Anandan, T., et al. (2010). Characterization of a chromosomal toxin–antitoxin, Rv1102c-Rv1103c system in *Mycobacterium tuberculosis*. *Biochemical & Biophysical Research Communications, 400*, 293–298. doi:10.1016/j.bbrc.2010.08.023.

Hayes, F. (2003). Toxins-antitoxins: Plasmid maintenance, programmed cell death, and cell cycle arrest. *Science*, *301*, 1496–1499.

Huang, F., & He, Z.-G. (2010). Characterization of an interplay between a *Mycobacterium tuberculosis* MazF homolog, Rv1495 and its sole DNA topoisomerase I. *Nucleic Acids Research, 38*, 8219–8230. doi:10.1093/nar/gkq737.

Jaffé, A., Ogura, T., & Hiraga, S. (1985). Effects of the ccd function of the F plasmid on bacterial growth. *Journal of Bacteriology, 163*, 841–849.

Kaneko, T., Nakajima, N., Okamoto, S., et al. (2007). Complete genomic structure of the bloom-forming toxic cyanobacterium *Microcystis aeruginosa* NIES-843. *DNA Research, 14*, 247–256. doi:10.1093/dnares/dsm026.

Kolodkin-Gal, I., Hazan, R., & Gaathon, A., et al. (2007). A linear pentapeptide is aquorum-sensing factor required for mazEF-mediated cell death in Escherichia coli. *Science*, *318*, 652–655. doi:10.1126/science.1147248.

Korch, S. B., Contreras, H., & Clark-Curtiss, J. E. (2009). Three *Mycobacterium tuberculosis* Rel toxin–antitoxin modules inhibit mycobacterial growth and are expressed in infected human macrophages. *Journal of Bacteriology, 191*, 1618–1630. doi:10.1128/JB.01318-08.

Lee, B., Holkenbrink, C., Treuner-Lange, A., & Higgs, P. I. (2012). *Myxococcus xanthus* developmental cell fate production: Heterogeneous accumulation of developmental regulatory proteins and reexamination of the role of MazF in developmental lysis. *Journal of Bacteriology, 194*, 3058–3068. doi:10.1128/JB.06756-11.

Levin, I., Schwarzenbacher, R., Page, R., et al. (2004). Crystal structure of a PIN (PilT N-terminus) domain (AF0591) from *Archaeoglobus fulgidus* at 1.90 Å resolution. *Proteins, 56*, 404–408. doi:10.1002/prot.20090.

Lewis, K. (2007). Persister cells, dormancy and infectious disease. *Nature Reviews Microbiology, 5*, 48–56. doi:10.1038/nrmicro1557.

Lewis, K. (2010). Persister cells. *Annual Review of Microbiology, 64*, 357–372. doi:10.1146/annurev.micro.112408.134306.

Magnuson, R. D. (2007). Hypothetical functions of toxin-antitoxin systems. *Journal of Bacteriology*, *189*, 6089–6092. doi:10.1128/JB.00958-07.

Maisonneuve, E., Shakespeare, L. J., Jørgensen, M. G., & Gerdes, K. (2011). Bacterial persistence by RNA endonucleases. *Proceedings of the National academy of Sciences of the United States of America, 108*, 13206–13211. doi:10.1073/pnas.1100186108.

Makarova, K. S., Wolf, Y. I., & Koonin, E. V. (2009). Comprehensive comparative-genomic analysis of Type 2 toxin–antitoxin systems and related mobile stress response systems in prokaryotes. *Biology Direct, 4*, 19. doi:10.1186/1745-6150-4-19.

Mattison, K., Wilbur, J. S., So, M., & Brennan, R. G. (2006). Structure of FitAB from *Neisseria gonorrhoeae* bound to DNA reveals a tetramer of toxin–antitoxin heterodimers containing PIN-domains and ribbon-helix-helix motifs. *Journal of Biological Chemistry, 281*, 37942–37951. doi:10.1074/jbc.M605198200.

Mckenzie, J. L., Duyvestyn, J. M., Smith, T., et al. (2012a). Determination of ribonuclease sequence-specificity using pentaprobes and mass spectrometry. *RNA, 18*, 1267–1278. doi:10.1261/rna.031229.111.

Mckenzie, J. L., Robson, J., Berney, M., et al. (2012b). A VapBC toxin–antitoxin module is a post-transcriptional regulator of metabolic flux in mycobacteria. *Journal of Bacteriology, 194*, 2189–2204. doi:10.1128/JB.06790-11.

Miallau, L., Faller, M., & Chiang, J., et al. (2009). Structure and proposed activity of a member of the VapBC family of toxin-antitoxin systems. VapBC-5 from Mycobacterium tuberculosis. *Journal of Biological Chemistry*, *284*, 276–283. doi:10.1074/jbc.M805061200.

Nariya, H., & Inouye, M. (2008). MazF, an mRNA interferase, mediates programmed cell death during multicellular Myxococcus development. *Cell, 132*, 55–66. doi:10.1016/j.cell.2007.11.044.

Niederweis, M. (2008). Nutrient acquisition by mycobacteria. *Microbiology, 154*, 679–692. doi:10.1099/mic.0.2007/012872-0.

Pandey, D. P., & Gerdes, K. (2005). Toxin–antitoxin loci are highly abundant in free-living but lost from host-associated prokaryotes. *Nucleic Acids Research, 33*, 966–976. doi:10.1093/nar/gki201.

Provvedi, R., Boldrin, F., Falciani, F., et al. (2009). Global transcriptional response to vancomycin in *Mycobacterium tuberculosis*. *Microbiology, 155*, 1093–1102. doi:10.1099/mic.0.024802-0.

Ramage, H. R., Connolly, L. E., & Cox, J. S. (2009). Comprehensive functional analysis of *Mycobacterium tuberculosis* toxin–antitoxin systems: Implications for pathogenesis, stress responses, and evolution. *PLoS Genetics, 5*, e1000767. doi:10.1371/journal.pgen.1000767.

Robson, J., Mckenzie, J. L., Cursons, R. T., et al. (2009). The *vapBC* operon from *Mycobacterium smegmatis* is an autoregulated toxin–antitoxin module that controls growth via inhibition of translation. *Journal of Molecular Biology, 390*, 353–367. doi:10.1016/j.jmb.2009.05.006.

Schumacher, M. A., Piro, K. M., Xu, W., et al. (2009). Molecular mechanisms of HipA-mediated multidrug tolerance and its neutralization by HipB. *Science, 323*, 396–401. doi:10.1126/science.1163806.

Schuster-Böckler, B., Schultz, J., & Rahmann, S. (2004). HMM logos for visualization of protein families. *BMC Bioinformatics, 5*, 7. doi:10.1186/1471-2105-5-7.

Sevin, E. W., & Barloy-Hubler, F. (2007). RASTA-Bacteria: A web-based tool for identifying toxin–antitoxin loci in prokaryotes. *Genome Biology, 8*, R155. doi:10.1186/gb-2007-8-8-r155.

Shao, Y., Harrison, E. M., Bi, D., et al. (2011). TADB: A web-based resource for type 2 toxin–antitoxin loci in bacteria and archaea. *Nucleic Acids Research, 39*, D606–D611. doi:10.1093/nar/gkq908.

Sharp, J. D., Cruz, J. W., Raman, S., et al. (2012). Growth and translation inhibition through sequence specific RNA binding by a *Mycobacterium tuberculosis* VapC toxin. *Journal of Biological Chemistry, 287*, 12835–12847. doi:10.1074/jbc.M112.340109.

Shleeva, M., Mukamolova, G. V., Young, M., et al. (2004). Formation of "non-culturable" cells of *Mycobacterium smegmatis* in stationary phase in response to growth under suboptimal conditions and their Rpf-mediated resuscitation. *Microbiology, 150*, 1687–1697. doi:10.1099/mic.0.26893-0.

Singh, R., Barry, C. E., & Boshoff, H. I. M. (2010). The three RelE homologs of *Mycobacterium tuberculosis* have individual, drug-specific effects on bacterial antibiotic tolerance. *Journal of Bacteriology, 192*, 1279–1291. doi:10.1128/JB.01285-09.

Szekeres, S., Dauti, M., & Wilde, C., et al. (2007). Chromosomal toxin-antitoxin loci can diminish large-scale genome reductions in the absence of selection. *Molecular Microbiology, 63*, 1588–1605. doi:10.1111/j.1365-2958.2007.05613.x.

Titgemeyer, F., Amon, J., Parche, S., et al. (2007). A genomic view of sugar transport in *Mycobacterium smegmatis* and *Mycobacterium tuberculosis*. *Journal of Bacteriology, 189*, 5903–5915. doi:10.1128/JB.00257-07.

Tsilibaris, V., Maenhaut-Michel, G., Mine, N., & van Melderen, L. (2007). What is the benefit to *Escherichia coli* of having multiple toxin–antitoxin systems in its genome? *Journal of Bacteriology, 189*, 6101–6108. doi:10.1128/JB.00527-07.

van Melderen, L., & Saavedra De Bast, M. (2009). Bacterial toxin-antitoxin systems: More than selfish entities? *PLoS Genetics*, *5*, e1000437. doi:10.1371/journal.pgen.1000437.

Winther, K. S., & Gerdes, K. (2011). Enteric virulence associated protein VapC inhibits translation by cleavage of initiator tRNA. *Proceedings of the National academy of Sciences of the United States of America, 108*, 7403–7407. doi:10.1073/pnas.1019587108.

Yang, M., Gao, C., Wang, Y., et al. (2010). Characterization of the interaction and cross-regulation of three *Mycobacterium tuberculosis* RelBE modules. *PLoS ONE, 5*, e10672. doi:10.1371/journal.pone.0010672.

Zhao, L., & Zhang, J. (2008). Biochemical characterization of a chromosomal toxin–antitoxin system in *Mycobacterium tuberculosis*. *FEBS Letters, 582*, 710–714. doi:10.1016/j.febslet.2008.01.045.

Zhu, L., Zhang, Y., Teh, J., et al. (2006). Characterization of mRNA interferases from *Mycobacterium tuberculosis*. *Journal of Biological Chemistry, 281*, 18638–18643. doi:10.1074/jbc.M512693200.

Zhu, L., Phadtare, S., Nariya, H., et al. (2008). The mRNA interferases, MazF-mt3 and MazF-mt7 from *Mycobacterium tuberculosis* target unique pentad sequences in single-stranded RNA. *Molecular Microbiology, 69*, 559–569. doi:10.1111/j.1365-2958.2008.06284.x.

Zhu, L., Sharp, J. D., Kobayashi, H., et al. (2010). Noncognate *Mycobacterium tuberculosis* toxin–antitoxins can physically and functionally interact. *Journal of Biological Chemistry, 285*, 39732–39738. doi:10.1074/jbc.M110.163105.

Chapter 9
Type II Toxin-Antitoxin Loci: The *phd/doc* Family

Abel Garcia-Pino, Yann Sterckx, Roy D. Magnuson and Remy Loris

Abstract The *phd/doc* family is one the smallest families of toxin–antitoxin modules and was first discovered as a plasmid addiction module on *E. coli* bacteriophage P1. The toxin Doc interacts with the ribosome, competes with hygromycin, and inhibits translation. Structurally, Doc resembles Fic domains, which are known to transfer an AMP moiety to the hydroxyphenyl group of a tyrosine in the target protein. Although much of the AMP/ATP binding site of Fic is conserved in Doc, no specific enzymatic activity has yet been linked to Doc. Nevertheless, mutations in the loop corresponding to the active site loop of Fic render Doc inactive. Regulation of the P1 *phd/doc* operon is understood in terms of a detailed molecular mechanism and involves conditional cooperativity. The N-terminal domain of Phd forms the DNA-binding unit and represents a common DNA-binding fold that is also shared with a number of antitoxins from different TA families, among which YefM is the best studied. Enhancement of the DNA-binding affinity of Phd by Doc stems both from allosteric coupling between the Doc- and DNA-binding sites on Phd and from avidity effects due to Doc-mediated bridging of two Phd dimers bound to the operator site. Activation of the system at high Doc-to-Phd ratios stems from a low-to-high affinity switch in the interaction between Phd and Doc.

A. Garcia-Pino · Y. Sterckx · R. Loris (✉)
Department of Structural Biology,VIB and Structural Biology Brussels,
Vrije Universiteit Brussel, Pleinlaan 2, B-1050 Brussels, Belgium
e-mail: Remy.Loris@VIB-VUB.be

R. D. Magnuson
Department of Biological Sciences, University of Alabama in Huntsville,
301 Sparkman Drive, SC 369RHuntsville, AL 35758, USA

K. Gerdes (ed.), *Prokaryotic Toxin-Antitoxins*,
DOI: 10.1007/978-3-642-33253-1_9,

Fig. 9.1 Maintenance of the bacteriophage P1 genome within a bacterial host cell population by post-segregational killing. **a** Upon its injection into a bacterial host cell, the genome of bacteriophage P1 circularizes and behaves as a low-copy-number plasmid. Upon cell division, plasmid partitioning systems encoded on the P1 genome ensure that each daughter cell receives one copy of the circular genome. The type II TA system *phd/doc* eliminates plasmid-free daughter cells from the population by the post-segregational killing mechanism. **b** Map of the circularized bacteriophage P1 genome based on the work of Lobocka et al. (2004). *Dark gray* genes are involved in the architecture of the bacteriophage's coat and *green* genes bear a function in the phage's lytic life cycle. The *yellow genes* are involved in plasmid partitioning. The *phd/doc* plasmid stabilization module is indicated in purple. All other genes are shown in *light gray*. **c** Enlarged view of the bacteriophage P1 *phd/doc* operon. The *gray boxes* indicate the −35 region (TGGTGC), the −10 region (TATAAT), and the Ribosome Binding Site (RBS, CACGAGGTGT), respectively. The *brown box* indicates the *phd/doc* operator region (OR) with its two imperfect palindromes (GTGTACAC and GAGTACAC). The *cyan and red boxes* highlight the respective open reading frames of *phd* and *doc* ▶

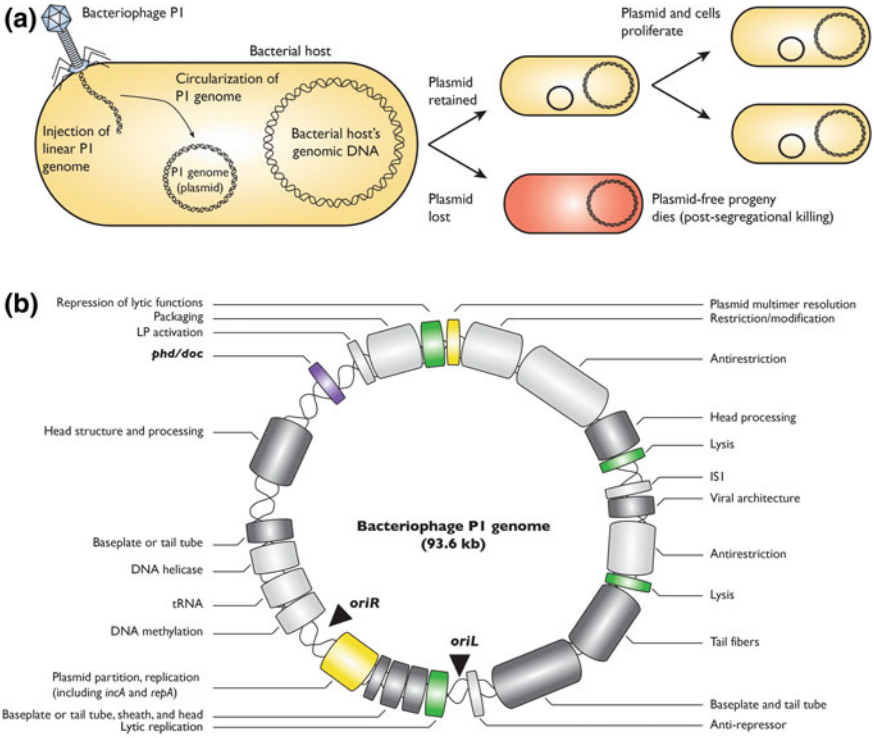
9.1 Discovery of the *phd–doc* Module and its Role in P1 Biology

The discovery and early work on *phd/doc* is closely linked with the replication and life cycle of *Escherichia coli* bacteriophage P1 (Fig. 9.1). This bacteriophage can be transmitted horizontally, as a virus, or vertically, as a fairly autonomous plasmid prophage (Fig. 9.1a). Loss of the plasmid prophage is low: the apparent loss rate is circa 10^{-5} per cell per generation. Four phage-encoded systems (replication, partition, dimer resolution, and addiction) contribute to the stability of the plasmid prophage (reviewed in Lobocka et al. 2004).

A map of the bacteriophage P1 genome is given in Fig. 9.1b. The replication control system of bacteriophage P1 includes as its principle components *oriR* (the origin of replication), RepA (an initiator/regulator protein), various host factors including DnaA, and *incA* (a set of *repA*-binding sites involved in regulating copy number). The main job of the replication system is to ensure that there are at least two plasmid molecules available when the cell is about to divide, certainly not fewer (which would make partition difficult) and preferably not many more (which might represent an intolerable metabolic burden to the host).

In 1993, the lab of Michael Yarmolinsky at NIH discovered a small operon on bacteriophage P1 that helps to stabilize the prophage as a low copy number plasmid (Lehnherr et al. 1993, Fig. 9.1c). The discovery of the P1 plasmid addiction operon was motivated in part by the discrepancy between the loss rate of the natural P1 plasmid (circa 10^{-5} per cell per generation) and the higher loss rates observed in smaller, but partition proficient derivatives. This discrepancy suggested that the P1 genome contains an "addiction" locus that might function by postsegregational killing, similar to addiction systems previously identified on certain low copy number plasmids, such as F and R1.

To demonstrate the addictive nature of bacteriophage P1, the replication of P1 was rendered strictly thermosensitive by the use of a RepA amber allele, in



combination with a thermosensitive amber suppressor. Additionally, an extra source of repressor was provided (in order to repress the lytic replicon). At the permissive temperature, P1 plasmid replication proceeded normally. At the restrictive temperature, plasmid replication was arrested. The production of viable cells, as indicated by production of colony-forming units, was similarly restricted. Optical density and other measures of cell mass nevertheless continued to rise for a while, indicating that the plasmid-containing cells were able to continue spitting

off plasmid-free daughter cells, which were of limited replicative potential. Interestingly, the plasmid addiction system cannot only confer the addictive phenotype, but also relieve it. Progressively smaller fragments of P1 were identified first by their ability to prevent arrest of plasmid-free daughter cells, and then later by their ability to produce such arrest when placed on a thermosensitive plasmid replicon. A combination of deletion mapping, transposon mutagenesis, and site-directed mutation was used to map toxin and antitoxin functions to their respective open reading frames. The toxin was named Doc for "Death on curing" since the loss of the prophage from a lysogen is traditionally called "curing" the lysogen). The corresponding antitoxin or antidote was named Phd for "Prevents host death" (Lehnherr et al. 1993).

The addictive nature of the *phd/doc* operon crucially depends on the action of ClpXP, which continuously degrades Phd. In presence of a functional ClpXP, the half-life of Phd is about 120 min, or two bacterial generations under the conditions used (Lehnherr and Yarmolinsky 1995). This relatively slow degradation probably helps to avoid premature activation of Doc. In absence of a functional copy of ClpXP, Phd levels remain stable over several generations, and the cells survive loss of the P1 plasmid.

Although the *phd/doc* module on bacteriophage P1 is a member of a family of TA modules, almost all genetic and certainly all biochemical and biophysical studies have concentrated on the Phd and Doc proteins encoded in the P1 genome. Therefore, in the remainder of this chapter, the names "Phd" and "Doc" will generally refer to these two particular proteins. Phd consists of 73 amino acids and is functional as a dimer while Doc encompasses 126 amino acids and is functional as a monomer.

9.2 Distribution of Phd, Doc, and *phd/doc* Modules

At the time of its discovery in 1993, the *phd/doc* module on bacteriophage P1 was an isolated "plasmid addiction" operon for which no homologs could be found among the known bacterial gene or protein sequences. At that time, "plasmid addiction modules" had only been identified on a few low copy number plasmids, and most of these bore no relationship to each other (Hiraga et al. 1986; Ruiz-Echevarria et al. 1991a, b; Roberts et al. 1994;). Their origin was unknown and postsegregational killing of plasmid-free segregants was the only suggested function. In 1997, four operons closely related to P1 *phd/doc* were discovered on three large *E. coli* conjugative plasmids and on the chromosome of *E. coli* K-12 DH5α (*hsdr recA*) (Tyndall et al. 1997). These are located in segments with high sequence identity to bacteriophage P1 and flanking type IC family of DNA restriction and modification systems. All but one of them are functional. Other homologs were found in bacteriophage P7 (Lehnherr et al. 1993), on the chromosome of *Salmonella enterica* serovar Typhimurium (Zhao and Magnuson

2005). *Francisella* plasmid pFNL10 codes for a homolog of Phd (Pomerantsev et al. 2001).

More information on the distribution of *phd/doc* modules in prokaryotes had to wait the availability of full genome sequences. In a survey of 126 prokaryotic genomes, a search for Doc homologs resulted in the identification of 25 *phd/doc* modules and three lone *doc* genes not associated with an identifiable antitoxin (Pandey and Gerdes 2005). These are distributed over Gram-positive and Gram-negative *Bacteria* as well as *Archae*a. The *phd/doc* family appears to be a relatively minor family of TA modules and after *ccd* the least widely distributed among the seven classic TA families (*ccd*, *mazEF*, *relBE*, *vapBC*, *phd/doc*, *parDE*, and *higBA*).

In recent years, it has become clear that classifications of TA modules based on connected pairs of toxins and antitoxins makes little sense. Indeed, essentially every type of known toxin can be found associated with virtually every type of known antitoxin, leading to "chimeric" TA modules (Anantharaman and Aravind 2003; Leplae et al. 2011). Thus, Doc-like toxins can be found associated not only with Phd-like antitoxins but also with MazE-like antitoxins (known as *tasAB* modules—Fico and Mahillon 2006) and several previously unrecognized antitoxins identified via the "guilt by association" principle (if a known TA toxin or antitoxin is found associated with the gene of another unknown protein in a two-gene operon, the second gene is assumed to encode for a corresponding antitoxin or toxin, respectively).

Similarly, Phd-type antitoxins are further associated with toxins belonging to the RelE/ParE superfamily, VapC toxins, YafO family members, Zeta-toxins, and mostly yet not validated toxins identified via "guilt by association" (Leplae et al. 2011). It should be noted that Doc-type toxins form only a very small percentage of the Phd partners, the major ones belonging to the RelE/ParE superfamily. Among these, *E. coli* YefM, the antitoxin pairing with the RelE-like YoeB has been well studied (sometimes, *yoeB/yefM* is referred to as *axe-txe*—Grady and Hayes 2003). YefM was shown to be a largely unfolded protein in solution that folds upon interacting with YoeB (Gazit and Sauer 1999a; Cherny and Gazit 2004). It was also the first member of the Phd family for which a crystal structure became available, both in complex with YoeB (*E. coli* version—Kamada and Hanaoka 2005) and in its isolated state (*M. tuberculosis* version—Kumar et al. 2008). The latter caused a small controversy since this crystal structure was seen to disprove the unfolded nature of YefM in solution. As will be discussed later in this chapter, further data on P1 Phd are able to resolve this apparent contradiction.

The breakpoints in the chimeric TA systems appear to map not between the toxin and the antitoxin, but within the antitoxin protein, and in particular between the well-conserved repressor domains and the more variable toxin-neutralizing domains (Smith and Magnuson 2004; McKinley and Magnuson 2005; Magnuson 2007). The DNA-binding domains in the antitoxins can, thus, be exchanged easily without compromising functionality.

9.3 Phd and Doc Homologs Outside the TA Context

An exhaustive BLAST search by Anantharaman and Aravind (2003) identified three groups of *Doc*-related genes. The first and obvious group corresponds to Doc homologs that are part of bona fide TA modules. Two other groups of Doc-related proteins or protein domains occur outside the TA context. One of them concerns the Fic proteins. These are single domain proteins with weak sequence identity to doc but which cluster together phylogenetically and are characterized by a 12 amino acid insertion before amino acid position 28 of P1 doc (just before helix α2—see below). They do not in general appear to be part of a TA operon and are not found in any sort of conserved operon architecture. The archetype here is the *E. coli* Fic protein, mutations in which cause filamentation (hence their name: filamentation induced by cyclic AMP) (Utsumi et al. 1982; Kawamukai et al. 1988; Komano et al. 1991). Their function remains unknown.

The other doc-related family outside an obvious TA context concerns Fic domains that are part of larger multidomain proteins. They are present not only in *Bacteria* and *Achaea* but also in animals. They are typified by the animal HYPE protein, of which the human version is known to interact with the Huntingtin protein. No genes encoding putative transcriptional regulators are linked to the genes for the prokaryotic HYPE orthologs, although in one case (a protein from *Pasteurella* annotated as PfhB2) a putative DNA-binding domain is present as part of the HYPE protein itself.

Also Phd, or more specifically the DNA-binding domain of Phd occurs outside the TA context. This is not surprising as the DNA-binding domains of other TA antitoxin families constitute common prokaryotic transcription repressor domains such as the helix-turn-helix, ribbon-helix-helix, and AbrB-type domains. Thus, domains potentially homologous to Phd have been identified in larger DNA-binding proteins, particularly in members of the *E. coli* DnaT protein family. DnaT associates with PriA/B and DnaC and in the presence of an intact leading strand, reassemble DnaB helicase and DNA polymerase III on the replication fork (Heller and Marians 2005).

9.4 Crystal Structure of P1 Doc and its Interaction with Phd

The first crystal structure of Doc to become available was that of an inactive mutant (His66Tyr) in complex with a 22 C-terminal peptide of Phd (Garcia-Pino et al. 2008). This structure revealed Doc to be unrelated to all other TA toxins with known structure such as CcdB, RelE, MazF, ParE, and PezT. Doc consists of a globular domain formed by a bundle of six α-helices (Fig. 9.2a). It can be seen as the stacking of three pairs of antiparallel helix-loop-helix building blocks.

Conserved surface residues cluster on a single patch that also harbors the known surface mutations that render Doc inactive (Fig. 9.2a). This conserved surface

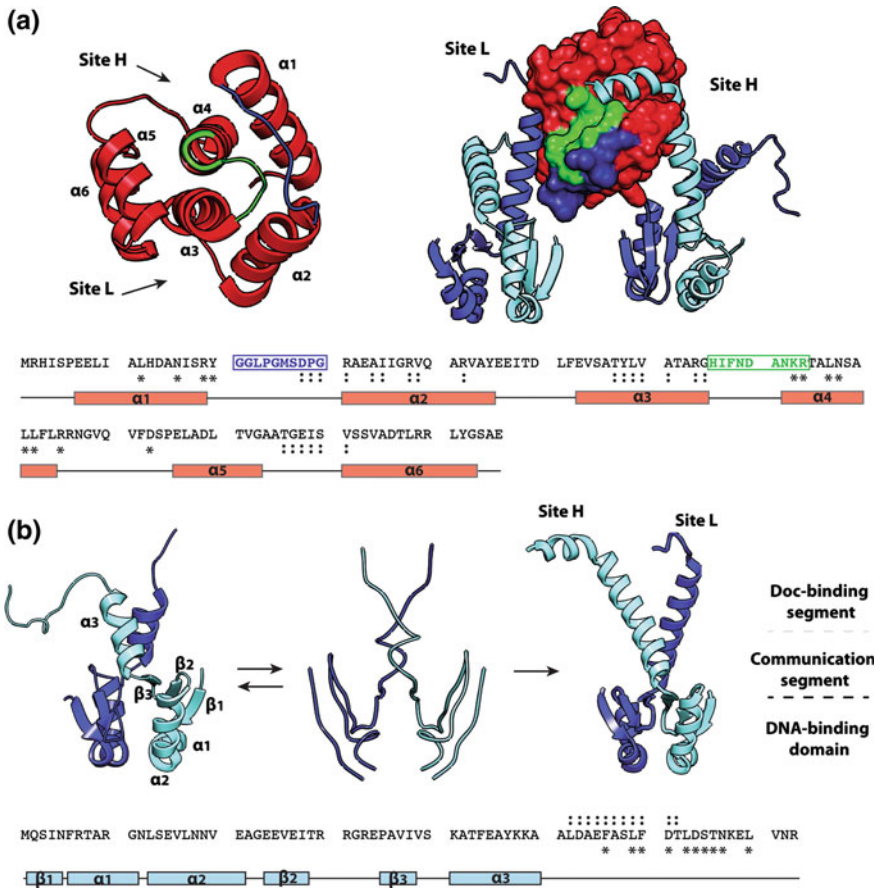


Fig. 9.2 Crystal structures of Phd and Doc**. a** The *left panel* shows the crystal structure of bacteriophage P1 Doc (PDB 3DD7) in a cartoon representation. Doc consists only of α-helices (six in total) shown in *red*. The loop region between helices α1 and α2 (also termed 'flap') is colored *dark blue*, whereas the conserved sequence motif (HXFX[D/E]GNGR) between helices α3 and α4 is colored *green*. The *right panel* displays the crystal structure of the Phd–Doc–Phd complex (PDB 3K33) in which the Doc monomer (*red*) is shown in a surface representation and the two Phd dimers (respective monomers colored *blue and cyan*) are depicted in a cartoon representation. The first Phd dimer is bound to the high-affinity site on Doc (Site H), whereas the second Phd dimer is bound Doc's low-affinity site (Site L). The *green* and *dark blue patches* on Doc's surface represent the α1–α2 loop region and the conserved sequence motif between helices α3 and α4, respectively. These regions are highlighted in the same colors in the amino acid sequence with a match to Doc's secondary structure elements. The residues indicated by an *asterisk* (*) or a *colon* (:) constitute Site H or Site L for Phd binding, respectively. **b** The crystal structure of Phd (PDB 3HRY) reveals that the antitoxin's DNA-binding domain and communication segment reside in an equilibrium between two states as shown in the *left panel*. Upon binding to Doc, the toxin binding segment of Phd becomes fully structured (*right panel*). The Phd amino acid sequence displays the matching secondary structure elements of Phd's DNA-binding domain and communication segment. The residues indicated by an *asterisk* (*) or a *colon* (:) constitute Site H or Site L for Phd binding, respectively

Fig. 9.3 The incomplete Fic-fold of P1 Doc provides evidence for the possible origins of the *Phd–doc* module. **a** The *left, middle, and right panels* show the crystal structures of Bacteriophage P1 Doc–Phd$^{52–73}$ (PDB 3DD7), *Shewanella oneidensis* Fic (PDB 3EQX), and *Neisseria meningitidis* Fic (PDB 2F6S), respectively. Depending on the way the inhibitory α-helix (α_{inh} with consensus sequence [S/T]XXXE[G/N]) in the Fic-fold (colored in *cyan* in all structures) is provided, the Fic-domain proteins are classified in three groups: class I (α_{inh} is provided by a complementary antitoxin, class II (α_{inh} is provided by the protein's N-terminal moiety), or class III (α_{inh} is provided by the protein's C-terminal moiety). The conserved Fic active site motif (HXFX[D/E]GNGRXXR) and flap region are, respectively, shown in *green and dark blue* in all structures. A sequence alignment of all three proteins (generated using SALIGN (Marti-Renom et al. 2004)) indeed show that the flap and active site regions (indicated in *green* and *blue*, respectively) coincide from a structural point-of-view. Furthermore, the sequence alignment reveals that α_{inh} (highlighted in *cyan*) are provided by the protein's N-terminal part, the protein's C-terminal part, and by a complementary antitoxin for *S. oneidensis* Fic, *N. meningitidis* Fic, and Bacteriophage P1 Doc, respectively. **b** Schematic representation of the possible evolutionary origin of the *Phd–doc* module. During the course of evolution, the α_{inh}-segment of an ancestral Fic domain is fused with a DNA-binding domain by a gene shuffling event resulting in a α_{inh}-less toxic Fic-domain protein (*1*). Hereby, a toxin–antitoxin module is created in which the antitoxin alleviates toxin action through fold complementation when the antitoxin binds the toxin (*2*) ▶

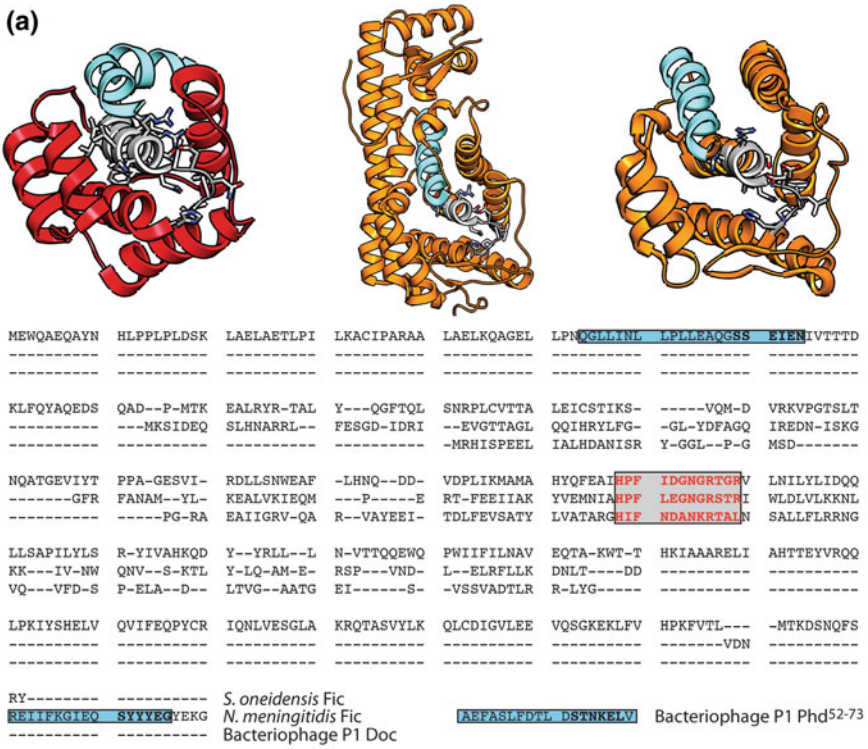
essentially centers on the loop between helices α3 and α4 (His66-Asn72) extends into the beginning of helix α4. Several mutations, such as H66Y, D70N, and A76E that specifically disrupt toxicity, but preserve other, regulatory activities of Doc (Magnuson and Yarmolinsky 1998) map to this region.

The structure of wild-type Doc, obtained in 2010 via co-expression and co-crystallization with full length Phd, is identical to that of the His66Tyr mutant showing that inactivation of Doc through this mutation is not due to global or local misfolding (Garcia-Pino et al. 2010). This complex further shows two distinct and different ways in which Doc interacts with two Phd antitoxin dimers, simultaneously (Fig. 9.2a). In the higher affinity interaction, involves a deep groove created by the relative arrangements of helices α1, α4, and α5. The C-terminal 22 amino acids of Phd fill this groove while adopting a kinked α-helical structure.

The second, lower affinity binding site for the C-terminal domain of Phd is formed by a shallow groove involving helices α2 and α3 (Garcia-Pino et al. 2010). Here, Phd remains unstructured from residue 65 onwards, and the corresponding interaction surface is considerably smaller than for the high-affinity site (Fig. 9.2a).

The conserved and functionally relevant region of Doc formed by loops α1–α2 and α3–α4, which harbor most known Doc-inactivating mutations, is sandwiched between the two Phd binding areas, but is not physically blocked by either Phd molecules (Fig. 9.2a).

Despite this wealth of structural information, a structure of Doc in absence of any Phd fragment or bound to the ribosome nevertheless remains missing. The latter would be of interest to verify the presence or absence of changes in conformational dynamics in the protein upon binding of Phd, which could be relevant for its mechanism of action. Similarly, a complex of Doc with its target, the 30 S ribosomal subunit, is also eagerly awaited.


(a)
MEWQAEQAYN HLPPLPLDSK LAELAETLPI LKACIPARAA LAELKQAGEL LPNQGLLINL LPLLEAQGSS EIENIVTTTD
KLFQYAQEDS QAD--P-MTK EALRYR-TAL Y---QGFTQL SNRPLCVTTA LEICSTIKS- -----VQM-D VRKVPGTSLT
---MKSIDEQ SLHNARRL-- FESGD-IDRI --EVGTTAGL QQIHRYLFG- -GL-YDFAGQ IREDN-ISKG
-MRHISPEEL IALHDANISR Y-GGL--P-G MSD-------
NQATGEVIYT PPA-GESVI- RDLLSNWEAF -LHNQ--DD- VDPLIKMAMA HYQFEAIHPF IDGNGRTGRV LNILYLIDQQ
-------GFR FANAM--YL- KEALVKIEQM ---P-----E RT-FEEIIAK YVEMNIAHPF LEGNGRSTRI WLDLVLKKNL
----PG-RA EAIIGRV-QA R--VAYEEI- TDLFEVSATY LVATARGHIF NDANKRTALN SALLFLRRNG
LLSAPILYLS R-YIVAHKQD Y--YRLL--L N-VTTQQEWQ PWIIFILNAV EQTA-KWT-T HKIAAARELI AHTTEYVRQQ
KK---IV-NW QNV--S-KTL Y-LQ-AM-E- RSP---VND- L--ELRFLLK DNLT----DD
VQ---VFD-S P-ELA--D-- LTVG--AATG EI------S- -VSSVADTLR R-LYG-----
LPKIYSHELV QVIFEQPYCR IQNLVESGLA KRQTASVYLK QLCDIGVLEE VQSGKEKLFV HPKFVTL--- -MTKDSNQFS
-------VDN
RY-------- S. oneidensis Fic
REIIFKGIEQ SYYYEGYEKG N. meningitidis Fic
Bacteriophage P1 Doc
AEFASLFDTL DSTNKELV Bacteriophage P1 Phd52-73
(b)
(1)
(2)
(3)
Fic-domain
ainh
Fic-domain protein
DBD-domain
ainh
Fic-domain
TA module
Fic-domain
ainh
Fic-domain protein

9.5 A possible Origin for the *phd/doc* Modules

As predicted by Anantharaman and Aravind (2003), the crystal structure of Doc determined in 2008 shows a strong resemblance to a series of protein domains present in both prokaryotes and eukaryotes: the "cAMP-induced filamentation" (Fic) domain family (Fig. 9.3a). The Doc and Fic families show very little sequence identity, but a possible evolutionary relationship was nevertheless suggested based upon close similarities between both folds and the presence of a conserved sequence motif that contains residues crucial for the activities of both protein families.

Among the Fic-domain structures present in the Protein Data Bank, the Fic homolog from *Neisseria meningitidis* (Fic_Nm— PDB entry 2G03) is structurally most similar to P1 Doc. Each α-helix of Doc has an equivalent in Fic_Nm and their topologies are also identical. Nevertheless, Fic_Nm has an additional α-helix at its C-terminus that is conspicuously absent in P1 Doc. Superposition of Fic_Nm on the Phd–doc complex shows that the C-terminal 22 amino acids of Phd bound to the high-affinity site of Doc superimpose on the C-terminal α-helix of Fic_Nm. Thus, a protein resembling a full length Fic can be reconstituted by connecting the Phd-derived helix to the C-terminus of Doc (Fig. 9.3a).

Other Fic proteins such as the one from *Shewanella oneidensi* (Fic_So—PDB entry 3EQX—Das et al. 2009) are larger and contain additional α-helices not present in Doc. Fic_So represents the second class of Fic proteins where the Phd-mimicking helix is located at the N-terminus rather than at the C-terminus of the domain.

Finally, Fic-related proteins have been identified that are inhibited by a helical segment that is provided *in trans* by a partner protein. While reminiscent of TA modules, these Fic family members are distinct from the Doc family in that their active site and specificity loops much closer resemble other Fic proteins than Doc and that they, like other Fic's, possess adenylation activity that is presumed absent in Doc. Thus in total, three related ways of inhibiting Fic proteins have been identified, which led to a classification of the Fic family (Engel et al. 2012). The phd/doc module then can be seen as a subgroup of the type I Fic regulatory class.

Together, these data suggest a possible origin for the *Phd–doc* family (Garcia-Pino et al. 2008, Fig. 9.3b). We propose that a Fic-like ancestor became associated with a common DNA-binding domain through illegitimate recombination. During the same or subsequent event, the C-terminal α-helix of the Fic-like ancestor was transferred to the C-terminus of the DNA-binding protein to create a *phd–doc* module. Additional evidence for such a scenario can be found in the Phd–Doc interface, which is quite hydrophobic. P1 Doc in absence of Phd is not too soluble and the high-affinity binding site for Phd exposes a significant amount of hydrophobic surface. This feature might be a relic of the disruption of the hydrophobic core of the Fic-like ancestor when its C-terminal α-helix was removed.

While the above-described scenario for the origin of *phd/doc* described above is plausible, direct evidence for similar scenarios for other TA modules is not as

clear-cut. A circumstantial argument in favor of the helix-transfer hypothesis is that next to the regular two-component TA modules, one can find two other types of TA modules. The first ones are three-component systems where the DNA-binding domains and toxin-neutralizing domains are present as separate proteins. Such situations have been encountered in several families of TA modules. The best-known case here is the ω-ε-ζ module on the low copy number plasmid pSM19035 from *Streptococcus pyogenes* (Zielenkiewicz and Ceglowski 2005), which is related to the *pesAT* modules. Another well-documented case is the *paaR-paaA-parE* modules on the chromosome of *E. coli* O157:H7 and which are related to the *parDE* modules (Hallez et al. 2010). Interestingly, there are also a few cases of TA modules where the DNA-binding domain is completely absent. The most clear-cut case here is formed by the *mazEF* modules found in *Staphylococcus aureus* and closely related species (Fu et al. 2007). The toxin is a clear member of the MazF family while the antitoxin is unusually short, but bears some (weak) sequence resemblance to the C-terminal half of regular MazE proteins. Thus, it seems that here the MazE protein either lost its DNA-binding module or alternatively never gained one. As a consequence, the *S. aureus mazEF* is not autoregulated and transcription is controlled via alternative sigma factor σ^B (Donegan and Cheung 2009). The atypical *relBE* module from the hyperthermophilic archaeon *Pyrococcus horikoshii* OT3 may be another example (Takagi et al. 2005).

The second type of non-classic TA modules that support complex recombination events and the exchange of secondary structure elements between precursor proteins as the origin of TA modules are "fused" TA modules (Leplae et al., 2011). These are genes that encode a single protein that contains an N-terminal DNA-binding domain and a C-terminal toxin domain.

9.6 Molecular Flexibility and the DNA-Binding Properties of Phd

As is generally observed in TA modules, the *phd/doc* operon is autoregulated with Phd acting as a repressor and Doc modulating this activity without itself interacting with the operator DNA. In the absence of Doc, Phd binds and protects a palindromic site in the promoter region and, at higher concentrations, the second and adjacent, imperfect palindromic site (Magnuson et al. 1996; Gazit and Sauer 1999a). In the presence of Doc, a similar pattern of protection is observed, but the sites are filled and protected simultaneously, rather than consecutively. EMSA experiments, utilizing the nontoxic DocH66Y variant, showed that subsaturating concentrations of DocH66Y, relative to Phd, dramatically enhance the affinity of Phd for a double site DNA fragment, but not for single site DNA. Stoichiometry experiments indicated that a single molecule of Doc is sufficient to increase the binding, the cooperativity, and the half-life of the double site complex. This led to the hypothesis that a single molecule of Doc might contact or "bridge" two

molecules of Phd, a model which is supported by the crystal structure of the Phd–Doc complex (Garcia-Pino et al. 2010). EMSA experiments further indicated that one or two additional molecules of Doc could further enhance DNA binding but that saturating concentrations of Doc relative to Phd lead to partial derepression in vivo, to a loss of specificity for the double site relative to the single site, and to supershifted (Doc–Phd) complexes on single site DNA. This phenomenon, which is also observed in a number of other TA modules, is known as "transcription regulation by conditional co-operativity".

Like most TA antitoxins, Phd is a small dimeric protein consisting of an N-terminal DNA-binding domain and a C-terminal toxin-binding domain (Garcia-Pino et al. 2010, Fig. 9.2b). The N-terminal domain consists of a three-stranded antiparallel β-sheet flanked by two short α-helices. The α-helices harbor the DNA-binding site as evidenced by mutational studies. No structure of a Phd–DNA complex is available yet, and therefore the details of this interaction including the basis of operator specificity remain unknown. Dimerization occurs essentially through sandwiching of the β-sheets, which also creates most of the hydrophobic core of the protein. Thus, dimer formation is coupled to folding.

Phd-like DNA-binding antitoxin domains also occur in association with a number of other TA toxins. Most notable here is YefM, which associates with the RelE-related toxin YoeB and for which data on regulation and structure are available (Kamada and Hanaoka 2005; Kumar et al. 2008; Kedzierska et al. 2007).

The N-terminal domain of Phd was found to be only marginally stable and to be highly flexible (Cherny and Gazit 2004; Garcia-Pino et al. 2010). This flexibility is heavily dependent on ionic strength, with high ionic strengths leading to a more rigid structure. Thus, in its crystalline state, the N-terminal domain of Phd was found to exist in two conformations: a fully folded one where all secondary structure elements are formed and a partially folded one where only the β-sheet is intact and most of the α-helix is either unwound into loop structure or fully disordered (Fig. 9.2b). In this partially ordered state, the presumed DNA-binding site is not formed.

In solution, a well-ordered state is observed by NMR and CD spectroscopy at high ionic strength. While the NMR-derived structure of Phd in solution closely matches the fully ordered state found in the crystals, its thermodynamic stability is only marginal. In addition, the NMR order parameters indicate that while the protein is folded, it remains significantly more flexible than typical globular proteins. The major contributors to this low stability, and likely its high flexibility as well, are the small size of the hydrophobic core that is formed almost entirely though dimerization and the presence of a cavity of significant size in this core.

The C-terminal Doc-binding domain of Phd is intrinsically disordered in solution both as an isolated peptide and in the context of full length Phd. In the crystal structures of Phd in absence of Doc, much of this segment remains disordered (Garcia-Pino et al. 2010). Nevertheless, residues Lys41–Lys49 adopt an α-helical conformation while for residues Ala50–Arg73 different irregular structures are observed that depend on lattice contacts. Upon interacting with Doc, residues Lys49-Arg73 fold into a kinked α-helix when bound to the high-affinity binding

site. When bound to the low-affinity site, residues Lys48–Thr62 of Phd fold into an α-helix, and residues Ser65–Arg73 remain entirely disordered.

The N- and C-terminal domains of Phd are not independent. Binding of Doc to Phd increases its affinity for DNA binding, most likely through stabilization of the N-terminal domain (Garcia-Pino et al. 2010). Residues Lys41–Lys48 of Phd can be regarded as a communication segment that transmits information between both domains (Fig. 9.2b). The main structural features of this segment (a small hydrophobic cage surrounded by a salt-bridge network) are conserved throughout the Phd/YefM superfamily. Mutations that affect this communication segment have severe effects in the thermodynamics of DNA binding, probably due to the impact on the overall stability of the protein.

9.7 Autoregulation by Conditional Cooperativity

An almost general property of TA modules is that transcription regulation of TA operons requires the combined action of both toxin and antitoxin. Only the antitoxin harbors a DNA-binding domain, but the toxin is required for efficient autorepression. For a long time, the role of the toxin in autoregulation was considered to increase the affinity of the antitoxin for its operator DNA. In 1998 however, in a study on the autoregulation of the P1 *phd/doc* operon, Magnuson and Yarmolinsky observed that Doc increases the affinity of PhD for its operator DNA at low Doc-to-Phd ratios (Magnuson and Yarmolinsky 1998). However, for high Doc-to-Phd ratios, the affinity decreased again. This phenomenon was later on observed for a number of other TA modules such as F-plasmid *ccd*, *E. coli relBE* and *Salmonalla enterica vapBC* and was termed "conditional co-operativity" (Afif et al. 2001; Overgaard et al. 2008; Winther and Gerdes 2012) Although its molecular basis remained obscure, conditional cooperativity was considered a crucial feature of TA modules as it allows to restrain the cellular ratios of toxin and antitoxin in such a way that it prevents runaway activation of toxins as a consequence of stochastic fluctuations during normal cell growth. Thus when either protein is present in excess, transcription is activated and will stop at the moment that the correct ratio is again achieved. Phd–Doc complexes were reported to display a very short, subsecond half-life of complex dissociation, indicating a dynamic equilibrium among the different species formed (Gazit and Sauer 1999b). The latter may help toward a fast response of the system toward changing external conditions or fluctuations in the Phd:Doc ratio.

Recently, the molecular basis of conditional cooperativity in the P1 *phd/doc* module was unraveled (Garcia-Pino et al. 2010, Fig. 9.4a). The key property of Phd–doc is that the Doc monomer contains two binding sites for Phd, the affinities of which differ two orders in magnitude. This allows Phd (which is a dimer) and Doc to form linear chains of the type…-Doc-Phd_2-Doc-Phd_2-… In these complexes, the DNA-binding domains are separated in such a way that two Doc-bridged Phd dimers can simultaneously bind to the two palindromes of the *phd/doc* operator sequence.

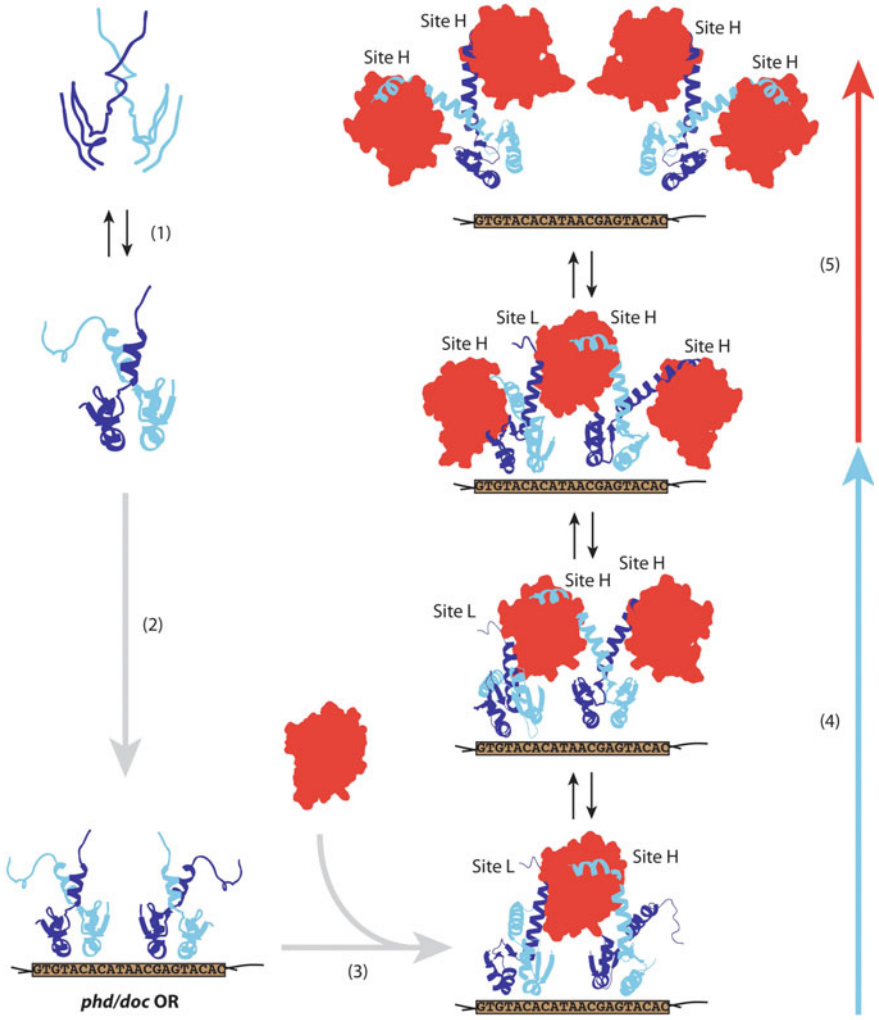


Fig. 9.4 Intrinsic disorder, allostery, and regulation of transcription are coupled within the *phd/doc* module. The antitoxin's DNA-binding domain and communication segment reside in an equilibrium between two states (*1*). Only the folded state can bind DNA (2): one Phd dimer may bind the first (GTGTACAC) or the second (GAGTACAC) binding site in the operator region. However, under these conditions, the binding of Phd to the DNA is weak and the operon remains unrepressed. Upon Doc binding, the Doc-binding segments of both Phd dimers become folded and a Phd–Doc–Phd complex exists in which one Phd binds Doc at Site H and the other Phd binds Doc at Site L (*3*). Doc forms a physical bridge between two Phd dimers which increases the affinity of Phd for its operator DNA which in turn results in repression of the operon. Two Doc-binding segments remain available for Doc binding on each Phd dimer in this configuration. Further binding of Doc results in higher order Phd–Doc complexes which further tighten the Phd–DNA interaction leading to a full stop in transcription (*4*). At increasing Doc concentrations, Doc–Phd–Doc complexes are formed which have a significantly lower affinity for the *phd/doc* operator region: the result is that the Doc–Phd–Doc complexes dissociate from the DNA which leads to derepression of the operon (*5*)

Thus, the *phd/doc* operon is fully repressed by a Phd_2-Doc-Phd_2 complex that covers both palindromes and that employs both the high and the low-affinity binding sites of Doc. The increase in repression compared to the effects of free Phd stems from a combination of the avidity affect caused by bridging two Phd dimers and the stabilizing effect of Doc on the DNA-binding domain of Phd.

The Phd_2-Doc-Phd_2 complex (Fig. 9.2a) contains two Doc-binding domains of Phd which are not involved in any interaction. Therefore, upon increasing the Doc-to-Phd ratio, e.g., because of ClpXP-mediated Phd degradation, these will act as a buffer to sequester two further Doc molecules without significant change in repression and interactions will occur via the high-affinity site of Doc. Upon still further increasing relative amounts of Doc, no free Phd is available to neutralize Doc. But since one of the interactions in the operator complex involves the low-affinity site of Doc, this interaction can be exchanged for a high-affinity interaction. This exchange allows neutralizing one further molecule of Doc, but disrupts the operator complex. The resulting Doc-Phd_2-Doc does not allow for efficient repression, and fresh Phd can be produced to restore normal steady state without activating Doc.

The structure of the non-repressing Doc-Phd_2-Doc complex has been determined both via SAXS (Garcia-Pino et al. 2010) and via single crystal X-ray diffraction (Arbing et al. 2010, Fig. 9.4b). Both models are very similar and show a complex that is highly rigid: the two Doc monomers adopt a fixed position and orientation around the Phd dimer, which is fully folded. This rigidity of the Doc-Phd_2-Doc complex prevents that two such complexes can position themselves next to each other on the *phd/doc* operator sites. Derepression at high toxin to antitoxin ratios is, thus, based on a steric exclusion mechanism.

The conditional cooperativity mechanisms thus allows for stochastic fluctuations in the Phd:Doc ratio without risking uncontrolled activation of Doc. The detailed molecular mechanisms for conditional cooperativity are likely to different for different types of TA modules. However, the low-to-high-affinity switch is probably a crucial conserved feature, as is suggested by recent studied on the *ccd* module on plasmid F (De Jonge et al. 2009). Here, an extra layer of regulation is added via a rejuvenation mechanism that kicks in after toxin activation. Whether a related level of regulation exists for *phd/doc* is currently unknown.

9.8 Mechanism of Action of Doc

Like the apparent majority of TA toxins, P1 Doc inhibits translation (Liu et al. 2008; Garcia-Pino et al. 2008). Upon induction of ectopic Doc expression in *E. coli*, leucine incorporation is rapidly halted, indicating a halt of translation. Incorporation of uracil and thymine is affected less markedly, suggesting that transcription and replication are not directly affected. In agreement with these observations, P1 Doc inhibits protein synthesis in cell-free expression systems. Purified P1 Doc does not inhibit gyrase in cleavage assays and inhibition of transcription and replication in vivo seem to be secondary effects following translation arrest.

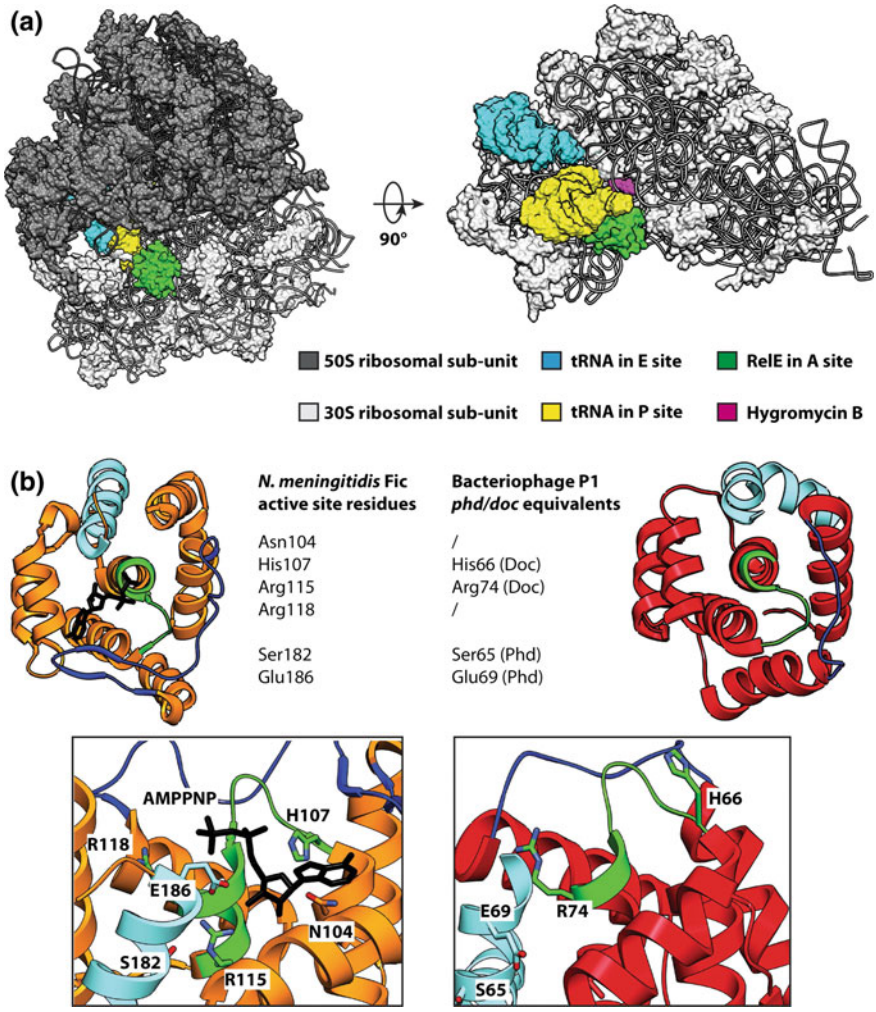


Fig. 9.5 Doc is a ribosome inhibitor. **a** The *left panel* depicts the crystal structure of the RelE-arrested *Thermus thermophilus* 70S ribosome in the post-translational state (PDBs 3KIU, 3KIW, 3KIX, 3KIY). The 50S and 30S subunits are shown in dark and light gray, respectively. The tRNAs in the ribosome's E and P site are shown in *cyan and yellow*, respectively. The ribosome's A site is occupied by the ribosome-dependent mRNase RelE (*green surface)*. A 90° rotation of the ribosome along the x-axis reveals the binding pocket of the ribosome-arresting antibiotic hygromycin B (magenta surface, PDB 1HNZ). **b** The *left panel* displays the crystal structure of *N. meningitidis* Fic in complex with ATP analog AMPPNP (PDB 3S6A). The Fic protein is depicted as before: the inhibitory α-helix is shown in *cyan*, the conserved active site motif in *green*, and the flap region in *dark blue*. The AMPPNP is shown in a *black stick* representation. The *right panel* shows P1 Doc in complex with $Phd^{52–73}$ (PDB 3DD7). Both insets show a close-up of the putative active sites of both proteins. Several well-conserved residues (Asn104, His107, Arg115, Arg118, Ser182, and Glu186 for *N. menigitidis* Fic, and His66, Arg74, Ser65, Glu69 for Doc-$Phd^{52–73}$) are shown in a *stick representation*

Induction of Doc induces translation-dependent mRNA cleavage in *E. coli* strain MG1655 (Garcia-Pino et al. 2008). The observed cleavage pattern is essentially identical to that observed on RelE activation and is not observed in *relE* knockouts. A similar dependence of Doc on *E. coli* MazF remains debated (Hazan et al. 2001). Thus, mRNA cleavage by Doc is a secondary effect and does not involve an enzymatic reaction of Doc itself. Rather, the ribosome inhibition by Doc induces RelE activation. This phenomenon, which is likely a chain reaction resulting from establishment of general stress conditions in the cell upon induction of the first TA module, may be of importance for establishing the persister state. Indeed, it was shown that TA modules on the *E. coli* chromosome have a cumulative effect on persister cell formation (Maisonneuve et al. 2011). Nevertheless, there seems to be some specificity in TA cross-activation, as there is in their responses to individual stresses, two observations for which there is currently no satisfactory model.

In contrast to other TA toxins that interfere with translation, Doc on itself stabilizes mRNA in BL21(DE3) (Liu et al. 2008). This commonly used expression strain lacks *lon*, and therefore Doc activation probably fails to induce RelE mRNA cleavage activity, revealing the direct action of Doc. Northern blot analysis of four distinct transcripts revealed that their mRNA levels remain constant up to 120 min after induction of Doc. When RNA polymerase activity is blocked with rifampicin, the half-life of mRNAs is significantly enlarged, with the most pronounced effects on short-lived mRNA such as *tufA* (half-life increases from 4 to 48 min). Ribosomal profiles from cells overexpressing Doc show that Doc induction leads to stabilization of polysomes just as in cells treated with hygromycin B. Interestingly, a hygromycin B-resistant mutant was reported to be non-sensitive to Doc indicating not only a similar mechanism of action but also a common binding site on the ribosome (Liu et al. 2008). It, therefore, seems that Doc binds to the ribosome at a position overlapping with the hygromycin B binding site and distinct from the known RelE binding site (Fig. 9.5a, Neubauer et al. 2009).

An interesting and yet unexplored viewing angle regarding the activity of Doc at the molecular level is its close resemblance to Fic (Fig. 9.5b) (Garcia-Pino et al. 2008). Besides their general structural similarities, Doc and Fic family members share a single catalytic motif (HXFX(D/E)(A/G)N(K/G)R) (in the loop between helices α3 and α4) that is highly conserved in both families. In Fic, this loop is an active site loop harboring several catalytic residues essential for its AMPylation activity (Kinch et al. 2009; Luong et al. 2010; Worby et al. 2009; Xiao et al. 2010; Yarbrough et al. 2009). Several known inactivating mutations of Doc cluster in this loop and involve residues which in Fic have a catalytic function (e. g. His66 and Asp70). This not only shows that this loop is crucial for Doc activity but also suggests the possibility that Doc possesses a yet unknown but crucial enzymatic function. Further inspection of the Fic ATP binding site shows that in the case of doc, binding of ATP is unlikely due to steric clashes with the β- and γ-phosphate moieties. Thus, any putative enzymatic activity of Doc is likely to be distinct from that of Fic, although probably related.

References

Afif, H., Allali, N., Couturier, M., & Van Melderen, L. (2001). The ratio between CcdA and CcdB modulates the transcriptional repression of the ccd poison-antidote system. *Molecular Microbiology, 41*, 73–82.

Anantharaman, V., & Aravind, L. (2003). New connections in the prokaryotic toxin-antitoxin network: relationship with the eukaryotic nonsense-mediated RNA decay system. *Genome Biology, 4*, R81.

Arbing, M. A., Handelman, S. K., Kuzin, A. P., Verdon, G., Wang, C., Su, M., et al. (2010). Crystal structures of Phd-Doc, HigA, and YeeU establish multiple evolutionary links between microbial growth-regulating toxin-antitoxin systems. *Structure, 18*, 996–1010.

Cherny, I., & Gazit, E. (2004). The YefM antitoxin defines a family of natively unfolded proteins: implications as a novel antibacterial target. *Journal of Biological Chemistry, 279*, 8252–8261.

Das, D., Krishna, S. S., McMullan, D., Miller, M. D., Xu, Q., Abdubek, P., et al. (2009). Crystal structure of the Fic (Filamentation induced by cAMP) family protein SO4266 (gil24375750) from S*hewanella oneidensis* MR-1 at 1.6 A resolution. *Proteins, 75*, 264–271.

De Jonge, N., Garcia-Pino, A., Buts, L., Haesaerts, S., Charlier, D., Zangger, K., et al. (2009). Rejuvenation of CcdB-poisoned gyrase by an intrinsically disordered protein domain. *Molecular Cell, 35*, 154–163.

Donegan, N. P., & Cheung, A. L. (2009). Regulation of the *mazEF* toxin-antitoxin module in *Staphylococcus aureus* and its impact on *sigB* expression. *Journal of Bacteriology, 191*, 2795–2805.

Engel, P., Goepfert, A., Stanger, F. V., Harms, A., Schmidt, A., Schirmer, T., et al. (2012). Adenylylation control by intra- or intermolecular active-site obstruction in Fic proteins. *Nature, 482*, 107–110.

Fico, S., & Mahillon, J. (2006). *TasA-tasB*, a new putative toxin-antitoxin (TA) system from *Bacillus thuringiensis* pGI1 plasmid is a widely distributed composite *mazE-doc* TA system. *BMC Genomics, 7*, 259.

Fu, Z., Donegan, N. P., Memmi, G., & Cheung, A. L. (2007). Characterization of MazFSa, an endoribonuclease from *Staphylococcus aureus*. *Journal of Bacteriology, 189*, 8871–8879.

Garcia-Pino, A., Christensen-Dalsgaard, M., Wyns, L., Yarmolinsky, M., Magnuson, R. D., Gerdes, K., et al. (2008). Doc of prophage P1 is inhibited by its antitoxin partner Phd through fold complementation. *Journal of Biological Chemistry, 283*, 30821–30827.

Garcia-Pino, A., Balasubramanian, S., Wyns, L., Gazit, E., De Greve, H., Magnuson, R. D., et al. (2010). Allostery and intrinsic disorder mediate transcription regulation by conditional cooperativity. *Cell, 142*, 101–111.

Gazit, E., & Sauer, R. T. (1999a). Stability and DNA binding of the Phd protein of the phage P1 plasmid addiction system. *Journal of Biological Chemistry, 274*, 2652–2657.

Gazit, E., & Sauer, R. T. (1999b). The Doc toxin and Phd antidote proteins of the bacteriophage P1 plasmid addiction system form a heterotrimeric complex. *Journal of Biological Chemistry, 274*, 16813–16818.

Grady, R., & Hayes, F. (2003). Axe-Txe, a broad-spectrum proteic toxin-antitoxin system specified by a multidrug-resistant, clinical isolate of *Enterococcus faecium*. *Molecular Microbiology, 47*, 1419–1432.

Hallez, R., Geeraerts, D., Sterckx, Y., Mine, N., Loris, R., & Van Melderen, L. (2010). New bactericidal toxins homologous to ParE belonging to 3-component toxin-antitoxin systems in *Escherichia coli* O157:H7. *Molecular Microbiology, 76*, 719–732.

Hazan, R., Sat, B., Reches, M., & Engelberg-Kulka, H. (2001). Postsegregational killing mediated by the P1 phage "addiction module" *phd-doc* requires the Escherichia coli programmed cell death system *mazEF*. *Journal of Bacteriology, 183*, 2046–2050.

Heller, R. C., & Marians, K. J. (2005). The disposition of nascent strands at stalled replication forks dictates the pathway of replisome loading during restart. *Molecular Cell, 17*, 733–743.

Hiraga, S., Jaffé, A., Ogura, T., Mori, H., & Takahashi, H. (1986). F plasmid *ccd* mechanism in *Escherichia coli*. *Journal of Bacteriology, 166*, 100–104.

Kamada, K., & Hanaoka, F. (2005). Conformational change in the catalytic site of the ribonuclease YoeB toxin by YefM antitoxin. *Molecular Cell, 19*, 497–509.

Kawamukai, M., Matsuda, H., Fujii, W., Nishida, T., Izumoto, Y., Himeno, M., et al. (1988). Cloning of the *fic-1* gene involved in cell filamentation induced by cyclic AMP and construction of a delta *fic Escherichia coli* strain. *Journal of Bacteriology, 170*, 3864–3869.

Kedzierska, B., Lian, L. Y., & Hayes, F. (2007). Toxin-antitoxin regulation: bimodal interaction of YefM-YoeB with paired DNA palindromes exerts transcriptional autorepression. *Nucleic Acids Research, 35*, 325–339.

Kinch, L. N., Yarbrough, M. L., Orth, K., & Grishin, N. V. (2009). Fido, a novel AMPylation domain common to Fic, Doc, and AvrB. *PLoS ONE, 4*, e5818.

Komano, T., Utsumi, R., & Kawamukai, M. (1991). Functional analysis of the *fic* gene involved in regulation of cell division. *Research in Microbiology, 142*, 269–277.

Kumar, P., Issac, B., Dodson, E. J., Turkenburg, J. P., & Mande, S. C. (2008). Crystal structure of *Mycobacterium tuberculosis* YefM antitoxin reveals that it is not an intrinsically unstructured protein. *Journal of Molecular Biology, 383*, 482–493.

Lehnherr, H., Maguin, E., Jafri, S., & Yarmolinsky, M. B. (1993). Plasmid addiction genes of bacteriophage P1: *doc*, which causes cell death on curing of prophage, and *phd*, which prevents host death when prophage is retained. *Journal of Molecular Biology, 233*, 414–428.

Lehnherr, H., & Yarmolinsky, M. B. (1995). Addiction protein Phd of plasmid prophage P1 is a substrate of the ClpXP serine protease of *Escherichia coli*. *Proceedings of National Academic Science USA, 92*, 3274–3277.

Leplae, R., Geeraerts, D., Hallez, R., Guglielmini, J., Drèze, P., & Van Melderen, L. (2011). Diversity of bacterial type II toxin-antitoxin systems: a comprehensive search and functional analysis of novel families. *Nucleic Acids Research, 39*, 5513–5525.

Liu, M., Zhang, Y., Inouye, M., & Woychik, N. A. (2008). Bacterial addiction module toxin Doc inhibits translation elongation through its association with the 30S ribosomal subunit. *Proceedings of National Academic Science USA, 105*, 5885–5890.

Lobocka, M. B., Rose, D. J., Plunkett, G, 3rd, Rusin, M., Samojedny, A., Lehnherr, H., et al. (2004). Genome of bacteriophage P1. *Journal of Bacteriology, 186*, 7032–7068.

Luong, P., Kinch, L. N., Brautigam, C. A., Grishin, N. V., Tomchick, D. R., & Orth, K. (2010). Kinetic and structural insights into the mechanism of AMPylation by VopS Fic domain. *Journal of Biological Chemistry, 285*, 20155–20163.

Magnuson, R., Lehnherr, H., Mukhopadhyay, G., & Yarmolinsky, M. B. (1996). Autoregulation of the plasmid addiction operon of bacteriophage P1. *Journal of Biological Chemistry, 271*, 18705–18710.

Magnuson, R., & Yarmolinsky, M. B. (1998). Corepression of the P1 addiction operon by Phd and Doc. *Journal of Bacteriology, 180*, 6342–6351.

Magnuson, R. D. (2007). Hypothetical functions of toxin-antitoxin systems. *Journal of Bacteriology, 189*, 6089–6092.

Maisonneuve, E., Shakespeare, L. J., Jørgensen, M. G., & Gerdes, K. (2011). Bacterial persistence by RNA endonucleases. *Proceedings of the National Academy of Science* USA, *108*, 13206-13211.

Marti-Renom, M. A., Madhusudhan, M. S., & Sali, A. (2004). Alignment of protein sequences by their profiles. *Protein Science, 13*, 1071–1087.

McKinley, J. E., & Magnuson, R. D. (2005). Characterization of the Phd repressor-antitoxin boundary. *Journal of Bacteriology, 187*, 765–770.

Neubauer, C., Gao, Y. G., Andersen, K. R., Dunham, C. M., Kelley, A. C., Hentschel, J., et al. (2009). The structural basis for mRNA recognition and cleavage by the ribosome-dependent endonuclease RelE. *Cell, 139*, 1084–1095.

Overgaard, M., Borch, J., Jørgensen, M. G., & Gerdes, K. (2008). Messenger RNA interferase RelE controls *relBE* transcription by conditional co-operativity. *Molecular Microbiology, 69*, 841–857.

Pandey, D. P., & Gerdes, K. (2005). Toxin–antitoxin loci are highly abundant in free-living but lost from host-associated prokaryotes. *Nucleic Acids Research, 33*, 966–976.

Pomerantsev, A. P., Golovliov, I. R., Ohara, Y., Mokrievich, A. N., Obuchi, M., Norqvist, A., et al. (2001). Genetic organization of the *Francisella* plasmid pFNL10. *Plasmid, 46*, 210–222.

Roberts, R. C., Ström, A. R., & Helinski, D. R. (1994). The *parDE* operon of the broad-host-range plasmid RK2 specifies growth inhibition associated with plasmid loss. *Journal of Molecular Biology, 237*, 35–51.

Ruiz-Echevarria, M. J., Berzal-Herranz, A., Gerdes, K., & Díaz-Orejas, R. (1991a). The *kis* and *kid* genes of the parD maintenance system of plasmid R1 form an operon that is autoregulated at the level of transcription by the co-ordinated action of the Kis and Kid proteins. *Molecular Microbiology, 5*, 2685–2693.

Ruiz-Echevarria, M. J., de Torrontegui, G., Giménez-Gallego, G., & Díaz-Orejas, R. (1991b). Structural and functional comparison between the stability systems ParD of plasmid R1 and Ccd of plasmid F. *Molecular and General Genetics, 225*, 355–362.

Smith, J. A., & Magnuson, R. D. (2004). Modular organization of the Phd repressor/antitoxin protein. *Journal of Bacteriology, 186*, 2692–2698.

Takagi, H., Kakuta, Y., Okada, T., Yao, M., Tanaka, I., & Kimura, M. (2005). Crystal structure of archaeal toxin-antitoxin RelE-RelB complex with implications for toxin activity and antitoxin effects. *Nature Structural and Molecular Biology, 12*, 327–331.

Tyndall, C., Lehnherr, H., Sandmeier, U., Kulik, E., & Bickle, T. A. (1997). The type IC *hsd* loci of the enterobacteria are flanked by DNA with high homology to the phage P1 genome: implications for the evolution and spread of DNA restriction systems. *Molecular Microbiology, 23*, 729–736.

Utsumi, R., Nakamoto, Y., Kawamukai, M., Himeno, M., & Komano, T. (1982). Involvement of cyclic AMP and its receptor protein in filamentation of an *Escherichia coli fic* mutant. *Journal of Bacteriology, 151*, 807–812.

Winther, K. S., & Gerdes, K. (2012). Regulation of enteric *vapBC* transcription: induction by VapC toxin dimer-breaking. *Nucleic Acids Research, 40*, 4347–4357.

Worby, C. A., Mattoo, S., Kruger, R. P., Corbeil, L. B., Koller, A., Mendez, J. C., et al. (2009). The Fic domain: regulation of cell signaling by adenylylation. *Molecular Cell, 34*, 93–103.

Xiao, J., Worby, C. A., Mattoo, S., Sankaran, B., & Dixon, J. E. (2010). Structural basis of Fic-mediated adenylylation. *Nature Structural and Molecular Biology, 17*, 1004–1010.

Yarbrough, M. L., Li, Y., Kinch, L. N., Grishin, N. V., Ball, H. L., & Orth, K. (2009). AMPylation of Rho GTPases by *Vibrio* VopS disrupts effector binding and downstream signaling. *Science, 323*, 269–272.

Zhao, X., & Magnuson, R. D. (2005). Percolation of the Phd repressor-operator interface. *Journal of Bacteriology, 187*, 1901–1912.

Zielenkiewicz, U., & Ceglowski, P. (2005). The toxin-antitoxin system of the streptococcal plasmid pSM19035. *Journal of Bacteriology, 187*, 6094–6105.

Chapter 10
Type II Toxin-Antitoxin Loci: The *fic* Family

Arnaud Goepfert, Alexander Harms, Tilman Schirmer and Christoph Dehio

Abstract FIC domain containing proteins (Fic proteins) are present in all domains of life but particularly widespread among prokaryotes. FIC domains with a fully conserved HxFx[D/E]GNGRxxR active site motif catalyze adenylylation (also known as AMPylation), the transfer of an adenosine 5′-monophosphate moiety onto target proteins. Adenylylation activity is tightly controlled by an inhibitory α-helix (α_{inh}) that can either be part of the Fic protein (intramolecular inhibition) or encoded on a different polypeptide chain (intermolecular inhibition), the latter constituting a novel class of type II toxin-antitoxin (TA) modules represented by VbhT-VbhA of *Bartonella schoenbuchensis* and FicT-FicA of *Escherichia coli*. The helix α_{inh} harbors a [S/T]xxxE[G/N] motif with the conserved glutamate partially obstructing the ATP-binding site and forcing ATP to bind in a catalytically incompetent conformation. Release of inhibition by removal of the antitoxin component or by mutation of the conserved glutamate in α_{inh} converts Fic proteins into toxins that severely impair bacterial growth.

10.1 Introduction

Fic proteins defined by carrying the conserved FIC domain are spread over all domains of life and are particularly abundant among prokaryotes, where several thousand of them have been identified in genome sequencing projects (PFAM

A. Goepfert · A. Harms · C. Dehio (✉)
Focal Area Infection Biology, Biozentrum, University of Basel,
Klingelbergstrasse 70, CH-4056 Basel, Switzerland
e-mail: christoph.dehio@unibas.ch

A. Goepfert · T. Schirmer
Focal Area Structural Biology and Biophysics, Biozentrum, University of Basel,
Klingelbergstrasse 70, CH-4056 Basel, Switzerland

K. Gerdes (ed.), *Prokaryotic Toxin-Antitoxins*,
DOI: 10.1007/978-3-642-33253-1_10,

pf02661). Only few of the bacterial Fic proteins present in the databases have known 3-D structures (Das et al. 2009; Luong et al. 2010; Xiao et al. 2010; Palanivelu et al. 2011), and very little is known about their biological role with the exception of some representatives that have secondarily evolved into virulence factors for the manipulation of eukaryotic host cells (Yarbrough et al. 2009; Worby et al. 2009).

The term *fic* (for "filamentation induced by cAMP") was first used by Utsumi et al. to describe the phenotype of an *Escherichia coli* mutant (Utsumi et al. 1982). This mutant showed a marked defect in cell division when grown at 43 °C and in the presence of cyclic AMP (cAMP) so that the bacteria appeared as long filaments under the light microscope. The mutant *fic-1* allele of the gene designated *fic* was found to code for a G55R substitution in the corresponding single-domain protein Fic that gave the name to the domain family (Kawamukai et al. 1989). Further experimentation revealed that a Δ*fic* deletion mutant strain was viable and showed no discernible phenotypes such as cell filamentation unless complemented with the *fic-1* allele in trans (Kawamukai et al. 1988), thus suggesting that *fic-1* is a gain-of-function mutation. However, though some potential links to carbon metabolism and cAMP signaling could be drawn, the biochemical and physiological basis of the *fic* phenotype has remained elusive until today.

Two decades after these initial studies, Yarbrough and coworkers reported that VopS, a type III secretion effector protein of *Vibrio parahaemolyticus* with two FIC domains, modifies small Rho family GTPases in the host cell by adenylylation, the covalent transfer of an adenosine 5′-monophosphate moiety (Yarbrough et al. 2009). The molecular mechanism involves a nucleophilic attack of the hydroxyl group of the modifiable target amino acid onto the α-phosphorous of ATP leading to transfer of the AMP moiety. Research in this context identified the consensus FIC signature sequence (HxFx[D/E]GNGRxxR) to form the active site motif of FIC domains with the histidine being essential for adenylylation activity. In addition, it was noted that Fic proteins usually exhibit autoadenylylation in addition to target protein adenylylation, even so the physiological role of this phenomenon remained elusive (Xiao et al. 2010; Kinch et al. 2009).

Although consecutive studies demonstrated activities virtually identical to those of VopS for the closely related FIC domains of host cell-translocated virulence factors of *Histophilus somni* and *Pasteurella multocida* (Mattoo et al. 2011), it is striking that the far majority of prokaryotic Fic proteins lack any obvious connection to virulence and therefore likely have a role in bacterial physiology. Exemplified by the VbhT-VbhA complex of *Bartonella schoenbuchensis*, it was recently demonstrated that the toxicity and adenylylation activity of a prokaryotic Fic protein can be specifically suppressed by complex formation with a cognate antitoxin, thus forming a toxin-antitoxin (TA) complex that shares several features of type II TA systems (Engel et al. 2012). In the following, we will briefly review our current knowledge on the evolution and the structure/function relation of this novel family of FIC domain-related TA modules, as well as the structural and functional conservation of the inhibitory mechanism in other Fic protein classes.

10.2 A Novel Class of FIC Domain-Associated Type II Toxin-Antitoxin Systems

Recently, we found that VbhT, a Fic protein of the ruminant-specific pathogen *Bartonella schoenbuchensis*, strongly inhibits bacterial growth when ectopically expressed in *E. coli*. This toxic effect was fully suppressed upon coexpression with VbhA, a small protein encoded by the *vbhA* gene located immediately upstream and likely transcriptionally/translationally coupled to *vbhT* (Engel et al. 2012). Toxicity of VbhT was apparently linked to adenylylation of *bona fide* target proteins in *E. coli* lysates and was dependent on the integrity of the FIC signature motif.

The VbhT-VbhA pair exhibits most of the features that qualify for a type II TA system (Yamaguchi et al. 2011) as shown by Engel et al. (2012).

- Endogenously expressed, VbhT is toxic to the cell.
- VbhA inhibits VbhT toxicity by forming a tight complex.
- VbhT and VbhA are encoded by an operon structure with an upstream antitoxin gene and a downstream toxin gene with presumable transcriptional and translational coupling.

Only transcriptional autoregulation by cooperative promoter binding, a salient feature of other type II TA systems, has not yet been addressed for the VbhT-VbhA module.

A search for *vbhA* homologs encoded immediately upstream of FIC domain-encoding genes identified approximately 150 genes (Engel et al. 2012). Alignment of these VbhA homologous revealed a consensus [S/TxxxE[G/N] motif with a strictly conserved glutamate (Fig. 10.1) that turned out to play a critical role in the inhibitory mechanism.

FicA from *E. coli* is one of the VbhA homologs and is encoded in front of the *fic* gene. In fact, FicA has recently been reported to physically interact with the *E. coli* Fic protein by tandem-affinity purification (Hu et al. 2009) and it was proposed that, indeed, the FicT-FicA module represents a novel TA system (Makarova et al. 2009). Whereas the presence of the [S/TxxxE[G/N] motif in FicA probably indicates a similar mode of inhibition as for the VbhA antitoxin, the slight degeneration of the consensus FIC motif (Fig. 10.1) leaves uncertain the catalytic activity of the toxin as is the case for other Fic proteins (see Sect. 10.7).

Taken together, we propose that VbhT-VbhA and FicT-FicA are the representative members of a novel class of the type II TA system.

Interestingly, gene pairs homologous to *vbhA-vbhT* and *FicT-FicA* are often genetically linked to mobile genetic elements, though this is not the case for the *E. coli ficA-ficT* and close homologs in the *Enterobacteriaceae* where syntenic loci in the chromosome are indicative of vertical transmission.

It is unknown today how the FIC domain-containing toxins of these TA systems get activated in the bacterial cell, but it is well conceivable that—just as described for other TA systems—VbhA-like antitoxins may be more labile than the Fic toxins or prone to faster elimination by intracellular proteases upon stress conditions or other physiological stimuli.

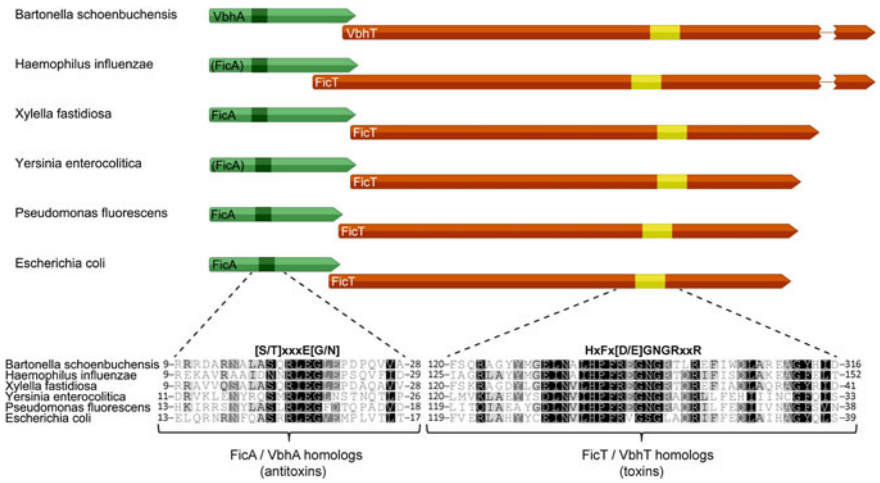


Fig. 10.1 Representatives of the VbhT-VbhA/FicT-FicA toxin-antitoxin family. The genetic organization of six representatives antitoxin-toxin operons (Sequences were taken from *Bartonella schoenbuchensis* R1 (*vbhT* is B11C_100026), *Haemophilus influenza* biotype *aegyptius* plasmid pF1947 (*ficT* is pF1947_03), *Xylella fastidiosa* 9a5c (*ficT* is XF1657), *Yersinia enterocolitica* subsp. *enterocolitica* 8081 (*ficT* is YE1667), *Pseudomonas fluorescens* SBW25 (*ficT* is PFLU0270), and *Escherichia coli* str. K-12 substr. MG1655 (*ficT* is b3661).) is shown at the *top* and the alignment of the signature motifs of VbhA/FicA-homologous antitoxins and VbhT/Fic-homologous toxins at the *bottom*. FicA homologs that were manually annotated are shown in *brackets*. Note that the active site of *Escherichia coli* FicT slightly diverges from the consensus

10.3 Crystal Structure of the VbhT-VbhA TA Module

Figure 10.2a shows the 1.5 Å structure of the FIC domain of VbhT [VbhT(FIC)] in a 1:1 complex with its cognate antitoxin VbhA (Engel et al. 2012). The core of the FIC domain, as defined by Pfam (Punta et al. 2012), is formed by a four-helix bundle (α2–α5). This fold is complemented by the N-terminal helix α1 and two antiparallel helices (α6 and α7) that are arranged perpendicularly to the bundle. Helices α4–α6 form the ATP-binding site. The FIC domain signature motif HxFx[D/E]GNGRxxR is located in the loop that follows helix α4 and extends to the N-cap of helix α5 (Fig. 10.2a). The histidine is indispensible for AMP transferase activity and most probably acts as the general base for the deprotonation of the target side-chain hydroxyl (Luong et al. 2010; Palanivelu et al. 2011). The phenylalanine of the motif is buried in the hydrophobic core of the protein and is important for the folding and the active loop conformation. The remaining residues accommodate the ATP triphosphate moiety. The carboxylic side-chain of the [D/E] residue interacts with a magnesium ion that, in turn, coordinates the α- and β-phosphates of ATP (Fig. 10.3b) (Engel et al. 2012). These two phosphates are

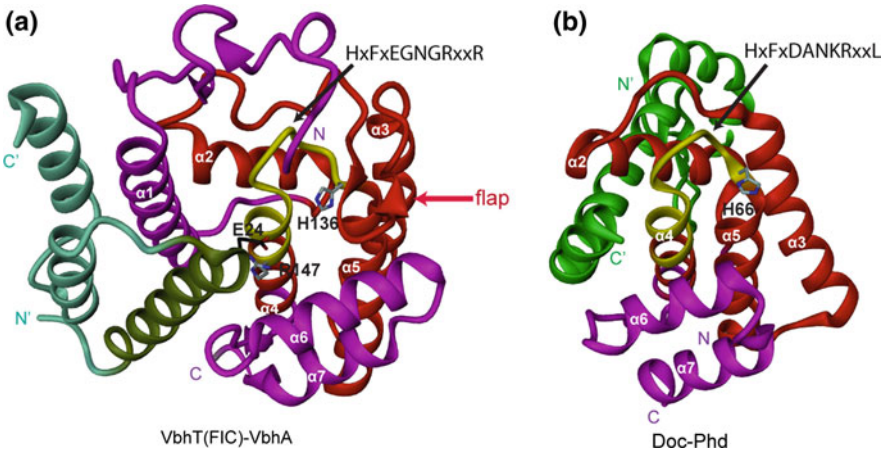


Fig. 10.2 Toxins VbhT (*left*) and Doc (*right*) and their interaction with cognate antitoxins VbhA and Phd, respectively. The toxins exhibit the FIC fold and are shown in *red* (FIC core as defined by PFAM (Punta et al. 2012)) and *magenta*. Their signature motif segments are highlighted in *yellow*. The antitoxins are shown in *green*. **a** Crystal structure of the complex of VbhT(FIC) with antitoxin VbhA (PDB code 3SHG) (Engel et al. 2012) with important residues in full. E24 of the inhibitory helix (α_{inh}, *dark green*) obstructs part of the ATP binding site by forming an intermolecular salt bridge with R147. **b** Crystal structure of the complex of Doc with antitoxin Phd (PDB code 3KH2) (Arbing et al. 2010). The C-terminal helix of Phd is located at a similar position relative to Doc as the N-terminal end of the VbhA helix α_{inh} in relation to VbhT. Thus, the Phd antitoxin may excert its inhibitory role analogously to VbhA, though the function of Doc is still unknown and binding of a substrate to its signature motif is hypothetical

bound to the compound anion binding nest (Watson and Milner-White 2002) at the N-cap of helix α5 with sequence GNGR and form several H-bonds with its main-chain amide groups. The first arginine of the motif interacts with the β-phosphate of ATP, whereas the second arginine serves as a binding site for the γ-phosphate. Furthermore, close to the active site, there is the so-called "flap", a β-hairpin or loop structure that closes onto the adenine ring and creates a suitable target docking site that allows the correct positioning of the modifiable target residue in the FIC active site via sequence-independent main chain–main chain interactions (Palanivelu et al. 2011).

VbhA is a small (62 residues) protein that folds into three anti-parallel helices and wraps around helix α1 of VbhT(FIC), mainly through hydrophobic interactions (Engel et al. 2012). Most relevant, the N-terminal α-helix (α_{inh}), that contains the [S/T]xxxE[G/N] motif is close to the FIC active site (as defined by its signature motif). In particular, residue glutamate E24, which is part of the [S/T]xxxE[G/N] signature motif (see below), forms an intermolecular salt bridge with the second arginine of the FIC motif.

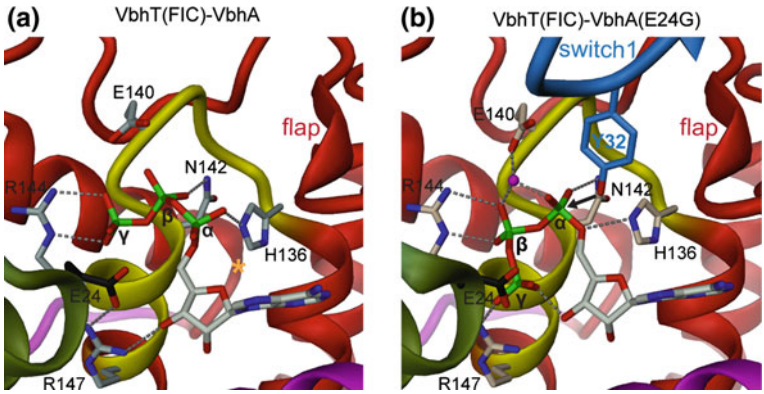


Fig. 10.3 Close-up view of ATP substrate binding to VbhT(FIC) in complex with wild-type VbhA (*left*) and the inhibition-relieved VbhA mutant E24G. Note that the position adopted by E24 of wild-type VbhA (*left*) is taken by the ATP γ-phosphate upon truncation of the glutamate finger (*right*). Furthermore, the location and orientation of the α-phosphate is distinct in the two complexes. In the wild-type situation (*left*), the position that would have to be adopted by a target hydroxyl group (*orange star*) is not accessible. Only in complex with the VbhA mutant (*right*), the α-phosphorous can be attacked in-line with the scissile bond as demonstrated by the superimposed model of the modifiable tyrosine 32 of the cdc42 target protein taken from Xiao et al. (2010). For further details and a movie see Engel et al. (2012)

10.4 VbhA Antitoxin Prevents Productive Binding of ATP to the Catalytic Site of VbhT

How VbhA inhibits the adenylylation activity of the VbhT toxin has been revealed by a detailed crystallographic investigation. It turned out that productive binding of the ATP substrate is impeded by the presence of the conserved glutamate of the antitoxin (E24) which forms an ionic interaction with the second FIC arginine (R147, see Fig. 10.3a).

Only upon truncation of the glutamate (mutant E24G), the ATP γ-phosphate attains the binding subsite close to the second arginine with concomitant rotation of the α-phosphate into an orientation competent for the in-line attack of the target side-chain hydroxyl (Fig. 10.3b). Obviously, the same productive mode of ATP binding is expected for isolated VbhT, explaining its toxicity. Thus, in summary, the antitoxin partly obstructs the nucleotide-binding site of the toxin which impedes competent ATP binding and accordingly inhibits the enzyme's action.

10.5 VbhT-VbhA is Structurally Related to Doc-Phd

The FIC domain fold is very similar to the structure of the well-known Doc toxin of the Doc-PhD type II TA system (see Chap. 9), and both protein families are classified together in the Pfam protein database (Punta et al. 2012) to form the Fic/

Doc family (Pfam: PF02661), also called Fido (Kinch et al. 2009). The Doc-Phd TA locus was discovered as a post-segregational killing system (PSK) that enforces stable maintenance of the P1 phage genome during lysogeny (Lehnherr et al. 1993). In the absence of inhibition by its PhD antitoxin, Doc inhibits translation elongation via an interaction with the ribosome (Liu et al. 2008), but details of the mechanism are not known.

Doc shows the same overall α-helical topology (α2–α7) as Fic proteins, but with helix α1 missing (Fig. 10.2). It has a divergent signature motif (HxFxDANKRxxL) but the histidine that is required for catalysis in Fic proteins is conserved in Doc and, in fact indispensible for its toxicity (Magnuson and Yarmolinsky 1998). Neither adenylylation nor any other enzymatic activity has yet been reported for Doc. However, based on the conservation of their particular active site motif, Doc family proteins are also likely to bind a small molecule substrate and/or catalyze a different enzymatic reaction. Furthermore, the absence of the flap that is essential for the positioning of target protein in the FIC active site and thus critical for adenylylation catalysis also suggests a distinct activity for Doc. Given the similar topology and functional active site location, Fic and Doc are likely to have a common origin and have adopted distinct functional properties later during evolution.

The relative arrangement of the Phd antitoxin with Doc is similar as between VbhA and VbhT (Fig. 10.2). This may indicate a similar mode of inhibition by partial active site obstruction. However, the absence of the [S/T]xxxE[G/N] inhibitory motif in Phd implies differences in the detailed mechanism. In addition to its C-terminal segment that neutralizes Doc activity, Phd carries an N-terminal DNA-binding domain involved in transcriptional regulation (Fig. 10.2b). In contrast, VbhA has no additional domain and a putative role in transcriptional autorepression remains to be elucidated (Fig. 10.2a).

10.6 The Versatile Arrangement of the Inhibitory Helix

Surprisingly, structural comparison of the VbhT-VbhA complex with the structures of other Fic proteins deposited in the PDB databank revealed that a structural equivalent of the inhibitory α-helix of VbhA including the inhibitory motif was present as part of the Fic polypeptide chain itself in several other structures (Engel et al. 2012) (Fig. 10.4). A comparative sequence analysis showed that in the far majority of potentially adenylylating Fic proteins (i.e., Fic proteins with a strictly conserved HxFx[D/E]GNGRxxR FIC active site motif), the inhibitory helix is not provided by a VbhA-like antitoxin but fused N-terminally to the FIC domain. Moreover, in some Fic proteins α_{inh} is found as a carboxy-terminal helix. We therefore proposed a tripartite classification system for Fic proteins: class I in which α_{inh} is contributed by an antitoxin, class II and III in which the inhibitory helix is part of the Fic protein itself as an N- or C-terminal extension of the FIC domain, respectively. We managed to construct inhibition-relieved and thus activated

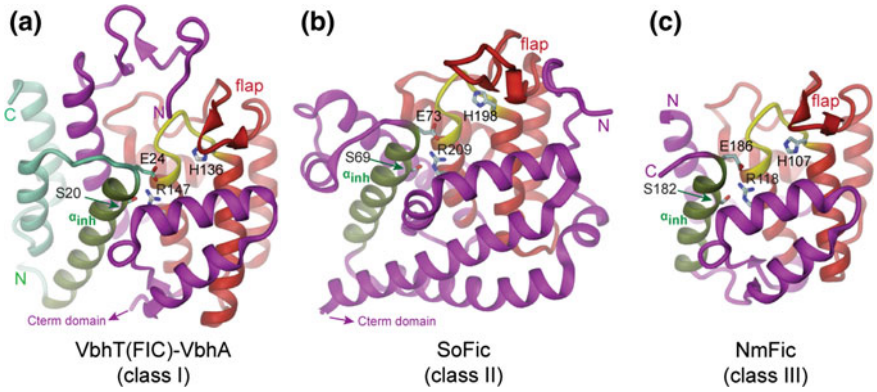


Fig. 10.4 The inhibitory motif can inhibit Fic-catalyzed adenylylation in an inter- (class I) or intra-molecular (class II and III) fashion. Representative crystal structures for **a** Class I, VbhT(FIC)-VbhA as shown in Fig. 10.2. **b** Class II, Fic protein of *S. oneidensis* (SoFic, PDB code 3EQX) (Das et al. 2009) with C-terminal domain omitted, **c** Class III, Fic protein of *N. meningitidis* (NmFic, PDB code 2G03). Whereas, in class I, the inhibitory helix (α_{inh}) with the glutamate finger is provided by a separate antitoxin protein, in class II and class III the helix is part of the FIC domain itself (Engel et al. 2012)

mutants of bacterial Fic proteins of all three classes and demonstrate that they exhibit considerable toxicity accompanied by cell filamentation in ectopically expressing *E. coli* (Engel et al. 2012). Hence, Fic proteins of class II and III—not being associated with a separate antitoxin—are not real TA systems, but could be seen as something conceptually related.

Fic proteins of class II typically display a two-domain architecture with an N-terminal FIC domain and a C-terminal DNA-binding domain, though several other domain architectures occur as well. Our model class II Fic protein So4266 of *Shewanella oneidensis* showed considerable toxicity and strongly increased autoadenylylation upon inactivation of the inhibitory helix, but no target adenylylation was detected (Engel et al. 2012). The authors of the initial publication on the crystal structure of this protein speculated that it may be a nuclease (Das et al. 2009), and it is conceivable to assume that autoadenylylation may induce conformational changes or modify the oligomeric state of the protein that would then trigger enzymatic activity. Other evidence for a function in the context of DNA is the finding that homologs of this Fic protein are often encoded within type I restriction loci (Miller et al. 2005). Like class II Fic proteins, those of class III also seem to act via the adenylylation of target proteins. But the presence of the inhibitory helix within the same polypeptide chain raises the question of how these proteins could be activated under physiological conditions in vivo.

10.7 Other Enzymatic Activities of Fic Proteins: Versatility in Function

Although adenylylation appears to be the prominent enzymatic activity of Fic proteins, other biochemical activities have been detected or suggested for members of this protein family. Whereas the majority of FIC domain-containing proteins seem to preferentially utilize ATP for target modification (Mattoo et al. 2011), VopS is capable of transferring GTP with virtually equal efficiency and may function as a guanylyl transferase. Yet, the FIC domain-containing type III secretion effector XopAC of the plant pathogen *Xanthomonas campestris* was found to specifically inactivate two host kinases by uridylylation, i.e., the transfer of an UMP moiety (Feng et al. 2012). AvrB, a type III secretion effector of *Pseudomonas syringae*, displays the canonical FIC fold and shares the same active site location as Fic proteins but with a highly degenerated FIC active site motif (MxDxRGSAAxxE). While it has been biologically characterized, the molecular mechanism underlying its function remains unknown though a possible kinase activity has been proposed (Desveaux et al. 2007). The *Legionella pneumophila* type IV secretion effector protein AnkX harbors an active site motif (HxFxDANGRxxV) slightly diverging from the canonical HxFx[D/E]GNGRxxR. This active site plasticity enables AnkX to bind CDP-choline at the active site and to catalyze phosphocholination of two small Rab GTPases in the host cell (Mukherjee et al. 2011).

10.8 Conclusions and Outlook

The VbhT-VbhA/FicT-FicA family shares several features characteristic of type II TA systems and is thus proposed to extend this class of TA systems. These novel TA system members are widespread among bacteria and are often associated with mobile DNA suggesting propagation by horizontal gene transfer. Based on the conservation of the canonical FIC active site motif in most Fic toxin components, they appear to mediate target protein adenylylation as exemplarily demonstrated for VbhT. Whether the *E. coli* Fic and close homologs in the Enterobacteriaceae that bear a slightly divergent active site motif also catalyze adenylylation or whether they exert other activities remains to be tested. The majority of Fic proteins do not belong to the VbhT-VbhA/FicT-FicA TA system and have no associated antitoxins. In fact, most of them are predicted to be autoregulated (Engel et al. 2012), i.e., inhibited intramolecularly by a mechanism analogous to that seen for the VbhT/VbhA members. Mutation of the strictly conserved glutamic acid of α_{inh} releaves the inhibition and thus provides a means to study Fic protein function. Future work on the large group of Fic proteins in general, and on FIC domain-related type II TA systems in particular should address the physiological modes of activation (stochastic and/or induced proteolysis), the nature of

the target molecules (proteins and possibly other macromolecules), and the nature of the inflicted post-translational modification (adenylylation, uridinylation, phosphocholination, etc.). In the context of the topic of this book, it will be particularly revealing to compare the molecular activities of the toxin components and their physiological regulation by the respective antitoxin components of the structurally closely related adenylylation-competent VbhT-VbhA module and the likely adenylylation incompetent FicT-FicA module, as well as representatives of the more distantly related Doc-PhD family.

References

Arbing, M.A., et al. (2010). Crystal structures of Phd-Doc, HigA, and YeeU establish multiple evolutionary links between microbial growth-regulating toxin-antitoxin systems. Structure 18, *8*, 996–1010.

Das, D. (2009). Crystal structure of the Fic (Filamentation induced by cAMP) family protein SO4266 (gil24375750) from *Shewanella oneidensis* MR-1 at 1.6 A resolution. *Proteins, 75*(1), 264–271.

Desveaux, D., Singer, A. U., Wu, A. J., McNulty, B. C., Musselwhite, L., Nimchuk, Z., et al. (2007). Type III effector activation via nucleotide binding, phosphorylation, and host target interaction. *PLoS Pathogens, 3*(3), e48.

Engel, P., Goepfert, A., Stanger, F. V., Harms, A., Schmidt, A., Schirmer, T., et al. (2012). Adenylylation control by intra- or intermolecular active-site obstruction in Fic proteins. *Nature, 482*(7383), 107–110.

Feng, F., Yang, F., Rong, W., Wu, X., Zhang, J., Chen, S., et al. (2012). A Xanthomonas uridine 5′-monophosphate transferase inhibits plant immune kinases. *Nature, 485*(7396), 114–118.

Hu, P., et al. (2009). Global functional atlas of *Escherichia coli* encompassing previously uncharacterized proteins. *PLoS Biology, 7*(4), e96.

Kawamukai, M., Matsuda, H., Fujii, W., Nishida, T., Izumoto, Y., Himeno, M., et al. (1988). Cloning of the *fic-1* gene involved in cell filamentation induced by cyclic AMP and construction of a Δ*fic Escherichia coli* strain. *Journal of Bacteriology, 170*(9), 3864–3869.

Kawamukai, M., Matsuda, H., Fujii, W., Utsumi, R., & Komano, T. (1989). Nucleotide sequences of *fic* and *fic-1* genes involved in cell filamentation induced by cyclic AMP in *Escherichia coli*. *Journal of Bacteriology, 171*(8), 4525–4529.

Kinch, L. N., Yarbrough, M. L., Orth, K., & Grishin, N. V. (2009). Fido, a novel AMPylation domain common to fic, doc, and AvrB. *PloS one, 4*(6), e5818.

Lehnherr, H., Maguin, E., Jafri, S., & Yarmolinsky, M. B. (1993). Plasmid addiction genes of bacteriophage P1: Doc, which causes cell death on curing of prophage, and phd, which prevents host death when prophage is retained. *Journal of Molecular Biology, 233*(3), 414–428.

Liu, M., Zhang, Y., Inouye, M., & Woychik, N. A. (2008). Bacterial addiction module toxin Doc inhibits translation elongation through its association with the 30S ribosomal subunit. *Proceedings of the National academy of Sciences of the United States of America, 105*(15), 5885–5890.

Luong, P., Kinch, L. N., Brautigam, C. A., Grishin, N. V., Tomchick, D. R., & Orth, K. (2010). Kinetic and structural insights into the mechanism of AMPylation by VopS Fic domain. *The Journal of Biological Chemistry, 285*(26), 20155–20163.

Magnuson, R., & Yarmolinsky, M. B. (1998). Corepression of the P1 addiction operon by Phd and Doc. *Journal of Bacteriology, 180*(23), 6342–6351.

Makarova, K. S., Wolf, Y. I., & Koonin, E. V. (2009). Comprehensive comparative-genomic analysis of type 2 toxin-antitoxin systems and related mobile stress response systems in prokaryotes. *Biology Direct, 4*, 19.

Mattoo, S., Durrant, E., Chen, M. J., Xiao, J., Lazar, C. S., Manning, G., et al. (2011). Comparative analysis of *Histophilus somni* immunoglobulin-binding protein A (IbpA) with other fic domain-containing enzymes reveals differences in substrate and nucleotide specificities. *The Journal of Biological Chemistry, 286*(37), 32834–32842.

Miller, W. G., Pearson, B. M., Wells, J. M., Parker, C. T., Kapitonov, V. V., & Mandrell, R. E. (2005). Diversity within the *Campylobacter jejuni* type I restriction-modification loci. *Microbiology, 151*(Pt 2), 337–351.

Mukherjee, S., Liu, X., Arasaki, K., McDonough, J., Galan, J. E., & Roy, C. R. (2011). Modulation of Rab GTPase function by a protein phosphocholine transferase. *Nature, 477*(7362), 103–106.

Palanivelu, D. V., Goepfert, A., Meury, M., Guye, P., Dehio, C., & Schirmer, T. (2011). Fic domain-catalyzed adenylylation: Insight provided by the structural analysis of the type IV secretion system effector BepA. *Protein science : a publication of the Protein Society, 20*(3), 492–499.

Punta, M. et al. (2012). The Pfam protein families database. Nucleic acids research 40 (Database issue):D290–301.

Utsumi, R., Nakamoto, Y., Kawamukai, M., Himeno, M., & Komano, T. (1982). Involvement of cyclic AMP and its receptor protein in filamentation of an *Escherichia coli fic* mutant. *Journal of Bacteriology, 151*(2), 807–812.

Watson, J. D., & Milner-White, E. J. (2002). A novel main-chain anion-binding site in proteins: The nest. A particular combination of phi, psi values in successive residues gives rise to anion-binding sites that occur commonly and are found often at functionally important regions. *Journal of Molecular Biology, 315*(2), 171–182.

Worby, C. A., Mattoo, S., Kruger, R. P., Corbeil, L. B., Koller, A., Mendez, J. C., et al. (2009). The fic domain: Regulation of cell signaling by adenylylation. *Molecular Cell, 34*(1), 93–103.

Xiao, J., Worby, C. A., Mattoo, S., Sankaran, B., & Dixon, J. E. (2010). Structural basis of Fic-mediated adenylylation. *Nature Structural & Molecular Biology, 17*(8), 1004–1010.

Yamaguchi, Y., Park, J. H., & Inouye, M. (2011). Toxin-antitoxin systems in bacteria and archaea. *Annual Review of Genetics, 45*, 61–79.

Yarbrough, M. L., Li, Y., Kinch, L. N., Grishin, N. V., Ball, H. L., & Orth, K. (2009). AMPylation of Rho GTPases by *Vibrio* VopS disrupts effector binding and downstream signaling. *Science, 323*(5911), 269–272.

Chapter 11
Type II Toxin-Antitoxin Loci, *hipBA* and Persisters

Kim Lewis and Sonja Hansen

Abstract Persisters are dormant cells responsible for drug tolerance of chronic infections. Persisters can be produced stochastically and make up a small subpopulation of cells. Mechanisms of dormancy appear to be highly redundant, and Toxin/Antitoxin (TA) modules have been linked to persister formation. Interferases (mRNA endonucleases) RelE or MazF cause stasis and multidrug tolerance when expressed topically, and deletion of multiple interferases in *E. coli* decreases the level of persisters. In this chapter, we mainly focus on description of the HipA and TisB toxins. HipA has served as a model for the study of persisters. HipA is a protein kinase, which phosphorylates EF-Tu, inhibiting protein synthesis and causing dormancy. TisB is a typical antimicrobial peptide and is induced by the SOS response. TisB forms anion channels in the *E. coli* membrane, causing a drop in pmf and ATP, leading to dormancy.

11.1 Introduction

First reported in 1944 (Bigger 1944), persisters attracted little attention until their rediscovery as a principal component of biofilm drug tolerance (Brooun et al. 2000; Spoering and Lewis 2001). While acute infections have been reasonably well contained in the developed world, chronic infections have been on the rise as a direct consequence of medical intervention—the widespread use of indwelling devices, and the increase in immune compromised patients due to cancer

K. Lewis (✉) · S. Hansen
Antimicrobial Discovery Center, Department of Biology,
Northeastern University, Boston, MA, USA
e-mail: k.lewis@neu.edu

K. Gerdes (ed.), *Prokaryotic Toxin-Antitoxins*,
DOI: 10.1007/978-3-642-33253-1_11,

chemotherapy, organ transplantation, and the increase in the elderly population due to advances in medicine. Many chronic infections are associated with the ability of the pathogen to form a biofilm (Costerton et al. 1999). Biofilms show a surprising ability to resist killing by antibiotics, without having any obvious drug resistance mechanisms. Our study of dose-dependent killing of a *P. aeruginosa* biofilms showed the presence of a small subpopulation of cells completely tolerant to antibiotics (Brooun et al. 2000; Spoering and Lewis 2001). We recently linked the ability of a pathogen to make persisters to the clinical manifestation of chronic disease (Lafleur et al. 2010; Mulcahy et al. 2010; Lewis 2010). In patients with cystic fibrosis, a prolonged infection with *P. aeruginosa* ultimately leads to death. The pathogens do not necessarily carry antibiotic resistance. What we found is that in late isolates from these patients, antibiotic treatment selects for mutants that make 1,000-fold more persisters (Mulcahy et al. 2010). The ability to produce antibiotic-tolerant persisters is apparently the culprit responsible for death of patients with cystic fibrosis. Similarly, in patients with unresolved oral thrush infection of *C. albicans*, isolates were invariably high persister (*hip*) mutants (Lafleur et al. 2010).

The finding of persisters provided a rational explanation to the paradox of chronic infections, which resist killing by antibiotics without exhibiting any resistance mechanisms. We proposed that persisters are dormant, and tolerance stems from inactivity of antibiotic targets in these cells (Lewis 2007; Keren et al. 2004). Indeed, bactericidal antibiotics kill not by inhibiting functions, but by corrupting their targets. Thus, aminoglycoside cause mistranslation, producing toxic misfolded peptides (Davis et al. 1986); cells wall synthesis inhibitors cause induction of autolysins and cell death by a process that is still poorly understood (Bayles 2007); and fluoroquinolones bind to the DNA gyrase and topoisomerase and inhibit the relegation step without affecting the nicking. As a result, the enzymes are converted into DNA endonucleases (Hooper 2001).

Dormancy of persisters was confirmed in an experiment designed to isolate these cells with a temporary phenotype. *E. coli* cells carrying a ribosomal promoter fusion to degradable GFP were sorted based on fluorescence intensity. The population of these strains split into two groups—bright and dim. Apparently, a cell entering into dormancy would have a diminished rate of protein synthesis and degradation of GFP would make it dim. These dim cells had an increased antibiotic tolerance (Shah et al. 2006). Using this method and isolation of persisters by simply lysing a culture with ampicillin, we obtained persister transcriptome which pointed to the expression of TA modules (Keren et al. 2004; Shah et al. 2006). Overexpression of RelE (Keren et al. 2004), MazF (Vazquez-Laslop et al. 2006), and HipA (Falla and Chopra 1998; Korch et al. 2003; Correia et al. 2006; Vazquez-Laslop et al. 2006) toxins produced cells that emulated persisters—these were nongrowing and multidrug tolerant. In *E. coli*, there are at least 11 TA modules, suggesting a high degree of redundancy of function. A recent study by Kenn Gerdes and coworkers shows that deletion of at least five TA modules is required for a decrease in persister formation (Maisonneuve et al. 2011). There are at least 62 TA modules in *M. tuberculosis* (Ramage et al. 2009), a pathogen which

is very difficult to eradicate, probably because of the presence of a dormant, persister form (Wayne and Sohaskey 2001; Garton et al. 2008; Wayne and Hayes 1996). Knocking out multiple TAs in *M. tuberculosis* is probably not realistic, but our recent isolation of persisters from this organism points to some candidate TA modules that may be involved in dormancy (Keren et al. 2011).

One possible explanation for the multiplicity of TA modules is that each of them is induced, and plays a major role in persister formation under particular conditions. In this chapter, we will describe in more detail the HipA toxin, which has served as a model for studying persister formation; and the TisB toxin, provides an example of an inducible persister protein.

11.2 Type II Toxin-Antitoxin A loci: *hipBA*

11.2.1 Discovery of hipBA

The *hipA* gene was discovered by Harris Moyed in a targeted search for *E. coli* mutants with an increased frequency of persister cells (Moyed and Bertrand 1983). In this study, an ethyl methane sulfate (EMS) treated *E. coli* culture was subjected to several cycles of ampicillin exposure followed by outgrowth in fresh medium. Upon plating on solid medium, mutants were identified that produced a higher number of persister cells than the parent strain, but were unable to grow in the presence of ampicillin. One of these *hip* mutations was mapped to 33.8 min on the *E. coli* chromosome, the position of the *hipBA* operon, which is located in the terminus region near *terC3* (Moyed and Bertrand 1983). The persister fraction of the strain harboring the mutation in the *hip* locus (*hipA7* allele) was at least 1,000-times higher compared to the isogenic wild type after treatment with ampicillin (Moyed and Bertrand 1983). Similar results were obtained for phosphomycin and cycloserine treatment and diaminopimelic acid (DAP) starvation, which inhibit different steps of peptidoglycan synthesis (Moyed and Bertrand 1983). Due to the increase in survival upon exposure to cell wall acting antibiotics, Moyed suggested that *HipA* is involved in cell division (Moyed and Bertrand 1983). However, the *hipA7* allele also appears to affect DNA synthesis as it protected cells from thymine-less death, the quinolone naladixic acid and lethal heat shock (Scherrer and Moyed 1988). In that sense *hipA7* is a typical persister gene since expression confers multidrug tolerance.

The *hipA7* mutant has a severe growth defect at low temperature, which increases with the decrease of temperature. Persistence and cold sensitivity appear to be linked (Scherrer and Moyed 1988). At 20 °C, only colonies that have lost the high persistence phenotype are able to grow (Scherrer and Moyed 1988). However, exposure to non-permissive temperature is not lethal, after a shift from 20 °C back to 37 °C comparable CFUs are recovered for *hipA7* and *hipA*$^{+}$ strains (Scherrer and Moyed 1988). The block in cell division becomes more rapid with the age of the culture that is shifted to low temperature (Scherrer and Moyed

1988). Revertant that can be recovered at a low frequency (0.1 %) have partially lost both cold sensitivity and high persistence. Scherrer and Moyed were unable to separate the two phenotypes by complementation using plasmids containing *hipA* fragments. However, they were able to place the location of the *hipA* mutation near the 5′ end of *hipA* (Scherrer and Moyed 1988). Years later *hipA7* was sequenced and one mutation near the 5′ end, G22S, was indeed confirmed, as well as a second mutation, D291A (Korch et al. 2003).

11.2.2 Structure and Organization of hipBA

Falla and Chopra first recognized the similarity between the *hipBA* operon and poison/antidote systems of plasmid stabilization systems involved in post-segregational killing (Falla and Chopra 1999). The *hipA* gene exhibits similarity to *y4mE* and *y4dL* (28 and 27 % identity, respectively) (Falla and Chopra 1998). Both genes are located on a plasmid in *Rhizobium* species NGR234 and are controlled by negative transcriptional regulators, *y4mF* and *y4dL*. At that time, the notion that TA systems have chromosomal counterparts was still fairly young. The *pem* homologs *chpA* and *chpB* and the *relBEF* operon were among the first TA loci identified on the chromosome (Masuda et al. 1993; Gotfredsen and Gerdes 1998) because they shared the basic characteristics of plasmid-encoded TA systems: organization in an operon in which the toxin usually follows the antitoxin; the antitoxin functions as a transcriptional repressor and the toxin as the co-repressor; and free toxin is inhibitory or lethal to the host cell (Jensen and Gerdes 1995; Tian et al. 1996). Also, the *hip* locus shares these characteristics (Fig. 11.1). The antitoxin *hipB* and the toxin *hipA* form an operon (Black et al. 1991). The stop codon of *hipB* overlaps the start codon of *hipA* by one base pair (bp) (Black et al. 1991). *hipB* encodes a 10 kDa protein and forms a dimer in solution, *hipA* encodes a 50 kDa protein, which is exceptionally large for a TA toxin. HipB is a Cro-like DNA-binding protein with a helix-turn-helix DNA-binding motif (Black et al. 1991). HipA is toxic in the absence of HipB, since *hipB* cannot be deleted in $hipA^+$ background, and *hipA*-containing plasmids cannot be transformed in $hipB^-$ background (Black et al. 1991, 1994). Using transcriptional *lacZ*-fusions and gel shift assays it was demonstrated that HipB is the repressor of the *hipBA* operon (Black et al. 1994). Expression of the *hipBA* operon is tightly regulated. Three regions of dyad symmetry with the consensus sequence $TATCC(N)_8GGATA$, a fourth longer inverted repeat located and an integration host factor (IHF) consensus binding site are present upstream of the operon (Black et al. 1991, 1994). Both HipB alone and the HipBA complex bind cooperatively to these operator sequences (Black et al. 1994).

The HipB dimer is sandwiched by two HipA molecules (Schumacher et al. 2009). The HipA N domains interact with one HipB subunit, while the C domains mainly interact with the other HipB subunit (Schumacher et al. 2009). In contrast to many other antitoxins, HipB binding does not occlude the active site of HipA

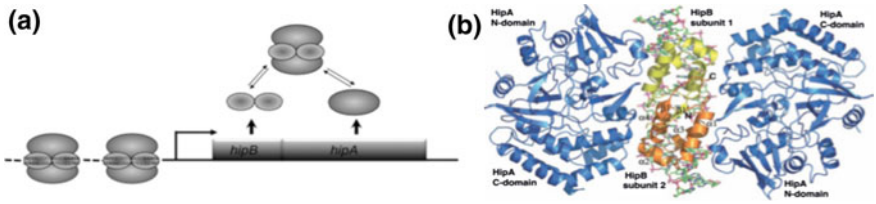


Fig. 11.1 Genetic organization and components of *E. coli hipBA*. **a** Organization of the *HipB*A operon. Two of the four inverted repeats are depicted. *HipB* neutralizes *HipA* and represses transcription of the *hipBA* operon. **b** Ribbon diagram of the *HipA-HipB*-DNA operator complex. Two *HipA* monomers are pictured in *blue*, and one subunit of the *HipB* dimer is *yellow* and other is *orange*. The DNA is shown as sticks with carbon, nitrogen, oxygen, and phosphorus atoms colored in *green*, *blue*, *red*, and *magenta*, respectively (Schumacher et al. 2009)

(Schumacher et al. 2009). It has been proposed that HipB contacts lock HipA in an open inactive conformation which still allows ATP to bind but prevents protein substrates from binding which would cause a bigger conformational change and domain closure (Schumacher et al. 2009). HipB specifically interacts with DNA through major groove contacts. Binding of HipB induces a large 70° bend in the operator which aligns the recognition helices for binding to consecutive major grooves. Ser^{29}, Gln^{39}, Ala^{40}, Ser^{43}, and Lys^{38} extend contact to the motif $T_2A_3T_4C_5C_6$ (Schumacher et al. 2009). HipB also makes 11 phosphate contacts to each half site of the inverted repeat (Schumacher et al. 2009). DNA contacts are not limited to HipB, Lys^{379}, and Arg^{382} of HipA also interacts with DNA via phosphate backbone contacts (Schumacher et al. 2009).

11.2.3 HipA-Mediated Growth Inhibition and Drug Tolerance

The increase in tolerance to multiple antibiotics is not restricted to the *hipA7* allele, because overexpression of HipA also leads to a high frequency of persister cells (Falla and Chopra 1998; Korch et al. 2003; Correia et al. 2006; Vazquez-Laslop et al. 2006). In contrast, a deletion of *hipA* or *hipBA* does not produce any appreciable phenotype (Falla and Chopra 1999; Korch et al. 2003; Hansen et al. 2008). Ectopic expression of *hipA* is toxic (Falla and Chopra 1998; Korch et al. 2003; Correia et al. 2006; Rotem et al. 2010). Upon induction of HipA, DNA replication, as well as RNA, and protein synthesis are inhibited (Korch and Hill 2006), while overexpression of *hipA7* does not lead to growth inhibition even in the absence of HipB (Korch et al. 2003). Similar to other TA systems (Pedersen et al. 2002; Christensen et al. 2003) the static effect of HipA overproduction can be readily reversed by induction of *hipB* expression (Korch and Hill 2006). Though growth is eventually fully restored, colonies appear over an extended period of time depending on the level of induction (Rotem et al. 2010). At intermediate induction levels, HipA is neutralized by HipB immediately in part of the population, and growth continues like in wild type, while other cells are arrested for as

long as several days (Rotem et al. 2010). Using fluorescent reporters it was demonstrated that HipA expression has to pass a threshold level in order to have an effect on growth (Rotem et al. 2010).

11.2.4 Mechanism of Action of HipA Kinase

Comparative sequence analysis places HipA in the phosphatidylinositol (PI) 3/4-kinase superfamily with closest similarity to the PI 4-kinase II family (Correia et al. 2006). The superfamily includes bacterial and eukaryotic members but is lacking any archaeal representatives (Correia et al. 2006). Motifs that are shared with eukaryotic superfamily members include the catalytic residue of PI 3-kinases, aspartate 309, Mg^{2+}-binding residues asparagine 314 and aspartate 332, and the ATP-binding lysine 181 (Correia et al. 2006). The catalytic core domain resides in the C-terminal part of HipA-like proteins. In addition, the N-terminus harbors a globular domain spanning 100 amino acids that does not have any detectable homology outside the HipA family (Correia et al. 2006). Rare examples are found in which the N-terminal domain is encoded by a separate gene that tends to be adjacent to the gene encoding the HipA-like catalytic domain (e.g. HI0666 in *Haemophilus influenza* RdKW20) and is also found at a distal location (e.g., c5296 in *E. coli* CFT073) suggesting an important still unknown role of the N-terminal domain in the function of HipA (Correia et al. 2006).

In vitro phosphorylation of HipA confirmed the results obtained by the bioinformatics approach (Correia et al. 2006). A single site, Ser150, was identified that was phosphorylated. No phosphorylation was detected in a S150A point mutant, and neither in a D309Q (predicted active site) and G332Q (Mg^{2+}-binding site) point mutant (Correia et al. 2006). Residues S150, D309, and D322 were also necessary for HipA toxicity and HipA-mediated persistence (Correia et al. 2006).

HipA binds ATP with high affinity (Schumacher et al. 2009). HipA has a globular fold that can be divided into an N-terminal α/β-domain and a C-terminal α-helical-domain (Schumacher et al. 2009). HipA, in particular the C-terminal domain, shows the highest structural similarity to human CDK2/cyclin A kinase (Schumacher et al. 2009). Together, these findings suggest that HipA acts on one or more targets via phosphorylation and that the phosphorylated state of its target(s) causes shut down of cellular metabolism (Schumacher et al. 2009). Elongation factor Tu (EF-Tu) is the first validated target, which was identified in a pull down assay (Schumacher et al. 2009). EF-Tu mediates aminoacyl-tRNA binding to the ribosome. Upon GTP hydrolysis to GDP, EF-Tu undergoes a conformational change from a closed to an open form, which can no longer bind the ribosome. Thr^{382} stabilizes the closed conformation, upon phosphorylation of Thr^{382}, EF-Tu is locked in its open confirmation and thus translation halts (Lippmann et al. 1993; Alexander et al. 1995). Phosphorylation of EF-Tu by HipA was confirmed in vitro and also by HipA-IREGGR*T*VGA interaction, an EF-Tu peptide encompassing Thr^{382} (Schumacher et al. 2009).

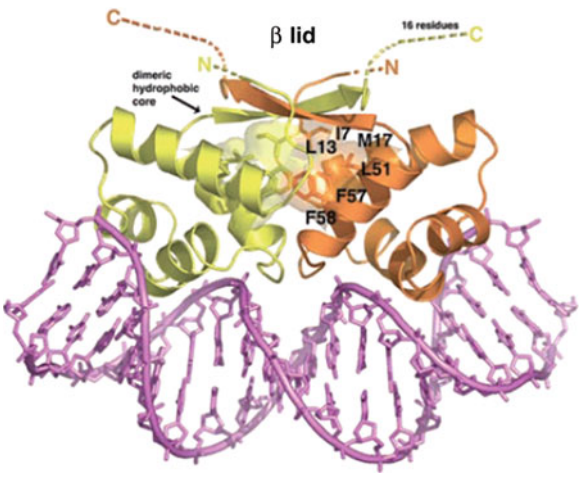


Fig. 11.2 *HipB* pictured as a dimer with one *yellow* and one *orange* subunit. DNA is shown as *pink* sticks. The *β*-lid is composed of a short two-stranded *β* sheet. Disordered C-terminal residues extend from the lid and are depicted as colored dashes (Schumacher et al. 2009)

11.2.5 Degradation of HipB by Lon Protease

To activate HipA, the antitoxin HipB has to be removed or degraded from the HipA-$HipB_2$-HipA complex. HipB is rapidly degraded in wild-type background but is stabilized in *lon*$^-$ background (Hansen et al. 2012). These findings were confirmed biochemically, showing that Lon is the main protease responsible for HipB proteolysis (Hansen et al. 2012). The first 3 and the last 16 residues of HipB are disordered and are located near a *β*-lid (composed of *β*1 from one to *β*1′ from the other subunit) that covers the hydrophobic core of the protein (Fig. 11.2) (Schumacher et al. 2009). Rapid turnover of HipB is dependent on the presence of the unstructured carboxy-terminal stretch of HipB (Hansen et al. 2012), suggesting a site for protease attack (Schumacher et al. 2009). It was suggested that the three C-terminal residues Leu^{84}, Glu^{85}, and Trp^{86} also bind a HipA pocket, which lowers the chance for proteolytic attack and may alter HipA function (Evdokimov et al. 2009). However, the affinity of HipB for HipA or for *hipBA* operator DNA was not altered when the 16-residue fragment was removed or when the conserved C-terminal tryptophan of HipB was substituted to alanine, indicating that the major role of this region of HipB is indeed to control HipB degradation (Hansen et al. 2012).

11.3 Type I TA loci: *tisB*

As mentioned in the Introduction, preferential induction under given conditions may be the reason for the redundancy of TA modules. In order to test this possibility, we turned to TAs that carry a Lex box in their promoter region and appear to be controlled by the SOS response. These are *symER*, *hokE*, *yafN*/*yafO*, and *tisAB*/*istr1* (Courcelle et al. 2001; Fernandez De Henestrosa et al. 2000; Kawano

et al. 2007; McKenzie et al. 2003; Motiejunaite et al. 2007; Pedersen and Gerdes 1999; Singletary et al. 2009; Vogel et al. 2004). Another locus, *dinJ*/*yafQ* contains a Lex box but is not believed to be under SOS control (Courcelle et al. 2001; Fernandez De Henestrosa et al. 2000). Fluoroquinolones induce the SOS response (Phillips et al. 1987), and we tested the ability of ciprofloxacin to induce persister formation (Dorr et al. 2009; Dorr et al. 2010). Examination of deletion strains showed that the level of persisters dropped dramatically, 10–100-fold, in a Δ*tisAB* mutant. This suggests that TisB was responsible for the formation of the majority of persisters under conditions of SOS induction. The level of persisters was unaffected in strains deleted in the other Lex box containing TA modules. The *tisAB* promoter was also the strongest responding to SOS induction among the Lex box TA modules. Persister levels observed in time-dependent killing experiments with ampicillin or streptomycin that do not cause DNA damage were unchanged in the Δ*tisAB* strain. TisB only had a phenotype in the presence of a functional RecA protein, confirming the dependence on the SOS pathway (Dorr et al. 2010).

Ectopic overexpression of TisB sharply increased the level of persisters. Drop in persisters in a deletion strain and increase upon overexpression gives reasonable confidence in functionality of a persister gene. The dependence of TisB-induced persisters on a particular regulatory pathway, the SOS response, further strengthens the case for TisB as a specialized persister protein (Dorr et al. 2010).

The role of TisB in persister formation is unexpected based on what we know about this type of proteins. TisB is a typical antimicrobial peptide—it is short, 29 amino acids long, hydrophobic, and carries a net positive charge. We find that it forms an anion channel in a model lipid bilayer membrane (Gurnev et al. 2012). Ectopic expression of TisB was reported to decrease the level of ATP in the cells (Unoson and Wagner 2008). Bacteria, plants, and animals all produce antimicrobial membrane-acting peptides (Garcia-Olmedo et al. 1998; Sahl and Bierbaum 1998; Wolfson et al. 1990). Toxins of many TA loci found on plasmids belong to this type as well. If a daughter cell does not inherit a plasmid, the concentration of a labile antitoxin decreases, and the toxin such as the membrane-acting hok kills the cell (Gerdes et al. 1986). High-level artificial overexpression of TisB also causes cell death (Unoson and Wagner 2008). It is remarkable from this perspective that the membrane-acting TisB under conditions of natural expression has the exact opposite effect of protecting the cell from antibiotics. Cells expressing TisB stop growing, and the drop in pmf and ATP levels will shut down the targets of bactericidal antibiotics. Ciprofloxacin kills cells primarily by converting its target proteins, DNA topoisomerases, into DNA endonucleases (Chen et al. 1996; Drlica et al. 2008). A drop in ATP will then prevent topoisomerases from damaging the DNA. Tolerance to β-lactams and aminoglycosides will similarly result from the inactivity of their targets. By creating a dormant state, TisB will cause a shutdown of antibiotic targets and multidrug tolerance (Fig. 11.3).

Fluoroquinolones such as ciprofloxacin are widely used broad-spectrum antibiotics, and their ability to induce multidrug-tolerant cells is unexpected and a cause of considerable concern. Induction of persister formation by fluoroquinolones may contribute to the ineffectiveness of antibiotics in eradicating infections.

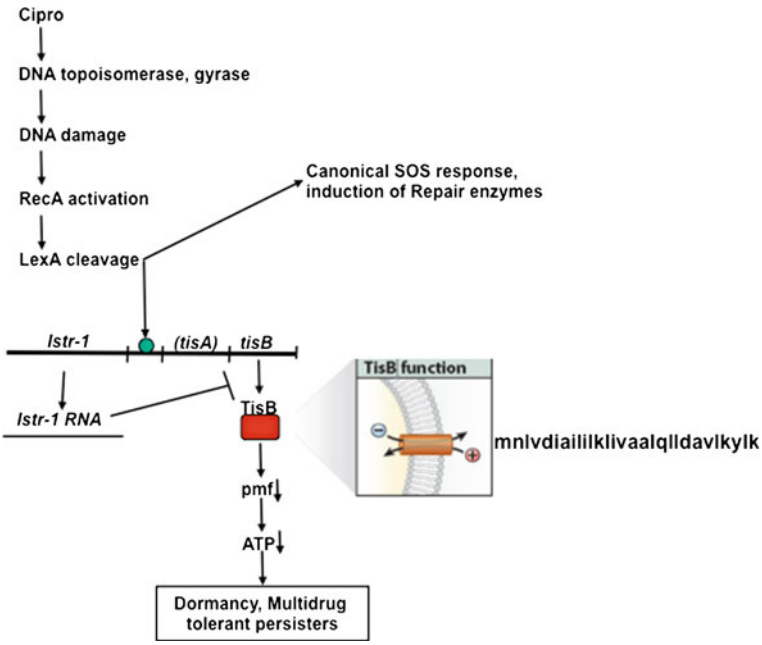


Fig. 11.3 A model of TisB-dependent persister formation in *E. coli*. A fluoroquinolone antibiotic causes DNA damage by converting the DNA gyrase and topoisomerase into endonucleases. This activates the RecA protein which in turn activates the LexA repressor, causing it to cleave. The canonical SOS response is induced, and repair enzymes that contain lex boxes in their promoter regions are transcribed. The Lex repressor also controls the expression of the TisB toxin, a small cationic membrane-acting agent. Decrease in the pmf and ATP shut down target functions, including DNA topoisomerase and gyrase, and a dormant persister is formed

Indeed, pre-exposure with a low dose of ciprofloxacin drastically increased tolerance to subsequent exposure with a high dose, and TisB persisters are multidrug tolerant.

Induction of persisters by the SOS-induced TisB toxin links together two seemingly opposite strategies of survival—active repair, and systems shutdown in a dormant state. It seems that in the presence of DNA damaging factors, the optimal strategy is to both induce repair and increase the number of dormant cells, which will survive when everything else fails. Indeed, a progressive increase in the exposure to fluoroquinolones kills regular cells but has little effect on the survival of persisters. This means that it is the dormant persisters rather than regular cells with induced repair that will ultimately survive the DNA damaging antibiotic.

The finding of the role of TisB in tolerance opens an intriguing possibility of a wider link between other stress responses and persister formation. Pathogens are exposed to many stress factors in the host environment apart from DNA damaging agents—oxidants, high temperature, low pH, membrane-acting agents. It is possible that all stress responses induce the formation of surviving persisters.

11.4 Other Toxin/Antitoxins and Persister Formation

Deletion of *hipBA* had no affect on persistence; therefore, other genes are likely involved in persister formation. To identify additional genes affecting persister formation, a gene expression profile from persister cells was obtained using HM22, which has the *hipA7* allele, and produces 1 % persisters in exponentially growing cells (Keren et al. 2004). HM22 was treated with ampicillin, and persisters were collected over the course of 180 min (Keren et al. 2004). A gene cluster comprising approximately 300 genes was identified, which was upregulated over time concomitant with the enrichment of persisters. This group contained several stress response genes from phage-shock, cold-shock, heat-shock SOS response, and also genes belonging to the TAs *dinJ/yafQ*, *yefM*, *relBE*, and *mazEF* (Keren et al. 2004). Overexpression of RelE caused a 10,000-fold increase in the persister fraction after cefotaxime and ofloxacin treatment and a 10-fold increase after tobramycin treatment, but had no effect on mitomycin treatment, while deletion of *relBE* had no effect on persistence (Keren et al. 2004). A second transcriptome was obtained from naïve persisters using *E. coli*-ASV, a strain with an unstable GFP fused to a ribosomal promoter (Shah et al. 2006). Persister cells were sorted out using fluorescent-activated cell sorting (FACS) under the assumption that an exponentially growing culture produces persisters that are dormant or non-growing and therefore will appear dim, whereas the growing population appears bright (Shah et al. 2006). In accordance with the earlier transcriptome of persister enrichment of the *hipA7* mutant, the TA systems *dinJ/yafQ* and *yefM/yoeB* were upregulated in persister cell. However, *relBE* and *mazEF* were not significantly upregulated under the conditions tested. The most highly expressed gene in persister cells was *ygiU* (*mqsR*) *of the ygiUT* TA system (Shah et al. 2006), see also Chap. 6. Overexpression of YgiU led to nearly complete protection after ofloxacin and cefotaxime treatment, but did not result in protection to mitomycin C and tobramycin treatment (Shah et al. 2006). A *ygiUT* deletion had no phenotype (Shah et al. 2006), One study reports a modest seven-fold reduction in survival of the *ygiUT* (*mqsRA*) deletion at a single 2 h time point after ampicillin treatment, and unfortunately but does not characterize the phenotype of the *mqsRA* deletion beyond the time point and the antibiotic (Kim and Wood 2010). Overexpression of YafQ in biofilms selectively protected from cefazolin to tobramycin treatment, but not doxycycline and rifampicin (Harrison et al. 2009). The deletion of *yafQ* did not affect survival of planktonic cells, but caused reduced persister levels in biofilms after cefazolin and tobramycin treatment pointing to the importance of *dinJ/yafQ* for persister cell formation under specific conditions (Harrison et al. 2009). Asking whether overexpression of any mRNA endonuclease (Maisonneuve et al. 2011) will cause an increase in tolerance, Maisonneuve et al. (2011) tested five of the ten known mRNA endonucleases in *E. coli*, RelE, YafO, MqsR, HigB, and MazF, for their survival after treatment with ampicillin and ciprofloxacin and find elevated persister levels in all cases (Maisonneuve et al. 2011). One concern about overexpression of proteins is that also unrelated

proteins, that are toxic upon overexpression from plasmids, lead to an increase in antibiotic tolerance (Vazquez-Laslop et al. 2006). Using a non-toxic RelE mutant to titrate endogenous RelB, which in turn leads to unbound and hence active endogenous RelE, it was possible to show that endogenous RelE is sufficient for the increase in tolerance (Maisonneuve et al. 2011). Single deletions of any toxin gene did not produce a robust phenotype. However, a strain in which all ten endonuclease-encoding genes were deleted caused a 100–200-fold reduction in the persister level (Maisonneuve et al. 2011). The phenotype could not be attributed to any particular toxin. Deletion of five or more toxin genes in any order resulted in a phenotype, which was more pronounced the more toxins were deleted (Maisonneuve et al. 2011).

11.5 Concluding Remarks

The principal role of persisters in biofilm antibiotic tolerance in vitro, and the recent findings linking persisters to chronic infections in cystic fibrosis and oropharyngeal candidiasis suggest that persisters are responsible for recalcitrance of chronic infections. This largely addresses the significance question, and future studies will detail the role of persisters in all chronic infections. Linking of persisters to the function of TA modules is an important development, but so far has been restricted to studies of *E. coli*. In spite of the redundancy of TAs in *E. coli*, there is good evidence that additional mechanisms are likely to produce dormant cells (Hansen et al. 2008). Identification of persister genes in *E. coli* and other microorganisms has only started and will undoubtedly bring many interesting discoveries.

These encouraging advances provide a good starting position to tackle the remaining unsolved problems in this emerging field: the nature of noise generation in persister formation; the mechanisms of persister resuscitation; the identity of mutations in clinical isolates of *hip* strains; the interplay between resistance, tolerance, and protection by stress responses; and, finally, the challenging problem of persister eradication.

References

Alexander, C., Bilgin, N., Lindschau, C., Mesters, J. R., Kraal, B., Hilgenfeld, R., et al. (1995). Phosphorylation of elongation factor Tu prevents ternary complex formation. *Journal of Biological Chemistry, 270*(24), 14541–14547.

Bayles, K. W. (2007). The biological role of death and lysis in biofilm development. *Nature Reviews, 5*(9), 721–726.

Bigger, J.W. (1944). Treatment of staphylococcal infections with penicillin. Lancet ii: 497–500.

Black, D. S., Irwin, B., & Moyed, H. S. (1994). Autoregulation of *hip*, an operon that affects lethality due to inhibition of peptidoglycan or DNA synthesis. *Journal of Bacteriology, 176*(13), 4081–4091.

Black, D. S., Kelly, A. J., Mardis, M. J., & Moyed, H. S. (1991). Structure and organization of *hip*, an operon that affects lethality due to inhibition of peptidoglycan or DNA synthesis. *Journal of Bacteriology, 173*(18), 5732–5739.

Brooun, A., Liu, S., & Lewis, K. (2000). A dose-response study of antibiotic resistance in *Pseudomonas aeruginosa* biofilms. *Antimicrobial Agents and Chemotherapy, 44*(3), 640–646.

Chen, C. R., Malik, M., Snyder, M., & Drlica, K. (1996). DNA gyrase and topoisomerase IV on the bacterial chromosome: quinolone-induced DNA cleavage. *Journal of Molecular Biology, 258*(4), 627–637. doi:10.1006/jmbi.1996.0274.

Christensen, S. K., Pedersen, K., Hansen, F. G., & Gerdes, K. (2003). Toxin-antitoxin loci as stress-response-elements: ChpAK/MazF and ChpBK cleave translated RNAs and are counteracted by tmRNA. *Journal of Molecular Biology, 332*(4), 809–819. doi:S0022283603009227. [pii].

Correia, F. F., D'Onofrio, A., Rejtar, T., Li, L., Karger, B. L., Makarova, K., et al. (2006). Kinase activity of overexpressed *HipA* is required for growth arrest and multidrug tolerance in *Escherichia coli*. *Journal of Bacteriology, 188*(24), 8360–8367.

Costerton, J. W., Stewart, P. S., & Greenberg, E. P. (1999). Bacterial biofilms: A common cause of persistent infections. *Science, 284*(5418), 1318–1322.

Courcelle, J., Khodursky, A., Peter, B., Brown, P. O., & Hanawalt, P. C. (2001). Comparative gene expression profiles following UV exposure in wild-type and SOS-deficient Escherichia coli. *Genetics, 158*(1), 41–64.

Davis, B. D., Chen, L. L., & Tai, P. C. (1986). Misread protein creates membrane channels: an essential step in the bactericidal action of aminoglycosides. *Proceedings of the National academy of Sciences of the United States of America, 83*(16), 6164–6168.

Dorr, T., Lewis, K., & Vulic, M. (2009). SOS response induces persistence to fluoroquinolones in Escherichia coli. *PLoS Genetics, 5*(12), e1000760. doi:10.1371/journal.pgen.1000760.

Dorr, T., Vulic, M., & Lewis, K. (2010). Ciprofloxacin causes persister formation by inducing the TisB toxin in Escherichia coli. *PLoS Biology, 8*(2), e1000317. doi:10.1371/journal.pbio.1000317.

Drlica, K., Malik, M., Kerns, R. J., & Zhao, X. (2008). Quinolone-mediated bacterial death. *Antimicrobial Agents and Chemotherapy, 52*(2), 385–392. doi:10.1128/AAC.01617-06.

Evdokimov, A., Voznesensky, I., Fennell, K., Anderson, M., Smith, J. F., & Fisher, D. A. (2009). New kinase regulation mechanism found in *HipB*A: a bacterial persistence switch. *Acta Crystallographica Section D Biological Crystallography, 65*(Pt 8), 875–879.

Falla, T. J., & Chopra, I. (1998). Joint tolerance to beta-lactam and fluoroquinolone antibiotics in *Escherichia coli* results from overexpression of *hipA*. *Antimicrobial Agents and Chemotherapy, 42*(12), 3282–3284.

Falla, T. J., & Chopra, I. (1999). Stabilization of *Rhizobium* symbiosis plasmids. *Microbiology, 145*(Pt 3), 515–516.

Fernandez De Henestrosa, A. R., Ogi, T., Aoyagi, S., Chafin, D., Hayes, J. J., Ohmori, H., et al. (2000). Identification of additional genes belonging to the LexA regulon in Escherichia coli. *Molecular Microbiology, 35*(6), 1560–1572.

Garcia-Olmedo, F., Molina, A., Alamillo, J. M., & Rodriguez-Palenzuela, P. (1998). Plant defense peptides. *Biopolymers, 47*(6), 479–491. doi:10.1002/(SICI)1097-0282(1998)47:6<479:AID-BIP6>3.0.CO;2-K.

Garton, N. J., Waddell, S. J., Sherratt, A. L., Lee, S. M., Smith, R. J., Senner, C., et al. (2008). Cytological and transcript analyses reveal fat and lazy persister-like bacilli in tuberculous sputum. *PLoS Med, 5*(4), e75. doi:10.1371/journal.pmed.0050075.

Gerdes, K., Bech, F. W., Jorgensen, S. T., Lobner-Olesen, A., Rasmussen, P. B., Atlung, T., et al. (1986). Mechanism of postsegregational killing by the hok gene product of the parB system of plasmid R1 and its homology with the relF gene product of the E. coli relB operon. *EMBO Journal, 5*(8), 2023–2029.

Gotfredsen, M., & Gerdes, K. (1998). The Escherichia coli relBE genes belong to a new toxin-antitoxin gene family. *Molecular Microbiology, 29*(4), 1065–1076.

Gurnev, P. A., Ortenberg, R., Dorr, T., Lewis, K., Bezrukov, S. M. (2012). Persister-promoting bacterialtoxin TisB produces anion-selective pores in planar lipid bilayers. FEBS letters *586*(16), 2529–2534. doi:10.1016/j.febslet.2012.06.021

Hansen, S., Lewis, K., & Vulic, M. (2008). Role of global regulators and nucleotide metabolism in antibiotic tolerance in Escherichia coli. *Antimicrobial Agents and Chemotherapy, 52*(8), 2718–2726. doi:AAC.00144-0810.1128/AAC.00144-08. [pii].

Hansen, S., Vulic, M., Min, J., Yen, T. J., Schumacher, M. A., Brennan, R. G., Lewis, K. (2012). Regulation of the Escherichia coli HipBA toxin-antitoxin system by proteolysis. PLoS ONE *7*(6), e39185.doi:10.1371/journal.pone.0039185

Harrison, J. J., Wade, W. D., Akierman, S., Vacchi-Suzzi, C., Stremick, C. A., Turner, R. J., et al. (2009). The chromosomal toxin gene yafQ is a determinant of multidrug tolerance for Escherichia coli growing in a biofilm. *Antimicrobial Agents and Chemotherapy, 53*(6), 2253–2258. doi:10.1128/AAC.00043-09.

Hooper, D. (2001). Mechanism of action of antimicrobials: Focus on fluoroquinolones. *Clinical Infectious Diseases, 32*, S9–S15.

Jensen, R. B., & Gerdes, K. (1995). Programmed cell death in bacteria: proteic plasmid stabilization systems. *Molecular Microbiology, 17*(2), 205–210.

Kawano, M., Aravind, L., & Storz, G. (2007). An antisense RNA controls synthesis of an SOS-induced toxin evolved from an antitoxin. *Molecular Microbiology, 64*(3), 738–754. doi:10.1111/j.1365-2958.2007.05688.x.

Keren, I., Minami, S., Rubin, E., & Lewis, K. (2011). Characterization and transcriptome analysis of mycobacterium tuberculosis persisters. *mBio, 2*(3), e00100–e00111. doi:10.1128/mBio.00100-11.

Keren, I., Shah, D., Spoering, A., Kaldalu, N., & Lewis, K. (2004). Specialized persister cells and the mechanism of multidrug tolerance in *Escherichia coli*. *Journal of Bacteriology, 186*(24), 8172–8180.

Kim, Y., & Wood, T. K. (2010). Toxins Hha and CspD and small RNA regulator Hfq are involved in persister cell formation through MqsR in Escherichia coli. *Biochemical and Biophysical Research Communications, 391*(1), 209–213. doi:10.1016/j.bbrc.2009.11.033.

Korch, S. B., Henderson, T. A., & Hill, T. M. (2003). Characterization of the *hipA7* allele of *Escherichia coli* and evidence that high persistence is governed by (p)ppGpp synthesis. *Molecular Microbiology, 50*(4), 1199–1213.

Korch, S. B., & Hill, T. M. (2006). Ectopic overexpression of wild-type and mutant *hipA* genes in *Escherichia coli*: Effects on macromolecular synthesis and persister formation. *Journal of Bacteriology, 188*(11), 3826–3836.

Lafleur, M. D., Qi, Q., & Lewis, K. (2010). Patients with long-term oral carriage harbor high-persister mutants of Candida albicans. *Antimicrobial Agents and Chemotherapy, 54*(1), 39–44. doi:10.1128/AAC.00860-09.

Lewis, K. (2007). Persister cells, dormancy and infectious disease. *Nature Reviews, 5*(1), 48–56. doi:10.1038/nrmicro1557.

Lewis, K. (2010). Persister cells. *Annual Review of Microbiology, 64*, 357–372. doi:10.1146/annurev.micro.112408.134306.

Lippmann, C., Lindschau, C., Vijgenboom, E., Schroder, W., Bosch, L., & Erdmann, V. A. (1993). Prokaryotic elongation factor Tu is phosphorylated in vivo. *Journal of Biological Chemistry, 268*(1), 601–607.

Maisonneuve, E., Shakespeare, L. J., Jorgensen, M. G., & Gerdes, K. (2011). Bacterial persistence by RNA endonucleases. *Proceedings of the National academy of Sciences of the USA, 108*(32), 13206–13211. doi:10.1073/pnas.1100186108.

Masuda, Y., Miyakawa, K., Nishimura, Y., & Ohtsubo, E. (1993). chpA and chpB, Escherichia coli chromosomal homologs of the pem locus responsible for stable maintenance of plasmid R100. *Journal of Bacteriology, 175*(21), 6850–6856.

McKenzie, G. J., Magner, D. B., Lee, P. L., & Rosenberg, S. M. (2003). The dinB operon and spontaneous mutation in Escherichia coli. *Journal of Bacteriology, 185*(13), 3972–3977.

Motiejunaite, R., Armalyte, J., Markuckas, A., & Suziedeliene, E. (2007). Escherichia coli dinJ-yafQ genes act as a toxin-antitoxin module. *FEMS Microbiology Letters, 268*(1), 112–119. doi:10.1111/j.1574-6968.2006.00563.x.

Moyed, H. S., & Bertrand, K. P. (1983). *hipA*, a newly recognized gene of *Escherichia coli* K-12 that affects frequency of persistence after inhibition of murein synthesis. *Journal of Bacteriology, 155*(2), 768–775.

Mulcahy, L. R., Burns, J. L., Lory, S., & Lewis, K. (2010). Emergence of Pseudomonas aeruginosa strains producing high levels of persister cells in patients with cystic fibrosis. *Journal of Bacteriology, 192*(23), 6191–6199. doi:10.1128/JB.01651-09.

Pedersen, K., Christensen, S. K., & Gerdes, K. (2002). Rapid induction and reversal of a bacteriostatic condition by controlled expression of toxins and antitoxins. *Molecular Microbiology, 45*(2), 501–510.

Pedersen, K., & Gerdes, K. (1999). Multiple hok genes on the chromosome of Escherichia coli. *Molecular Microbiology, 32*(5), 1090–1102.

Phillips, I., Culebras, E., Moreno, F., & Baquero, F. (1987). Induction of the SOS response by new 4-quinolones. *Journal of Antimicrobial Chemotherapy, 20*(5), 631–638.

Ramage, H. R., Connolly, L. E., & Cox, J. S. (2009). Comprehensive functional analysis of mycobacterium tuberculosis toxin-antitoxin systems: implications for pathogenesis, stress responses, and evolution. *PLoS Genetics, 5*(12), e1000767. doi:10.1371/journal.pgen.1000767.

Rotem, E., Loinger, A., Ronin, I., Levin-Reisman, I., Gabay, C., Shoresh, N., et al. (2010). Regulation of phenotypic variability by a threshold-based mechanism underlies bacterial persistence. *Proceedings of the National academy of Sciences of the United States of America, 107*(28), 12541–12546.

Sahl, H. G., & Bierbaum, G. (1998). Lantibiotics: Biosynthesis and biological activities of uniquely modified peptides from gram-positive bacteria. *Annual Review of Microbiology, 52*, 41–79. doi:10.1146/annurev.micro.52.1.41.

Scherrer, R., & Moyed, H. S. (1988). Conditional impairment of cell division and altered lethality in *hipA* mutants of *Escherichia coli* K-12. *Journal of Bacteriology, 170*(8), 3321–3326.

Schumacher, M. A., Piro, K. M., Xu, W., Hansen, S., Lewis, K., & Brennan, R. G. (2009). Molecular mechanisms of *HipA*-mediated multidrug tolerance and its neutralization by *HipB*. *Science, 323*(5912), 396–401.

Shah, D., Zhang, Z., Khodursky, A., Kaldalu, N., Kurg, K., & Lewis, K. (2006). Persisters: A distinct physiological state of *E. coli*. *BMC Microbiology, 6*(1), 53–61.

Singletary, L. A., Gibson, J. L., Tanner, E. J., McKenzie, G. J., Lee, P. L., Gonzalez, C., et al. (2009). An SOS-regulated type 2 toxin-antitoxin system. *Journal of Bacteriology, 191*(24), 7456–7465. doi:10.1128/JB.00963-09.

Spoering, A. L., & Lewis, K. (2001). Biofilms and planktonic cells of *Pseudomonas aeruginosa* have similar resistance to killing by antimicrobials. *Journal of Bacteriology, 183*(23), 6746–6751.

Tian, Q. B., Hayashi, T., Murata, T., & Terawaki, Y. (1996). Gene product identification and promoter analysis of hig locus of plasmid Rts1. *Biochemical and Biophysical Research Communications, 225*(2), 679–684. doi:10.1006/bbrc.1996.1229.

Unoson, C., & Wagner, E. G. (2008). A small SOS-induced toxin is targeted against the inner membrane in Escherichia coli. *Molecular Microbiology, 70*(1), 258–270. doi:10.1111/j.1365-2958.2008.06416.x.

Vazquez-Laslop, N., Lee, H., & Neyfakh, A. A. (2006). Increased persistence in *Escherichia coli* caused by controlled expression of toxins or other unrelated proteins. *Journal of Bacteriology, 188*(10), 3494–3497.

Vogel, J., Argaman, L., Wagner, E. G., & Altuvia, S. (2004). The small RNA IstR inhibits synthesis of an SOS-induced toxic peptide. *Current Biology, 14*(24), 2271–2276. doi:10.1016/j.cub.2004.12.003.

Wayne, L. G., & Hayes, L. G. (1996). An in vitro model for sequential study of shiftdown of mycobacterium tuberculosis through two stages of nonreplicating persistence. *Infection and Immunity, 64*(6), 2062–2069.

Wayne, L. G., & Sohaskey, C. D. (2001). Nonreplicating persistence of mycobacterium tuberculosis. *Annual Review of Microbiology, 55*, 139–163. doi:10.1146/annurev.micro.55.1.139.

Wolfson, J. S., Hooper, D. C., McHugh, G. L., Bozza, M. A., & Swartz, M. N. (1990). Mutants of Escherichia coli K-12 exhibiting reduced killing by both quinolone and beta-lactam antimicrobial agents. *Antimicrobial Agents and Chemotherapy, 34*(10), 1938–1943.

Chapter 12
Type II Toxin-Antitoxin Loci: The *Epsilon/zeta* Family

Hannes Mutschler and Anton Meinhart

Abstract Epsilon/zeta is a widespread TA gene family, members of which stabilise resistance plasmids in Gram-positive and -negative bacteria. Additionally, chromosomally encoded *epsilon/zeta* loci are virulence determinants in highly pathogenic *Streptococcus pneumoniae* strains. Here, we provide an overview of the unique mechanism of cell-poisoning by the toxin component, toxin inhibition by antitoxin, regulation of TA protein expression and possible biological functions of this system apart from plasmid maintenance.

12.1 Introduction

Members of the *epsilon/zeta* family are type II TA loci that predominately prevail on plasmids and chromosomes of pathogenic bacteria. Whilst members of this TA family were initially identified as plasmid maintenance systems that ensure stable inheritance similar to classical addiction modules (Ceglowski et al. 1993a, b), recent data imply that chromosomally encoded systems participate in more complex cellular and host-specific functions such as modulation of virulence (Brown et al. 2004; Harvey et al. 2011). The key to this versatility seems to lie not only in the very distinct modes of toxin activation, but also in the remarkable functional mechanism by which zeta toxins evoke programmed cell death (PCD): whereas toxins from other type II TA systems are known to interfere with translation or replication, zeta toxins are the first members that directly target bacterial cell wall synthesis. Dependent on the

H. Mutschler · A. Meinhart (✉)
Department of Biomolecular Mechanisms, Max Planck Institute for Medical Research, Jahnstr. 29, 69120, Heidelberg, Germany
e-mail: anton.meinhart@mpimf-heidelberg.mpg.de

K. Gerdes (ed.), *Prokaryotic Toxin-Antitoxins*,
DOI: 10.1007/978-3-642-33253-1_12,

physiological state of bacteria, zeta poisoning can either evoke a lytic or bacteriostatic phenotype that is reminiscent to that observed for antibiotics targeting cell wall synthesis. In the following chapter, we will discuss the current level of knowledge about regulation of this unique TA family as well as the functional mechanism of zeta toxins action and their potential biological implications.

12.2 The Spread of *Epsilon/Zeta* Loci

12.2.1 Plasmid Encoded Epsilon/Zeta Loci

Epsilon/zeta systems were initially discovered on the low-copy-number plasmid pSM19035 (Ceglowski et al. 1993a, b), which was found in a clinical isolate of *Streptococcus pyogenes* that was reported first in 1972 (Dixon and Lipinski 1972). Notably, this source strain was the first reported multi-drug resistant streptococcal serotype, since pSM19035 was found to confer *S. pyogenes* with resistances to high levels of erythromycin and lincomycins. Plasmid pSM19035 belongs to the *inc*18 group of resistance factors that despite their low copy number (2–5 plasmid molecules per cell (Clewell 1981)) were found to be >1,000-fold more stable than one would predict assuming that random segregation is responsible for plasmid inheritance (Ceglowski et al. 1993a, b). Since then, pSM19035 and its different ORFs (which have been named with Greek letters from α to ζ) were used as model system to study stable inheritance of conjugable low-copy-number plasmids.

Stable maintenance of pSM19035 is ensured by different segregation (*seg*) loci that are encoded on two inverted repeats and whose gene products control plasmid copy number, improve plasmid partitioning after replication, ensure a better-than-random *seg* during cell division, and evoke post-segregational killing (PSK; for a detailed review see (Lioy et al. 2010)). Among the different *seg* loci, mainly the *segB1* and *segB2* loci were found to encode the key-determinants responsible for plasmid stability (Ceglowski et al. 1993b). The *segB2* locus leads to a 50-fold increase in stability of pSM19035 (Pratto et al. 2008) and encodes for a partitioning system that is considered to be the bacterial analogue of the eukaryotic centromere. Furthermore, Alonso and co-workers found that the *segB1* locus is the main effector required for the observed stable inheritance of pSM19035 in *Bacillus subtilis* (Ceglowski et al. 1993b). *segB1* encodes for a transcriptional repressor, ω, which is a global regulator of pSM19035 (de la Hoz et al. 2000), and the bicistronic operon ε/ζ (Fig. 12.1a). Removal of a 3 kb segment encoding for the ε and ζ ORFs strongly increased the rate of plasmid loss per generation (Ceglowski et al. 1993a). Later work revealed that the ε/ζ locus encodes for a genuine type II TA system, in which the unstable 10 kDa antitoxin ε inactivates ζ, a 32 kDa protein with an unknown cytotoxic activity (Camacho et al. 2002; Meinhart et al. 2003; Zielenkiewicz and Ceglowski 2001, 2005). Due to their first discovery on pSM19035, the Greek letters ε/ζ have been used eponymously for all plasmid encoded members of this TA family.

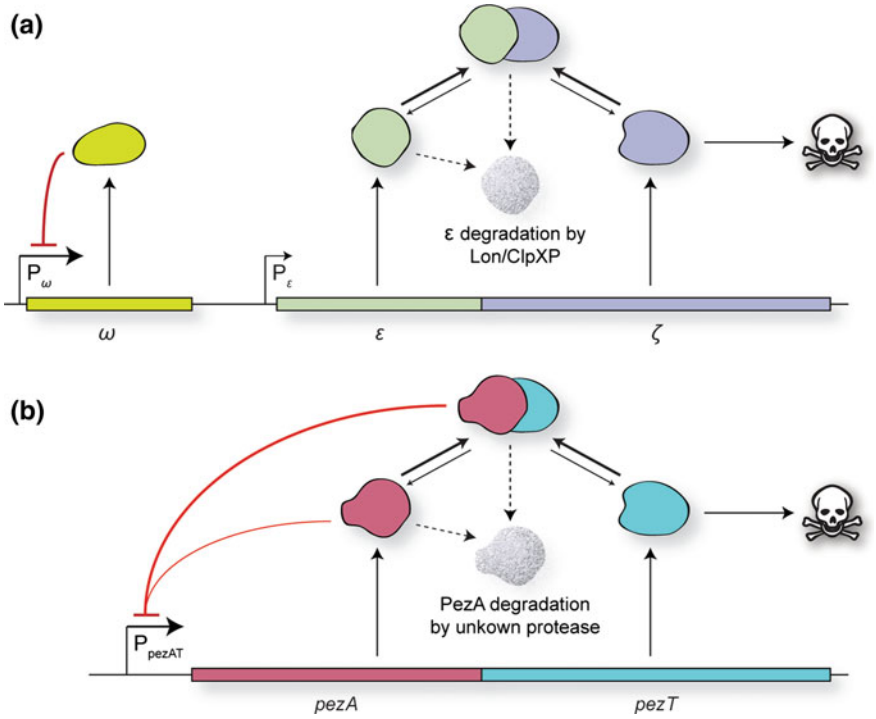


Fig. 12.1 Genetic organisation of the ω-ε-ζ and the *pezAT* locus. **a** Architecture of the *segB1* locus that was found to stabilise the multi-resistance plasmid pSM19035 in *Streptococcus pyogenes*. The locus contains two discrete promoters P_ω and P_ε. P_ω regulates synthesis of either the ω, the ω/ε or the ω/ε/ζ transcripts and is autorepressed by ω the repressor (*yellow*) and P_ε is a weak constitutive promoter. Whereas the ε antitoxin (*green*) and ζ toxin (*blue*) form an inactive complex after translation, the ω repressor does not interact either with ε or with ζ on a protein level. *segB1* is thought to sense low plasmid levels through the concentration of ω protein. At low ω protein levels, more ε and ζ proteins are produced, which establishes plasmid addiction. In case the plasmid is lost, PSK is achieved by release of the ζ protein from the ε/ζ protein complex through continuous degradation of the ε protein by the Lon and ClpXP proteases. **b** Architecture of the *pezAT* operon from *S. pneumoniae*. In contrast to the *segB1* system, no ORF for a individual ω-like transcriptional repressor is present. Autoregulation of the operon is achieved by the repressor and antitoxin PezA (*red*) together with the toxin PezT (*cyan*) acting as co-repressor. Like ε/ζ proteins, both proteins form a stable complex after translation. When and how the toxin PezT is released from the complex by PezA degradation is unknown

We suggest using "*epsilon/zeta*" as the umbrella term for all *epsilon/zeta* TA systems, independently of whether they are chromosomal or plasmid borne and to restrict the usage of Greek letters exclusively to plasmid-encoded members.

Homologues of *segB1* were also found on conjugative broad-host-range plasmids that are partially related to the *inc*18 incompatibility group such as pIP501, pRE25 (Rosvoll et al. 2010; Schwarz et al. 2001), pIP402, pAMβ1 (Brantl et al. 1990), pBM407 (Holden et al. 2009) and plasmids of the pVEF

series (Sletvold et al. 2007, 2008). Although plasmid-encoded *epsilon/zeta* systems were presumed to be confined to Gram-positive bacteria, ε/ζ homologues were also recently identified on plasmids of Gram-negative bacteria such as IncP1 plasmids from *Neisseria gonorrhoeae* (Pachulec and van der Does 2010; Van Melderen and Saavedra De Bast 2009). Interestingly, in the tetracycline determinant *tetM* of these plasmids, two different ε/ζ loci were identified, whilst no ORF for a ω transcriptional repressor was found. This implies that the ε/ζ system from IncP1 plasmids are regulated differently than the ones found in *inc*18-like plasmids.

12.2.2 Chromosomally Encoded Epsilon/Zeta Systems

In addition to the plasmid encoded ε/ζ systems, chromosomally encoded *epsilon/zeta* homologues were found on different integrative elements found in *Streptococcus spp*. (Brown et al. 2001; Croucher et al. 2009). The first chromosomal locus described is generally known as *pezAT* (pneumococcal epsilon zeta antitoxin toxin and originally named SP1050/SP1051) and was identified on the Pneumococcal Pathogenicity Island 1 (PPI-1), a 27 kb genomic island found in the chromosome of pathogenic strains of *S. pneumoniae*. PPI-1 positive strains show increased pathogenicity, consistent with the fact that PPI-1 encodes for multiple putative virulence factors (Tettelin et al. 2001). Furthermore, pathogenicity seems to correlate with the gene content of a recombination hot-spot region termed PPI-1v at the 3'-end of PPI-1. Notably, PPI-1v derivatives of highly virulent pneumococcal isolates universally contain the *pezAT* operon, whereas the other gene content is highly variable (Harvey et al. 2011). In contrast, *pezAT* was not observed in intermediately virulent or non-invasive clinical isolates.

PezT and ζ from pSM19035 share 42 % amino acid sequence identity, and similar to ζ, PezT was found to cause growth arrest and PCD in *Escherichia coli* in absence of its antitoxin PezA (Khoo et al. 2007). The antitoxin PezA is a multi-domain protein whose carboxy-terminal three-helix bundle domain shows 21 % amino acid sequence identity to ε from pSM19035. In contrast to the ω-ε-ζ system, *pezAT* is a genuine bicistronic operon, in which the antitoxin domain of PezA is fused with an additional amino-terminal helix-turn-helix DNA binding domain that acts as a repressor for transcription from *pezAT* (Fig. 12.1b) (Khoo et al. 2007). Homologues of *pezAT* were also identified on the integrative conjugable elements (ICE) ICE*Sp*23FST81 of the multi-drug resistant pandemic clone *S. pneumoniae*$^{\text{Spain23F}}$ ST81 (Croucher et al. 2009) and a general cross-species distribution is indicated by its presence on genetic elements found in *Streptococcus suis* and *Streptococcus agalactiae* (Croucher et al. 2009). A broad phylogenetic spread of *epsilon/zeta* systems is further supported by a recent bioinformatic analysis based on the 'guilt by association' principle by van Melderen and co-workers which suggests that *epsilon/zeta* systems are widespread in all bacterial

kingdoms, and strikingly mostly reside on bacterial chromosomes rather than on plasmids (Leplae et al. 2011).

12.3 Mechanism of Zeta Toxin Action

Whereas the mechanism of action of most other TA toxins was known, the bacteriotoxic activity of zeta toxins had remained elusive until very recently. One of the main hurdles for predicting the activity of this toxic protein was that neither the primary amino acid sequence nor any tertiary structure provided sufficient information to formulate detailed predictions of its toxic activity. Nevertheless, the identification of a genuine Walker A amino acid sequence motif, which is generally required for ATP/GTP binding, suggested that the toxic mechanism of zeta toxins is distinct from other TA toxins. In fact, the three-dimensional crystal structures of the ε/ζ complex from *S. pyogenes* and the PezA/PezT complex from *S. pneumoniae* as well as mutational studies implied that zeta toxins are kinases that phosphorylate an unknown target using ATP or GTP (Khoo et al. 2007; Meinhart et al. 2003; Nowakowska et al. 2005). This hypothesis was based on the tertiary structure of zeta toxins that shows striking similarities to kinases such as the nucleoside monophosphate kinase family, polynucleotide kinases and chloramphenicol phosphotransferases (Fig. 12.2a). Indeed, structure based, site-directed mutations in the putative active site resulted in inactive toxin variants (Khoo et al. 2007; Meinhart et al. 2003). A requirement of ATP/GTP binding for zeta toxins functional mechanism was further supported by the mode of toxin inactivation by the antitoxin: in the ε/ζ as well as in the PezA/PezT complex, the antitoxin inactivates the toxin by occlusion of the ATP/GTP-binding site (Khoo et al. 2007; Meinhart et al. 2003).

First instructive attempts to decipher the functional mechanism of ζ toxins were made by Zielenkiewicz and Ceglowski who showed that ζ is bactericidic in the Gram-positive *B. subtilis,* but causes growth arrest and filamentation in the Gram-negative *E. coli* (Zielenkiewicz and Ceglowski 2005). Later, Lioy et al. showed that a less toxic variant of ζ (ζY83C), which can be cloned in absence of ε antitoxin, inhibits cell proliferation and causes cell lysis in a subpopulation of *B. subtilis* (Lioy et al. 2006). In the latter report, a pleiotropic effect of ζ expression in *E. coli* that led to inhibition of DNA-replication, transcription and translation without specifically targeting any of these cellular events was described.

A structure-based educated guess enabled us to construct an attenuated variant of the ζ homologue PezT that made detailed studies of the phenotype of zeta toxin overexpression in *E. coli* feasible (Mutschler et al. 2011). The delayed nature of the phenotype allowed us to finally identify the metabolic process that is poisoned by zeta toxins as well as the substrate molecule of their predicted kinase activity. Expression of the carboxy-terminally truncated PezT variant in liquid culture caused cells to arrest during cell division followed by formation of membrane bulges at their septum and eventually caused cell lysis (Mutschler et al. 2011). In

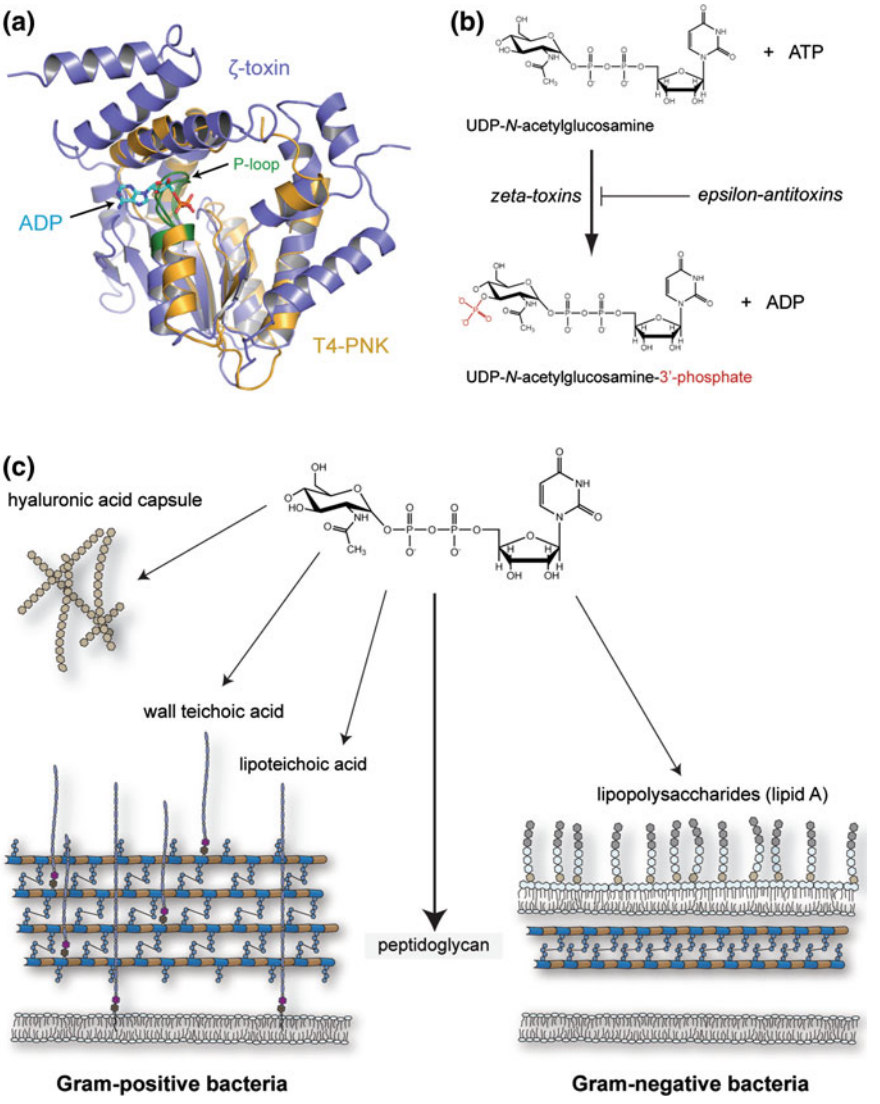


Fig. 12.2 Zeta toxins phosphorylate the fundamental metabolite UNAG. **a** The α/β-fold of zeta toxins (PDB ID: 1GVN and shown as ribbon model coloured in *blue*) resembles kinases such as T4-polynucleotide kinase (PDB ID: 1LTQ and shown as ribbon model coloured in *orange*). The kinase domain of the T4 polynucleotide kinase bound to ADP was superposed with the ζ toxin. The characteristic P-loop motifs of T4-polynucleotide kinase and ζ superimpose well (highlighted in *green*), which suggests a similar NTP binding mode for the toxin. **b** Zeta toxins phosphorylate the nucleotide sugar UDP-*N*-acetylglucosamine (UNAG) at the 3′-hydroxyl group of the *N*-acetylglucosamine moiety using ATP as phosphate donor. **c** UNAG is an essential precursor for synthesis of various polysaccharides that are part of the bacterial cell wall

contrast, cells adherently growing on LB agar plates showed severe morphological changes and cells adopted a bloated sphere-like shape upon toxin expression (Mutschler et al. 2011). Both observations suggested that dramatic changes in the peptidoglycan layer that usually confines the rod-like morphology of *E. coli* cells were caused by zeta toxins activity. Additionally, the phenotype closely resembled that which is commonly observed upon treatment with cell wall synthesis targeting antibiotics such as ampicillin.

Based on these observations, we ultimately showed that zeta toxins convert the essential cell wall precursor UDP-*N*-acetylglucosamine (UNAG) to UDP-*N*-acetyl glucosamine-3'-phosphate (UNAG-3P) using ATP, both in vitro and in vivo (Mutschler et al. 2011) (Fig. 12.2b). UNAG is a central metabolite required for peptidoglycan synthesis and various other glycosyl transfer reactions (Fig. 12.2c). UNAG is in particular essential for formation of *N*-acetylglucosamine-*N*-acetyl-muramic acid (muropeptide), the basic disaccharide unit of the peptidoglycan scaffold. The cytosolic biosynthetic pathway of muropeptide synthesis consists of eight distinct steps in *E. coli*, during the course of which a pentapeptide is linked to the 3'-hydroxyl group of the *N*-acetylglucosamine moiety of UNAG (Fig. 12.3) (Barreteau et al. 2008). Therefore, phosphorylation of UNAG at this hydroxyl group by zeta toxins converts this important building block into a metabolite unusable for muropeptide synthesis. Zeta toxin induced PCD results not only from the depletion of the pool of UNAG molecules in the cytosol, but also from UNAG-3P-mediated competitive inhibition of MurA, the enzyme that catalyses the first committed step in peptidoglycan synthesis (Fig. 12.3) (Mutschler et al. 2011). The entire cellular pathway of muropeptide synthesis and in particular the MurA enzymes themselves are conserved among Gram-negative and -positive bacteria (Du et al. 2000). In a recent attempt, we could show that similar to MurA from *E. coli* the homologous proteins from *S. pneumoniae* are inhibited by UNAG-3P as well (unpublished results). In agreement with our finding in *E. coli*, a recent detailed study by Alonso and co-workers also showed that low expression levels of ζ in *B. subtilis* evoke a very similar phenotype to the one we observed in *E. coli* (Lioy et al. 2012): after an early onset, synthesis of macromolecules ceased, ATP and GTP synthesis was reduced, and membrane permeability increased. Interestingly, the same study also showed that a very small subpopulation of cells survived ζ poisoning and seemed to be tolerant towards expression of ζ. Apparently, the toxicity of zeta toxins is strongly dependent on the initial growth rate of poisoned cells also in the Gram-positive bacterium *B. subtilis*, as observed for the Gram-negative *E. coli* (Mutschler et al. 2011): whereas cells with high growth rates and consequently a large demand on peptidoglycan precursors are very susceptible towards UNAG-3P formation, slowly dividing cells can survive PezT expression for a considerable amount of time. Interestingly, a significant number of cells that were poisoned by ζ for 120 min could still be rescued by ε antitoxin expression, showing that some cells can tolerate UNAG-3P production for a considerable amount of time. However, it remains to be verified whether plasmid-borne ζ-toxins

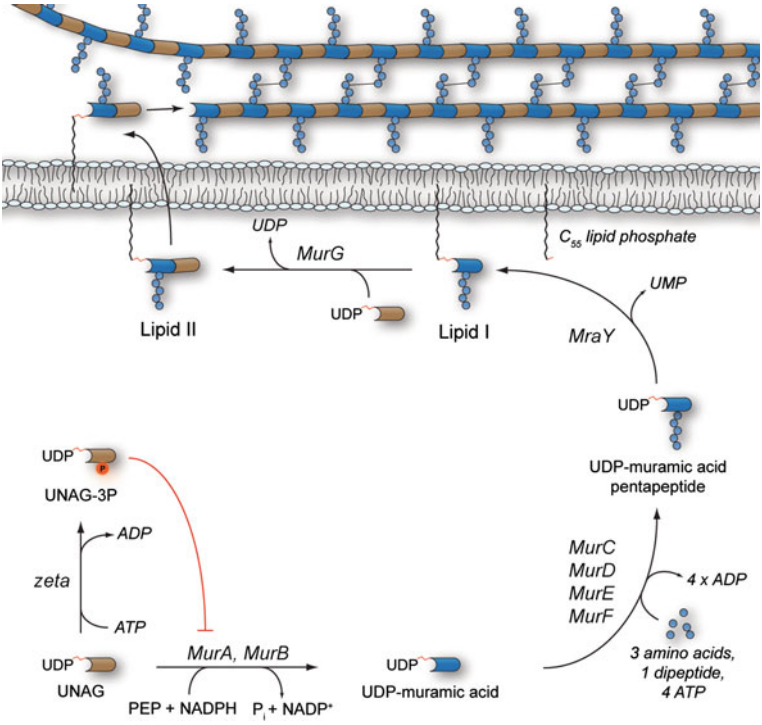


Fig. 12.3 Mechanism of cell poisoning by the UNAG-kinase reaction of zeta toxins. Zeta toxins convert UNAG into UNAG-3P using ATP. In contrast to UNAG, UNAG-3P cannot be utilised for peptidoglycan synthesis, since the 3'-hydroxyl of the *N*-acetylglucosamine moiety is blocked by the phosphate. Additionally, UNAG-3P directly inhibits peptidoglycan synthesis through formation of a reversible non-productive complex with MurA. This inhibition leads to an imbalance of murein degradation and resynthesis, which eventually causes lysis due to the high internal turgor pressure of bacterial cells. In absence of free zeta toxins, UNAG is converted into UDP-muramic acid by the enzymes MurA and MurB. Muramic acid (*blue*) is then used as acceptor for the addition of a five amino acids long peptide by four enzymes (MurC-MurF). The sequence of this muropeptide is species dependent. Subsequently, peptidyl-muramic acid is transferred from UDP to undecaprenol phosphate forming lipid I. The glycosyltransferase MurG finishes the cytosolic part of peptidoglycan synthesis by the linkage of an additional *N*-acetylglucosamine with lipid I and forming lipid II. This lipopolysaccharide is transferred to the outer membrane layer where different glycotransferases add the disaccharide unit of lipid II to the existing peptidoglycan macromolecule. During this process, the muropeptides of the different layers are cross-linked which strongly increases the stability of the bacterial sacculus

poisonous activity serves any additional cellular processes such as induction of dormant states or persistence aside from plasmid maintenance as proposed recently (Lioy et al. 2006, 2012). It is also plausible that ζ-toxins simply select for already dormant or persistent cells. In conclusion, zeta toxins are the first TA toxins known to poison bacterial cell wall synthesis. Moreover, zeta toxins are the first TA toxins that target a small molecule metabolite.

12.4 Toxin Inhibition and Substrate Binding

12.4.1 Toxin Inhibition

The crystal structure of the toxin–antitoxin ε/ζ complex from pSM19035 as well as the PezAT complex revealed that both *epsilon/zeta* systems form a heterotetrameric arrangement in which two toxin molecules bind to a central dimeric three-helix bundle of antitoxins (Fig. 12.4a, b) (Khoo et al. 2007; Meinhart et al. 2001, 2003). Toxin inhibition is mainly performed by the amino-terminal helix of the antitoxin that covers the nucleotide binding site of the toxin (Fig. 12.4c). In common, ε as well as PezA insert four side chain groups – a tyrosine, a phenylalanine (ε)/leucine (PezA) and two glutamate moieties into the ATP-binding pocket of their cognate toxins and thereby occlude the binding site. Interestingly, most other residues that are part of the TA interface are not conserved (Khoo et al. 2007). This suggests that the main evolutionary constraint on the antitoxin is to abolish the toxin activity, but residues involved in protein–protein interaction can diverge and the TA binding affinity is thereby modulated.

12.4.2 Toxin Catalysis and Substrate Specificity

Our current structural understanding of zeta toxins substrate specificity and the mechanism of enzyme catalysis is restricted to the structure of the ε/ζ-UNAG complex, in which the ATP-binding site of ζ is occluded by the antitoxin ε (Mutschler et al. 2011) (Fig. 12.4d). However, based on the structural similarity of the active site architecture of zeta toxins with nucleotide kinases, predictions concerning the mechanism of enzyme catalysis were possible and have been verified by site-directed mutagenesis studies (Khoo et al. 2007; Meinhart et al. 2003). Similarly to all related phosphoryltransferases, the Walker-A motif containing P-loop of zeta toxins will coordinate the β-phosphate group of an ATP molecule. Side chains of the degenerate Walker B motif, which is also commonly found in kinases that phosphorylate low-molecular-weight substrates (Leipe et al. 2003), will coordinate the catalytically important divalent metal ion, which most likely is Mg^{2+}. In all zeta toxins, an invariant aspartate residue (Asp-66 in PezT and Asp-67 in ζ, Fig. 12.4d) most likely deprotonates the 3′-hydroxyl group to form an oxyanion that performs the nucleophilic attack on the γ-phosphate group of ATP. In a so called "inline mechanism" this group will then be transferred onto the 3'-oxyanion of UNAG to form UNAG-3P. Furthermore, positively charged residues such as arginine side chains are positioned to compensate for emerging negative charges occurring during the phosphoryltransfer reaction.

In contrast to the ATP-binding site and the site where the phosphoryltransfer reaction takes place, the UNAG binding site of zeta toxins is unique. Residues that form the UNAG binding cleft are highly conserved in all zeta toxins that have been

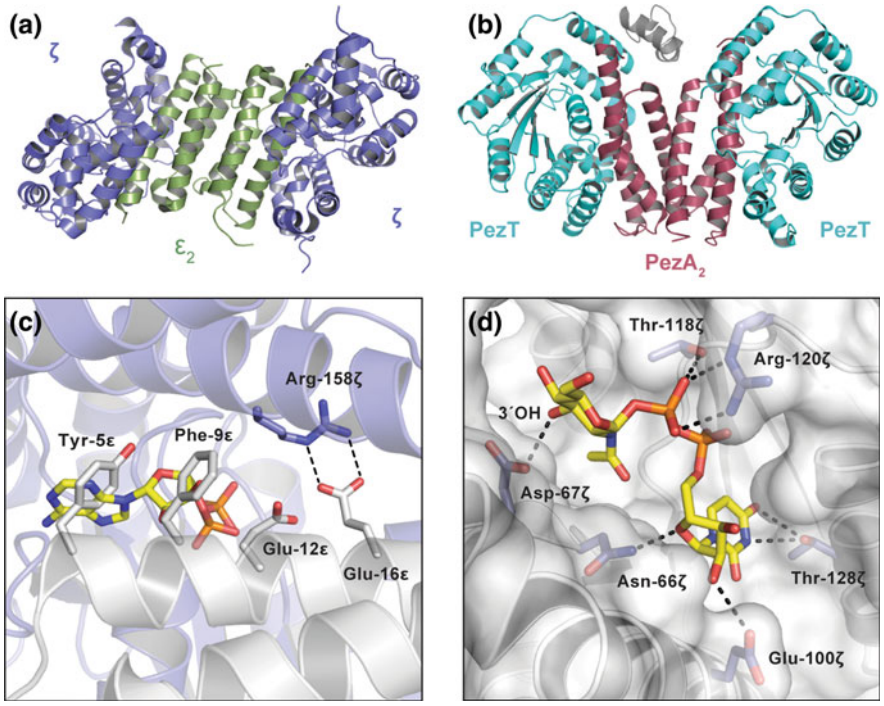


Fig. 12.4 Mechanism of zeta toxin inhibition by heterotetramer formation. **a** Architecture of the heterotetrameric ε/ζ complex from *S. pyogenes* (PDB ID: 1GVN). The ε-antitoxin is a three-helix-bundle, whose homodimer (shown as ribbon model in *green*) form binds two ζ-toxins (shown as ribbon model in *blue*) at each side. **b** The overall architecture of the analogous PezAT complex from *S. pneumoniae* (PDB ID: 2P5T). Similar as in the *S. pyogenes* ε/ζ complex, the C-terminal domain of the antitoxin PezA forms a homodimer (shown as ribbon model in *red*) that inhibits two PezT toxins (shown as ribbon model in *cyan*). Note that only one amino-terminal helix-turn-helix motif of PezA (shown as ribbon model in *grey*) could be identified in the structure. However, the linker to a PezA three-helix-bundle is missing. **c** Epsilon-antitoxins block the ATP-binding site of zeta toxins exemplified by the ε/ζ complex structure. The predominant contacts between ε antitoxin (shown as ribbon model in *grey*) and ζ toxin (shown as ribbon model in *blue*) involve the amino-terminal helix of ε antitoxin. Tyr-5 and Phe-9 of ε antitoxin point into the nucleotide binding pocket and impede binding of the nucleoside moiety. Glu-12 and Glu-16 of ε antitoxin establish a repulsive negative potential impairing phosphate binding. Moreover, in the complex, the side chain of Glu-16 of ε antitoxin forms hydrogen bonds and salt bridges to Arg-158 of ζ toxin, which is involved in enzyme catalysis in free zeta toxins. **d** A transparent molecular surface representation of ζ toxin with a ribbon representation shown underneath. UNAG (shown as a stick model) is embedded in a deep cleft. Hydrogen bonds relevant for substrate binding are illustrated as *black* dashed lines. Residues of ζ toxin that are important for substrate binding are depicted as stick model. Tyr-118 and Arg-120 from the zeta-toxin specific GTXR binding motif recognise the diphosphate of the UDP moiety. Asn-66 and Glu-100 form specific contacts with the ribose moiety, whilst Thr-128 forms hydrogen bonds to the uracil base that otherwise binds into a hydrophobic cleft. Asp-67 of ζ is thought to deprotonate the 3'-hydroxylgroup of the N-acetylglucosamine moiety prior to nucleophilic attack on the γ-phosphate group of ATP

identified so far (Mutschler and Meinhart 2011). In particular, a short conserved stretch of amino acids with the consensus sequence GTXR (where X can by any residue) is located immediately adjacent to the degenerate Walker B motif (Fig. 12.4d). The existence of this short motif defines the family of zeta toxins and predictions whether an ORF frame encodes for a putative UNAG-kinase or not can already be made based on the primary amino acid sequence (Mutschler and Meinhart 2011). The structure of the ζ-toxin in complex with UNAG revealed that the threonine residue (Thr-118) of this motif is essential for correct positioning of phosphate groups of the UDP moiety of UNAG within the active site. Similarly, the guanidinium group of the arginine residue (Arg-120) is required for binding the phosphate groups of UNAG and this side chain is also important for compensating their negative charge. The uracil base of UNAG is sandwiched into a hydrophobic pocket whose amino acid sequence, irrespective from its hydrophobic character, is less conserved (Mutschler et al. 2011).

Whereas the structure of the ζ-toxin UNAG complex did not provide much insight into the specificity for the sugar moiety of UNAG, biochemical experiments showed that zeta toxins are highly specific for UNAG: even chemically very closely related molecules such as UDP-glucose or UDP-galactose are not or only poorly phosphorylated (Mutschler et al. 2011).

12.5 Transcriptional Regulation of *Epsilon/Zeta* Operons

12.5.1 Regulation of Transcription of the ω-ε-ζ Operon

In contrast to most type II TA systems, expression of the ω-ε-ζ operon is not regulated by the antitoxin ε but by the ω transcriptional repressor (Ceglowski et al. 1993a; de la Hoz et al. 2000). ORFs for all three proteins are located on the *segB1* operon of pSM19035 and related plasmids with the ω locus upstream of the ε-ζ bicistron (Fig. 12.1a, de la Hoz et al. 2000). Although *segB1* contains two discrete promoter regions 5′ to the ω ORF (P_ω) and 5′ to the ε ORF (P_ε), transcription initiation mostly starts at P_ω from a σ^{70}-like promoter region (de la Hoz et al. 2000). In contrast, P_ε seems to be a weak constitutive and orphan promoter that only marginally contributes to transcription of the ε-ζ operon (de la Hoz et al. 2000). P_ω contains conserved heptad repeats in head-to-tail and head-to-head configuration to which ω binds and auto-represses synthesis of its own ω but also of ω-ε or full ω-ε-ζ transcripts (Ceglowski et al. 1993a; de la Hoz et al. 2000). The transcriptional regulator ω belongs to the MetJ/Arc repressor family featuring a ribbon-helix-helix DNA-binding motif (Murayama et al. 2001). A single ω_2 dimer binds to one heptad repeat and cooperative binding of ω_2 dimer was suggested to be achieved by polymerisation of ω_2 on arrays of repeated DNA elements (Weihofen et al. 2006).

An unsolved conundrum is the regulation of the ω-ε-ζ operon during fluctuating expression levels. Similar to conventional TA systems, the antitoxin ε was reported

to be proteolytically less stable than the ζ-toxin thereby leading to a classical PSK (Camacho et al. 2002). However, no physical interaction has been reported that could explain a cross-communication between the ω repressor and any component of the ε/ζ TA system. A recently proposed model suggests that plasmid "addiction" is achieved by direct sensing of the plasmid copy number through the expression levels of ω (Lioy et al. 2010): at low plasmid-number and thus low ω levels, repression of the P_ω promoter is reduced and consequently transcription of ε and ζ is increased. Through this mechanism, addiction to the plasmid is established and in case of total plasmid loss, the ongoing degradation of ε triggers release of ζ toxin that causes PCD of plasmid free cells. Using this mechanism, the ω-ε-ζ operon most likely confers stability also to other multi-resistance plasmids and their transposon derivatives in pathogens such as *Enterococcus faecalis* (Schwarz et al. 2001), *Enterococcus faecium* (Rosvoll et al. 2010; Sletvold et al. 2008), *S. agalactiae* (Brantl et al. 1990) and *Bacteroides spp.* (Schlesinger et al. 2007).

12.5.2 Regulation of Transcription of the pezAT *Operon*

In contrast to the ε/ζ system, transcription regulation of the *pezAT* operon is performed by the PezA antitoxin, which is a weak repressor on its own and establishes full repression only together with the PezT toxin (Fig. 12.1b): the bicistronic *pezAT* operon is transcribed from a σ^{70}-like promoter (P_{pezAT}) and repressed by binding of the PezA/PezT complex to a palindromic sequence that overlaps with the promoter region (Khoo et al. 2007). PezA contains an amino-terminal helix-turn-helix DNA binding motif as repressor domain, which is fused with the three-helix-bundle domain that binds and inhibits the PezT toxin (Khoo et al. 2007). Based on the primary structure, it is evident that ω and the repressor domain of PezA have different evolutionary origins. Moreover, upon closer inspection of the coding sequence for the linker region that connects the helix-turn-helix DNA binding with the three-helix-bundle domain we could identify an orphan Shine-Dalgarno box in appropriate spacing to a start codon within the PezA coding region. However, no translation from this internal ribosomal binding site has been observed (unpublished results). This suggests that the evolutionary origin of *pezA* is most likely a fusion event of an unrelated transcriptional repressor ORF to the 5′-end of the ORF of an ε orthologue. Notably, the nanomolar affinity leading to PezA homodimer formation and the femtomolar affinity of the PezAT interaction (Mutschler et al. 2010) imply that also $PezA_2T$ heterotrimers can be formed under conditions where PezA is in excess over PezT. It seems plausible, that regulation of transcription from the *pezAT* operon could be fine tuned by a "conditional cooperativity" mechanism similar as described for other TA systems (Overgaard et al. 2008; Garcia-Pino et al. 2010; Winther and Gerdes 2012). However, whilst it has been shown that the antitoxin homodimer $PezA_2$ is only a weak repressor (Khoo et al. 2007), it remains to be shown whether a $PezA_2T_2$ heterotetramer or a $PezA_2T$ heterotrimer differs significantly in its affinity for the regulatory DNA element of *pezAT* operon.

12.6 Zeta Toxin Release and Activation

12.6.1 The ε/ζ Locus

Release of ζ toxins from the neutralised ε/ζ complex leading to toxin activation is achieved by continuous proteolytic degradation of the ε antitoxin exceeding the rate of ζ toxin depletion. Whereas the ζ toxin has a half-life of more than 60 min in *B. subtilis*, the antitoxin ε is short lived with an average half life of only ∼18 min at near-physiological expression levels (Camacho et al. 2002). Furthermore, using *B. subtilis* chaperon-protease knockout strains, the ubiquitous ATP-dependent chaperone-proteases Lon and to a certain extent the ClpXP complex have been suggested to continuously degrade the ε antitoxin (Lioy et al. 2006). An unexpected property of the *epsilon/zeta* system, however, is their thermodynamic and proteolytic stability in vitro: unfolding studies as well as limited proteolysis experiments showed that ε was much more resistant to urea unfolding or trypsin digests than ζ and the latter is stabilised once it is protected in the TA complex (Camacho et al. 2002). It remains to be shown whether the Lon or ClpXP proteases degrade ε only in its free form or whether they can actively disassemble the ε/ζ complex. Both scenarios are plausible, since the moderate affinity of 1.1 μM for the ε-ζ interaction (Camacho et al. 2002) suggests a dynamic binding equilibrium.

A recent study by Ceglowski and co-workers, who monitored inheritance of the ω-ε-ζ system containing pBT233-1 plasmid in unrelated Gram-positive bacteria, revealed that this minimal plasmid-maintenance system of pSM19035 is sufficient for stable maintenance in most firmicutes (Brozozowska et al. 2012). This finding is in favour of a model that suggests that the broad range occurrence of pSM19035 is due to general conservation of chaperone-protease systems in Gram-positive bacteria, which can activate the ε/ζ TA system. In contrast, ω-ε-ζ was shown to be very inefficient in stabilising the same pBT233-1 plasmid in *E. coli* (Zielenkiewicz and Ceglowski 2005). This observation could either be explained by differences in the proteolytic pathways in Gram-positive and -negative bacteria or a higher tolerance of Gram-negative bacteria to poisoning of peptidoglycan synthesis by ζ toxins.

12.6.2 The pezAT *Locus*

In contrast to the plasmid-encoded ε/ζ locus, toxin release and activation of the chromosomally encoded *pezAT* has remained mostly unexplored so far. Similarly to ε/ζ, we found PezT to be less stable when compared to the free PezA or the PezAT complex in unfolding experiments in vitro (Mutschler et al. 2010). Despite these minor similarities, the kinetic properties of the PezAT complex are very distinct from those of ε/ζ. PezA and PezT proteins associate within milliseconds, which results in quasi-irreversible binding with femtomolar affinity, and thus

distinguishes PezA/PezT as one of the tightest protein–protein complexes described in the literature so far (Mutschler et al. 2010). Thus, a toxin release and activation mechanism by simple dissociation is rather unlikely under physiological conditions. Moreover, this tight binding suggests that release of PezT can only be achieved by an active disassembly by a chaperone-protease complex. Most likely, release of PezT from the PezAT complex is a host-adaptive feature, since we could not activate the toxin either in *E. coli*, or in *B. subtilis*. Further experiments using the original host environment are required to identify the mechanism and the conditions under which PezT becomes activated. However, it is conceivable that the *pezAT* gene system has been fine tuned for a certain host environment and has rather a narrow-range distribution among different bacteria.

12.7 Implications of Zeta Toxins for Virulence

While the major biological function of the ε/ζ locus seems to be stable plasmid inheritance through PSK, the role of PCD induced by chromosomally encoded PezAT systems is still unclear. Because of their prevalence on mobile streptococcal DNA-elements that have been acquired through horizontal gene transfer, a putative function such as the stabilisation of integrative elements similarly to what has been proposed for other TA systems found on mobile genetic elements seems plausible (Holden et al. 2009). Moreover, the *pezAT* locus could even be a simple selfish entity, since it usually resides in hypervariable regions that are joined together in a mosaic-like fashion. Similar selfish functions have been proposed for most other chromosomal TA systems, which apparently only secure their own stable inheritance but not that of any other proximal gene clusters (Van Melderen and Saavedra De Bast 2009). However, apart from a putative function as stable maintenance system using PSK, hints for a more complex function of PezAT-systems came from work by Brown et al. who compared clinical serotype 3 isolates of *S. pneumoniae* with PezT knockout strains for their performance in different mouse infection models (Brown et al. 2004). Notably, PezT knockout strains showed a considerably attenuated infection after intranasal inoculations of mice when compared with wild-type strains. Moreover, pneumococcal strains in which the PezT ORF had been disrupted were unable to compete with wild-type strains in mouse challenge models. Most likely, a PezAT system has a supportive, virulence-specific effect in *S. pneumoniea*, since PezT negative strains had no general growth defect in laboratory broth, serum or blood.

Additional evidence for a supporting role of PezAT in virulence came from Harvey et al. who compared the gene content of virulent serotype 1 pneumococcal isolates from indigenous Australian patients (Harvey et al. 2011). In fact, the *pezAT* ORF was exclusively contained in the recombination hot-spot PPI-1v of hypervirulent isolates, whereas it was absent in the PPI-1v of non-invasive and intermediately virulent strains. Similarly as described by Brown et al., these hypervirulent strains caused rapid lethal infection in intranasally challenged mice,

whereas the infection progress of the *pezAT* negative, intermediately virulent isolates was strongly attenuated. Most importantly, genetically modified *S. pneumoniae* D39 strains were readily outcompeted by wild-type strains in nasal lavage fluid, lungs and blood models when their *pezAT* positive PPI-1v element had been replaced by a *pezAT* negative PPI-1v element from a non-invasive isolate (Harvey et al. 2011). Along this line, replacing PPI-1v of D39 with one from another hypervirulent serotype 1 clinical isolate resulted in a hybrid strain that was as effective in infection as the wild-type strain. In conclusion, the exclusive occurrence of the *pezAT* locus in hypervirulent strains of *S. pneumoniae* strongly suggests, that PezAT contributes to virulence. Although it can not be excluded that genes other than *pezAT* within PPI-1v contribute to virulence as well, it is noteworthy that the residual gene content is highly variable (Harvey et al. 2011).

But how can the UNAG-kinase activity of PezT boost bacterial infections? One possible scenario is that PezT might be released from PezA during environmental stresses such as lack of nutrients, presence of antibiotics or the infected host immune response, similarly as observed for other TA systems (Bordes et al. 2011; Christensen et al. 2003; Ramage et al. 2009; Wang et al. 2011). The accumulation of UNAG-3P will result in inhibition of cell wall synthesis, and thus will induce lysis of a pneumococcal subpopulation. Virulence of pneumococci relies on their ability to undergo lysis under stress conditions, which is evoked either by self-induced autolysis (Pinas et al. 2008; Regev-Yochay et al. 2007) or by fratricide in which competent cells induce lysis of their non-competent sister cells (Guiral et al. 2005). Similary, PezAT could provide such an alternative pathway for autolysis. Generally, lysis of Gram-positive bacteria has great impact on virulence, since it leads to accelerated release of cellular components that are detrimental to the infected host. Such components are teichoic acids, lipoteichoic acids and bacterial DNA which are well established as virulence factors (Nau and Eiffert 2002). In addition, cell fragments from lysed pneumococci inhibit phagocytosis and impair phagocyte-mediated defense against living pneumococci (Martner et al. 2009). Most importantly, pneumococcal lysis leads to triggered release of pneumolysin (Ply), which has multiple detrimental effects for the host organism (see (Hirst et al. 2004)). Low Ply concentrations are already sufficient to trigger apoptosis (Braun et al. 2002), to activate host complement (Mitchell et al. 1991) and to induce proinflammatory reactions (Cockeran et al. 2001). At elevated concentrations, Ply causes general cellular damage in the host organism by its cholesterol-dependent ability to form large membrane pores (Gilbert et al. 1999). However, unlike other bacterial exotoxins, Ply is not directly secreted into the bacterial environment and its release relies on pneumococcal lysis after peptidoglycan degradation as, for example, caused by the major pneumococcal autolysin LytA (Martner et al. 2008). Notably, similarly to what has been described for PezT, impairment of Ply and LytA function also causes attenuated virulence in infection models (Berry et al. 1989a, b; Berry and Paton 2000; Lock et al. 1992). Thus, it is likely that PezT indeed supports pneumococcal lysis, causing cellular fragmentation and triggered release of virulence factors.

The remaining conundrum is how activation of PezT is accomplished during infection and more experimental data are required to decipher the molecular mechanisms of PezATs contribution to pneumococcal virulence. Likewise, the role of PezAT homologues in other streptococcal pathogens such as *S. agalactiae* and *S. suis* (Croucher et al. 2009) remains completely unexplored. Moreover, the presence of zeta toxin homologues in the chromosomes of pathogenic *E. coli* and *Capnocytophaga* strains raises interesting questions about a general implication of the UNAG-kinase family in virulence of Gram-negative pathogens.

References

Barreteau, H., Kovac, A., Boniface, A., Sova, M., Gobec, S., & Blanot, D. (2008). Cytoplasmic steps of peptidoglycan biosynthesis. *FEMS Microbiology Reviews, 32*, 168–207.

Berry, A. M., & Paton, J. C. (2000). Additive attenuation of virulence of *Streptococcus pneumoniae* by mutation of the genes encoding pneumolysin and other putative pneumococcal virulence proteins. *Infection and Immunity, 68*, 133–140.

Berry, A. M., Lock, R. A., Hansman, D., & Paton, J. C. (1989a). Contribution of autolysin to virulence of *Streptococcus pneumoniae*. *Infection and Immunity, 57*, 2324–2330.

Berry, A. M., Yother, J., Briles, D. E., Hansman, D., & Paton, J. C. (1989b). Reduced virulence of a defined pneumolysin-negative mutant of *Streptococcus pneumoniae*. *Infection and Immunity, 57*, 2037–2042.

Bordes, P., Cirinesi, A. M., Ummels, R., Sala, A., Sakr, S., Bitter, W., et al. (2011). SecB-like chaperone controls a toxin–antitoxin stress-responsive system in *Mycobacterium tuberculosis*. *Proceedings of the National Academy of Sciences of the United States of America, 108*, 8438–8443.

Brantl, S., Behnke, D., & Alonso, J. C. (1990). Molecular analysis of the replication region of the conjugative *Streptococcus agalactiae* plasmid pIP501 in *Bacillus subtilis*. Comparison with plasmids pAM beta 1 and pSM19035. *Nucleic Acids Research, 18*, 4783–4790.

Braun, J. S., Sublett, J. E., Freyer, D., Mitchell, T. J., Cleveland, J. L., Tuomanen, E. I., et al. (2002). Pneumococcal pneumolysin and $H_{(2)}O_{(2)}$ mediate brain cell apoptosis during meningitis. *The Journal of Clinical Investigation, 109*, 19–27.

Brown, J. S., Gilliland, S. M., & Holden, D. W. (2001). A *Streptococcus pneumoniae* pathogenicity island encoding an ABC transporter involved in iron uptake and virulence. *Molecular Microbiology, 40*, 572–585.

Brown, J. S., Gilliland, S. M., Spratt, B. G., & Holden, D. W. (2004). A locus contained within a variable region of pneumococcal pathogenicity island 1 contributes to virulence in mice. *Infection and Immunity, 72*, 1587–1593.

Brzozowska, I., Brzozowska, K., Zielenkiewicz, U. (2012). Functioning of the TA cassette of streptococcal plasmid pSM19035 in various Gram-positive bacteria. *Plasmid,68*, 51–60.

Camacho, A. G., Misselwitz, R., Behlke, J., Ayora, S., Welfle, K., Meinhart, A., et al. (2002). In vitro and in vivo stability of the $\varepsilon_2\zeta_2$ protein complex of the broad host-range *Streptococcus pyogenes* pSM19035 addiction system. *Biological Chemistry, 383*, 1701–1713.

Ceglowski, P., Boitsov, A., Chai, S., & Alonso, J. C. (1993a). Analysis of the stabilization system of pSM19035-derived plasmid pBT233 in *Bacillus subtilis*. *Gene, 136*, 1–12.

Ceglowski, P., Boitsov, A., Karamyan, N., Chai, S., & Alonso, J. C. (1993b). Characterization of the effectors required for stable inheritance of *Streptococcus pyogenes* pSM19035-derived plasmids in *Bacillus subtilis*. *Molecular Genetics and Genomics, 241*, 579–585.

Christensen, S. K., Pedersen, K., Hansen, F. G., & Gerdes, K. (2003). Toxin–antitoxin loci as stress-response-elements: ChpAK/MazF and ChpBK cleave translated RNAs and are counteracted by tmRNA. *Journal of Molecular Biology, 332*, 809–819.

Clewell, D. B. (1981). Plasmids, drug resistance, and gene transfer in the genus *Streptococcus*. *Microbiological Reviews, 45*, 409–436.

Cockeran, R., Theron, A. J., Steel, H. C., Matlola, N. M., Mitchell, T. J., Feldman, C., et al. (2001). Proinflammatory interactions of pneumolysin with human neutrophils. *Journal of Infectious Diseases, 183*, 604–611.

Croucher, N. J., Walker, D., Romero, P., Lennard, N., Paterson, G. K., Bason, N. C., et al. (2009). Role of conjugative elements in the evolution of the multidrug-resistant pandemic clone *Streptococcus pneumoniae*^{Spain23F} ST81. *Journal of Bacteriology, 191*, 1480–1489.

de la Hoz, A. B., Ayora, S., Sitkiewicz, I., Fernandez, S., Pankiewicz, R., Alonso, J. C., et al. (2000). Plasmid copy-number control and better-than-random segregation genes of pSM19035 share a common regulator. *Proceedings of the National Academy of Sciences of the United States of America, 97*, 728–733.

Dixon, J. M., & Lipinski, A. E. (1972). Resistance of group A beta-hemolytic streptococci to lincomycin and erythromycin. *Antimicrobial Agents and Chemotherapy, 1*, 333–339.

Du, W., Brown, J. R., Sylvester, D. R., Huang, J., Chalker, A. F., So, C. Y., et al. (2000). Two active forms of UDP-*N*-acetylglucosamine enolpyruvyl transferase in Gram-positive bacteria. *Journal of Bacteriology, 182*, 4146–4152.

Garcia-Pino, A., Balasubramanian, S., Wyns, L., Gazit, E., De Greve, H., Magnuson, R. D., et al. (2010). Allostery and intrinsic disorder mediate transcription regulation by conditional cooperativity. *Cell, 142*, 101–111.

Gilbert, R. J., Jimenez, J. L., Chen, S., Tickle, I. J., Rossjohn, J., Parker, M., et al. (1999). Two structural transitions in membrane pore formation by pneumolysin, the pore-forming toxin of *Streptococcus pneumoniae*. *Cell, 97*, 647–655.

Guiral, S., Mitchell, T. J., Martin, B., & Claverys, J. P. (2005). Competence-programmed predation of noncompetent cells in the human pathogen *Streptococcus pneumoniae*: Genetic requirements. *Proceedings of the National Academy of Sciences of the United States of America, 102*, 8710–8715.

Harvey, R. M., Stroeher, U. H., Ogunniyi, A. D., Smith-Vaughan, H. C., Leach, A. J., & Paton, J. C. (2011). A variable region within the genome of *Streptococcus pneumoniae* contributes to strain–strain variation in virulence. *PLoS ONE, 6*, e19650.

Hirst, R. A., Kadioglu, A., O'Callaghan, C., & Andrew, P. W. (2004). The role of pneumolysin in pneumococcal pneumonia and meningitis. *Clinical and Experimental Immunology, 138*, 195–201.

Holden, M. T., Hauser, H., Sanders, M., Ngo, T. H., Cherevach, I., Cronin, A., et al. (2009). Rapid evolution of virulence and drug resistance in the emerging zoonotic pathogen *Streptococcus suis*. *PLoS ONE, 4*, e6072.

Khoo, S. K., Loll, B., Chan, W. T., Shoeman, R. L., Ngoo, L., Yeo, C. C., et al. (2007). Molecular and structural characterization of the PezAT chromosomal toxin–antitoxin system of the human pathogen *Streptococcus pneumoniae*. *Journal of Biological Chemistry, 282*, 19606–19618.

Leipe, D. D., Koonin, E. V., & Aravind, L. (2003). Evolution and classification of P-loop kinases and related proteins. *Journal of Molecular Biology, 333*, 781–815.

Leplae, R., Geeraerts, D., Hallez, R., Guglielmini, J., Dreze, P., & Van Melderen, L. (2011). Diversity of bacterial type II toxin–antitoxin systems: A comprehensive search and functional analysis of novel families. *Nucleic Acids Research, 39*, 5513–5525.

Lioy, V. S., Martin, M. T., Camacho, A. G., Lurz, R., Antelmann, H., Hecker, M., et al. (2006). pSM19035-encoded zeta toxin induces stasis followed by death in a subpopulation of cells. *Microbiology, 152*, 2365–2379.

Lioy, V. S., Pratto, F., de la Hoz, A. B., Ayora, S., & Alonso, J. C. (2010). Plasmid pSM19035, a model to study stable maintenance in firmicutes. *Plasmid, 64*, 1–17.

Lioy, V. S., Machon, C., Tabone, M., Gonzalez-Pastor, J. E., Daugelavicius, R., Ayora, S., et al. (2012). The zeta toxin induces a set of protective responses and dormancy. *PLoS ONE, 7*, e30282.

Lock, R. A., Hansman, D., & Paton, J. C. (1992). Comparative efficacy of autolysin and pneumolysin as immunogens protecting mice against infection by *Streptococcus pneumoniae*. *Microbial Pathogenesis, 12*, 137–143.

Martner, A., Dahlgren, C., Paton, J. C., & Wold, A. E. (2008). Pneumolysin released during *Streptococcus pneumoniae* autolysis is a potent activator of intracellular oxygen radical production in neutrophils. *Infection and Immunity, 76*, 4079–4087.

Martner, A., Skovbjerg, S., Paton, J. C., & Wold, A. E. (2009). *Streptococcus pneumoniae* autolysis prevents phagocytosis and production of phagocyte-activating cytokines. *Infection and Immunity, 77*, 3826–3837.

Meinhart, A., Alings, C., Strater, N., Camacho, A. G., Alonso, J. C., & Saenger, W. (2001). Crystallization and preliminary X-ray diffraction studies of the $\varepsilon\zeta$ addiction system encoded by *Streptococcus pyogenes* plasmid pSM19035. *Acta Crystallographica. Section D, Biological Crystallography, 57*, 745–747.

Meinhart, A., Alonso, J. C., Strater, N., & Saenger, W. (2003). Crystal structure of the plasmid maintenance system ε/ζ: Functional mechanism of toxin ζ and inactivation by $\varepsilon_2\zeta_2$ complex formation. *Proceedings of the National Academy of Sciences of the United States of America, 100*, 1661–1666.

Mitchell, T. J., Andrew, P. W., Saunders, F. K., Smith, A. N., & Boulnois, G. J. (1991). Complement activation and antibody binding by pneumolysin via a region of the toxin homologous to a human acute-phase protein. *Molecular Microbiology, 5*, 1883–1888.

Murayama, K., Orth, P., de la Hoz, A. B., Alonso, J. C., & Saenger, W. (2001). Crystal structure of omega transcriptional repressor encoded by *Streptococcus pyogenes* plasmid pSM19035 at 1.5 Å resolution. *Journal of Molecular Biology, 314*, 789–796.

Mutschler, H., & Meinhart, A. (2011). Epsilon/zeta systems: Their role in resistance, virulence, and their potential for antibiotic development. *Journal of Molecular Medicine, 89*, 1183–1194.

Mutschler, H., Reinstein, J., & Meinhart, A. (2010). Assembly dynamics and stability of the pneumococcal epsilon zeta antitoxin toxin (PezAT) system from *Streptococcus pneumoniae*. *Journal of Biological Chemistry, 285*, 21797–21806.

Mutschler, H., Gebhardt, M., Shoeman, R. L., & Meinhart, A. (2011). A novel mechanism of programmed cell death in bacteria by toxin–antitoxin systems corrupts peptidoglycan synthesis. *PLoS Biology, 9*, e1001033.

Nau, R., & Eiffert, H. (2002). Modulation of release of proinflammatory bacterial compounds by antibacterials: Potential impact on course of inflammation and outcome in sepsis and meningitis. *Clinical Microbiology Reviews, 15*, 95–110.

Nowakowska, B., Kern-Zdanowicz, I., Zielenkiewicz, U., & Ceglowski, P. (2005). Characterization of *Bacillus subtilis* clones surviving overproduction of Zeta, a pSM19035 plasmid-encoded toxin. *Acta Biochimica Polonica, 52*, 99–107.

Overgaard, M., Borch, J., Jorgensen, M. G., & Gerdes, K. (2008). Messenger RNA interferase RelE controls relBE transcription by conditional cooperativity. *Molecular Microbiology, 69*, 841–857.

Pachulec, E., & van der Does, C. (2010). Conjugative plasmids of *Neisseria gonorrhoeae*. *PLoS ONE, 5*, e9962.

Pinas, G. E., Cortes, P. R., Orio, A. G., & Echenique, J. (2008). Acidic stress induces autolysis by a CSP-independent ComE pathway in *Streptococcus pneumoniae*. *Microbiology, 154*, 1300–1308.

Pratto, F., Cicek, A., Weihofen, W. A., Lurz, R., Saenger, W., & Alonso, J. C. (2008). *Streptococcus pyogenes* pSM19035 requires dynamic assembly of ATP-bound ParA and ParB on parS DNA during plasmid segregation. *Nucleic Acids Research, 36*, 3676–3689.

Ramage, H. R., Connolly, L. E., & Cox, J. S. (2009). Comprehensive functional analysis of *Mycobacterium tuberculosis* toxin–antitoxin systems: Implications for pathogenesis, stress responses, and evolution. *PLoS Genetics, 5*, e1000767.

Regev-Yochay, G., Trzcinski, K., Thompson, C. M., Lipsitch, M., & Malley, R. (2007). SpxB is a suicide gene of *Streptococcus pneumoniae* and confers a selective advantage in an in vivo competitive colonization model. *Journal of Bacteriology, 189*, 6532–6539.

Rosvoll, T. C., Pedersen, T., Sletvold, H., Johnsen, P. J., Sollid, J. E., Simonsen, G. S., et al. (2010). PCR-based plasmid typing in *Enterococcus faecium* strains reveals widely distributed pRE25-, pRUM-, pIP501- and pHTbeta-related replicons associated with glycopeptide resistance and stabilizing toxin–antitoxin systems. *FEMS Immunology and Medical Microbiology, 58*, 254–268.

Schlesinger, D. J., Shoemaker, N. B., & Salyers, A. A. (2007). Possible origins of CTnBST, a conjugative transposon found recently in a human colonic Bacteroides strain. *Applied and Environment Microbiology, 73*, 4226–4233.

Schwarz, F. V., Perreten, V., & Teuber, M. (2001). Sequence of the 50-kb conjugative multiresistance plasmid pRE25 from *Enterococcus faecalis* RE25. *Plasmid, 46*, 170–187.

Sletvold, H., Johnsen, P. J., Simonsen, G. S., Aasnaes, B., Sundsfjord, A., & Nielsen, K. M. (2007). Comparative DNA analysis of two vanA plasmids from *Enterococcus faecium* strains isolated from poultry and a poultry farmer in Norway. *Antimicrobial Agents and Chemotherapy, 51*, 736–739.

Sletvold, H., Johnsen, P. J., Hamre, I., Simonsen, G. S., Sundsfjord, A., & Nielsen, K. M. (2008). Complete sequence of *Enterococcus faecium* pVEF3 and the detection of an omega-epsilon-zeta toxin–antitoxin module and an ABC transporter. *Plasmid, 60*, 75–85.

Tettelin, H., Nelson, K. E., Paulsen, I. T., Eisen, J. A., Read, T. D., Peterson, S., et al. (2001). Complete genome sequence of a virulent isolate of *Streptococcus pneumoniae*. *Science, 293*, 498–506.

Van Melderen, L., & Saavedra De Bast, M. (2009). Bacterial toxin–antitoxin systems: More than selfish entities? *PLoS Genetics, 5*, e1000437.

Wang, X., Kim, Y., Hong, S. H., Ma, Q., Brown, B. L., Pu, M., et al. (2011). Antitoxin MqsA helps mediate the bacterial general stress response. *Nature Chemical Biology, 7*, 359–366.

Weihofen, W. A., Cicek, A., Pratto, F., Alonso, J. C., & Saenger, W. (2006). Structures of omega repressors bound to direct and inverted DNA repeats explain modulation of transcription. *Nucleic Acids Research, 34*, 1450–1458.

Winther, K. S., & Gerdes, K. (2012). Regulation of enteric vapBC transcription: Induction by VapC toxin dimer-breaking. *Nucleic Acids Research, 40*, 4347–4357.

Zielenkiewicz, U., & Ceglowski, P. (2001). Mechanisms of plasmid stable maintenance with special focus on plasmid addiction systems. *Acta Biochimica Polonica, 48*, 1003–1023.

Zielenkiewicz, U., & Ceglowski, P. (2005). The toxin–antitoxin system of the streptococcal plasmid pSM19035. *Journal of Bacteriology, 187*, 6094–6105.

Chapter 13
Archaeal Type II Toxin-Antitoxins

Shiraz A. Shah and Roger A. Garrett

Abstract A few of the bacterial type II TA systems, primarily those involved in translational inhibition, occur widely throughout the archaeal domain. Using a bioinformatic approach, the frequency and distribution of these diverse TA loci were examined within completed genomes of 124 archaea that are distributed fairly evenly throughout the major archaeal phyla. Results for the frequency and diversity of TA loci are summarised for archaea isolated from environmental niches generally characterised by extreme conditions including high temperature, high salt concentrations, high pressures, extremes of pH or strictly anaerobic conditions. No clear correlations were found between the number of TA loci present and either the genome size or particular environmental conditions. Multiple TA loci tend to be concentrated in variable genomic regions where the occurrence of intra- or inter-genomic gene transfer are most prevalent. For members of the Sulfolobales which are uniformly rich in TA loci, a case is made for some TA systems facilitating maintenance of important genomic regions.

13.1 Introduction

Until recently, type II TA systems have received relatively little attention in comparative genomic studies of archaea. This reflects a general uncertainty regarding their functions, the significance of their structural diversity and, to some

S. A. Shah · R. A. Garrett (✉)
Archaea Centre, Department of Biology, University of Copenhagen, Ole Maaløes Vej 5 DK-2200 Copenhagen N, Denmark
e-mail: garrett@bio.ku.dk

S. A. Shah
e-mail: sashah@bio.ku.dk

K. Gerdes (ed.), *Prokaryotic Toxin-Antitoxins*,
DOI: 10.1007/978-3-642-33253-1_13,

degree, their identities. Moreover, this uncertainty was compounded by the small gene sizes, especially for the antitoxins, which rendered their annotation difficult. This widespread deficiency was first highlighted by Gerdes' group who identified large numbers of non-annotated TA loci in archaeal and bacterial genomes and demonstrated the structural diversity of the protein components within different TA families (Pandey and Gerdes 2005; Gerdes et al. 2005; Jørgensen et al. 2009). This development, combined with contemporary insights gained into molecular mechanisms of toxin inhibitory activity (reviewed in Gerdes et al. 2005), served to focus attention on the profound importance of TA systems for cellular viability and survival.

Genome-based surveys of bacterial type II TA systems, carrying two genes, have identified eight major families denoted *vapBC, relBE, hicBA, mazEF, phd/doc, parDE, ccdAB* and *higBA* with an additional system in a *Streptococcus* plasmid carrying three genes (ω, ε and ζ, a repressor, antitoxin and toxin, respectively). VapC, RelE, MazE and HicA toxins have all been demonstrated experimentally to inhibit translation and Doc has also been implicated, at least indirectly, in affecting translation. In contrast, ParE and CcdB target the bacterial DNA gyrase thereby blocking DNA replication (reviewed in Gerdes et al. 2005).

Only three of these toxin families VapC, RelE and HicA, each targeting translation, occur commonly among archaea and this chapter is mainly focussed on these three TA systems. In the bacterium *Shigella flexneri* VapC toxins act by cleaving initiator tRNA within the anticodon loop thereby inhibiting translational initiation (Dienemann et al. 2011; Winther and Gerdes 2011), while RelE binds at the ribosomal A-site cutting the bound mRNA within the codon (Neubauer et al. 2009), and HicA is a translation-dependent mRNA transferase (Jørgensen et al. 2009; Makarova et al. 2009a). MazF and Doc have also been implicated in targeting translation, but their homologs are rarely found among archaea (Pandey and Gerdes 2005; Makarova et al. 2009b). Most archaea do not carry a homolog of the bacterial gyrase, the target of the ParE and CcdB toxins, employing instead the archaea-specific topoisomerase VI (Gadelle et al. 2003; Yamashiro and Yamagishi 2005).

An extensive genomic survey of bacterial and archaeal type II TA systems by Makarova et al. (2009b) that did not take into account the many non-annotated genes, reinforced the considerable structural diversity of the major TA families and identified additional subtypes, especially of the antitoxin components. This study also provided bioinformatical evidence for a possible additional TA locus encoding MNT (Minimal Nucleotidyl Transferase) and HEPN (Higher Eukaryote and Prokaryote Nucleotide binding). Although there is currently no experimental support for any toxin activity (Makarova et al. 2009b), we nevertheless included MNT/HEPN gene pairs in the present analysis because they occur commonly in archaea, especially among the hyperthermophiles and, moreover, their frequency of genome occurrence mirrors partially that of *vapBC* gene pairs.

A bioinformatical approach was employed to identify archaeal type II TA loci within 124 completed archaeal genomes. Exhaustive searches were made for the major families of TA gene loci, *vapBC, relBE* and *hicAB* and for the HEPN/MNT gene pairs and attempts were made to identify non-annotated antitoxin and toxin genes.

13.2 The Archaeal Perspective

Archaea differ from bacteria in their cellular biology in fundamental ways and they share many cellular processes exclusively with eukaryotes albeit generally in less complex forms. Although the evolutionary history of archaea and their relationship to early eukarya remains enigmatic (Gribaldi et al. 2010; Kurland et al. 2006), the maintenance of unique cellular properties among archaea is likely to be due to their successful adaptation to extreme environmental conditions. These include high temperature, extremes of pH, high salt, high pressures and strictly anaerobic conditions, and such environments that also tend to be low in energy sources (Kletzin 2007). It has been argued that some of the properties unique to archaea arose from adaptation to chronic energy stress through modifying catabolic pathways and by conserving energy via their low permeability ether-linked lipid membranes (Valentine 2007). Thus, stress in bacteria and archaea cannot simply be equated when considering the modes of action of toxins.

TA systems that are shared between bacteria and archaea appear primarily to inhibit translation, cleaving either mRNA bound in the ribosomal A-site (RelE), the anticodon of the initiator tRNA (VapC) or mRNA directly (HicA). The ribosomal tRNA binding sites, decoding site and peptidyl transferase centre constitute the most conserved regions of the translational apparatus, in both bacteria and archaea (and also in eukarya), as judged by their shared sensitivities to a wide range of antibiotics which specifically target these sites in both Domains (e.g. Rodriguez-Fonseca et al. 1995). Experimental studies indicate that bacterial TAs have alternative cellular targets, including the bacterial DNA gyrase, but it remains unknown whether there are unidentified archaeal toxins which bind to archaea-specific cellular sites.

13.3 A Bioinformatical Approach

All archaeal genomes publicly available at the beginning of 2012 were screened for the presence of type TA loci of the superfamilies *vapBC, relBE and hicAB*, as well as gene pairs of the predicted HEPN/MNT TA locus, by first constructing toxin-specific hidden Markov models (HMMs), using the jackhmmer program (Eddy 2011), against the genomes using known toxin genes as queries. Subsequently, all open reading frames (ORFs) between 50 and 250 aa that did not overlap previously annotated ORFs above 250 aa in length were extracted and screened using the constructed HMMs. Every upstream or downstream ORF, depending on the TA family type, located within a fixed distance of the matching toxin ORF, was extracted and clustered according to sequence similarity. Some of these clusters were judged to comprise antitoxin gene families based on manual inspection of their genomic contexts, and they were paired with the corresponding toxin genes to generate TA loci. Subsequently, we found that significant numbers

of TA loci were partially overlapping with larger annotated ORFs, particularly for members of the Thermococcales and, therefore, we extended the analyses to include these genes, which involved extensive manual inspection of the genomes.

13.4 Phylogenetic Distribution and Frequency of Archaeal TA Loci

A phylogenetic tree based on 16S rRNA sequences was generated for 124 archaea for which genome sequences were available. The genome size and natural habitat are given for each organism, and thermophiles are distinguished from mesophiles with a border for optimal growth of 50 °C (Table 13.1). More details of the natural environments and optimal growth conditions for many of the organisms are given by Kletzin (2007). Some orders, including the hyperthermophilic Sulfolobales, Thermoproteales and Thermococcales, are relatively overrepresented by closely related organisms including several *Sulfolobus islandicus, Pyrobaculum* and *Thermococcus* strains, while the less well-characterised Korarchaea (K) and Thaumarchaea (T) are underrepresented. This bias primarily reflects that the former group is relatively easy to isolate and culture and that some of them have been employed as model organisms for molecular, cellular and genetic studies. The total numbers of identified TA loci are given for *vapBC*, *relBE* and *hicAB* families and for the HEPN/MNT gene pairs in Table 13.1.

The results reveal a wide range of type II TA contents. Several organisms carry 30 or more TA loci but many have very few or no detectable loci. VapBC constitutes the dominant TA family and they are most prevalent among thermophiles, in particular in members of the thermoacidophilic Sulfolobales (Pandey and Gerdes 2005; Guo et al. 2011) and in some *Thermococcus* species. In contrast, *relBE* or *hicAB* gene pairs are quite rare especially among the 40 crenarchaeal genomes. For the euryarchaea *relBE* gene pairs were observed in about half of the genomes and several of these carried 1–9 copies. Similarly, *hicAB* pairs were identified in about half the euryarchaeal genomes with multiple copies occurring mainly among the Methanomicrobiales and Methanosarcinales. MNT/HEPN gene pairs occur much more frequently but are irregularly distributed. They are most common among crenarchaeal thermoacidophiles and thermoneutrophiles and the euryarchaeal hyperthermophiles (Table 13.1).

13.5 TA Loci Frequency and their Relationship to Genome Size and Environmental Factors

Generally, there is no simple correlation between genome size and TA locus frequency for the different archaeal phyla. For example, for most members of the Sulfolobales, the estimated number of TA loci varies from 17 to 49 but the

Table 13.1 Type II TA loci in archaeal genomes phylogenetic tree of archaea for which complete genome sequences are available together with the estimated number of TA loci of the *vapBC, relBE* and *hicAB* families, and the number of MNT/HEPN gene pairs

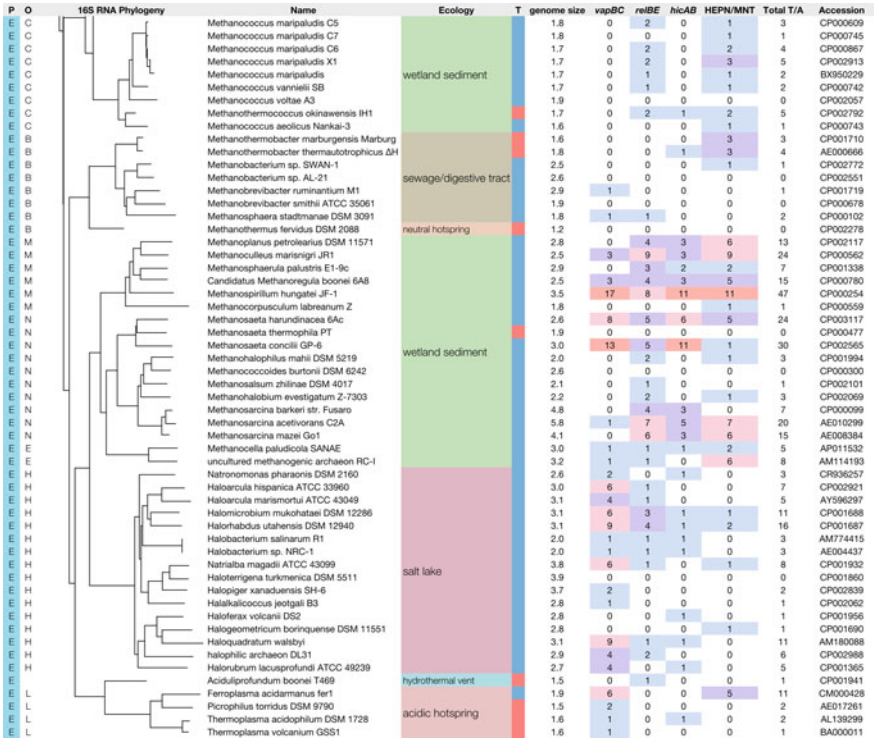
P	O	16S RNA Phylogeny	Name	Ecology	T	genome size	*vapBC*	*relBE*	*hicAB*	HEPN/MNT	Total T/A	Accession
C	S		Sulfolobus islandicus HVE10/4			2.7	19	3	0	5	27	CP002426
C	S		Sulfolobus islandicus REY15A			2.5	18	2	0	7	27	CP002425
C	S		Sulfolobus islandicus Y.G.57.14			2.7	21	1	0	6	28	CP001403
C	S		Sulfolobus islandicus M.16.27			2.7	19	1	0	7	27	CP001401
C	S		Sulfolobus solfataricus 98/2			2.7	20	2	0	8	30	CP001800
C	S		Sulfolobus solfataricus P2			3.0	22	1	0	10	33	AE006641
C	S		Sulfolobus islandicus M.14.25			2.6	19	1	0	6	26	CP001400
C	S		Sulfolobus islandicus L.D.8.5			2.7	24	5	0	7	36	CP001731
C	S		Sulfolobus islandicus L.S.2.15	acidic hotspring		2.7	24	2	0	6	32	CP001399
C	S		Sulfolobus islandicus M.16.4			2.6	22	0	0	7	29	CP001402
C	S		Sulfolobus islandicus Y.N.15.51			2.8	21	1	0	8	30	CP001404
C	S		Sulfolobus acidocaldarius DSM 639			2.2	16	0	0	1	17	CP000077
C	S		Sulfolobus tokodaii str. 7			2.7	25	3	0	21	49	BA000023
C	S		Acidianus hospitalis W1			2.1	24	2	0	12	38	CP002535
C	S		Acidianus brierleyi				15	4	0	8	27	
C	S		Metallosphaera cuprina Ar-4			1.8	2	0	0	0	2	CP002656
C	S		Metallosphaera sedula DSM 5348			2.2	12	0	0	5	17	CP000682
C	D		Pyrolobus fumarii 1A	hydrothermal vent		1.8	23	0	0	16	39	CP002838
C	D		Hyperthermus butylicus DSM 5456			1.7	17	0	0	1	18	CP000493
C	D		Desulfurococcus mucosus DSM 2162			1.3	0	0	0	0	0	CP002363
C	D		Desulfurococcus kamchatkensis 1221n	neutral hotspring		1.4	2	0	0	0	2	CP001140
C	D		Thermosphaera aggregans DSM 11486			1.3	1	0	0	0	1	CP001939
C	D		Staphylothermus hellenicus DSM 12710			1.6	5	0	0	7	12	CP002051
C	D		Staphylothermus marinus F1			1.6	4	0	0	7	11	CP000575
C	D		Ignicoccus hospitalis KIN4/I	hydrothermal vent		1.3	0	0	0	0	0	CP000816
C	D		Aeropyrum pernix K1			1.7	10	0	0	2	12	BA000002
C	D		Ignisphaera aggregans DSM 17230			1.9	7	0	0	5	12	CP002098
C	O		Acidilobus saccharovorans 345-15	acidic hotspring		1.5	0	0	0	0	0	CP001742
C	P		Pyrobaculum aerophilum str. IM2			2.2	8	0	0	17	25	AE009441
C	P		Pyrobaculum arsenaticum DSM 13514			2.1	9	0	0	21	30	CP000660
C	P		Pyrobaculum calidifontis JCM 11548	neutral hotspring		2.0	2	0	0	6	8	CP000561
C	P		Pyrobaculum islandicum DSM 4184			1.8	2	0	0	2	4	CP000504
C	P		Thermoproteus neutrophilus V24Sta			1.8	2	0	0	3	5	CP001014
C	P		Pyrobaculum sp. 1860			2.5	16	0	0	14	30	CP003098
C	P		Thermoproteus tenax Kra 1			1.8	1	0	0	4	5	FN869859
C	P		Thermoproteus uzoniensis 768-20			1.9	11	0	0	18	29	CP002590
C	P		Vulcanisaeta moutnovskia 768-28			2.3	7	0	0	7	14	CP002529
C	P		Vulcanisaeta distributa DSM 14429	acidic hotspring		2.4	6	0	0	16	22	CP002100
C	P		Caldivirga maquilingensis IC-167			2.1	0	0	0	6	6	CP000852
C	P		Thermofilum pendens Hrk 5			1.8	11	0	0	15	26	CP000505
A			Caldiarchaeum subterraneum			1.7	10	1	0	10	21	BA000048
K			Korarchaeum cryptofilum OPF8			1.6	6	0	0	3	9	CP000968
T			Cenarchaeum symbiosum A	marine		2.0	0	0	0	0	0	DP000238
T			Nitrosopumilus maritimus SCM1			1.6	0	0	0	0	0	CP000866
N			Nanoarchaeum equitans Kin4-M	hydrothermal vent		0.5	0	0	0	0	0	AE017199
E	Y		Methanopyrus kandleri AV19			1.7	0	0	0	0	0	AE009439
E	T		Pyrococcus yayanosii CH1			1.7	6	0	0	0	6	CP002779
E	T		Pyrococcus furiosus DSM 3638			1.9	7	0	1	1	9	AE009950
E	T		Pyrococcus abyssi GE5			1.8	17	0	0	5	22	AL096836
E	T		Pyrococcus horikoshii OT3			1.7	10	1	0	5	16	BA000001
E	T		Pyrococcus sp. NA2			1.9	7	0	0	3	10	CP002670
E	T		Thermococcus sp. AM4	deep sea vent		2.1	39	5	2	12	58	CP002952
E	T		Thermococcus gammatolerans EJ3			2.0	14	4	0	7	25	CP001398
E	T		Thermococcus kodakarensis KOD1			2.1	35	2	2	8	47	AP006878
E	T		Thermococcus onnurineus NA1			1.8	4	0	0	0	4	CP000855
E	T		Thermococcus sp. 4557			2.0	16	2	1	7	26	CP002920
E	T		Thermococcus barophilus MP			2.0	17	0	1	6	24	CP002372
E	T		Thermococcus sibiricus MM 739			1.8	5	0	0	5	10	CP001463
E	A		Archaeoglobus veneficus SNP6			1.9	13	0	2	6	21	CP002588
E	A		Archaeoglobus fulgidus DSM 4304	hydrothermal vent		2.2	24	2	2	8	36	AE000782
E	A		Ferroglobus placidus DSM 10642			2.2	21	3	1	17	42	CP001899
E	A		Archaeoglobus profundus DSM 5631			1.6	6	1	0	1	8	CP001857
E	C		Methanocaldococcus fervens AG86			1.5	0	0	0	0	0	CP001696
E	C		Methanocaldococcus jannaschii DSM 2661			1.7	2	4	0	7	13	L77117
E	C		Methanocaldococcus sp. FS406-22	deep sea vent		1.8	2	2	1	0	5	CP001901
E	C		Methanocaldococcus vulcanius M7			1.7	2	1	0	4	7	CP001787
E	C		Methanocaldococcus infernus ME			1.3	0	0	1	2	3	CP002009
E	C		Methanotorris igneus Kol 5			1.9	0	0	1	2	3	CP002737

(continued)

minimal genome of *Acidianus hospitalis* (2.2 Mb) carries 38 while the largest genome of *Sulfolobus solfataricus* P2 (3 Mb) contains 33. For other phyla, a clearer picture emerges when comparing strains within the same genus, e.g. the seven *Thermococcus* strains which have genome sizes ranging from 1.8 to 2.1 Mb. When ordered according to increasing approximate size (Table 13.1), these genomes carry 4, 10, 24, 25, 26, 47 and 58 TA loci, respectively, showing that the TA frequency increases disproportionately with genome size. A similar pattern is seen with *Pyrobaculum* strains. These results also underline the often large differences in the TA contents of pairs of closely related organisms.

There is little correlation between TA loci numbers and optimum growth temperatures. Although *Hyperthermus butylicus* which can grow up to 108 °C has

Table 13.1 (continued)



P	O	Name	Ecology	T	genome size	*vapBC*	*relBE*	*hicAB*	HEPN/MNT	Total T/A	Accession
E	C	Methanococcus maripaludis C5	wetland sediment		1.8	0	2	0	1	3	CP000609
E	C	Methanococcus maripaludis C7			1.8	0	0	0	1	1	CP000745
E	C	Methanococcus maripaludis C6			1.7	0	2	0	2	4	CP000867
E	C	Methanococcus maripaludis X1			1.7	0	2	0	3	5	CP002913
E	C	Methanococcus maripaludis			1.7	0	1	0	1	2	BX950229
E	C	Methanococcus vannielii SB			1.7	0	1	0	1	2	CP000742
E	C	Methanococcus voltae A3			1.9	0	0	0	0	0	CP002057
E	C	Methanothermococcus okinawensis IH1			1.7	0	2	1	2	5	CP002792
E	C	Methanococcus aeolicus Nankai-3			1.6	0	0	0	1	1	CP000743
E	B	Methanothermobacter marburgensis Marburg			1.6	0	0	0	3	3	CP001710
E	B	Methanothermobacter thermautotrophicus ΔH			1.8	0	0	1	3	4	AE000666
E	B	Methanobacterium sp. SWAN-1			2.5	0	0	0	1	1	CP002772
E	B	Methanobacterium sp. AL-21	sewage/digestive tract		2.6	0	0	0	0	0	CP002551
E	B	Methanobrevibacter ruminantium M1			2.9	1	0	0	0	1	CP001719
E	B	Methanobrevibacter smithii ATCC 35061			1.9	0	0	0	0	0	CP000678
E	B	Methanosphaera stadtmanae DSM 3091			1.8	1	1	0	0	2	CP000102
E	B	Methanothermus fervidus DSM 2088	neutral hotspring		1.2	0	0	0	0	0	CP002278
E	M	Methanoplanus petrolearius DSM 11571	wetland sediment		2.8	0	4	3	6	13	CP002117
E	M	Methanoculleus marisnigri JR1			2.5	3	9	3	9	24	CP000562
E	M	Methanosphaerula palustris E1-9c			2.9	0	3	2	2	7	CP001338
E	M	Candidatus Methanoregula boonei 6A8			2.5	3	4	3	5	15	CP000780
E	M	Methanospirillum hungatei JF-1			3.5	17	8	11	11	47	CP000254
E	M	Methanocorpusculum labreanum Z			1.8	0	0	0	1	1	CP000559
E	N	Methanosaeta harundinacea 6Ac			2.6	8	5	6	5	24	CP003117
E	N	Methanosaeta thermophila PT			1.9	0	0	0	0	0	CP000477
E	N	Methanosaeta concilii GP-6			3.0	13	5	11	1	30	CP002565
E	N	Methanohalophilus mahii DSM 5219			2.0	0	2	0	1	3	CP001994
E	N	Methanococcoides burtonii DSM 6242			2.6	0	0	0	0	0	CP000300
E	N	Methanosalsum zhilinae DSM 4017			2.1	0	1	0	0	1	CP002101
E	N	Methanohalobium evestigatum Z-7303			2.2	0	2	0	1	3	CP002069
E	N	Methanosarcina barkeri str. Fusaro			4.8	0	4	3	0	7	CP000099
E	N	Methanosarcina acetivorans C2A			5.8	1	7	5	7	20	AE010299
E	N	Methanosarcina mazei Go1			4.1	0	6	3	6	15	AE008384
E	E	Methanocella paludicola SANAE			3.0	1	1	1	2	5	AP011532
E	E	uncultured methanogenic archaeon RC-I			3.2	1	1	0	6	8	AM114193
E	H	Natronomonas pharaonis DSM 2160	salt lake		2.6	2	0	1	0	3	CR936257
E	H	Haloarcula hispanica ATCC 33960			3.0	6	1	0	0	7	CP002921
E	H	Haloarcula marismortui ATCC 43049			3.1	4	1	0	0	5	AY596297
E	H	Halomicrobium mukohataei DSM 12286			3.1	6	3	1	1	11	CP001688
E	H	Halorhabdus utahensis DSM 12940			3.1	9	4	1	2	16	CP001687
E	H	Halobacterium salinarum R1			2.0	1	1	1	0	3	AM774415
E	H	Halobacterium sp. NRC-1			2.0	1	1	1	0	3	AE004437
E	H	Natrialba magadii ATCC 43099			3.8	6	1	0	1	8	CP001932
E	H	Haloterrigena turkmenica DSM 5511			3.9	0	0	0	0	0	CP001860
E	H	Halopiger xanaduensis SH-6			3.7	2	0	0	0	2	CP002839
E	H	Halalkalicoccus jeotgali B3			2.8	1	0	0	0	1	CP002062
E	H	Haloferax volcanii DS2			2.8	0	0	1	0	1	CP001956
E	H	Halogeometricum borinquense DSM 11551			2.8	0	0	0	1	1	CP001690
E	H	Haloquadratum walsbyi			3.1	9	1	1	0	11	AM180088
E	H	halophilic archaeon DL31			2.9	4	2	0	0	6	CP002988
E	H	Halorubrum lacusprofundi ATCC 49239			2.7	4	0	1	0	5	CP001365
E		Aciduliprofundum boonei T469	hydrothermal vent		1.5	0	1	0	0	1	CP001941
E	L	Ferroplasma acidarmanus fer1	acidic hotspring		1.9	6	0	0	5	11	CM000428
E	L	Picrophilus torridus DSM 9790			1.5	2	0	0	0	2	AE017261
E	L	Thermoplasma acidophilum DSM 1728			1.6	1	0	1	0	2	AL139299
E	L	Thermoplasma volcanium GSS1			1.6	1	0	0	0	1	BA000011

In the kingdom/phyla column (P) *C* denotes Crenarchaeota, *E*—Euryarchaeota, *T*—Thaumarchaeota, *K*—Korarchaeota *A*—Aigarchaeota and *N*—Nanoarchaeota. In the Order column (O) *S* denotes Sulfolobales, *D*—Desulfurococcales, *O*—Acidolobales, *P*—Thermoproteales, *Y*—Methanopyrales, *T*—Thermococcales, *A*—Archaeoglobales, *C*—Methanococcales, *B*—Methanobacteriales, *M*—Methanomicrobiales, *N*—Methanosarcinales, *E*—Methanocellales, *H*—Halobacteriales and *L*—Thermoplasmatales. The ecological niches of the different organisms are indicated together with their degree of thermophilicity in column T, with a border of optimal growth of 50°C (blue below and red above). The numbers of the different TA loci and MNT/HEPN gene pairs are colour-shaded extending from *bright red* (>20), *light red* (20–11), *pink* (10–6), *violet* (5–3) and *light blue* (2–1). Approximate genome sizes and the Genbank/EMBL accession numbers are given for the genomes

a relatively high TA locus content of 18 (mainly *vapBC* loci) for a member of the Thermoproteales, *Methanopyrus kandleri* growing up to 110 °C has no detectable TA loci and some of the hyperthermophilic *Methanocaldococcus* strains also exhibit few TA loci.

More difficult to assess is the impact of the natural environments and the available nutrients, although in this respect the *S. islandicus* strains may be informative (Reno et al. 2009; Guo et al. 2011). They were all isolated from terrestial acidic hot springs with similar maximum growth temperatures and pH ranges but widely separated, and isolated, geographically; on Iceland, in

Kamchatka, Russia and in Yellowstone and Lassen National Parks, USA while the related *S. solfataricus* P2 strain derives from Naples, Italy. Each of the strains carries 26–36 TA loci which suggests that the nature of the environment is important. Moreover, active terrestial hot springs are likely to be particularly challenging for cells because temperatures can continuously change from maxima of around 80 °C to 0 °C if surrounded by ice, and pH values and nutrient availability can also change rapidly. A definitive answer to the effect of environmental factors on TA activity would require detailed and time-consuming experimental analyses of archaea cultivated under a wide range of environmental conditions.

13.6 Orphan Toxin and Antitoxin Genes

Many orphan toxins and some orphan antitoxin genes were detected in the genomes and the numbers tend to be proportional to the numbers of type II TA loci. For example, there are many orphan toxin genes among the Sulfolobales. Some of these may have been classed as orphans because the adjacent antitoxin protein gene was not identified (Pandey and Gerdes 2005) and others may be located adjacent to unidentified type III RNA antitoxin genes (see Chap. 15).

Presumably, over time, antitoxins or toxins may become associated with other cellular functions by selection. One such example could be provided by a single *vapC*-like gene (Ahos0712) of *A. hospitalis*. It lies in an operon with genes encoding proteins involved in transcription and initiator tRNA binding to the ribosome (You et al. 2011). This gene cassette is highly conserved in gene content, gene synteny and sequence in other *Sulfolobus* genomes (Guo et al. 2011). A possible explanation is that this orphan VapC-like protein acts as a VapC competitor and may regulate or inhibit initiator tRNA cleavage.

13.7 Locations within Genomes

Earlier comparative genomic analyses of closely related *Sulfolobus* species indicated that TA gene pairs tend to be concentrated in relatively large genomic regions (0.7–1 Mbp). These regions are the most variable in gene synteny and gene content (Guo et al. 2011) consistent with the extensive exchange of genes having occurred intra- and/or inter-genomically. This is illustrated for the genomes of *S. islandicus* REY15A and the related *S. solfataricus* P2 where a high level of gene synteny is maintained throughout about two-thirds of the genome while the remaining one-third is extensively shuffled (Fig. 13.1a). Most of the TA loci of both species fall within the variable region. Although few pairs of genome sequences from closely related archaeal species are available which show extensive gene synteny, a comparable analysis was possible for the genomes of two *Thermococcus* species. *T. kodakarensis* carrying many TA loci and *T. onnurineus*

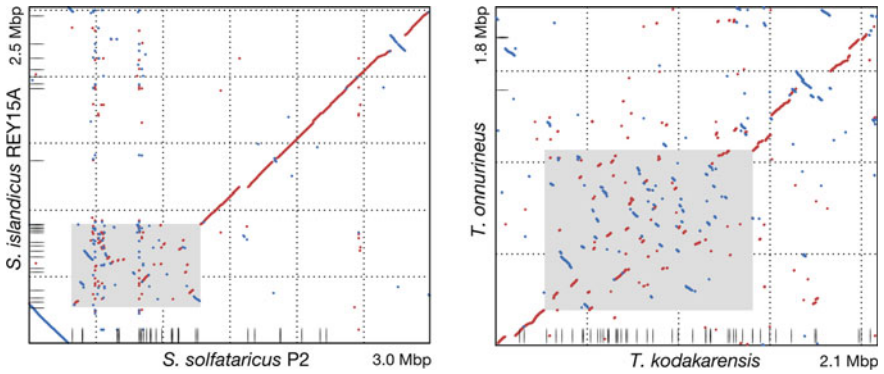


Fig. 13.1 Comparison of genomes from pairs of closely related archaea. Dot plots of **a** *Sulfolobus* species *S. islandicus* REY15A and *S. solfataricus* P2, and **b** *Thermococcus* species *T. onnurineus* and *T. kodakarensis* showing regions of gene synteny (*red*) and inverted synteny (*blue*). The total genome sizes are given and the large variable regions in each genome are shaded. TA loci are denoted by black lines along the corresponding genome axes

that exhibits very few TA loci (Fig. 13.1b). Here, the gene synteny is more limited and extends only over about one half of the genome but again the TA loci of *T. kodakarensis* are concentrated in the shuffled genome region. The latter example also illustrates the stark differences in the number of TA loci between some fairly closely related species.

Although several genomes, including some *Sulfolobus* species, contain many transposable elements and TA loci there is no general proportionality between the two. For example, both the above *Thermococcus* genomes carry few IS elements but one of the species, *T. kadakarensis,* exhibits several TA loci (Fig. 13.1b). Moreover, several other genomes carry many transposable elements but few TA loci (e.g. *Pyrococcus furiosus, Halobacterium* NRC1 and *Thermoplasma volcanium*) while some exhibit few transposable elements but contain many TA loci (e.g. *Sulfolobus acidocaldarius, H. butylicus* and *Thermococcus* sp. AM4) (Brügger et al. 2002; Filée et al. 2007).

13.8 TA Sequence Diversity Within Genomes

A. hospitalis genome carries 24 *vapBC* loci concentrated within the genomic regions 350–410 kb and 1,374–1,912 kb (You et al. 2011). While the VapC toxins are all PIN domain proteins (PilT N-terminal domain), the VapB antitoxins were classified into three families of transcriptional regulators, AbrB, CcdA/CopG and DUF217 (Fig. 13.2a, You et al. 2011). Tree building based on sequence alignments demonstrated that the sequences of these antitoxins and toxins are all highly diverse, with sequence identities between them rarely exceeding 30 %, as indicated by all the long tree branches for each protein (Fig. 13.2). A parallel tree-building study of the closely related *S. islandicus* strains REY15A and HVE10/4

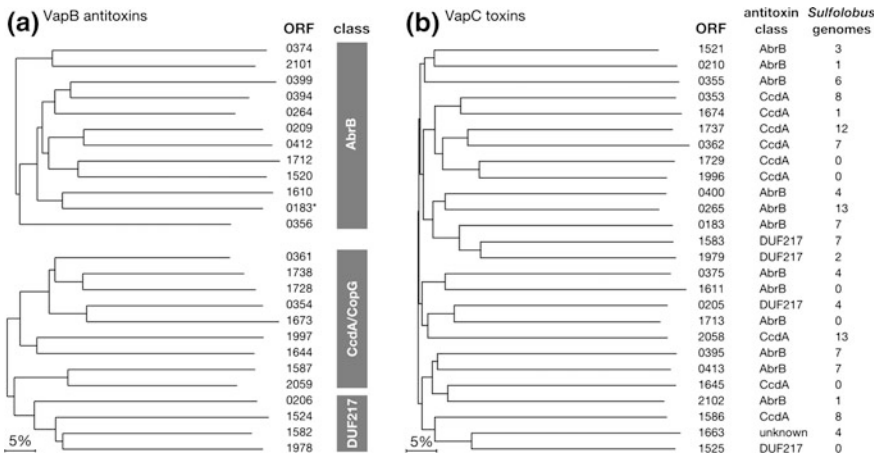


Fig. 13.2 Phylogenetic trees for VapB antitoxins and VapC toxins of the acidothermophile *A. hospitalis* W1. **a** The VapB tree demonstrates that the highly diverse antitoxins can be classified into three main subfamilies AbrB, CcdA/CopG and DUF217. In **b** the VapC tree shows the highly diverse toxin sequences falling into one major grouping. The VapB subfamily linked to each VapC is given. Moreover, the number of closely similar VapC proteins present in the available 13 Sulfolobales genomes (Table 13.1) is listed—*0* indicates that no VapC with a similar sequence is encoded in the genomes, while *13* indicates that a VapC with a closely similar sequence is encoded in each the genomes. Ahos Genbank numbers are given for each protein. Modified from You et al. (2011)

carrying 18 and 19 *vapBC* gene pairs, respectively, yielded a similar pattern of long branches for each VapB and VapC protein (Guo et al. 2011). Thus, all antitoxins and toxins within each archaeon are highly diverse in sequence.

In contrast, when intergenomic comparisons were made for other members of the Sulfolobales, isolated from both closely and distantly separated geographical terrestrial hot springs, several VapBC complexes showed high sequence similarity. For example, 11 of the 24 VapBC protein pairs identified in *A. hospitalis* (Fig. 13.2) exhibit closely similar sequences to homologs encoded in at least 7 of the 13 available *Sulfolobus* genomes (You et al. 2011). A further example is illustrated for the VapC toxins of *Pyrococcus* species (Fig. 13.3a) and for the predicted MNT toxin of the MNT/HEPN gene pairs for *Pyrobaculum* species (Fig. 13.3b). The result shows that the VapC and MNT sequences within each cluster of short branches derive from different species. Thus, there is apparently selection against the uptake of closely similar *vapBC* loci or MNT/HEPN gene pairs in a given genome, despite the abundance of many similar gene pairs in the environment.

The tree-building results of the analysis demonstrated further that for given gene pairs the subtypes of VapB and VapC do not always correspond implying that gene pairs exchange partners thereby potentially creating increased functional diversity of the TA systems (You et al. 2011), consistent with an earlier hypothesis (Gerdes et al. 2005).

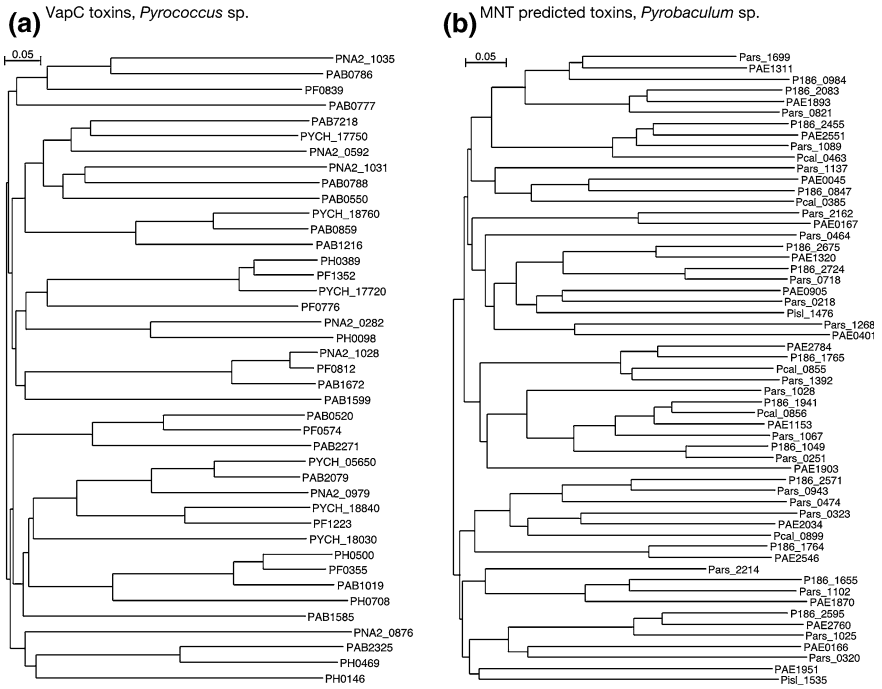


Fig. 13.3 Phylogenetic trees for the toxin VapC and the predicted toxin MNT. **a** VapC proteins encoded in different *Pyrococcus* species, and **b** MNT proteins encoded in diverse *Pyrobaculum* species. Proteins that fall within the small clusters of short branches all derive from different organisms. Trees generated for proteins deriving from one organism invariably yield long branches. Gene numbers are given for each of the genomes analysed (see Table 13.1)

13.9 Stress Response

Antitoxin-toxins were originally shown to enhance plasmid maintenance as a consequence of the growth of plasmid-free cells being preferentially inhibited, post segregation, by free toxins that are inherently more stable than antitoxins (Gerdes et al. 2005). To date, relatively few archaeal plasmids have been sequenced and there is no current evidence for type II TA loci occurring widely in plasmids. Nevertheless, the plasmid maintenance mechanism led to the hypothesis that the TA systems encoded widely in chromosomes facilitate retention of local DNA regions carrying important genes that might otherwise be prone to loss (Magnuson 2007; Van Melderen 2010).

This hypothesis receives support from the observation that *vapBC* loci and the HEPN/MNT gene pairs are concentrated within variable genomic regions of members of the Sulfolobales and Thermococcales, where intergenomic DNA exchange appears to be most active (Fig. 13.1). Furthermore, the hypothesis is reinforced by the high sequence diversity of each of the numerous VapC proteins

encoded within these genomes, exemplified for *A. hospitalis* (Fig. 13.2). For any pair of similar VapBC complexes, the loss of one would be compensated for by the presence of the other, thereby undermining any DNA maintenance capability.

For bacteria which grow slowly in nutrient-poor environments, multiple toxins are strongly implicated in responding to different types of nutrient deficiency and/or in enhancing quality control (Gerdes 2000; Pandey and Gerdes 2005). Involvement in stress response entails that the toxins inhibit growth, allowing the host to lie in a dormant state during the period of environmental stress (Pedersen et al. 2002; Gerdes et al. 2005). In this context, toxins have also been implicated in producing persistent cells which are able to remain dormant for longer periods and to withstand prolonged exposure to stress factors including antibiotics (Maisonneuve et al. 2011).

There may well be a negative effect on host growth as a consequence of carrying large numbers of TA loci (30–40 TA loci for some *Sulfolobus* species and a few other archaea (Table 13.1)) because of the likelihood of the continuous presence of low levels of free toxin (Wilbur et al. 2005). Although only highly diverse *vapBC* loci are present, presumably in order to avoid redundancy, the total number of TA loci present per genome may reflect a compromise between the ability to maintain important genes and to survive different environmental stresses while retaining an adequate growth rate under normal conditions.

In conclusion, there is a major deficit in experimental work on archaeal TA systems, especially with regard to stress responses. Almost all research to date has focussed on bacteria. Exceptions include the demonstration that the mode of action of a bacterial RelE toxin in *M. jannaschii* and bacteria were similar in vitro (Christensen and Gerdes 2003). Moreover, a recent study demonstrated that VapC proteins from *Pyrobaculum aerophilum* and *Mycobacterium* tuberculosis both cleave RNA at guanosine-rich motifs (McKenzie et al. 2012). Furthermore, heat shock of *S. solfataricus* (from 80 to 90 °C) was shown to induce expression of some TA loci while knockout of a single *vapBC* locus increased heat shock lability (Cooper et al. 2009; Maezato et al. 2011). Clearly, however, many challenging experiments remain to be performed in this rapidly developing field.

13.10 Type II TA Systems and Viral Defence

It has been proposed that bacterial TA systems could be involved in combating bacteriophage infection by, for example, blocking ribosomes and preventing the viruses from dominating the translational apparatus, prior to their propagating and lysing cells (see Chaps. 5 and 15). The inferred result would be that only the phage-infected cells would die. In principle, archaeal TA systems which primarily target translation could act similarly. However, most archaeal viruses, and especially those from extremely thermophilic and halophilic environments, show morphotypes and genomic properties distinct from bacterial and eukaryal viruses and they generally exist in stable relationships with their hosts at low copy numbers, infrequently, if ever, causing cell lysis (Prangishvili et al. 2006; Porter

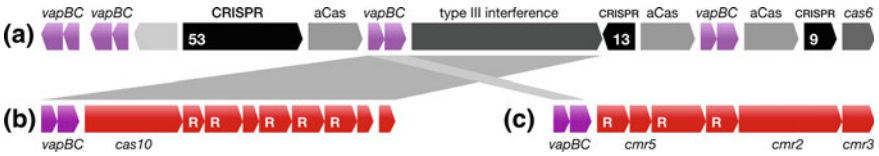


Fig. 13.4 Type III CRISPR systems linked to *vapBC* gene pairs. **a** CRISPR loci and genes of the acidothermophile *A. hospitalis* W1. CRISPR loci (*black*) show the numbers of repeats present. Genes encode proteins involved in uptake of new spacers (adaptation) labelled aCas, a gene encoding the RNA processing enzyme Cas6, and a gene cassette encoding type III interference proteins. Four *vapBC* gene pairs that are highly divergent in sequence are also present. **b** Expansion of the type III interference cassette of *A. hospitalis*, and **c** location of a highly similar *vapBC* gene pair located next to a different class of type III CRISPR interference cassette (denoted Cmr) in *S. islandicus* HVE10/4. Numbers of repeats are indicated for each CRISPR locus

et al. 2007). Consistent with these properties, circumstantial evidence suggests that the level of free viruses, at least in extreme thermoacidophilic environments, tend to be low relative to cellular levels suggesting that these viruses prefer to remain within cells under these challenging conditions (Snyder et al. 2010).

Another intriguing possibility arises from juxtapositioning of TA loci and CRISPR loci (Clustered Regularly Interspaced Short Palindromic Repeats) in some archaea. CRISPR-based adaptive immune systems target invading genetic elements, primarily viruses and conjugative plasmids, and they have been classified into three major types, of which only two (types I and III) occur in archaea, often with both major types present in the same archaeon (Garrett et al. 2011). The CRISPR arrays carry spacer regions taken up from invading genetic elements and their processed transcripts are able to facilitate targeting and cleavage of genetic elements with matching sequences. An example of a complex assembly of a type III CRISPR-based system, present in the *A. hospitalis* genome, is shown in Fig. 13.4. The CRISPR arrays and associated gene cassettes are interwoven with four *vapBC* loci for which all the antitoxins and toxins carry highly divergent sequences (Fig. 13.2). Thus, these CRISPR-associated TA systems could play a secondary role in combating invading genetic elements by helping to maintain the functional CRISPR immune systems, which also tend to be located within the variable chromosomal regions. Another interesting aspect of this system is that one *vapBC* locus associated with the type III interference system in *A. hospitalis* (Fig. 13.4b) shows a high level of sequence identity with *vapBC* loci specifically associated with a different subclass of type III interference systems found in the *S. islandicus* strains REY15A and HVE10/4 (Fig. 13.4c, Guo et al. 2011) suggesting that individual types of TA loci may coevolve with genes exhibiting specific functions.

13.11 Conclusions

Clearly, these are early days for studies of archaeal TA loci. Almost all of the experimental work to date has been performed on different bacterial TA systems some of which have no equivalent among archaea. Support is provided here for a

role in maintaining important regions of chromosomal DNA for those organisms, particularly members of the Sulfolobales and Thermococcales, which exhibit large variable genomic regions and often carry many TA loci. Involvement in response to nutrient deficiency and other stress factors are highly probable and these potential functional roles are not mutually exclusive. A rationale is provided for parts of the highly conserved translational apparatus being the primary target for some toxins that archaea share with bacteria. Finally, it remains to be seen whether there are undiscovered archaea-specific TA systems, or possibly hybrid systems with bacterial and archaeal antitoxin-toxin components, which exclusively target archaea-specific cellular components.

References

Brügger, K., Redder, P., She, Q., Confalonieri, F., Zivanovic, Y., & Garrett, R. A. (2002). Mobile elements in archaeal genomes. *FEMS Microbiology Letters, 206*, 131–141.

Christensen, S. K., & Gerdes, K. (2003). RelE toxins from bacteria and archaea cleave mRNAs on translating ribosomes which are rescued by tmRNA. *Molecular Microbiology, 48*, 1389–1400.

Cooper, C. R., Daugherty, A. J., Tachdjian, S., Blum, P. H., & Kelly, R. M. (2009). Role of *vapBC* toxin-antitoxin loci in the thermal stress response of *Sulfolobus solfataricus*. *Biochemical Society Transactions, 37*, 123–126.

Dienemann, C., Bøggild, A., Winther, K. S., Gerdes, K., & Brodersen, D. (2011). Crystal structure of VapBC toxin-antitoxin complex from *Shigella flexneri* reveals a hetero-octameric DNA-binding assembly. *Journal of Molecular Biology, 414*, 713–722.

Eddy, S. R. (2011). Accelerated profile HMM searches. *PLoS Computational Biology, 7*, 10.

Filée, J., Siguier, P., & Chandler, M. (2007). Insertion sequence diversity in archaea. *Microbiological and Molecular Biological Reviews, 71*, 121–157.

Gadelle, D., Filee, J., Buhler, C., & Forterre, P. (2003). Phylogenomics of type II DNA topoisomerases. *BioEssays, 25*, 232–242.

Garrett, R. A., Vestergaard, G., & Shah, S. A. (2011). Archaeal CRISPR-based immune systems: exchangeable functional modules. *Trends in Microbiology, 19*, 549–556.

Gerdes, K. (2000). Toxin-antitoxin modules may regulate synthsis of macromolecules during nutritional stress. *Journal of Bacteriology, 182*, 561–572.

Gerdes, K., Christensen, S. K., & Lobner-Olesen, A. (2005). Prokaryotic toxin-antitoxin stress response loci. *Nature Reviews Microbiology, 3*, 371–382.

Gribaldo, S., Poole, A. M., Daubin, V., Forterre, P., & Brochier-Armanet, C. (2010). The origin of eukaryotes and their relationship with the archaea: are we at a phylogenomic impasse? *Nature Reviews Microbiology, 8*, 743–752.

Guo, L., Brügger, K., Liu, C., Shah, S. A., Zheng, H., Zhu, Y., et al. (2011). Genome analyses of Icelandic strains of *Sulfolobus islandicus*: Model organisms for genetic and virus-host interaction studies. *Journal of Bacteriology, 193*, 1672–1680.

Jørgensen, M. G., Pandey, D. P., Jaskolska, M., & Gerdes, K. (2009). HicA of *Escherichia coli* defines a novel family of translation-independent mRNA transferases in bacteria and archaea. *Journal of Bacteriology, 191*, 1191–1199.

Kletzin, A. (2007). General characteristics and important model organisms. In R. Cavicchioli (Ed.), *Molecular and Cellular Biology* (pp. 14–92). USA: ASM press.

Kurland, C. G., Collins, L. J., & Penny, D. (2006). Genomics and the irreducible nature of eukaryotic cells. *Science, 312*, 1011–1014.

Maezato, Y., Daugherty, A., Dana, K., Soo, E., Cooper, C., Tachdjian S., et al. (2011) VapC6, a ribonucleolytic toxin regulates thermophilicity in the crenarchaeote *Sulfolobus solfataricus. RNA, 17*, 1381–1392.

Magnuson, R. D. (2007). Hypothetical functions of toxin-antitoxin systems. *Journal of Bacteriology, 189*, 6089–6092.

Maisonneuve, E., Shakespeare, L. J., Jørgensen, M. G., & Gerdes, K. (2011). Bacterial persistence by RNA endonucleases. *Proceedings of National Academic of Sciences USA, 108*, 13206–13211.

Makarova, K. S., Grishin, N. V., & Koonin, E. V. (2009a). The HicAB cassette, a putative novel, RNA targeting toxin-antitoxin system in archaea and bacteria. *Bioinformatics, 22*, 2581–2584.

Makarova, K. S., Wolf, Y. I., & Koonin, E. V. (2009b). Comprehensive comparative-genomic analysis of type 2 toxin-antitoxin systems and related mobile stress response systems in prokaryotes. *Biology Direct, 4*, 19.

McKenzie, J. L., Duyvestyn, J. M., Smith, T., Bendak, K., Mackay, J., Cursons, R. et al. (2012). Determination of ribonuclease sequence-specificity using Pentaprobes and mass spectrometry. *RNA, 18*, 1267–1278

Melderen, L. V. (2010). Toxin-antitoxin systems: why so many, what for? *Current Opinion in Microbiology, 13*, 781–785.

Neubauer, C., Gao, Y. G., Andersen, K. R., Dunham, C. M., Kelley, A. C., Hentschel, J., et al. (2009). The structural basis for mRNA recognition and cleavage by the ribosome-dependent endonuclease RelE. *Cell, 139*, 1084–1095.

Pandey, D. P., & Gerdes, K. (2005). Toxin-antitoxin loci are highly abundant in free-living but lost from host-associated prokaryotes. *Nucleic Acids Research, 33*, 966–976.

Pedersen, K., Christensen, S. K., & Gerdes, K. (2002). Rapid induction and reversal of a bacteriostatic condition by controlled expression of toxins and antitoxins. *Molecular Microbiology, 45*, 501–510.

Porter, K., Russ, B. E., & Dyall-Smith, M. L. (2007). Virus-host interactions in salt lakes. *Current Opinion in Microbiology, 10*, 418–424.

Prangishvili, D., Forterre, P., & Garrett, R. A. (2006). Viruses of the archaea: A unifying view. *Nature Reviews Microbiology, 11*, 837–848.

Reno, M. L., Held, N. L., Fields, C. J., Burke, P. V., & Whitaker, R. J. (2009). *Sulfolobus islandicus* pan-genome. *Proceedings of National Academic of Sciences, USA, 106*, 8605–8610.

Rodriguez-Fonseca, C., Amils, R., & Garrett, R. A. (1995). Fine structure of the peptidyl transferase centre on 23 S-like rRNAs deduced from chemical probing of antibiotic-ribosome complexes. *Journal of Molecular Biology, 247*, 224–235.

Snyder, J. C., Bateson, M. M., Lavin, M., & Young, M. J. (2010). Use of cellular CRISPR (clusters of regularly interspaced short palindromic repeats) spacer-based microarrays for detection of viruses in environmental samples. *Applied and Environment Microbiology, 76*, 7251–7258.

Valentine, D. L. (2007). Adaptations to energy stress dictate the ecology and evolution of archaea. *Nature Reviews Microbiology, 5*, 316–323.

Wilbur, J. S., Chivers, P. T., Mattison, K., Potter, L., Brennan, R. G., & So, M. (2005). *Neisseria gonorrheae* FitA interacts with FitB to bind DNA through its ribbon-helix-helix motif. *Biochemistry, 44*, 12515–12524.

Winther, K. S., & Gerdes, K. (2011). Enteric virulence associated protein VapC inhibits translation by cleavage of initiator tRNA. *Proceedings of National Academic of Sciences, USA, 108*, 7403–7407.

Yamashiro, K., & Yamagishi, A. (2005). Characterization of the DNA gyrase from the thermoacidophilic archaeon *Thermoplasma acidophilum*. *Journal of Bacteriology, 187*, 8531–8536.

You, X.-Y., Liu, C., Wang, S.-Y., Jiang, C.-Y., Shah, S. A., Prangishvili, D., et al. (2011). Genomic studies of *Acidianus hospitalis* W1 a host for studying crenarchaeal virus and plasmid life cycles. *Extremophiles, 15*, 487–497.

Chapter 14
Type II Toxin-Antitoxin Loci: Phylogeny

Hong-Yu Ou, Yiqing Wei and Dexi Bi

Abstract The web-based database TADB offers a comprehensive compilation of both predicted and experimentally supported type II toxin-antitoxin (TA) locus data and genetic features. Here, we examine the distribution and diversity of TADB-recorded TA loci throughout bacterial and archaeal genomes; in particular, those carried by the chromosomal mobile elements, such as prophages and integrative and conjugative elements (ICEs). Then the discussion covers the toxin protein-based TA family classification and the conserved domain pairs found in TA proteins.

14.1 Introduction

Bacterial and archaeal type II toxin-antitoxin (TA) loci have been hypothesized or demonstrated to play key roles in stress response, bacterial physiology, and stabilization of horizontally acquired genetic elements (Gerdes et al. 2005). To date, numerous phylogenetically and functionally distinct type II TA systems have been identified through experimental and bioinformatics approaches (Pandey and Gerdes 2005; Sevin and Barloy-Hubler 2007; Shao et al. 2011; Leplae et al. 2011; Ramage et al. 2009). In particular, the advent of the genomics era witnesses the widespread abundance of type II TA loci throughout the genomes of almost all free-living bacteria and archaea. When present on plasmids or other horizontally acquired elements on the chromosomes, TA systems are frequently viewed as 'selfish DNA'

H.-Y. Ou (✉) · Y. Wei · D. Bi
State Key Laboratory of Microbial Metabolism and School of Life Sciences & Biotechnology, Shanghai Jiaotong University, 200030, Shanghai, China
e-mail: hyou@sjtu.edu.cn

K. Gerdes (ed.), *Prokaryotic Toxin-Antitoxins*,
DOI: 10.1007/978-3-642-33253-1_14,

as many of these addiction systems will elicit post-segregational killing in the element-deprived offspring cells (Van Melderen and Saavedra De Bast 2009). The various type II TA systems may well contribute towards stochastic variation within the population, gearing up individual subsets of organisms distinctly to favor survival of the population itself in the face of a likely wide diversity of unpredictable environmental, competitive, and hostile challenges ahead.

14.2 Web-Based Database for Type II TA Loci

The TA loci data derived from computationally predicted datasets and/or reports of experimentally validated TA genes have been recently organized as a web-based database called TADB (Shao et al. 2011). The BLASTP-identified 921 TA loci present in 126 sequenced genomes reported by Gerdes and co-workers (Pandey and Gerdes 2005) were first uploaded. Next, the 5,806 TA loci found in 604 genomes that had been assigned to 44 conserved TA domain pairs by Makarova et al. (2009) were archived in TADB. The database was then further complemented with the predicted data in 883 annotated genomes by the RPSBLAST-based tool RASTA-Bacteria (Sevin and Barloy-Hubler 2007). Subsequently, the three datasets were compared to ensure that only de-duplicated entries were included in TADB. In addition, supplementary TA loci had been checked by the literature search. Currently, TADB version 1.1 contains 10,753 type II TA gene pairs distributed within 1,240 genome sequences representative of 962 strains of phylogenetically diverse bacteria and archaea, including details of 106 experimentally validated TA loci. A broad range of similarity search, sequence alignment, genome context browser, and phylogenetic tools are readily accessible via TADB.

14.3 Occurrence of Type II TA Loci in Bacteria and Archaea

Conserved genetic organization of known type II TA systems typically contains two but occasionally three tandem genes coding for cognate partners. Comprehensive genome analyses have revealed that the type II TA systems are diverse and widespread in the prokaryotic organisms with at least one TA locus (Pandey and Gerdes 2005; Sevin and Barloy-Hubler 2007; Makarova et al. 2009; Leplae et al. 2011). An individual bacterial or archael genome may generally harbor $0 \sim > 50$ TA loci. However, 23 strains possess more than 50 putative TA loci, such as the pathogenic *Mycobacterium tuberculosis* H37Rv (66 TA loci), the symbiotic *Photorhabdus luminescens* TTO155 (79 TA loci), and the hyperthermophilic archaeon *Thermococcus kodakarensis* KOD1 (55 TA loci). No simple linear correlation between the number of TA loci and the genome size was observed (Leplae et al. 2011); for example, there are 20 TADB-recorded TA loci in the *Thermus thermophilus* HB8 chromosome

(1.8 Mb in size), 15 in the *Escherichia coli* K-12 MG1655 chromosome (4.6 Mb), 167 in the *Microcystis aeruginosa* NIES-843 chromosome (5.8 Mb), and 22 in the *Streptomyces coelicolor* A3(2) chromosome (8.6 Mb). Similarly, no simple correlation was deduced between the number of TA loci and the membership of a phylum or the growth rate (Van Melderen and Saavedra De Bast 2009). A slight bias of TA loci occurrence may be present on the bacterial life cycle, *i.e.*, absent or very few in small symbionts versus free-living prokaryotes. BLASTP against the TADB-recorded toxin and antitoxin proteins failed to detect any TA pairs in 45 % of the 122 completely sequenced genomes of symbiotic bacteria taken from the GOLD database, like *Buchnera aphidicola* and *Mycoplasma hyopneumoniae*; whereas, only 12 % of the 1,097 sequenced free-living prokaryotes lack TA loci. Whether the TA systems promote the bacterial fitness and facilitate the evolution of free-living organisms awaits investigation within the population.

14.4 Chromosomal Mobile Element-Encoding TA System

TA loci distribution varies greatly in different isolates of one bacterial species (see an example in Fig. 14.1), suggesting that horizontal transfer of TA loci may be frequent. A few plasmid-encoded TA systems have been well documented, and shown to give a plasmid-carrying competitive advantage. In addition, chromosomal TA pairs have been shown or postulated to play multiple other functions. These systems can act to stabilize large horizontally acquired chromosomal regions by selecting against their loss (Van Melderen and Saavedra De Bast 2009). TA pairs also provide the host cell immunity to the same or related toxin encoded by other co-resident mobile genetic elements, thus acting as anti-addiction modules by allowing for the loss of one or more 'redundant' anti-toxin gene and its carrier mobile genetic element (Van Melderen and Saavedra De Bast 2009).

TA loci have increasingly being discovered within or close to the mobile elements on bacterial chromosomes, including super-integrons, prophages, integrative and conjugative elements (ICEs), and a growing list of genomic islands (Pandey and Gerdes 2005; Van Melderen and Saavedra De Bast 2009). The super-integron on the chromosome II of *Vibrio cholerae* O1 El Tor strain N16961 codes for 16 putative TA systems, of which three copies of *parDE* contribute to the maintenance of the integrity of this 130-Kb mobile element (Yuan et al. 2011). SXT is a 100-Kb ICE that confers resistance to multiple antibiotics upon many *Vibrio cholerae* clinical isolates and the embedded *mosAT* TA locus stabilizes this element (Wozniak and Waldor 2009).

Remarkably, out of the 272 ICEs with the whole sequences described in the ICEberg database (Bi et al. 2012), 81 (81/272 = 30 %) ICEs have one or two type II TA loci. As an example, the putative TA pairs with conserved TA domains identified in four completely sequenced *Pseudomonas aeruginosa* genomes (strain PAO1, UCBPP-PA14, LESB58, and PA7) are shown in Fig. 14.1. As for the UCBPP-PA14 strain, a TA locus (coding for a RelE-RHH domain pair) was found

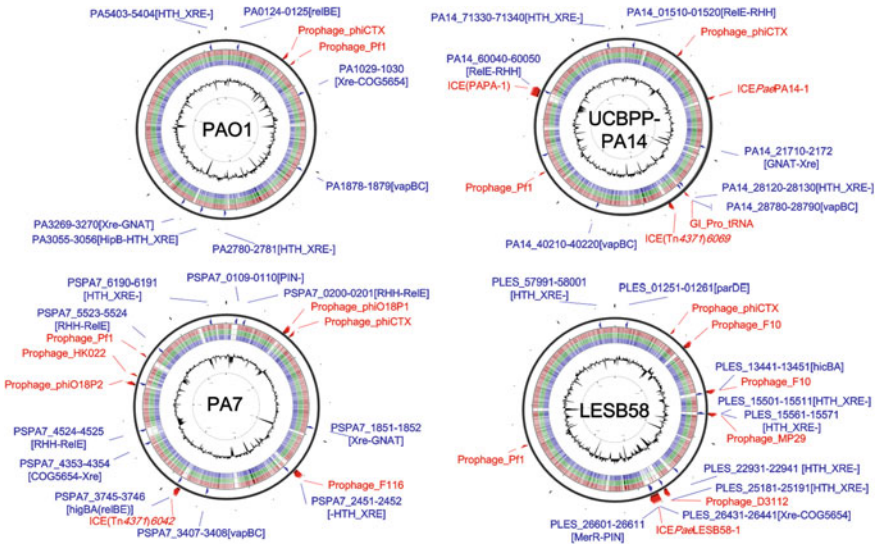


Fig. 14.1 Chromosomal maps of type II TA loci in the complete genomes of four *Pseudomonas aeruginosa* strains (PAO1, UCBPP-PA14, LESB58, and PA7). *Blue labels* represent TADB-recorded putative TA loci, and the corresponding TA family or TA domain pair is shown in *square brackets*. The *red* regions on the *outer ring* correspond to prophages predicted by PHAST or the integrative and conjugative elements (ICEs) as described in ICEberg. The BLASTN-derived comparison maps of one genome with the three others under study were created using the CGview server. The *inner rings* represent the GC skews

in the 108-Kb PAPI-1 ICE that conducts acute pneumonia and bacteremia in murine models (Harrison et al. 2010), and two TA loci within the *pro* tRNA gene-associated island. Of the LESB58 strain, the 111-Kb ICE*clc*-like element ICE*Pae*LESB58-1 carries two TA loci (with the MerR-PIN and Xre-COG5654 TA domain pairs) and mercuric resistance genes, while each of the three prophages (F10, MP29, and D3112) encodes one TA system. The 52-Kb Tn*4371*-like element ICE (Tn4371)*6042* of the PA7 strain codes for a TA system belonging to the *higBA* (*relBE*) family, and the streptomycin, bleomycin, and aminoglycoside resistances. The prophage F116 of the PA7 strain owns a putative TA locus. These TA systems may play key roles in stabilization of the self-transmissible integrative elements, which thus contribute to horizontal transfer of virulence determinants, antibiotic-resistant genes, and many other bacterial traits.

14.5 Type II TA Family Classification Based on Toxin Protein

In general, classification of the known type II TA loci is based on sequence similarity of toxin protein, associated with a specific cognate antitoxin protein (Gerdes et al. 2005; Van Melderen and Saavedra De Bast 2009). Recently,

phylogenetically related toxins sharing three-dimensional structural similarity were proposed to be grouped together in single families, as exemplified by the example of the RelE family (Fig. 14.2). As of now, 949 TADB-archived TA loci in 159 genomes have been such assigned into 11 two-component and 3 three-component TA families on the basis of the toxin protein sequence similarity and tertiary structure. The 11 two-component TA families are listed as follows: *ccd* (containing 9 TA loci archived by TADB), *hicBA* (60 loci), *hipBA* (1 locus), *mazEF/chpA* (89 loci), *mosAT* (1 locus), *parD/PemKI* (2 loci), *parDE* (62 loci), *phd-doc* (27 loci), *relBE* (322 loci), *vapBC* (383 loci), and *yeeUV* (2 loci). The *relBE* family contains five subfamilies based on the tertiary structure comparisons of the toxin proteins as described below: *relBE* (204 loci), *higBA* (111 loci), *yfeM-yoeB* (3 loci), *ygiTU/mgsAR* (1 locus), and *prlF-yhaV* (3 loci). In addition, the *bat-bto* locus has now been grouped into the *vapBC* family. The Bto toxin encoded by *bll2434* in *Bradyrhizobium japonicum* USDA 110 contains the C-terminal PIN domain commonly present in VapC-like endoribonucleases (Miclea et al. 2010). The 4 three-component families are *omega-epsilon-zeta* (1 locus), *pasABC* (1 locus), *paaR-paaA-parE* (2 loci) and *higBA secB* (see Chap. 17). Additional details about the TA families are available through the 'Browse by TA family' webpage of the TADB database.

We have performed pair-wise structure alignments between the ten toxin proteins with determined crystal structures using jFATCAT due to no significant sequence similarity found by BLASTP. The jFATCAT-defined P value provides a measurement of the chance of two random structures being identified as similar. The smaller the P value, the more statistically significant similarity between corresponding structures; a pair of proteins with P value $< 5E-02$ are considered to be significantly similar. By this criterion, YoeB (P value $= 6.92E-08$) and MqsR(YgiU) (P value $= 3.33-E03$) exhibit significant tertiary folds based on the similarity to the reference RelE in the RelBE family. As such the YoeB and MqsR (YgiT) were assigned to the RelE family. Additionally, as HigB and YhaV have been reported as RelE-homologues (Schmidt et al. 2007; Christensen-Dalsgaard et al. 2010), the *higBA and prlF-yhaV* locus have been grouped into the *relBE* family but as individual subfamilies, given the unique gene order and/or cellular activity associated with each of these *relBE* subfamilies. Therefore, the *relBE* family in TADB now includes five subfamilies: *relBE*, *higBA*, *yoeB-yefM*, *ygiTU*, *and prlF-yhaV*. However, although ParE also shows significant tertiary fold-based similarity to RelE (P value $= 1.29E-06$), ParE possesses a unique cellular target and activity (Jiang et al. 2002), hence *parDE* loci has been retained as a separate TA family.

14.6 TA Domain Pair-Based Classification

We have also grouped TADB entries by a second TA domain pair-based classification system. The structurally and functionally distinct TA domain pairs had been characterized by bioinformatics analysis (Makarova et al. 2009). Eighteen

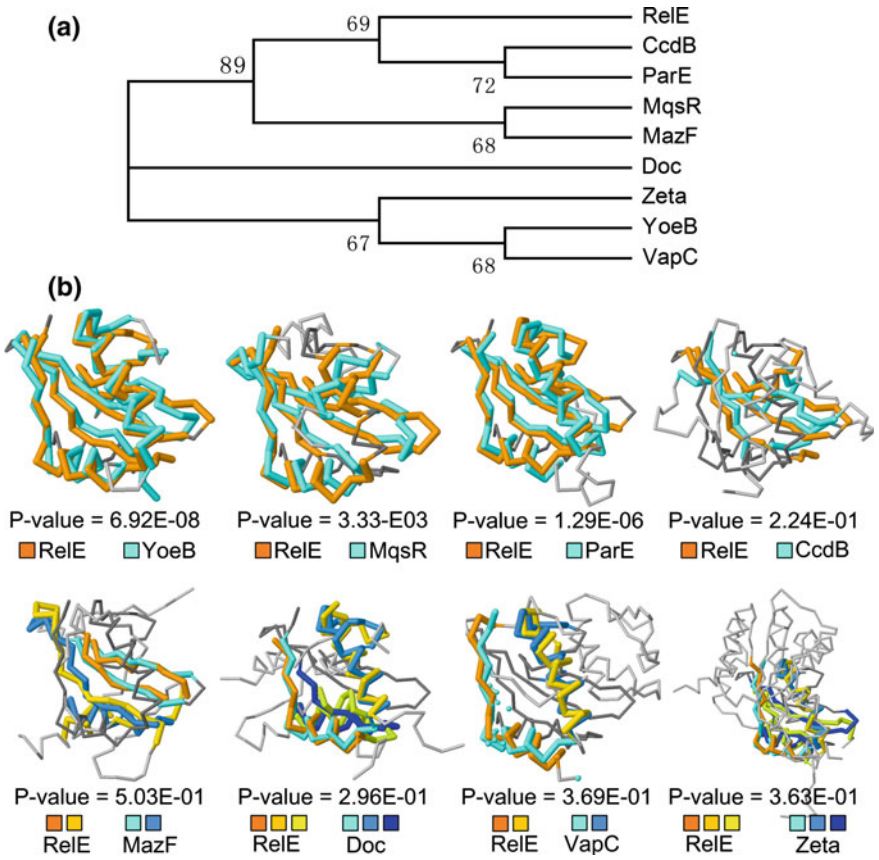


Fig. 14.2 Phylogeny and structure alignment of the ten toxin proteins with determined 3-D structures. The RelE toxin (PDB ID: 3KHA.A) is severed as reference of pair-wise structure alignment while the other toxin protein is YoeB (2A6Q.A), MqsR (3HI2.A), ParE (3KXE.A), CcdB (3VUB.A), MazF (1UB4.A), Doc (3K33.A), VapC (3DBO.A), Zeta (1GVN.A), respectively. **a** Phylogenetic relationships of the ten toxin proteins under study. The phylogenetic tree was generated in MEGA5.1 using the maximum-likelihood method with default settings. Bootstrap percentage values (1,000 replicates) are shown at the nodes. **b** Tertiary fold-based pair-wise structure alignments using the jFATCAT-flexible tool. P values listed denote the degree of significance of the match; a pair of proteins with P-value $< 5E - 02$ are considered to be significantly similar

conserved domains were found among the toxin proteins, including RelE, YgiU, SMa0917, PIN, MNT, MazFn, MazF, HipA, HicA, GNAT, Fic, DUF397, COG5654, COG4679, COG3832, COG2929, COG2856, and Bro. Similarly, eighteen domains of antitoxin proteins: Xre, AbrB, ArsR, COG1753, COG2442, COG2880, COG2886, COG5304, COG5606, COG5642, HEPN, HicB, MerR, PHD, RHH, XF1863, and YhfG. The detailed information is available at the CDD database using the domain name as the search terms. (Table 14.1)

Table 14.1 The relationship of type II TA family[a] and TA domain pair[b] classification systems of identified and/or predicted TA loci

#	TA domain pairname	TA family name	Number of TA loci previously assigned to TA family AND TA domain pair sets[c]	Total number of TA loci in TADB mapping to each classification category[d]
1	Xre-HipA	*hipBA*	1	258
2	Xre-YgiU	*ygiT-mqsR(relBE)*	1	45
3	Xre-PIN	*vapBC*	11	46
4	Xre-MazF	*mazEF*	2	18
5	Xre-RelE	*relBE*	100	645
6	COG5606-RelE	*higBA(relBE)*	14	53
7	RHH-RelE	*relBE*	83	
8	RHH-RelE	*parDE*	36	510
9	RHH-RelE	*higBA(relBE)*	1	
10	RHH-RelE	*paaR-paaA-parE*	2	
11	AbrB-RelE	*relBE*	13	48
12	PHD-RelE	*relBE*	65	
13	PHD-RelE	*parDE*	12	306
14	PHD-RelE	*yefM-yoeB(relBE)*	2	
15	RHH-Fic	*phd-doc*	2	20
16	RHH-PIN	*vapBC*	163	535
17	RHH-MazF	*relBE*	1	
18	RHH-MazF	*mazEF*	35	161
19	RHH-MazF	*ccd*	8	
20	XF1863-MazF	*mazEF*	2	27
21	AbrB-MazF	*mazEF*	26	99
22	AbrB-Fic	*phd-doc*	12	58
23	AbrB-PIN	*vapBC*	94	379
24	COG2886-PIN	*vapBC*	4	58
25	PHD-PIN	*vapBC*	55	298
26	PHD-Fic	*phd-doc*	2	14
27	HicB-HicA	*hicBA*	46	247
		Total number of classified TA loci	793	3825

[a] The type II TA family classification system is based on toxin protein sequence similarity as described in the reviews by Gerdes et al. (2005) and Van Melderen et al. (2009)

[b] The type II TA domain pair classification system as suggested recently by Makarova et al. (2009) is based on identification of TA pairs sharing cognate toxin and anti-toxin domains and is independent of wider protein-level similarity .

[c] The relationships noted in this table were defined by analysis of 793 TA loci that had previously been classified by both systems into individual TA families and TA domain pairs.

[d] The remaining 3,032 TA loci presently in TADB that had been classified by the TA domain pair system alone were then classified into corresponding TA families using this same mapping approach. Numbers indicated by curly brackets represent TA loci of known TA domain pair assignment that could not be matched reliably to a unique TA family.

When 5,877 TA loci described by TADB were scanned by specific toxin and antitoxin domains, 44 different TA domain pairs were obtained (see the 'Browse by toxin/antitoxin-related domain' webpage of the TADB database), such as 647 TA loci with Xre-RelE domains (11 %), 563 with HEPN-MNT (9 %) and 541 with RHH-PIN (9 %). The relationship of TA family with TA domain pairs, where recognizable, is shown in Table 14.1. Interestingly, there are 'hybrid' associations of a toxin belonging to one family and an antitoxin from another (Schmidt et al. 2007). Leplae et al. (2011) recently proposed that toxin families as well as antitoxin families originate from distinct ancestors, and such refer to antitoxin and toxin families independently instead of TA system families.

Internet Resources

CDD, conserved domain database at NCBI (http://www.ncbi.nlm.nih.gov/sites/cdd/)
CG view server, comparative genomics tool for circular genomes (http://stothard.afns.ualberta.ca/cgview_server/)
GOLD, genomes online database(http://www.genomesonline.org)
ICEberg, database of integrative and conjugative elements found in bacteria (http://db-mml.sjtu.edu.cn/ICEberg/)
MEGA5, phylogenetic analysis tool (http://www.megasoftware.net)
PHAST, fast phage search tool(http://phast.wishartlab.com)
jFATCAT, protein 3-D structure comparison tool available at RCSB PDB (http://www.pdb.org/pdb/workbench/workbench.do)
RASTA-Bacteria, prediction tool of type II TA loci in prokaryotes(http://genoweb.univ-rennes1.fr/duals/RASTA-Bacteria/)
TADB, database of type II TA loci in bacteria and archaea (http://bioinfo-mml.sjtu.edu.cn/TADB/)

Acknowledgments This study was supported by grants from the 973 program (2012CB721002), Ministry of Science and Technology, China, the National Natural Science Foundation of China (31170082), and the Program for New Century Excellent Talents in University, Ministry of Education, China (NCET-10-0572).

References

Bi, D., Xu, Z., Harrison, E.M., Tai, C., Wei, Y., He, X., Jia, S., Deng, Z., Rajakumar, K., Ou, H.Y. (2012). ICEberg: A web-based resource for integrative and conjugative elements found in Bacteria. *Nucleic Acids Research*, *40*(Database issue), D621–D626.

Christensen-Dalsgaard, M., Jorgensen, M. G., & Gerdes, K. (2010). Three new RelE-homologous mRNA interferases of *Escherichia coli* differentially induced by environmental stresses. *Molecular Microbiology, 75*(2), 333–348.

Gerdes, K., Christensen, S. K., & Lobner-Olesen, A. (2005). Prokaryotic toxin-antitoxin stress response loci. *Nature Reviews Microbiology, 3*(5), 371–382.

Harrison, E. M., Carter, M. E., Luck, S., Ou, H. Y., He, X., Deng, Z., et al. (2010). Pathogenicity islands PAPI-1 and PAPI-2 contribute individually and synergistically to the virulence of *Pseudomonas aeruginosa* strain PA14. *Infection and Immunity, 78*(4), 1437–1446.

Jiang, Y., Pogliano, J., Helinski, D. R., & Konieczny, I. (2002). ParE toxin encoded by the broad-host-range plasmid RK2 is an inhibitor of *Escherichia coli* gyrase. *Molecular Microbiology, 44*(4), 971–979.

Leplae, R., Geeraerts, D., Hallez, R., Guglielmini, J., Dreze, P., & Van Melderen, L. (2011). Diversity of bacterial type II toxin-antitoxin systems: a comprehensive search and functional analysis of novel families. *Nucleic Acids Research, 39*(13), 5513–5525.

Makarova, K. S., Wolf, Y. I., & Koonin, E. V. (2009). Comprehensive comparative-genomic analysis of type 2 toxin-antitoxin systems and related mobile stress response systems in prokaryotes. *Biol Direct, 4*, 19.

Miclea, P. S., Peter, M., Vegh, G., Cinege, G., Kiss, E., Varo, G., et al. (2010). Atypical transcriptional regulation and role of a new toxin-antitoxin-like module and its effect on the lipid composition of *Bradyrhizobium japonicum*. *Molecular Plant-Microbe Interactions, 23*(5), 638–650.

Pandey, D. P., & Gerdes, K. (2005). Toxin-antitoxin loci are highly abundant in free-living but lost from host-associated prokaryotes. *Nucleic Acids Research, 33*(3), 966–976.

Ramage, H. R., Connolly, L. E., & Cox, J. S. (2009). Comprehensive functional analysis of *Mycobacterium tuberculosis* toxin-antitoxin systems: Implications for pathogenesis, stress responses, and evolution. *PLoS Genetics, 5*(12), e1000767.

Schmidt, O., Schuenemann, V. J., Hand, N. J., Silhavy, T. J., Martin, J., Lupas, A. N., et al. (2007). *prlF* and *yha*V encode a new toxin-antitoxin system in *Escherichia coli*. *Journal of Molecular Biology, 372*(4), 894–905.

Sevin, E. W., & Barloy-Hubler, F. (2007). RASTA-Bacteria: A web-based tool for identifying toxin-antitoxin loci in prokaryotes. *Genome Biology, 8*(8), 155.

Shao, Y., Harrison, E.M., Bi, D., Tai, C., He, X., Ou, H.Y., Rajakumar, K., Deng, Z. (2011). TADB: A web-based resource for Type 2 toxin-antitoxin loci in bacteria and archaea. *Nucleic Acids Research*, *39*(Database issue), D606–D611.

Van Melderen, L., & Saavedra De Bast, M. (2009). Bacterial toxin-antitoxin systems: More than selfish entities? *PLoS Genetics, 5*(3), e1000437.

Wozniak, R. A., & Waldor, M. K. (2009). A toxin-antitoxin system promotes the maintenance of an integrative conjugative element. *PLoS Genetics, 5*(3), e1000439.

Yuan, J., Yamaichi, Y., & Waldor, M. K. (2011). The three *Vibrio cholerae* chromosome II-encoded ParE toxins degrade chromosome I following loss of chromosome II. *Journal of Bacteriology, 193*(3), 611–619.

Chapter 15
Type III Toxin-Antitoxin Loci

Tim R. Blower, Francesca L. Short and George P. C. Salmond

Abstract The *toxIN* locus from a cryptic plasmid of the plant pathogen *Pectobacterium atrosepticum* was the first type III toxin–antitoxin (TA) system to be identified. This new paradigm describes how an RNA antitoxin directly interacts with and inhibits a protein toxin. Within this chapter, we discuss the discovery of *toxIN* as a potent bacteriophage resistance mechanism and the subsequent genetic, biochemical and structural studies that allowed construction of generalised working models for the activities of these novel systems. Furthermore, the recent identification of three independent families of type III systems, suggests the possibility of wider biological roles. Although type III TA loci had perhaps been initially considered a niche field, it is now clear that these systems are widespread, and thus presumably important in the biology of prokaryotes.

15.1 Introduction

The recent concept of "type III" toxin–antitoxin (TA) systems was first generated following studies of the *toxIN* locus, identified on a cryptic plasmid of *Pectobacterium atrosepticum* (formerly *Erwinia carotovora* subspecies *atroseptica*) strain SCRI1039 (Fineran et al. 2009). This addition to the TA field arose from a serendipitous discovery made through studying the bacteriophage (phage) resistance conferred by *toxIN*. Unlike type I, where the antitoxin and toxin interact as RNAs, or type II, where they interact as proteins, type III describes how

T. R. Blower · F. L. Short · G. P. C. Salmond (✉)
Department of Biochemistry, University of Cambridge,
Hopkins Building Tennis Court Road, Cambridge, CB2 1QW, UK
e-mail: gpcs2@cam.ac.uk

K. Gerdes (ed.), *Prokaryotic Toxin-Antitoxins*,
DOI: 10.1007/978-3-642-33253-1_15,

an RNA antitoxin can inhibit a proteinaceous toxin (Blower et al. 2011b). The definition of the new TA type was underpinned by genetic studies (Fineran et al. 2009). The activity of the type III ToxN toxic protein was shown to be essentially linked to the observed bacteriophage resistance phenotype (Blower et al. 2009). Structural studies later revealed that the macromolecular complex formed from pseudoknots of antitoxic RNA coiling around and inhibiting the endoribonuclease toxin (Blower et al. 2011a). These type III systems are now known to be much more widespread than previously thought, with three independent families identified throughout taxonomically distant bacterial species. The relevant loci have been found in bacterial chromosomes and plasmids and a predicted type III system has also been identified within a bacteriophage genome (Blower et al. 2012). This widespread distribution suggests that type III systems may play more important roles in bacterial physiology and evolution than previously considered.

15.2 Abortive Infection and the Discovery of *toxIN*

Though bacteria are considered the most abundant organisms on the planet, they are outnumbered ten-to-one by their viral parasites, bacteriophages (phages). With an estimated $\geq 10^{30}$ phages on the planet (Wommack and Colwell 2000; Chibani-Chennoufi et al. 2004), the infection rate is predicted to be around 10^{25} infectious events per second (Lima-Mendez et al. 2007). The sheer scale of such lethal events provides a constant turnover of biological material, which has an impact on global climate and nutrient cycling (Whitman et al. 1998; Fuhrman 1999). Over millennia, this virus–host interplay has generated multiple back-and-forth exchanges of bacterial phage-resistance mechanisms and counter-mechanisms (Labrie et al. 2010). From the bacterial perspective, protective steps can include avoiding phage adsorption by altering the cell surface, preventing DNA injection, using restriction-modification systems to degrade incoming DNA, acquiring phage-specific immunity through clustered regularly interspaced short palindromic repeats (CRISPRs) (Sorek et al. 2008) and abortive infection (Abi) (Chopin et al. 2005).

"Abi" is a catch-all term that covers many disparate systems. The hallmark of each is that an infected cell carrying an Abi system will undergo cellular suicide after infection with a bacteriophage—akin to a prokaryotic "apoptosis". This also destroys the viral particle and thus protects the clonal bacterial population, in what could be considered an altruistic act on the part of the single cell (Shub 1994) Abi systems have been identified in multiple genera of bacteria, within both Gram-negative species, such as *Shigella dysenteriae*, *Vibrio cholerae*, *Escherichia coli* and *P. atrosepticum* (Smith et al. 1969; Chowdhury et al. 1989; Snyder 1995; Fineran et al. 2009), and also Gram-positive species including *Lactococcus lactis*, where they have been most closely studied (Forde and Fitzgerald 1999; Chopin et al. 2005).

15.2.1 Lactococcal Abortive Infection Systems

Modern industries rely on bacterial fermentation to make a wide range of products (Kutter and Sulakvelidze 2005). The dairy industry openly discusses fermentation failure and has been trying to eradicate the problem for over 70 years (McGrath and van Sinderen 2007). Pasteurised milk is not sterile and the threat from phages present in the starter material has encouraged extensive study of phage resistance mechanisms within lactic acid bacteria (Hill 1993; Forde and Fitzgerald 1999).

There are currently over 20 Abi systems, isolated and defined from *L. lactis*, which act at various stages in the phage life cycle (Chopin et al. 2005). At least six systems are known to be bicistronic, with one component often toxic to the host cell upon over-expression. Through characterisation of phage mutants that have spontaneously "escaped" the effects of an Abi system, it has been possible to begin studies of the modes of Abi action and phage interactions in some systems (Bidnenko et al. 2009; Haaber et al. 2010; Scaltriti et al. 2011). One system (AbiQ) was shown to encode a small protein, which was preceded by a peculiar set of contiguous short tandem repeats in the upstream DNA sequence (Emond et al. 1998).

15.2.2 Analysis of Cryptic Phage Resistance Plasmid, pECA1039

A cryptic plasmid was identified within DNA samples prepared from the plant pathogen *P. atrosepticum* SCRI1039. Analysis of the 5620 bp sequence for this plasmid (pECA1039) identified a ColE1-type origin and approximately 10 potential open reading frames (Fineran et al. 2009). One specific open reading frame encoded a homologue of the AbiQ system from *L. lactis* described above, and was also preceded by a set of tandem repeats. This observation led to experiments to test the potential ability of pECA1039 to provide phage resistance in a related strain of *P. atrosepticum*, the genomically sequenced reference strain, SCRI1043 (Bell et al. 2004). Plasmid pECA1039 very effectively aborted the infection of multiple *Pectobacterium* phages and this ability was abrogated by a frame-shift mutation in the *abiQ* homologue (Fineran et al. 2009) confirming that the *Pectobacterium* AbiQ homologue was indeed the functional Abi protein.

Sub-cloned regions of pECA1039 were then used to test the Abi phenotype in multiple host backgrounds, against suites of phages (Table 15.1). The pECA1039-encoded system provided phage resistance in *Pectobacterium*, *E. coli* and *Serratia* host backgrounds, against a wide range of phages that it has not, previously encountered (Fineran et al. 2009; Blower et al. 2009; Blower et al. 2012). The extent of the viral abortion effect was phage dependent. In some cases, phages remained unaffected. In contrast, the Efficiency of Plating (EOP) of other phages was $<1 \times 10^{-9}$. It was also noted that some phages generated spontaneous "escape" phages, which bred true to form a totally Abi-resistant mutant of that phage (Blower et al. 2009).

Table 15.1 Bacteriophage resistance provided by ToxIN from plasmid pECA1039

Host species for sub-cloned pECA1039		*Pectobacterium atrosepticum* SCRI1043	*Escherichia coli* DH5α	*Serratia marcescens* Db11	*Serratia* 39006
Number of phages tested		26	15	11	14
Number of phages Abi sensitive[a]		16	8	3	13
Range of aborted EOPs[b]	Weakest Abi	3.8×10^{-2}	2.6×10^{-7}	5.0×10^{-2}	4.1×10^{-6}
	Strongest Abi	$<3.6 \times 10^{-10}$	$<4.3 \times 10^{-9}$	$<1.0 \times 10^{-9}$	3.3×10^{-9}

[a] Defined by a reduction in plaque size and/or greater than ten-fold reduction in efficiency of plating

[b] EOP, Efficiency of plating, calculated by (number of plaques on test strain/number of plaques on control strain)

15.3 TA Analysis of the *toxIN* Locus

Initial attempts to clone the pECA1039 *abiQ* homologue in isolation failed. Recombinant plasmids were only obtained when the preceding tandem repeats were included, or a regulated promoter, such as that of pBAD30 (Guzman et al. 1995) was used to repress the expression of the *abiQ* homologue. This result prompted further study of these two components and their regulation.

TA systems are often described through the ability of the antitoxin to negate the toxic effects of the toxin when both components are differentially expressed from separate replicons carrying promoters that can be independently induced. The repetitive nucleotide sequence encoded upstream of the *abiQ* homologue was able to inhibit the bacteriostatic toxicity of the AbiQ homologue itself (Fineran et al. 2009), when both were over expressed. This was the first evidence to suggest that this single Abi protein was in fact encoded as part of a bicistronic TA system. The *abiQ* homologue was then renamed *toxN* (toxin) and the inhibitory upstream repetitive element was named *toxI* (ToxN inhibitor).

15.3.1 Regulation of TA Stoichiometry in the toxIN *Locus*

The *toxN* gene encodes a protein of 171 amino acids. Examination of the upstream DNA showed an inverted repeat sequence immediately preceding the *toxN* ribosome binding site. This Rho-independent transcriptional terminator is preceded by 5.5, 36 nucleotide tandem repeats. This bicistronic arrangement is under the control of a near-consensus sigma 70-type promoter (Fig. 15.1a). Studies using a *lacZ*-reporter plasmid showed that this was a constitutively active promoter (Fineran et al. 2009), and this was further corroborated by the identification of the

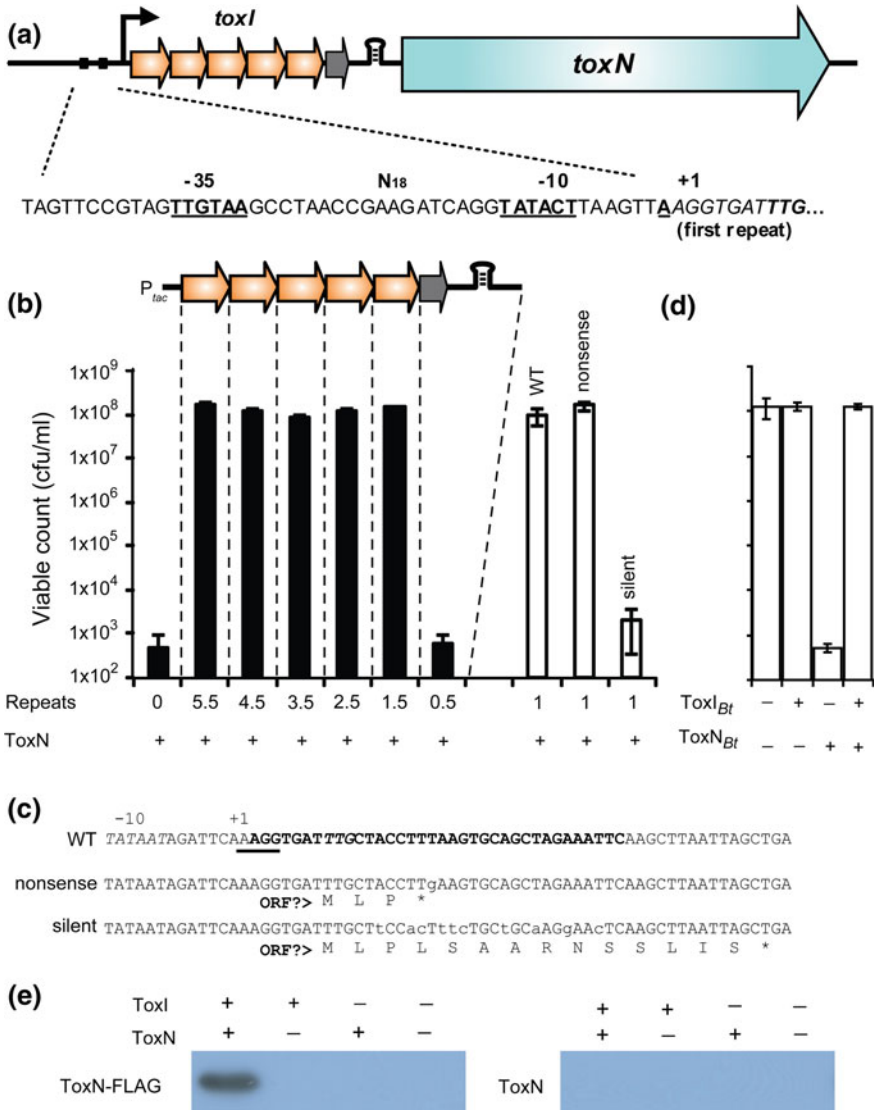


Fig. 15.1 The *toxIN* locus forms a type III, protein–RNA, TA system. **a** Schematic representation of the *toxIN* locus. The transcriptional start (*black kinked arrow* and +1), *toxI* tandem repeats (*orange arrows* for full repeats, *grey arrow* for half-repeat), Rho-independent terminator (hairpin), *toxN* gene (*cyan arrow*), and promoter -35 and -10 elements are indicated. **b** Protection of *E. coli* DH5α from ToxN inhibition by expression of *toxI* deletions composed of 5.5, 4.5, 3.5, 2.5, 1.5 and 0.5 repeats, a wild-type single 36 nucleotide repeat, a nonsense 36 nucleotide repeat mutant or a silent 36 nucleotide repeat mutant. **c** Sequences of the wild type, nonsense and silent construct inserts and the putative short peptides they might encode. A putative ribosome-binding site is underscored, the possible start codon is in italic type and the single repeat is in bold type. **d** Protection of *E. coli* DH5α from ToxN of *B. thuringiensis* by transcription of the cognate ToxI sequence. **e** Expression of ToxI RNA results in the stable production of ToxN-FLAG protein as detected by Western blot. The symbols "+" and "−" refer to induction or repression of the components described in each panel. This figure was adapted from Fineran et al. (2009)

ToxN protein in cells, even prior to phage infection (Blower et al. 2009). Furthermore, the *lacZ*-reporter assays indicated that 90 % of transcripts end at the terminator between the repetitive and protein-coding elements. This is thought to be a vital step in regulating the antitoxin/toxin stoichiometry, prior to activation of the system (Fineran et al. 2009; Blower et al. 2009).

15.3.2 Identification of ToxI as an RNA Antitoxin

Whereas type I TA systems are defined by the two components interacting as RNAs, and type II TA systems interact as proteins (as discussed in other chapters), it was not possible to classify ToxIN as either of these two types. The repetitive antitoxic sequence of ToxI also had the theoretical capacity to encode a short repetitive peptide. A series of experiments were therefore designed to identify whether it was the predicted encoded peptide, or the underlying RNA transcript, that was involved in the inhibition of ToxN (Fineran et al. 2009).

First, the putative *toxI* peptide was tagged with C-terminal DNA extension that could encode a hexahistidine sequence in the translated message. These constructs remained functional as antitoxins, but no ToxI protein was detected by Western blot. Secondly, systematic ToxI deletion constructs were generated in order to assess the importance of each repeat in inhibiting ToxN. Constructs of 5.5–1.5 repeats could inhibit ToxN, as could a singular lone repeat, though 0.5 repeats could not (Fig. 15.1b). Then, focussing upon a single repeat, specific constructs were designed to differentiate between the antitoxic effects of the RNA and any resulting small peptide. The antitoxic activity of a nonsense construct, where translation of the encoded peptide would be truncated, was compared to that from a "silent" construct, in which the peptide was still encoded through redundancy, but the RNA contained multiple base changes (Fig. 15.1c). The nonsense construct remained active, indicating that the RNA sequence was vital for ToxN inhibition (Fig. 15.1b). Furthermore, a homologous ToxIN system from a plasmid of *Bacillus thuringiensis*, in which the *toxI* repeats could not encode a peptide, was also shown to operate as a TA system (Fig. 15.1d).

These results strongly suggested that ToxI was an RNA. To identify whether the ToxN protein was present when inhibited by ToxI, a C-terminal FLAG tag was added to the ToxN protein. This protein was then observed by Western blot during co-overexpression of ToxI and ToxN (Fig. 15.1e). This indicated that both ToxI RNA and ToxN protein co-existed whilst the toxin was inhibited, though at this stage it was not known whether the interaction was direct or indirect. These results led to the formal classification of ToxIN as the first member of the novel, protein-RNA, type III TA systems (Fineran et al. 2009).

15.4 Structural Analysis of the ToxIN Complex

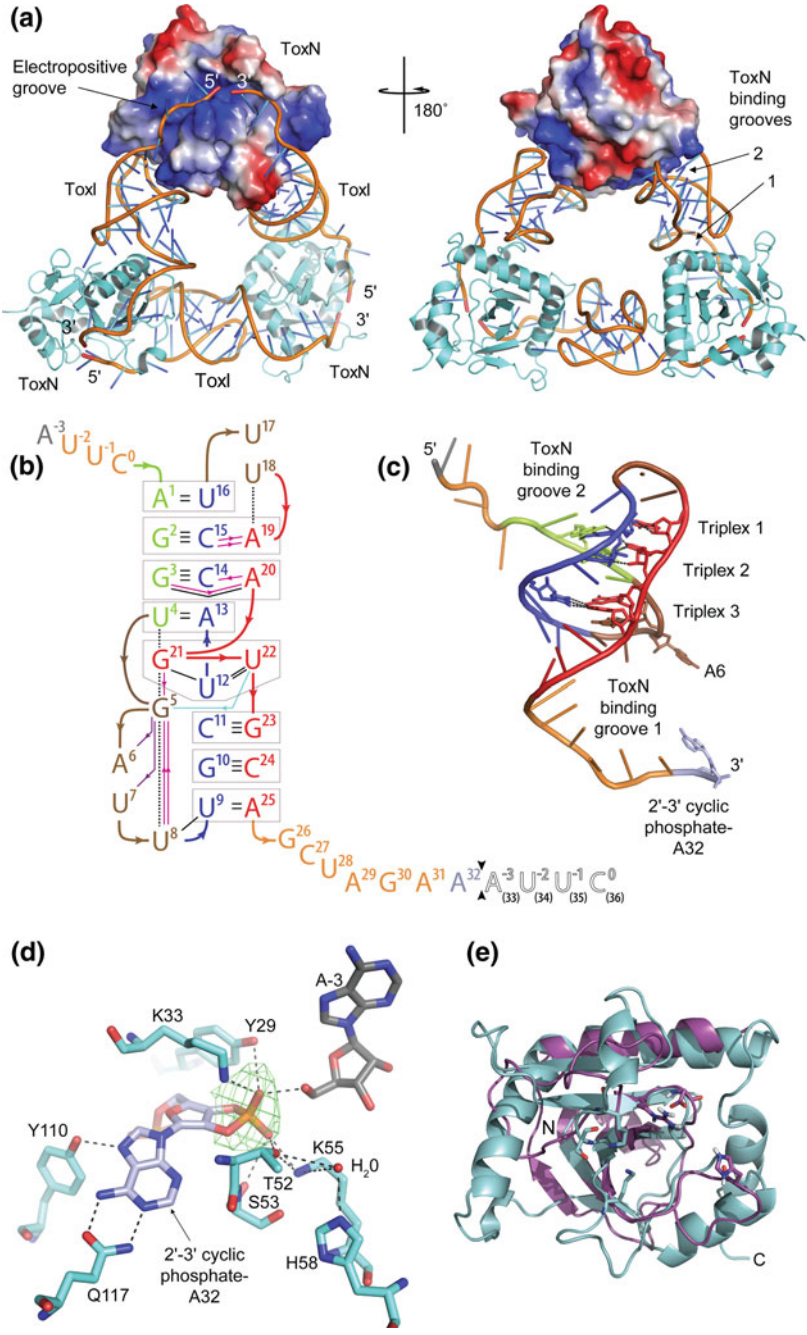
As it had been possible to observe ToxN when co-overexpressed with ToxI, this suggested a route to generate sufficient quantities of the protein sample for structural and biochemical analyses. Following expression and purification, both native and selenomethionine forms of the ToxIN complex were crystallised and the structure solved (Blower et al. 2011a).

15.4.1 ToxIN Forms a Heterohexameric Assembly

The observed repeating structural unit consisted of a ToxN monomer in complex with a 36 nucleotide RNA oligomer, itself derived from a full-length *toxI* transcript. This heterodimeric complex formed into trimers, thereby generating heterohexamers with a triangular architecture (Fig. 15.2a). The 3′ end of each ToxI pseudoknot was adjacent to the 5′ end of the next, in a pseudo-continuous head-to-tail manner, and each ToxI oligomer interacted extensively with two ToxN molecules—one at each terminus (ToxN binding grooves 1 and 2, Fig. 15.2a). Every ToxN molecule thereby interacted with two ToxI molecules over an extended electropositive surface. The buried surface area of ToxN at each protein–RNA interface was roughly 2,000 $Å^2$, corresponding with an avid macromolecular interaction (Chakrabarti and Janin 2002), which is unlikely to occur through crystal contacts alone. Furthermore, the ToxIN trimer was observed in each of three crystal forms, by both crystallographic and non-crystallographic symmetries. Analytical gel filtration was also used to confirm that ToxIN forms a high-molecular weight complex. Altogether, this indicated that the heterohexameric ToxIN was the biologically relevant macromolecular complex (Blower et al. 2011a). In this complex, ToxN has a compact globular fold with a highly twisted, six-stranded, antiparallel β-sheet core surrounded by four α-helices, whilst ToxI forms a convoluted RNA pseudoknot fold.

15.4.2 ToxI Inhibits ToxN Through Formation of a Pseudoknot

The manner by which ToxI inhibits ToxN could not have been predicted prior to solving the complex structure. However, previous genetic analysis did predict that a single 36 nucleotide repeat was sufficient for antitoxicity (Fineran et al. 2009). The structure showed individual 36 nucleotide ToxI RNAs adopting a convoluted pseudoknot structure, caging ToxN prior to activation (Blower et al. 2011a). When the sequence of the pseudoknot was determined, it appeared that each functional 36 nucleotide unit differed from the 36 nucleotide repeat encoded within the DNA of pECA1039, by starting four nucleotides upstream of the DNA repeat


(a)
Electropositive groove
ToxN
5'
3'
ToxI
ToxI
180°
ToxN binding grooves
2
1
5'
3'
3'
ToxN
5'
ToxI
ToxN
(b)
A-3 U-2 U-1 C0
U17
U18
A1 = U16
G2 ≡ C15 A19
G3 ≡ C14 A20
U4 = A13
G21 U22
U12
G5
C11 ≡ G23
A6
G10 ≡ C24
U7
U9 = A25
U8
G26 C27 U28 A29 G30 A31 A32 A-3 U-2 U-1 C0
(33) (34) (35) (36)
(c)
5'
ToxN binding groove 2
Triplex 1
Triplex 2
Triplex 3
A6
ToxN binding groove 1
3'
2'-3' cyclic phosphate-A32
(d)
A-3
K33
Y29
Y110
K55
H2O
T52
S53
2'-3' cyclic phosphate-A32
H58
Q117
(e)
N
C

◀ **Fig. 15.2** ToxIN is formed from ToxI pseudoknots and ToxN endoribonucleases. **a** Overview of the ToxIN heterohexameric macromolecular complex (PDB: 2XDD). Two ToxN protomers are shown as cartoons (*coloured cyan*) and the third as a surface representation, with the positively charged surface in *blue* and the negatively charged surface in *red*. The phosphate sugar backbones of each ToxI are shown as *orange cartoons*, with bases as single sticks. Right, 180° rotation about the vertical axis. **b** Overview of hydrogen bonding in the ToxI pseudoknot. Nucleotides −3 to 32 correspond to one 36 nucleotide RNA oligomer in the crystal structure. Nucleotides 1–36 correspond to a single consensus *toxI* 36 nucleotide repeat. Black arrows between nucleotides 32 and 33 indicate the putative ToxN cutting site in ToxI. The *black-outlined*, open letters for nucleotides 33–36 represent the 5′-most four nucleotides of a second 36 nucleotide ToxI oligomer from the crystal structure. Three interacting sections are shown in *green* (nucleotides 1–4), *blue* (nucleotides 9–16) and *red* (nucleotides 19–25), separated by *brown loops* (nucleotides 5–8 and 17–18). Duplex and triplex base pairs are highlighted by *grey* boxes. The single-stranded RNA tail nucleotides are shown in *orange*, except the termini, with A3 in *grey* and A32 in *violet*. Base–base hydrogen bonds are shown as *black lines*. Ribose 2′OH-base hydrogen bonds are shown as *magenta lines*, ribose 2′OH-phosphate hydrogen bonds as *violet lines* and a phosphate-base hydrogen bond as a light *blue line*; arrows point from the first partner. *Black dashed lines* depict selected stacking interactions. **c** ToxI pseudoknot structure, showing the locations of base triplexes. The main chain backbone and bases are coloured as in (b). **d** ToxN active site interactions with ToxI. A simulated annealing map omitting the electron density for the proposed 2′,3′-cyclic phosphate group is shown in *green*. Hydrogen bonding interactions within 2.6–3.5 Å are indicated by *black dashed lines*, and a water molecule is shown as a *red sphere*. Colouring is as in (a) and (c). (e) Structural overlay of ToxN (PDB: 2XDB) in *cyan* and Kid (PDB: 2C06) in *magenta*. This figure was adapted from Blower et al. (2011a, b)

(Fig. 15.2b). The physiological relevance of this nuance is that each transcript containing 5.5 repeats can be cleaved into four active pseudoknots. It was therefore proposed that the ToxIN heterohexamer folds and assembles following, or in concert with, multiple ribonuclease (RNase) steps that generate the ToxI pseudoknots.

Each ToxI monomer folds as an interdigitated hairpin-type pseudoknot (Puglisi et al. 1988), with two single-stranded tails (Fig. 15.2b, c). The pseudoknot forms from two pairs of complementary stem regions, which together make a central double helical core supported by three triplex base-pair interactions (Fig. 15.2b). The base-pairing strands are connected at either end of the stacked core by looped regions. The first loop provides an additional interaction surface with ToxN, inserting base A6 into a complementary hydrophobic pocket (Fig. 15.2c).

The non-canonical base pairing within the triplex base pairs of ToxI is stabilised through recognised folds such as the A-minor motifs (Wimberly et al. 2000; Nissen et al. 2001). Together with the overall fold, these motifs cause ToxI to closely resemble an artificial RNA aptamer generated to bind vitamin B_{12} (Sussman et al. 2000), and the naturally occurring queuosine riboswitch (Klein et al. 2009). It is therefore a seductive (and testable) hypothesis that ToxI, or analogous type III RNA antitoxin sequences, might bind to small metabolites providing another route to toxin activation under particular physiological or environmental conditions.

15.4.3 ToxN is an Endoribonuclease Related to Type II Toxins Kid and MazF

A 2′,3′-cyclic phosphate was identified by a region of increased electron density around the 3′ end of each ToxI pseudoknot. This suggested that each ToxI can be cleaved from the precursor RNA by intramolecular cleavage, potentially catalysed by ToxN. In vitro analysis of the RNase activity of wild-type ToxN and a less active mutant, identified in previous mutational studies (Blower et al. 2009), confirmed that ToxN was indeed an active endoribonuclease able to cut ToxI, ToxN mRNA and other vital cellular mRNAs (such as that encoding the primary sigma factor RpoD) (Blower et al. 2011a). The active site of ToxN appears to co-ordinate the precursor ToxI RNA for activation of the 2′OH group, leading to cleavage (Fig. 15.2d). The cleaved product would then remain bound to ToxN, occluding the active site and preventing ToxN from inducing bacteriostasis through cleavage of essential cellular mRNAs. ToxN is therefore an interesting example of a small-RNA processing enzyme inhibited by its own non-coding RNA product.

Having solved the structure of ToxN, it was possible to use bioinformatic tools, such as DALI (Holm and Sander 1993), to identify structural homologues held within the Protein Data Bank. It was initially surprising that the highest scoring hits were in fact type II toxins and all members of the Kid/MazF family of endoribonucleases (Hargreaves et al. 2002; Kamada et al. 2003; Blower et al. 2011a. An overlay of ToxN and Kid shows how the core regions are clearly conserved, whilst the decorative helices are significantly different (Fig. 15.2e). This implies the structural core is maintained to provide the ability to bind and cleave varying RNA substrates, whilst the structural variations are likely to provide ToxN with the ability to interact with the RNA pseudoknot antitoxin, ToxI, whereas Kid interacts with its corresponding protein antitoxin, Kis (Blower et al. 2011b).

15.5 Phylogenetic FUGUE Searches and Expansion of Type III System Numbers

When ToxIN of pECA1039 was first characterised, simple BLASTp (Altschul et al. 1990) searches against the ToxN sequence identified 19 homologues in the NCBI sequence database, some of which were associated with ToxI-like tandem repeats (Fineran et al. 2009). With the structure of ToxIN solved (Blower et al. 2011a), it was possible to greatly expand the search for new type III TA systems by using structural, rather than sequence, similarity to ToxN as the initial “hook” to detect potential TA loci. Distant structural homologues of ToxN were identified using FUGUE, which uses specific substitution profiles and gap penalties based on the target structure to detect protein sequences likely to adopt similar folds

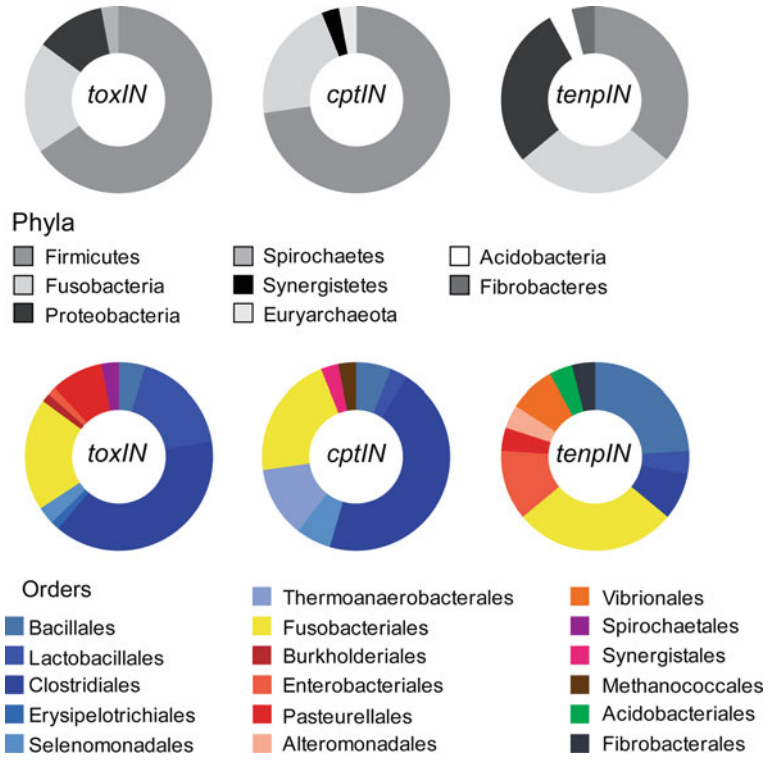


Fig. 15.3 Taxonomy of Type III TA loci. The taxonomic distribution of the identified members from each Type III TA family is shown as pie charts. The *upper row* represents the different Phyla and the *lower row* represents the sub-division of these Phyla into respective Orders. Class has been omitted for clarity. Colours are as indicated in the inset keys. This figure was adapted from Blower et al. (2012)

(Shi et al. 2001). Each hit was then examined for the features of a type III TA locus established for the *toxIN* system; namely, a series of tandem repeats to function as the antitoxin, a terminator to regulate relative levels of toxin and antitoxin synthesis, and a common promoter preceding the bicistronic organisation. Thirty-six putative type III TA systems were identified in this way. Comparison of these new loci showed that the toxin protein sequences fell into three distinct families, so a representative exemplar from each family was used for a further iteration of searches using BLASTp to detect homologous proteins (Altschul et al. 1990). The combined hits from both iterations of bioinformatic searches were collected in a database of type III TA systems, which currently contains 125 loci from three independent families, designated *toxIN*, *cptIN* and *tenpIN* (Blower et al. 2012).

Type III loci were identified in a wide range of prokaryotic genomes, with all three families dominated by entries from the Firmicutes and Fusobacteria (Fig. 15.3). Regarding the lifestyle of the host organism, over half of the 125 type

Table 15.2 Overview of type III TA family properties[a]

Type III Family		*toxIN*	*cptIN*	*tenpIN*
Number of members	Chromosome	57	33	20
	Plasmid	9	0	5
	Phage	1	0	0
	Total	67	33	25
Antitoxin repeat length (nucleotides)	Mean	38.1	44.8	50.9
	Range	31–62	40–48	39–57
Number of repeats	Mean	2.8	2.4	2.2
	Range	1.9–5.6	1.9–3.4	1.9–3.0
Toxin length (amino acids)	Mean	168.0	148.6	164.2
	Range	70–288	45–213	70–259

[a] This data table is adapted from Blower et al. (2012)

III systems are from human commensals sequenced in the Human Microbiome Project. Type III systems were also identified in extremophiles, human and animal pathogens, bacteria from soil and water environments, and one aquatic archaebacterium. Several bacteria contained more than one type III locus on their chromosome (in some cases these were from within the same type III family) while other bacteria contained type III systems from two or all three of the *toxIN*, *cptIN* and *tenpIN* families (Blower et al. 2012). Therefore, it is clear that these systems are widespread and this might imply that they fulfil important and, perhaps, diverse physiological and ecological roles.

To assess the impact of horizontal gene transfer on the dissemination of type III TA loci, phylogenetic analysis of the collected type III toxin proteins and their hosts was performed. The protein phylogenetic tree showed a clear separation of the *toxIN*, *cptIN* and *tenpIN* families (Blower et al. 2012). The divergence of protein sequences was not closely linked with host phylogeny, suggesting that the current distribution of type III TA systems is in part due to horizontal transfer of these modules between taxonomically unrelated bacteria.

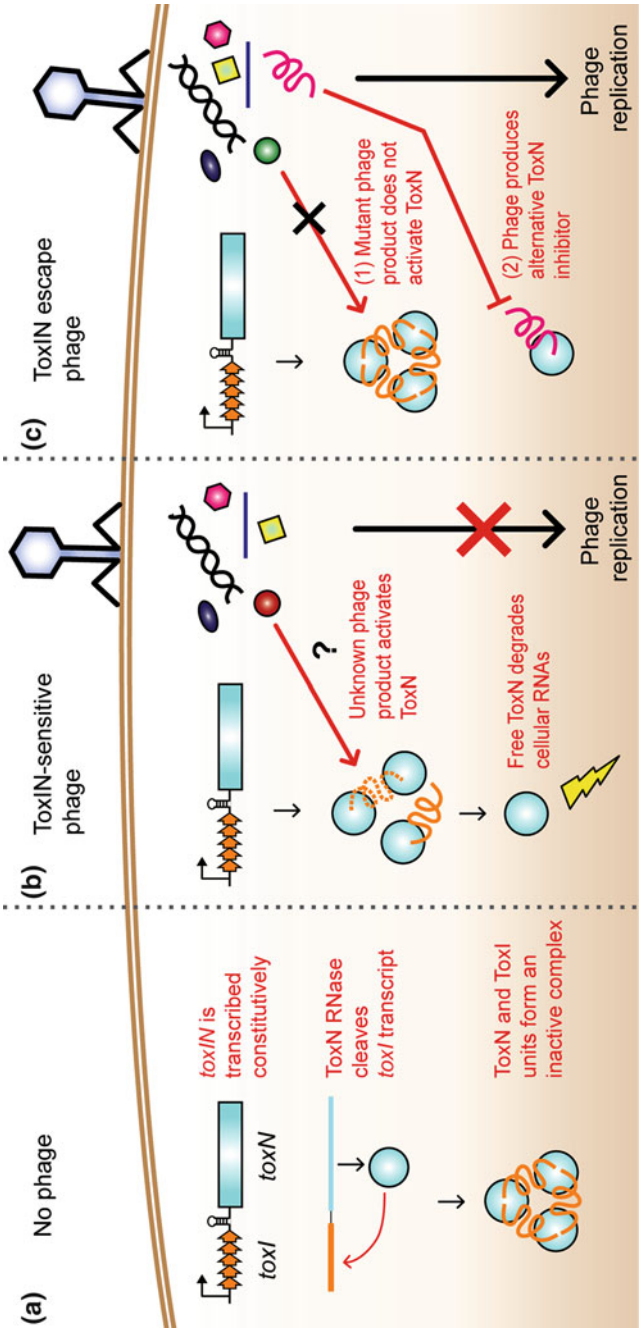
From comparing the features of the *toxIN*, *cptIN* and *tenpIN* families, it appears that there is a shift towards longer length, but lower copy number, in the antitoxic tandem repeats associated with each family (Table 15.2). Despite having entirely unrelated nucleotide sequences, antitoxin sequences from the *toxI* family aligned well onto the structural template of the *P. atrosepticum* ToxI RNA, indicating that the pseudoknot fold of the RNA is likely to be conserved within this family (Blower et al. 2012). In contrast, the *cptI* and *tenpI* antitoxin sequences could not be aligned, to each other or onto the structural template of ToxI, although most of these sequences also contained potential pseudoknot base-pairing regions. In the absence of a structural template it is difficult to predict how these new families of RNA inhibitors might fold and bind to their cognate toxin proteins; new type III systems may show novel RNA folds and complex architectures and it is tempting to speculate that these would vary in a family specific manner. Structural characterisation of more type III TA systems is now required as this is likely to add significantly to current knowledge of protein–RNA interactions.

The TA function of representatives from the new *cptIN* and *tenpIN* families was confirmed in a simple kill/rescue assay outlined earlier. To further define the functional overlap between different type III TA systems, the potential for these loci to mediate phage resistance was also examined. The prototype TenpIN system, from *Photorhabdus luminescens*, conferred high resistance to several coliphage, while two different CptIN systems did not (Blower et al. 2012). At present, therefore, the ability to protect from bacteriophages is one established function for some type III TA systems, but this activity does not appear to be universal. More work will be required to assess the true functional repertoire of type III TA systems, and the physiological, ecological and evolutionary contributions they make to their diverse hosts.

15.6 A Working Model for ToxIN Abortive Infection

Using the combined information on ToxIN function and ToxN toxicity, it is possible to build a model for the activity of ToxIN during a phage infection cycle (Fig. 15.4). In this working model, the *toxIN* locus is transcribed constitutively under normal conditions, and the terminator between *toxI* and *toxN* ensures an excess of antitoxin transcript. ToxN processes the repetitive *toxI* transcript to generate the final inhibitory ToxI RNA pseudoknots, which remain bound and sequester ToxN in an inactive complex (Fig. 15.4a). Infection by phage results in the removal of ToxI inhibition, by an unknown mechanism. The free ToxN protein then degrades cellular RNAs, which arrests cell growth and replication of the phage—eventually leading to bacterial cell death (Fig. 15.4b).

As discussed earlier, phages that are normally sensitive to Abi systems can accumulate mutations allowing them to "escape" the phage resistance mechanism (Bidnenko et al. 2009; Haaber et al. 2010; Scaltriti et al. 2010). In the case of ToxIN-mediated Abi, this model suggests two obvious mechanisms for how escape might occur (Fig. 15.4c). Firstly, the phage product responsible for "activation" of ToxN could be altered or removed, such that phage infection goes undetected by the ToxIN system. Secondly, the bacteriophage could produce an alternative inhibitor to prevent ToxN toxicity following activation by the phage. This second type of interaction mechanism has already been observed between the T4 bacteriophage Dmd protein and the type II toxin LsoA, which is released from its cognate antitoxin following phage infection but is then directly counteracted by Dmd (Otsuka and Yonesaki 2012). As escape mutants of *Pectobacterium* phages have already been isolated (Blower et al. 2009), it will be possible, through their molecular characterisation, to further investigate the less well-understood steps of this proposed model.


(a)
No phage
toxI
toxN
toxIN is transcribed constitutively
ToxN RNase cleaves toxI transcript
ToxN and ToxI units form an inactive complex
(b)
ToxIN-sensitive phage
?
Unknown phage product activates ToxN
Free ToxN degrades cellular RNAs
Phage replication
(c)
ToxIN escape phage
(1) Mutant phage product does not activate ToxN
(2) Phage produces alternative ToxN inhibitor
Phage replication

◀ **Fig. 15.4** Model for abortive infection by ToxIN from *P. atrosepticum*. Schematic of ToxIN activity under normal growth conditions, and following infection by a ToxIN-sensitive or escape mutant bacteriophage. ToxI and ToxN are shown in orange and cyan, respectively. **a** ToxN and ToxI exist as an inert complex under normal conditions. **b** Following phage infection, ToxN is released to degrade cellular RNAs. **c** Phages could escape from ToxIN-mediated abortive infection by either (1) preventing ToxN activation, or (2) by producing an alternative ToxN inhibitor

15.7 Conclusion

The ToxIN locus on a cryptic plasmid of *P. atrosepticum* was first identified as an Abi system. Although the type I *hok/sok* locus from *E. coli* plasmid R1 excludes T4 (Pecota and Wood 1996), and the chromosomal type II *mazEF* system can provide some protection from phage P1 (Hazan and Engelberg-Kulka 2004), the degree of phage resistance provided by ToxIN is much greater, and it can affect a wide range of phages in multiple bacterial host backgrounds. This Abi phenotype was also shown by the type III TenpIN systems (Blower et al. 2012), although has not yet been observed for CptIN loci.

With the recent identification of three entire families of type III loci, and demonstration of their widespread nature (Blower et al. 2012), many questions remain as to the possible biological roles of these systems. As with other types of TA systems, it is likely that the type III systems may fulfil many proposed roles (Magnuson 2007), depending upon host strain and environmental conditions. This is therefore a very exciting time to be exploring the biological impacts of these type III systems. In summary, it seems that Evolution has taken a simple motif that "works"—inhibition of a potentially lethal protein through binding a co-regulated RNA pseudoknot—and applied and exploited this widely in bacteria. This basic molecular mechanism might have been one engine helping to drive adaptive evolution of prokaryotes to overcome, or adapt to, some of the multitude of environmental challenges that face bacterial populations; not least, the "life or death" issue of viral predation.

Acknowledgments Funding is provided by the Biotechnology and Biological Sciences Research Council (UK), and the Department for Business, Innovation and Skills (UK) together with the Cambridge Commonwealth Trust (UK).

References

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology, 215*, 403–410.

Bell, K. S., Sebaihia, M., Pritchard, L., Holden, M. T., Hyman, L. J., Holeva, M. C., et al. (2004). Genome sequence of the enterobacterial phytopathogen *Erwinia carotovora* subsp. *atroseptica* and characterization of virulence factors. *Proceedings of the National Academy of Sciences of the United States of America, 101*, 11105–11110.

Bidnenko, E., Chopin, A., Ehrlich, S. D., & Chopin, M. C. (2009). Activation of mRNA translation by phage protein and low temperature: The case of *Lactococcus lactis* abortive infection system AbiD1. *BMC Molecular Biology, 10*, 4.

Blower, T. R., Fineran, P. C., Johnson, M. J., Toth, I. K., Humphreys, D. P., & Salmond, G. P. (2009). Mutagenesis and functional characterisation of the RNA and protein components of the *toxIN* abortive infection and toxin–antitoxin locus of *Erwinia*. *Journal of Bacteriology, 191*, 6029–6039.

Blower, T. R., Pei, X. Y., Short, F. L., Fineran, P. C., Humphreys, D. P., Luisi, B. F., et al. (2011a). A processed non-coding RNA regulates an altruistic bacterial antiviral system. *Nature Structural and Molecular Biology, 18*, 185–190.

Blower, T. R., Salmond, G. P., & Luisi, B. F. (2011b). Balancing at survival's edge: The structure and adaptive benefits of prokaryotic toxin–antitoxin partners. *Current Opinion in Structural Biology, 21*, 109–118.

Blower, T. R., Short, F. L., Rao, F., Mizuguchi, K., Pei, X. Y., Fineran, P. C., Luisi, B. F., & Salmond, G. P. (2012). Identification and classification of bacterial Type III toxin–antitoxin systems encoded in chromosomal and plasmid genomes. *Nucleic Acids Research*. doi:10.1093/nar/gks231.

Chakrabarti, P., & Janin, J. (2002). Dissecting protein–protein recognition sites. *Proteins, 47*, 334–343.

Chibani-Chennoufi, S., Bruttin, A., Dillmann, M. L., & Brussow, H. (2004). Phage-host interaction: An ecological perspective. *Journal of Bacteriology, 186*, 3677–3686.

Chopin, M. C., Chopin, A., & Bidnenko, E. (2005). Phage abortive infection in lactococci: Variations on a theme. *Current Opinion in Microbiology, 8*, 473–479.

Chowdhury, R., Biswas, S. K., & Das, J. (1989). Abortive replication of choleraphage Φ149 in *Vibrio cholerae* biotype el tor. *Journal of Virology, 63*, 392–397.

Emond, E., Dion, E., Walker, S. A., Vedamuthu, E. R., Kondo, J. K., & Moineau, S. (1998). AbiQ, an abortive infection mechanism from *Lactococcus lactis*. *Applied and Environment Microbiology, 64*, 4748–4756.

Fineran, P. C., Blower, T. R., Foulds, I. J., Humphreys, D. P., Lilley, K. S., & Salmond, G. P. (2009). The phage abortive infection system, ToxIN, functions as a protein-RNA toxin–antitoxin pair. *Proceedings of the National Academy of Sciences of the United States of America, 106*, 894–899.

Forde, A., & Fitzgerald, G. F. (1999). Bacteriophage defence systems in lactic acid bacteria. *Antonie van Leeuwenhoek, 76*, 89–113.

Fuhrman, J. A. (1999). Marine viruses and their biogeochemical and ecological effects. *Nature, 399*, 541–548.

Guzman, L. M., Belin, D., Carson, M. J., & Beckwith, J. (1995). Tight regulation, modulation, and high-level expression by vectors containing the arabinose P_{BAD} promoter. *Journal of Bacteriology, 177*, 4121–4130.

Haaber, J., Samson, J. E., Labrie, S. J., Campanacci, V., Cambillau, C., Moineau, S., et al. (2010). Lactococcal abortive infection protein AbiV interacts directly with the phage protein SaV and prevents translation of phage proteins. *Applied and Environment Microbiology, 76*, 7085–7092.

Hargreaves, D., Santos-Sierra, S., Giraldo, R., Sabariegos-Jareno, R., de la Cueva-Mendez, G., Boelens, R., et al. (2002). Structural and functional analysis of the kid toxin protein from *E. coli* plasmid R1. *Structure, 10*, 1425–1433.

Hazan, R., & Engelberg-Kulka, H. (2004). *Escherichia coli mazEF*-mediated cell death as a defense mechanism that inhibits the spread of phage P1. *Molecular Genetics and Genomics, 272*, 227–234.

Hill, C. (1993). Bacteriophage and bacteriophage resistance in lactic acid bacteria. *FEMS Microbiology Reviews, 12*, 87–108.

Holm, L., & Sander, C. (1993). Protein structure comparison by alignment of distance matrices. *Journal of Molecular Biology, 233*, 123–138.

Kamada, K., Hanaoka, F., & Burley, S. K. (2003). Crystal structure of the MazE/MazF complex: Molecular bases of antidote-toxin recognition. *Molecular Cell, 11*, 875–884.

Klein, D. J., Edwards, T. E., & Ferre-D'Amare, A. R. (2009). Cocrystal structure of a class I preQ$_1$ riboswitch reveals a pseudoknot recognizing an essential hypermodified nucleobase. *Nature Structural and Molecular Biology, 16*, 343–344.

Kutter, E., & Sulakvelidze, A. (2005). *Bacteriophages: Biology and applications*. Boca Raton, Florida, USA: CRC Press.

Labrie, S. J., Samson, J. E., & Moineau, S. (2010). Bacteriophage resistance mechanisms. *Nature Reviews Microbiology, 8*, 317–327.

Lima-Mendez, G., Toussaint, A., & Leplae, R. (2007). Analysis of the phage sequence space: The benefit of structured information. *Virology, 365*, 241–249.

Magnuson, R. D. (2007). Hypothetical functions of toxin–antitoxin systems. *Journal of Bacteriology, 189*, 6089–6092.

McGrath, S., & van Sinderen, D. (2007). *Bacteriophage genetics and molecular biology*. Norfolk: Caister Academic Press.

Nissen, P., Ippolito, J. A., Ban, N., Moore, P. B., & Steitz, T. A. (2001). RNA tertiary interactions in the large ribosomal subunit: The A-minor motif. *Proceedings of the National Academy of Sciences of the United States of America, 98*, 4899–4903.

Otsuka, Y., & Yonesaki, T. (2012). Dmd of bacteriophage T4 functions as an antitoxin against *Escherichia coli* LsoA and RnlA toxins. *Molecular Microbiology, 83*, 669–681.

Pecota, D. C., & Wood, T. K. (1996). Exclusion of T4 phage by the *hok/sok* killer locus from plasmid R1. *Journal of Bacteriology, 178*, 2044–2050.

Puglisi, J. D., Wyatt, J. R., & Tinoco, I, Jr. (1988). A pseudoknotted RNA oligonucleotide. *Nature, 331*, 283–286.

Scaltriti, E., Moineau, S., Launay, H., Masson, J. Y., Rivetti, C., Ramoni, R., et al. (2010). Deciphering the function of lactococcal phage ul36 Sak domains. *Journal of Structural Biology, 170*, 462–469.

Scaltriti, E., Launay, H., Genois, M. M., Bron, P., Rivetti, C., Grolli, S., et al. (2011). Lactococcal phage p2 ORF35-Sak3 is an ATPase involved in DNA recombination and AbiK mechanism. *Molecular Microbiology, 80*, 102–116.

Shi, J., Blundell, T. L., & Mizuguchi, K. (2001). FUGUE: Sequence-structure homology recognition using environment-specific substitution tables and structure-dependent gap penalties. *Journal of Molecular Biology, 310*, 243–257.

Shub, D. A. (1994). Bacterial viruses. Bacterial altruism? *Current Biology, 4*, 555–556.

Smith, H. S., Pizer, L. I., Pylkas, L., & Lederberg, S. (1969). Abortive infection of *Shigella dysenteriae* P2 by T2 bacteriophage. *Journal of Virology, 4*, 162–168.

Snyder, L. (1995). Phage-exclusion enzymes: A bonanza of biochemical and cell biology reagents? *Molecular Microbiology, 15*, 415–420.

Sorek, R., Kunin, V., & Hugenholtz, P. (2008). CRISPR–a widespread system that provides acquired resistance against phages in bacteria and archaea. *Nature Reviews Microbiology, 6*, 181–186.

Sussman, D., Nix, J. C., & Wilson, C. (2000). The structural basis for molecular recognition by the vitamin B$_{12}$ RNA aptamer. *Natural Structural Biology, 7*, 53–57.

Whitman, W. B., Coleman, D. C., & Wiebe, W. J. (1998). Prokaryotes: The unseen majority. *Proceedings of the National Academy of Sciences of the United States of America, 95*, 6578–6583.

Wimberly, B. T., Brodersen, D. E., Clemons, W. M, Jr, Morgan-Warren, R. J., Carter, A. P., Vonrhein, C., et al. (2000). Structure of the 30S ribosomal subunit. *Nature, 407*, 327–339.

Wommack, K. E., & Colwell, R. R. (2000). Virioplankton: Viruses in aquatic ecosystems. *Microbiology and Molecular Biology Reviews, 64*, 69–114.

Chapter 16
Type II Toxin-Antitoxin Loci Encoded by Plasmids

Elizabeth Diago-Navarro, Ana M. Hernández-Arriaga and Ramón Díaz-Orejas

Abstract Toxin–antitoxin (TA) loci were initially identified as auxiliary plasmid maintenance modules (Gerdes 2000). We review a few type II TA systems of the early list that are found in plasmids of Gram-negative bacteria and that have been more extensively characterized. These include the *ccd* system of plasmid F, the *kis-kid* system of plasmid R1, the *higBA* system of plasmid Rts1, and the *parDE* system of plasmid RK2. We also review two systems found in plasmids of Gram-positive bacteria: the ω-ε-ζ system found in a plasmid pSM19035 of *Streptococcus pyogenes* and the *axe-txe* system found in pRUM, a multidrug resistant factor of *Enterococcus faecium*. Most of these systems have been analyzed both at the functional and structural levels and on the whole they provide an insight on the essential features, diversity, relationships, and relevance of type II TA loci.

16.1 Introduction

TA loci were initially discovered in plasmids as small bicistronic maintenance modules. These gene systems encode an unstable antitoxin and a stable toxin, the concerted action of which leads to the selective elimination or growth inhibition of plasmid-free cells. The toxin is always a protein but the unstable antitoxin can be either an RNA (type I or type III TA loci) or a protein (type II TA loci). The first TA systems described were *ccd* of F (Ogura and Hiraga 1983), *hok-sok* (*parB* locus) of plasmid R1 (Gerdes et al. 1986), and *kis-kid* (*parD* locus) also from R1

E. Diago-Navarro · A. M. Hernández-Arriaga · R. Díaz-Orejas (✉)
Centro de Investigaciones Biológicas (CSIC), Ramiro de Maeztu 9,
28040 Madrid, Spain
e-mail: ramondiaz@cib.csic.es

K. Gerdes (ed.), *Prokaryotic Toxin-Antitoxins*,
DOI: 10.1007/978-3-642-33253-1_16,

Table 16.1 Plasmidic TA type II loci described

Operon	Plasmid	Components	Antitoxin/DNA binding motif	Toxin/activity	Protease degrading antitoxin
kis-kid	R1	Kis, Kid	Kis/LHH	Kid/ribonuclease activity	Lon (indirect evidence)
ccd	F	CcdA, CcdB	CcdA/RHH	CcdB/gyrase inhibition	Lon
parD-E	RK2/RP4	ParD, ParE	ParD/RHH	ParE/gyrase inhibition	ND
axe-txe	pRUM	Axe, Txe	Axe/ND	Txe/ribonuclease activity	ND
ω-ε-ζ	pSM19035	ω, ε, ζ	ε/RHH in ω regulatory protein	ζ/kinase	ND
higBA	Rts1	HigA, HigB	HigA/ND	HigB/ribonuclease activity	ND

ND not determined

(Bravo et al. 1987). *hok-sok* is the prototype of type I loci in which the antitoxin Sok is an antisense RNA that inhibits the synthesis of the Hok toxin. *ccd* and *kis-kid* were the two first type II TA loci described in which the antitoxins, CcdA and Kis, respectively, are proteins that interact with and neutralize the activity of their toxins, CcdB, and Kid, respectively. More recently, a novel TA system, ToxIN, has been identified in a plasmid of *Erwinia carotovora* as a phage abortive infection system (Fineran et al. 2009; Blower et al. 2011). ToxIN is the prototype of type III TA systems in which the antitoxin, ToxI, is a short RNA that interacts with and neutralizes the RNase activity of the toxin, ToxN.

Soon after the discovery of the *ccd* and *kis-kid*, further type II TA loci were identified on plasmids. The early list (Gerdes 2000), included the *pem* locus, identical to *parD*, found in plasmid R100 of enterobacteria (Tsuchimoto et al. 1988), the *parDE* system found in the broad-host-range plasmid RK2 (Roberts et al. 1990), the *stbDE* system found in plasmid R485 of enterococci (Hayes 1998), the *phd-doc* system identified in the bacteriophage P1 (Lehnherr et al. 1993), the three-component TA system, ω-ε-ζ found in plasmid pSM19025 of *Streptococcus pyogenes* (Ceglowski et al. 1993a), the *higBA* system of plasmid Rts1 of *Escherichia coli* (Tian et al. 1996b), the tri-component *pasABC* system found in plasmid pTF-CF2 of *Thiobacillus ferroxidans* (Smith and Rawlings 1997), the *stb* system found in plasmid pMYSH600 of *Shigella flexneri* (Radnedge et al. 1997), and the *relBE* system found in plasmid P307 of *E. coli* and in plasmids of other hosts (Gronlund and Gerdes 1999; Gerdes 2000).

In this chapter, we focus in type II TA loci from plasmids but rather than covering the whole available information, we review a few systems of the early list that have been more extensively characterized: these include the *ccd*, *parD*, *higBA*, *parDE*, *phd-doc*, and ω-ε-ζ systems. We also briefly review the *axe-txe* locus of pRUM of enterococci. Most of these systems have been analyzed both at the

functional and structural levels and on the whole they provide a global frame to understand the basic features, function, and relevance of type II TA loci. The *phd-doc* system, originally found in plasmid P1, is the subject of a full chapter of this book and will not be review here. More generally, information on type II systems can be found in (Pandey and Gerdes 2005; Buts et al. 2005; Fineran et al. 2009; Blower et al. 2011), and in (Leplae et al. 2011). Basic features of the systems analyzed in this chapter are given in the Table 16.1.

16.2 *ccd* Locus

ccd of plasmid F was identified as a plasmid maintenance module (Ogura and Hiraga 1983). The *ccd* system is organized as an operon of two genes, *ccdA* and *ccdB*, encoding respectively an antitoxin and a toxin. This system increases the number of plasmid containing cells in the culture due to the CcdB-dependent elimination of plasmid-free segregants (Jaffe et al. 1985). This mechanism, first reported in the *hok-sok* type I TA system of plasmid R1, is called postsegregational killing (PSK) (Fig. 16.1). Conditional activation of the CcdB toxin in these cells was due to the faster decay of the CcdA antitoxin mediated by the ATP-dependent Lon protease (Van Melderen et al. 1994, 1996). A comparative analysis indicated that the system stabilized a mini-R1 plasmid and that this stabilization was similar to the one determined by the *parD* (*kis, kid*) locus of R1 and 10-fold lower than the ones determined by *parDE* of RK2/RP4 and *hok-sok* of R1 (Jensen et al. 1995).

16.2.1 Toxin Target Interactions: Structure and Mechanism of Action of CcdB

The crystal structure of CcdB toxin was the first one reported of a TA locus (Loris et al. 1999) (Fig. 16.2b and d). CcdB is a dimeric protein in which the two subunits are related by a noncrystallographic 2-fold rotation axis. In each monomer there are eight β-strands, five of them arranged in antiparallel orientation forming a main β-sheet in which a minor β-sheet formed by three antiparallel β-strands is inserted. An extended α-helix is located at the C-terminal region. CcdB contains a hairpin at the N-terminal region. The CcdB structure established a reference for further structural studies in this system and, jointly with the structure of the Kid toxin, revealed a common structural module shared by toxins that reach different targets (see below).

CcdB targets the GyrA subunit of DNA-gyrase and inhibits this type II topoisomerase (Bernard and Couturier 1992; Miki et al. 1992; Maki et al. 1996; Critchlow et al. 1997). Bacterial DNA-gyrase is a heterotetrameric protein formed by two dimeric subunits, GyrA and GyrB. This enzyme is unique because it is able to introduce negative supercoiling into the DNA.

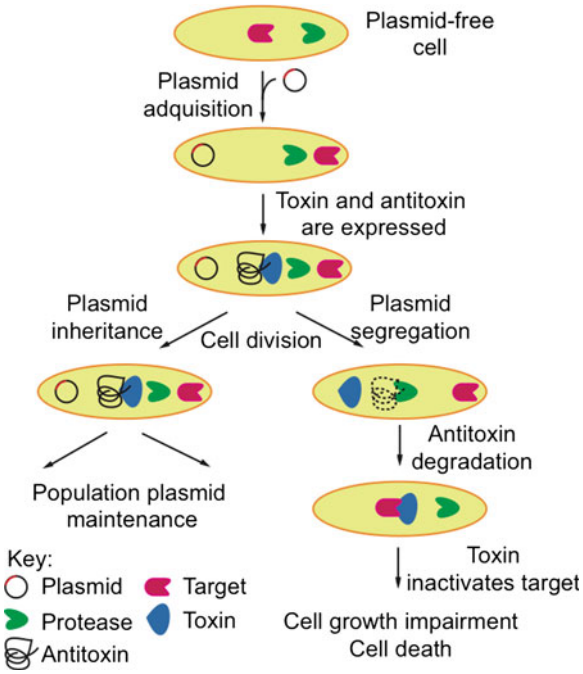


Fig. 16.1 Postsegregational killing mechanism. When a bacterial cell inherits a plasmid containing a TA system, the toxin and antitoxin are produced; the antitoxin neutralizes the toxin and the cell can grow and divide. If the plasmid is inherited at cell division, the continuous synthesis of the antitoxin will keep the toxin neutralized in these cells and this allows cell proliferation. In cells loosing the plasmid at cell division, the continuous degradation of the antitoxin by a specific cell proteases releases the toxin that can inactive its target, thus preventing cell proliferation. The selective proliferation of plasmid containing cells will maintain the extra-chromosomal element in the bacterial population. Adapted from Buts et al. (2005)

The interaction of CcdB with GyrA requires a particular conformation of DNA-gyrase that occurs when the DNA strand passage is taking place (Smith and Maxwell 2006). This interaction is accompanied by an extensive rearrangement affecting the tower and the catalytic domains of the dimeric GyrA subunit (Simic et al. 2009). As a result of the interactions, the enzymatic cycle of DNA-gyrase is frozen at the stage when the DNA strands are cleaved, which leads to DNA lesions and ultimately, cell death (Bernard et al. 1993). Due to the DNA damage CcdB triggers the SOS response and induces lytic propagation of the λ prophage. Trp99, Gly100 and Ile101, the three terminal residues of CcdB, are involved in this process (Bahassi et al. 1995). More recently, the molecular basis of gyrase poisoning by CcdB was established by the resolution of the crystal structure of CcdB in complex with the dimerization domain of GyrA (Dao-Thi et al. 2005) (Fig. 16.2e). In this complex, the C-terminal residues of CcdB are located in close proximity to Arg462 of GyrA, a key residue of the dimerization, and DNA exit gate of this gyrase subunit suggesting that the toxin

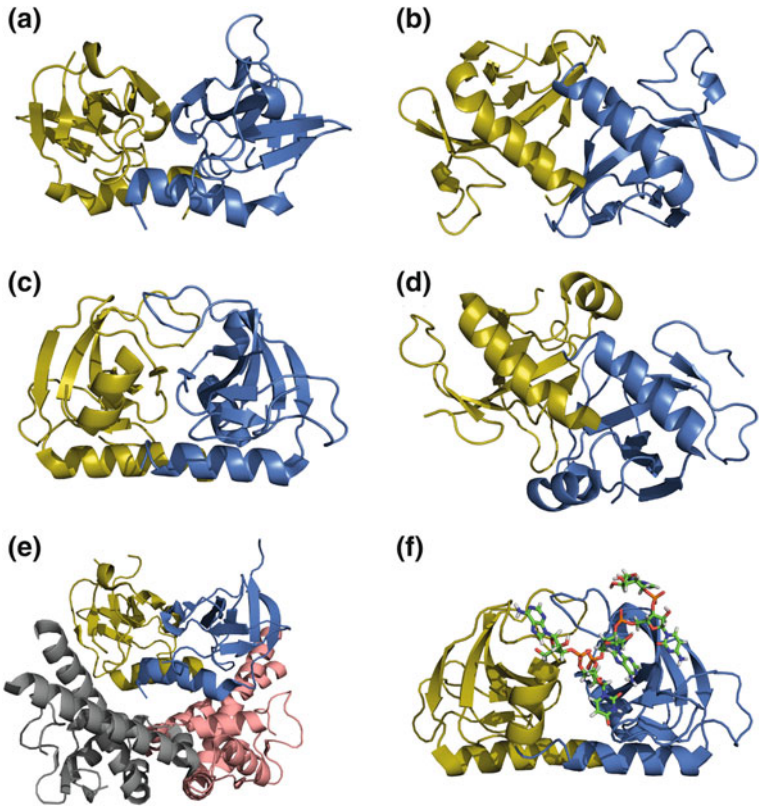


Fig. 16.2 Crystal structures of CcdB and Kid and its interaction with their targets. **a** and **b** show different views (*ribbon* representation) of the dimeric CcdB toxin (PDB code:1VUB) (Loris et al. 1999). **c** and **d** show different views of the crystal structure of the dimeric Kid toxin (PDB code: 1M1F) (Hargreaves et al. 2002). Monomers in the dimers are colored in *olive yellow* and *marine blue*. **e** Shows the dimeric CcdB toxin bound to the GyrA dimerization domain (*gray* and *salmon ribbons*) (PDB code:1X75) (Dao-Thi et al. 2005). **f** Shows the dimeric Kid toxin bound to a short RNA susbtrate (AdUACA, *sticks*) (PDB code: 2C06) (Kamphuis et al. 2006). Representation was generated using PYMOL, version 0.99rc6 (De Lano 2002)

interferes with the function of this dynamic region. Consistently with this, mutants R462C, R462S, and R462A in GyrA confer resistance to the action of the toxin (Bernard and Couturier 1992; Bahassi et al. 1995).

16.2.2 CcdA-CcdB Interactions

CcdA forms tight complexes with CcdB and neutralizes the potential of this protein to interact with GyrA and to inhibit the DNA-gyrase activity. Genetic analysis and the available structure of the C-terminal region of $CcdA_2$ in complex

with $CcdB_2$ indicate that this neutralization is due to the C-terminal region of the antitoxin (Bernard and Couturier 1991; Madl et al. 2006). Upon interaction with CcdB, the disordered C-terminal region of CcdA becomes structured (Madl et al. 2006) and CcdA degradation by the Lon protease is prevented (Van Melderen et al. 1994, 1996).

CcdA, in addition to neutralize the CcdB toxin, is able to remove the toxin from the CcdB-GyrA complex thus "rejuvenating" the DNA gyrase. This leads to the formation of free GyrA and a $CcdB_2$-CcdA complex (Maki et al. 1996; Bahassi et al. 1999). The CcdB toxin has two partially overlapping sites with picomolar and micromolar affinities for CcdA (De Jonge et al. 2009): the picomolar affinity site would play an specific role in "rejuvenation" of DNA gyrase by CcdA, whereas the micromolar affinity site would play a role in the formation of the $CcdB_2$-$CcdA_2$ repressor complexes that occur when CcdA is in excess (see below).

16.2.3 Regulation of ccd *Expression*

Transcriptional regulation of *ccd* is due to specific interactions of the CcdA antitoxin within the promoter–operator (PO) region of the system. This region has three CcdA binding sites, I, II and III, where the antitoxin binds cooperatively with different affinities. One CcdA dimer specifically recognizes the 6 bp palindromic sequence of the high affinity site I. Binding of two additional dimers of CcdA to the lower affinity sites II and III allows direct interactions between these dimers which could explain the observed cooperativity (Madl et al. 2006).

NMR analysis of CcdA-DNA complexes (Madl et al. 2006) shows that the antitoxin binds into the major groove of DNA by insertion of the positively charged N-terminal β-sheet of the ribbon-helix-helix (RHH) motif located in the N-terminal region of the antitoxin (PDB accession numbers of different CcdA-DNA conformers: 2H3A and 2H3C). This interaction is similar to the one found for the MetJ and Arc repressors/DNA complexes (Somers et al. 1994). CcdA-DNA interactions involve three residues of the N-terminal β-sheet (R4, T6 and T8). The DNA-binding region of CcdA is clearly separated from its toxin-neutralizing region: a mutated CcdA protein containing only 41 C-terminal residues was able to neutralize the toxin, but not to autoregulate transcription or to bind to the DNA promoter region (Afif et al. 2001).

CcdB does not bind DNA but interacts with CcdA and this interaction increases the affinity, specificity, and stability of the CcdA-CcdB regulatory complex.

16.2.3.1 Conditional Cooperativity

CcdA and CcdB form different complexes depending on the relative TA ratios (Afif et al. 2001; Dao-Thi et al. 2002). When the amount of antitoxin equals or is in excess of the toxin, $CcdA_2$-$CcdB_2$ complexes are formed that bind to the *ccd* PO

region. A stable $(CcdA_2\text{-}CcdB_2)_n$ regulatory complex, containing multiple DNA binding sites, is assembled as an spiral around this DNA region (Dao-Thi et al. 2002). When further CcdB toxin is added, the protein-DNA complexes are destabilized and a hexamer, $CcdB_2\text{-}CcdA_4$, is formed that fails to bind to DNA. This mechanism has been termed conditional cooperativity, i.e., depending on the ratio of both proteins; the toxin can either enhance or inhibit transcription of this system. This singular regulatory pattern depends on the different affinities of the disordered C-terminal domain of CcdA for two sites in CcdB (De Jonge et al. 2009). The highest affinity site in CcdB is responsible for the formation of the hexameric $CcdB_2\text{-}CcdA_4$ complex which leads to deregulation. The lower affinity site in CcdB is responsible for the formation of the regulatory $CcdA_2\text{-}CcdB_2$ complex (De Jonge et al. 2009). This transcriptional regulatory mechanism has been later found in other TA systems and analyzed in deep in the *phd-doc* system (Garcia-Pino et al. 2010).

16.3 *parD* (*kis-kid*) Locus

Studies carried out by Nordström and by former collaborators, on the replication of the antibiotic resistance factor R1 and its control and on the auxiliary maintenance systems of this plasmid, contribute to understand how extra-chromosomal elements are stably propagated and maintained in the cell (Nordström 2006). Plasmid R1 codes for three maintenance systems: the *par* locus encoding *parMRC* centromer-like stability system, *parB* encoding the *hok-sok* type I TA system, and *parD* encoding the *kis-kid* type II TA system. These systems are dispensable and, in principle, independent of the basic replicon of the plasmid, i.e., they stabilize the plasmid without substantially altering its copy number.

kis-kid was discovered in the $parA^-$ *and* $parB^-$ miniR1-derivative pKN1562, due to the fortuitous isolation of a point mutation in the antitoxin gene that greatly increased plasmid stability without substantially affecting plasmid copy number (Bravo et al. 1987, 1988). The antitoxin and toxin genes of the system, called *kis* (killer suppressor) and *kid* (killing determinant), form an operon that is located in the proximity of the basic replicon of the plasmid. The mutation that activated the system was located within the antitoxin gene and promoted an interference with the growth of the host, particularly at high temperature. Deletions or insertions affecting the toxin gene prevented this interference. Further analysis indicated that the Kis antitoxin repressed transcription of the operon; this repression was an order of magnitude more efficient in the presence of the Kid toxin, implying a coordinated action of the two proteins in the regulation of the *kis-kid* operon (Ruiz-Echevarria et al. 1991a).

16.3.1 Function: Plasmid Maintenance and Replication Rescue

The wild-type (wt) *kis-kid* system mediated a low but detectable plasmid stabilization (Bravo et al. 1987, 1988). In a comparative analysis, it was found that this system, as *ccd*, reduced 10-fold the loss frequency of the plasmid and that displacement of the wt pKN1562 plasmid by the incompatibility determinant CopA, was followed by inhibition of cell growth in plasmid-free segregants (Jensen et al. 1995). This suggested that plasmid stabilization mediated by this system was due to the decay of Kis antitoxin and to the subsequent activation of Kid toxin in plasmid-free cells. A role of ATP-dependent Lon protease in the decay of the Kis antitoxin has been proposed (Tsuchimoto et al. 1992). Surprisingly, it was found that the *kis-kid* system interfered with the isolation of plasmid replication mutants (Ortega et al. 1989). This unexpected result indicated a functional coupling between the plasmid replication and the *kis-kid* TA maintenance modules. Further analysis indicated that *kis-kid* was deregulated and the toxin partially activated in response to inefficient replication and that this leads to a rescue of plasmid stability; surprisingly, a partial recovery of plasmid replication was also observed (Ruiz-Echevarria et al. 1995b). Recent data suggested that the coupling between the replication and *kis-kid* TA modules is probably due to a decrease in the Kis antitoxin level associated with a decrease in plasmid copy number (Lopez-Villarejo et al. 2012). A basal replication rescue of plasmid R1 when this process is inefficient is determined by the gene dosage-dependent decrease in the levels of CopB, a repressor of plasmid replication (Riise and Molin 1986). It was later on found that the RNase activity of Kid selectively reduced the level of *copBmRNA* thus further contributing to the replication rescue mediated by CopB (Pimentel et al. 2005; Lopez-Villarejo et al. 2012) (see below).

16.3.2 The Structure of the Kid Toxin

The resolution of the crystal structure of the Kid toxin was an important reference to further define key interactions of this toxin with its RNA target and its antitoxin Kis (see Fig. 16.2a and c). The structure was solved at a 1.4 Å resolution (Hargreaves et al. 2002) and indicated that, in the crystal, Kid is a dimer formed by two monomers related by a 2-fold noncrystallographic symmetry axis. In each monomer there is a five-stranded antiparallel β-sheet, which forms the core of the monomer. The C-terminal region is formed by a 12 residues organized as a α-helix that makes hydrophobic contacts with the five-stranded β-sheet core. An additional short α-helix makes contacts with this core. Comparison of this structure with the structure of the CcdB toxin indicated that in spite of their functional differences, Kid and CcdB share a common structural module (see later on).

16.3.3 The RNase Activity of the Kid Toxin

The discovery of the endoribonuclease activity of the Kid was independently reported by the laboratory of Inouye and by our laboratory (Zhang et al. 2004; Muñoz-Gomez et al. 2005) following the pioneer discovery of the ribosome-associated RNase activity of the RelE toxin (Pedersen et al. 2003). In contrast to RelE, Kid (PemK) acts as an ribosome-independent endoribonuclease. Kid recognizes and cleaves in vivo and in vitro RNA at 5′-UA(A/C/U)-3′ sequence, either 5′ or 3′ to the first A (Muñoz-Gomez et al. 2005; Zhang et al. 2004). Flanking Us 5′ or 3′ to the core sequence increases cleavage efficiency by Kid, both in vivo and in vitro, but RNA cleavage at core sequences not flanked by Us have also been observed (Zhang et al. 2004; Pimentel et al. 2005; Muñoz-Gomez et al. 2005; Kamphuis et al. 2006; Diago-Navarro et al. 2009).

Based on NMR analysis of Kid-RNA interactions and on the location of mutants affecting specifically the toxicity of Kid, the structure of a complex of the toxin and a 5′-AdUACA-3′ RNA substrate was modelled (Fig. 16.2f). Key residues involved in RNA binding were identified within a concatenated RNA-binding surface formed by residues of both monomers and the catalytic acid (R73), catalytic base (D75) and stabilizing residue (H17) of the active site of Kid were identified (Kamphuis et al. 2006, 2007). Similar to the RNases A and T1, the RNA cleavage by Kid involved the uracil-2′-OH group and yielded two fragments with a 2′:3′-cyclic phosphate group and a free 5′-OH group. The dimeric Kid toxin contains two symmetric RNA binding and catalytic sites but native mass spectrometry (NMS) analysis of Kid-RNA complexes indicates that a dimer of the Kid toxin interacts with a single RNA substrate. Thus, RNA binding to one of the two symmetric binding surfaces interferes with the binding of a second RNA substrate (Kamphuis et al. 2006). The predictions of this model have been further assessed and validated analyzing a significant collection of Kid mutants involved in RNA binding and RNA cleavage (Diago-Navarro et al. 2009).

Due to its RNase activity, the Kid toxin inhibits protein synthesis both in prokaryotes and eukaryotes. Early assays indicated that Kid, at difference of CcdB, does not induce the SOS system but inhibits the infection of the cells by bacteriophage λ and prevents the lytic induction of the λ prophage (Ruiz-Echevarria et al. 1991b). Purified Kid inhibits initiation of ColE1 replication (Ruiz-Echevarría et al. 1995). This inhibition was specific as it was prevented by Kis antitoxin and by mutations that inactivated Kid toxicity. Inhibition of ColE1 replication by Kid correlates with the cleavage of the RNA transcript required for priming replication initiation of this plasmid (Muñoz-Gomez et al. 2005; Zhang et al. 2004). Kid inhibits de novo initiation of bacteriophage λ replication but it does not affect replication of the DNA copy that inherited the initiation complex. It was suggested this inhibition could be related to the decay of the unstable λO initiation protein as the consequence of the inhibition of protein synthesis by the toxin (Potrykus et al. 2002).

16.3.4 *Comparison of* ccd *and* kis-kid

Sequence alignments revealed similarities between CcdA and Kis antitoxins (Ruiz-Echevarria et al. 1991b). CcdA and Kis also share a common organization: both proteins contain an N-terminal DNA binding domain (CcdA a RHH and Kis a loop-helix-helix (LHH) domains), and a partially disordered C-terminal region that can neutralize the toxin (Santos-Sierra et al. 2002; Bernard and Couturier 1991; Salmon et al. 1994; Kamphuis et al. 2007; Madl et al. 2006). Attempts to cross-neutralize the Kid toxin with the wt CcdA antitoxin and to select single mutations that enhanced this neutralization failed, indicating a functional compartmentalization of these systems. Furthermore, significant homology between Kid and CcdB toxins could not be traced by an amino acidic sequence alignment, a result that was apparently consistent with the different targets of these toxins (Ruiz-Echevarria et al. 1991b). However, when the structure of the dimeric Kid toxin was obtained and compared with the structure of the CcdB toxin, it was clear that these toxins share related tertiary folds (Hargreaves et al. 2002) (Fig. 16.2a–c, b–d). Key residues and structural elements required for RNA binding and RNase activity found in Kid where absent in CcdB. The available information on the structures of the toxins with their targets and the different locations in the common module of mutants that specifically affect Kid or CcdB toxicity open the possibility that a common ancestor could have both activities. These data established the first toxin superfamily in TA systems.

16.3.5 *Neutralization of the Kid Toxin by the Kis Antitoxin*

Genetic analysis indicates that Kis antitoxin has two clearly differentiated regions involved in regulation of the system (the N-terminal region) and in neutralization of Kid toxin (C-terminal region) (Santos-Sierra et al. 2002). NMS analysis indicated that depending on the antitoxin–toxin ratios, both proteins can form complexes with different stoichiometries (Fig. 16.3) (Kamphuis et al. 2007). In moderate excess of the toxin the main specie formed is a linear Kid_2-Kis_2-Kid_2 heterohexamer in which a single antitoxin dimer neutralizes two toxin dimers. TA interaction regions were identified by NMR analysis in the heterohexamer which structure was modelled on the available crystal structure of the homologous $MazF_2$-$MazE_2$-$MazF_2$ heterohexamer (Kamphuis et al. 2007; Kamada et al. 2003). The model indicates that the Kis antitoxin is structured at its N-terminal region as a dimer containing a LHH DNA-binding domain; two C-terminal tails are projected divergently from this dimeric amino terminal part; each one of these tails invades key regions of a Kid dimer required for RNA binding and cleavage (site I), thus inactivating simultaneously the two active sites of the toxin. As in the MazE-MazF heterohexamer, there is an additional interaction interface (site IV) involving

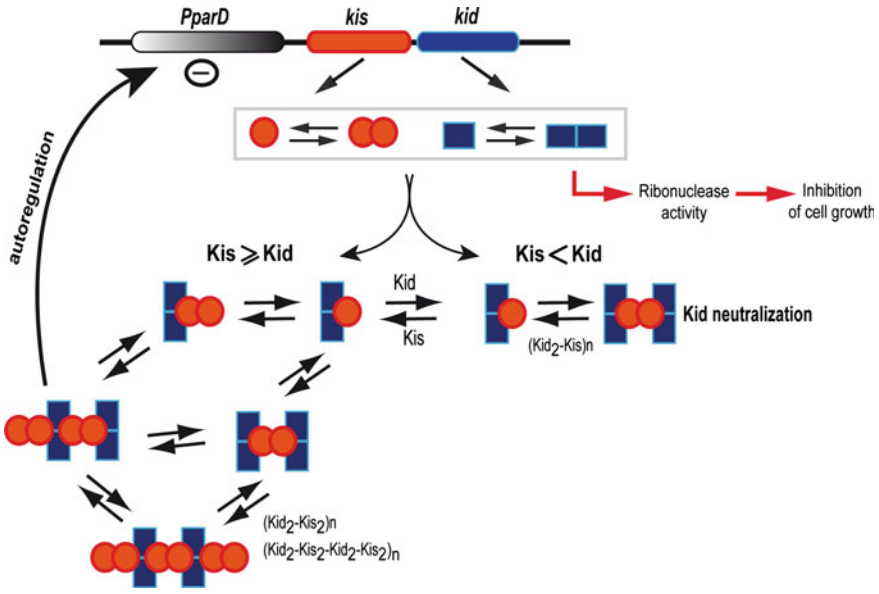


Fig. 16.3 Regulation of the *kis-kid* operon. Model for the transcriptional regulation of the *kis-kid* operon based on the Kis-Kid ratio. Both free Kis and Kid proteins exist in equilibrium between monomeric and dimeric forms. Kis and Kid interact to form Kis-Kid complexes of different stoichiometries. In excess of the toxin, a heterohexameric complex Kid_2-Kis_2-Kid_2 is the main species detected although higher order $(Kid_2\text{-}Kis)_n$ complexes can also be formed. These complexes have a low binding affinity for the specific sequences of the PO region and this increases transcription activity of the *kis-kid* operon. In excess of the antitoxin different complexes such as $(Kid_2\text{-}Kis)_n$, $(Kid_2\text{-}Kis_2)_n$ and $(Kid_2\text{-}Kis_2\text{-}Kid_2\text{-}Kis_2)_n$ can be detected but in the presence of the specific sequences of the PO region, a Kid_2-Kis_2-Kid_2-Kis_2 heterooctamer bound to these sequences is the only complex detected. The heterooctamer contains two dimers of the antitoxin, properly spaced by a dimmer of the toxin, that interact with the arms of palindromic sequences present in the two repressor binding sites of the PO region. These complexes repress transcription of the operon and keep the toxin neutralized

residues of the Kid dimer and other residues of the structured amino-terminal region of the antitoxin (Kamphuis et al. 2006).

Functional and structural information on the cross neutralization of Kid toxin by the ChpAI/MazE antitoxin have been obtained in order to understand the molecular basis of the compartmentalization of the *kis-kid* and *chpA*/*mazEF* systems. At the basal level of expression, the neutralization of Kid toxin by the homologous ChpAI antitoxin was undetectable. However, increasing the gene dosage of this system by cloning it in the multicopy pUC18 plasmid leads to partial neutralization of the Kid toxin (Santos-Sierra et al. 1997). Cross interactions in the ChpAI-Kid complex were further analyzed by NMR and NMS (Kamphuis et al. 2007). Surprisingly, a ChpAI-Kid heterotetramer, but no a heterohexamer, was formed in excess of the Kid toxin and interactions at sites I and IV detected in the Kis-Kid complex were lost. The analysis further indicated that

in this heterotetramer, the two C-terminal regions of the dimeric ChpAI antitoxin were required to neutralize the two active sites of a single dimeric Kid toxin. This structural characterization established on a molecular basis the differences that lead to the functional compartmentalization of the two homologous systems.

16.3.6 Regulation of the kis-kid *System*

kis-kid is regulated at the transcriptional level by the concerted action of the Kis and Kid proteins. In excess of Kis, a Kis-Kid complex represses *kis-kid* transcription. This was assessed combining EMSA, NMS assays of protein-DNA complexes, and footprinting analysis of protein-DNA interactions (Monti et al. 2007). EMSA assays indicated that Kis alone binds with low affinity to a fragment containing the *kis-kid* PO region. However, in the presence of Kid, the affinity increases one order of magnitude. In addition, the relative TA ratios influence this interaction, i.e., in excess of Kis antitoxin a main Kis-Kid/DNA complex could be observed by EMSA assays that was more stable than the larger Kis-Kid/DNA complexes observed in excess of the Kid toxin. NMS assays indicated that the antitoxin contacts the DNA as a dimer and that the main specie formed on the DNA, when the concentration of the antitoxin equals or exceeds the concentration of the toxin, is a Kid_2-Kis_2-Kid_2-Kis_2 heterooctamer. Tandem mass spectrometry analysis showed two antitoxin dimers of this heterooctamer are stabilized by DNA interactions. At difference, a DNA free Kid_2-Kis_2-Kid_2 heterohexamer (see above) is the main specie formed in moderate excess of the toxin and further analysis indicated that it binds poorly to the *kis-kid* PO region (Fig. 16.3). This behavior is the signature of conditional cooperativity that was first reported in the *ccd* system (see Sect. 2.3.1).

Footprinting analysis indicates the repressor complex contacts two repetitions of the PO region, one of them (site I) overlapping with the −10 promoter box and the other (site II) spaced by 33 bp and located upstream the −35 promoter box (Monti et al. 2007). Contacts at sites I and II occur at two discrete regions within the arms of an inverted sequence present in these repetitions. The antitoxin makes the same discrete contacts than the TA complex implying that it pilots the interaction of the Kis-Kid repressor complex with the DNA.

The integration of these data indicates that dimers of the toxin bridge two antitoxin dimers, stabilize the regulatory complex on the DNA and act as an spacer that positions favorably the antitoxin dimers to make specific contacts with the two arms of an internal inverted repetition present in sites I and II (Monti et al. 2007).

In addition to the above, the Kis and Kid levels are also regulated by translational coupling, by selective degradation of the polycistronic *kis-kid* mRNA and by the selective action of ATP-dependent proteases on Kis. Translational coupling between *kis* and *kid* was suggested by the overlap of the last codons of *kis* gene with the translation initiation signals of *kid* gene (Bravo et al. 1987). A detailed analysis, using *kid* translational fusions and different constructs that truncated the

synthesis of Kis antitoxin, indicated that synthesis of Kid is indeed dependent on the synthesis of Kis (Ruiz-Echevarria et al. 1995a). Analysis of *kis-kid* transcripts indicated that degradation of the polycistronic *kis-kid* mRNA molecule lead to the accumulation of a short mRNA containing the *kis* coding sequence but not the *kid* gene, therefore increasing the relative dosage of the *kis-mRNA* (Ruiz-Echevarria et al. 1995a). This could contribute to maintain the excess of the antitoxin required to keep the system regulated. Preferential activity of particular ATP-dependent protease on the Kis antitoxin is another level at which the TA ratio, and therefore the regulation of the system, could be modulated. Although the involvement of Lon protease in Kis stability has been proposed, a precise evaluation of the role of different ATP-dependent proteases of *E. coli* on the stability of this antitoxin remains to be done. A review of the *kis-kid* system and of its relationships with other TA systems has been recently published (Diago-Navarro et al. 2010).

16.4 *parDE* Locus

parDE was initially identified in the broad-host-range plasmid RK2 as a maintenance system (Roberts and Helinski 1992; Roberts et al. 1994). ParD forms a tight complex with ParE and thereby neutralizes the toxin (Roberts et al. 1994; Jensen et al. 1995; Johnson et al. 1996); in addition, ParD plays a role in the regulation of the *parDE* operon (Davis et al. 1992). ParE inhibits cell growth, causes cell filamentation and cell death. The *parDE* system contributes efficiently to the stable maintenance of the broad-host-range plasmid RK2 in *E. coli* and in different Gram-negative bacteria (Roberts and Helinski 1992; Thomas and Helinski 1989). A comparative analysis showed that the system conferred a 100-fold stabilization of a mini R1 replicon in *E. coli*, an effect that was similar to the one mediated by the *hok-sok* system and 10-fold higher than the one conferred by the *ccd* or *kis-kid* systems (Jensen et al. 1995). *parDE* homologues loci have also been found in the chromosome of bacteria as *Caucolobacter crescentus* and *Vibrio cholerae* (Pandey and Gerdes 2005).

16.4.1 Transcriptional Regulation of parDE

ParD antitoxin binds as a dimer to specific sequences of the *parDE* promoter region and thereby represses *parDE* transcription (Davis et al. 1992; Eberl et al. 1992; Roberts et al. 1993). The stoichiometry of ParD–DNA complex has been estimated to be 4:1 (Oberer et al. 1999). ParD protein binds to a 33 bp sequence extending from position −32 to +2, respect to the transcriptional start site of the *parDE* operon (Roberts et al. 1993; Johnson et al. 1996). A palindromic DNA sequence (5′-CACAT/ATGTG-3′) in this region has been proposed as the main operator. Increasing the concentration of ParD extends the protection to 19 bp

downstream of the transcription initiation site. Partial sequence similarity with the palindromic sequence is found in this region, suggesting a second weaker binding site for ParD (Roberts et al. 1993). The specific interaction of ParD in the *parDE* PO region overlaps with the RNA polymerase binding sites (from −45 to +20 positions). Thus, the authors suggested that the binding of this antitoxin to its operator DNA sequences could inhibit transcription of *parDE* by steric hindrance of RNA polymerase binding (Roberts et al. 1993). ParD and ParE form a heterotetrameric $ParD_2$–$ParE_2$ repressor complex, whereas the toxin on its own does not binds to the promoter region. ParE does not alter the DNA binding specificity of ParD. Interestingly, the presence of the toxin has, if any, a minor effect in repression of the *parDE* promoter, implying that the antitoxin plays the main role in regulation of the system (Johnson et al. 1996).

16.4.2 Toxicity of ParE

ParE of RK2 inhibits cell growth and causes cell filamentation and death of plasmid-free segregants. The toxic activity is located at the C-terminal region of the protein (Roberts and Helinski 1992; Fiebig et al. 2010). ParE targets the *E. coli* DNA-gyrase, unwinds supercoiled DNA and can inhibit DNA replication of an *oriC* supercoiled template in a *E. coli* cell extract (Jiang et al. 2002). ParE interacts with one of the subunits of the gyrase and stabilizes a gyrase-DNA cleavable complex (Jiang et al. 2002). This activity is similar to the one of CcdB toxin and quinolone antibiotics (Couturier et al. 1998; Maxwell 1997). The toxic activity of ParE in vitro is dependent on the presence of ATP and results in a cleavable gyrase-DNA complex. The inactivation of gyrase by ParE is prevented by the presence of the specific antitoxin, ParD, that can also rescue the ParE toxin in complex with the DNA gyrase and "rejuvenate" this key replication protein (Jiang et al. 2002). The cell filamentation induced by ParE is likely due to the induction of the SOS system as consequence of the DNA cleavage promoted by the anti-topoisomerase activity of this toxin.

16.4.3 Structural Information on **parDE**

NMR and CD spectroscopy analysis indicated that ParD consists of a well-structured N-terminal domain and a relatively unstructured C-terminal domain (Oberer et al. 2002). The N-terminal region is involved in *parDE* transcriptional regulation and the C-terminal region in neutralization of the ParE toxin. The unstructured C-terminal end of the protein (Roberts et al. 1993) acquires a higher degree of order upon interaction with the ParE toxin (Oberer et al. 2007).

The repressor activity of the dimeric ParD antitoxin is mediated by a RHH motif located in the N-terminus. This motif is similar to the one found in the Arc,

CopG or MetJ repressor proteins (Breg et al. 1990; del Solar et al. 2002; Rafferty et al. 1989). Based on comparison with the Arc-repressor-DNA complex (Raumann et al. 1994), a model of a ParD dimer bound to a 10 bp inverted repeat sequence of the *parDE* operator has been proposed (Oberer et al. 2007). According to this model, the β-sheet of the RHH domain of ParD is inserted into the major groove of the DNA with the protein 2-fold axis matching the center of the sequence. ParD protein is a dimer in solution (Johnson et al. 1996; Oberer et al. 1999, 2007), and binds to a dimer of ParE forming a tetrameric complex, $ParD_2$-$ParE_2$, in which the toxin is neutralized (Johnson et al. 1996). Two dimers of ParD could bind via the RHH domain to two adjacent major grooves in the *parDE* operator DNA.

ParE of RK2 belongs to the RelE/ParE superfamily which includes, in addition to RelE and ParE the RNases YoeB, YafQ, HigB, Txe, and YhaV toxins (Francuski and Saenger 2009). While the structure of several members of this family has been solved (Li et al. 2008; Kamada and Hanaoka 2005), the crystallographic structure of the ParE toxin is not available. However, a structural model of ParE has been obtained using the structure of RelE of *E. coli* as template (Barbosa et al. 2010). According to this model, ParE contains a core of four-stranded β-sheets flanked by three α-helices. A surface topography analysis identified a cavity formed by the region delimited by the C-terminal α-helix and the loop connecting two of the β-sheets, $\beta 2$ and $\beta 3$, which is indicative of a possible RNase active site (Barbosa et al. 2010). However, up to now, this activity has not been detected and instead an anti-topoisomerase activity of this protein has been demonstrated (see above).

Surrogated information on the structure of the ParD and ParE proteins and on the ParD-ParE complex of plasmid RK2 was given by the crystal structure of the ParD1-ParE1 complex from the homologous system found in the chromosome of *C. crescentus* (Dalton and Crosson 2010). As ParD-ParE of RK2, the ParD1-ParE1 complex is a heterotetramer formed by two antitoxins and two toxins monomers. The two ParD1 antitoxin monomers dimerize by the amino-terminal region forming a RHH DNA binding motif. Two long C-terminal helices are projected divergently from this structured region and bind each one, to a groove present in each monomer of the ParE toxin dimmer. Hydrophobic and also complementary charged residues of these proteins are involved in these interactions. ParD1 antitoxin, as ParD of RK2, is a dimer in solution, but in contrast to it is largely α-helical and stably folded. Furthermore, at difference to ParD of RK2, the binding of ParD1 to ParE1 does not undergo any detectable or important conformational change in the antitoxin structure (Dalton and Crosson 2010).

The *C. crescentus* ParE1 shows structural homology with *E. coli* YoeB and RelE endoribonucleases, in spite of differences in the length of the terminal α-helices. However, ParE1 does not contain the catalytic residues required for mRNA cleavage on the ribosome. Instead, ParE1 toxin, as ParE from RK2 (Jiang et al. 2002), might act as an inhibitor of the DNA gyrase.

An interesting variant of ParE from RK2 is found in the ParE2 toxin encoded by *V. cholerae* chromosome. This toxin shows a modest similarity to ParE of RK2 and it inactivates DNA gyrase by targeting GryA subunit and stalling the DNA-gyrase

cleavage complex. However, this inactivation occurs by a mechanism different to other gyrase inhibitors, as ParE2 interferes with gyrase-dependent DNA supercoiling but not DNA relaxation and it can inhibit a set of strains resistant to other gyrase inhibitors (Yuan et al. 2010).

16.5 *higBA* Locus

Plasmid Rts1 present in *Proteus vulgaris* confers temperature sensitive growth to host cells. The *hig* locus is an operon of two genes, *higB* and *higA*, coding for two small proteins HigB and HigA, and was found to be responsible of the thermosensitive phenotype. HigB protein seems to interfere with the propagation of plasmid-free cells, while the HigA supressed, both in *cis* and in *trans*, this interference (Tian et al. 1996b). It is proposed that this locus is responsible for selectively killing cells that have lost the plasmid at the nonpermissive temperature and that this killing was due to the decay of the antitoxin and the subsequent activation of the toxin.

16.5.1 Regulation of the higBA *Operon*

Uncommon to most TA operons, the toxin gene precedes the antitoxin gene. The operon contains two promoters, a strong one preceding *higB*, the first gene of the operon, and a second weaker promoter within the *higB* coding region. HigB and HigA regulate coordinately the first promoter but not the second one (Tian et al. 1996a). A GST-tagged HigA dimeric protein interacts with the region containing the strong promoter covering its −10 and −35 promoter sequences but does not bind to the second promoter. HigA and HigB interact to form a highly stable complex in which the proteins are in equimolar ratios (Tian et al. 2001).

16.5.2 Toxicity of HigB

HigB is an endoribonuclease belonging to the RelE family. It binds the 50S subunit of the ribosome where it cleaves both in-frame and out-of-frame AAA RNA sequences. As the *Proteus* genome is AT rich and AAA, coding for lysin, is one of the most frequently used codons in this specie, it has been suggested that the specificity of the HigB RNase reflects a close relationship between the Rst1 plasmid and this host (Hurley and Woychik 2009).

Interestingly, two functional *higBA* loci were later identified in the superintegron of *V. cholera* (Christensen-Dalsgaard and Gerdes 2006). The toxins of these systems are ribosomal-dependent RNases that act by inhibiting cell growth rather

than by killing the host cells. Both systems are induced by amino acid starvation in *E. coli* and in *V. cholerae*, and are able to stabilize plasmids thus suggesting that they may contribute to the maintenance of the superintegron (Christensen-Dalsgaard and Gerdes 2006; Budde et al. 2007). A functional *higBA* system was also reported in *Mycobacteriun tuberculosis*. The activity of this system seems to be controlled by a SecB-like chaperone that prevents the inactivation of the antitoxin by aggregation or proteolysis (Bordes et al. 2011).

16.6 *ω-ε-ξ* Locus

ω-ε-ξ was originally discovered in pSM19035, a low copy plasmid of the Inc18 group found in a clinical isolates of the Gram-positive human pathogen *Streptococcus pyogenes* and conferring high resistance to erythromycin and lyncomicins (Ceglowski et al. 1993b; Dixon and Lipinski 1972; Zielenkiewicz and Ceglowski 2005). Later on, this system was found in plasmids carrying antibiotic resistances in other Gram-positive pathogens, including vancomycin resistant enterococci (Klare et al. 1995; Kuhn et al. 2005; Mutschler and Meinhart 2011).

16.6.1 Genetic Organization and Function

ω-ε-ξ is an atypical proteic plasmid addiction system. At variance with most other TA systems, *ω-ε-ξ* is a tricistronic operon repressed by a global regulator. Epsilon and Zeta are respectively the antitoxin and toxin proteins of the system and omega is a third component that regulates the system at the transcriptional level independently of the epsilon antitoxin and the zeta toxin (Ceglowski et al. 1993a, b). As in other proteic TA systems, the antitoxin is less stable than the toxin and this is important for the activation of the toxin in plasmid-free segregants, and therefore for *ω-ε-ξ* mediated plasmid stabilization via PSK (Lioy et al. 2006; Zielenkiewicz and Ceglowski 2005; Camacho et al. 2002). The three components are required for this stabilization: inactivation of the toxin or overexpression of the antitoxin that occurs in deregulated variants of the system leads to increase level of the antitoxin that neutralizes the toxin and abolishes PSK stabilization (Zielenkiewicz and Ceglowski 2005).

The efficiency of the epsilon-zeta system as a postsegregational stabilization cassette depends on the differential stability of the two proteins, on the intracellular levels of these proteins and also on the stability of the TA complexes. Interestingly, there are two versions of the epsilon-zeta system in plasmid of *Enterococcus faecalis* having differences in the TA interface, but not in the active site of the toxin. These versions have different plasmid stabilization efficiencies and it has been suggested that the differences could be related to the different TA interfaces and could also reflect a host-adaptative feature (Sletvold et al. 2008; Rosvoll et al.

2010). It has also been pointed that in some systems the affinity of the epsilon antitoxin for the zeta toxin is exceptionally high (within the fentomolar range) and that activation of the toxin in such conditions would require a specific intervention (Mutschler et al. 2010).

16.6.2 Regulation of ω-ε-ξ

The structure of the dimeric omega protein has been obtained and its DNA binding region, a RHH DNA binding motif, identified (Murayama et al. 2001). The structure of this global regulator gave important clues to understanding its ability to recognize heptamers arranged in direct and also inverse orientation that can be found in the promoter–operator regions regulated by this protein (de la Hoz et al. 2004). Omega protein, in addition to regulate transcription of the delta protein, interacts with the heptad repeats of *parS*, the centromer-like, anchors the delta motor protein to this site and plays an active role in the dynamic interactions that lead to the coordinate assembly and disassembly of plasmid pairs during the partitioning process (Soberon et al. 2011). The coordinated regulation of the replication, partitioning and TA systems mediated by the omega protein achieves almost complete stability of the plasmid minimizing its burden on the growth of the host (de la Hoz et al. 2000; Lioy et al. 2010). Interestingly, a shorter version of the *ω*-*ε*-*ξ* system containing the standard bicistronic and autoregulated arrangement typical of proteic TA operons was found in plasmids of the IncP1 family in *Neiseria gonorrhoeae* (Pachulec and van der Does 2010). In these systems, the antitoxin harbors a DNA binding domain, in addition to the toxin neutralizing one, which is involved in autoregulation; the toxin acts as a corregulator. The epsilon-zeta systems are widespread in all bacterial kingdoms both in plasmids and, preferentially, in the chromosomes (Leplae et al. 2011). The *ω*-*ε*-*ζ* version present in the chromosome of *Streptococcus pneumoniae* is also a bicistronic operon that has been identified as a virulence determinant (Brown et al. 2004).

16.6.3 Zeta Activity and Structure

The first important clue to understanding the mechanisms of action of the toxin and the antitoxin of this system was given by the determination and characterization of the crystal structure of the hetero-tetramer formed by dimers of the zeta toxin in complex with dimers of the epsilon antitoxin (Meinhart et al. 2003; Mutschler et al. 2010) (PDB code: 1GVN). Analysis of this complex revealed the presence of an ATP/GTP-binding site in the toxin and the potential of this protein to act as a kinase, phosphorylating, and inactivating a target important for cell growth or viability. Mutations of residues of the putative ATP/GTP-binding site inactivated the toxin. This nucleotide-binding site was protected in this structure

by the α-helical N-terminal region of the epsilon antitoxin, thus explaining the neutralization potential of this protein. Following this first important hint, the target of the toxin was identified using an attenuated version of the toxin (Mutschler et al. 2011). The zeta toxin phosphorylates UNAG, at its 3′-OH group, the substrate used by the MurA protein for the synthesis of peptidoglycan. The phosphorylation inhibits competitively the activity of the MurA protein, thus corrupting bacterial cell wall synthesis and eventually promoting cell lysis. This lysis explains the activity of the pathogen as the consequence of the release of virulence effectors of the bacterial in the target cells (reviewed and discussed by Mutschler and Meinhart (2011)). This activity also opens a way of controlling bacterial virulence using inhibitors of the Zeta toxin.

16.7 *axe-txe* Locus

axe-txe was originally found in pRUM plasmid of *Enterococcus faecium* (Grady and Hayes 2003), a multidrug resistance plasmid that coexists with a conjugative vancomicyn resistant plasmid (Rice et al. 1998). pRUM confers resistance to chloramphenicol, erythromycin, streptomycin, and spectinomycin.

The *axe-txe* operon encodes two small proteins, the antitoxin Axe (*a*ntito*x*in from *e*nterococcus) and the toxin Txe (*to*xin from *e*nterococcus). This system seems to be broadly disseminated among clinical isolates of enterococci that are resistant to vancomycin (VRE); *axe-txe* genes were found in 56/75 of the VRE isolates (Moritz and Hergenrother 2007). A recent study has found that clinical isolates of *E. faecium* containing pRUM plasmid were also positive for *axe-txe* (Rosvoll et al. 2010). There is also a correlation and physical linkage between *axe-txe* and *vanA* operon (Moritz and Hergenrother 2007; Rosvoll et al. 2010). It is proposed the system contributes to the stability of plasmid-encoded antibiotic resistance genes in enterococci.

pRUM-like plasmids were shown to be widely distributed in CC17-related strains of *E. faecium* (Rosvoll et al. 2010). In these strains, a Tn5382 transposon containing a *vanB2* subtype conferring resistance to vancomycin was found integrated into a pRUM-like plasmid containing the *axe-txe* system (Bjorkeng et al. 2011). It is argued that the linkage between *vanB2*-Tn5382 and the pRUM-like *repA* and *axe-txe* system had facilitated the dissemination and persistence of VRE.

The system contributes to plasmid stability not only in *E. faecium* but also in evolutionary diverged bacteria as *Bacillus thuringiensis* and *E. coli* (Grady and Hayes 2003). Axe is homologous to YefM and Phd of *yefM-yoeB* and *phd-doc* of *E. coli* (Grady and Hayes 2003). The toxin Txe is related to the YoeB branch of the RelE superfamily. Overproduction experiments suggested that Txe stalls growth while YoeB affected cell viability (Grady and Hayes 2003).

Recently, *axe-txe* of pS177 from *E. faecium* has been characterized. This locus is identical to the one encoded by pRUM. Similar to other members of the RelE superfamily, Txe inhibits translation by cleaving of mRNA (Halvorsen et al.

2011). mRNA cleavage occurs at the first base downstream of the translational start codon AUG, as described for the YoeB of *E. coli* or *Staphylococcus aureus* (Zhang and Inouye 2009; Yoshizumi et al. 2009).

Due to its broad distribution in clinical isolates of enterococci resistant to vancomycin, *axe-txe* has been proposed as a possible antibacterial target to impact enterococci infections (Halvorsen et al. 2011).

16.8 Final Remarks

Work on type I and type II loci from plasmids contributed to establish the TA field at an early stage. In this chapter, we have described orginal type II TA loci from plasmids. With some variation, a common theme has emerged. In most cases, TA operons consist of an upstream antitoxin gene and a downstream toxin gene, although reverse gene orders are not as uncommon as originally thought. In almost all cases, the operons are autoregulated by the TA complex with the antitoxin being the DNA binding factor. A common but singular feature of this autorregulation is the response of transcription regulation to the relative levels of the toxins and antitoxins. In several systems analyzed, TA repression is efficient in excess of the antitoxins, the component that specifically binds to operator DNA. Surprisingly, however, in excess of the toxin, the complex does not bind operator DNA and transcription becomes derepressed. This regulatory pattern was termed conditional cooperativity. Conditional cooperativity may function to correct fluctuations in the levels of the antitoxin, and thereby rescue the cells from the effects of a detrimental excess of toxin. Regulation of TA systems is also modulated at the post-transcriptional level and it is interlinked with the mechanisms that regulate the action of specific ATP-dependent cell proteases that act on the antitoxins.

TA systems can be considered as programs that regulate cell growth or viability and that can be activated selectively due to the differential decay of their antitoxins. In TA systems carried by plasmids, the selective activation of toxins in cells that lose the plasmid at cell division contributes to the effective maintenance of these genetic elements in bacterial populations. This is the so-called PSK mechanism that can also explains the advantage of plasmids containing TA systems when they are competing with plasmids devoid of these addition modules (Cooper and Heinemann 2000).

Members of the early TA list show variations on the common patterns described above. Thus, the *higBA* system has the antitoxin–toxin order inverted and two systems, *pasABC* and ω-ε-ζ, are tricistronic operons. In the *pasABC* system, the *pasC* component enhances the neutralization of the PasB toxin by the PasA antitoxin. In the three-component ω-ε-ζ system, a global regulator, the omega protein, but not the toxin and antitoxin components, represses the operon and coordinates this maintenance system with the replication and partitioning systems of the plasmid thus achieving almost complete plasmid stability. In plasmid R1 coupling between the replication and *kis-kid* maintenance modules has been observed. This coupling can

contribute, jointly with the copy number control protein CopB, to rescue replication of the plasmid in situations in which its copy number is anomalously low.

Although toxins achieve the same general objective, that is, to interfere with the proliferation of the cell, they do it targeting different essential processes. Translation is a frequent target of toxins (Kid, HigA, Txe, and Doc) and different toxins can act at different stages to inhibit protein synthesis. The repertoire of the processes inhibited extends also to the DNA gyrase (CcdB and ParE) and to the cell wall synthesis (zeta toxin). Interestingly, toxins targeting RNA or DNA gyrase can share a common structural module (Kid and CcdB or ParE and RelE-type toxin). Specific effects linked to the particular mode of action of the toxins, such as the SOS induction by the CcdB toxin or the inhibition of ColE1 and bacteriophage λ propagation by the Kid toxin, can be expected.

Plasmidic TA systems influence the maintenance of these extrachromosomal elements, and therefore favor the prevalence of the genetic traits carried by plasmids and their possible contribution to the adaptation and evolution of bacterial populations. While plasmid maintenance seems a measurable signature of TA systems carried by these extrachromosomal elements, the activity of their proteins, notably their toxins, can influence and even activate other cellular functions. This challenges the view that plasmidic TA systems play only a role in plasmid maintenance. Indeed, as proteolysis of antitoxins can trigger the toxins of homologous plasmidic and chromosomic TA systems and these toxins can inhibit cell growth or viability by similar mechanisms, a functional overlap (for instance, in stress response or in bacterial persistence), can be expected.

Acknowledgments Work in the laboratory was supported by current projects, BFU 2008-01566, CDS2008-0013 (INTERMODS), BFU2011-25939 and by the networks BFU 2008-0079-E/BNC and BFU2011-14145-E. We acknowledge frequent discussions with Juan López-Villarejo and Damian Lobato-Márquez, members of our laboratory. Discussions with Manuel Espinosa and with other members or advisors of the INTERMODS project are also acknowledged.

References

Afif, H., Allali, N., Couturier, M., & Van Melderen, L. (2001). The ratio between CcdA and CcdB modulates the transcriptional repression of the ccd poison-antidote system. *Molecular Microbiology, 41*, 73–82.

Bahassi, E. M., Salmon, M. A., Van Melderen, L., Bernard, P., & Couturier, M. (1995). F plasmid CcdB killer protein: ccdB gene mutants coding for non-cytotoxic proteins which retain their regulatory functions. *Molecular Microbiology, 15*, 1031–1037.

Bahassi, E. M., O'Dea, M. H., Allali, N., Messens, J., Gellert, M., & Couturier, M. (1999). Interactions of CcdB with DNA gyrase. Inactivation of Gyra, poisoning of the gyrase-DNA complex, and the antidote action of CcdA. *The Journal of Biological Chemistry, 274*, 10936–10944.

Barbosa, L. C., Garrido, S. S., Garcia, A., Delfino, D. B., & Marchetto, R. (2010). Function inferences from a molecular structural model of bacterial ParE toxin. *Bioinformation, 4*, 438–440.

Bernard, P., & Couturier, M. (1991). The 41 carboxy-terminal residues of the miniF plasmid CcdA protein are sufficient to antagonize the killer activity of the CcdB protein. *Molecular and General Genetics, 226*, 297–304.

Bernard, P., & Couturier, M. (1992). Cell killing by the F plasmid CcdB protein involves poisoning of DNA-topoisomerase II complexes. *Journal of Molecular Biology, 226*, 735–745.

Bernard, P., Kezdy, K. E., Van Melderen, L., Steyaert, J., Wyns, L., Pato, M. L., et al. (1993). The F plasmid CcdB protein induces efficient ATP-dependent DNA cleavage by gyrase. *Journal of Molecular Biology, 234*, 534–541.

Bjorkeng, E., Rasmussen, G., Sundsfjord, A., Sjoberg, L., Hegstad, K., & Soderquist, B. (2011). Clustering of polyclonal VanB-type vancomycin-resistant *Enterococcus faecium* in a low-endemic area was associated with CC17-genogroup strains harbouring transferable vanB2-Tn5382 and pRUM-like repA containing plasmids with axe-txe plasmid addiction systems. *Apmis, 119*, 247–258.

Blower, T. R., Salmond, G. P., & Luisi, B. F. (2011). Balancing at survival's edge: The structure and adaptive benefits of prokaryotic toxin–antitoxin partners. *Current Opinion in Structural Biology, 21*, 109–118.

Bordes, P., Cirinesi, A. M., Ummels, R., Sala, A., Sakr, S., Bitter, W., et al. (2011). SecB-like chaperone controls a toxin–antitoxin stress-responsive system in *Mycobacterium tuberculosis*. *Proceedings of the National Academy of Sciences of the United States of America, 108*, 8438–8443.

Bravo, A., de Torrontegui, G., & Diaz, R. (1987). Identification of components of a new stability system of plasmid R1, ParD, that is close to the origin of replication of this plasmid. *Molecular and General Genetics, 210*, 101–110.

Bravo, A., Ortega, S., de Torrontegui, G., & Diaz, R. (1988). Killing of *Escherichia coli* cells modulated by components of the stability system ParD of plasmid R1. *Molecular and General Genetics, 215*, 146–151.

Breg, J. N., van Opheusden, J. H., Burgering, M. J., Boelens, R., & Kaptein, R. (1990). Structure of Arc repressor in solution: Evidence for a family of beta-sheet DNA-binding proteins. *Nature, 346*, 586–589.

Brown, J. S., Gilliland, S. M., Spratt, B. G., & Holden, D. W. (2004). A locus contained within a variable region of pneumococcal pathogenicity island 1 contributes to virulence in mice. *Infection and Immunity, 72*, 1587–1593.

Budde, P. P., Davis, B. M., Yuan, J., & Waldor, M. K. (2007). Characterization of a higBA toxin–antitoxin locus in *Vibrio cholerae*. *Journal of Bacteriology, 189*, 491–500.

Buts, L., Lah, J., Dao-Thi, M. H., Wyns, L., & Loris, R. (2005). Toxin–antitoxin modules as bacterial metabolic stress managers. *Trends in Biochemical Sciences, 30*, 672–679.

Camacho, A. G., Misselwitz, R., Behlke, J., Ayora, S., Welfle, K., Meinhart, A., et al. (2002). In vitro and in vivo stability of the epsilon2zeta2 protein complex of the broad host-range Streptococcus pyogenes pSM19035 addiction system. *Biological Chemistry, 383*, 1701–1713.

Ceglowski, P., Boitsov, A., Chai, S., & Alonso, J. C. (1993a). Analysis of the stabilization system of pSM19035-derived plasmid pBT233 in *Bacillus subtilis*. *Gene, 136*, 1–12.

Ceglowski, P., Boitsov, A., Karamyan, N., Chai, S., & Alonso, J. C. (1993b). Characterization of the effectors required for stable inheritance of *Streptococcus pyogenes* pSM19035-derived plasmids in *Bacillus subtilis*. *Molecular and General Genetics, 241*, 579–585.

Christensen-Dalsgaard, M., & Gerdes, K. (2006). Two higBA loci in the *Vibrio cholerae* superintegron encode mRNA cleaving enzymes and can stabilize plasmids. *Molecular Microbiology, 62*, 397–411.

Cooper, T. F., & Heinemann, J. A. (2000). Postsegregational killing does not increase plasmid stability but acts to mediate the exclusion of competing plasmids. *Proceedings of the National Academy of Sciences of the United States of America, 97*, 12643–12648.

Couturier, M., Bahassi M. el & Van Melderen, L. (1998) Bacterial death by DNA gyrase poisoning. *Trends in Microbiology 6*, 269–275.

Critchlow, S. E., O'Dea, M. H., Howells, A. J., Couturier, M., Gellert, M., & Maxwell, A. (1997). The interaction of the F plasmid killer protein, CcdB, with DNA gyrase: Induction of DNA cleavage and blocking of transcription. *Journal of Molecular Biology, 273*, 826–839.

Dalton, K. M., & Crosson, S. (2010). A conserved mode of protein recognition and binding in a ParD–ParE toxin–antitoxin complex. *Biochemistry, 49*, 2205–2216.

Dao-Thi, M. H., Charlier, D., Loris, R., Maes, D., Messens, J., Wyns, L., et al. (2002). Intricate interactions within the ccd plasmid addiction system. *The Journal of Biological Chemistry, 277*, 3733–3742.

Dao-Thi, M. H., Van Melderen, L., De Genst, E., Afif, H., Buts, L., Wyns, L., et al. (2005). Molecular basis of gyrase poisoning by the addiction toxin CcdB. *Journal of Molecular Biology, 348*, 1091–1102.

Davis, T. L., Helinski, D. R., & Roberts, R. C. (1992). Transcription and autoregulation of the stabilizing functions of broad-host-range plasmid RK2 in *Escherichia coli*, *Agrobacterium tumefaciens* and *Pseudomonas aeruginosa*. *Molecular Microbiology, 6*, 1981–1994.

De Jonge, N., Garcia-Pino, A., Buts, L., Haesaerts, S., Charlier, D., Zangger, K., et al. (2009). Rejuvenation of CcdB-poisoned gyrase by an intrinsically disordered protein domain. *Molecular Cell, 35*, 154–163.

de la Hoz, A. B., Ayora, S., Sitkiewicz, I., Fernandez, S., Pankiewicz, R., Alonso, J. C., et al. (2000). Plasmid copy-number control and better-than-random segregation genes of pSM19035 share a common regulator. *Proceedings of the National Academy of Sciences of the United States of America, 97*, 728–733.

de la Hoz, A. B., Pratto, F., Misselwitz, R., Speck, C., Weihofen, W., Welfle, K., et al. (2004). Recognition of DNA by omega protein from the broad-host range *Streptococcus pyogenes* plasmid pSM19035: analysis of binding to operator DNA with one to four heptad repeats. *Nucleic Acids Research, 32*, 3136–3147.

De Lano, W. L. (2002). *The Pymol Molecular Graphics System*. San Carlos: De Lano Scientific.

del Solar, G., Hernandez-Arriaga, A. M., Gomis-Ruth, F. X., Coll, M., & Espinosa, M. (2002). A genetically economical family of plasmid-encoded transcriptional repressors involved in control of plasmid copy number. *Journal of Bacteriology, 184*, 4943–4951.

Diago-Navarro, E., Kamphuis, M. B., Boelens, R., Barendregt, A., Heck, A. J., van den Heuvel, R. H., et al. (2009). A mutagenic analysis of the RNase mechanism of the bacterial Kid toxin by mass spectrometry. *The FEBS Journal, 276*, 4973–4986.

Diago-Navarro, E., Hernandez-Arriaga, A. M., Lopez-Villarejo, J., Munoz-Gomez, A. J., Kamphuis, M. B., Boelens, R., et al. (2010). parD toxin–antitoxin system of plasmid R1—basic contributions, biotechnological applications and relationships with closely-related toxin–antitoxin systems. *The FEBS Journal, 277*, 3097–3117.

Dixon, J. M., & Lipinski, A. E. (1972). Resistance of group A beta-hemolytic streptococci to lincomycin and erythromycin. *Antimicrobial Agents and Chemotherapy, 1*, 333–339.

Eberl, L., Givskov, M., & Schwab, H. (1992). The divergent promoters mediating transcription of the par locus of plasmid RP4 are subject to autoregulation. *Molecular Microbiology, 6*, 1969–1979.

Fiebig, A., Castro Rojas, C. M., Siegal-Gaskins, D., & Crosson, S. (2010). Interaction specificity, toxicity and regulation of a paralogous set of ParE/RelE-family toxin–antitoxin systems. *Molecular Microbiology, 77*, 236–251.

Fineran, P. C., Blower, T. R., Foulds, I. J., Humphreys, D. P., Lilley, K. S., & Salmond, G. P. (2009). The phage abortive infection system, ToxIN, functions as a protein-RNA toxin–antitoxin pair. *Proceedings of the National Academy of Sciences of the United States of America, 106*, 894–899.

Francuski, D., & Saenger, W. (2009). Crystal structure of the antitoxin–toxin protein complex RelB–RelE from *Methanococcus jannaschii*. *Journal of Molecular Biology, 393*, 898–908.

Garcia-Pino, A., Balasubramanian, S., Wyns, L., Gazit, E., De Greve, H., Magnuson, R. D., et al. (2010). Allostery and intrinsic disorder mediate transcription regulation by conditional cooperativity. *Cell, 142*, 101–111.

Gerdes, K. (2000). Toxin–antitoxin modules may regulate synthesis of macromolecules during nutritional stress. *Journal of Bacteriology, 182*, 561–572.

Gerdes, K., Rasmussen, P. B., & Molin, S. (1986). Unique type of plasmid maintenance function: Postsegregational killing of plasmid-free cells. *Proceedings of the National Academy of Sciences of the United States of America, 83*, 3116–3120.

Grady, R., & Hayes, F. (2003). Axe-Txe, a broad-spectrum proteic toxin–antitoxin system specified by a multidrug-resistant, clinical isolate of *Enterococcus faecium*. *Molecular Microbiology, 47*, 1419–1432.

Gronlund, H., & Gerdes, K. (1999). Toxin–antitoxin systems homologous with relBE of Escherichia coli plasmid P307 are ubiquitous in prokaryotes. *Journal of Molecular Biology, 285*, 1401–1416.

Halvorsen, E.M., Williams, J.J., Bhimani, A.J., Billings, E.A., & Hergenrother, P.J. (2011) Txe, an endoribonuclease of the enterococcal Axe-Txe toxin–antitoxin system, cleaves mRNA and inhibits protein synthesis. *Microbiology* (*Reading*, *England*) *157*, 387–397.

Hargreaves, D., Santos-Sierra, S., Giraldo, R., Sabariegos-Jareno, R., de la Cueva-Mendez, G., Boelens, R., et al. (2002). Structural and functional analysis of the kid toxin protein from *E. coli* plasmid R1. *Structure, 10*, 1425–1433.

Hayes, F. (1998). A family of stability determinants in pathogenic bacteria. *Journal of Bacteriology, 180*, 6415–6418.

Hurley, J. M., & Woychik, N. A. (2009). Bacterial toxin HigB associates with ribosomes and mediates translation-dependent mRNA cleavage at A-rich sites. *The Journal of Biological Chemistry, 284*, 18605–18613.

Jaffe, A., Ogura, T., & Hiraga, S. (1985). Effects of the ccd function of the F plasmid on bacterial growth. *Journal of Bacteriology, 163*, 841–849.

Jensen, R. B., Grohmann, E., Schwab, H., Diaz-Orejas, R., & Gerdes, K. (1995). Comparison of ccd of F, parDE of RP4, and parD of R1 using a novel conditional replication control system of plasmid R1. *Molecular Microbiology, 17*, 211–220.

Jiang, Y., Pogliano, J., Helinski, D. R., & Konieczny, I. (2002). ParE toxin encoded by the broad-host-range plasmid RK2 is an inhibitor of *Escherichia coli* gyrase. *Molecular Microbiology, 44*, 971–979.

Johnson, E. P., Strom, A. R., & Helinski, D. R. (1996). Plasmid RK2 toxin protein ParE: Purification and interaction with the ParD antitoxin protein. *Journal of Bacteriology, 178*, 1420–1429.

Kamada, K., & Hanaoka, F. (2005). Conformational change in the catalytic site of the ribonuclease YoeB toxin by YefM antitoxin. *Molecular Cell, 19*, 497–509.

Kamada, K., Hanaoka, F., & Burley, S. K. (2003). Crystal structure of the MazE/MazF complex: Molecular bases of antidote-toxin recognition. *Molecular Cell, 11*, 875–884.

Kamphuis, M. B., Bonvin, A. M., Monti, M. C., Lemonnier, M., Munoz-Gomez, A., van den Heuvel, R. H., et al. (2006). Model for RNA binding and the catalytic site of the RNase Kid of the bacterial parD toxin–antitoxin system. *Journal of Molecular Biology, 357*, 115–126.

Kamphuis, M. B., Monti, M. C., van den Heuvel, R. H., Santos-Sierra, S., Folkers, G. E., Lemonnier, M., et al. (2007). Interactions between the toxin Kid of the bacterial parD system and the antitoxins Kis and MazE. *Proteins, 67*, 219–231.

Klare, I., Heier, H., Claus, H., Bohme, G., Marin, S., Seltmann, G., Hakenbeck, R., Antanassova, V., & Witte, W. (1995) *Enterococcus faecium* strains with vanA-mediated high-level glycopeptide resistance isolated from animal foodstuffs and fecal samples of humans in the community. *Microbial Drug Resistance* (*Larchmont*, *N.Y*) *1*, 265–272.

Kuhn, I., Iversen, A., Finn, M., Greko, C., Burman, L. G., Blanch, A. R., et al. (2005). Occurrence and relatedness of vancomycin-resistant enterococci in animals, humans, and the environment in different European regions. *Applied and Environmental Microbiology, 71*, 5383–5390.

Lehnherr, H., Maguin, E., Jafri, S., & Yarmolinsky, M. B. (1993). Plasmid addiction genes of bacteriophage P1: doc, which causes cell death on curing of prophage, and phd, which prevents host death when prophage is retained. *Journal of Molecular Biology, 233*, 414–428.

Leplae, R., Geeraerts, D., Hallez, R., Guglielmini, J., Dreze, P., & Van Melderen, L. (2011). Diversity of bacterial type II toxin–antitoxin systems: A comprehensive search and functional analysis of novel families. *Nucleic Acids Research, 39*, 5513–5525.

Li, G. Y., Zhang, Y., Inouye, M., & Ikura, M. (2008). Structural mechanism of transcriptional autorepression of the *Escherichia coli* RelB/RelE antitoxin/toxin module. *Journal of Molecular Biology, 380*, 107–119.

Lioy, V.S., Martin, M.T., Camacho, A.G., Lurz, R., Antelmann, H., Hecker, M., Hitchin, E., Ridge, Y., Wells, J.M., & Alonso, J.C. (2006) pSM19035-encoded zeta toxin induces stasis followed by death in a subpopulation of cells. *Microbiology* (*Reading*, *England*) *152*, 2365–2379.

Lioy, V. S., Pratto, F., de la Hoz, A. B., Ayora, S., & Alonso, J. C. (2010). Plasmid pSM19035, a model to study stable maintenance in Firmicutes. *Plasmid, 64*, 1–17.

Lopez-Villarejo, J., Diago-Navarro, E., Hernandez-Arriaga, A.M., & Diaz-Orejas, R. (2012) Kis antitoxin couples plasmid R1 replication and parD (kis,kid) maintenance modules. *Plasmid, 67*, 118–127.

Loris, R., Dao-Thi, M. H., Bahassi, E. M., Van Melderen, L., Poortmans, F., Liddington, R., et al. (1999). Crystal structure of CcdB, a topoisomerase poison from *E. coli*. *Journal of Molecular Biology, 285*, 1667–1677.

Madl, T., Van Melderen, L., Mine, N., Respondek, M., Oberer, M., Keller, W., et al. (2006). Structural basis for nucleic acid and toxin recognition of the bacterial antitoxin CcdA. *Journal of Molecular Biology, 364*, 170–185.

Maki, S., Takiguchi, S., Horiuchi, T., Sekimizu, K., & Miki, T. (1996). Partner switching mechanisms in inactivation and rejuvenation of *Escherichia coli* DNA gyrase by F plasmid proteins LetD (CcdB) and LetA (CcdA). *Journal of Molecular Biology, 256*, 473–482.

Maxwell, A. (1997). DNA gyrase as a drug target. *Trends in Microbiology, 5*, 102–109.

Meinhart, A., Alonso, J. C., Strater, N., & Saenger, W. (2003). Crystal structure of the plasmid maintenance system epsilon/zeta: functional mechanism of toxin zeta and inactivation by epsilon 2 zeta 2 complex formation. *Proceedings of the National Academy of Sciences of the United States of America, 100*, 1661–1666.

Miki, T., Park, J. A., Nagao, K., Murayama, N., & Horiuchi, T. (1992). Control of segregation of chromosomal DNA by sex factor F in *Escherichia coli*. Mutants of DNA gyrase subunit A suppress letD (ccdB) product growth inhibition. *Journal of Molecular Biology, 225*, 39–52.

Monti, M. C., Hernandez-Arriaga, A. M., Kamphuis, M. B., Lopez-Villarejo, J., Heck, A. J., Boelens, R., et al. (2007). Interactions of Kid-Kis toxin–antitoxin complexes with the parD operator–promoter region of plasmid R1 are piloted by the Kis antitoxin and tuned by the stoichiometry of Kid-Kis oligomers. *Nucleic Acids Research, 35*, 1737–1749.

Moritz, E. M., & Hergenrother, P. J. (2007). Toxin–antitoxin systems are ubiquitous and plasmid-encoded in vancomycin-resistant enterococci. *Proceedings of the National Academy of Sciences of the United States of America, 104*, 311–316.

Muñoz-Gomez, A. J., Lemonnier, M., Santos-Sierra, S., Berzal-Herranz, A., & Diaz-Orejas, R. (2005). RNase/anti-RNase activities of the bacterial parD toxin–antitoxin system. *Journal of Bacteriology, 187*, 3151–3157.

Murayama, K., Orth, P., de la Hoz, A. B., Alonso, J. C., & Saenger, W. (2001). Crystal structure of omega transcriptional repressor encoded by *Streptococcus pyogenes* plasmid pSM19035 at 1.5 Å resolution. *Journal of Molecular Biology, 314*, 789–796.

Mutschler, H., & Meinhart, A. (2011) Epsilon/zeta systems: Their role in resistance, virulence, and their potential for antibiotic development. *Journal of Molecular Medicine* (*Berlin*, *Germany*) *89,* 1183–1194.

Mutschler, H., Reinstein, J., & Meinhart, A. (2010). Assembly dynamics and stability of the pneumococcal epsilon zeta antitoxin toxin (PezAT) system from *Streptococcus pneumoniae*. *The Journal of Biological Chemistry, 285*, 21797–21806.

Mutschler, H., Gebhardt, M., Shoeman, R. L., & Meinhart, A. (2011). A novel mechanism of programmed cell death in bacteria by toxin–antitoxin systems corrupts peptidoglycan synthesis. *PLoS Biology, 9*, e1001033.

Nordström, K. (2006). Plasmid R1—replication and its control. *Plasmid, 55*, 1–26.

Oberer, M., Lindner, H., Glatter, O., Kratky, C., & Keller, W. (1999). Thermodynamic properties and DNA binding of the ParD protein from the broad host-range plasmid RK2/RP4 killing system. *Biological Chemistry, 380*, 1413–1420.

Oberer, M., Zangger, K., Prytulla, S., & Keller, W. (2002). The anti-toxin ParD of plasmid RK2 consists of two structurally distinct moieties and belongs to the ribbon-helix-helix family of DNA-binding proteins. *The Biochemical Journal, 361*, 41–47.

Oberer, M., Zangger, K., Gruber, K., & Keller, W. (2007). The solution structure of ParD, the antidote of the ParDE toxin antitoxin module, provides the structural basis for DNA and toxin binding. *Protein Science, 16*, 1676–1688.

Ogura, T., & Hiraga, S. (1983). Mini-F plasmid genes that couple host cell division to plasmid proliferation. *Proceedings of the National Academy of Sciences of the United States of America, 80*, 4784–4788.

Ortega, S., de Torrontegui, G., & Diaz, R. (1989). Isolation and characterization of a conditional replication mutant of the antibiotic resistance factor R1 affected in the gene of the replication protein repA. *Molecular and General Genetics, 217*, 111–117.

Pachulec, E., & van der Does, C. (2010). Conjugative plasmids of *Neisseria gonorrhoeae*. *PLoS ONE, 5*, e9962.

Pandey, D. P., & Gerdes, K. (2005). Toxin–antitoxin loci are highly abundant in free-living but lost from host-associated prokaryotes. *Nucleic Acids Research, 33*, 966–976.

Pedersen, K., Zavialov, A. V., Pavlov, M. Y., Elf, J., Gerdes, K., & Ehrenberg, M. (2003). The bacterial toxin RelE displays codon-specific cleavage of mRNAs in the ribosomal A site. *Cell, 112*, 131–140.

Pimentel, B., Madine, M. A., & de la Cueva-Mendez, G. (2005). Kid cleaves specific mRNAs at UUACU sites to rescue the copy number of plasmid R1. *The EMBO Journal, 24*, 3459–3469.

Potrykus, K., Santos, S., Lemonnier, M., Diaz-Orejas, R., & Wegrzyn, G. (2002) Differential effects of Kid toxin on two modes of replication of lambdoid plasmids suggest that this toxin acts before, but not after, the assembly of the replication complex. *Microbiology* (*Reading, England*) *148*, 2489–2495.

Radnedge, L., Davis, M. A., Youngren, B., & Austin, S. J. (1997). Plasmid maintenance functions of the large virulence plasmid of *Shigella flexneri*. *Journal of Bacteriology, 179*, 3670–3675.

Rafferty, J. B., Somers, W. S., Saint-Girons, I., & Phillips, S. E. (1989). Three-dimensional crystal structures of *Escherichia coli* met repressor with and without corepressor. *Nature, 341*, 705–710.

Raumann, B. E., Rould, M. A., Pabo, C. O., & Sauer, R. T. (1994). DNA recognition by beta-sheets in the Arc repressor–operator crystal structure. *Nature, 367*, 754–757.

Rice, L. B., Carias, L. L., Donskey, C. L., & Rudin, S. D. (1998). Transferable, plasmid-mediated vanB-type glycopeptide resistance in *Enterococcus faecium*. *Antimicrobial Agents and Chemotherapy, 42*, 963–964.

Riise, E., & Molin, S. (1986). Purification and characterization of the CopB replication control protein, and precise mapping of its target site in the R1 plasmid. *Plasmid, 15*, 163–171.

Roberts, R. C., & Helinski, D. R. (1992). Definition of a minimal plasmid stabilization system from the broad-host-range plasmid RK2. *Journal of Bacteriology, 174*, 8119–8132.

Roberts, R. C., Burioni, R., & Helinski, D. R. (1990). Genetic characterization of the stabilizing functions of a region of broad-host-range plasmid RK2. *Journal of Bacteriology, 172*, 6204–6216.

Roberts, R. C., Spangler, C., & Helinski, D. R. (1993). Characteristics and significance of DNA binding activity of plasmid stabilization protein ParD from the broad host-range plasmid RK2. *The Journal of Biological Chemistry, 268*, 27109–27117.

Roberts, R. C., Strom, A. R., & Helinski, D. R. (1994). The parDE operon of the broad-host-range plasmid RK2 specifies growth inhibition associated with plasmid loss. *Journal of Molecular Biology, 237*, 35–51.

Rosvoll, T. C., Pedersen, T., Sletvold, H., Johnsen, P. J., Sollid, J. E., Simonsen, G. S., et al. (2010). PCR-based plasmid typing in *Enterococcus faecium* strains reveals widely distributed pRE25-, pRUM-, pIP501- and pHTbeta-related replicons associated with glycopeptide resistance and stabilizing toxin–antitoxin systems. *FEMS Immunology and Medical Microbiology, 58*, 254–268.

Ruiz-Echevarria, M. J., Berzal-Herranz, A., Gerdes, K., & Diaz-Orejas, R. (1991a). The kis and kid genes of the parD maintenance system of plasmid R1 form an operon that is autoregulated at the level of transcription by the co-ordinated action of the Kis and Kid proteins. *Molecular Microbiology, 5*, 2685–2693.

Ruiz-Echevarria, M. J., de Torrontegui, G., Gimenez-Gallego, G., & Diaz-Orejas, R. (1991b). Structural and functional comparison between the stability systems ParD of plasmid R1 and Ccd of plasmid F. *Molecular and General Genetics, 225*, 355–362.

Ruiz-Echevarria, M. J., de la Cueva, G., & Diaz-Orejas, R. (1995a). Translational coupling and limited degradation of a polycistronic messenger modulate differential gene expression in the parD stability system of plasmid R1. *Molecular and General Genetics, 248*, 599–609.

Ruiz-Echevarria, M. J., de la Torre, M. A., & Diaz-Orejas, R. (1995b). A mutation that decreases the efficiency of plasmid R1 replication leads to the activation of parD, a killer stability system of the plasmid. *FEMS Microbiology Letters, 130*, 129–135.

Salmon, M. A., Van Melderen, L., Bernard, P., & Couturier, M. (1994). The antidote and autoregulatory functions of the F plasmid CcdA protein: A genetic and biochemical survey. *Molecular and General Genetics, 244*, 530–538.

Santos-Sierra, S., Giraldo, R., & Diaz-Orejas, R. (1997). Functional interactions between homologous conditional killer systems of plasmid and chromosomal origin. *FEMS Microbiology Letters, 152*, 51–56.

Santos-Sierra, S., Pardo-Abarrio, C., Giraldo, R., & Diaz-Orejas, R. (2002). Genetic identification of two functional regions in the antitoxin of the parD killer system of plasmid R1. *FEMS Microbiology Letters, 206*, 115–119.

Simic, M., De Jonge, N., Loris, R., Vesnaver, G., & Lah, J. (2009). Driving forces of gyrase recognition by the addiction toxin CcdB. *The Journal of Biological Chemistry, 284*, 20002–20010.

Sletvold, H., Johnsen, P. J., Hamre, I., Simonsen, G. S., Sundsfjord, A., & Nielsen, K. M. (2008). Complete sequence of *Enterococcus faecium* pVEF3 and the detection of an omega-epsilon-zeta toxin–antitoxin module and an ABC transporter. *Plasmid, 60*, 75–85.

Smith, A. B., & Maxwell, A. (2006). A strand-passage conformation of DNA gyrase is required to allow the bacterial toxin, CcdB, to access its binding site. *Nucleic Acids Research, 34*, 4667–4676.

Smith, A. S., & Rawlings, D. E. (1997). The poison-antidote stability system of the broad-host-range *Thiobacillus ferrooxidans* plasmid pTF-FC2. *Molecular Microbiology, 26*, 961–970.

Soberon, N. E., Lioy, V. S., Pratto, F., Volante, A., & Alonso, J. C. (2011). Molecular anatomy of the *Streptococcus pyogenes* pSM19035 partition and segrosome complexes. *Nucleic Acids Research, 39*, 2624–2637.

Somers, W. S., Rafferty, J. B., Phillips, K., Strathdee, S., He, Y. Y., McNally, T., et al. (1994). The Met repressor–operator complex: DNA recognition by beta-strands. *Annals of the New York Academy of Sciences, 726*, 105–117.

Thomas, C. M., & Helinski, D. R. (1989). Vegetative replication and stable inheritance of IncP plasmids. In C. M. Thomas (Ed.), *Promiscuous Plasmids of Gram-Negative Bacteria* (pp. 1–25). San Diego: Academic Press.

Tian, Q. B., Hayashi, T., Murata, T., & Terawaki, Y. (1996a). Gene product identification and promoter analysis of hig locus of plasmid Rts1. *Biochemical and Biophysical Research Communications, 225*, 679–684.

Tian, Q. B., Ohnishi, M., Tabuchi, A., & Terawaki, Y. (1996b). A new plasmid-encoded proteic killer gene system: Cloning, sequencing, and analyzing hig locus of plasmid Rts1. *Biochemical and Biophysical Research Communications, 220*, 280–284.

Tian, Q. B., Ohnishi, M., Murata, T., Nakayama, K., Terawaki, Y., & Hayashi, T. (2001). Specific protein-DNA and protein–protein interaction in the hig gene system, a plasmid-borne proteic killer gene system of plasmid Rts1. *Plasmid, 45*, 63–74.

Tsuchimoto, S., Ohtsubo, H., & Ohtsubo, E. (1988). Two genes, pemK and pemI, responsible for stable maintenance of resistance plasmid R100. *Journal of Bacteriology, 170*, 1461–1466.

Tsuchimoto, S., Nishimura, Y., & Ohtsubo, E. (1992). The stable maintenance system pem of plasmid R100: Degradation of PemI protein may allow PemK protein to inhibit cell growth. *Journal of Bacteriology, 174*, 4205–4211.

Van Melderen, L., Bernard, P., & Couturier, M. (1994). Lon-dependent proteolysis of CcdA is the key control for activation of CcdB in plasmid-free segregant bacteria. *Molecular Microbiology, 11*, 1151–1157.

Van Melderen, L., Thi, M. H., Lecchi, P., Gottesman, S., Couturier, M., & Maurizi, M. R. (1996). ATP-dependent degradation of CcdA by Lon protease. Effects of secondary structure and heterologous subunit interactions. *The Journal of Biological Chemistry, 271*, 27730–27738.

Yoshizumi, S., Zhang, Y., Yamaguchi, Y., Chen, L., Kreiswirth, B. N., & Inouye, M. (2009). Staphylococcus aureus YoeB homologues inhibit translation initiation. *Journal of Bacteriology, 191*, 5868–5872.

Yuan, J., Sterckx, Y., Mitchenall, L. A., Maxwell, A., Loris, R., & Waldor, M. K. (2010). Vibrio cholerae ParE2 poisons DNA gyrase via a mechanism distinct from other gyrase inhibitors. *The Journal of Biological Chemistry, 285*, 40397–40408.

Zhang, Y., & Inouye, M. (2009). The inhibitory mechanism of protein synthesis by YoeB, an *Escherichia coli* toxin. *The Journal of Biological Chemistry, 284*, 6627–6638.

Zhang, J., Zhang, Y., Zhu, L., Suzuki, M., & Inouye, M. (2004). Interference of mRNA function by sequence-specific endoribonuclease PemK. *The Journal of Biological Chemistry, 279*, 20678–20684.

Zielenkiewicz, U., & Ceglowski, P. (2005). The toxin–antitoxin system of the streptococcal plasmid pSM19035. *Journal of Bacteriology, 187*, 6094–6105.

Chapter 17
Toxin-Antitoxin Loci in *Mycobacterium tuberculosis*

Ambre Sala, Patricia Bordes, Gwennaele Fichant and Pierre Genevaux

Abstract Chromosomally encoded type II toxin–antitoxin (TA) systems generally consist of two adjacent genes in an operon encoding a stable toxin and a less stable, protease-sensitive cognate antitoxin. While the toxin and the antitoxin form a stable complex under normal growth conditions, the degradation of the antitoxin by stress-proteases under certain conditions leads to activation of the toxin and subsequent growth inhibition. Such stress-responsive TA systems have been associated with various cellular processes, including stabilization of genomic regions, protection against foreign DNA, biofilm formation, persistence, and control of the stress response. Yet, the contribution of chromosomal TA to bacterial virulence is presently unknown. Herein, we investigate the potential role of multiple chromosomally encoded TA systems in virulence, focusing on the tuberculosis agent *Mycobacterium tuberculosis*, which contains more than 70 TA loci in its genome. We describe what is currently known about the multiple TA families present in this bacterium, with emphasis on the recently discovered atypical stress-responsive toxin–antitoxin-chaperone (TAC) system, a TA system controlled by a SecB-like chaperone.

17.1 Introduction

Toxin–antitoxin (TA) systems are two partner modules that were first discovered as plasmid-borne loci involved in plasmid maintenance by a mechanism called "postsegregational killing". This phenomenon relies on a stable toxin that inhibits

A. Sala · P. Bordes · G. Fichant · P. Genevaux (✉)
Laboratoire de Microbiologie et Génétique Moléculaire (LMGM), Centre National de la Recherche Scientifique (CNRS) and Université Paul Sabatier, Toulouse, France
e-mail: pierre.genevaux@ibcg.biotoul.fr

K. Gerdes (ed.), *Prokaryotic Toxin-Antitoxins*,
DOI: 10.1007/978-3-642-33253-1_17,

cell growth and a labile protease-sensitive antitoxin that antagonizes the toxin function, both encoded by the plasmid. When the plasmid is lost, the antitoxin is degraded resulting in toxin-mediated cell growth inhibition and eventual cell death (Yamaguchi et al. 2011). Since then, many TA loci have been found in numerous bacterial and archeal chromosomes, suggesting alternative functions, such as stabilization of chromosomal regions or growth modulation in response to environmental insults (Gerdes et al. 2005).

Three types of TA systems are defined based on the nature of the antitoxin and on the mechanism of neutralization of the toxin, which is always a protein. In type I TA systems, the antitoxin is a small antisense RNA that forms a duplex with a specific region of the toxin mRNA resulting in the inhibition of toxin production. Type II antitoxins are proteins that inhibit toxin function by a protein–protein interaction. The type III TA system is represented by only one characterized system to date, ToxIN from *Petrobacterium atrosepticum*, whose RNA antitoxin ToxI inactivates the ToxN toxin by direct binding (Yamaguchi et al. 2011). This chapter will focus on chromosomally encoded type II TA systems, with emphasis on the multiple TA systems present in the major human pathogen *M. tuberculosis* and their potential implication in virulence.

17.2 General Features of Type II TA Systems

Type II TA systems are encoded in an operon consisting of two small genes whose products, a toxin and its cognate antitoxin, form a complex in which the toxin is inactive under normal growth conditions. Type II antitoxins generally carry a DNA binding domain which binds directly on the promoter region of the operon, and acts as a repressor, often together with the toxin. Under certain conditions, the antitoxin is degraded by stress-proteases, resulting in toxin activation, growth inhibition, and eventual cell death (Gerdes et al. 2005). Free toxins from type II systems have been shown to target essential cellular processes such as replication, cell wall synthesis, cell division, or translation (Gerdes et al. 2005; Yamaguchi et al. 2011, Mutschler et al. 2011).

Chromosomal type II systems are widely distributed in prokaryotic genomes and often clustered in genomic islands (Makarova et al. 2009). This is in agreement with the fact that TA modules are often found on mobile genetic elements such as plasmid, phages, or transposable elements, thus mostly belonging to the prokaryotic mobilome (Makarova et al. 2009). For this reason, the content of type II TA systems is highly variable between species and also from one strain to another (http://bioinfo-mml.sjtu.edu.cn/TADB/). A study of 126 sequenced prokaryotic genomes revealed that free-living organisms have many TA loci, whereas obligate intracellular organisms seem to have fewer (Pandey and Gerdes 2005). However, some *Rickettsia* species, which are strict intracellular bacteria, contain

multiple TA systems (Audoly et al. 2011). In addition, it has been proposed that human pathogens generally contain more TA systems than their non-pathogenic counterparts (Georgiades and Raoult 2011), but the differences are often subtle and multiple TA modules are found in non-pathogenic species as well (Makarova et al. 2009). Thus, to date there is no obvious correlation between TA abundance and organism lifestyle or virulence.

17.3 Role of Chromosomal Type II TA Systems and Possible Implication in Virulence

To date, the biological role of chromosomal TA systems is still unclear and under debate. Given their apparent dissemination by horizontal gene transfer, it has been proposed that TA systems could in some cases be junk DNA conserved by virtue of their intrinsic properties of addiction (Van Melderen 2010). Nevertheless, many chromosomal TA systems are induced under stress conditions and some seem to have important physiological functions, such as stabilization of genomic regions, anti-addiction against similar plasmid-borne toxins, protection against phage infection, biofilm formation, persistence, and control of the stress response (Van Melderen 2010; Yamaguchi and Inouye 2011). Therefore, the presence of multiple chromosomal TA systems in many bacteria suggests that these systems could function as intricate cellular networks of reactive elements that facilitate adaptive response to various environmental changes.

There is little or no evidence of a potential implication of TA systems in virulence. However, several important pathogens including *Yersinia pestis*, *Streptococcus pneumoniae*, *Staphylococcus aureus*, *Pseudomonas aeruginosa*, *Vibrio cholerae*, *Rickettsia felis*, and *M. tuberculosis* encode functional TA systems (Goulard et al. 2010; Mutschler et al. 2011; Yoshizumi et al. 2009; Hood et al. 2010; Wozniak and Waldor 2009; Audoly et al. 2011; Ramage et al. 2009). Interestingly, the *pezAT* locus of *S. pneumoniae* is located in a pathogenicity island and the PezT toxin was demonstrated to accelerate the progression of pneumococcal infections in mice, suggesting a role in virulence (Mutschler et al. 2011). In the case of *V. cholerae*, the *mosAT* TA locus was shown to promote the maintenance of SXT, an integrative and conjugative element that confers resistance to multiple antibiotics (Wozniak and Waldor 2009). In *P. aeruginosa*, the functional Tse2 toxin is addressed to other bacteria via the type VI secretion system, suggesting a potential role in biofilm formation, a key process in the chronic infection of cystic fibrosis patients by this bacterium (Hood et al. 2010). Remarkably, a recent study showed that VapC toxins from *Rickettsia* species could trigger apoptotic response in human cells, implying that TA toxins could be directly involved in virulence by poisoning host cells (Audoly et al. 2011).

Several studies suggest that TA systems could be involved in the formation of persisters, a phenomenon thought to be responsible for chronic infections (Lewis 2010). Persisters are specialized cells which constitute a subset of any bacterial population and are phenotypically drug-tolerant due to a slowed metabolism (Lewis 2010). Remarkably, the antitoxin gene *mqsA* is the most persistence-induced gene in *Escherichia coli* (Shah et al. 2006). The MqsA antitoxin from *E. coli* is known to antagonize the ribosome-independent RNase toxin MqsR and to regulate other stress genes than the *mqsRA* operon, including the major stress response regulator encoding gene *rpoS* (Wang et al. 2011). Consistent with the idea that TA modules contribute to persistence, an *E. coli* strain deleted for 10 TA loci was shown to produce reduced levels of persister cells (Maisonneuve et al. 2011). In *M. tuberculosis*, the causal agent of tuberculosis, a recent study of the drug-tolerant persisters transcriptome revealed that in addition to a massive shutdown of expression of the metabolic and biosynthetic pathways, a set of 10 TA systems were significantly upregulated together with other genes potentially involved in persisters formation (Keren et al. 2011). For this pathogen, persistence in hypoxic macrophages is a key step in the infectious cycle. This phenomenon is thought to be responsible for latent tuberculosis, an asymptomatic form of the disease concerning one third of the human population according to the WHO, which constitutes the major tuberculosis reservoir (www.who.int; Barry et al. 2009). The particularly high number of stress-responsive TA loci found in the *M. tuberculosis* chromosome and the ability of toxins to slow down bacterial growth raises the question of their implication in the setting up and/or the maintenance of the persistent state.

17.4 The Multiple TA Systems of *Mycobacterium tuberculosis*

The *M. tuberculosis* H37Rv chromosome potentially contains 75 TA systems (Fig. 17.1; Ramage et al. 2009; Gupta 2009; http://bioinfo-mml.sjtu.edu.cn/TADB/index.php). As it is generally the case, these systems are not randomly distributed across the *M. tuberculosis* chromosome but mostly clustered in hot spot regions (Makarova et al. 2009). Six well described TA pair families are represented: *vapBC* (50 systems), *mazEF* (10 systems), *yefM/yoeB* (1 system), *relBE* (2 systems), *higBA* (3 systems), and *parDE* (2 systems), while 7 systems could not be classified. A considerable subset of these TA systems (i.e., 63) has been experimentally tested, essentially in the bacterial models *E. coli* and *M. smegmatis*, for growth inhibition by the putative toxin and its neutralization by the putative antitoxin. The results are summarized in Fig. 17.1 (Gupta 2009; Ramage et al. 2009; Huang and He 2010; Singh et al. 2010; Zhu et al. 2010; Ahidjo et al. 2011). These studies identified a total of 37 TA systems that are functional in at least one of the tested conditions. In this part, we will describe each TA family present in *M. tuberculosis*, their mechanism of action and their potential implication in persistence.

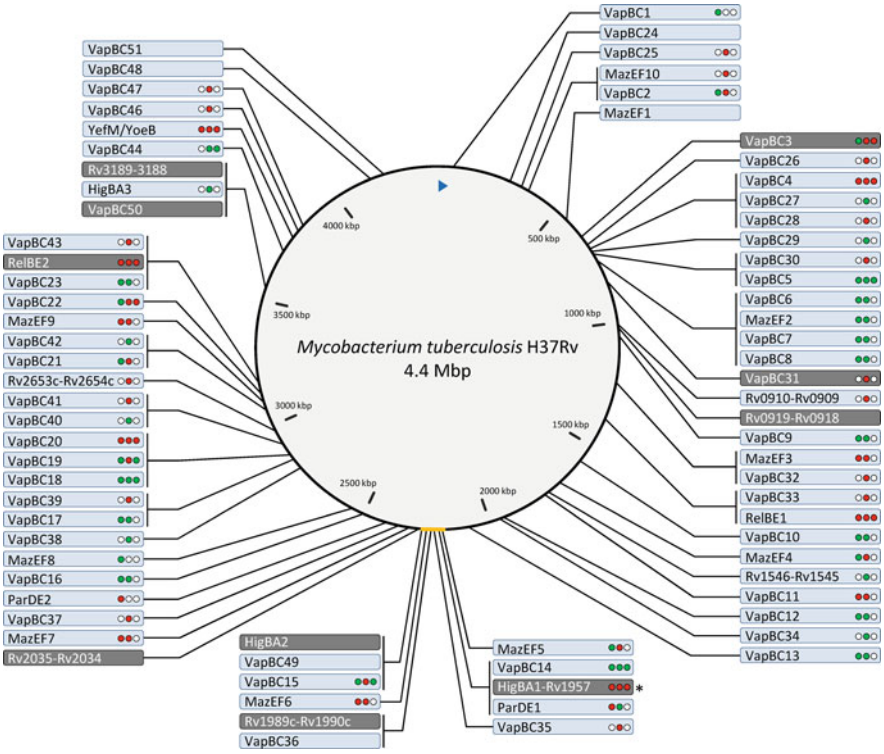


Fig. 17.1 Chromosomal map of *M. tuberculosis* H37Rv TA systems. Type II TA systems are annotated according to the tuberculist database (http://tuberculist.epfl.ch/index.html) except from VapBC49 (Rv2018-Rv2019), VapBC50 (Rv3181c-Rv3180c), VapBC51 (Rv3750c-Rv3749c), HigBA2 (Rv2022c-Rv2021c), HigBA3 (Rv3182-Rv3183), YefM/YoeB (Rv3357-Rv3358) and MazEF10 (Rv0298-Rv0299). For each system, the activity of the toxin is represented by *colored circles*. The first circle from the left represents experiments performed in *E. coli*, the second in *M. smegmatis,* and the third in *M. tuberculosis*; *red color* stands for "inhibition of growth", *green* for "no inhibition of growth" and *white* for "not tested". Systems that have never been tested experimentally display no circle. The 10 most induced TA systems in drug-tolerant persister cells are highlighted on *dark gray background*. The large genomic island containing TAC is depicted as a wide *orange line* and TAC is indicated by an *asterisk*

17.4.1 The vapBC Family

The most abundant TA loci in *M. tuberculosis* are from the VapBC (Virulence associated protein) family, characterized by a toxin with a PIN domain (homologous to PilT N-terminal domain). PIN domain proteins are found in all kingdoms of life and are generally involved in mRNA metabolism (Arcus et al. 2011). VapC toxins from diverse organisms, including *M. tuberculosis* VapC1, VapC2, VapC5, VapC11, and VapC29, were shown to have ribonuclease activity in vitro (Miallau et al. 2009; Ahidjo et al. 2011; Ramage et al. 2009), whereas *M. tuberculosis*

VapC4 toxicity seems to be due to stable RNA binding (Sharp et al. 2012). Winther and Gerdes (2011) showed that two VapC from the entero-pathogenic bacteria *Shigella flexneri* and *Salmonella enterica* are specific tRNAses that cleave the initiator tRNA. Thus, VapC toxins appear to act by distinct mechanisms, all of them targeting RNA.

The structure of the *M. tuberculosis* VapBC5 complex was solved and revealed an asymmetric complex with a ratio of one antitoxin for one toxin (Miallau et al. 2009). Yet, the VapB5 antitoxin inhibits the toxin by a mechanism which seems to be conserved, where the antitoxin prevents efficient binding of Mg^{2+} at the active site of VapC (Miallau et al. 2009). The *M. tuberculosis* VapB antitoxins are generally related to families of transcriptional regulators or DNA binding domains already reported to be associated with VapC toxins (Makarova et al. 2009), i.e., 33 RHH, 8 Phd, 2 ArbR, and 1 MerR. Yet, 5 other VapB do not present any conserved domain.

In *M. tuberculosis* several VapBC were activated in response to stresses encountered during infection: hypoxia (VapBC15, VapBC7, and VapBC25) and in IFN-γ-stimulated murine bone marrow-derived macrophages (VapBC11, VapBC3, and VapBC47) (Ramage et al. 2009). Moreover, VapBC3, VapBC31, and VapBC50 were shown to be among the 10 TA systems the most upregulated in drug-tolerant persisters, suggesting a possible role in this physiological state (Keren et al. 2011). Remarkably, a recent study in the closely related bacterium *M. smegmatis* revealed that its unique VapC toxin is an RNase involved in coupling the rate of glycerol utilization to bacterial growth via post-transcriptional regulation of genes involved in sugar transport (McKenzie et al. 2012). Therefore, it is conceivable that the multiple stress-responsive VapBC present in the *M. tuberculosis* chromosome could act as global regulators of translation in response to very diverse environmental stresses.

17.4.2 The mazEF *Family*

In *E. coli*, overexpression of the sequence-specific endoribonuclease MazF leads to growth arrest via cleavage of almost all cellular mRNAs (Yamaguchi and Inouye 2011). The *E. coli* MazEF complex, in which the MazF toxin is inactive, is a linear hetero-hexamer composed of a dimer of the MazE antitoxin symmetrically bound to two MazF dimers (Blower et al. 2011). Many MazF specific cleavage sites have been reported, for example *E. coli* MazF cleaves at ACA sequences (Zhang et al. 2005) and *Bacillus subtilis* MazF (EndoA) cleaves UACAU sequences (Park et al. 2011). From the 10 MazF family members present in *M. tuberculosis*, 7 were shown to affect *E. coli* and/or *M. smegmatis* growth, 2 exhibited non-observable phenotype, and one was never tested (Fig. 17.1; Ramage et al. 2009; Gupta 2009; Zhu et al. 2006). It has been shown that MazF9 cleaves UAC sequences, MazF3 cleaves U rich regions (Zhu et al. 2006), MazF6 cleaves at UUCCU or CUCCU sequences, and MazF4 at UCGCU (Zhu et al. 2008). Such differences in the size of

the cleavage sites imply a variation in the cleavage frequency. This suggests that recognition of 3 or 5 bases by MazF may not trigger the same type of response, the first leading more likely to a global translation inhibition, whereas the latter, having less frequent recognition sites, could regulate specific mRNAs in the cell. In this case, differential activation of TA systems depending on environmental conditions would lead to the appropriate adaptive response. In addition to mRNA recognition, the *E. coli* MazF was further shown to generate specialized ribosomes by cleavage of the 3′ region of the 16S rRNA, which contains the anti-Shine-Dalgarno sequence, and subsequently favor leaderless mRNAs translation (Vesper et al. 2011). Such a mechanism would lead to an alternative translation pathway, which could promote stress adaptation.

The expression of *mazEF* from *E. coli* was shown to be induced in various stress conditions, including DNA damage, heat shock and oxidative stress in a RelA dependent manner (Hazan et al. 2004). Similarly, *M. tuberculosis* MazF2 is downregulated after four hours of nutrient starvation, whereas, MazF3 is upregulated during amino acid starvation in a *relA* deleted strain (Betts et al. 2002; Dahl et al. 2003). These results suggest that *mazEF* modules from *M. tuberculosis* could be involved in the stringent response to amino acid starvation.

The *M. tuberculosis* MazF9 toxin was shown to be neutralized by the non-cognate antitoxins MazE6, VapB27, and VapB40 (Zhu et al. 2010). This is in agreement with the very low conservation between the different MazEs from *M. tuberculosis* (4–22 % of identity) and with the fact that VapB27 and VapB40 are closer to the *E. coli* MazE than any other MazE present in *M. tuberculosis*. In contrast, *M. tuberculosis* MazF toxins are well conserved with up to 46 % sequence identity. Together, these intriguing findings suggest that complex and fine-tuned cellular networks of toxin–antitoxin systems may exist in this pathogen. Finally, it has been shown that the mycobacterial MazF4 toxin and DNA topoisomerase I interact and mutually inhibit each other (Huang and He 2010). Since the DNA topoisomerase I gene is essential in *M. tuberculosis* (Sassetti et al. 2003), this suggests that MazF4 could act by two alternative mechanisms to inhibit mycobacterial growth.

17.4.3 The yefM/yoeB *and* relBE *Families*

The YefM/YoeB family, named after its *E. coli* member, was first described as homologous of the Axe-Txe addiction system of the multi-drug resistance plasmid pRUM of *Enterococcus faecium* (Grady and Hayes 2003). The YefM antitoxins are homologous to Phd, the antitoxin of the Phd-Doc module of phage P1, whereas the YoeB toxins belong to the RelE superfamily (Yamaguchi and Inouye 2011). In *M. tuberculosis*, the TA modules Rv1247c-Rv1246c, Rv2865-Rv2866, and Rv3357-Rv3358 are often referred to as RelBE1, RelBE2, and RelBE3, respectively, since the toxins belong to the RelE superfamily. Yet, the RelBE3 protein sequences are very close to those of the *E. coli yefM*/yoeB system (58 and 68 %

sequence similarities on protein levels, respectively), with the mycobacterial RelE3 toxin containing the conserved C-terminal histidine and tyrosine residues of YoeB involved in RNase activity (Kamada and Hanaoka 2005). Therefore, in contrast with *relBE1* and *2*, the *relBE3* system clearly belongs to the *yefM/yoeB* family and was thus renamed *yefM/yoeB*. In the cases of *relBE1* and *relBE2*, even if the antitoxins present YefM domains (PF02604), the toxins do not contain the YoeB conserved residues and the original appellation was thus retained.

Both RelE and YoeB from *E. coli* are ribosome-dependent ribonucleases that cleave mRNA at the ribosomal A site (Yamaguchi et al. 2011). Nevertheless, despite sequence and structural homology, *E. coli* RelE and YoeB seem to act by distinct mechanisms: RelE binds to the 30S subunit of 70S ribosomes and inhibits translation elongation, whereas YoeB binds to the 50S subunit and inhibits translation initiation (Yamaguchi et al. 2011). Gel filtration analyses of the *M. tuberculosis* YefM-YoeB complex suggest a heterotrimeric complex, as it is the case for the *E. coli* homologues (Kumar et al. 2008). The structure of *M. tuberculosis* YefM was solved and revealed a well-structured protein with a flexible C-terminal domain which superimposes well with the *E. coli* YefM from the complex structure (Kumar et al. 2008). As for YoeB, inactivation of *E. coli* RelE upon binding to RelB seems to occur via a conformational shift in the catalytic site of the toxin (Blower et al. 2011; Kamada and Hanaoka 2005). Both RelE and YoeB exhibit a microbial RNase fold, called RelE-like fold, also shared by other RNase toxins such as MqsR (Brown et al. 2009).

yefM/yoeB and the two *relBE* of *M. tuberculosis* were shown to be functional *bona fide* TA systems in *E. coli*, *M. smegmatis* as well as in *M. tuberculosis* (Korch et al. 2009; Singh et al. 2010). Yang and colleagues (2010) showed that there is complex cross regulation between these three systems with different patterns of cross interaction and thus of promoter regulation. For example, they showed that YefM interacts only with its cognate toxin YoeB, whereas RelB1 and RelB2 are able to interact with any of the RelE1, RelE2, or YoeB toxins. Moreover, YefM can bind its promoter on its own, whereas RelB1 and RelB2 need to be part of the TA complex, but not necessarily with their cognate toxin (Yang et al. 2010).

Interestingly, *relBE2* is among the 10 most induced TA systems in *M. tuberculosis* persister cells (Keren et al. 2011). Transcriptional analyzes also revealed that *yefM/yoeB*, *relBE1,* and *relBE2* genes are all expressed during growth in broth medium, whereas only *relE1*, *relB2,* and *yoeB* transcripts were detected in human macrophages in late stages of infection (Korch et al. 2009). The *relE1, relE2* and *yoeB* genes were also shown to be upregulated in response to antibiotic treatment and in lung tissues of infected mice (Singh et al. 2010). Remarkably, individual overexpression of the three toxin genes led to increased levels of persisters after antibiotic treatment, and deletion of *relE2* or *yoeB* affected persisters formation by 4–9 fold, while *relE1* deletion had no effect (Singh et al. 2010). However, single mutant strains of each toxin were not affected in in vivo growth or survival, indicating that they do not individually have direct roles in the infection (Singh et al. 2010).

17.4.4 The parDE *Family*

parDE was first identified on the broad-host-range plasmid RK2 as a stabilization system (Roberts and Helinski 1992). The ParE toxin acts by inhibiting DNA gyrase, thereby blocking DNA replication (Jiang et al. 2002). The crystal structure of the ParDE1 complex from *Caulobacter crescentus* revealed a heterotetrameric complex composed of two homodimers of toxin and antitoxin (Blower et al. 2011). As *C. crescentus* ParD1, *M. tuberculosis* ParD1 is predicted to have an RHH DNA binding domain, whereas ParD2 does not contain any conserved domains. With the exception of the unclassified toxins (Fig. 17.1) it is thus likely that ParE1 and ParE2 are the only toxins of *M. tuberculosis* that do not target RNA. Ectopic expression of ParE2 inhibited both *E. coli* and *M. smegmatis* growth whereas under the same conditions, ParE1 only affected *E. coli* growth (Ramage et al. 2009; Gupta 2009). So far, none of these systems were tested in *M. tuberculosis*.

17.4.5 Unclassified TA Modules

Among the 7 unclassified TA systems of *M. tuberculosis*, three were experimentally tested and two were functional in *M. smegmatis* (Ramage et al. 2009; Fig. 17.1). Four of these unclassified TA modules were identified as part of the 10 most induced TA systems in persister cells, but none of them have been tested for TA functions (Keren et al. 2011; Fig. 17.1). Except from the Rv2653c-Rv2654c system, all the unclassified modules contain conserved domains, and are paired together as follows (NCD is for No Conserved Domain): SRPBCC-NCD (Rv0910-Rv0909, Rv1546-Rv1545), GNAT-RHH (Rv0919-Rv0918), COG5654-Xre (Rv1989c-Rv1990, Rv3189-Rv3188), and COG3832-ArsR (Rv2035-Rv2034). The last three pairs of conserved domains have been predicted as putative TA families by Makarova and colleagues (2009).

17.4.6 The higBA *Family*

The first *higBA* locus (*h*ost *i*nhibition of *g*rowth) was found on the Rts1 plasmid of *Proteus vulgaris* where it functions as an addiction module (Tian et al. 1996). The HigB toxin from Rts1 plasmid was shown to act as a ribosome-dependent ribonuclease with a mechanism distinct from RelE and YoeB, in which the toxin binds the ribosome 50S subunit and cleaves at AAA sequences on the processing mRNAs (Hurley and Woychik 2009). HigB1 and HigB2 from *V. cholerae* were shown to have different, less specific cleavage patterns (Christensen-Dalsgaard and Gerdes 2006). The HigA antitoxins present a C-terminal DNA binding domain

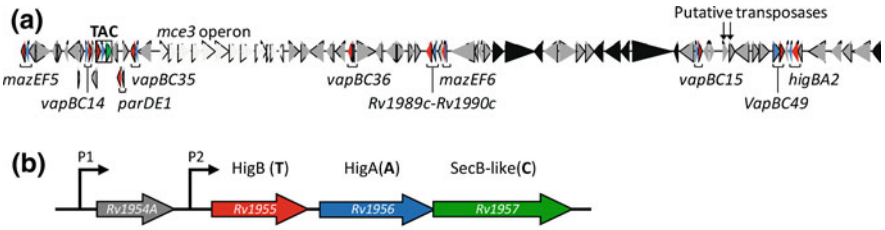


Fig. 17.2 Genomic context of the *M. tuberculosis* TAC system. **a** To-scale representation of the genomic island containing the TAC system, extending from the gene *Rv1942* to the gene *Rv2028*. TA loci are reported under the gene symbols colored in *red* for the toxin gene and *blue* for the antitoxin gene. The chaperone encoding gene *Rv1957* is colored in *green*. Genes with potential implication in virulence are highlighted in *light gray* for the *mce3* operon and in *black* for genes belonging to the dormancy regulon. Genes encoding putative transposases are also shown. **b** To-scale representation of the operon organization of the genes encoding the TAC system members: *higB* (378 bp), *higA* (450 bp) and *Rv1957* (576 bp). The promoters P1 and P2 are depicted, as well as the *Rv1954a* gene of unknown function

from the HTH-Xre family, clearly visible in the crystal structure of *E. coli* HigA (Arbing et al. 2010).

Inspection of TA systems in *M. tuberculosis* revealed that two systems predicted as *relBE* (http://bioinfo-mml.sjtu.edu.cn/TADB/) are in fact homologous of the hitherto unique *M. tuberculosis* HigBA1 system, Rv1955-Rv1956, and were thus herein renamed HigBA2 (Rv2022c-Rv2021c) and HigBA3 (Rv3182-Rv3183). The three *higBA* loci of *M. tuberculosis* present homology with the *gp49-gp48* locus from linear prophage N15 of *E. coli* which consists of a RelE-like toxin located upstream of an HTH Xre-domain containing antitoxin in the operon, an organization characteristic of HigBA modules (Gerdes et al. 2005; Makarova et al. 2009). While HigB2 has not been investigated yet, conditional expression of the HigB3 toxin in *M. smegmatis* displayed no apparent inhibition of growth (Fig. 17.1; Ramage et al. 2009). In contrast, HigB1 severely inhibited *E. coli*, *M. smegmatis* and *M. tuberculosis* growth (Gupta 2009; Ramage et al. 2009; Fivian-Hugues and Davis 2010). Remarkably, both *higBA1* and *higBA2* are among the 10 most upregulated TA systems in *M. tuberculosis* drug-tolerant persisters and are part of a large genomic island comprising other genes potentially involved in dormancy (Fig. 17.2a; Keren et al. 2011; Stinear et al. 2008). The *higBA1* locus is atypical in that it is part of an operon comprising two other genes: the upstream *Rv1954a* gene of unknown function and the downstream *Rv1957* gene encoding a SecB-like molecular chaperone (Fig. 17.2b; Smollet et al. 2009; Bordes et al. 2011). The proposed mechanism of action and activation as well as the evolutionary history of such novel tripartite toxin-antitoxin-chaperone, named TAC system, are presented below.

17.5 The Atypical Toxin-Antitoxin-Chaperone System of *Mycobacterium tuberculosis*

The TAC system HigB1-HigA1-Rv1957 of *M. tuberculosis* is part of a genomic island of approximately 80,760 bp (Stinear et al. 2008), which contains nine other TA systems, including five *vapBC*, 2 *mazEF*, 1 unclassified, *parDE1*, and *higBA2* (Fig. 17.2a). Among these, five are functional, and three are among the 10 most induced TA systems in persister cells (Fig. 17.1; Keren et al. 2011). Noticeably, this genomic island also contains potential pathogenicity determinants: one of the four *mce* operons of *M. tuberculosis*, *mce3*, which contains genes thought to be involved in lipid transport, as well as nine genes identified as being part of the dormancy regulon (Fig. 17.2a; de la Paz Santangelo et al. 2009; Voskuil et al. 2003).

17.5.1 Chaperone-Mediated Control of HigB1 Activation Cascade

It was recently shown that both in *E. coli* and in mycobacteria, the strong inhibition of growth induced by the mycobacterial toxin HigB1 from TAC is fully suppressed by the concerted action of the HigA1 antitoxin and the SecB-like chaperone Rv1957 (Bordes et al. 2011). In this case, the chaperone directly interacts with the HigA1 antitoxin and efficiently protects it from both aggregation and degradation by unknown protease(s) in vivo (Bordes et al. 2011). In agreement with such data, single deletion of the antitoxin gene *higA1* in *M. tuberculosis* is lethal (Fivian-Hughes and Davis 2010), while disruption of *Rv1957* alone exhibits a slow growth phenotype most likely due to a reduced antitoxin activity (Sassetti et al. 2003). These results indicate that the SecB-like chaperone Rv1957 indeed facilitates folding of the antitoxin and its subsequent interaction with the toxin (Fig. 17.3a). On the basis of these observations, it seems very likely that the HigB1 activation cascade could be triggered either by a decrease in expression or by direct hijacking of the chaperone at the post-translational level (see proposed model in Fig. 17.3b).

As stated above, Rv1957 shares significant sequence similarities with the export chaperone SecB. In most Gram-negative bacteria, the homotetrameric chaperone SecB binds non-native precursor proteins and specifically transfers them to the SecA subunit of the Sec translocase, thus facilitating their export (Randall and Hardy 2002). Remarkably, the SecB-like chaperone Rv1957 from TAC functionally replaces the chaperone SecB during protein export in *E. coli* (Bordes et al. 2011). Furthermore, as observed for SecB, Rv1957 is a homotetrameric chaperone, thus reinforcing the fact that both proteins might be evolutionarily related (Bordes et al. 2011). The presence of a well-defined outer membrane and a remarkably large number of putative outer membrane proteins in *M. tuberculosis* suggests that this bacterium could make use of a functional export chaperone under certain circumstances (Zuber et al. 2008; Niederweis et al. 2010). An attractive hypothesis is that the

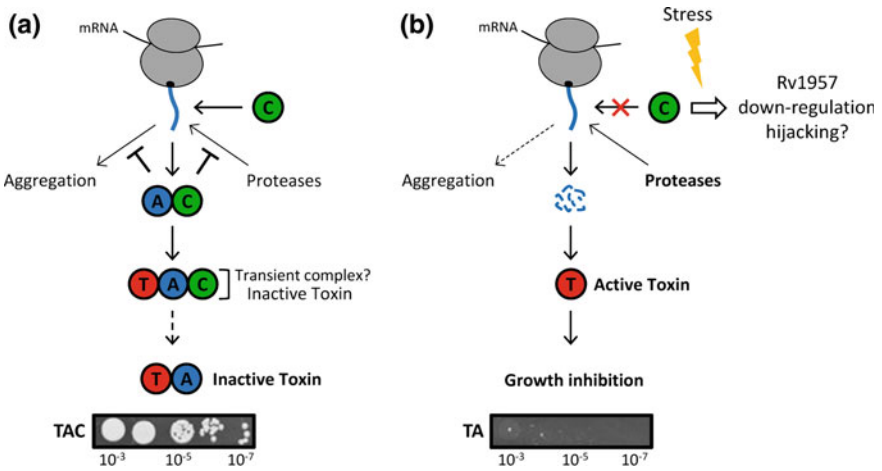


Fig. 17.3 Proposed model for TAC function in *M. tuberculosis*. **a** Under permissive conditions, the chaperone (*C*, *green*) interacts early with the nascent antitoxin (*A*, *blue*), protecting it from both aggregation and degradation by unknown protease(s). The chaperone-bound antitoxin is then competent for toxin (*T*, *red*) binding and neutralization. Yet, it is not known whether the chaperone Rv1957 is part of the inactive TA complex or not. As shown at the bottom of the figure, the growth of an *E. coli* strain expressing the three TAC partners (*TAC*) in a spot test experiment (dilutions 10^{-3}–10^{-7}) was not affected. **b** Specific stress conditions could induce a decrease in chaperone level or availability, rapidly resulting in antitoxin degradation and subsequent toxin activation (see text for details). Note that previous results suggest that aggregation may not be physiologically relevant in the presence of active degradation, hence the *dotted arrow*. A spot test experiment showing the growth inhibition of *E. coli* expressing *higB-higA* without chaperone (*TA*) is shown at the bottom of the figure (adapted from Bordes et al. 2011)

SecB-like chaperone represents an intimate link between export and activation of stress-responsive toxin–antitoxin systems. In this case, the cytoplasmic accumulation of precursor proteins or the synthesis of stress-specific preproteins could result in recruitment of Rv1957 to assist their efficient targeting to the Sec translocon. Increasing presecretory clients could compete with HigA1 for binding to the chaperone, thus facilitating degradation of the free antitoxin and the subsequent activation of the toxin until normal conditions resume (see proposed model in Fig. 17.3b). Yet, how the mycobacterial SecB-like chaperone responds to stress and controls the HigB1 toxin activation cascade, and to what extent such activation is important for *M. tuberculosis* virulence and pathogenesis are so far unresolved questions.

17.5.2 Stress Regulation of the TAC Operon

A number of experimental evidences point toward a role for TAC in *M. tuberculosis* stress adaptive response. Indeed, it has been shown that transcription of the *higB1-higA1-Rv1957* operon is induced by several relevant stress conditions

including DNA damage (Smollett et al. 2009), heat shock (Stewart et al. 2002), nutrient starvation (Betts et al. 2002), hypoxia (Ramage et al. 2009), drug persistence (Keren et al. 2011), and in host phagocytes (Tailleux et al. 2008).

The *TAC* operon is under the control of two promoters identified by 5'-RACE (Smollet et al. 2009): *higB*P1 located 51 nucleotides upstream of the start codon of *higB*, thus controlling the expression of *higB1-higA1-Rv1957*, and the more distal promoter *higB*P2 located 29 nucleotides upstream of the ATG start codon of the upstream *Rv1954a* gene of unknown function, to control the whole operon (Fig. 17.2b). The DNA damage inducible P1 promoter possesses a RecA-NDp motif characteristic of LexA/RecA-independent genes in *M. tuberculosis* (Smollett et al. 2009). Interestingly, the Clp protease gene regulator (ClpR) Rv2745c of *M. tuberculosis* was shown to specifically bind the conserved RecA-NDp motif of several DNA repair genes and to activate transcription of *clpC1* and *clpP1P2* protease genes as well (Sherrid et al. 2010, Estorninho et al. 2010). These results suggest that ClpR could be part of the regulatory network that controls TAC activation via degradation of the antitoxin (see below). Noticeably, such a mechanism could be of particular interest since in many cases, stress conditions that induce ClpR expression in *M. tuberculosis* often induce *higB* (Rustad et al. 2008; Mehra et al. 2010).

As with most TA loci, *TAC* is autorepressed and this is likely to occur via the binding of HigA1 to a palindromic motif overlapping the −35 element of the *higB*P2 promoter (Fivian-Hugues and Davis 2010). Yet, it is not established whether HigB1 and/or Rv1957 also participate in this process. The activity of HigA as a global transcriptional regulator was also investigated. Microarray analysis of the Δ*TAC* mutant revealed that the *fadB*, *Rv3173c,* and *Rv3662c* genes might be directly or indirectly activated by HigA (Fivian-Hugues and Davis 2010). In addition, a one-hybrid reporter system showed that HigA was able to interact specifically with the promoter region of gene clusters containing stress response genes *whiB7*, *rubA*, *rubB,* and fatty acid metabolism genes *echA20*, *fadE28*, *fadE29*, *ufaA2,* and *fad26* (Guo et al. 2009). Finally, ChIP-Seq analysis using FLAG-tagged HigA expressed in *M. tuberculosis* H37Rv revealed 30 potential HigA binding sites in intergenic regions, including the promoter region of *higB*P2 and another HigBA locus, *higBA3* (http://genome.tbdb.org/annotation/genome/tbdb/RegulatoryNetwork.html). Taken together, these data suggest that as for MqsA in *E. coli* (Wang et al. 2011), HigA could specifically regulate targets other than its own operon and thus be part of a network of stress-associated regulons involved in *M. tuberculosis* pathogenesis.

17.5.3 Role of Proteases in TAC Activation

Classical type II toxins are activated following degradation of their cognate antitoxins by stress proteases Lon and/or Clp (Gerdes et al. 2005). However, in the case of TAC, the additional presence of a SecB-like chaperone that assists the

antitoxin suggests a more competitive mode of regulation (Fig. 17.3). One of the main questions is thus how the HigB toxin is activated and to what extent stress proteases are involved. Previous work showed that in the absence of the chaperone, the HigA antitoxin was hardly detected in cell extracts, indicating that depletion of the chaperone triggers antitoxin degradation, even in the absence of activating stress conditions (Bordes et al. 2011).

The major stress protease families present in *M. tuberculosis* significantly differ from those of *E. coli*. The most striking difference is that *M. tuberculosis* possesses a 20S proteasome-like degradation machinery to which proteins are addressed by addition of an ubiquitine-like tag called Pup (Prokaryotic ubiquitin-like protein), giving the name of "Pupylation" to this pathway (Pearce et al. 2008). Yet, no evidence for a role of the proteasome in controlling TA activation has been described and none of the TAC components appeared to be pupylated (Poulsen et al. 2010), thus suggesting that the proteasome may not be involved in this case.

In contrast with *E. coli*, *M. tuberculosis* and closely related species do not have the AAA+ proteases Lon or ClpYQ. However, they do have two *clpP* genes, *clpP1* and *clpP2*, encoding the proteolytic subunit, and three ClpP-associated regulatory subunits encoding genes, *clpX, clpC1,* and *clpC2*, involved in substrate binding and specificity (Ribeiro-Guimaraes and Pessolani 2007). Remarkably, as for *higA1* and *higB1*, the *clpC2* gene is strongly induced in drug-tolerant persisters from *M. tuberculosis* (Keren et al. 2011), suggesting that ClpC2ClpP AAA+ stress protease could participate in antitoxin degradation upon chaperone depletion or hijacking, as proposed (Fig. 17.2b). Yet, the understanding of the underlying mechanism will first require the identification of such protease(s) involved in the degradation of the HigA1 antitoxin.

17.5.4 Evolutionary Analysis of TAC Systems

In order to study the evolutionary history of the *M. tuberculosis TAC* system and to analyze the occurrence of such system in the taxonomy, we searched for similar systems using iterative BLASTp, on the basis of sequence similarity and conserved organization and size. Rapidly, we noticed that sequence similarities were often weak (e-value around 10^{-03}). Nevertheless, when the organization and size criteria were filled and when at least two of the three partners showed even weak sequence similarities, the genes were then annotated as a putative *TAC* system. This approach allowed us to identify 50 putative TAC systems in both the complete and in progress genomes available in the NCBI database (http://www.ncbi.nlm.nih.gov/; Table 17.1). *TAC* systems are distributed in six phyla: Proteobacteria, Firmicutes, Actinobacteria, Thermotogae, Verrumicrobia, and Synergistetes. Strains containing *TAC* systems are always in the minority when analyzing phyla, except in Thermotogae. This wide and sparse pattern of distribution through the taxonomy cannot be easily explained by events of vertical transfer and loss, but rather suggests that the *TAC* systems were acquired by

Table 17.1 Taxonomic repartition of TAC systems

Taxonomic group	Strain	T/A/C	Genome Ref.	
β-protéobactéries	*Achromobacter xylosoxidans* C54	HMPREF0005_02907/08/09	ACRC01000563.1	*
	Cupriavidus taiwanensis LMG19424	pRALTA_0477/78/79	**CU633751.1**	
	Gallionella capsiferriformans ES-2	Galf_0727/28/29	CP002159.1	
	Ralstonia solanacearum CFBP2957	RCFBP_mp30545/44/43	**FP885907.1**	*
δ-protéobactéries	*Desulfatibacillum alkenivorans* AK-01	Dalk_1596/97/98	CP001322.1	
	Desulfobacca acetoxidans DSM 11109	Desac_2259/60/61	CP002629.1	
	Desulfohalobium retbaense DSM 5692	Dret_0389/90/91	CP001734.1	
	Geobacter lovleyi ATCC BAA-1151	Glov_3492/91/90	CP001089.1	
	Geobacter uraniireducens Rf4	Gura_1411/10/09	CP000698.1	
	Pelobacter carbinolicus DSM 2380	Pcar_0788/89/90	CP000142.2	
	Syntrophus aciditrophicus SB	SYN_00016/15/14	CP000252.1	
	uncultured Desulfobacterium sp. env.	N47_E41910/20/30	FR695877.1	
γ-protéobactéries	*Acinetobacter johnsonii* SH046	HMPREF0016_03309/10/11	GG704981.1	
	Methylomonas methanica MC09	Metme_2813/12/11	CP002738.1	
	Nitrosococcus halophilus Nc4	Nhal_3521/22/23	CP001798.1	
	Pseudomonas fluorescens SBW25	PFLU_4126A/B/C	AM181176.4	
	Vibrio cholerae RC385	VCRC385_03678/79/80	GG774561.1	
	Vibrio cholerae HE-09	VCHE09_0202/03/04	AFOP01000111.1	*
Actinobacteria	*Clavibacter michiganensis subsp. Sepedonicus* ATCC 33113	CMS2954/55/56	AM849034.1	*
		pCS0043/44/45	**AM849035.1**	*
	Corynebacterium glutamicum ATCC 13032	Cgl1746/45/44	BA000036.3	
	Microlunatus phosphovorus NM-1	MLP_49450/60/70	AP012204.1	
	Mycobacterium bovis AF2122/97	Mb1990/91/92	BX248340.1	*
	Mycobacterium bovis BCG Pasteur 1173P2	BCG_1994/95/96	AM408590.1	
	Mycobacterium bovis Tokyo 172	JTY_1978/79/80	AP010918.1	
	Mycobacterium gilvum PYR-GCK	Mflv_3208/07/06	CP000656.1	
	Mycobacterium tuberculosis H37Rv	Rv1955/56/57	BX842578.1	*
	Mycobacterium tuberculosis CDC1551	MT2004/05/06	AE000516.2	*
	Mycobacterium tuberculosis KZN 1435	TBMG_02037/36/35	CP001658.1	*
	Mycobacterium tuberculosis H37Ra	MRA_1964/65/66	CP000611.1	*
	Mycobacterium tuberculosis F11	TBFG_01945/46/47	CP000717.1	*
	Propionibacterium freudenreichii subsp. shermanii CIRM-BIA1	PFREUD_02980/70/60	FN806773.1	
	Rhodococcus opacus B4	ROP_pROB01-01190/80/70	**AP011116.1**	
	Xylanimonas cellulosilytica DSM 15894	Xcel_2923/22/21	CP001821.1	
Firmicutes	*Alkaliphilus metalliredigens* QYMF	Amet_3134/33/32	CP000724.1	
	Caldicellulosiruptor bescii DSM 6725	Athe_2320/19/18	CP001393.1	
	Caldicellulosiruptor hydrothermalis DSM 18901	Calhy_0207/08/09	CP002219.1	
	Caldicellulosiruptor kronotskyensis DSM 18902	Calkro_0234/35/36	CP002330.1	
	Caldicellulosiruptor saccharolyticus DSM 8903	Csac_0625/26/27	CP000679.1	
	Lactobacillus plantarum subsp. plantarum ATCC 14917	HMPREF0531_10367/68/69	ACGZ02000005.1	
	Thermoanaerobacterium xylanolyticum LX-11	Thexy_1763/62/61	CP002739.1	
Thermotogae	*Petrotoga mobilis* SJ95	Pmob_1270/71/72	CP000879.1	
	Thermotoga lettingae TMO	Tlet_0094/93/92	CP000812.1	
	Thermotoga maritima MSB8	TM1329/30/31	AE000512.1	
	Thermotoga naphthophila RKU-10	Tnap_1472/71/70	CP001839.1	
	Thermotoga neapolitana DSM 4359	CTN_1254/53/52	CP000916.1	
	Thermotoga petrophila RKU-1	Tpet_1452/51/50	CP000702.1	
	Thermotogales bacterium mesG1.Ag.4.2 ctg90	ThebaDRAFT_1153/54/55	AEDC01000002.1	
Synergistetes	*Aminobacterium colombiense* DSM 12261	Amico_0944/45/46	CP001997.1	
Verrumicrobia	*Opitutus terrae* DSM 11246	Oter_0529/30/31	CP001032.1	

For each identified TAC system the locus tag of each gene is given (T/A/C) as well as the corresponding EMBL genome reference. The codes shown in bold characters indicate that TAC is present on a plasmid. Asterisks designate known pathogenic bacteria

horizontal gene transfer. This is in agreement with *TAC* localization on a large genomic island in strains from the *M. tuberculosis* complex (Fig. 17.2a). Moreover, 4 *TAC* systems are plasmid-borne and thus potentially involved in plasmid stability (Table 17.1). Given these observations, the evolutionary history of *TAC* systems is likely to be highly complex and largely mediated by mobile genetic elements (manuscript in preparation).

We explored existing links between solitary SecB and chaperones from *TAC* and found that all *TAC* chaperones present significant homology links with the SecB chaperone family (Pfam02556; http://pfam.sanger.ac.uk). In addition, we noticed that in some δ-proteobacteria *TAC* chaperones were very similar to the known SecB consensus and were already annotated as SecB on these genomes (manuscript in preparation). Taken together, these observations further confirm the evolutionarily link between the SecB and *TAC* chaperones.

We also classified the *TAC* antitoxins according to their sequence similarity and found that three well-defined antitoxin families could be associated with SecB-like chaperones, namely HigA, MqsA, and HicB (manuscript in preparation). Interestingly, all the TA systems associated with a SecB-like chaperone identified to date present an unusual genetic organization with the toxin gene being upstream of the antitoxin gene in the operon. This organization is thought to be unfavorable to the formation of a "protective" TA complex since the stable toxin is synthesized before the unstable and protease-sensitive antitoxin. In the case of TAC, the dedicated chaperone could maintain a pool of native antitoxin protected from proteolysis as part of the antitoxin-chaperone complex and thus competent for the neutralization of the toxin as soon as it emerges from the ribosome.

17.6 Concluding Remarks

The remarkable prevalence of multiple chromosomally encoded type II TA systems in bacterial genomes raises questions about their role in cell physiology. The fact that many TA systems respond to stress conditions encountered during the infection process and that expression of toxins often leads to persistence phenomena alludes to possible roles in bacterial virulence. Remarkably, one of the main features of *M. tuberculosis* pathogenesis is its ability to persist long term in the host granulomas in a non-replicating and drug-tolerant state, and later awaken to cause disease. The presence of at least 75 TA systems in its chromosome, of which 37 appear to be functional in vivo (Fig. 17.1), suggests that such intricate networks of stress-responsive elements could indeed participate in *M. tuberculosis* virulence and thus represent potentially promising new drug targets. Further studies are clearly warranted to address such fundamental questions.

Acknowledgments We thank Virginie Calderon and all the members of the Genevaux laboratory for insightful discussions. This work was supported by a French MENRT fellowship to AS and an ATIP-CNRS grant to PG.

References

Ahidjo, B. A., Kuhnert, D., McKenzie, J. L., Machowski, E. E., Gordhan, B. G., Arcus, V., et al. (2011). VapC toxins from *Mycobacterium tuberculosis* are ribonucleases that differentially inhibit growth and are neutralized by cognate VapB antitoxins. *PLoS ONE, 6*, e21738.

Arbing, M. A., Handelman, S. K., Kuzin, A. P., Verdon, G., Wang, C., Su, M., et al. (2010). Crystal structures of Phd-Doc, HigA, and YeeU establish multiple evolutionary links between microbial growth-regulating toxin–antitoxin systems. *Structure, 18*, 996–1010.

Arcus, V. L., McKenzie, J. L., Robson, J., & Cook, G. M. (2011). The PIN-domain ribonucleases and the prokaryotic VapBC toxin–antitoxin array. *Protein Engineering, Design & Selection, 24*, 33–40.

Audoly, G., Vincentelli, R., Edouard, S., Georgiades, K., Mediannikov, O., Gimenez, G., et al. (2011). Effect of rickettsial toxin VapC on its eukaryotic host. *PLoS ONE, 6*, e26528.

Barry, C. E., 3rd, Boshoff, H. I., Dartois, V., Dick, T., Ehrt, S., Flynn, J., et al. (2009). The spectrum of latent tuberculosis: rethinking the biology and intervention strategies. *Nature Reviews Microbiology, 7*, 845–855.

Betts, J. C., Lukey, P. T., Robb, L. C., McAdam, R. A., & Duncan, K. (2002). Evaluation of a nutrient starvation model of *Mycobacterium tuberculosis* persistence by gene and protein expression profiling. *Molecular Microbiology, 43*, 717–731.

Blower, T. R., Salmond, G. P., & Luisi, B. F. (2011). Balancing at survival's edge: the structure and adaptive benefits of prokaryotic toxin–antitoxin partners. *Current Opinion in Structural Biology, 21*, 109–118.

Bordes, P., Cirinesi, A. M., Ummels, R., Sala, A., Sakr, S., Bitter, W., et al. (2011). SecB-like chaperone controls a toxin–antitoxin stress-responsive system in *Mycobacterium tuberculosis*. *Proceedings of the National Academy of Sciences of the United States of America, 108*, 8438–8443.

Brown, B. L., Grigoriu, S., Kim, Y., Arruda, J. M., Davenport, A., Wood, T. K., et al. (2009). Three dimensional structure of the MqsR:MqsA complex: a novel TA pair comprised of a toxin homologous to RelE and an antitoxin with unique properties. *PLoS Pathogens, 5*, e1000706.

Christensen-Dalsgaard, M., & Gerdes, K. (2006). Two *higBA* loci in the *Vibrio cholerae* superintegron encode mRNA cleaving enzymes and can stabilize plasmids. *Molecular Microbiology, 62*, 397–411.

Dahl, J. L., Kraus, C. N., Boshoff, H. I., Doan, B., Foley, K., Avarbock, D., et al. (2003). The role of RelMtb-mediated adaptation to stationary phase in long-term persistence of *Mycobacterium tuberculosis* in mice. *Proceedings of the National Academy of Sciences of the United States of America, 100*, 10026–10031.

de la Paz Santangelo, M., Klepp, L., Nunez-Garcia, J., Blanco, F. C., Soria, M., Garcia-Pelayo, M. C., et al. (2009). Mce3R, a TetR-type transcriptional repressor, controls the expression of a regulon involved in lipid metabolism in *Mycobacterium tuberculosis*. *Microbiology, 155*, 2245–2255.

Estorninho, M., Smith, H., Thole, J., Harders-Westerveen, J., Kierzek, A., Butler, R. E., et al. (2010). ClgR regulation of chaperone and protease systems is essential for *Mycobacterium tuberculosis* parasitism of the macrophage. *Microbiology, 156*, 3445–3455.

Fivian-Hughes, A. S., & Davis, E. O. (2010). Analyzing the regulatory role of the HigA antitoxin within *Mycobacterium tuberculosis*. *Journal of Bacteriology, 192*, 4348–4356.

Georgiades, K., & Raoult, D. (2011). Genomes of the most dangerous epidemic bacteria have a virulence repertoire characterized by fewer genes but more toxin–antitoxin modules. *PLoS ONE, 6*, e17962.

Gerdes, K., Christensen, S. K., & Lobner-Olesen, A. (2005). Prokaryotic toxin–antitoxin stress response loci. *Nature Reviews Microbiology, 3*, 371–382.

Goulard, C., Langrand, S., Carniel, E., & Chauvaux, S. (2010). The *Yersinia pestis* chromosome encodes active addiction toxins. *Journal of Bacteriology, 192*, 3669–3677.

Grady, R., & Hayes, F. (2003). Axe-Txe, a broad-spectrum proteic toxin–antitoxin system specified by a multidrug-resistant, clinical isolate of *Enterococcus faecium*. *Molecular Microbiology, 47*, 1419–1432.

Guo, M., Feng, H., Zhang, J., Wang, W., Wang, Y., Li, Y., et al. (2009). Dissecting transcription regulatory pathways through a new bacterial one-hybrid reporter system. *Genome Research, 19*, 1301–1308.

Gupta, A. (2009). Killing activity and rescue function of genome-wide toxin–antitoxin loci of *Mycobacterium tuberculosis*. *FEMS Microbiology Letters, 290*, 45–53.

Hazan, R., Sat, B., & Engelberg-Kulka, H. (2004). *Escherichia coli* mazEF-mediated cell death is triggered by various stressful conditions. *Journal of Bacteriology, 186*, 3663–3669.

Hood, R. D., Singh, P., Hsu, F., Guvener, T., Carl, M. A., Trinidad, R. R., et al. (2010). A type VI secretion system of *Pseudomonas aeruginosa* targets a toxin to bacteria. *Cell Host & Microbe, 7*, 25–37.

Huang, F., & He, Z. G. (2010). Characterization of an interplay between a *Mycobacterium tuberculosis* MazF homolog, Rv1495 and its sole DNA topoisomerase I. *Nucleic Acids Research, 38*, 8219–8230.

Hurley, J. M., & Woychik, N. A. (2009). Bacterial toxin HigB associates with ribosomes and mediates translation-dependent mRNA cleavage at A-rich sites. *Journal of Biological Chemistry, 284*, 18605–18613.

Jiang, Y., Pogliano, J., Helinski, D. R., & Konieczny, I. (2002). ParE toxin encoded by the broad-host-range plasmid RK2 is an inhibitor of *Escherichia coli* gyrase. *Molecular Microbiology, 44*, 971–979.

Kamada, K., & Hanaoka, F. (2005). Conformational change in the catalytic site of the ribonuclease YoeB toxin by YefM antitoxin. *Molecular Cell, 19*, 497–509.

Keren, I., Minami, S., Rubin, E., & Lewis, K. (2011). Characterization and transcriptome analysis of *Mycobacterium tuberculosis* persisters. *MBio, 2*, 00100–00111.

Korch, S. B., Contreras, H., & Clark-Curtiss, J. E. (2009). Three *Mycobacterium tuberculosis* Rel toxin–antitoxin modules inhibit mycobacterial growth and are expressed in infected human macrophages. *Journal of Bacteriology, 191*, 1618–1630.

Kumar, P., Issac, B., Dodson, E. J., Turkenburg, J. P., & Mande, S. C. (2008). Crystal structure of *Mycobacterium tuberculosis* YefM antitoxin reveals that it is not an intrinsically unstructured protein. *Journal of Molecular Biology, 383*, 482–493.

Lewis, K. (2010). Persister cells. *Annual Review of Microbiology, 64*, 357–372.

Maisonneuve, E., Shakespeare, L. J., Jorgensen, M. G., & Gerdes, K. (2011). Bacterial persistence by RNA endonucleases. *Proceedings of the National Academy of Sciences of the United States of America, 108*, 13206–13211.

Makarova, K. S., Wolf, Y. I., & Koonin, E. V. (2009). Comprehensive comparative-genomic analysis of type 2 toxin–antitoxin systems and related mobile stress response systems in prokaryotes. *Biology Direct, 4*, 19.

McKenzie, J.L., Robson, J., Berney, M., Smith, T.C., Ruthe, A., Gardner, P.P., et al. (2012) A VapBC toxin-antitoxin module is a post-transcriptional regulator of metabolic flux in mycobacteria. *Journal of Bacteriology, 194*, 2189–21204.

Mehra, S., Pahar, B., Dutta, N. K., Conerly, C. N., Philippi-Falkenstein, K., Alvarez, X., et al. (2010). Transcriptional reprogramming in nonhuman primate (rhesus macaque) tuberculosis granulomas. *PLoS ONE, 5*, e12266.

Miallau, L., Faller, M., Chiang, J., Arbing, M., Guo, F., Cascio, D., et al. (2009). Structure and proposed activity of a member of the VapBC family of toxin–antitoxin systems. VapBC-5 from *Mycobacterium tuberculosis*. *Journal of Biological Chemistry, 284*, 276–283.

Mutschler, H., Gebhardt, M., Shoeman, R. L., & Meinhart, A. (2011). A novel mechanism of programmed cell death in bacteria by toxin–antitoxin systems corrupts peptidoglycan synthesis. *PLoS Biology, 9*, e1001033.

Niederweis, M., Danilchanka, O., Huff, J., Hoffmann, C., & Engelhardt, H. (2010). Mycobacterial outer membranes: in search of proteins. *Trends in Microbiology, 18*, 109–116.

Pandey, D. P., & Gerdes, K. (2005). Toxin–antitoxin loci are highly abundant in free-living but lost from host-associated prokaryotes. *Nucleic Acids Research, 33*, 966–976.

Park, J. H., Yamaguchi, Y., & Inouye, M. (2011). *Bacillus subtilis* MazF-bs (EndoA) is a UACAU-specific mRNA interferase. *FEBS Letters, 585*, 2526–2532.

Pearce, M. J., Mintseris, J., Ferreyra, J., Gygi, S. P., & Darwin, K. H. (2008). Ubiquitin-like protein involved in the proteasome pathway of *Mycobacterium tuberculosis*. *Science, 322*, 1104–1107.

Poulsen, C., Akhter, Y., Jeon, A. H., Schmitt-Ulms, G., Meyer, H. E., Stefanski, A., et al. (2010). Proteome-wide identification of mycobacterial pupylation targets. *Molecular System Biology, 6*, 386.

Ramage, H. R., Connolly, L. E., & Cox, J. S. (2009). Comprehensive functional analysis of *Mycobacterium tuberculosis* toxin–antitoxin systems: implications for pathogenesis, stress responses, and evolution. *PLoS Genetics, 5*, e1000767.

Randall, L. L., & Hardy, S. J. (2002). SecB, one small chaperone in the complex milieu of the cell. *Cellular & Molecular Life Sciences, 59*, 1617–1623.

Ribeiro-Guimaraes, M. L., & Pessolani, M. C. (2007). Comparative genomics of mycobacterial proteases. *Microbial Pathogenesis, 43*, 173–178.

Roberts, R. C., & Helinski, D. R. (1992). Definition of a minimal plasmid stabilization system from the broad-host-range plasmid RK2. *Journal of Bacteriology, 174*, 8119–8132.

Rustad, T. R., Harrell, M. I., Liao, R., & Sherman, D. R. (2008). The enduring hypoxic response of *Mycobacterium tuberculosis*. *PLoS ONE, 3*, e1502.

Sassetti, C. M., Boyd, D. H., & Rubin, E. J. (2003). Genes required for mycobacterial growth defined by high density mutagenesis. *Molecular Microbiology, 48*, 77–84.

Shah, D., Zhang, Z., Khodursky, A., Kaldalu, N., Kurg, K., & Lewis, K. (2006). Persisters: a distinct physiological state of *E. coli*. *BMC Microbiology, 6*, 53.

Sharp, J.D., Cruz, J.W., Raman, S., Inouye, M., Husson, R.N. and Woychik, N.A. (2012) Growth and translation inhibition through sequence specific RNA binding by a *Mycobacterium tuberculosis* VAPC toxin. *Journal of Biological Chemistry, 287*, 12835–12847.

Sherrid, A. M., Rustad, T. R., Cangelosi, G. A., & Sherman, D. R. (2010). Characterization of a Clp protease gene regulator and the reaeration response in *Mycobacterium tuberculosis*. *PLoS ONE, 5*, e11622.

Singh, R., Barry, C. E., 3rd, & Boshoff, H. I. (2010). The three RelE homologs of *Mycobacterium tuberculosis* have individual, drug-specific effects on bacterial antibiotic tolerance. *Journal of Bacteriology, 192*, 1279–1291.

Smollett, K. L., Fivian-Hughes, A. S., Smith, J. E., Chang, A., Rao, T., & Davis, E. O. (2009). Experimental determination of translational start sites resolves uncertainties in genomic open reading frame predictions—application to *Mycobacterium tuberculosis*. *Microbiology, 155*, 186–197.

Stewart, G. R., Wernisch, L., Stabler, R., Mangan, J. A., Hinds, J., Laing, K. G., et al. (2002). Dissection of the heat-shock response in *Mycobacterium tuberculosis* using mutants and microarrays. *Microbiology, 148*, 3129–3138.

Stinear, T. P., Seemann, T., Harrison, P. F., Jenkin, G. A., Davies, J. K., Johnson, P. D., et al. (2008). Insights from the complete genome sequence of *Mycobacterium marinum* on the evolution of *Mycobacterium tuberculosis*. *Genome Research, 18*, 729–741.

Tailleux, L., Waddell, S. J., Pelizzola, M., Mortellaro, A., Withers, M., Tanne, A., et al. (2008). Probing host pathogen cross-talk by transcriptional profiling of both *Mycobacterium tuberculosis* and infected human dendritic cells and macrophages. *PLoS ONE, 3*, e1403.

Tian, Q. B., Ohnishi, M., Tabuchi, A., & Terawaki, Y. (1996). A new plasmid-encoded proteic killer gene system: cloning, sequencing, and analyzing *hig* locus of plasmid Rts1. *Biochemical & Biophysical Research Communications, 220*, 280–284.

Van Melderen, L. (2010). Toxin–antitoxin systems: why so many, what for? *Current Opinion in Microbiology, 13*, 781–785.

Vesper, O., Amitai, S., Belitsky, M., Byrgazov, K., Kaberdina, A. C., Engelberg-Kulka, H., et al. (2011). Selective translation of leaderless mRNAs by specialized ribosomes generated by MazF in *Escherichia coli*. *Cell, 147*, 147–157.

Voskuil, M. I., Schnappinger, D., Visconti, K. C., Harrell, M. I., Dolganov, G. M., Sherman, D. R., et al. (2003). Inhibition of respiration by nitric oxide induces a *Mycobacterium tuberculosis* dormancy program. *Journal of Experimental Medicine, 198*, 705–713.

Wang, X., Kim, Y., Hong, S. H., Ma, Q., Brown, B. L., Pu, M., et al. (2011). Antitoxin MqsA helps mediate the bacterial general stress response. *Nature Chemical Biology, 7*, 359–366.

Winther, K. S., & Gerdes, K. (2011). Enteric virulence associated protein VapC inhibits translation by cleavage of initiator tRNA. *Proceedings of the National Academy of Sciences of the United States of America, 108*, 7403–7407.

Wozniak, R. A., & Waldor, M. K. (2009). A toxin–antitoxin system promotes the maintenance of an integrative conjugative element. *PLoS Genetics, 5*, e1000439.

Yamaguchi, Y., & Inouye, M. (2011). Regulation of growth and death in *Escherichia coli* by toxin–antitoxin systems. *Nature Reviews Microbiology, 9*, 779–790.

Yamaguchi, Y., Park, J. H., & Inouye, M. (2011). Toxin–antitoxin systems in bacteria and archaea. *Annual Review of Genetics, 45*, 61–79.

Yang, M., Gao, C., Wang, Y., Zhang, H., & He, Z. G. (2010). Characterization of the interaction and cross-regulation of three *Mycobacterium tuberculosis* RelBE modules. *PLoS ONE, 5*, e10672.

Yoshizumi, S., Zhang, Y., Yamaguchi, Y., Chen, L., Kreiswirth, B. N., & Inouye, M. (2009). *Staphylococcus aureus* YoeB homologues inhibit translation initiation. *Journal of Bacteriology, 191*, 5868–5872.

Zhang, Y., Zhu, L., Zhang, J., & Inouye, M. (2005). Characterization of ChpBK, an mRNA interferase from *Escherichia coli*. *Journal of Biological Chemistry, 280*, 26080–26088.

Zhu, L., Phadtare, S., Nariya, H., Ouyang, M., Husson, R. N., & Inouye, M. (2008). The mRNA interferases, MazF-mt3 and MazF-mt7 from *Mycobacterium tuberculosis* target unique pentad sequences in single-stranded RNA. *Molecular Microbiology, 69*, 559–569.

Zhu, L., Sharp, J. D., Kobayashi, H., Woychik, N. A., & Inouye, M. (2010). Noncognate *Mycobacterium tuberculosis* toxin–antitoxins can physically and functionally interact. *Journal of Biological Chemistry, 285*, 39732–39738.

Zhu, L., Zhang, Y., Teh, J. S., Zhang, J., Connell, N., Rubin, H., et al. (2006). Characterization of mRNA interferases from *Mycobacterium tuberculosis*. *Journal of Biological Chemistry, 281*, 18638–18643.

Zuber, B., Chami, M., Houssin, C., Dubochet, J., Griffiths, G., & Daffe, M. (2008). Direct visualization of the outer membrane of mycobacteria and corynebacteria in their native state. *Journal of Bacteriology, 190*, 5672–5680.

Chapter 18
Toxin-Antitoxin Loci in *Streptococcus pneumoniae*

Wai Ting Chan, Inma Moreno-Córdoba, Chew Chieng Yeo and Manuel Espinosa

Abstract The Gram-positive human pathogen *Streptococcus pneumoniae* harbours a number of type II toxin–antitoxin loci, three of which have been functionally characterised. The pneumococcal *relBE2* and *yefM-yoeB* TA systems are described here in detail in terms of their genetic organisation, gene regulation and protein structure. A number of putative TA loci with unusual genetic organisations are also presented. The possible roles of the pneumococcal TA loci in persistence and the phenomenon of bistability are discussed.

Abbreviations

(TAS)	Toxin–antitoxin system(s)
(nt)	Nucleotides
(IR-1 and IR-2)	Inverted repeats
(ORF)	Open reading frame

18.1 Introduction

Streptococcus pneumoniae (the pneumococcus) is a Gram-positive pathogenic bacterium that is responsible for a number of infections from symptomatic to acute that still pose a relentless threat to mankind not only in developing countries, but also in industrialised countries. Pneumococcal infections cause high rates of

W. T. Chan · I. Moreno-Córdoba · M. Espinosa (✉)
Centro de Investigaciones Biológicas, Consejo Superior de Investigaciones Científicas, Ramiro de Maeztu 9, E28040 Madrid, Spain
e-mail: mespinosa@cib.csic.es

C. C. Yeo
Faculty of Agriculture and Biotechnology, Universiti Sultan Zainal Abidin, Jalan Sultan Mahmud, 20400 Kuala Terengganu, Malaysia

K. Gerdes (ed.), *Prokaryotic Toxin-Antitoxins*,
DOI: 10.1007/978-3-642-33253-1_18,

morbidity and mortality among children, adults and the elderly and represent a huge socioeconomic burden. *S. pneumoniae* is the causal agent of pneumococcal pneumonia, of more than 50 % of meningitis, and also of sepsis, acute otitis media and pronounced nosocomial diseases in hospitals. Pneumococcus is responsible for the deaths of nearly 2 million children below 5 years annually, that is more than AIDS, tuberculosis and malaria combined (O'Brien et al. 2009). The bacterium is naturally found in the surface of the upper respiratory tract where it colonises between 5 and 70 % of healthy individuals, depending on geographical, seasonal and socioeconomical conditions (Bogaert et al. 2004). From its natural niche, the pneumococci spread to invade new tissues or individuals that would develop illness when their immunological system is compromised. In addition to the high death rate caused by pneumococcal infections, there is a huge economical impact due to non-fatal illness from the loss of working hours, drug treatments and hospital expenses, which may amount to the thousands of millions € per year.

S. pneumoniae is naturally transformable with exogenous DNA and in conjunction with its hyper-recombinogenic nature, has converted the bacterium into a reservoir of horizontally transferred antibiotic resistance and virulence genes (Hanage et al. 2009). Horizontal DNA transfer among closely related streptococci has been elegantly shown by determination of the mosaic organisation of some virulence factors like the choline- and penicillin-binding proteins (Moscoso et al. 2010). The efficacy of treatment of pneumococcal diseases has been reduced due to the increasing number of clinical isolates of *S. pneumoniae* that exhibit antibiotic resistances (Gamez and Hammerschmidt 2012) and the characterisation of capsule-deficient virulent strains that are non-typeable (Park et al. 2012). In spite of the success achieved after implementation of a wide vaccination programme with the heptavalent vaccine, there has been a new increase in resistance rates due to the spread of serotype 19A, especially the multidrug-resistant ST320 and ST276 clones, due to replacement of vaccine-strain serotypes by non-vaccine serotypes (Perez-Trallero et al. 2012). The development of novel combat strategies is urgently needed and is a big challenge for the international pneumococcal research community. There are three prerequisites to achieve this goal. It is important firstly to understand the distribution of serotypes, the occurrence of antibiotic resistance, and the frequency of non-typeable strains; second, to identify and characterise new intervention targets by studying host-pathogen interactions, and third to validate these targets in in vitro and in vivo infection models in order to identify potential candidates for prevention and therapy. It is thus urgent to find new strategies to tackle pneumococcal infections, and a number of approaches (mostly complementary with each other) have been suggested such as: (1) rotation schemes in antibiotic treatments; (2) vaccination with more polyvalent vaccines (13- and 20-valent, for instance) (O'Brien et al. 2009); (3) targeting key virulence factors that are conserved in all sequenced genomes; (4) bacteriophage therapy; (5) functional determination of the human microbiome of a large cohort of patients; (6) synthesis of novel molecules with anti-pneumococcal activity and (7) employment of new strategies like the use of toxin proteins from toxin–antitoxin system(s) (TAS) as potential targets for development of new antimicrobials (Alonso et al. 2007; Mutschler and Meinhart 2011).

18.2 Pneumococcal Type II TA genes

The number of type II TA genes reported in *S. pneumoniae* was initially relatively low: three copies in non-capsulated strain R6 and five copies in the capsulated and virulent strain TIGR (Pandey and Gerdes 2005; Nieto et al. 2006). However, the discovery of more novel TAs in the past decade and the increase in the number of sequenced pneumococcal strains led to the discovery of more pneumococcal TAs (Leplae et al. 2011). A more recent in-depth study (Chan et al.,2013) showed that pneumococcal TA could go up to 10 copies in certain strains (Table 18.1) and this number is perhaps still an underestimation). To date, only three chromosomally encoded pneumococcal TA have been proven to be functional and studied in some detail, namely *relBE2* (Nieto et al. 2006), *pezAT* (Khoo et al. 2007) and *yefM-yoeB* (Nieto et al. 2007; Chan et al. 2011). The *pezAT* TAS is reviewed comprehensively in Chap. 12, and therefore only *relBE2* and *yefM-yoeB* will be discussed in this chapter, in terms of the genetic organisation of the operons, transcriptional regulation, functional activities and, finally, the available details on their structures.

18.2.1 relBE2

The *relBE* family is one of the best studied TA (Gotfredsen and Gerdes 1998; Gronlund and Gerdes 1999; Christensen and Gerdes 2004; Cherny et al. 2007; Overgaard et al. 2008; Hurley et al. 2011) and analyses performed on the genome of *S. pneumoniae* indicated the presence of two *relBE* homologues, which were renamed as *relBE1* and *relBE2* (Pandey and Gerdes 2005). The functionality of *relBE1* had been previously precluded in the absence of any mRNA cleavage pattern following induction of *relE1* transcription in *Escherichia coli*, as inhibition of translation by mRNA cleavage is one of the well defined characteristic of the RelE family (Christensen and Gerdes 2003). Sequence comparison of the putative RelE1 pneumococcal toxin with the *E. coli* RelE showed that the amino acids that were likely involved in the toxic activity were absent in RelE1, which could explain why overexpression of the pneumococcal *relE1* gene in *E. coli* did not reveal any RNase activity (Christensen and Gerdes 2003). On the other hand, expression of pneumococcal *relE2* yielded very similar cleavage pattern as *E. coli* RelE, which was at the tmRNA stop and UUA codons (Christensen and Gerdes 2003). In a growth profile assay, expression of *relE2* led to cell growth arrest in a pneumococcal wild-type strain and the toxicity was more severe when the chromosomal *relB2* was inactivated. Moreover, in a heterologous Gram-negative host *E. coli*, cell growth arrest caused by the expression of *relE2* could be rescued by expression of its cognate antitoxin *relB2*. However, prolonged exposure to *relE2* induction will lead to most of the cells being unable to form colonies, indicating that the antitoxin salvage is only effectual within a window period of time (Nieto et al. 2006). These observations contrasted with the *E. coli* RelBE in its natural host, where the rescue mission by the *E. coli* RelB antitoxin was still effective after long periods of exposure to the *E. coli*

Table 18.1 TA homologues in *S. pneumoniae* strains

S. pneumoniae strains	HicBA	PhD-Doc	RelBE1	RelBE2	HigBA	YefM-YoeB	PezAT	Xre-COG2856AC	Xre-COG2856B	Bro1-Xre	Xre-Bro2	Total
AP200	1	1	1	1		1		1			1	7
70585	1	1	1	1			1	1	1		1	8
670-6B	1	1	1	1		1	1	1			1	8
ATCC 700669	1	1	1	2			2	1	2			10
BS397	1	1	1	1		1	1	1	1		1	9
BS455	1	1	1	1		1	1	1	1		1	9
BS457	1	1	1	1		1	1	1	1		1	9
BS458	1	1	1	1		1	1	1	1		1	9
CCRI 1974	1	1	1	1				1				5
CCRI 1974M2	1	1	1	1				1				5
CDC0288-04	1		1	1		1	1	1				6
CDC1087-00	1	1	1	1		1		1				6
CDC1873-00	1	1	1	2		1	1	1	1		1	10
CDC3059-06	1	1	1	2			1	1	2		1	10
CGSP14	1	1	1	1		1	2	1				8
D39	1		1	1		1	1	1				6
G54	1	1	1	1				1				5
GA04375	1	1		1				1				4
GA17545	1	1	1	1				1				5
GA17570	1	1	1	1		1	1		1			7
GA41301	1	1	1	2		1	1	1	1			9
GA41317	1		1	1		1		1	1		1	7
GA47368	1	1	1	2		1	1	1	1		1	10
GA47901	1	1	1	1				1				5
Hungary19A-6	1	1	1	1		1	1	1	1	1	1	10

(continued)

Table 18.1 (continued)

S. pneumoniae strains	HicBA	PhD-Doc	RelBE1	RelBE2	HigBA	YefM-YoeB	PezAT	Xre-COG2856AC	Xre-COG2856B	Bro1-Xre	Xre-Bro2	Total
INV104	1	1	1	1		1		1				6
INV200	1	1	1	1		1	1	1				7
JJA	1	1	1	2		1	1	1	1			9
MLV-016	1	1	1	1		1		1				6
OXC141	1	1	1	1			1	1	1		1	8
P1031	1	1	1	2			2	1	1		1	10
R6	1		1	1		1	1	1				6
SP11-BS70	1	1	1	1		1		1	1		1	8
SP14-BS292	1	1	1	1		1	1	1	1		1	9
SP14-BS69	1	1	1	1			1	1	1		1	8
SP18-BS74	1	1	1	1				1	1	1		7
SP195	1	1	1	1		1	1		2		1	9
SP19-BS75	1	1	1	1		1	1	1				7
SP23-BS72	1		1	1		1	1	1				6
SP3-BS71	1	1	1	1			1	1	1		1	8
SP6-BS73	1	1	1	1			1	1	1			7
SP9-BS68	1	1	1	1		1	1		1			7
SP-BS293	1	1	1	1		1	1	1	1		1	9
str. Canada MDR_19A	1	1	1	1				1				5
str. Canada MDR_19F	1	1	1	1				1				5
Taiwan19F-14	1	1	1	1				1				5
TCH8431/19A	1	1	1	1				1				5
TIGR4	1	1	1	1	1	1	1	1				8
Total	48	43	47	55	1	29	34	45	28	2	20	352

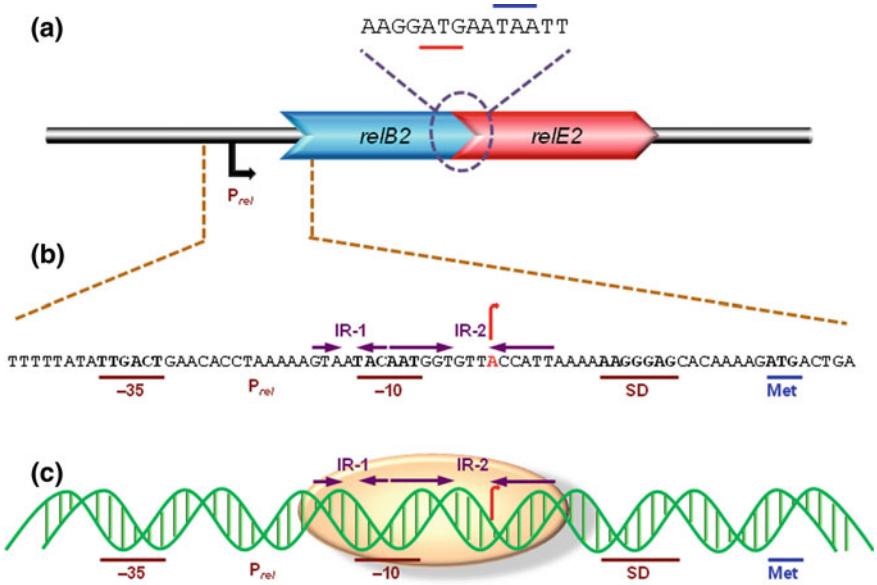


Fig. 18.1 Molecular characteristics of pneumococcal *relBE2* TA. **a** The *relBE2* locus is organised as an operon with the *relB2* antitoxin gene preceding the *relE2* toxin gene. Both genes overlap by 8 nt, and are co-transcribed from a promoter, P_{rel}, located upstream of *relB2*. The *blue line* indicates the stop codon of *relB2*, whereas the *red line* depicts the start codon of *relE2*. **b** Promoter region of *relBE2* operon. Two IRs were identified within the promoter region of *relBE2*, with both IRs partially overlapping with the −10 sequence of P_{rel}. The transcription start site is highlighted (the 'A' residue in *red*) and also represented by a *red arrow*. Shine-Dalgarno (*SD*) sequences and the start codon (Met) of *relB2* are indicated. **c** Schematic drawing of DNA helices of the *relBE2* promoter region. The possible binding site of RelB2 or RelBE2 protein complex (indicated by beige oval) is indicated (Moreno-Córdoba et al. 2012)

RelE toxin (Pedersen et al. 2002). Anyhow, all these results are evident that the pneumococcal *relBE2* is a *bona fide* broad-range spectrum TAS, likely with different level of toxicity in different hosts.

As in other TAS, the pneumococcal *relB2* antitoxin gene precedes its cognate *relE2* toxin gene, and both genes overlap by eight nucleotides (nt) (Fig. 18.1a). The TA genes are co-transcribed as a single mRNA from a single promoter, termed P_{rel}, identified upstream of *relB2* and the structure of the operon is such that is suggestive of the occurrence of translation coupling (Nieto et al. 2006). Two short inverted repeats (IR-1 and IR-2) were identified within the P_{rel} promoter region, spanning from nt −16 to +6 (with +1, the transcription start site). The right arm of IR-1 and part of the left arm of IR-2 overlap the −10 region promoter P_{rel} (Fig. 18.1b, c), and provide a potential binding site for repressor proteins. Indeed, electrophoretic mobility shift assays demonstrated that RelB2 antitoxin and RelB2-RelE2 TA protein complex were bound to a DNA fragment containing this region, hinting at the existence of an operator for this operon, and autoregulated by its own proteins, but this has yet to be firmly established (Moreno-Córdoba et al. 2012).

Fig. 18.2 Sequences alignment and ribbon-helix-helix motifs. **a** Pneumococcal RelB2, which is also a repressor of its own transcription, has a ribbon-helix-helix motif that is similar to the well-studied CopG repressor protein (Gomis-Ruth et al. 1998). **b** Multiple sequence alignment of YefM homologues of four different bacteria (YefMSpn: *S. pneumoniae*; YefMEco: *E. coli*; YefMMtu: *M. tuberculosis* and AxeEfa: *E. faecium*) and their secondary structures showed that although the sequences were not well conserved among each other, all of them seem to possess the ribbon-helix-helix motifs with similar length, hinting at the importance of this motif. Residues that would participate in beta sheets are highlighted in *green*, and residues participating in the formation of alpha helices are represented by *orange* and the rest of the residues in *black* would be random coils

Analytical ultracentrifuge analyses indicated that the RelB2 antitoxin preferentially formed dimers in solution, whereas the RelB2-RelE2 TA protein complex, in its vast majority, behaved as a heterohexamer $(RelB2_2\text{-}RelE2)_2$, even though other combinations ($RelB2_2$-RelE2) and $(RelB2_2\text{-}RelE2)_4$ were also observed in smaller quantities (Moreno-Córdoba et al. 2012). To assess the stoichiometry of the proteins when they bind to the DNA fragment that encompassed the IRs, native mass spectrometry experiments were conducted in addition to analytical ultracentrifuge analyses. The RelB2 dimer was found to bind weakly to the DNA target, whereas the binding was a lot stronger when the RelB2-RelE2 TA protein complex was included, with the molecular complex that could be assigned to a heterohexamer $(RelB2_2\text{-}RelE2)_2$ (Moreno-Córdoba et al. 2012). These observations were somewhat dissimilar to the two archaeal (*Pyrococcus horikoshii* and *Methanococcus jannaschii*) RelBE TA protein complexes, where heterotetramers $(RelB\text{-}RelE)_2$ were observed instead (Takagi et al. 2005; Francuski and Saenger 2009). On the other hand, two different observations were found for *E. coli* RelBE: a heterotetramer $(RelB\text{-}RelE)_2$ (Li et al. 2008) or a heterotrimer in solution ($RelB_2$-RelE) (Overgaard et al. 2008); further structural data might solve this apparent discrepancy.

Circular dichroism analyses showed that the pneumococcal RelB2 antitoxin has a relatively high average of secondary structures (34 % of α-helices), demonstrating the protein is partially folded as its *E. coli* counterpart. The overall α-helices increased to 40 % during the formation of a TA protein complex with its cognate RelE2 (Cherny et al. 2007; Moreno-Córdoba et al. 2012). Molecular modelling suggested that the N-terminus of each pneumococcal RelB2 monomer could have a more complete folding upon dimerisation. The RelB2 monomer has a ribbon-helix-helix motif, that is similar to the well-studied Arc-CopG-Omega family of transcriptional repressors (Fig. 18.2a) (Pabo and Sauer 1992; Gomis-

Ruth et al. 1998; del Solar et al. 2002; Murayama et al. 2001), and when dimerisation occurred, the RelB2 dimer could have been bound to the DNA target in a manner similar to CopG, that is through an antiparallel beta sheet. Nevertheless, the 28 N-terminus amino acid residues of pneumococcal RelB2 shares 64 % similarity with the 45-residue CopG encoded by the streptococcal plasmid pMV158 (Fig. 18.2a) (del Solar et al. 2002; Moreno-Córdoba et al. 2012). Combining all the analyses that we have obtained so far, we propose that the RelB2 dimer serves as a repressor that binds weakly to the operator site (the IRs which overlap with the −10 sequence of the promoter P_{rel}) of the *relBE2* operon, to repress its own transcription; whereas RelE2 serves as a co-repressor, bridging two RelB2 dimers to form a heterohexamer $(RelB2_2\text{-}RelE2)_2$ complex, which then binds more avidly to the operator to further repress the transcription of the *relBE2* TA operon (Fig. 18.1c).

18.2.2 yefM-yoeB*: BOX Makes the Difference*

The pneumococcal *yefM-yoeB* TAS was discovered due to its sequence similarity with the *E. coli* counterpart (Pandey and Gerdes 2005; Nieto et al. 2007). The functionality of this TAS had been shown in *E. coli*: overproduction of the pneumococcal YoeB toxin led to cell growth arrest and reduction of culturable cells, with the toxicity more prominent in B strains than K-12 strains (Nieto et al. 2007). The cells could then be resuscitated by production of pneumococcal YefM antitoxin to the level dependent on the YoeB toxin exposure period: the number of colony-forming units was reduced to 60 % after 2 h and 4 % after 4 h of YoeB production, compared to a viability of only 0.4 % without YefM production, indicating that there was a time frame for YefM rescue as was observed for the pneumococcal *relBE2* TA (Nieto et al. 2007). Interestingly, the production of YoeB in a Lon-deficient strain led to 85 % of cells being able to form colonies after 1 h of YoeB overexpression, as compared to only 25 % of the colonies formed by the wild-type strain (Nieto et al. 2007). This observation signified that Lon protease could play a role in YoeB-mediated toxicity. In *E. coli*, Lon-mediated lethality was found to be partially dependent on the *E. coli yefM-yoeB*, but not other Lon substrates like MazE and RelB antitoxins. Nonetheless, there was another pathway which led to Lon-dependent lethality, as deletion of the *E. coli yefM-yoeB* only delayed lethality but not fully suppressed it (Christensen et al. 2004). The target of pneumococcal YoeB has not been determined so far, but the *E. coli* YoeB was shown to inhibit translation initiation by causing cleavage of translating mRNAs (Christensen et al. 2004; Zhang and Inouye 2009). *E. coli* YoeB alone does not have RNase activity but it was found to be associated with the 50S ribosomal subunit in 70S ribosomes and interacted with the ribosomal A site leading to cleavage of the mRNA (Zhang and Inouye 2009). It is also worth to mention that overproduction of Lon could also induce *E. coli* YoeB-dependent cleavage of translating mRNAs (Christensen et al. 2004).

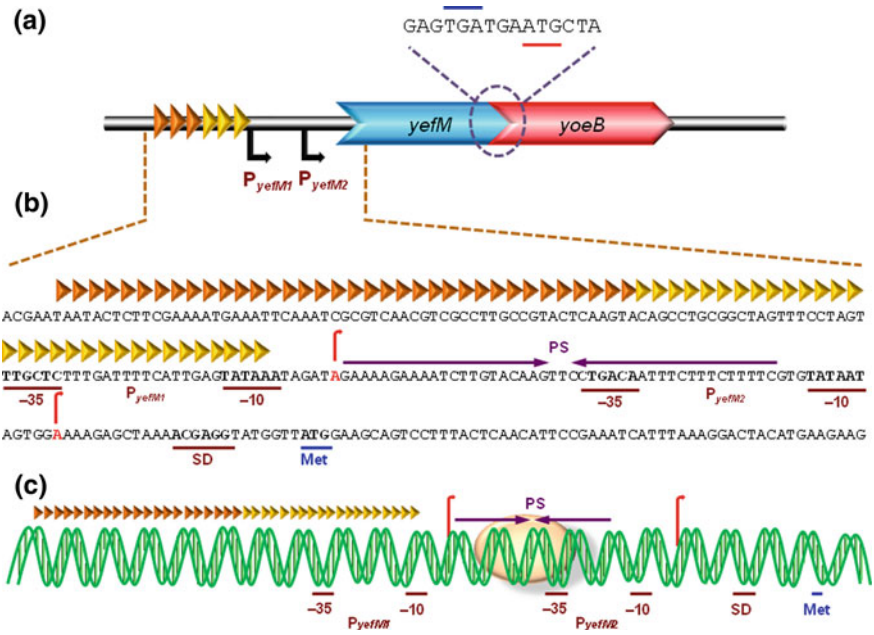


Fig. 18.3 Molecular characteristics of pneumococcal YefM-YoeB TA. **a** Like *relBE2*, the *yefM* antitoxin gene is located upstream of the *yoeB* toxin gene. Both genes are separated by 3 nt, and they are co-transcribed from promoter(s) upstream of *yefM*. **b** Two promoters were identified, the 'native' promoter P_{yefM2}, and an additional upstream P_{yefM1} promoter. A pneumococcal mobile repeat sequence, designated boxAC element (boxA indicated by *orange triangle* and boxC depicted by *yellow triangle*), was identified further upstream, with the 3′-portion of boxC overlapping with the −35 sequence of P_{yefM1} promoter by 5 bp, indicating that the insertion of the boxAC element had created the additional P_{yefM1} promoter. Two transcription start sites, (indicated by *red arrows*) as determined by primer extension, had confirmed the functionality of these two promoters. Moreover, a palindrome sequence (*PS*) was found that overlapped with the −35 sequence of P_{yefM2} promoter, and served as the binding site of the YefM or YefM-YoeB protein complex (beige oval), as shown in the schematic drawing of DNA helices in **c**. Shine-Dalgarno sequences (*SD*) and the start codon (Met) of *yefM* were also indicated

Like *relBE2*, the pneumococcal *yefM-yoeB* also possesses typical characteristics of TAS: (1) *yefM* antitoxin gene is located upstream of *yoeB* toxin gene (Fig. 18.3a) and (2) both genes are co-transcribed as a single mRNA transcript even though both genes do not overlap but they are separated by three nt, i.e., a stop codon TGA (Fig. 18.3a) (Chan et al. 2011). However, the regulation of this TAS is atypical due to insertion of a piece of relatively short DNA sequences upstream of the promoter P_{yefM2} of this TA operon. Without this piece of DNA, the regulation is like the norm, where the YefM antitoxin binds loosely to the operator, a palindrome sequence that overlaps with the −35 consensus of promoter P_{yefM2}, and the YefM-YoeB TA protein complex binds tightly to its cognate region to further suppress its transcription (Fig. 18.3b, c) (Chan et al. 2011). The piece of DNA that changes the regulation norm of this TAS is called a BOX element. BOX elements are supposedly mobile repeat

sequences that are found exclusively in *S. pneumoniae* and its closely related species. BOX elements are abundant and distributed, apparently in a random fashion, in the intergenic spaces in the genomes of pneumococci (127 copies in strain TIGR4 (Mrazek et al. 2002) and 115 copies in strain R6 (Hoskins et al. 2001)). Each BOX element usually consists of three different modules designated boxA, boxB and boxC, with approximately 50 bp for each module (Martin et al. 1992). BoxB resides in between boxA and boxC, and can be absent in some BOX elements or can be found up to eight copies in others. BOX elements which contain a boxA and a boxC module have the potential to form a stable stem-loop structure which could affect the expression of neighbouring genes (Knutsen et al. 2006). The sequence of the BOX elements can differ from the consensus, but the secondary structures remain the same in most, if not all cases (Martin et al. 1992). Some studies have shown that the BOX elements enhanced gene expression by stabilizing mRNA (the BOX element was co-transcribed along with the downstream gene) via increasing its half life, or serving as binding sites for regulatory proteins (Knutsen et al. 2006; Croucher et al. 2009). They were also suggested to be involved in the regulation of virulence, genetic competence (Knutsen et al. 2006) as well as phase variation (Saluja and Weiser 1995).

In the case of the pneumococcal *yefM-yoeB* TAS, the incorporation of the BOX element (which was composed of just a boxA and a boxC subelements and was therefore termed as boxAC element) upstream of promoter P_{yefM2}, had added an additional functional promoter P_{yefM1}. Transcriptional fusion assays in *E. coli* showed that promoter P_{yefM1} was 15-fold weaker than promoter P_{yefM2}, and was not regulated by the YefM or YefM-YoeB protein complex (i.e. the proteins did not bind there), as observed in footprinting assays (Fig. 18.4a) (Chan et al. 2011). Transcriptional fusions also indicated that the overall activity with the presence of both promoters had been reduced to half of the strength of the 'natural' promoter P_{yefM2} alone. Moreover, only slight repression was evident when YefM antitoxin or YefM-YoeB TA complex was provided in *trans* (Fig. 18.4a). It seemed like the boxAC element had thwarted the repression perhaps by forming a stem loop secondary structure together with the P_{yefM2} promoter in the DNA, as predicted via DNA secondary structure prediction program mfold (Fig. 18.4b) (Chan et al. 2011). However, strikingly, when *yefM* antitoxin gene was present in *cis* along with both promoters, a prominent de-repression was observed, perhaps by destabilising the predicted stem-loop structure. More surprisingly, the promoter activity was even higher than the promoter P_{yefM2} alone, indicating the promoter activity had been activated (note: no functional promoter was found within *yefM* reading frame) (Fig. 18.4a). The presence of *yoeB* toxin gene along with the two promoters and *yefM* antitoxin gene somehow neutralised this activation. It has to be emphasised that the activation was not due to the YefM or YoeB proteins, as when the translation of both proteins was abolished, the activation still persisted (Fig. 18.4a) (Chan et al. 2011). We suggested that this activation phenomenon could be due to either an unknown DNA *cis*-acting element or the boxAC sequences providing a binding site for host binding factors that have yet to be identified. The boxAC element is highly conserved in all the sequenced pneumococcal strains that harboured intact *yefM-yoeB* locus, but absent in other *yefM-yoeB* homologues of different bacteria. The conservation of the

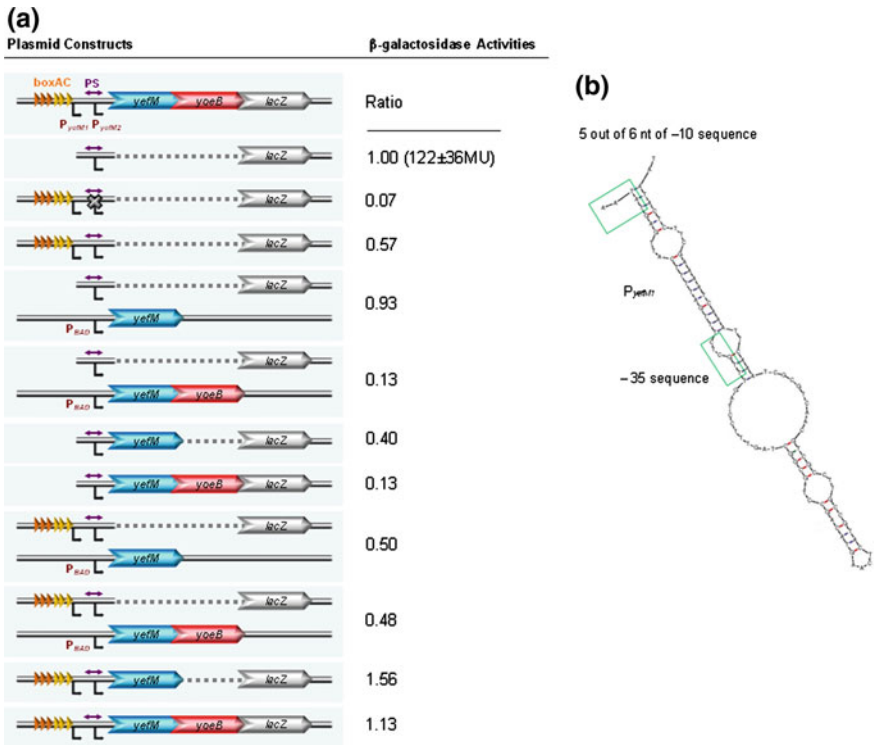


Fig. 18.4 *β*-galactosidase activities and the stem-loop structure of boxAC element. **a** Various pneumococcal *yefM-yoeB* genes and their promoter region were cloned upstream of a promoterless *lacZ* plasmid, pQF52, in order to measure the promoter activities. In certain cases, the *yefM* or *yefM-yoeB* genes were constructed downstream of a pBAD promoter, inducible by the addition of arabinose, in pLNBAD plasmid, to produce the proteins in *trans*. All the constructs were transformed into *E. coli* and subjected to *β*-galactosidase assays (Chan et al. 2011). All the readings were normalised to the construct with just the P_{yefM2} promoter alone (i.e., the 'native' promoter for *yefM-yoeB*), which gave a reading of 122 ± 36 Miller Units (*MU*) or a ratio of 1. The P_{yefM1} promoter was 15-fold weaker than P_{yefM2}, which was ratio of 0.07. Nonetheless, when both promoters were present, the overall promoter activity had a ratio of 0.57, which was lower than the activity of the P_{yefM2} promoter alone. With the construct of just the P_{yefM2} promoter alone, and when YefM protein was provided in *trans*, the promoter activity had slightly reduced to ratio of 0.93, and a more prominent reduction (ratio of 0.13) was observed when both the YefM-YoeB protein complex was produced in *trans*. Whereas in *cis*, a more prominent decrease was observed when the *yefM* gene was present along with the promoter P_{yefM2} (ratio of 0.40), and the promoter activity was further reduced to ratio of 0.13 in the presence of both *yefM* and *yoeB* genes. On the other hand, in the presence of both P_{yefM1} and P_{yefM2} promoters, when YefM and YefM-YoeB protein complex were provided in *trans*, no prominent or only very slight repression (ratio of 0.50 and 0.48, respectively) was observed, compared to the activity of both promoters without the proteins (ratio of 0.57). Strikingly, activation instead of repression was evident when *yefM* (ratio of 1.56) and *yefM-yoeB* (ratio of 1.13) genes were present in *cis* along with the promoters. The activation still persisted even when the translation of both proteins was abolished (data not shown). **b** The boxAC element tends to form a stem-loop structure, which could possibly affect the transcription efficiency of the overall promoter activity as a whole. The boxAC element might also provide a binding site for a hitherto-unidentified host factor(s), which led to the activation as shown in (a)

boxAC-*yefM-yoeB* arrangement in the pneumococcal genomes indicated that the insertion of boxAC element was not recent and is likely beneficial to its host. By increasing the overall basal transcription level of the *yefM-yoeB* locus, the boxAC element could possibly enable the pneumococcal cells to better adapt and survive unfavourable conditions (see Sect. 18.5.2).

The crystal structure of the pneumococcal YefM-YoeB has yet to be resolved; nevertheless, we may get some hints from the resolved structures from *E. coli* and *Mycobacterium tuberculosis*. The crystal structure of *E. coli* YefM-YoeB was a heterotrimer which consisted of two *E. coli* YefM monomers and one *E. coli* YoeB monomer. The N-terminal regions of each of the two *E. coli* YefM monomers bound to each other and formed a symmetric dimer; one C-terminal region of the *E. coli* YefM monomer was disordered, whereas another C-terminal end of the *E. coli* YefM monomer formed a helix and was bound to the YoeB monomer (the RNase active site, and thus suppressed the toxicity of YoeB), causing the orientation of the structure to be unparalleled, and therefore sterically hindered the binding of a second *E. coli* YoeB monomer to the C-terminal region of the other *E. coli* YefM monomer (Kamada and Hanaoka 2005). On the other hand, crystallographic and biophysical studies showed that the *M. tuberculosis* YoeB-freeYefM had a well-defined tetrameric structure rather than an intrinsically unfolded protein. When the *M. tuberculosis* YefM and *E. coli* YefM-YoeB structures were compared, it appeared that the tetramer interface of *M. tuberculosis* YefM mimicked the *E. coli* YefM-YoeB interface, with the *E. coli* YoeB replaced by the *M. tuberculosis* YefM dimer (Kumar et al. 2008). The presence of a ribbon-helix-helix motif was observed in YefM among various bacterial species (Fig. 18.2b). The N-terminal region of YefM was more conserved, whereas the C-terminal end was highly pliable, which perhaps adopted different conformations in different monomers. This could explain why the *M. tuberculosis* YefM was able to counteract the toxicity of *E. coli* YoeB (Kumar et al. 2008). In another example, the *E. coli* YefM-YoeB and its homologue (namely Axe-Txe) in *Enterococcus faecium* could also complement each other. However, this was not the case for the pneumococcal YefM-YoeB and *E. coli* YefM-YoeB (Nieto et al. 2007). Based on molecular modelling, a positively charged moiety defined by an R72 residue in *E. coli* YefM which was responsible for the interaction with the negatively charged pocket on the interface of the *E. coli* YoeB was observed in Axe but not in pneumococcal YefM, thus inferring the possible reason for the failure of pneumococcal YefM to neutralise the *E. coli* YoeB toxin. The pneumococcal YoeB structural model also showed a larger exposed surface than the *E. coli* YoeB and Txe toxins, which perhaps increased the potential interaction surface that led to the ineffectiveness of *E. coli* YefM to counteract the pneumococcal toxin (Nieto et al. 2007).

18.3 Other Pneumococcal TAS

We had done an exhaustive bioinfomatic search of TAS in the genome of 48 sequenced *S. pneumoniae* strains that were available in the NCBI database (Chan et al.,2013). The search was based on the expanded and combined data from

previous reports (Pandey and Gerdes 2005; Makarova et al. 2009; Shao et al. 2010). The TAS that were included were CcdAB, HicAB, HipAB, MazEF, PemIK, ParDE, Phd-Doc, RelBE, HigBA, YefM-YoeB, VapBC, MosAT, YeeUV, PezAT, Xre-COG2856 and Xre-Bro; and only complete TA pairs were included. However, only HicAB, PhD-Doc, RelBE, HigBA, YefM-YoeB, PezAT, Xre-COG2856 and Xre-Bro were identified in the genomes of *S. pneumoniae*, whereas the highly abundant bacterial and archaeal TAS like MazEF and VapBC were absent (Table 18.1). The total number of TAS ranged from 4 to 10 for each individual strain, with the average number being 7. Of all the TAS, HicBA was found as one copy consistently in all the 48 strains, whereas the functional RelBE2 could be found either as one or two copies in all the strains. PhD-Doc, RelBE1 and Xre-COG2856CA were present in nearly all the strains. The other two functional TAS, PezAT and YefM-YoeB, were found in more than half of the strains examined. HigBA was present only as one copy in one strain, the TIGR4 (Table 18.1). We have chosen a few TAS that are interesting for further discussion. It has to be emphasised here again that except for RelBE2, PezAT and YefM-YoeB, the rest of the pneumococcal TAS mentioned here are putative with no experimental information available so far.

18.3.1 Genetic Organisation of the Known Orthologs: RelBE1 and RelBE2

As mentioned earlier, there were two RelBE homologues identified in the genome of *S. pneumoniae*: RelBE1 and RelBE2. The pneumococcal RelBE1 was present in all the strains except GA04375 (Table 18.1). In addition, the RelBE1 sequences were highly conserved among all the strains. These observations made us rethink our earlier inference that RelBE1 was not functional as mentioned in Sect. 18.2.1, as the assessment of the cell growth profile after overexpression of the *relE*1 gene was never conducted (Christensen and Gerdes 2003; Nieto et al. 2006).

Conversely, RelBE2 is a functional TAS in both *S. pneumoniae* and *E. coli* (Nieto et al. 2006). The *relBE2* TAS is present in all the 48 strains analysed here, and two copies were found in seven strains (Table 18.1). A recent comprehensive study using polymerase chain reaction had shown that pneumococcal *relBE2* was present in all the 100 strains tested (70 from Spanish and Polish clinical isolates, and 30 from database-available genomes), which was in agreement with our in silico results, whereas the other two functional pneumococcal TAS *yefM-yoeB* and *pezAT* were not so much conserved (Khoo et al. 2007; Nieto et al. 2010). Interestingly, rearrangements were detected around the chromosomal genes that were adjacent to *relBE2*, and these rearrangements were classified into three groups (Fig. 18.5). The first group displayed a gene organisation where the *relBE2* operon was located in between a metal-dependent hydrolase gene (*vicX*) upstream, and a putative type II restriction-modification system downstream, followed by the pneumococcal repeats Repeat Unit of Pneumococcus (RUP) (Fig. 18.5a)

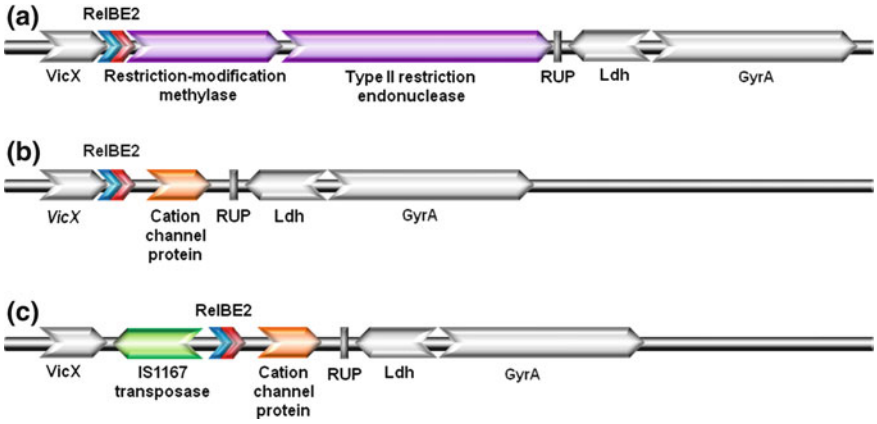


Fig. 18.5 Polymorphisms in the genetic organisation of the pneumococcal *relBE2* locus and its neighbouring genes found in clinical isolates of *S. pneumoniae*. In general, *relBE2* is flanked by *vicX* (metal-dependent hydrolase) at its upstream and a RUP element, *ldh* (lactate dehydrogenase) (in opposite orientation), as well as *gyrA* (the A subunit of DNA gyrase) at its downstream. Three different organisations were observed among all the strains analysed (70 clinical isolates and 30 strains with sequenced genomes in the database; see Nieto et al., 2010): (**a**) genes encoding a putative restriction-modification methylase family protein and a putative type II restriction endonuclease were found immediately downstream of *relE2*; (**b**) a gene encoding a cation channel protein was identified immediately downstream of *relE2* instead; and (**c**) same as in (**b**) but an IS1167 element had inserted upstream of *relB2* in the opposite orientation

(Croucher et al. 2011). The second group was similar to the first group, but the putative type II restriction-modification system genes were replaced by a cation channel protein (Fig. 18.5b). The third group had the gene organisation similar to the second group, but with an insertion of an IS1167 element organised in a divergent orientation upstream of the *relBE2* operon (Fig. 18.5c). The polymorphisms observed in these clinical pneumococcal isolates implied that these strains, which have been subjected to the high antibiotic selection pressure during infection treatment, might have accelerated the evolution process. The advantage of conserving *relBE2* in the genome of *S. pneumoniae* is currently unknown, but it could serve as a potential candidate for the development of novel antimicrobials due to its apparent ubiquity in all pneumococcal isolates.

18.3.2 The Uncommon TAS (1): Xre-COG2856

A computational study done by Makarova et al. (2009) revealed some potential novel TAS, and among one of them found in the pneumococcal genome was Xre-COG2856 (Makarova et al. 2009). The Clusters of Orthologous Gene (COG2856) putative toxin belongs to the family of metzincin Zn-dependent

protease (Stocker and Bode 1995; Makarova et al. 2009), and it is usually associated with the helix-turn-helix domain-containing protein of the Xre family (the putative antitoxin and also a transcription repressor). In our bioinfomatic study of the 48 *S. pneumoniae* strains, we had subdivided this putative toxin COG2856 into three groups due to low sequence similarities among each other, namely COG2856A, COG2856B and COG2856C. The similarities among them were 42 % for COG2856A and COG2856B; 30 % for COG2856A and COG2856C; and 26 % for COG2856B and COG2856C. The COG2856A and COG2856C genes were found in almost all the strains (45 out of 48 strains) with COG2856C preceding COG2856A. Both putative toxin genes shared a common antitoxin *xre* gene that overlaps with the 5'-end of the COG2856C gene (Fig. 18.6a). Therefore, we designated this putative TA pair as Xre-COG2856CA. A few different annotations were found among the strains for this putative TAS even though the sequences of these three genes were highly conserved. In most cases (19 out of 45), *xre* and COG2856C genes overlap by 14 nt, whereas in strain R6, both genes overlap by 50 nt, and for the rest of the strains, COG2856C (22 out of 45) or both *xre*-COG2856C (3 out of 45) genes were not annotated. Genes overlapping by 14 bp or 50 bp are somewhat uncommon for TAS as in general, the TA genes overlap by 1 or 4 nt (Pandey and Gerdes 2005), so that both genes are co-transcribed and the proteins are also co-expressed by translational coupling (Engelberg-Kulka and Glaser 1999). For the putative toxin genes, COG2856C is either 72 nt (12 out of 45 strains) or 50 nt (6 out of 45 strains) apart from the gene of COG2856A, and there were also cases where either COG2856C (22 out of 45) or both COG2856C-COG2856A (3 out of 45) was not annotated. Besides, there is a strain (AP200) in which both genes were annotated as 157 nt apart. Analysis of the intergenic region sequences led us to identify a possible promoter upstream of the gene of COG2856A. However, the distance between the putative Shine-Dalgarno sequence and the possible start codon TTG is 18 nt, which is perhaps a bit too far from the norm, and thus if the gene of COG2856A is expressed, the translation could be less efficient. Surprisingly, in strain SP6-BS73, there is a point mutation at the COG2856C stop codon **T**AA → **C**AA, which leads to both COG2856C and COG2856A genes combined as a single open reading frame (ORF) that could potentially encode a 298-amino acid residues protein. In normal conditions, the antitoxin protein should remain in excess to counteract the toxin protein. Having two toxins sharing a common antitoxin is not optimal unless the transcription or translation of the toxins is weak, or the stoichiometry of the TA complex is two (or more) toxins to one antitoxin. The stoichiometry of the TA protein production is important to the fate of the cells, to maintain growth, stay dormant or die in certain cases.

On the other hand, the *xre*-COG2856B putative TA pair is less prevalent as it is present in only 25 out of 48 strains. In most of the strains, it is present as one copy, but in strains ATCC 700669, CDC3059-06 and SP195, two copies of these genes were found (Table 18.1). The gene sequences of *xre*-COG2856B were not as conserved as seen in *xre*-COG2856CA, but the interesting part about this putative TA pair is the gene arrangement, which could be grouped into four categories

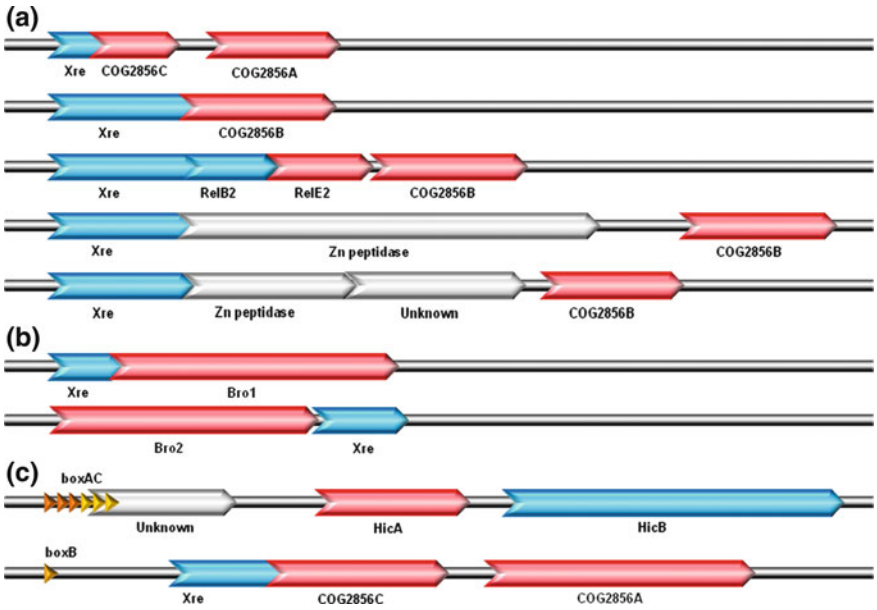


Fig. 18.6 The genetic organisations of the pneumococcal putative TAS and pneumococcal TAS that are associated with the BOX elements. **a** The putative Xre-GOG2856 TAS. Three homologues of COG2856 putative toxins were found, namely COG2856A, COG2856B and COG2856C. In all the cases, the COG2856C gene was located upstream of COG2856A gene and both putative toxin genes shared a common putative antitoxin *xre* gene. As for the COG2856B putative toxin, four different gene organisations were found: (1) In most cases, the *xre* gene was located upstream of the COG2856B gene; (2) *relBE2* was located in between *xre* and COG2856 genes; (3) A gene encoding a putative Zn peptidase was located between the *xre* and COG2856 genes; (4) A gene encoding a putative Zn peptidase, followed by an unknown ORF were located between the *xre* and COG2856 genes. **b** The putative Xre-Bro TAS. Two homologues of Bros were observed and the gene organisations of this family were categorised into two groups: (1) The *xre* putative antitoxin gene precedes the Bro1 putative toxin gene; and (2) the Bro2 putative toxin gene precedes the *xre* putative antitoxin gene instead. **c** Pneumococcal TAS with BOX elements. Besides the pneumococcal *yefM-yoeB* TAS, a boxAC element was also found possibly associated with another TAS, *hicAB*. However, there was an unknown ORF in between the boxAC element and *hicA* putative toxin, and this ORF overlapped with the 3′-portion of boxAC element. In another case, a small boxB element was identified upstream of *xre* in *xre*-COG2856CA putative TAS in strain R6

(Fig. 18.6a). First, which was also the most common one and found in most strains, the putative *xre* antitoxin precedes the COG2856B gene and both genes were either 12 nt (14 out of 28 cases) or 21 nt (3 out of 28 cases) apart. Second, *xre* and COG2856B genes were found flanking the *relBE2* TA genes in the same orientation, in 7 out of 28 cases; both *xre* and *relB2* or *relB2* and *relE2* were overlapped by 1 nt respectively, whereas *relE2* is 12 nt apart from COG2856B gene. This operon-like organisation gives an indication that both TAS could be interrelated. Interactions between TAS have been previously reported, e.g. over-production of the Doc toxin (from the PhD-Doc TAS) has been shown to activate RelE-mediated mRNA cleavage in *E. coli* (Garcia-Pino et al. 2008). In the third

category, a gene that encodes a Zn^{2+}-dependent peptidase-like protein appeared to be located between *xre* and COG2856B (3 out of 28 cases) genes. Lastly, a gene that encodes a Zn^{2+} peptidase-like protein and an ORF of unknown function were located between *xre* and COG2856B (1 out of 28 cases) genes (Fig. 18.6a). The diversity observed in the gene arrangements of these putative TAS are likely due to the highly recombinogenic nature of *S. pneumoniae* and the functionality and role of these TAS as well as the possible interrelations between them are new exciting avenues to be explored.

18.3.3 The Uncommon TAS (2): Xre-Bro

Besides *xre*-COG2856, *xre*-Bro (Baculovirus Repeat ORFs) is another putative TA pair that could be found in the genome of *S. pneumoniae* (Makarova et al. 2009). Two relatively different Bro putative toxin sequences (34 % similarity) were identified, termed Bro1 and Bro2 here, with their respective *xre* putative antitoxin genes. These two *xre* genes shared low sequence similarity with each other, as well as with the *xre* in the *xre*-COG2856AC and *xre*-COG2856B putative TAS. The *xre*-Bro1 genes were only present in two strains (Hungary 19A-6 and SP18-BS74) and its organisation is that of a typical TAS in which the putative *xre* antitoxin gene precedes the Bro1 putative toxin, and both genes are separated by 1 nt (Fig. 18.6b).

Conversely, Bro2-*xre* is present as one copy in 20 strains (Table 18.1). Strikingly, even though they are homologues, the Bro2-*xre* had an opposite orientation to *xre*-Bro1, in which the Bro2 putative toxin gene is located upstream of the *xre* putative antitoxin gene (Fig. 18.6b). Both Bro2 and *xre* genes were annotated to be 11–13 nt apart eight copies for 11 nt, seven copies for 12 nt (three copies for 13 nt) (note: two strains, GA41314 and OXC141, were not annotated). Bro proteins are a family of DNA-binding proteins encoding transcription regulators of DNA viruses that have been also reported to exist in bacterial TAS (Makarova et al. 2009; Iyer et al. 2002). The Xre-Bro is similar to phage repressor–antirepressor proteins that determine the state of the phage by regulating the expression of lytic genes (Heinrich et al. 1995; Makarova et al. 2009). The phage antirepressor protein usually consists of an N-terminal Bro-N and a C-terminal Ant (antirepressor) or KilA-C family domains (Makarova et al. 2009; Iyer et al. 2002). It was postulated that the Xre and Bro might regulate the expression of a third component toxin gene (Heinrich et al. 1995; Makarova et al. 2009).

18.4 BOX Elements

The influence of the BOX subelement, boxAC, on the expression of the *yefM-yoeB* TAS of *S. pneumoniae* and its conservation among pneumococcal strains had been narrated in Sect. 18.2.2 (Chan et al. 2011). Besides *yefM-yoeB*, a boxAC element

was also identified 264 nt upstream of the putative *hicAB* TAS in R6 strain. However, there is an ORF coding an unknown protein of 59 amino acid residues that overlaps with the 3′-portion of the boxAC sequences by 25 nt, and this ORF is also 103 nt apart from the putative *hicA* gene (Fig. 18.6c). No prominent promoter sequence was observed for this ORF, and thus whether this putative gene is transcribed and translated remains to be determined. In another case, a BOX element was also found to be associated with the *xre*-COG2856CA putative TAS. In strain R6, a small boxB (21 nt) was identified 138 nt upstream of *xre* (Fig. 18.6c). However, discrepancies in the database annotations were found which is not surprising due to the lack of consensus in some of these BOX elements. Whether these elements play any role in the expression of the associated TAS as was observed for *yefM-yoeB* awaits future research.

18.5 The Possible Roles of TAS in the Lifestyle of the Pneumococci

Many possible biological functions have been ascribed for TAS over the years and these have been presented in detail elsewhere in this book. Here, we would like to discuss two possible processes that pertain to changes in otherwise homogeneous populations of bacteria but that lead to a separation into sub-populations that are unrelated to genetic changes, as persistence formation and bistability. We would argue in favour of some scarce experiments that could relate either process to the triggering of TAS. The possible role of the PezAT TAS in the virulence of *S. pneumoniae* has been described in Chap. 12.

18.5.1 Persistence

Bacterial persistence seems to be a stochastic process in which individual cells enter into dormancy and are thus protected against harmful substances. Persisters can, then, overcome harsh conditions while enhancing the chances of the population survival (Evdokimov et al. 2009). The persisters are just a fraction of cells within a homogeneous population, and they play a key role in the decision making of bacterial populations (Elowitz et al. 2002). Persisters are genetically identical to the rest of the non-tolerant cells, and they are able to exit dormancy by mechanisms that are not fully understood (Epstein 2009), perhaps by the stochastic appearance of ‘scouts’ which would check the environmental conditions and would decide whether the rest of the population should continue to be dormant (and the scouts would die) or the entire population can resume growth. Persisters are most frequently found in biofilms, the microbial preferred lifestyle, so that the accessibility of harmful compounds is reduced and cells can be safer (Keren et al. 2004).

Chromosomally encoded TAS have been proposed to act as circuits devoted to bacterial response to stress, so that activation of one or more of these TAS would lead to survival of at least part of the bacterial population under conditions of nutrient limitation by inducing dormancy (Christensen et al. 2001; Pedersen et al. 2002; Maisonneuve et al. 2011). In the case of *S. pneumoniae* toxin-activation could contribute to survival of the bacterium within the human host during infection by suppression of cellular growth that could lead to (1) tolerance to antibiotics (Nieto et al. 2010) and (2) escape from the host immune response (Baquero 2009). In this respect, it has been proposed that factors contributing to pneumococcal persistence during colonisation might increase its virulence at later stages (Kadioglu et al. 2008). Mutations that abolished synthesis of the pneumoccocal RelE2 toxin led to a phenotype of erythromycin tolerance, and activation of the *relBE2* operon upon erythromycin challenge induced suppression of protein synthesis, leaving the bacterial cells unable to recover viability, although persistence in pneumococcal biofilms was not tested (Nieto et al. 2010).

Persistence has also been reported to exist in other streptococcal species like *S. epidermidis,* and *S. mutans*, in addition to *S. pneumoniae*. In *S. mutans*, overproduction of the RelBE and MazEF proteins led to 10-fold increase in the number of persisters, although a double deletion mutant in these TAS lacked a phenotype in persister formation (Leung and Lévesque 2012). Screening of an expression library showed the presence of other candidate genes for persistence, including some involved in transcription and replication, sugar and energy metabolism and cell wall synthesis (Leung and Lévesque 2012), pointing to redundancy of genes involved in the persister phenotype. It was also shown that the intra-species *S. mutans* quorum sensing system, the CSP-ComDE regulatory circuit, was directly involved in persistence (Leung and Lévesque 2012). It would thus appear that persistence, just as competence, is a bistable process (below) that reveals a neat way to ensure the survival of part of the population under environmental adverse circumstances, and that TAS may play an important role in it.

18.5.2 Bistability

A bistable behaviour takes place when, under stressful conditions, a bacterial population of genetically identical cells growing in the same conditions undergoes a stochastic separation into two different subpopulations, so that the population may have the choice between two semi-stable, co-existing situations (Dubnau and Losick 2006). Any process leading to bistability is ruled by switching on- and -off of several regulatory feedback mechanisms that would result in benefits for the entire population, although not necessarily for individuals or sub-groups (Rao et al. 2002). Bistability in Gram-positive bacteria has been reported in processes like expression of type I pilus genes, genetic competence, persistence, biofilm formation, phase variation and escape from sporulation (Basset et al. 2012; Claverys and Havarstein 2007; Claverys et al. 2006; Dubnau and Losick 2006; Leung and Lévesque 2012; Saluja and Weiser 1995). Most of the genes participating in

bistable behaviour also are involved in virulence with the exception of sporulation (at least in *Bacillus subtilis*) which is a point of no return process once some stages are passed (Dubnau and Losick 2006). However, it has been suggested that, at least in *B. subtilis*, sporulation could be a '*molecular race*' competing in time, but not ruled by cross-regulatory mechanisms that would occur at the beginning of the process, that is when the decision to sporulate or to escape from it is still reversible (Kuchina et al. 2011). To adapt to a changing environment, the bacterial cells have to optimise their fitness by expressing different sets of genes within a population to improve their chances of survival. This can be achieved by a bet-hedging strategy, which postulates that the individuals must lower their variance in fitness, so that the entire population would benefit as a whole (Olofsson et al. 2009). Bet-hedging contemplates the strategic choices of individual cells within a changing environment, and it implies the stochastic switching-on and -off of a relatively large number of genes to maximise the fitness of the entire population (Cooper and Kaplan 1982). To achieve this, the population would need to combine different bet-hedging strategies, namely conservative ('better safe than sorry'), diversified ('not put all your eggs in one basket'), and adaptive ('flip the coin') (Cooper and Kaplan 1982). The population heterogeneity can incur a fitness cost, since part of the population would always be maladaptive in a given environment. Theoretical studies have indicated, however, that under conditions of environmental fluctuations, stochastic switching led to a net fitness gain for the population (Thattai and van Oudenaarden 2004; Gaál et al. 2010).

An interesting example of bistable behaviour was obtained by performing a global analysis of transcription during onset of competence in pneumococci. The results showed activation of two choline-binding proteins, and a functional role of one of them in competence could be demonstrated (Rimini et al. 2000). In addition, a relationship between competence and autolysis has been demonstrated for pneumococci (Claverys et al. 2006), and a suggestion that at least one of these choline-binding proteins could play a role in autolysis or in changes in the cell wall during competence was made (Rimini et al. 2000). These findings could be used by us to argue in favour of a connection between competence and the pneumococcal PezAT system within bistable behaviour (see Chap. 12).

The discovery of a boxAC element associated with transcription of the pneumococcal *yefM-yoeB* and the location of a boxAC element upstream of the putative *hicAB* and a boxB element upstream of *xre*-COG2856CA TAS made us speculate that these TAS may have a role related to other cell processes. We have preliminary data indicating that the former TA influences, albeit slightly, pneumococcal biofilm formation, a role that has been shown for TAs of *E. coli* (Kim et al. 2009). We would like to discuss briefly the possibility that the role of the BOX element upstream of the pneumococcal *yefM-yoeB* operon might be related to bistability. An early report by the Clavery's laboratory showed that BOX elements are located in the vicinity of competence-related genes like *comA*, *hexB* and possibly *mmsA* genes (Martin et al. 1992). For instance, a BOX element is located 17-bp upstream of the *comA* start codon, indicating that the *comA* promoter is either located within or upstream of the BOX element. Moreover, insertion of heterologous DNA within the BOX element confers a Com$^-$ phenotype

(Chandler and Morrison 1988), which supports the significant influence of the BOX element in *comA* gene expression. Besides competence (a bistable process), BOX elements are also found close to genes like *neuA*, which encodes the pneumococcal virulence factor neuraminidase (Camara et al. 1991), and *ply*, which encodes another virulence factor termed pneumolysin (Walker et al. 1987). The location of these BOX elements in the immediate vicinity of genes related to competence and virulence raises the intriguing possibility of BOX elements involved in coordinating the expression and modulation of these genes depending, perhaps, on the copy numbers of the boxB subelement and its orientation (Knutsen et al. 2006). Competence and virulence in pneumococcus are involved in bistability, and also considered as global responses of the bacteria upon stress, the same as the TAS. Since BOX elements have been associated with: (1) genetic competence (Martin et al. 1992), (2) virulence-related genes (Martin et al. 1992), and (3) phenotypic phase variation (Saluja and Weiser 1995), it would appear that the hyper-recombinogenic nature of *S. pneumoniae* provides the bacteria with selective advantages when subjected to stressful conditions, such as the incorporation of an integrative piece, like BOX, into an operative genome piece, such as *comA*, *neuA*, *ply* and the *yefM-yoeB* operon (Baquero 2004, 2009). This '*molecular attunement*' of the cells to their changing environment may have led to a better fitness of the bacteria, although strains totally devoid of all BOX elements have not been constructed (Croucher et al. 2011). Interestingly, a recent in silico evolutionary simulation showed that under fluctuating environments, the emergence and maintenance of bistability occur only in the presence of noise, and the emergence of bistability enables faster adaptation to the fluctuating environment (Kuwahara and Soyer 2012). Selective forces that lead to higher/lower noise might also enhance/inhibit the evolution of bistability and nonlinearity in gene regulation (Kuwahara and Soyer 2012).

Our final conclusion is the proposal, speculative as it may be at present, that some of the pneumococcal TAS may be involved in (or related to) processes such as competence and persister development, i.e. biofilm formation. These processes belong to the category of bistable processes, would constitute a general response to stress, and would be mediated by killing part of the population though fratricide (Claverys and Havarstein 2007; Claverys et al. 2006). The presence of BOX elements placed upstream of competence genes (Knutsen et al. 2006), and the location of the boxAC element upstream of the *yefM-yoeB* operon (Chan et al. 2011), leads us to propose that this TA pertains to the bistable response of pneumococcus, and that the coupling of boxAC to *yefM-yoeB* may lead to enhancement of the evolution of the *S. pneumoniae* strains that carry them, so that the adaptability of the bacteria to their niches within the human host will be improved.

Acknowledgments Thanks are due to members of the INTERMODS Consortium for critical reading of the manuscript, and to members of Espinosa's lab for many fruitful discussions. C. C. Y. would like to thank A. Meinhart, J. A. Harikrishna and past members of his lab for the productive work and discussions as well as the camaraderie. While this chapter was written, our labs were supported with grants from: the Spanish Ministry of Economy and Competitiveness (grants CSD-2008-00013, INTERMODS, and BFU2010-19597 to M.E.) and the Malaysian Ministry of Science, Technology and Innovation (grants 09-99-06-0104-EA001, 02-02-05-SF0019, 02-02-05-SF0026 and SAGA grant M18 to C.C.Y.)

References

Alonso, J. C., Balsa, D., Cherny, I., Christensen, S. K., Espinosa, M., Francuski, D., et al. (2007). Bacterial toxin-antitoxin systems as targets for the development of novel antibiotics. In R. A. Bonomo & M. E. Tolmasky (Eds.), *Enzyme-mediated resistance to antibiotics: Mechanisms, dissemination, and prospects for inhibition* (pp. 313–329). Washington: ASM Press.

Baquero, F. (2004). From pieces to patterns: Evolutionary engineering in bacterial pathogens. *Nature Reviews Microbiology, 2*, 510–518.

Baquero, F. (2009). Environmental stress and evolvability in microbial systems. *Clinical Microbiology and Infection, 15*, 5–10.

Basset, A., Turner, K. H., Boush, E., Sayeed, S., Dove, S. L., & Malley, R. (2012). An epigenetic switch mediates bistable expression of the type 1 pilus genes in *Streptococcus pneumoniae*. *Journal of Bacteriology, 194*, 1088–1091.

Bogaert, D., van Belkum, A., Sluijter, M., Luijendijk, A., de Groot, R., Rümke, H. C., et al. (2004). Colonisation by *Streptococcus pneumoniae* and *Staphylococcus aureus* in healthy children. *Lancet, 363*, 1871–1872.

Camara, M., Mitchell, T. J., Andrew, P. W., & Boulnois, G. J. (1991). *Streptococcus pneumoniae* produces at least two distinct enzymes with neuraminidase activity: Cloning and expression of a second neuraminidase gene in *Escherichia coli*. *Infection and Immunity, 59*, 2856–2858.

Claverys, J. P., & Havarstein, L. S. (2007). Cannibalism and fratricide: Mechanisms and raisons d'être. *Nature Reviews Microbiology, 5*, 219–229.

Claverys, J. P., Prudhomme, M., & Martin, B. (2006). Induction of competence regulons as general stress responses in Gram-positive bacteria. *Annual Review of Microbiology, 60*, 451–475.

Cooper, W. S., & Kaplan, R. H. (1982). Adaptive 'coin-flipping': A decision-theoretical examination of natural selection for a random individual variation. *Journal of Theoretical Biology, 94*, 135–151.

Croucher, N. J., Vernikos, G. S., Parkhill, J., & Bentley, S. D. (2011). Identification, variation and transcription of pneumococcal repeat sequences. *BMC Genomics, 12*, 120.

Croucher, N. J., Walker, D. R., Romero, P., Lennard, N., Paterson, G. K., Bason, N. C., et al. (2009). Role of conjugative elements in the evolution of the multidrug-resistant pandemic clone *Streptococcus pneumoniae* Spain23F ST81. *Journal of Bacteriology, 191*, 1480–1489.

Chan, W. T., Nieto, C., Harikrishna, J. A., Khoo, S. K., Yasmin Othman, R., Espinosa, M., et al. (2011). Genetic regulation of the *yefM-yoeB*$_{Spn}$ toxin-antitoxin locus of *Streptococcus pneumoniae*. *Journal of Bacteriology, 193*, 4612–4625.

Chan, W. T., Moreno-Córdoba, I., Yeo, C. C., & Espinosa, M. (2013). Manuscript in preparation.

Chandler, M. S., & Morrison, D. A. (1988). Identification of two proteins encoded by *com*, a competence control locus of *Streptococcus pneumoniae*. *Journal of Bacteriology, 170*, 3136–3141.

Cherny, I., Overgaard, M., Borch, J., Bram, Y., Gerdes, K., & Gazit, E. (2007). Structural and thermodynamic characterization of the *Escherichia coli* RelBE toxin-antitoxin system: Indication for a functional role of differential stability. *Biochemistry, 46*, 12152–12163.

Christensen, K. S., Maenhauf-Michel, G., Mine, N., Gothesman, S., Gerdes, K., & Van Melderen, L. (2004). Overproduction of the Lon protease triggers inhibition of translation in *Escherichia coli*: Involvement of the *yefM-yoeB* toxin-antitoxin system. *Molecular Microbiology, 51*, 1705–1717.

Christensen, S. K., & Gerdes, K. (2003). RelE toxins from bacteria and archaea cleave mRNAs on translating ribosomes, which are rescued by tmRNA. *Molecular Microbiology, 48*, 1389–1400.

Christensen, S. K., & Gerdes, K. (2004). Delayed-relaxed response explained by hyperactivation of RelE. *Molecular Microbiology, 53*, 587–597.

Christensen, S. K., Mikkelsen, M., Pedersen, K., & Gerdes, K. (2001). RelE, a global inhibitor of translation, is activated during nutritional stress. *Proceedings of the National academy of Sciences of the United States of America, 98*, 14328–14333.

del Solar, G., Hernández-Arriaga, A. M., Gomis-Rüth, F. X., Coll, M., & Espinosa, M. (2002). A genetically economical family of plasmid-encoded transcriptional repressors in control of plasmid copy number. *Journal of Bacteriology, 184*, 4943–4951.

Dubnau, D., & Losick, R. (2006). Bistability in bacteria. *Molecular Microbiology, 61*, 564–572.
Elowitz, M. B., Levine, A. J., Siggia, E. D., & Swain, P. S. (2002). Stochastic gene expression in a single cell. *Science, 297*, 1183–1186.
Engelberg-Kulka, H., & Glaser, G. (1999). Addiction modules and programmed cell death and antideath in bacterial cultures. *Annual Review of Microbiology, 53*, 43–70.
Epstein, S. S. (2009). Microbial awakenings. *Nature, 457*, 1083.
Evdokimov, A., Voznesensky, I., Fennell, K., Anderson, M., Smith, J. F., & Fisher, D. A. (2009). New kinase regulation mechanism found in HipBA: A bacterial persistence switch. *Acta Crystallographica. Section D, Biological Crystallography, 65*, 875–879.
Francuski, D., & Saenger, W. (2009). Crystal structure of the antitoxin-toxin protein complex RelB-RelE from *Methanococcus jannaschii*. *Journal of Molecular Biology, 393*, 898–908.
Gaál, B., Pitchford, J. W., & Wood, A. J. (2010). Exact results for the evolution of stochastic switching in variable asymmetric environments. *Genetics, 184*, 1113–1119.
Gamez, G., & Hammerschmidt, S. (2012). Combat pneumococcal infections: Adhesins as candidates for protein-based vaccine development. *Current Drug Targets, 13*, 323–337.
Garcia-Pino, A., Christensen-Dalsgaard, M., Wyns, L., Yarmolinsky, M. B., Magnuson, R. D., Gerdes, K., et al. (2008). Doc of prophage P1 is inhibited by its antitoxin partner Phd through fold complementation. *Journal of Biological Chemistry, 283*, 30821–30827.
Gomis-Ruth, F. X., Sola, M., Acebo, P., Parraga, A., Guasch, A., Eritja, R., et al. (1998). The structure of plasmid-encoded transcriptional repressor CopG unliganded and bound to its operator. *EMBO Journal, 17*, 7404–7415.
Gotfredsen, M., & Gerdes, K. (1998). The *Escherichia coli relBE* genes belong to a new toxin-antitoxin gene family. *Molecular Microbiology, 29*, 1065–1076.
Gronlund, H., & Gerdes, K. (1999). Toxin-antitoxin systems homologous with *relBE* of *Escherichia coli* plasmid P307 are ubiquitous in prokaryotes. *Journal of Molecular Biology, 285*, 1401–1415.
Hanage, W. P., Fraser, C., Tang, J., Connor, T. R., & Corander, J. (2009). Hyper-Recombination, diversity, and antibiotic resistance in pneumococcus. *Science, 324*, 1454–1457.
Heinrich, J., Velleman, M., & Schuster, H. (1995). The tripartite immunity system of phages P1 and P7. *FEMS Microbiology Reviews, 17*, 121–126.
Hoskins, J., Alborn, W. E, Jr, Arnold, J., Blaszczak, L. C., Burgett, S., DeHoff, B. S., et al. (2001). Genome of the bacterium *Streptococcus pneumoniae* strain R6. *Journal of Bacteriology, 183*, 5709–5717.
Hurley, J. M., Cruz, J. W., Ouyang, M., & Woychik, N. A. (2011). Bacterial toxin RelE mediates frequent codon-independent mRNA cleavage from the 5′ end of coding regions in vivo. *Journal of Biological Chemistry, 286*, 14770–14778.
Iyer, L.M., Koonin, E.V., Aravind, L. (2002). Extensive domain shuffling in transcription regulators of DNA viruses and implications for the origin of fungal APSES transcription factors. *Genome Biologyl*, *3*, RESEARCH0012.
Kadioglu, A., Weiser, J. N., Paton, J. C., & Andrew, P. W. (2008). The role of *Streptococcus pneumoniae* virulence factors in host respiratory colonization and disease. *Nature Reviews Microbiology, 6*, 288–301.
Kamada, K., & Hanaoka, F. (2005). Conformational change in the catalytic site of the ribonuclease YoeB toxin by YefM antitoxin. *Molecular Cell, 19*, 497–509.
Keren, I., Shah, D., Spoering, A., Kaldalu, N., & Lewis, K. (2004). Specialized persister cells and the mechanism of multidrug tolerance in *Escherichia coli*. *Journal of Bacteriology, 186*, 8172–8180.
Khoo, S. K., Loll, B., Chan, W. T., Shoeman, R. L., Ngoo, L., Yeo, C. C., et al. (2007). Molecular and structural characterization of the PezAT chromosomal toxin-antitoxin system of the human pathogen *Streptococcus pneumoniae*. *Journal of Biological Chemistry, 282*, 19606–19618.
Kim, Y., Wang, X., Ma, Q., Zhang, X. S., & Wood, T. K. (2009). Toxin-antitoxin systems in *Escherichia coli* influence biofilm formation through YjgK (TabA) and fimbriae. *Journal of Bacteriology, 191*, 1258–1267.

Knutsen, E., Johnsborg, O., Quentin, Y., Claverys, J. P., & Havarstein, L. S. (2006). BOX elements modulate gene expression in *Streptococcus pneumoniae*: Impact on the fine-tuning of competence development. *Journal of Bacteriology, 188*, 8307–8312.

Kuchina, A., Espinar, L., Cagatay, T., Balbin, A. O., Zhang, F., Alvarado, A., et al. (2011). Temporal competition between differentiation programs determines cell fate choice. *Molecular Systems Biology, 7*, 557.

Kumar, P., Issac, B., Dodson, E. J., Turkenburg, J. P., & Mande, S. C. (2008). Crystal structure of *Mycobacterium tuberculosis* YefM antitoxin reveals that it is not an intrinsically unstructured protein. *Journal of Molecular Biology, 383*, 482–493.

Kuwahara, H., & Soyer, O. S. (2012). Bistability in feedback circuits as a byproduct of evolution of evolvability. *Molecular Systems Biology, 8*, 564.

Leplae, R., Geeraerts, D., Hallez, R., Guglielmini, J., Drèze, P., & Van Melderen, L. (2011). Diversity of bacterial type II toxin–antitoxin systems: A comprehensive search and functional analysis of novel families. *Nucleic Acids Research, 39*, 5513–5525.

Leung, V., & Lévesque, C. M. (2012). A stress inducible quorum sensing peptide mediates the formation of non-inherited multidrug tolerant persister cells. *Journal of Bacteriology, 194*, 2265–2274.

Li, G.-Y., Zhang, Y., Inouye, M., & Ikura, M. (2008). Structural mechanism of transcriptional autorepression of the *Escherichia coli* RelB/RelE antitoxin/toxin module. *Journal of Molecular Biology, 380*, 107–119.

Maisonneuve, E., Shakespeare, L. J., Jørgensen, M. G., & Gerdes, K. (2011). Bacterial persistence by RNA endonucleases. *Proceedings of the National academy of Sciences of the United States of America, 108*, 13206–13211.

Makarova, K., Wolf, Y. I., & Koonin, E. V. (2009). Comprehensive comparative-genomic analysis of type 2 toxin-antitoxin systems and related mobile stress response systems in prokaryotes. *Biology Direct, 4*, 19.

Martin, B., Humbert, O., Camara, M., Guenzi, E., Walker, J. R., Mitchell, T., et al. (1992). A highly conserved repeated DNA element located in the chromosome of *Streptococcus pneumoniae*. *Nucleic Acid Research, 20*, 3479–3483.

Moreno-Córdoba, I., Diago-Navarro, E., Barendregt, A., Heck, A. J. R., Alfonso, C., Díaz-Orejas, R., et al. (2012). The toxin-antitoxin proteins RelBE2*Spn* of *Streptococcus pneumoniae*: Characterization and association to their DNA target. *Proteins, 80*, 1834–1846.

Moscoso, M., Domenech, M., & García, E. (2010). Vancomycin tolerance in clinical and laboratory *Streptococcus pneumoniae* isolates depends on reduced enzyme activity of the major LytA autolysin or cooperation between CiaH histidine kinase and capsular polysaccharide. *Molecular Microbiology, 77*, 1052–1064.

Mrazek, J., Gaynon, L. H., & Karlin, S. (2002). Frequent oligonucleotide motifs in genomes of three Streptococci. *Nucleic Acids Research, 30*, 4216–4221.

Murayama, K., Orth, P., de la Hoz, A. B., Alonso, J. C., & Saenger, W. (2001). Crystal structure of omega transcriptional repressor encoded by *Streptococcus pyogenes* plasmid pSM19035 at 1.5 A resolution. *Journal of Molecular Biology, 314*, 789–796.

Mutschler, H., & Meinhart, A. (2011). ε/ζ systems: Their role in resistance, virulence, and their potential for antibiotic development. *Journal of Molecular Medicine, 89*, 1183–1194.

Nieto, C., Cherny, I., Khoo, S. K., de García Lacoba, M., Chan, W. T., Yeo, C. C., et al. (2007). The *yefM-yoeB* toxin-antitoxin systems of *Escherichia coli* and *Streptococcus pneumoniae*: Functional and structural correlation. *Journal of Bacteriology, 189*, 1266–1278.

Nieto, C., Pellicer, T., Balsa, D., Christensen, S. K., Gerdes, K., & Espinosa, M. (2006). The chromosomal *relBE*2 toxin-antitoxin locus of *Streptococcus pneumoniae*: Characterization and use of a bioluminescence resonance energy transfer assay to detect toxin-antitoxin interaction. *Molecular Microbiology, 59*, 1280–1296.

Nieto, C., Sadowy, E., de la Campa, A. G., Hryniewicz, W., & Espinosa, M. (2010). The *relBE2Spn* toxin-antitoxin system of *Streptococcus pneumoniae*: Role in antibiotic tolerance and functional conservation in clinical isolates. *PLoS ONE, 5*, e11289.

O'Brien, K. L., Wolfson, L. J., Watt, J. P., Henkle, E., Deloria-Knoll, M., McCall, N., et al. (2009). Burden of disease caused by *Streptococcus pneumoniae* in children younger than 5 years: Global estimates. *Lancet, 374*, 893–902.

Olofsson, H., & Ripa, J. N. J. (2009). Bet-hedging as an evolutionary game: The trade-off between egg size and number. *Proceedings of the Royal Society B, 276*, 2963–2969.

Overgaard, M., Borch, J., Jørgensen, M. G., & Gerdes, K. (2008). Messenger RNA interferase RelE controls *relBE* transcription by conditional cooperativity. *Molecular Microbiology, 69*, 841–857.

Pabo, C. O., & Sauer, R. T. (1992). Transcription factors: Structural families and principles of DNA recognition. *Annual Review of Biochemistry, 61*, 1053–1095.

Pandey, D. P., & Gerdes, K. (2005). Toxin-antitoxin loci are highly abundant in free-living but lost from host-associated prokaryotes. *Nucleic Acids Research, 33*, 966–976.

Park, I.H., Kim, K.-H., Andrade, A.L., Briles, D.E., McDaniel, L.S., Nahm, M.H. (2012). Nontypeable pneumococci can be divided into multiple *cps* types, including one type expressing the novel gene *pspK*. *mBio*, *3*, e00035-00012.

Pedersen, K., Christensen, K. S., & Gerdes, K. (2002). Rapid induction and reversal of a bacteriostatic conditions by controlled expression of toxins and antitoxins. *Molecular Microbiology, 45*, 501–510.

Perez-Trallero, E., Marimón, J. M., Alonso, M., Ercibengoa, M., & García-Arenzana, J. M. (2012). Decline and rise of the antimicrobial susceptibility of *Streptococcus pneumoniae* isolated from middle ear fluid in children: Influence of changes in circulating serotypes. *Antimicrobial Agents and Chemotherapy, 56*, 3989–3991.

Rao, C. V., Wolf, D. M., & Arkin, A. P. (2002). Control, exploitation and tolerance of intracellular noise. *Nature, 420*, 231–237.

Rimini, R., Jansson, B., Feger, G., Roberts, T. C., de Francesco, M., Gozzi, A., et al. (2000). Global analysis of transcription kinetics during competence development in *Streptococcus pneumoniae* using high density DNA arrays. *Molecular Microbiology, 36*, 1279–1292.

Saluja, S. K., & Weiser, J. N. (1995). The genetic basis of colony opacity in *Streptococcus pneumoniae*: Evidence for the effect of box elements on the frequency of phenotypic variation. *Molecular Microbiology, 16*, 215–227.

Shao, Y., Harrison, E. M., Bi, D., Tai, C., He, X., Ou, H.-Y., et al. (2010). TADB: A web-based resource for Type 2 toxin–antitoxin loci in bacteria and archaea. *Nucleic Acids Research, 39*, D606–D611.

Stocker, W., & Bode, W. (1995). Structural features of a superfamily of zinc-endopeptidases: The metzincins. *Current Opinion in Structural Biology, 5*, 383–390.

Takagi, H., Kakuta, Y., Okada, T., Yao, M., Tanaka, I., & Kimura, M. (2005). Crystal structure of archaeal toxin-antitoxin RelE-RelB complex with implications for toxin activity and antitoxin effects. *Nature Structural and Molecular Biology, 12*, 327–331.

Thattai, M., & van Oudenaarden, A. (2004). Stochastic gene expression in fluctuating environments. *Genetics, 167*, 523–530.

Walker, J. A., Allen, R. L., Falmagne, P., Johnson, M. K., & Boulnois, G. J. (1987). Molecular cloning, characterization, and complete nucleotide sequence of the gene for pneumolysin, the sulfhydryl-activated toxin of *Streptococcus pneumoniae*. *Infection and Immunity, 55*, 1184–1189.

Zhang, Y., & Inouye, M. (2009). The inhibitory mechanism of protein synthesis by YoeB, an *Escherichia coli* toxin. *Journal of Biological Chemistry, 284*, 6627–6638.

Chapter 19
Biotechnological and Medical Exploitations of Toxin-Antitoxin Genes and Their Components

Guillermo de la Cueva-Méndez and Belén Pimentel

Abstract We review here the state of the art on the application of toxin–antitoxin pairs in biotechnology and medicine, touching on technologies that range from simple selection of recombinant DNA in the laboratory, to complex and ambitious therapeutic strategies that may become routine in the future.

19.1 TA Pairs as Selection Markers in Bacteria

19.1.1 Selection of Recombinant DNA Clones by Toxin Inactivation

Recombinant DNA technologies have become of paramount importance in most biology laboratories today. One inconvenience of DNA cloning is the low frequency with which DNA inserts are successfully introduced into target plasmid molecules in ligation reactions. This means that most target DNA plasmids may be circularized back during ligation without carrying an insert, and this "empty" vector population can generate a predominant background of false positive clones in subsequent selection steps. Distinguishing bacterial colonies carrying the pursued recombinant DNA from those carrying only the empty vector is costly and time-consuming. TA systems have been exploited to eliminate false positives from this mixture of clones, circumventing the problem above by only allowing the growth of bacterial cells that receive insert-bearing plasmids during transformation.

G. de la Cueva-Méndez (✉) · B. Pimentel
Centro Andaluz de Nanomedicina y Biotecnología (BIONAND),
Parque Tecnológico de Andalucía, C/Severo Ochoa, 35,
29590 Campanillas, Malaga, Spain
e-mail: gdelacueva@bionand.es

K. Gerdes (ed.), *Prokaryotic Toxin-Antitoxins*,
DOI: 10.1007/978-3-642-33253-1_19,

The first of these systems was developed using the CcdA and CcdB TA pair from plasmid F (Bernard 1995, 1996). These proteins contribute to plasmid F stability by killing newborn daughter cells that have not inherited a plasmid copy at cell division (Ogura and Hiraga 1983; Jaffe et al. 1985). CcdB toxin kills cells by inhibiting Gyrase A, one of the subunits of *E. coli's* topoisomerase II, and this poisoning activity is neutralized by CcdA (Miki et al. 1984). The *ccdB* gene was fused to the multiple cloning site (MCS) of plasmid pUC18 (Vieira and Messing 1982), creating a fusion gene encoding for the N-terminal fragment of beta-galactosidase and CcdB itself (Bernard et al. 1994). The resulting CcdB variant produced from this vector is sufficiently toxic to kill bacteria in the absence of CcdA. However, if *ccdB* is inactivated by insertion of a foreign DNA fragment in the MCS above, the recombinant plasmid cannot produce CcdB and no longer interferes with host viability. This positive selection tool for recombinant clones is highly efficient and simplifies cloning procedures, as only cells containing recombinant plasmids give rise to colonies. Similar technologies were implemented later on using toxin Kid from plasmid R1 (Gabant et al. 2000) and toxin ParE from plasmid RK2 (Kim et al. 2004). In all cases, empty vectors can be amplified using *E. coli* strains that express the corresponding antitoxin partner (Kis or ParD) from chromosomal locations, or that produce a CcdB-insensitive GyrA mutant.

The technology based on CcdB was initially marketed by Delphi Genetics SA, a spin-off company of the Université Libre de Bruxelles (ULB), and subsequently licensed to Invitrogen Inc. The latter company has included the technology in a large portfolio of positive selection vectors with different copy numbers, host range, and/or transfer properties, making it a standard tool in many molecular biology laboratories worldwide. For instance the technology is included in Invitrogen's Gateway vectors, which exploit site-specific recombination to shuttle DNA inserts from donor vectors into all sorts of recipient vectors rapidly and efficiently. In this technology positive recombinants are selected because site-specific recombination replaces the *ccdB* gene in recipient vectors by the gene of interest transferred from donor vectors (Walhout et al. 2000).

19.1.2 Selection of DNA Recombinant Clones by Antitoxin Reconstitution

A new generation of positive selection vectors based on CcdB has recently been generated by Delphi Genetics SA to produce the StabyCloning™ system. In this case, the *ccdB* gene is inserted in the chromosome of *E. coli* cells, and a truncated inactive version of the antitoxin gene is present in the cloning vector. Insert DNAs that include in their 5′ end the 14 base pairs missing from the antitoxin gene can reconstitute the inactive *ccdA* gene if they are integrated in the vector in the right orientation. This selection strategy is well suited for cloning blunt-ended PCR products produced using 5′ oligonucleotides that include the missing base pairs

above, an approach that has been used recently for the generation of microsatellite libraries (Tzika et al. 2008). One drawback of this strategy is that it requires modification of the strain used for selection. An advantage is that there is no need to include antibiotic resistance genes in the DNA vector, as reconstituted *ccdA* functions as the selective marker for plasmid maintenance. This eliminates some of the pitfalls associated with the use of some antibiotics during selection, like growth of satellite plasmid-free colonies.

19.1.3 Non-Antibiotic Markers for Selection of Plasmid-Bearing Cells

Antibiotics are commonly used during bacterial fermentation, and the vast majority of expression vectors contain antibiotic resistance genes to enable counter-selection of plasmid-free cells and avoid that they dominate the culture. The use of these genes raises safety concerns that are increasingly highlighted by regulatory authorities worldwide (Vandermeulen et al. 2011). Thus, the expectation is that a "zero tolerance" toward antibiotic-based selection and production systems will be the general consensus in the near future (Peubez et al. 2010). Given the tremendous impact that this decision would have on both academic and industrial settings, other approaches for selection of plasmid-containing cells are required.

TA pairs found in plasmids function as stability systems and therefore they offer an attractive alternative to the use of antibiotic resistance genes for plasmid selection. The introduction of TA pair *hok-sok* (Gerdes et al. 1986) or *parDE* (Roberts et al. 1993) in a high-copy number DNA vector was shown to reduce the plasmid loss frequency from 1.4×10^{-2} cells per generation to 2.4×10^{-8} and 2×10^{-5} cells per generation, respectively, in *E. coli*. Placing both toxin–antitoxin pairs on the same plasmid reduced loss frequency even further, down to 5.4×10^{-11} per cell and generation. This level of stability enhancement increases from 3 to 43 h the time required to accumulate 10 % of plasmid free cells in exponentially growing cultures, without affecting cell growth rates, or their efficiency producing recombinant proteins. This stability enhancement is well suited for batch and fed-batch fermentations but not for continuous operations because TA pairs only delay, but do not prevent, the takeover of a culture by plasmid-free cells. For instance, in the case above only 10 % of cells in the culture still retain the plasmid after 90 h of growth. In view of the stabilizing synergy displayed by *hoksok* and *parDE* it would be interesting to test whether other TA combinations, or the addition of other TA systems to the plasmid above, would enhance its stability even further.

Physical separation of toxin and antitoxin partners has also been used to achieve plasmid stabilization in the absence of antibiotics (Szpirer and Milinkovitch 2005). As with the StabyCloning technology, in this strategy toxin CcdB is expressed from the cell's chromosome while antitoxin CcdA is expressed from the bacterial plasmid. Plasmid maintenance, and therefore CcdA expression, ensures neutralization of CcdB expressed in host cells. Cells losing the plasmid cannot produce the antitoxin, and continuous production of CcdB eliminates the plasmid-free cell. This approach

makes it very difficult for the latter cells to escape cell death and, when combined with stabilization systems found in high-copy number plasmids, like the *cer* locus of ColE1 (Summers and Sherratt 1984), provides an excellent stabilization enhancement of plasmids in bacteria. One pitfall of this approach is that it requires a modification of the host strain, but the system improves recombinant protein expression and/or plasmid recovery from those modified cells and therefore provides a safer and more efficient manufacturing alternative for the production of DNA and protein vaccines (Peubez et al. 2010).

19.2 Artificial Activation of TA Pairs as Antibiotic Strategies

The discovery of antibiotics is a major medical achievement of the past century. The availability of penicillin, streptomycin, and sulfonamides in the late 1940s, and of chloramphenicol, erythromycin, and tetracycline in the following decade raised the expectation that deaths associated to bacterial infections would rapidly become a problem of the past. However, such anticipated hopes were shadowed by the overwhelming adaptability of microorganisms, and the rate at which they developed resistance to antibiotics (Davies 2007). The problem persists today despite a considerable expansion of our antibiotic weaponry, and both the number of resistant microorganisms and the breadth of resistance in single bacteria are unprecedented and mounting (Levy and Marshall 2004). The consequences of this are dramatic; 1.7 million U.S. citizens acquired bacterial infections in hospitals in 2002, causing 99,000 deaths and health care costs between $5 and 10 billion, with similar statistics reported from the EU, and reaching a disproportionate scale in developing countries (Peleg and Hooper 2010; Okeke et al. 2005).

Most bacteria become insensitive to antibiotics through the acquisition of resistance determinants that accumulate in plasmids called R-factors (Alekshun and Levy 2007; Nikaido 2009). R-factors are mobile genetic elements that can promote their conjugative transfer between bacteria of different genera. This, and their ability to exchange genetic material with co-existing plasmids and host cell chromosomes, explains the facility with which antibiotic resistance evolves and spreads in both Gram-positive and -negative bacteria (Martinez and Baquero 2002; Carattoli 2009; Norman et al. 2009; Jensen et al. 2010).

The emergence and spread of pathogenic bacteria that have become resistant to multiple antibiotics through lateral gene transfer have increased the need for novel antimicrobials. Interestingly, TA pairs are ubiquitous among plasmids conferring antibiotic resistance and chromosomes of pathogenic bacteria. In the latter case TA pairs tend to be clustered and linked to mobile genetic elements suggesting that, as described for antibiotic resistance genes, TA pairs are mobile elements that move frequently within and between plasmids and chromosomes (Pandey and Gerdes 2005; Guglielmini and Van Melderen 2011; Leplae et al. 2011; Shao et al. 2011). These observations, the apparent absence of TA pairs from eukaryotic organisms, and the fact that these toxins inhibit bacterial growth and/or induce bacterial cell

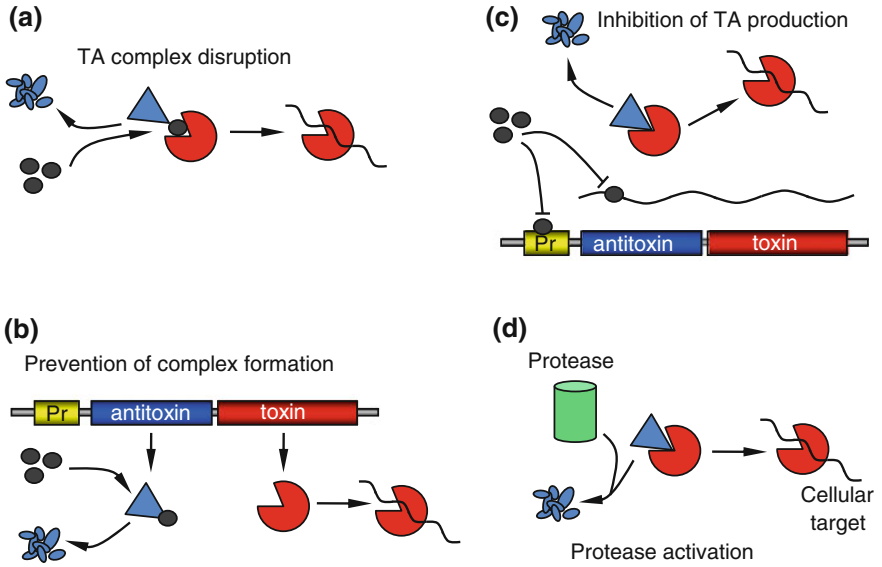


Fig. 19.1 Possible ways of inducing artificial activation of TA pairs. TA pairs may be activated artificially following direct (**a**, **b**) and indirect (**c**, **d**) approaches. Direct activation may be achieved with molecules (*gray circles*) that disrupt preformed TA complexes (**a**) or that prevent complex formation (**b**). Indirect activation may be achieved by inhibiting transcription or translation of TA genes (**c**) or by activating cellular proteases responsible for antitoxin degradation (**d**). Adapted from Williams and Hergenrother (2012)

death (Yamaguchi and Inouye 2011) has increased efforts directed at inducing artificial activation of toxins as an antibiotic strategy.

19.2.1 Indirect Approaches for Toxin Activation

Direct and indirect approaches have been proposed to induce toxin activation (Williams and Hergenrother 2012) (Fig. 19.1). Indirect activation relies on the observation that antitoxins are more labile than their cognate toxins (Gerdes et al. 2005). Antitoxins are degraded by cellular proteases and therefore these proteins must be constantly replenished in cells to keep their toxic partners neutralized. Thus, any molecule that inhibits transcription or translation at a TA locus would prevent replenishment of its encoded antitoxin and would lead to toxin activation. The same outcome would be expected from a molecule that increases the activity of proteases degrading antitoxins. Thus, it has been proposed that molecules like these could constitute valuable therapeutic agents to fight infections caused by antibiotic resistant pathogens that bear TA loci in their chromosomes and/or resident plasmids (Williams and Hergenrother 2012).

19.2.1.1 Inhibitors of Toxin–Antitoxin Production

The report above pointed out that sequence-specific DNA binders that inhibit transcription of target genes have been recently designed (Raskatov et al. 2012) and that molecules like this, able to bind promoters at TA operons, may be designed to repress transcription at these loci. We feel that this strategy may not be suitable in the latter case as sequence-specific DNA binders are designed to inhibit transcription of target genes by competing with transcriptional activators (Raskatov et al. 2012). Most TA proteins function as transcriptional auto-repressors, so inhibiting transcription of these operons would require 'stapling' TA complexes to their DNA binding sites, rather than inhibiting their interaction with them. Intuitively, developing 'TA-DNA staplers' might be a very challenging task.

An alternative would be to design antisense (as) RNAs that provoke the specific degradation of the mRNAs encoding type II TA pairs, or that outcompetes antitoxin RNAs in the case of type I TA pairs. The suitability of the latter approach has already been validated experimentally and has proved to be an efficient way of inducing toxin activation and cell killing in *E. coli* cells bearing plasmids that encode the *hok-sok* TA locus (Gerdes et al. 1986; Faridani et al. 2006). Delivering asRNA can be achieved using liposomes or self-penetrating peptide nucleic acids (Faridani et al. 2006; Meng et al. 2012; Bai et al. 2012). Of note, these delivery systems can also translocate mRNA through mammalian cell membranes, so the effects in human cells of any asRNAs designed to activate toxin expression should be carefully evaluated before hand, to discard that it silences the expression of essential genes in humans.

A pitfall of any approach directed at inhibiting the expression of type II TA pairs is that they would only induce the activation of toxin molecules that were produced in cells before treatment. The concentration of these proteins is kept relatively low in cells due to feedback loop regulatory mechanisms, like transcriptional auto-repression (Yamaguchi and Inouye 2011). Thus, permanent inhibition of toxin and antitoxin production is likely to induce only transient and mild toxicity. This may not be enough to kill cells, as many of these toxins only kill bacterial cells if activation is sustained long enough (Amitai et al. 2004; Lioy et al. 2006, 2012).

19.2.1.2 Increased Degradation of Antitoxins

Activation of antitoxin-degrading proteases (i.e. ClpP and Lon) is more likely to trigger sustained toxin activation, as constant antitoxin degradation would impede transcriptional auto-repression by TA complexes, overcoming the limitation above. Compounds that stimulate the activity of ClpP have been developed (Sass et al. 2011; Leung et al. 2011) and it would be interesting to test whether they induce activation of toxins degraded by this protease. In any case, it should be noted that this strategy is likely to induce indiscriminate killing of bacterial cells, independently of whether they contain TA pairs or not, as suggested by results obtained with Lon and ClpP (Goff and Goldberg 1987; Christensen et al. 2004; Sass et al. 2011). Furthermore, human cells encode homologues of these proteases

and the possibility that activators may lead to undesired side effects in treated patients should be discarded appropriately.

19.2.2 Direct Approaches for Toxin Activation

Direct approaches constitute a more straightforward approach to induce toxin activation, where a drug directly targets toxin and/or antitoxin proteins to either disrupt preformed complexes or to prevent their formation in the first instance. So far, only disruption of preformed complexes has been explored. In one work, hepta- and octapeptides representing fragments of a region in *Bacillus anthracis* antitoxin PemI that were predicted to interact with its toxin partner, PemK, were examined for their ability to disrupt PemK-PemI complexes (Agarwal et al. 2010). Some of these peptides managed to disrupt the TA complex with considerable efficiency. However, they also reduced the endoribonucleolytic activity of PemK, which is responsible for its toxic effect in cells. Most probably this is due to the fact that screened peptides were derived from the stretches of the antitoxin known to interact with the toxin. Although these results are encouraging and constitute a valuable proof of principle, partial activation of the toxin may not be enough to achieve the pursued antibiotic effect. It would be desirable to find new molecules able to disrupt TA complexes completely by targeting the antitoxin component, as these should not interfere with the activity of the toxin. Peptides like these may be derived from PemK regions known to interact with PemI, a possibility that is worth testing.

In a different work a library of 6-, 14-, and 17-aminoacid long peptides was screened in search of molecules that could disrupt the interaction between toxin ζ and antitoxin ε encoded by plasmid pSM19035 from *Streptococcus pyogenes*. Disruption of toxin ζ-antitoxin ε complexes was accomplished using two different mixtures of 17-aminoacid peptides, although the effect was not observed when single peptides in those mixtures were used instead (Lioy et al. 2010). These results pointed out that complex disruption might have been achieved through the combined action of different peptides with weak disrupting activity. Whether these results could be exploited to generate single disruptor molecules remains to be established. Furthermore, the work was performed using an inactive derivative of toxin ζ, and did not include an analysis of whether these peptides interacted with the toxin or with the antitoxin. Thus, further work will be required to determine if these peptides disrupt the TA complex without inhibiting toxin ζ.

19.2.3 Selecting the Right Strategy

Although still too early to predict the therapeutic value of toxin activation in clinical settings, the results above suggest that development of TA disrupting agents is technically possible. Yet, drug discovery and development is a long and

expensive process, and selecting the right approach will be important to speed up progress toward new antibiotics based on TA disruptors. An ideal TA complex-disruptor should be small, to facilitate its pharmacological formulation. This may prove difficult to accomplish, given the high affinity with which most antitoxin interact with their partners and the extension and physical–chemical nature of these interactions (Kamada et al. 2003; Koga et al. 2011; Overgaard et al. 2009; Chopra et al. 2011; Fiebig et al. 2010; Yang et al. 2010; Zhu et al. 2010). Lead compound identification should rely on screenings designed to select for disruptors that interact with antitoxins, so that positive hits do not interfere with the toxin's activity. It will be essential to select the right TA targets, a decision that should be based on data concerning their clinical relevance (i.e. their prevalence in antibiotic resistance clinical isolates), (Moritz and Hergenrother 2007; Williams et al. 2011), their functionality, and evidence of their bactericidal effect. Furthermore, the mode of action of some toxins enables cells to develop resistance against them (Bernard and Couturier 1992). Some toxins induce persistence, a dormant, reversible state that protects cells from the action of antibiotics (Wang and Wood 2011; Maisonneuve et al. 2011), or promote the survival of cells under stress (Amitai et al. 2004; Lioy et al. 2006, 2012). Toxins known to induce any of these phenotypes should also be discarded as targets.

TA pairs encoded in plasmids are thought to function as post-segregational killing (PSK) systems (Van Melderen and Saavedra De Bast 2009) and, intuitively, they appear to be ideal candidates for the approaches above. However, it has been shown that the *kis-kid* TA pair encoded in plasmid R1 functions as a rescue system, and that its activation increases plasmid copy numbers and enforces retention of antibiotic resistance in host cells, without killing them (Pimentel et al. 2005; de la Cueva-Mendez and Pimentel 2007). It is possible that other plasmid-encoded TA pairs may function as rescue systems. If so, their artificial activation is likely to keep cells alive and at the same time increase quantitatively their antibiotic resistance. Moreover, the simultaneous increase in plasmid copy numbers and sustained TA disruption should also lead to higher production of both the toxin and the antitoxin, which could titrate out the complex disruptor. Therefore, TA pairs like these should also be discarded as potential targets for disruptor discovery programs.

19.3 TA Pairs as Facilitator of Protein Structural Analysis

Deciphering the three-dimensional structure of proteins and their complexes is in many cases essential to understand biological function, and constitutes a major requirement for the rational design of pharmacological agents directed against them. Unfortunately, protein structure determination is often a long and expensive process, with a success rate below 5 %. More than 75 % attempts fail because target proteins cannot be prepared in sufficient amounts or suitable forms for structural analysis, and 70 % of the remaining cases do not succeed at producing good crystals or high quality NMR spectra (Terwilliger et al. 2009). Membrane

proteins are particularly vulnerable to the limitations above, as evidenced by the fact that they constitute less than 1 % of the high-resolution protein structures solved to date (Watts 2005). This hampers drug development because membrane proteins comprise one-third of all cellular polypeptides and are involved in 85 % of cell signaling pathways, making them very attractive drug targets in the pharmaceutical industry.

New enabling technologies are required to overcome the problem above, and TA pairs are being used to implement some of them. For instance, solid-state (ss) in cell NMR facilitates the study of drug interactions with membrane proteins, and even the structural analysis of large functionally active complexes with atomic resolution. Importantly, in-cell NMR does not require sample purification and when applied in solid-state, the hydrophobic nature of membrane proteins is not an issue that affects sample preparation. However, in-cell NMR is not a very sensitive technique, and it requires that target proteins are over-expressed by cells, and that they incorporate isotopic labels both uniformly and exclusively, to distinguish them from other cellular proteins (Reckel et al. 2005). The recent development of a condensed single protein production (cSPP) system, which exploits *E. coli's* toxin MazF, constitutes an important step in this direction. cSPP is a highly effective method for high yield production of recombinant proteins that also enables their exclusive isotopic labeling in vivo (Suzuki et al. 2005).

19.3.1 Condensed Single Protein Production

cSPP relies on two separate technologies, both devised by Masayori Inouye's laboratory. The first one is an expression system (pCOLD) that achieves formidable yields of recombinant protein production in conditions that increase solubility and stability. pCOLD is based on the observation that, when exposed to low temperature, *E. coli* induces the expression of a handful proteins that help cells to adapt to cold-shock (Jones et al. 1987). CspA (cold-shock protein A) stands out among these proteins, and constitutes up to 2 % of total protein content in *E. coli* cells grown at low temperature (Goldstein et al. 1990). The expression of CspA is regulated post-transcriptionally. *cspA* mRNA is produced both at 37 and 15 °C, but its 5′ and 3′ untranslated regions (UTR) function as a thermosensor (Giuliodori et al. 2010). These UTR regions fold differently at each one of these temperatures, and the structures produced have opposite consequences for the stability and translatability of the mRNA. Folding at 37 °C hides the mRNA's Shine-Dalgarno (SD) sequence and initiation codon from ribosomes, and also turns the mRNA very unstable. In contrast, folding at 15 °C exposes the SD sequence and initiation codon in a structure that is avidly recognized by ribosomes, increasing translation rates, and also augments the stability of the transcript 100-fold (Xia et al. 2002).

These features have been exploited to generate a series of expression vectors (pCOLD) that are induced at low temperature, enabling production of large amounts of soluble proteins in *E. coli* cells by simply growing them at 15 °C (Qing

et al. 2004). In these conditions, most of the cell's translation machinery is devoted to produce the recombinant protein, which enables its preferential isotopic labeling. Using pCOLD it is possible to perform a basic NMR characterization of recombinant proteins directly from cleared cell lysates, thus skipping the need to purify them to homogeneity. However, although NMR spectra obtained this way provide valuable information about protein structural integrity they do not allow solving their 3D structures with atomic resolution.

The latter requires exclusive (rather than predominant) isotopic labeling of target polypeptides, and background protein synthesis in cells under cold-shock is sufficiently high to impede this happening. To sort this out, Inouye's laboratory has cleverly exploited the toxin component of the *mazEF* TA pair in *E. coli*. MazF is an mRNA interferase that cleaves RNA at single-stranded ACA sites, and that is neutralized by its antitoxin partner MazE (Zhang et al. 2003). ACA sites are abundant in genes and therefore MazF depletes cells from mRNAs, inducing a dramatic inhibition of protein synthesis in *E. coli* cells. Interestingly, MazF arrests cells in a "quasi-dormant" state that sustains ATP regeneration, transcription, and translation (of ACA-less mRNAs), and that can be prolonged for days. ACA sites can be eliminated silently from genes, and work in Inouye's laboratory demonstrated that concomitant expression of MazF and a target gene engineered to encode an ACA-less mRNA results in sustained and high-level target expression, in the virtual absence of background cellular protein synthesis (Suzuki et al. 2005). Inouye's group also showed that these features are maintained in *E. coli* cells grown at low temperature, making it possible to combine the effects of MazF and pCOLD vectors to achieve very high yield production of recombinant proteins from ACA-less mRNAs (20–30 % of total protein) and their almost exclusive isotopic labeling (90 % of the incorporated isotope) in vivo. Furthermore the system allows production of target proteins in the absence of cell growth, which has two important advantages. On the one hand, it may be used to express and characterize the structure of some toxic proteins. On the other hand, it allows up to 150-fold condensation of cultures at the expensive stage of isotopic labeling, reducing costs as much as 99 % without compromising yields, which still remain in the range of milligrams per milliliter of culture.

19.3.2 SPP and in-Cell NMR

The system has been used for the structural analysis of prokaryotic and eukaryotic membrane proteins, which could be incorporated into *E. coli* membranes and produced reasonably well-resolved NMR spectra using whole cells (Mao et al. 2009), and even better ones when performed on outer membrane fractions from these samples (Vaiphei et al. 2011). Furthermore, as the SPP system allows protein production in the absence of cell growth, cerulenin, an inhibitor of phospholipid biosynthesis (Heath et al. 2001) could be used without killing cells to eliminate the typical spurious signals in ssNMR spectra that arise by non-specific isotopic

labeling of phospholipids when ^{13}C-glucose is used as precursor, further improving NMR spectra (Mao et al. 2011).

19.3.3 Improvement of cSPP

The cSPP system is being subject of clever improvement. For instance, co-induction of MazF and the target protein results in 20 % of the latter protein being produced without isotopic enrichment, which reduces the amount of final sample that can be observed by NMR. This problem was eliminated by either inducing transcription of both proteins independently (Schneider et al. 2009) or by using tryptophan-or histidine-free MazF variants to perform cSPP in *E. coli* auxotrophs for one of those amino acids. This trick enables exclusive production of the mRNA interferase in media lacking tryptophan or histidine, and subsequent synthesis of target proteins in the presence of these aminoacids (Vaiphei et al. 2011). Using this approach more than 98 % of the target protein is labeled in the final sample. More recently, the system has also been modified to produce soluble, isotopically labeled, human proteins in *E. coli* that accumulate in the periplasmic space of *E. coli*, from where they can be recovered easily for their subsequent NMR analysis (Mao et al. 2010).

The most critical step in the cSPP system is the preparation of an ACA-less gene for target proteins. Although this can be achieved by oligonucleotide-directed site-specific mutagenesis ACA sites can be very abundant in genes, making the approach cumbersome and time-consuming. The problem can be circumvented by chemically synthesizing the entire gene. An attractive option would be to generate cSPP systems based on mRNA interferases that are more specific than *E coli's* MazF. For instance, Kid, an homologue of MazF encoded in plasmid R1, cleaves mRNA at single-stranded UUACU sites, and also inhibits protein synthesis and cell growth without killing *E. coli* cells (Pimentel et al. 2005). Similarly, some MazF homologues in *S. aureus*, *B. subtilis*, *M. xanthus*, *M. tuberculosis,* and *H. walsbyi*, cleave mRNA at specific pentanucleotide (even heptanucleotide) sequences (Yamaguchi and Inouye 2011). It would be interesting to test if these mRNA interferases can be used to develop cSPP systems that reduce (even eliminate) the hassle of mutating cleavage sites from target genes.

A limitation of in-cell NMR spectroscopy in bacteria is its inability to study post-translational protein modifications or to analyze interactions between eukaryotic targets and their ligands in more physiological environments. To overcome this limitation development of a eukaryotic SPP system would be ideal. Interestingly, mRNA interferases Kid and MazF are functional in eukaryotic cells (de la Cueva-Mendez et al. 2003; Shimazu et al. 2007). Both proteins induce apoptosis in human cells, but the observation that MazF still shuts off protein synthesis completely in NBK/BIK- or BAK-deficient human cells, without killing them, provides an interesting clue of how a human cell-based SPP system could be generated.

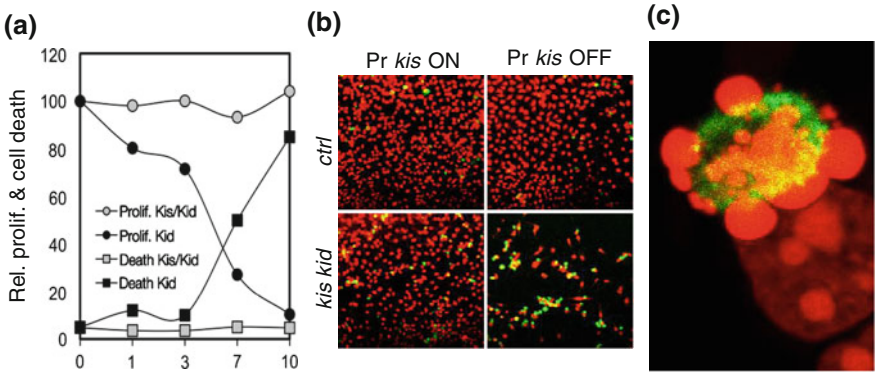


Fig. 19.2 Independent transcriptional control of *kis* and *kid* allows regulated inhibition of cell proliferation and cell death in human cells. **a** Time course comparing relative growth and apoptosis in HeLa cells stably transfected with a plasmid expressing Kid from a constitutive promoter and Kis from a doxycycline repressible promoter. *Curves* represent a comparison between these cells and empty control cells in conditions that allow production of Kis (Kis/Kid) or repress it (Kid). **b** Low magnification confocal images of samples analyzed in **a** after 10 days of Kis repression. Cells were stained with propidium iodide (*red*) and a membrane apoptotic marker (*green*). **c** Magnified image of one cell from the bottom right panel shown in **a** highlighting the characteristic apoptotic phenotype that Kid induces in human cells. Adapted from de la Cueva-Méndez et al. (2003)

19.4 Some TA Pairs are Functional in Eukaryotic Cells

As pointed out above, several groups have analyzed the effects of TA pairs in eukaryotic cells. Expression of RelE was shown to inhibit growth in *Saccharomyces cerevisiae*, an effect that was counteracted by co-expression of antitoxin RelB (Kristoffersen et al. 2000). Expression of RelE is toxic also in human cells, where it induces apoptosis, although in this species the neutralizing effect of RelB was not evaluated (Yamamoto et al. 2002). The results above led to the proposal that *relE* and *relB* genes could be used to restrict growth of genetically modified yeast strains to controlled environments. For instance, toxin RelE may be expressed constitutively in cells, independently of environmental conditions, while production of neutralizing amounts of antitoxin RelB would only be allowed in controlled environments, and not in cells that may escape from these.

Another report demonstrated that expression of toxin Kid from plasmid R1 also inhibits yeast cell growth and triggers apoptosis in higher eukaryotic cells, and that these effects are neutralized by co-expression of antitoxin Kis (de la Cueva-Mendez et al. 2003) (Fig. 19.2). In this work independent control of *kid* and *kis* gene expression was used to modulate cell proliferation and cell death in eukaryotes, and it was proposed that this strategy could be exploited to achieve selective cell killing in eukaryotic organisms. This approach would have multiple applications such as targeted gene- and protein-therapies against cancer, generation of cell lineage knock-outs for developmental studies, or of animal models for the study of degenerative disorders.

Similar results were obtained later on using toxin MazF and antitoxin MazE, the chromosomal homologues of Kid and Kis in *E. coli*, in human cells (Shimazu et al. 2007). This work showed that the mRNA interferase activity of MazF (and presumably also of Kid) is responsible for its lethal effect in human cells. Furthermore, it described that induction of apoptosis by MazF in human cells requires pro-apoptotic protein BAK and its upstream regulator NBK/BIK. Cells deficient in BAK do not induce apoptosis when exposed to MazF, although the protein still inhibits protein synthesis in these cells, completely. As mentioned earlier this offers the attractive possibility of developing an SPP system in mammalian cells, which would have great value for the structural characterization of human proteins by in-cell NMR.

19.5 Selection of High Transgene-Expressing Mammalian Cell Pools

The strength and stability of transgene expression in stably transfected mammalian cells depends on their chromosomal integration site. Integration often occurs randomly and the identification of cells that have integrated the transgene in chromosomal locations favoring sustained and high expression levels is time-consuming. Kid and Kis have been exploited to carry out effortless selection of clones with the characteristics above. For this, cells are stably transfected with a plasmid from which Kid expression can be transactivated with doxycycline. These cells are then transfected with a second plasmid from which the protein of interest and Kis are co-expressed using a bicistronic transcriptional unit. Strict co-expression of the transgene with antitoxin Kis in cells that express Kid allows enrichment of cells expressing high levels of the transgene. Progressive increase of Kid expression in these cells using doxycycline selects for cells that increase transgene and Kis expression appropriately. Extending this procedure for several weeks allows the selection of cells that express the transgene 120-fold higher than the initial population, in the absence of antibiotic selection (Nehlsen et al. 2010).

19.6 Selective Cell Killing in Multi-Cellular Organisms

Toxicity of Kid and protection by Kis occur in both somatic and embryonic cells. Experiments where single blastomeres of two-cell embryos of *X. laevis* were microinjected with Kid, or Kis, or both Kid and Kis proteins showed that the toxin inhibits development of the injected half of the embryo, and that this does not happen when Kis alone or both Kid and Kis are injected instead (de la Cueva-Mendez et al. 2003). These observations provided the grounds to analyze whether *kis* and *kid* could be exploited to kill specific cells in animals without harming any other cell in the organism. Targeted cell ablation is commonly used as a tool for studying the role of a particular cell line in the multicellular context of the entire organism. Genetic

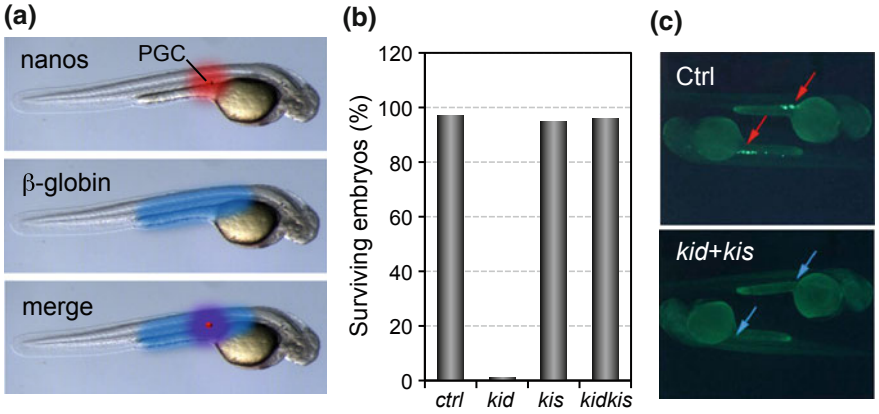


Fig. 19.3 Selective killing of primordial germ cells in animals using toxin *kid* and antitoxin *kis* mRNAs. **a** Scheme depicting the extension and level of Kid (*red*) and Kis (*blue*) expression in zebrafish embryos injected with *kid* and *kis* mRNAs fused to the 3′-UTR regions of *nanos1* and beta-globin, respectively. The *bottom image* is a merged representation of the top two panels. Kid/Kis ratios should protect any cells lying in *blue* or *purple regions*. PGC stands for primordial germ cells. **b** Percentage of surviving embryos 48 h after being injected the indicated mRNAs. **c** Control embryos injected with *kis* mRNA or with equimolecular amounts of *kis* and *kid* mRNAs developed without apparent somatic defects and are morphologically normal. Inspite of this, PGCs (*green fluorescene*) are only detected in control animals. *Red arrows* indicate the germ cells, and *blue arrows* indicate the regions where PGCs are normally found in the fluorescent images. Adapted from Slanchev et al. (2005)

methods for expressing toxic molecules under the control of tissue-specific promoters allow high specificity of cell targeting (Arase et al. 1999; Roman et al. 2001). Nevertheless, in many cases, even minute amounts of the toxin expressed outside of the targeted cells can compromise the viability of the organism. Obstacles of a similar nature often prevent the use of targeted cell ablation for medical purposes, such as in cancer treatment. It had been claimed that Kid and Kis could be used to refine the specificity of the toxic action to a well-defined cell population by protecting non-targeted cells using minimal background levels of Kis expression, but not enough to protect targeted cells from toxicity (de la Cueva-Mendez et al. 2003).

To verify this hypothesis an mRNA fusion between *kid* and the 3′-UTR of the zebrafish *nanos1* gene, was injected in one-cell fish embryos. Nanos1 3′-UTR is known to direct the expression of Nano1 to Primordial Germ Cells (PGC) (Koprunner et al. 2001) and therefore this treatment eliminated the PGCs in the embryos. This demonstrated that Kid is also functional in zebrafish cells, but leaky expression of Kid also occurred in other cells in the embryo, and this resulted in somatic defects and embryonic death. Injection of an mRNA fusion between *kis* and the 3′-UTR of ubiquitous beta-globin gene resulted in uniform somatic expression of Kis and did not lead to any visible effect on embryos or adults fish. Importantly, co-injection of the antidote mRNA effectively neutralized the deleterious effects of Kid on somatic development. Embryos injected with equimolecular ratios of toxin mRNA and antidote mRNA showed PGC loss but appeared normal at 24 and 48 h post-fertilization and could be raised to adulthood (Slanchev et al. 2005) (Fig. 19.3).

These findings confirmed that Kid and Kis could be applied for highly-specific ablation of targeted eukaryotic cells. It remains to be established if the same concept could be used to achieve selective killing of cancer cells.

19.7 Antiviral Therapies

MazF and MazE have also been used to implement novel antiviral strategies against Hepatitis C virus (HVC) and Human Immunodeficiency Virus (HIV). NS3 serine protease and PR aspartyl proteases are essential for HCV and HIV replication, respectively, and represent prime targets for developing antiviral therapies. In a recent work, the C-terminal 41-residue fragment of antitoxin MazE was fused to the N-terminal end of toxin MazF using linkers bearing a specific protease cleavage site for either HIV PR (HIV-1 protease) or NS3 protease (HCV protease). Incubation of these 'MazF-zymoxins' with the corresponding proteases induced proteolytic cleavage and release of the MazE peptide and enabled cleavage of single-stranded RNA in vitro (Park et al. 2012). Similar observations were made by a different group, which also fused MazF to a neutralizing fragment of MazE via an NS3-cleavable linker. Expression of this fusion protein was well tolerated in naive healthy cells but killed HCV-infected cells and cells expressing NS3 (Shapira et al. 2012).

Activation of *mazF* transcription directed by the long terminal repeat (LTR) of HIV-1 has also been used to develop an anti-viral therapy (Chono et al. 2011a). Transcriptional activation at the LTR requires the Tat protein, which is produced by HIV-1 infection. CD4+ T cells carrying *mazF* downstream of HIV1-LTR in their chromosome were infected with HIV-1. Tat expressed in these cells upon infection induced MazF expression and this resulted in degradation of ACA-rich HIV mRNA, and prevented viral replication. Moreover, the same result was observed when monkey primary CD4+ T cells carrying LTR-*mazF* were infected with simian/human immunodeficiency virus (SHIV). Strikingly, Tat-induced expression of MazF did not result in cleavage of cellular mRNA, and neither killed nor inhibited growth of CD4+ T cells. This result is somehow surprising considering the potent endoribonucleolytic and cytotoxic effect described for MazF in mammalian cells (Shimazu et al. 2007) and therefore deserves further investigation.

In a later work, MazF-transduced CD4+ T cells were infused autologously in monkeys (Chono et al. 2011b). These cells persisted for very long time in animals, and they did not develop antibodies against MazF, suggesting that the approach may constitute a suitable therapeutic approach against HIV. Supporting this, cells harvested from treated monkey more than 6 months post-infusion still suppressed the replication of SHIV ex vivo. All this supports that a MazF-based CD4+ T cell therapy may help the immune system to maintain a stable condition in HIV infected patients. As mentioned above understanding why mRNAs escape from MazF cleavage in these cells requires further investigation. This may happen because MazF is not expressed strongly enough upon infection. Accordingly, the

mRNA interferase could not be detected in cells transduced with HIV in the experiments above. However, detection was possible when these cells were transfected with retroviral vector overexpressing Tat. A multiplicity of infection (MOI) of 0.01 was used to transduce HIV in the experiments above. This may allow enough MazF expression to degrade HIV RNA, but not cellular mRNAs. Once HIV RNA is degraded, no more Tat (and therefore no more MazF) is produced, and the system is reset. It will be important to establish whether viral amplification taking place in cells of infected patients would lead to higher MOI than those tested in the work above, and whether this would increase MazF expression in cells to an extent that induces their death, If so, the population of infused mazF-CD4+ T cells may decrease more rapidly than expected, which would be detrimental from a therapeutic point of view.

Acknowledgments Work in Dr. de la Cueva-Méndez's laboratory is supported by Fundación Pública Andaluza Progreso y Salud, which depends on the Consejeria de Salud y Bienestar Social of the Junta de Andalucia.

References

Agarwal, S., Mishra, N. K., Bhatnagar, S., & Bhatnagar, R. (2010). PemK toxin of *Bacillus anthracis* is a ribonuclease: An insight into its active site, structure, and function. *Journal of Biological Chemistry, 285*, 7254–7270.

Alekshun, M. N., & Levy, S. B. (2007). Molecular mechanisms of antibacterial multidrug resistance. *Cell, 128*, 1037–1050.

Amitai, S., Yassin, Y., & Engelberg-Kulka, H. (2004). MazF-mediated cell death in *Escherichia coli*: A point of no return. *Journal of Bacteriology, 186*, 8295–8300.

Arase, K., Saijo, K., Watanabe, H., Konno, A., Arase, H., & Saito, T. (1999). Ablation of a specific cell population by the replacement of a uniquely expressed gene with a toxin gene. *Proceedings of the National Academy of Sciences of the United States of America, 96*, 9264–9268.

Bai, H., Sang, G., You, Y., Xue, X., Zhou, Y., Hou, Z., et al. (2012). Targeting RNA polymerase primary sigma70 as a therapeutic strategy against methicillin-resistant *Staphylococcus aureus* by antisense peptide nucleic acid. *PLoS ONE, 7*, e29886.

Bernard, P. (1995). New ccdB positive-selection cloning vectors with kanamycin or chloramphenicol selectable markers. *Gene, 162*, 159–160.

Bernard, P. (1996). Positive selection of recombinant DNA by CcdB. *BioTechniques, 21*, 320–323.

Bernard, P., & Couturier, M. (1992). Cell killing by the F plasmid CcdB protein involves poisoning of DNA-topoisomerase II complexes. *Journal of Molecular Biology, 226*, 735–745.

Bernard, P., Gabant, P., Bahassi, E. M., & Couturier, M. (1994). Positive-selection vectors using the F plasmid ccdB killer gene. *Gene, 148*, 71–74.

Carattoli, A. (2009). Resistance plasmid families in Enterobacteriaceae. *Antimicrobial Agents and Chemotherapy, 53*, 2227–2238.

Chono, H., Matsumoto, K., Tsuda, H., Saito, N., Lee, K., Kim, S., et al. (2011a). Acquisition of HIV-1 resistance in T lymphocytes using an ACA-specific *E. coli* mRNA interferase. *Human Gene Therapy, 22*, 35–43.

Chono, H., Saito, N., Tsuda, H., Shibata, H., Ageyama, N., Terao, K., et al. (2011b). In vivo safety and persistence of endoribonuclease gene-transduced CD4+ T cells in cynomolgus macaques for HIV-1 gene therapy model. *PLoS ONE, 6*, e23585.

Chopra, N., Agarwal, S., Verma, S., Bhatnagar, S., & Bhatnagar, R. (2011). Modeling of the structure and interactions of the *B. anthracis* antitoxin, MoxX: Deletion mutant studies

highlight its modular structure and repressor function. *Journal of Computer-Aided Molecular Design, 25*, 275–291.

Christensen, S. K., Maenhaut-Michel, G., Mine, N., Gottesman, S., Gerdes, K., & Van Melderen, L. (2004). Overproduction of the Lon protease triggers inhibition of translation in *Escherichia coli*: Involvement of the yefM-yoeB toxin–antitoxin system. *Molecular Microbiology, 51*, 1705–1717.

Davies, J. (2007). Microbes have the last word. A drastic re-evaluation of antimicrobial treatment is needed to overcome the threat of antibiotic-resistant bacteria. *EMBO Reports, 8*, 616–621.

de la Cueva-Mendez, G., Mills, A. D., Clay-Farrace, L., Diaz-Orejas, R., & Laskey, R. A. (2003). Regulatable killing of eukaryotic cells by the prokaryotic proteins Kid and Kis. *EMBO Journal, 22*, 246–251.

de la Cueva-Mendez, G., & Pimentel, B. (2007). Gene and cell survival: lessons from prokaryotic plasmid R1. *EMBO Reports, 8*, 458–464.

Faridani, O. R., Nikravesh, A., Pandey, D. P., Gerdes, K., & Good, L. (2006). Competitive inhibition of natural antisense Sok-RNA interactions activates Hok-mediated cell killing in *Escherichia coli*. *Nucleic Acids Research, 34*, 5915–5922.

Fiebig, A., Castro Rojas, C. M., Siegal-Gaskins, D., & Crosson, S. (2010). Interaction specificity, toxicity and regulation of a paralogous set of ParE/RelE-family toxin–antitoxin systems. *Molecular Microbiology, 77*, 236–251.

Gabant, P., Van Reeth, T., Dreze, P. L., Faelen, M., Szpirer, C., & Szpirer, J. (2000). New positive selection system based on the parD (kis/kid) system of the R1 plasmid. *BioTechniques, 28*, 784–788.

Gerdes, K., Christensen, S. K., & Lobner-Olesen, A. (2005). Prokaryotic toxin–antitoxin stress response loci. *Nature Reviews Microbiology, 3*, 371–382.

Gerdes, K., Rasmussen, P. B., & Molin, S. (1986). Unique type of plasmid maintenance function: Postsegregational killing of plasmid-free cells. *Proceedings of the National Academy of Sciences of the United States of America, 83*, 3116–3120.

Giuliodori, A. M., Di Pietro, F., Marzi, S., Masquida, B., Wagner, R., Romby, P., et al. (2010). The cspA mRNA is a thermosensor that modulates translation of the cold-shock protein CspA. *Molecular Cell, 37*, 21–33.

Goff, S. A., & Goldberg, A. L. (1987). An increased content of protease La, the lon gene product, increases protein degradation and blocks growth in *Escherichia coli*. *Journal of Biological Chemistry, 262*, 4508–4515.

Goldstein, J., Pollitt, N. S., & Inouye, M. (1990). Major cold shock protein of *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America, 87*, 283–287.

Guglielmini, J., & Van Melderen, L. (2011). Bacterial toxin–antitoxin systems: Translation inhibitors everywhere. *Mobile Genetic Elements, 1*, 283–290.

Heath, R. J., White, S. W., & Rock, C. O. (2001). Lipid biosynthesis as a target for antibacterial agents. *Progress in Lipid Research, 40*, 467–497.

Jaffe, A., Ogura, T., & Hiraga, S. (1985). Effects of the ccd function of the F plasmid on bacterial growth. *Journal of Bacteriology, 163*, 841–849.

Jensen, L. B., Garcia-Migura, L., Valenzuela, A. J., Lohr, M., Hasman, H., & Aarestrup, F. M. (2010). A classification system for plasmids from enterococci and other Gram-positive bacteria. *Journal of Microbiol Methods, 80*, 25–43.

Jones, P. G., VanBogelen, R. A., & Neidhardt, F. C. (1987). Induction of proteins in response to low temperature in *Escherichia coli*. *Journal of Bacteriology, 169*, 2092–2095.

Kamada, K., Hanaoka, F., & Burley, S. K. (2003). Crystal structure of the MazE/MazF complex: Molecular bases of antidote-toxin recognition. *Molecular Cell, 11*, 875–884.

Kim, H. G., Hwang, H. J., Kim, M. S., Lee, D. Y., Chung, S. K., Lee, J. M., et al. (2004). pTOC-KR: A positive selection cloning vector based on the ParE toxin. *BioTechniques, 36*(60–62), 64.

Koga, M., Otsuka, Y., Lemire, S., & Yonesaki, T. (2011). *Escherichia coli* rnlA and rnlB compose a novel toxin–antitoxin system. *Genetics, 187*, 123–130.

Koprunner, M., Thisse, C., Thisse, B., & Raz, E. (2001). A zebrafish nanos-related gene is essential for the development of primordial germ cells. *Genes & Development, 15*, 2877–2885.

Kristoffersen, P., Jensen, G. B., Gerdes, K., & Piskur, J. (2000). Bacterial toxin–antitoxin gene system as containment control in yeast cells. *Applied and Environment Microbiology, 66*, 5524–5526.

Leplae, R., Geeraerts, D., Hallez, R., Guglielmini, J., Dreze, P., & Van Melderen, L. (2011). Diversity of bacterial type II toxin–antitoxin systems: A comprehensive search and functional analysis of novel families. *Nucleic Acids Research, 39*, 5513–5525.

Leung, E., Datti, A., Cossette, M., Goodreid, J., McCaw, S. E., Mah, M., et al. (2011). Activators of cylindrical proteases as antimicrobials: Identification and development of small molecule activators of ClpP protease. *Chemistry & Biology, 18*, 1167–1178.

Levy, S. B., & Marshall, B. (2004). Antibacterial resistance worldwide: Causes, challenges and responses. *Nature Medicine, 10*, S122–S129.

Lioy, V. S., Machon, C., Tabone, M., Gonzalez-Pastor, J. E., Daugelavicius, R., Ayora, S., et al. (2012). The zeta toxin induces a set of protective responses and dormancy. *PLoS ONE, 7*, e30282.

Lioy, V. S., Martin, M. T., Camacho, A. G., Lurz, R., Antelmann, H., Hecker, M., et al. (2006). pSM19035-encoded zeta toxin induces stasis followed by death in a subpopulation of cells. *Microbiology, 152*, 2365–2379.

Lioy, V. S., Rey, O., Balsa, D., Pellicer, T., & Alonso, J. C. (2010). A toxin–antitoxin module as a target for antimicrobial development. *Plasmid, 63*, 31–39.

Maisonneuve, E., Shakespeare, L. J., Jorgensen, M. G., & Gerdes, K. (2011). Bacterial persistence by RNA endonucleases. *Proceedings of the National Academy of Sciences of the United States of America, 108*, 13206–13211.

Mao, L., Inoue, K., Tao, Y., Montelione, G. T., McDermott, A. E., & Inouye, M. (2011). Suppression of phospholipid biosynthesis by cerulenin in the condensed single-protein-production (cSPP) system. *Journal of Biomolecular NMR, 49*, 131–137.

Mao, L., Tang, Y., Vaiphei, S. T., Shimazu, T., Kim, S. G., Mani, R., et al. (2009). Production of membrane proteins for NMR studies using the condensed single protein (cSPP) production system. *Journal of Structural and Functional Genomics, 10*, 281–289.

Mao, L., Vaiphei, S. T., Shimazu, T., Schneider, W. M., Tang, Y., Mani, R., et al. (2010). The *E. coli* single protein production system for production and structural analysis of membrane proteins. *Journal of Structural and Functional Genomics, 11*, 81–84.

Martinez, J. L., & Baquero, F. (2002). Interactions among strategies associated with bacterial infection: Pathogenicity, epidemicity, and antibiotic resistance. *Clinical Microbiology Reviews, 15*, 647–679.

Meng, J., Kanzaki, G., Meas, D., Lam, C. K., Crummer, H., Tain, J., et al. (2012). A genome-wide inducible phenotypic screen identifies antisense RNA constructs silencing *Escherichia coli* essential genes. *FEMS Microbiology Letters, 329*, 45–53.

Miki, T., Yoshioka, K., & Horiuchi, T. (1984). Control of cell division by sex factor F in *Escherichia coli*. I. The 42.84-43.6 F segment couples cell division of the host bacteria with replication of plasmid DNA. *Journal of Molecular Biology, 174*, 605–625.

Moritz, E. M., & Hergenrother, P. J. (2007). Toxin–antitoxin systems are ubiquitous and plasmid-encoded in vancomycin-resistant enterococci. *Proceedings of the National Academy of Sciences of the United States of America, 104*, 311–316.

Nehlsen, K., Herrmann, S., Zauers, J., Hauser, H., & Wirth, D. (2010). Toxin–antitoxin based transgene expression in mammalian cells. *Nucleic Acids Research, 38*, e32.

Nikaido, H. (2009). Multidrug resistance in bacteria. *Annual Review of Biochemistry, 78*, 119–146.

Norman, A., Hansen, L. H., & Sorensen, S. J. (2009). Conjugative plasmids: Vessels of the communal gene pool. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 364*, 2275–2289.

Ogura, T., & Hiraga, S. (1983). Mini-F plasmid genes that couple host cell division to plasmid proliferation. *Proceedings of the National Academy of Sciences of the United States of America, 80*, 4784–4788.

Okeke, I. N., Laxminarayan, R., Bhutta, Z. A., Duse, A. G., Jenkins, P., O'Brien, T. F., et al. (2005). Antimicrobial resistance in developing countries. Part I: Recent trends and current status. *The Lancet Infectious Diseases, 5*, 481–493.

Overgaard, M., Borch, J., & Gerdes, K. (2009). RelB and RelE of *Escherichia coli* form a tight complex that represses transcription via the ribbon-helix-helix motif in RelB. *Journal of Molecular Biology, 394*, 183–196.

Pandey, D. P., & Gerdes, K. (2005). Toxin–antitoxin loci are highly abundant in free-living but lost from host-associated prokaryotes. *Nucleic Acids Research, 33*, 966–976.

Park, J. H., Yamaguchi, Y., & Inouye, M. (2012). Intramolecular regulation of the sequence-specific mRNA interferase activity of MazF fused to a MazE fragment with a linker cleavable by specific proteases. *Applied and Environment Microbiology, 78*, 3794–3799.

Peleg, A. Y., & Hooper, D. C. (2010). Hospital-acquired infections due to Gram-negative bacteria. *New England Journal of Medicine, 362*, 1804–1813.

Peubez, I., Chaudet, N., Mignon, C., Hild, G., Husson, S., Courtois, V., et al. (2010). Antibiotic-free selection in *E. coli*: New considerations for optimal design and improved production. *Microbial Cell Factories, 9*, 65.

Pimentel, B., Madine, M. A., & de la Cueva-Mendez, G. (2005). Kid cleaves specific mRNAs at UUACU sites to rescue the copy number of plasmid R1. *EMBO Journal, 24*, 3459–3469.

Qing, G., Ma, L. C., Khorchid, A., Swapna, G. V., Mal, T. K., Takayama, M. M., et al. (2004). Cold-shock induced high-yield protein production in *Escherichia coli*. *Nature Biotechnology, 22*, 877–882.

Raskatov, J. A., Meier, J. L., Puckett, J. W., Yang, F., Ramakrishnan, P., & Dervan, P. B. (2012). Modulation of NF-kappaB-dependent gene transcription using programmable DNA minor groove binders. *Proceedings of the National Academy of Sciences of the United States of America, 109*, 1023–1028.

Reckel, S., Lohr, F., & Dotsch, V. (2005). In-cell NMR spectroscopy. *ChemBioChem, 6*, 1601–1606.

Roberts, R. C., Spangler, C., & Helinski, D. R. (1993). Characteristics and significance of DNA binding activity of plasmid stabilization protein ParD from the broad host-range plasmid RK2. *Journal of Biological Chemistry, 268*, 27109–27117.

Roman, G., Endo, K., Zong, L., & Davis, R. L. (2001). P[Switch], a system for spatial and temporal control of gene expression in *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences of the United States of America, 98*, 12602–12607.

Sass, P., Josten, M., Famulla, K., Schiffer, G., Sahl, H. G., Hamoen, L., et al. (2011). Antibiotic acyldepsipeptides activate ClpP peptidase to degrade the cell division protein FtsZ. *Proceedings of the National Academy of Sciences of the United States of America, 108*, 17474–17479.

Schneider, W. M., Inouye, M., Montelione, G. T., & Roth, M. J. (2009). Independently inducible system of gene expression for condensed single protein production (cSPP) suitable for high efficiency isotope enrichment. *Journal of Structural and Functional Genomics, 10*, 219–225.

Shao, Y., Harrison, E. M., Bi, D., Tai, C., He, X., Ou, H. Y., et al. (2011). TADB: A web-based resource for Type 2 toxin–antitoxin loci in bacteria and archaea. *Nucleic Acids Research, 39*, D606–D611.

Shapira, A., Shapira, S., Gal-Tanamy, M., Zemel, R., Tur-Kaspa, R., & Benhar, I. (2012). Removal of hepatitis C virus-infected cells by a zymogenized bacterial toxin. *PLoS ONE, 7*, e32320.

Shimazu, T., Degenhardt, K., Nur, E. K. A., Zhang, J., Yoshida, T., Zhang, Y., et al. (2007). NBK/BIK antagonizes MCL-1 and BCL-XL and activates BAK-mediated apoptosis in response to protein synthesis inhibition. *Genes & Development, 21*, 929–941.

Slanchev, K., Stebler, J., de la Cueva-Mendez, G., & Raz, E. (2005). Development without germ cells: The role of the germ line in zebrafish sex differentiation. *Proceedings of the National Academy of Sciences of the United States of America, 102*, 4074–4079.

Summers, D. K., & Sherratt, D. J. (1984). Multimerization of high copy number plasmids causes instability: CoIE1 encodes a determinant essential for plasmid monomerization and stability. *Cell, 36*, 1097–1103.

Suzuki, M., Zhang, J., Liu, M., Woychik, N. A., & Inouye, M. (2005). Single protein production in living cells facilitated by an mRNA interferase. *Molecular Cell, 18*, 253–261.

Szpirer, C. Y., & Milinkovitch, M. C. (2005). Separate-component-stabilization system for protein and DNA production without the use of antibiotics. *BioTechniques, 38*, 775–781.

Terwilliger, T. C., Stuart, D., & Yokoyama, S. (2009). Lessons from structural genomics. *Annual Review of Biophysics, 38*, 371–383.

Tzika, A. C., Rosa, S. F., Fabiani, A., Snell, H. L., Snell, H. M., Marquez, C., et al. (2008). Population genetics of Galapagos land iguana (genus Conolophus) remnant populations. *Molecular Ecology, 17*, 4943–4952.

Vaiphei, S. T., Tang, Y., Montelione, G. T., & Inouye, M. (2011). The use of the condensed single protein production system for isotope-labeled outer membrane proteins, OmpA and OmpX in *E. coli*. *Molecular Biotechnology, 47*, 205–210.

Van Melderen, L., & Saavedra De Bast, M. (2009). Bacterial toxin–antitoxin systems: More than selfish entities? *PLoS Genetics, 5*(3), .

Vandermeulen, G., Marie, C., Scherman, D., & Preat, V. (2011). New generation of plasmid backbones devoid of antibiotic resistance marker for gene therapy trials. *Molecular Therapy, 19*, 1942–1949.

Vieira, J., & Messing, J. (1982). The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene, 19*, 259–268.

Walhout, A. J., Temple, G. F., Brasch, M. A., Hartley, J. L., Lorson, M. A., van den Heuvel, S., et al. (2000). GATEWAY recombinational cloning: application to the cloning of large numbers of open reading frames or ORFeomes. *Methods in Enzymology, 328*, 575–592.

Wang, X., & Wood, T. K. (2011). Toxin–antitoxin systems influence biofilm and persister cell formation and the general stress response. *Applied and Environment Microbiology, 77*, 5577–5583.

Watts, A. (2005). Solid-state NMR in drug design and discovery for membrane-embedded targets. *Nature Reviews Drug Discovery, 4*, 555–568.

Williams, J. J., Halvorsen, E. M., Dwyer, E. M., DiFazio, R. M., & Hergenrother, P. J. (2011). Toxin–antitoxin (TA) systems are prevalent and transcribed in clinical isolates of *Pseudomonas aeruginosa* and methicillin-resistant *Staphylococcus aureus*. *FEMS Microbiology Letters, 322*, 41–50.

Williams, J. J., & Hergenrother, P. J. (2012). Artificial activation of toxin–antitoxin systems as an antibacterial strategy. *Trends in Microbiology, 20*, 291–298.

Xia, B., Ke, H., Jiang, W., & Inouye, M. (2002). The Cold Box stem-loop proximal to the 5'-end of the *Escherichia coli* cspA gene stabilizes its mRNA at low temperature. *Journal of Biological Chemistry, 277*, 6005–6011.

Yamaguchi, Y., & Inouye, M. (2011). Regulation of growth and death in *Escherichia coli* by toxin–antitoxin systems. *Nature Reviews Microbiology, 9*, 779–790.

Yamamoto, T. A., Gerdes, K., & Tunnacliffe, A. (2002). Bacterial toxin RelE induces apoptosis in human cells. *FEBS Letters, 519*, 191–194.

Yang, M., Gao, C., Wang, Y., Zhang, H., & He, Z. G. (2010). Characterization of the interaction and cross-regulation of three *Mycobacterium tuberculosis* RelBE modules. *PLoS ONE, 5*, e10672.

Zhang, Y., Zhang, J., Hoeflich, K. P., Ikura, M., Qing, G., & Inouye, M. (2003). MazF cleaves cellular mRNAs specifically at ACA to block protein synthesis in *Escherichia coli*. *Molecular Cell, 12*, 913–923.

Zhu, L., Sharp, J. D., Kobayashi, H., Woychik, N. A., & Inouye, M. (2010). Noncognate *Mycobacterium tuberculosis* toxin–antitoxins can physically and functionally interact. *Journal of Biological Chemistry, 285*, 39732–39738.

Index

K. Gerdes (ed.), *Prokaryotic Toxin-Antitoxins*,
DOI: 10.1007/978-3-642-33253-1, © Springer-Verlag Berlin Heidelberg 2013