The GroE chaperonin machine is the major modulator of the CIRCE heat shock regulon of *Bacillus subtilis*

WOLFGANG SCHUMANN*, GEORG HOMUTH and AXEL MOGK

Institute of Genetics, University of Bayreuth, D-95440 Bayreuth, Germany *Corresponding author (Fax, 49-921-552710; Email, wolfgang.schumann@uni-bayreuth.de).

Regulation of the heat shock response in bacteria has been studied extensively in Escherichia coli where heat shock genes are classified into three classes and where each class is regulated by a different alternate sigma factor. Bacillus subtilis serves as a second model bacterium to study regulation of the heat shock response in detail. Here, four classes of heat shock genes have been described so far where two are controlled by two different repressor proteins and the third by the alternate sigma factor σ^{B} . Class I heat shock genes consists of two operons, the heptacistronic dnak and the bicistronic groE operon. Transcription of the dnak operon is complex involving two promoters, premature termination of transcription, mRNA processing and different stabilities of the processed transcripts to ensure the appropriate amounts of heat shock proteins under different growth conditions. The translation product of the hrcA gene, the first gene of the dnaK operon, binds to an operator designated CIRCE element, and its activity is modulated by the GroE chaperonin system. We assume that the HrcA protein, upon de novo synthesis and upon dissociation from its operator, is present in an inactive form and has to be activated by the GroE chaperonin system resulting in an HrcA-GroE reaction cycle. Induction of class I heat shock genes occurs by the appearance of denatured proteins within the cytoplasm which titrate the GroE system. This results in accumulation of inactive HrcA repressor and thereby in induction of class I heat shock genes. Upon removal of the non-native proteins from the cytoplasm, the GroE chaperonin will interact with HrcA and promote folding into its active conformation resulting in turning off of class I heat shock genes. This mechanism ensures adequate adjustment of class I heat shock proteins depending on their actual need.

1. Introduction

The heat shock response involves the transient increased production of a set of proteins designated heat shock proteins (Hsps). It occurs in all organisms studied so far upon thermal upshock or exposure to other stress factors which will lead to the accumulation of denatured proteins. Most Hsps act either as molecular chaperones or as proteases which protect the cells from stress-induced damage by preventing denaturation of cellular proteins, reactivating once-inactivated proteins, and regulating the degradation of irreversibly denatured proteins (Parsell and Lindquist 1993). Expression of all heat shock genes is controlled at the level of transcription by an heat shock factor which is either an alternate sigma factor or a repressor in prokaryotes or a transcriptional activator in eukaryotes. The activity of the heat shock factor is modulated by one or more Hsps through direct proteinprotein interaction.

The best-studied prokaryotic system is the Gramnegative bacterium *Escherichia coli* where three classes of heat shock genes are regulated by three different alternate sigma factors (for a recent review, see Missiakas *et al* 1996). Besides alternate sigma factors, the HrcA repressor is another important heat shock factor which has been discovered in the Gram-positive bacterium *Bacillus subtilis*. In this article, we will shortly review regulation of the three classes of heat shock genes in *E. coli* followed by a detailed description of the *dnaK* operon transcription and of the modulation of the HrcA activity.

Keywords. Heat shock response; CIRCE; HrcA; GroEL; DnaK; repressor

J. Biosci., 23, No. 4, October 1998, pp 415-422. © Indian Academy of Sciences

2. The E. coli heat shock stimulon

In E. coli three classes of heat shock genes, two major and one minor class, have been described which are all positively controlled at the transcriptional level by three different alternative sigma factors thereby constituting two regulons and one operon designated the heat shock stimulon (table 1; Missiakas et al 1996). The two major classes are induced upon accumulation of denatured proteins in either the cytoplasm or in the periplasm. Class I heat shock genes are under the control of σ^{32} and form the largest heat shock regulon of E. coli with some 35 members also called the sigma-32 regulon. The activity of σ^{32} is modulated by the DnaK chaperone machine (Gamer et al 1992; Liberek et al 1992; Liberek and Georgopoulos 1993). A detailed description of the regulation of the genes of this regulon is presented by Nakahigashi et al 1998.

Class II heat shock genes form the sigma-E regulon, and this regulon is induced when non-native proteins accumulate within the periplasm. There are several possibilities how denatured proteins might be formed: (i) severe heat stress (>45°C); (ii) overproduction of outer membrane proteins; (iii) mutations in genes whose products act as protein folding catalysts, e.g., *surA* encoding a PPIase (Rouvière and Gross 1996). Under physiological conditions, σ^{E} is sequestered by an anti-sigma factor which is an integral inner membrane protein. Upon accumulation of denatured proteins within the periplasm, σ^{E} is released from its anti-sigma factor, interacts with the RNA polymerase core enzyme and directs expression of about ten genes of the sigma-E regulon. Shut-off

 Table 1. The E. coli heat shock stimulon consists of three classes of heat shock genes which are regulated by three different alternate sigma factors.

Class	Sigma factor	Modulator	Number of heat shock genes
I	o ³²	DnaK chaperone machine	Some 35
II	$\sigma^{\rm E}$ (σ^{24})	Anti-sigma factor	About 10
III	0 ³⁴	?	pspABCDE

occurs upon removal of the non-native proteins from the periplasm and rebinding of $\sigma^{\rm E}$ to the anti-sigma factor. Two additional proteins are involved in these reactions, but their roles during the modulation of the activities of $\sigma^{\rm E}$ and its anti-sigma fcator are not well understood (Missiakas *et al* 1997; De Las Peñas *et al* 1997).

As already mentioned, class III heat shock genes consist of only one single operon, the *psp* operon (for a recent review, see Model *et al* 1997). This operon has been identified as being induced after infection with phages containing single-stranded DNA such as M13 or fd. Later, it could be shown that production of gene IV protein encoded by these phages is sufficient to induce this operon (*psp* stands for *phage shock protein*), which is under the positive control by o^{54} . Under physiological conditions, expression of this operon is prevented at the level of transcription by PspA acting as a repressor.

3. The B. subtilis heat shock stimulon

In *B. subtilis*, four different classes of heat shock genes have been described so far (table 2). While those of class I and III are under negative control by two different repressors, class II heat shock genes are positively controlled by the alternate sigma-factor σ^{B} . All those heat shock genes which neither belong to class I, nor to class II or III have been classified into class IV. Class IV genes are expressed from the vegetative promoter P_{A} , and *htpG* might be under negative control by a third repressor which is still unknown (Schulz *et al* 1997). This review will report in detail on the regulation of class I heat shock genes.

4. Transcriptional organization of the *dnaK* and *groE* operons

Class I heat shock genes are composed of nine genes organized in two operons, the *dnaK* and the *groE* operon, forming the CIRCE regulon. Both operons are preceded by vegetative promoters P_A followed by a perfect 9 bp inverted repeat which has been designated CIRCE element and functions as an operator (Zuber and Schumann 1994). Under physiological conditions, expression of both operons is partially repressed due to the product of the

Table 2. The *B. subtilis* heat shock stimulon consists of at least four different classes of heat shock genes which are regulated by different mechanisms.

Class	Promoter(s)	Regulation type	Regulatory sequence	Modulator	Heat shock genes
I	P_{A}	Repressor: HrcA	CIRCE	GroE chaperonin	dnaK, groE operons
II	P_{B}	Sigma-B	Promoter	Anti-sigma factor	50–100 genes
III	P_{A} or/and P_{B}	Repressor: CtsR	CtsR box	?	clpP
IV	P_{A}	Repressor?	?	?	clpX, lonA, lonB, htpG, ftsH, sigW, htrA, ahpC



Figure 1. Transcriptional organization of the *dnaK* operon of *B. subtilis.* Transcription of the heptacistronic *dnaK* operon starts at two different promoters, designated P_c and P, as indicated by the arrows; the thickness of the arrows is indicative of the relative amount of the different transcripts. Most of the transcripts initiated at P_c will be terminated at a potential ρ -independent terminator located between *dnaK* and *dnaJ* (stem-loop structure 2); these transcripts will be processed at a site located between *hrcA* and *grpE* (stem-loop structure 1). The stem-loop structures 3 through 5 indicate potential degradation sites, 2 a processing site and 6 another putative ρ -independent terminator.

first gene of the *dnaK* operon, the HrcA protein, which binds to the CIRCE element. Consequently, initiation of transcription at the promoters upstream of the *dnaK* and *groE* operons is partially prevented (Yuan and Wong 1995a,b; Schulz and Schumann 1996; Mogk *et al* 1997).

The groE operon consists of the two genes groES and groEL, encoding the Hsp60 chaperonin and its Hsp10 co-chaperonin, respectively; both genes are transcribed as a bicistronic mRNA (Schmidt et al 1992). The dnaK operon is heptacistronic and contains the genes hrcA, grpE, dnaK, dnaJ, orf35, orf28, and orf50 (figure 1). While genes grpE, dnaK and dnaJ code for the DnaK chaperone machine, the functions of the three distal genes are unknown. Knock-outs in these three latter genes neither evoke a temperature-sensitive phenotype nor do they influence the regulation of the heat shock response of the CIRCE regulon (Homuth et al 1997).

5. Transcription of the dnaK operon

Transcription of the *dnaK* operon is rather complex. The operon specifies three primary transcripts of 8.0, 3.6 and 4.3 kb. Two of these mRNAs (8.0 and 3.6 kb) are processed leading to processing products of 7.0, 1.0 and 2.6 kb. The amount of primary transcripts and processing products within the cells differs before and after thermal upshock. In addition, the stability of the different mRNA species varies considerably (figure 1; Homuth *et al* 1997; our unpublished results). The strongly heat-inducible 3.6 kb primary transcript initiated at the CIRCE-controlled promoter in front of *hrcA* is terminated at a potential

 ρ -independent terminator structure located between dnaKand dnaJ. This transcript is highly unstable, and its processing leads to the formation of two products of 1.0 and 2.6 kb. Furthermore, the CIRCE element located near the 5' end of this mRNA decreases the stability of the transcript as published first for the groE mRNA (Yuan and Wong 1995a,b). Our analyses of the dnaK transcripts confirmed their results. The bicistronic 2.6 kb dnaK-grpE processing product exists within the cells in large amounts already at physiological temperatures and is only weakly induced after a heat shock. This observation can be explained by the fact that the processing event taking place between hrcA and grpE occurs with a constant, but low rate. The bicistronic 2.6 kb grpE-dnaK processing product represents the most stable mRNA species of the whole dnaK operon (our unpublished results). The monocistronic 1.0 kb hrcA processing product is present within the cell in only small amounts before heat shock and is strongly induced afterwards. This induction can be explained by enhanced novel mRNA synthesis and, in addition, by a transient stabilization of this highly unstable processing product (unpublished results).

Partial readthrough at the terminator structure between dnaK and dnaJ leads to the appearance of two large mRNA species where the 7.0 kb mRNA molecule represents the processed form of the 8.0 kb primary transcript (figure 1). Both mRNAs are highly unstable, independent of the CIRCE element and most probably due to endoribonuclease-recognition sites near three potential secondary structures as shown in figure 1 (stem-loop structures designated 3, 4 and 5). While the transcripts initiated upstream of the CIRCE-controlled σ^A promoter are heat-inducible, the second σ^A -dependent promoter upstream of the four distal genes is constitutive (Homuth et al 1997).

The functional consequences of this transcriptional organization for expression of the different genes of the dnaK operon can be summarized as follows: Under physiological conditions, the hrcA gene is expressed at a low level being consistent with its function as a repressor. After thermal upshock, this gene is strongly induced; quantitative slot-blot analyses revealed an induction factor of 10-11 for the hrcA mRNA (figure 2) which is accompanied by a concomitant increase in the amount of HrcA protein by a factor of about ten (Mogk et al 1997). Enhanced amounts of HrcA are needed for switch-off of the heat shock response. The DnaK and GrpE proteins being components of the DnaK chaperone machine are present in large amounts already under physiological conditions; this is ensured by the appearance of the stable 2.6 kb processing product. After heat shock, the total amount of dnaK- and grpE-specific mRNA increases by a factor of six. The DnaJ protein represents the third component of the DnaK chaperone system. But



Figure 2. Schematic representation of the changes in mRNA level of the 7 genes of the *dnaK* operon before (0 min) and at different times after (5, 10, 15, and 30 min) heat shock from 37 to 48° C. Luminographs of slot-blot experiments were quantified with the WinCam software version 2.1 of Cybertech, Berlin, Germany.

in comparison to DnaK, it is needed in 3-fold lower amounts. This molar ratio is fine-tuned by partial termination between *dnaK* and *dnaJ* and by low stability of the *dnaJ*-encoding mRNA species; the *dnaJ* mRNA is 3-fold heat-induced. The three genes downstream of *dnaJ* are induced in a way similar to *dnaJ* (figure 2).

6. The GroE chaperonin machine modulates the activity of the HrcA repressor

Class I heat shock genes are under the negative control by a transcriptional repressor encoded by the *hrcA* gene, the first gene of the *dnaK* operon (figure 1). That *hrcA* codes for the repressor has been first shown by Yuan and Wong (1995). They screened for mutations affecting regulation of the *groE* operon and mapped them in *hrcA*. Furthermore, they could demonstrate that crude extracts prepared from an *E. coli* strain overproducing HrcA were able to specifically retard the mobility of a DNA fragment containing the operator, the CIRCE element. At about the same time, Schulz and Schumann (1996) could show that *hrcA* encodes a negative regulator of class I heat shock genes.

Transcriptional analysis of the class I operons indicated that their basal level is increased immediately after thermal shock. To allow a transient high level of expression of the two operons, the repressor must dissociate from its operator and be prevented from rebinding for at least 5 min. How is HrcA transiently kept in an inactive form? In E. coli and also in lower and higher eukaryotes it has been shown that non-native proteins formed after thermal upshock constitute the inducing signal. Therefore, we assume that in B. subtilis, too, preventing the repressor from binding to its operator is the consequence of accumulation of denatured proteins within the cytoplasm. Will these non-native proteins directly interact with HrcA thereby promoting its dissociation and keeping it in an inactive form or will they rather act in an indirect way? As mentioned above, in E. coli non-native proteins within the cytoplasm are recognized by the DnaK chaperone system. Therefore, we first asked whether in B. subtilis this system is also involved. We isolated knock-outs in grpE, dnaK, and *dnaJ* and tested them for their influence on the regulation of class I heat shock genes. It turned out that neither the basal level nor the induction behaviour in these three



Figure 3. The HrcA-GroEL reaction cycle. We assume that newly synthesized HrcA polypeptide chains have to interact with the GroE system to gain their active conformation allowing them to bind to their operator. Upon dissociation from its operator, HrcA is present in an inactive form and has to bind to GroEL first to become reactivated.

mutant strains was altered. This clearly ruled out any influence of the *B. subtilis* DnaK chaperone system on the regulation of the *dnaK* and *groE* operons (Mogk *et al* 1997).

Since the groE operon is also part of the CIRCE regulon, we next asked whether the GroE chaperonin system might be involved in modulating the activity of HrcA. Since neither groES nor groEL can be inactivated (both genes are essential; Li and Wong 1992), we replaced its own promoter by a controlable promoter thereby allowing depletion of both proteins upon removal of the inducer (Mogk et al 1997). To easily monitor expression of the dnaK and groE operons, transcriptional fusions were constructed between either the promoter region of dnaK or groE and the reporter gene bgaB which codes for a heat-stable β -galactosidase. These operon fusions were crossed separately into the B. subtilis strain carrying the groE operon under the controlable promoter. It turned out that upon depletion for the GroE proteins, both the dnaK and the groE operons were already expressed at physiological temperatures at a level exceeding that in the wild-type strain after heat induction (Mogk et al 1997). These data suggested to us that GroE acts as a modulator of the CIRCE regulon. To confirm these results we reasoned that overproduction of GroE proteins should affect expression of the class I operons adversely. The groEL gene was, cloned in a high-copy-number plasmid under the control of an IPTGinducible promoter, and expression of the dnaK operon was followed in the absence and in the presence of IPTG using the two transcriptional fusions. While a normal heat shock response was seen in the absence of IPTG, both the basal and the heat-induced level were severely reduced in the presence of IPTG (Mogk et al 1997).

To prove that HrcA interacts with its operator, we overproduced the protein in E. coli where it formed inclusion bodies. These inclusion bodies could only be dissolved in the presence of a chaotropic agent such as guanidinium hydrochloride or urea. Upon removal of the chaotropic agent either by dilution or by dialysis, the HrcA protein aggregated spontaneously. Therefore, it was almost impossible to obtain a bandshift between HrcA and a DNA fragment carrying the CIRCE element. Since GroEL seems to modulate the activity of HrcA (the role of GroES is not clear at the moment), we assumed a physical interaction between GroEL and HrcA. Therefore, we asked whether purified GroEL could prevent aggregation of HrcA. Indeed, upon dilution of HrcA from a high concentration of guanidinium hydrochloride into a chaotropic agent-free buffer, addition of GroEL in stoichiometric amounts largely prevented the formation of aggregates. This finding prompted us to test whether GroEL would also help to increase the amount of retarded DNA in a gel mobility shift assay. Indeed, the amount of retarded DNA was enhanced from 1 to 12% upon addition of GroEL (Mogk et al 1997). We concluded from these in vitro experiments that there is a physical interaction between HrcA and GroEL.

7. The HrcA-GroEL reaction cycle

The in vivo experiments revealed that after depletion of GroE proteins, high amounts of HrcA repressor accumulated within the cells, but the repressor was unable to interact with its operator to achieve repression. The in vitro experiments have demonstrated that GroEL prevents HrcA from forming aggregates and might in addition be involved in correct folding of the repressor. In the light of these results, the following scenario can be proposed for the molecular mechanism underlying the modulation of the HrcA activity. Whether the repressor binds as a dimer as shown in figure 3 is not yet known, but the structure of the operator consisting of two inverted repeats separated by a 9 bp spacer suggests binding as a dimet (or higher oligomer) where both monomers are on the same phase of the helix. Upon de novo synthesis, HrcA has to interact with GroEL to adopt its active conformation enabling it to dimerize and to bind to its operator. Upon dissociation, HrcA is unable to rebind to the CIRCE element, it needs activation through GroEL first. This model easily explains how the heat shock response is regulated at the molecular level. After thermal upshock, the amount of non-native proteins suddenly increases, thereby titrating GroEL which is, therefore, unable to reactivate HrcA. This model also explains why in the absence of GroEL, large amounts of inactive HrcA protein accumulate within the cells. Furthermore, it is postulated that the relative activity of HrcA is modulated by the GroE system in response to the amount of

145

. L	n)	
At	MGFSA#LSKDQASLLDERSREIFRRIVEGYLDTGEPLG8RSLSRLLPMSL 5 PASV R NVMSDLEELGLIYSPHIS-	agrlptqtglrffvdafmqvgdl-paderanidrqigpv
Bj	MAHHDPIHLIAPRAGLAQLNERSRDIFRQIVESYLATGEPVGSRNISRLIAMPLSFASVRNVMADLEQLGLIYAPHTS-	AGRLPTELGLRFFVDALMQVGDL-NDAERQSIQSQLTSV
Bs	KALLERALERALERALERALERALERALERALERALERALE	SGRVPSEKGYRYYVDHLLSPVKL-TKSDLDQIHSIFKEK
Bst	MLTERQLLILQVIIDDFIRSGQPVGSRTLSKKHEIALSSATIRNEMADLEELGYIEKTHVS-	SGRVPSEKGYRYYVDHLLSPQRL-TKDDIQKIKSIFAER
Ca	MEMEERKLKILQAIINDYINNGEPVGSRTIAKKYNLGISSATIRNEMADLEEMGYIEQLHTS-	SGRKPSDKGYRLYVDRLMEIPSM-SVEEEMLIKAIIDSA
Cc	mtqlfpgpivrtpglaeldarardifrrvvesyletgepvgsrtiskgg-valspasirntmqdlaqlglldaphts	AGRMPTHAGLRMFVDGFLEVGDV-AEQEKRAIEARLAVK
Ct	MENRIEMSQLRASKKDSKISYVLLMATKLYLESGQPVGSKLLKETYCSDLSSATIRNYFAQLETDGFLRKNHIS-	GGRIPTDLAFRYYADHNAPFLEQEEILAIQQKLTELPEY
Hр	-MVIDEIFQIMMLRRIKVGSNLNKKESLLDAFVKTYLQILEPISSKRLKELÄDLKISCATIRNYFQILSKEGMLYQAHSS-	GARLPTFKAFENYWQKSLRFETLKVNEKRLK
Li	MDLTERHKRILKALVDKFIQENRPVG\$KTLFDKHDIGLSPASIRTVLKDLEDFGYLASKHTS-	GGRIPTERGYRFYVDSLVILYEL-TLKEKQRIQQEYLKM
Ŀ1	SLLDSIQASSATIRNDMKALERLGLIQKEHTS-	SGRIPSVSGYKYFVENVIQLEEF-SQNDLFKVMKAFDGD
Mg	MKNLTPRQAQILKAIINEYIAYAIPVGSKLLTKXYFKNLSGGTLRNEMAALEKKGFLKKNHIS-	SGRVPSQIGYQYYVKVLNVSNTT-NDLKTRLRSVILQQH
Mp	MKNLTTRQAQILKAIINEYIAYPVPVG\$KLLTKKYFKNL\$GGTLRNEMAVLEKEGYLKKNHIS-	SGRIPSQLGYQYYVKLLTKNDDK-SNLKTRLRAIILQKH
Mt	DERRFEVLRAIVADFVATQEPIGSKSLVERHNLGVSSATVRNDMAVLEAEGYITQPHTS-	SGRVPTEKGYREFVDRLEDVKPL-SSAERRAIQSFLESG
Sa	MITDRQLSILNAIVEDYVDFGQPVGSKTLIERHNLNVSPATIRNEMKQLEDLNYIEKTHS-	SGRSPSQLGFRYYVNRLLEQTSH-QKTNKLRRLNQLLVE
Sal	RALSERRIEVLRAIVQDYVGTEE-VG\$KALTERHRLGV\$PATVRNDMAALEDEGYIAQPHTS-	AGPIPTDKGYRLFVDRLADVKPM-TAPERRAIHHFLDQA
Sm	TLQNSIASSRATIKNDMAALEKLGLLEKATTPP	AVVCPVKKAIRYFVEHSLNPDSL-DEQDVYQVIKAFUFE
syn	MVKPLRLNDRHQQILRATVQHYIATAEPVGSHTLAQEYQFAV3SATIRNALGQLEKAGLLYQPHVS-	AGRVPSDSGYRIYVDNLLTWSDRQSRTVKQRLENEINGD
	* * * *	•
* +		
R4	AGREGSUEGULTEASKALSGNSKGAGULTA-KNDVILKAVSTINLETINLETINLETINGDANQVENKITELEAGISSSQ	ITEAANT INARLOGUT LUEDROUT QIQITELUSELG
Bs	G-GAGSVERALDGALLALSGLANAAVVLIE-ASNARNALTEVALDEVALDVLUVGEDGVVENAVLILEFGVESSA- TBILEMVUKSAOTI.SDIJIVYSTVI.GDKISKIVVI.KOTOTI DODIMAVITVIMPALVRINKTVIŠPKMDI.SD	TEKIMUTINDELSGVENDELNERTEKENMUTEORT
Bst	T	LEKMVNTENERLIGVELVDLKDKTYKKVADVLEKHT
Ca	LYEIDKLVKQAMSLVSEMTKLTCVVKSLSARKSYIKSISLINIEPMMILCVFITDSGMIKNSIIRVKSNIENSS	LERIANILNSKLKGLTIEOINLEVINNIKKDLREYG
Cc	GRSFEEALAEASSILSGLAGGAGIVVTP-VREGGVKHVEFVPLGGGQVLAVMVFEDGQVENRLMRQAPGVTPSA	LOEASNFLNARLRGRTLTEARTEMGGELDAARROLN
Ct	SKNIVKDLQKASEVLSDILQLPVCFSSPRFESDSVINIQLVAIDDQRVVFVLSTEFGQVFTDVLwLPEQLPENS	LKRIEGFLQNYLRKQPSDSLLSQKEEDLGMVLYN
Hр	SASENFGLFTLLKK#S-*LERLERVIECEKRFLILDFLAFSCALGYSVXMEK	FLLELVGRSVKEVRSIAASFNALSLARQ-
Li	QFKLDQILKATASVLSSLSNAAGIVTGPAKNLDTLKHIELIHVRGDEILMILVMRSGTVLHRNIFVDQNYSQEA	LYQVSKYLNDNLKGYDIYEIQNVIIP-KLMIRKDGP
L1	FYRLSDLFKTAAKSLSELTGLTSFVLNAPQRDQQLVSFEMVMLDNHSVLSVITLGTGEVRTNQFILPKSMTEAD	LAVFSNLVKERLVGKKVIDIHYTLRTEIPQIVQRYF
Mg	KTIDEVIELGVKFINEIIN-LEVVLTNFSSDEVLKKIDLIILDKSFALFLLVSASGKVFKKTISYANQRQFED	IVICVRIFNDRIIDTRFSEINNQLEVLKEIIRTKVH
мр	KTIDEIIELGVKFINEMVN-LPV/DTHFSSDEVLKKIDLIMLDQSCALLELVSASGNVFKKTISYANQRQFED	INVCVRLENDRIIDTRECDIAQELDVLKEIIRSKVR
MC So	VDEDUVLRRAVRELAQUTRQVAVQIPTESTSTVRILEVIALI PARELAVVITUSGRVDQRIVELGUVIDDAQ	INDIANEMININAFENONI ODDINAENOESOFFI INDIANEMININAFENONI ODDINAENOESOFFI
28 591	N-QIDVSSALIIIFADELSSALSQIIIEVNINAVSELTAGAMDIKALIKANPALVINVIVESSGAVERVALASDIFESMDA	
Sm	3	
Syn	NWHFEALLQRMGQILAGLSGYIALITFPQTETVQLRHLQLMLLPSHQILIILVTDSYHTHSATLDLPAAMEAKEEGEL	EQELAIFSNFLNAQLRGKNLSELSHLNWQELDQKFSIYA
(I	3)	
At	MLAQDLVERGLAIWAGDNEEGKLGRLIVRGRSNLLEGLAGEEDIDRVRMLFDDLERKEN-LIEILNLAE	SGSGVRIFIGSENKLFSLSGSSLIV
Вј	QLTQKVISAGIASWSGGENEDRQLIVRGHANLLEDLHALEDLERVRLLFDDLETKRG-VIDLLGRAE	LSGSSTII
Bs	KNYDNILDALRSTFHSTNHVEKLFFGGKINMLNQPE-FHDITRVRSLLSLIEKEQD-VLKLVQSP	HT-GISIKIGKENDYEEMENCSLIT
Bst	RNYDSMLQTIVETLDIP-QEEKMFFAGKTNMLNQPE-FNDIQKIRPLMKMIEQEKD-FYRLLRKHN	RK-GIQVTIGRENQLSEMENC5LIT
Ca	HIFDCIMPNLYDILREADSTEVYKEGTMNIFNYPE-FXDIEKAKEFLSVIDDRRILDTLFNAS	ARDF5VVS
Cc	ETAARLVEDGLAAMSGGEGDARSLIVRGQANLLADARAREDIDRVRQLFDDLEQKGQ-LIGLLDDVR	DAEGVRIYIGAETRLFSLSGSSVIA
Ct	EVVVRYLTRYCHFSEED	
Кр	LERLEYSNTQITRFNLMGLKTLLNSYLFFDILGGKVLERLSKG	DCMLVT
Li	EDFIRIADLISSAMT?DNSEVTUYIDGFKNLYANFRDEEQQLSQVLSLLDDQGF-LKAFFSEYI	DODGVFTIIGKDG-DR\$MSGVSIIT
Ll	KVTSEVLQLFESIFDDLFKEHUTVAGHKNIFDYATDNLAELYKLFSDDERMLH-BIREITNND	EMRAVKFDNDEKFMKNLTIIS
Ма	EYOYVIDEILFKLFDLDQIEANKKIYGIQYLAKOPE-FANQEKLTKILNLLEDTSV-WQQMAFINQ	EISVAS
Mp	EYQYVIDEILFKLFNFEEFQQARKQVYGIHYLAQQPE-FANQERLTRILNLLEDTSV-WQQMAFMNQ	E-EVSVAS
Mt	GLGDAVGRAATVEHTEERLLLG-GTANLTRNAADFGGSLRSILEALEEQVV-VLRLLQQEA	GKVTVRIGHETASEQMVGTSMVS
Sa	FINKLINTMNNHISNQ3NSIYMGGKVKLIDALN-ESNVSSIQPILQYIESNRIAELLQDIS	SP-NINVKIGNEIDDSLSDISIVT
Sal	PTVSTVLSTLLETLVEETEER-~~LMIG-GTANLTRFGHDFP~-LTIRPVLEALBEQVV-LLKLLGEAK~~	DS-AMTVRIGHENAHEGLSSTSVVS
Sm	PRTDNVLDLFDHIFNPIFQEEVFISGKIKTLEFAGLDTYQFLENLQSVAL-EIRQSLPED	ELHRVQVADSK-EKSLADLTVIS
Syn	DFLKGLQQQIKPLLQRRMAGPLVVHGVSKVIQQPE-FSQLBQVQMLLSLLEQEQDKLFSLLFDPDNYGDWLANLGQE *	MNLLTGETMPKTRPVVTIRIGAENPLESMHPCTLVS
•		
AC.	AP INDEGRAV VGAVGVIGTTKLNIAKI V KNUU TAQIMAKLSKKQR	
6) D(SPIRUAWALIVEVUEVIETRUNIARVIFIVUIAARIVSHLUGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
DS Dat	ASISYMAA YUGSIALIGETKNIISKYVSILUHYTSULSKALTSULUE	
081		

Bs	ASYSVDQK-QIGSIAIIGPTRMNYSRVVSLLQHVTSDLSKALTSLYDE
Bst	ATYSIGDE-OLGTIALLGPTRMEYSRVITILNRVASDLSIALTKWYQNG
Ca	SVYKYNGR-PLGTIGIIGPTRIFYSKVIKVIMEVVDQINNNLDKMNNS
Cc	APYMTGROKVLGAIGVIGPARLNYARVIPLVDYTARVLGRMMDG
Ct	IPYYMDRT-PLGTFGVLGPMNLPY00VFGTLSLFTERLKVILT0SFYKFKLSFRRPCPTDPRCS0RPAELTRSSSIKLLPAKELS
Hp	RPVEFONKRMOLLCVGKLECDYEGFFOTISEEE
Li	SNYRMGEK-KIGALGIIGPORMDYNRALPLVDFTSKLVSEMVTRISK
ГJ	OKFVIPYR-GFGTLTVVGPVEMDYORTLSVLDLVAKVLTMKLSDYYRYLDGNHYEISK
Ma	TLINTTSE-AKHOLAIVGPTRMDYCXIKALLTTLKEEIEKYDKKIHNQT
Mo	TLINTINE-SKHOLAIVGPTRMDYOXVKALLLTLKEEIEEYDKOLHGGKTTSST
MŁ	TAYGTAHT-VYGCMGVVGPTRMDYEGTIASVAAVALYIGDVL
Sa	SOYHEDET-LKGOIAVIGETAMHYONVIOLLNRIW
Sal	VGYGSGRE-AVAKLGVVGPTRMDXPGTMGAVRAVARYVGOILAES
Sm	OKFLIPYR-GFGILTVIGPVDLDYORTISLINVISRVLAVKLGDFYRYLNSNHYEVH
Svn	AIYROOEL-PMGSVSILGPTRMVXCOTIPLVEOAAECLSEALSKN
-	

Figure 4. Alignment of HrcA amino acid sequences from 17 different bacterial species. Invariant amino acid residues are given in bold letters and marked by a star underneath the alignment whereas conserved residues are marked by a dot. At, Agrobacterium tumefaciens; Bj, Bradyrhizobium japonicum; Bs, Bacillus subtilis; Bst, Bacillus stearothermophilus; Ca, Clostridium acetobutylicum; Cc, Caulobacter crescentus; Ct, Chlamydia trachomatis; Hp, Helicobacter pylori; Li, Leptospira interrogans; Ll, Lactococcus lactus; Mg, Mycoplasma genitalium; Mp, Mycoplasma pneumoniae; Mt, Mycobacterium tuberculosis; Sa, Staphylococcus aureus; Sal, Streptomyces albus; Sm, Streptococcus mutans; Syn, Synechocystis PCC6803. denatured proteins accumulating within the intracellular compartment.

8. The HrcA-CIRCE mechanism is not specific for *B. subtilis*

As already outlined above, the CIRCE element has been described in more than 30 different bacterial species so far implying that all these species also inherit the *hrcA* gene (Hecker *et al* 1996; Schumann 1996). Therefore, the HrcA-CIRCE mechanism seems to be more wide-spread than the sigma-32 mechanism. Most interestingly, there seem to be bacterial species which encode both mechanisms, and where that occurs, the *dnaK* operon is under σ^{32} control, while the *groE* operon is either only controlled by HrcA-CIRCE or in addition preceded by a σ^{32} -type promoter, thereby ensuring negative autoregulation in both cases.

This regulation mechanism also allows specific induction of the groE operon when enhanced amounts of GroES and GroEL are needed. One example is *Rhodo*bacter sphaeroides which contains two groE operons where groESL is preceded by a CIRCE element (Lee et al 1997). GroES and GroEL synthesis parallels the regulated production of ribulose-1,5-bisphosphate carboxylase (RubisCO) under different growth conditions, suggesting important links among chaperonin synthesis, RubisCO control, and RubisCO function (Terlesky and Tabita 1991). We would like to suggest that in all those bacterial species where the groE operon is controlled by CIRCE and HrcA, it is needed for the assembly of some protein complex such as RubisCO in photosynthetic bacteria.

9. Conclusions and future work

In B. subtilis four classes of heat shock genes have been described to be regulated by different mechanisms. Class I genes consist of the dnaK and groE operons which constitute the CIRCE regulon. Both operons are regulated at the transcriptional level by the HrcA repressor which binds to the CIRCE element located in both operons between the transcriptional and translational start sites. The activity of the HrcA protein is modulated in a still unknown way by the GroE chaperonin system, and induction of the CIRCE regulon occurs by titration of the GroE system interacting with non-native proteins. Future work will concentrate on two different aspects: to elucidate (i) the domain structure of HrcA and (ii) the molecular mechanism of interaction between the GroEL and HrcA. To date, the primary sequence of HrcA from 17 different bacterial species is known. Alignment of the amino acid sequence of these proteins revealed about 20% identical and a further 10% similar amino acid residues bringing the overall similarity to

about 30% which is rather low for proteins of identical function. But all proteins contain two regions of increased similarity, one near the N-, the other near the C-terminus, designated box A and box B, respectively (Schulz *et al* 1995; see figure 4). Whereas box A exhibits similarity with a region within the DeoR family of DNA-binding proteins and might contain the DNA-binding motif (M Rudolph, unpublished results), box B might represent the multimerization domain. Experiments are in progress to introduce point mutations in both boxes by sequence-specific mutagenesis and to analyse the phenotype of these mutants.

Acknowledgments

We are indebted to K Nakahigashi and T Yura, F Narberhaus and P Mazodier for communicating their unpublished amino acid sequences of the HrcA proteins encoded by *Agrobacterium tumefaciens*, *Bradyrhizobium japonicum* and *Streptomyces albus*, respectively. The financial travel assistance of the German Ministry of Research to WS to attend the Workshop in Varanasi on "Molecular Biology of Stress responses" is gratefully acknowledged. Financial support for the work carried out in the laboratory is provided by the Deutsche Forschungsgemeinschaft and the Fonds der Chemischen Industrie.

References

- De Las Peñas A, Connolly L, Gross C A 1997 The o^{E} -mediated response to extracytoplasmic stress in *Escherichia coli* is transduced by RseA and RseB, two negative regulators of o^{E} ; *Mol. Micobiol.* 24 373-385
- Gamer J, Bujard H and Bukau B 1992 Physical interactions between heat shock proteins DnaK, DnaJ, and GnpE and the bacterial heat shock transcription factor σ^{32} ; Cell 69 833– 842
- Hecker M, Schumann W and Völker U 1996 Heat-shock and general stress response in *Bacillus subtilis*; *Mol. Microbiol.* 19 417–428
- Hornuth G, Masuda S, Mogk A, Kobayashi Y and Schumann W 1997 The *dnaK* operon of *Bacillus subtilis* is heptacistronic; J. Bacteriol, **179** 1153-1164
- Lee W T, Terlesky K C and Tabita F R 1997 Cloning and characterization of two groESL operons of *Rhodobacter* sphaeroides: Transcriptional regulation of the heat-induced groESL operon; J. Bacteriol. **179** 487-495
- Li M and Wong S-L 1992 Cloning and characterization of the groESL operon from Bacillus subtilis; J. Bacteriol. 174 3981– 3992
- Liberek K, Galitski T P, Zylicz M and Georgopoulos C 1992 The DnaK chaperone modulates the heat shock response of *Escherichia coli* by binding to the σ^{32} transcription factor; *Proc. Natl. Acad. Sci. USA* 89 3516-3520
- Liberek K and Georgopoulos C 1993 Autoregulation of the *Escherichia coli* heat shock response by the DnaK and DnaJ heat shock proteins; *Proc. Natl. Acad. Sci. USA* 90 11019-11023
- Missiakas D, Raina S and Georgopoulos C 1996 Heat shock regulation; in Regulation of gene expression in Escherichia

coli (eds) E C C Lin and A S Lynch (London: Chapman and Hall) pp 481-502

- Missiakas D, Mayer M P, Lemaire M, Georgopoulos C and Raina S 1997 Modulation of the *Escherichia coli* σ^{E} (RpoE) heat-shock transcription-factor activity by the RseA, RseB and RseC proteins; *Mol. Microbiol.* **24** 355–371
- Model P, Jovanovic G and Dworkin J 1997 The Escherichia coli phage-shock-protein (psp) operon; Mol. Microbiol. 24 255-261
- Mogk A, Hayward R and Schumann W 1996 Integrative vectors for constructing single-copy transcriptional fusions between *Bacillus subtilis* promoters and various reporter genes encoding heat-stable enzymes; *Gene* 182 33-36
- Mogk A, Homuth G, Scholz C, Kim L, Schmid F X and Schumann W 1997 The GroE chaperonin machine is a major modulator of the CIRCE heat shock regulon of *Bacillus* subtilis; EMBO J. 16 4579–4590
- Nakahigashi K, Kanemori M, Morita M, Yanagi H and Yura T 1998 Conserved function and regulation of σ^{32} homologues in Gram-negative bacteria; J. Biosci. 23 407–414
- Parsell D A and Lindquist S 1993 The function of heat-shock proteins in stress-tolerance: degradation and reactivation of damaged proteins; Annu. Rev. Genet. 27 437-496
- Rouvière P E and Gross C A 1996 SurA, a periplasmic protein with peptidyl-prolyl isomerase activity, participates in the assembly of outer membrane porins; *Genes Dev.* 10 3170–3182
- Schmidt A, Schiesswohl M, Völker U, Hecker M and Schumann W 1992 Cloning, sequencing, mapping, and transcriptional

analysis of the groESL operon from Bacillus subtilis; J. Bacteriol. 174 3993-3999

- Schulz A, Tzschaschel B and Schumann W 1995 Isolation and analysis of mutants of the *dnaK* operon of *Bacillus subtilis*; *Mol. Microbiol.* **15** 421–429
- Schulz A, Schwab S, Homuth G, Versteeg S and Schumann W 1997 The htpG gene of Bacillus subtilis belongs to class III heat shock genes and is under negative control; J. Bacteriol. 10 3103–3109
- Schulz A and Schumann W 1996 hrcA, the first gene of the Bacillus subtilis dnaK operon encodes a negative regulator of class I heat-shock genes; J. Bacteriol. 178 1088-1093
- Schumann W 1996 Regulation of the heat shock response in Escherichia coli and Bacillus subtilis; J. Biosci. 21 133-148
- Terlesky K C and Tabita F R 1991 Purification and characterization of the chaperonin 10 and chaperonin 60 proteins from *Rhodobacter sphaeroides*; *Biochemistry* 30 8181–8186
- Yuan G and Wong S-L 1995a Isolation and characterization of Bacillus subtilis regulatory mutants: evidence for orf39 in the dnaK operon as a repressor gene in regulating the expression of both groE and dnaK; J. Bacteriol. 177 6462-6468
- Yuan G and Wong S-L 1995b Regulation of groE expression in *Bacillus subtilis*: the involvement of the σ^{32} -like promoter and the roles of the inverted repeat sequence (CIRCE); *J. Bacteriol.* 177 5427-5433
- Zuber U and Schumann W 1994 CIRCE, a novel heat shock element involved in regulation of heat shock operon *dnaK* of *Bacillus subtilis*; J. Bacteriol. **176** 1359–1363