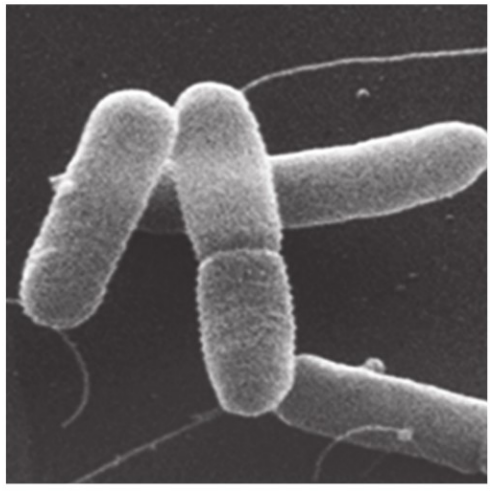
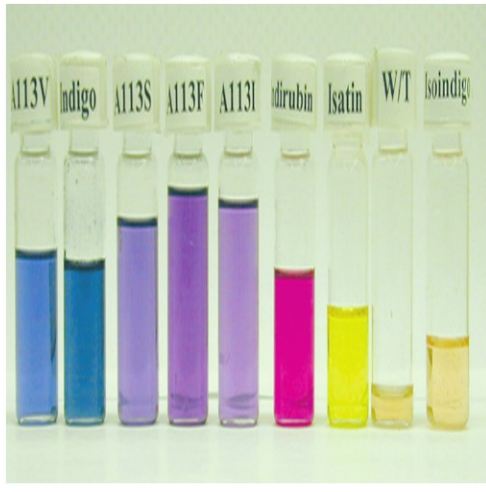
4

Pseudomonas

Molecular Biology of Emerging Issues





Edited by Juan-Luis Ramos and Roger C. Levesque

Pseudomonas

Pseudomonas

Edited by Juan-Luis Ramos, *CSIC, Granada, Spain*

Volume 1: Genomics, Life Style and Molecular Architecture
Volume 2: Virulence and Gene Regulation
Volume 3: Biosynthesis of Macromolecules and Molecular Metabolism
Volume 4: Molecular Biology of Emerging Issues

Pseudomonas

Volume 4
Molecular Biology of Emerging Issues

Edited by

Juan-Luis Ramos
CSIC
Granada, Spain

and

Roger C. Levesque
Université Laval
Québec, Canada

Springer

Library of Congress Cataloging-in-Publication Data

Pseudomonas / edited by Juan-Luis Ramos.
p. cm.
Includes bibliographical references and index.
ISBN 0-306-48375-0 (v. 1) – ISBN 0-306-48376-9 (v. 2) – ISBN 0-306-48377-7 (v. 3) – ISBN 0-306-48378-5 (set)
1. Pseudomonas. I. Ramos, Juan-Luis, 1956–

QR82.P78P772 2004
579.3′32 – dc22

2004043811

ISBN-10 0-306-48375-0 (Vol. 1)
ISBN-13 978-0-306-48375-2 (Vol. 1)
ISBN-10 0-306-48376-9 (Vol. 2)
ISBN-13 978-0-306-48376-9 (Vol. 2)
ISBN-10 0-306-48377-7 (Vol. 3)
ISBN-13 978-0-306-48377-6 (Vol. 3)
ISBN-10 0-387-28834-1 (Vol. 4)
ISBN-13 978-0-387-28834-5 (Vol. 4)
Also available as part of an indivisible set of all 3 volumes 0-306-48378-5

www.springer.com

10 9 8 7 6 5 4 3 2 1

A C.I.P. record for this book is available from the Library of Congress.

Printed in the Netherlands.

PREFACE

Twenty years have gone by since Jack Sokatch first published his outstanding *The Biology of Pseudomonas* back in 1986. This was followed by two books published by the ASM that contained the presentations of the *Pseudomonas* meetings held in Chicago in 1989 and Trieste in 1991. The earlier volume of these two was edited by Simon Silver, Al Chakrabarty, Barbara Iglewski, and Sam Kaplan, and the later one by Enrica Galli, Simon Silver, and Bernard Witholt. The time was ripe for a series of books on *Pseudomonas* because of its importance in human and plant pathogenesis, biofilms, soil and rhizosphere colonization, etc. Efforts were devoted to produce the first three volumes of the series on the biology of *Pseudomonas* after a meeting with Kluwer staff members in August 2002 during the XI IUMS conference in Paris (France). In less than a year a group of outstanding scientists in the field, after devoting much of their valuable time, managed to complete their chapters for the three volumes of the series. To ensure the high standard of each chapter, renowned scientists participated in the reviewing process. The three books collected part of the *"explosion"* of new vital information on the genus *Pseudomonas*.

A rapid search for articles containing the word "*Pseudomonas*" in the title in the last 10 years produces more than 6000 articles! Consequently, not all possible topics relevant to this genus were covered in the three previous volumes. This new volume, *Pseudomonas* volume IV edited by Roger Levesque from Université Laval in Canada and Juan L. Ramos from the CSIC in Spain, is intended to collect some of the most relevant emerging new issues. This fourth volume on *Pseudomonas* is organized in various topics grouped under a common heading: "*Pseudomona*s: Molecular Biology of Emerging Issues" and the chapters are organized in three sections: Virulence and Pathogens, Genomics and Proteomics, and Physiology, Metabolism and Biotechnology.

The section "Virulence and Pathogens" comprises a series of fascinating chapters on relevant issues that make bacteria of the species *Pseudomonas aeruginosa* pathogenic for humans and animals. Typing of strain collections using different molecular approaches have revealed that the current *P. aeruginosa* population is in linkage equilibrium and consists of a network of equivalent

genotypes. How these bacteria colonize the host tissues, which genes are specifically induced, cloning of pathogenesis determinants, how bacteria acquire iron or the role of phospholipases are some of the issues that are treated in this section.

The section "Genomics and proteomics" covers in two chapters issue related to how a genome-wide mutant library of *P. aeruginosa* is created, together with an update of the genome database of the PAO1 strain. This information is of an extremely high value not only for those working in *P. aeruginosa* but also for scientists working with other species of the genus *Pseudomonas*, an even to those working in other more general fields. This section also includes an authoritative review of type IV pili in species of the genus *Pseudomonas* and their role in twitching motility, bacteriophage sensitivity, attachment to surfaces and DNA uptake.

Finally the section on Physiology, Metabolism and Biotechnology covers the characteristics of a set of extremely important proteins in the metabolism of *Pseudomonas*: The oxygenases. The chapters included in this section cover structure, mechanisms of reaction and biotechnological potential of the enzymes that make *Pseudomonas* key in mineralization of many chemicals and in biotransformation processes. The chapter on the evolution of catabolic pathways explores some of the catabolic specialties of bacteria of this genus and how the information is spread in nature.

This fourth volume would never have seen the light if it were not for a group of outstanding scientists in the field, who after devoting much of their valuable time, have produced enlightening chapters to try to complete the story that began with the three previous volumes of the series. It has been an honor for us to work with them and we truly thank them.

The review process has also been of great importance to ensure the high standard of each chapter. Renowned scientists have participated in the review, correction, and editing of the chapters. Their assistance is immensely appreciated. We would like to express my most sincere gratitude to:

Arie Ben-Bassat	José-Luis García	Keith Poole
Lori Burrows	Bob Hancock	Burkhard Tuemmler
Pierre Cornelis	Karl-Erich Jaeger	Paolo Visca
Eduardo Díaz	José-Luis Martínez	Rolf-Michael Wittich
Estrella Duque	Soeren Molin	Thomas K. Wood
Kensuke Furukawa	George O'Toole	

We would also like to thank Carmen Lorente for her assistance and enthusiasm in the preparation of this fourth volume.

Juan L. Ramos and Roger C. Levesque

CONTENTS

LIST OF CONTRIBUTORS

Dr. Fiona Brinkman
Department of Molecular Biology and Biochemistry, Simon Fraser University, Burnaby, B.C., Canada, brinkman@sfu.ca

Dr. Ayelet Fischman
Department of Biotechnology and Food Engineering, Technion—Israel Institute of Technology, Haifa, 32000, Israel

Dr. Michael A. Jacobs
Department of Genome Sciences, Box 357730, University of Washington, Seattle, WA 98195, USA

Dr. Roger C. Levesque*
Microbiologie moléculaire et génie des protéines, Département de Biologie Médicale, et Pavillon Charles-Eugène Marchand, Faculté de Médecine Université Laval, Sainte-Foy, Québec, Canada G1K 7P4
E-mail: rclevesq@rsvs.ulaval.ca

Dr. Colin Manoil*
Department of Genome Sciences, Box 357730, University of Washington, Seattle, WA 98195, USA, E-mail: manoil@u.washington.edu

Dr. Rebecca Parales*
Section of Microbiology, 226 Briggs Hall, 1 Shields Avenue, University of California, Davis, CA 95616, USA, E-mail: reparales@ucdavis.edu

Dr. Sol M. Resnick
Biotechnology R&D. The DOW Chemical Company, San Diego CA 92121, USA

Dr. Lingyun Rui
Departments of Chemical Engineering and Molecular and Cell Biology, University of Connecticut, Storrs, CT 06269-3222, USA

Dr. Isabelle Schalk*
Département des Récepteurs et Protéines Membranaires, UMR 7176 – LC1 CNRS – Université Louis Pasteur Bld Sébastien Brant, F-67 413 Illkirc, Strasbourg, France, E-mail: schalk@esbs.u-strasbg.fr

Dr. Ying Tao
Departments of Chemical Engineering and Molecular and Cell Biology, University of Connecticut, Storrs, CT 06269-3222, USA

Dr. Burkhard Tümmler*
Klinische Forschergruppe, OE 6711, Medizinische Hochschule Hannover, D-30625 Hannover, Germany, E-mail: tuemmler.burkhard@mh-hannover.de

Dr. Jan Roelof van der Meer*
Department of Fundamental Microbiology, Bâtiment de Biologie, University of Lausanne, 1015 Lausanne, Switzerland,
E-mail: janroelof.vandermeer@unil.ch

Dr. Gönül Vardar
Departments of Chemical Engineering and Molecular and Cell Biology, University of Connecticut, Storrs, CT 06269-3222, USA

Dr. Michael L. Vasil*
University of Colorado at Denver and Health Sciences Center, Aurora, CO 80045, USA, E-mail: mike.vasil@uchsc.edu

Dr. Cynthia B. Whitchurch*
Department of Microbiology, Monash University, Clayton, Victoria 3800, Australia, E-mail: cynthia.whitchurch@med.monash.edu.au

Dr. Thomas K. Wood*
Departments of Chemical Engineering and Biology, TEXAS A&M University, College Station, TX 77843–3122 USA
E-mail: twood@engr.uconn.edu

Marked with an asterisk are the corresponding authors of each chapter.

1

NEW INSIGHTS ON IRON ACQUISITION MECHANISMS IN PATHOGENIC *PSEUDOMONAS*

Isabelle J. Schalk

Département des Récepteurs et Protéines Membranaires
UMR 7176 – LC1 CNRS, Université Louis Pasteur
ESBS, Bld Sébastien Brant
F-67 413 Illkirch, Strasbourg, France

Key Words: iron uptake, iron uptake regulation, siderophore, hemophore, outer membrane transporter

1. INTRODUCTION

Almost all bacteria require iron for growth and survival. Iron is a constituent of enzymes crucial in oxygen metabolism, electron transfer, and RNA synthesis. Despite its abundance in the earth's crust, the availability of iron is severely limited by the very high insolubility of iron(III) at physiological pH. In the presence of oxygen, iron(II) is rapidly oxidized to iron(III) which precipitates as a polymeric oxyhydroxide. The solubility of ferric hydroxide is extremely low (10^{-38} M) such that, at physiological pH, the concentration of free iron(III) is $<10^{-18}$ M. This value is far too low for sufficient iron to be acquired by passive diffusion of ions into the cells and to allow growth of aerobic microorganisms. In humans and animals, intracellular iron is mostly bound to ferritin or heme and the extracellular iron found in body fluids is attached to high-affinity iron-binding proteins, such as transferrin found in serum

Pseudomonas, Volume 4, edited by Juan-Luis Ramos and Roger C. Levesque

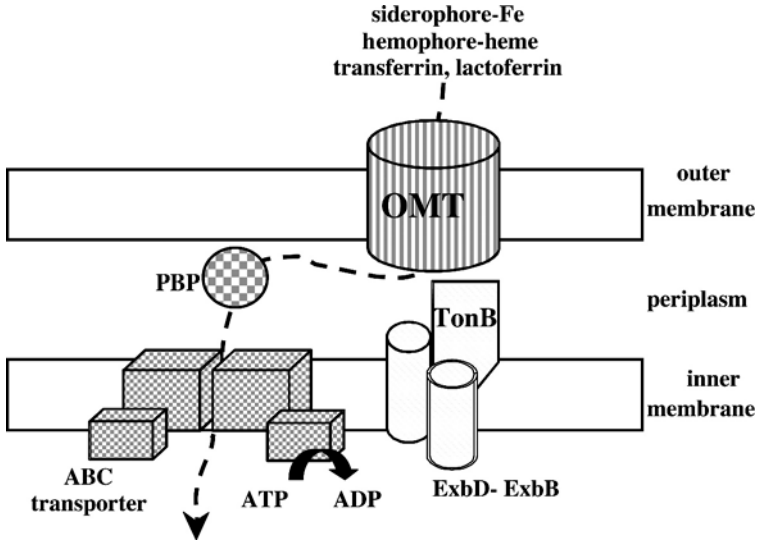


Figure 1. Iron uptake in Gram-negative bacteria. Iron is delivered by either lactoferrin or transferrin, heme or heme-hemophore, and homologous or heterologous siderophores to specific OMTs located in the outer membrane. The OMTs transport iron/iron ligand across the outer membrane and the transport is energized by the pmf of the cytoplasmic membrane via a complex formed of three proteins, TonB–ExbB–ExbD. In the periplasm, Fe^{3+}, heme, or siderophore-Fe^{3+}bind to periplasmic binding proteins (PBPs) that are specific for either Fe^{3+}, heme or ferric-siderophore. The PBPs deliver the iron compounds to an ABC transporter for the transport across the cytoplasmic membrane. The energy required for transport across the cytoplasmic membrane is provided by the hydrolysis of ATP.

and lymph and lactoferrin in secretions. The level of free iron in body fluids is therefore in the order of 10^{-18} M, which is far too low to support the growth of microorganisms. The infected host can therefore be considered as an iron-depleted environment.

In order to gain access to iron, bacteria have developed several strategies for the acquisition, solubilization, and transportation of iron (Figure 1). One of the most common methods of iron acquisition involves the synthesis and the secretion of low molecular weight iron ligands that are called siderophores.[20–22] After secretion, siderophores chelate iron in the extracellular environment with high affinity ($K_d = 10^{-49}$ M in some cases) and transport it back into the bacteria. This siderophore-mediated iron acquisition plays an important role in the virulence of pathogenic bacteria[19,128] including *Pseudomonas*.[102,151,166,167] It has been shown that siderophores are produced by clinical isolates during infection[62] and in the case of *Pseudomonas aeruginosa* in animal models of infection.[63] In parallel to the iron uptake via siderophores, certain bacteria (many of them pathogens) are able to use heme bound iron from hemoglobin,

hemopexin, and haptoglobin.[58,157,164] At last, some species can acquire iron from transferrins or lactoferrins also involving specific uptake systems in the cell envelope.[38,139] All iron uptake mechanisms involve a specific outer membrane transporter (OMT) and an ABC transporter for the transport across the outer and the inner membrane, respectively (Figure 1). The energy required is provided by the hydrolysis of ATP for the inner membrane transport and by the proton motive force (pmf) of the inner membrane for the transport through the OMT. In the case of the OMT, the energy is transferred from the inner membrane to the receptor located in the outer membrane by the TonB–ExbB–ExbD complex.

The different iron uptake systems of *Pseudomonas* have recently been reviewed in details[120] and will be summarized here. The major objective of this review is to address current information about the mechanism of the iron uptake at the molecular level in *Pseudomonas* and more specifically in *P. aeruginosa*, the best-studied member of this family of bacteria. The fluorescent properties of the siderophore pyoverdine (Pvd) of this bacterium have been used as a powerful tool to unfold the mechanism of interaction of the ferric-Pvd complex with its OMT FpvA. New insights on the iron uptake in the *Burkholderia* (formerly *Pseudomonas*) *cepacia* complex will be also presented. The different regulators of these iron uptake activities in *Pseudomonas* have been the topic of a chapter in *Pseudomonas* Vol III[52] and will be just summarized here.

2. IRON UPTAKE VIA SIDEROPHORES

2.1. Siderophores

In the last 40 years, structural studies have shown a great diversity of siderophore structures (for recent review on siderophores see refs [17,36,165]). But despite their great diversity, siderophore structures are always characterized by the presence of one, two, and in most cases, three bidentate chelating groups, which are generally oxygenated and crucial for the formation of very stable hexacoordinated octahedral complexes between the siderophore and iron(III). The chelating groups are generally either catecholates, hydroxamates or hydroxyacids, but can also be any other bidentate groups. Siderophores show strong affinity for only the higher oxidation state of iron, and the association constant for iron(III) for typical molecules containing three bidentate ligands is around 10^{30} M^{-1}, or greater. The affinity for gallium is also high, but the attraction for aluminum and for divalent ions is substantially less. Thus, among all naturally occurring metal ions of abundance, Fe(III) is the specific ion for the siderophore ligand.

Under iron deficient conditions, all fluorescent *Pseudomonas* are characterized by the production of yellow-green, water-soluble siderophores called Pvds. Pvds form a wide class of mixed siderophores with a great variety of structures and a high affinity for iron, with association constants as high as 10^{32} M^{-1}.[5] These siderophores are effective at acquiring iron from transferrin[102,166,167] and lactoferrin[167] and this has obvious implications *vis-à-vis in vivo* growth and pathogenesis. Mutants with deficiencies in Pvd biosynthesis or transport are effectively avirulent in animal models of infection.[102,151] In addition to their high-affinity iron chelators, fluorescent pseudomonads are known to produce lower-affinity siderophores, such as pyochelin, a derivative of salicylic acid, salicylic acid itself,[99,161] and quinolobactin.[111] These compounds can be considered as secondary siderophores since they are usually produced in much lower amounts than Pvds (a few mg/l vs. 100–200 mg/l for the Pvds, unpublished data) and are less efficient in iron binding and uptake. They can also be considered as compounds that provide rescue iron uptake systems, since their production (in the case of salicylic acid) and their role in iron uptake (in the case of pyochelin) have been shown to be greater in Pvd-deficient mutants than in wild-type cells.[59,98] Fluorescent *Pseudomonas* are also characterized by their capacity to take up heterologous siderophores like citrate, enterobactin,[122] ferrioxamine,[157] or Pvds produced by other pseudomonads,[37,98,122] which explains the large number of ferric-siderophore receptor homologs identified in the recently completed genome sequence[150] of *P. aeruginosa* (23 putative siderophore OMTs, http://www.pseudomonas.com/).

The *B. cepacia* complex has been shown to produce four siderophores: ornibactins, salicylic acid, pyochelin, and cepabactin.[101,104,144,145,161] Ornibactins and salicylic acid are the predominant siderophores found in cystic fibrosis isolates, but there is species and strain variation.[42]

2.1.1. Pyoverdine

Pvds, the main siderophores produced by all fluorescent pseudomonads, are constituted of a chromophore derived from 2,3-diamino-6,7-dihydroxyquinoline, a peptidic moiety bound to the chromophore and a side chain bound to the nitrogen atom at position C-3 of the chromophore. The side chain bound to the chromophore is most frequently a diacid of the Krebs cycle such as succinic, malic, or α-ketoglutaric acid or one of their respective amide derivatives. The composition and length of the peptide are unique to each strain. Based on the determination of about 30 primary structures of Pvd, the peptidic moiety is usually composed of 6–12 amino acids and can be either linear, partially cyclic or completely cyclic.[1] In the *P. aeruginosa* strain, three structurally different Pvds (with different peptide chains) have been identified (Pvd type I, II, and III), each recognized at the level of the outer membrane by a specific receptor, called FpvAI-III.[35,45,103] Since the chromophore is common

to each Pvd among fluorescent pseudomonads, the specificity of siderophore recognition by its OMT is attributed to the peptide region. Moreover, this diversity of Pvd structure among fluorescent *Pseudomonas* forms the basis of a novel typing scheme called siderotyping.[56,100]

The iron is chelated with a 1:1 stoichiometry by the chromophore, which forms the first bidentate group and by the peptide moiety, which contains the two other bidentate chelating groups, creating a complete and efficient iron(III) complexation. Three-dimensional structures of only three Pvds have been solved: the crystal structure of pseudobactin A from *Pseudomonas* B10 and the NMR structure of gallium(III) complexes of Pvd GM-II and Pvd G4R from *Pseudomonas fluorescens* and *Pseudomonas putida* G4R, respectively.[12,109,152,153] Comparison of these three Pvd structures shows that the coordination is not achieved in the same manner in all Pvds. Pseudobactin B10-Fe(III) and Pvd-Ga(III) have Λ configurations around the metal ion, while Pvd G4R-Ga(III) has a Δ configuration.

The diversity in the Pvd primary structures indicates that a large variety of enzymes are likely to be required for their synthesis. For all Pvds, the peptide and the chromophore are thought to be derived from amino acid precursors that are assembled by nonribosomal peptide synthetases (NRPSs) with other enzymes catalyzing additional reactions to complete the mature Pvds.[2,13,41,83,84,86,94,95,97,106,148,160] NRPSs are multimodular enzymes, with each module governing the insertion of a single amino acid into the peptide product. The synthesis of the chromophore, which is a condensation product of D-tyrosine and L-2,4-diaminobutyrate,[47] seems to involve the PvdL NRPS in *P. aeruginosa*.[112] PvdL is indeed, the only NRPS, which is highly conserved in all genomes analyzed. Another conserved enzyme among fluorescent *Pseudomonas* is PvdH. This enzyme, required also for the chromophore synthesis, is an aminotransferase that catalyzes the formation of L-2,4-diaminobutyrate from aspartate β-semialdehyde.[156] These enzymes involved in the synthesis of the chromophore probably work in concert with the other NRPS found in *P. aeruginosa* (PvdI, PvdJ, and PvdD[97]) to synthesize the entire Pvd backbone. The *pvc* genes have been described previously as being also involved in the synthesis of the chromophore of Pvd in *P. aeruginosa*.[147,148] These genes are however, not present in the genomes of the other fluorescent pseudomonads, suggesting that these genes play a role in Pvd biosynthesis only in the case of *P. aeruginosa* strains. It is likely that most of the genes required for the biosynthesis of Pvd in *P. aeruginosa* PAO1 have now been identified, however, more studies will be needed to clearly understand the complete enzymatic pathway of Pvd synthesis. A limited number of Pvd biosynthetic genes have also been reported in other fluorescent pseudomonads.[3,6,46,112] They show homology to known *P. aeruginosa pvd* genes and this is consistent with the presence of a conserved pathway for Pvd biosynthesis.

2.1.2. Ornibactin

Ornibactins are linear hydroxamate/hydroxycarboxylate siderophores similar in structure to the Pvds produced by fluorescent pseudomonads, yet they lack a chromophore. They are composed of a conserved tetrapeptide L-Orn^1(N^{δ}OH,N^{δ}acyl)-D-*threo*-Asp(β-OH)-L-Ser-L-Orn^4(N^{δ}OH,N^{δ}formyl)-1,4-diaminobutane to which is attached one of three possible acyl groups. These acyl groups, linked to Orn,[1] vary in length and include 3-hydroxybutanoic acid, 3-hydroxyhexanoic acid, and 3-hydroxyoctanoic acid, forming the three different ornibactins, which are designated ornibactin-C_4, ornibactin-C_6, and ornibactin-C_8 according to they acyl chain length.[145]

Two genes (*pvdA* and *pvdD*) involved in the synthesis of ornibactin have been identified in *B. cepacia*. The *pvdA* gene encodes the enzyme L-ornithine N^5-oxygenase, which is responsible for catalyzing the hydroxylation of L-ornithine and for the formation of the hydroxamate ligands.[143] The identification of the *B. cepacia pvdA* gene was based on its homology to the *P. aeruginosa pvdA* gene, which codes for the same enzyme and which is required for the synthesis of the siderophore.[160] A *pvdD* homolog that demonstrate homology to NRPSs, was also identified.[143]

2.1.3. Pyochelin

Pyochelin is a structurally unique phenolate siderophore which has neither hydroxamate nor cathecolate character. Pyochelin has been assigned the structure 2-[2-(*o*-hydroxyphenyl)-2-thiazolin-4-yl]-3-methyl-4-thiazolidinecarboxylic acid and apparently chelates Fe(III) in a 2:1 pyochelin-Fe(III) stoichiometry, with an affinity in the range of 2×10^5 M^{-1} in methanol.[8,39] The coordination site of pyochelin to ferric iron is not yet fully described but probably involves the phenolic-OH, the thiazoline, and perhaps the thiazolidine nitrogens. Pyochelin also binds other transition metals [e.g. Mo(IV), Ni(II), and Co(II)] with appreciable affinities and might be implicated in the delivery of both Co(II) and Mo(IV) to *P. aeruginosa* cells.[162]

Pyochelin is a condensation product of salicylic acid and two cysteins and its synthesis involves the *pchDEFG* genes.[57,129,130] The biosynthesis of this siderophore has recently been reviewed.[40]

2.2. Siderophore Outer Membrane Transporters

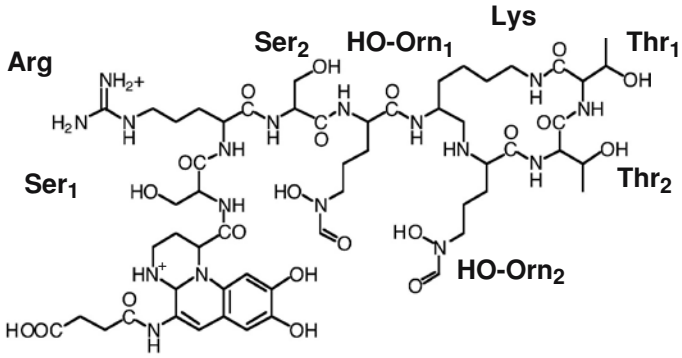
Once formed in the extracellular medium, the ferric-siderophore complexes are captured by their cognate OMT (MW between 75 and 90 kDa) at the bacterial cell surface, with K_d in the range of 0.3–50 nM. Twenty-three putative TonB-dependent OMT genes potentially involved in iron uptake were retrieved from the complete genome sequence of *P. aeruginosa* PAO1.[150] Only a few of these transporters have been identified: FptA,[9] PfeA,[122] FiuA,[157] FpvA,[121] and FpvB[60] for the iron uptake via pyochelin, enterobactine, ferrioxamine, and

Pvd, respectively. Recent studies have identified FpvB (gene PA4168) as being an alternative type of Pvd-Fe receptor of *P. aeruginosa*. Functional studies are available mostly on FpvA (PA2398), the OMT involved in the iron uptake via type I Pvd in *P. aeruginosa*. Three structurally different Pvds (with different peptide chains) have been identified (Pvd type I, II, and III) from *P. aeruginosa* strains, each recognized at the level of the outer membrane by a specific receptor, called FpvAI-III.[35,45,103] Each *P. aeruginosa* strain produces only one single receptor and uses one type of Pvd.[44] The receptor for *P. aeruginosa* PAO1 Pvd (type I Pvd, Figure 2) has been intensively characterized using physiological, immunological, molecular biology approaches, and also using the intrinsic fluorescence of Pvd. The spectral characteristics of iron-free Pvd or Pvd loaded with Ga or Al and FpvA are almost ideal to observe Fluorescence Resonance Energy Transfer (FRET) between Pvd and Trps of FpvA (Figure 2b), when they are at a short distance from each other, i.e. when Pvd is bound to its receptor.[53,136] FRET is one of the most powerful techniques to monitor protein ligand interactions as a function of time. It has been largely used to study the FpvA receptor generating many data that were used to define the first steps of the iron uptake mechanism via the siderophore Pvd. The different steps of FpvA dependent Pvd uptake will be discussed here after. Hybridization studies suggest that homologs of *fpvA* are present in a variety of *P. aeruginosa* strains as well as in *P. aureofaciens, P. putida*, and *P. fluorescens*,[154] and receptors for ferric-pseudobactin have been cloned and sequenced from *P. putida* WCS368 (*pupA*)[15] and *P. fluorescens* M114 (*pbuA*).[110] A receptor for Pvd/pseudobactin has also been described in *P. syringae*[143] although the gene has not been cloned. In *B. cepacia*, only the gene for the ferric-ornibactin receptor, *orbA*, has been cloned.

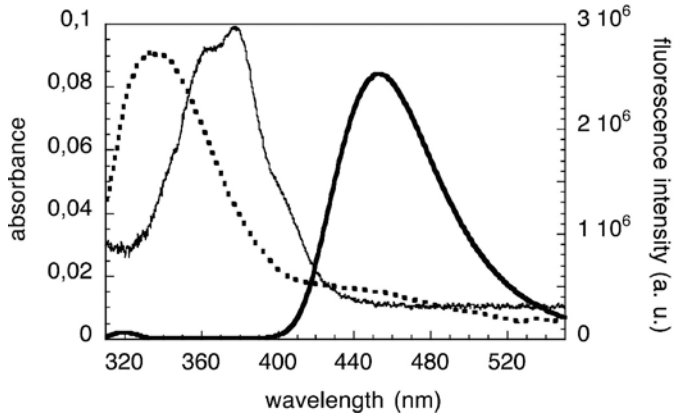
Until recently, only the crystal structures of OMTs of *E. coli* have been solved. The structure of FhuA,[52,92] the ferrichrome OMTs has been published in 1998, followed in 1999 by the structure of FepA, the ferric-enterobactin receptor[24] and in 2002 the structure of FecA, the ferric-citrate OMT.[51,168] Finally, recently the crystal structure of an OMT of *P. aeruginosa*, FpvA, has been solved[32,33] (Figure 3). Crystals of the FptA receptor have been also obtained and the crystal structure should be available soon.[34]

2.2.1. Structure of Siderophore OMTs

Based on sequence alignment,[54] all the siderophore OMTs of Gram-negative bacteria should have the same kind of structure as those described previously for the *E. coli* transporters and more recently for the FpvA receptor (Figure 3) of *P. aeruginosa*.[33] All these receptors are composed of a COOH-terminal β-barrel domain (22 β-strands) occluded by a N-terminal domain containing a mixed 4-stranded β-sheet, called plug (Figure 3). The globular N-terminal domain is firmly held in place by many hydrogen bonds made with the β-strands of the barrel (70 in the case of FecA[168]), blocking completely the



A.



B.

Figure 2. (a) Structure of Pvd. (b) Spectral properties of Pvd. UV spectrum of iron-free Pvd (thin line), fluorescent spectra of purified FpvA (dotted line), and iron-free Pvd (thick line). For the UV spectrum of iron-free Pvd (thin line), Pvd was dissolved in pyridine–acetic acid buffer pH 5.0. For the fluorescent spectra of purified FpvA (dotted line) and iron-free Pvd (thick line) the excitation wavelength was set at 290 and 380 nm, respectively. The protein and iron-free Pvd were dissolved in Tris–HCl buffer pH 8.0 and 1% (v/v) octyl-POE.

access to the periplasm. The major differences between all these structures are the sequence and the conformation of the extracellular loops of the β-barrel, which are specific to each OMT.

In all the OMT structures solved, the ferric-siderophore-binding site is located above the plug, well outside the membrane and is composed of residues of the plug and of the β-barrel domain. The electrostatic properties of the

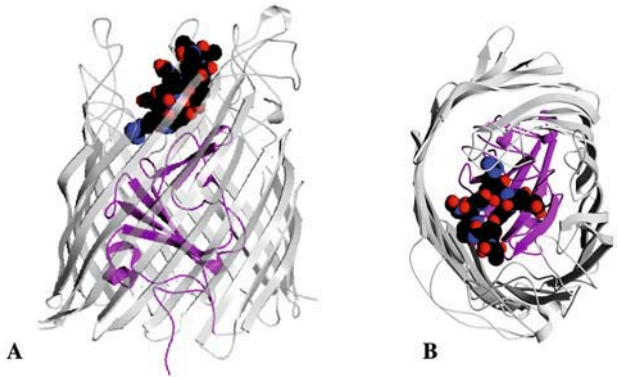


Figure 3. The FpvA-Pvd structure[33]. (a) Ribbon representation of the FpvA-Pvd. The plug domain and the transmembrane β-barrel domains are colored in purple and gray, respectively. The Pvd is depicted in space filling representation. (b) View along the β-barrel axis of the FpvA-Pvd structure drawn in ribbon. The Pvd is drawn in space filling representation.

siderophore-binding pocket confers specificity related to the chemical features of the siderophore. The binding pocket of iron-free Pvd on FpvA in *P. aeruginosa*, is mainly composed of aromatic residues.[33] In the FhuA-ferrichrome structure,[50,92] eight aromatic and two hydrophilic residues form the ferrichrome-binding site of FhuA. The ferric-citrate-binding pocket of FecA is composed of positively charged residues (5 Arg) interacting with the negatively charged ferric-citrate.[51,168] In the FepA structure, disorder in the electron density map in the region of the putative iron sites did not allow to model the ferric-enterobactin-binding site unambiguously.[24] However, the presence of two Fe density peaks and their relative weakness suggest both partial occupancy and multiple binding sites or ligand orientations for ferric-enterobactin.

Sequence alignment studies of all known siderophore OMTs showed the existence of a siderophore OMTs subfamily having an additional domain of around 70 amino acids at the N-terminus.[54,137,159] This region, of so far unknown structure, is localized in the periplasm and is involved in the regulation of the expression of the genes necessary for iron uptake in the bacteria.[163] Among others, belong to this subfamily the Pvd OMT FpvA of *P. aeruginosa*, the ferric-citrate OMT FecA of *E. coli*[18,83] and the hemophore OMT HasR of *Serratia marcescens*.[131] This subfamily and their special features have recently been reviewed.[137]

2.2.2. Binding Properties of Siderophore OMTs

The siderophore OMT subfamily characterized by an additional N-terminal end seem to have another special feature. At least two members of

the siderophore OMT family (FpvA and FecA, the archetypes of this subfamily) are able to bind with close affinities to a common or overlapping binding site, their iron-loaded and iron-free siderophore.[136,168]

This ability of a siderophore OMT to bind its iron-free siderophore has been shown for the first time on FpvA (the Pvd receptor in *P. aeruginosa*).[136] The fluorescent properties of the Pvd clearly showed that under iron limiting conditions, all the FpvA receptors at the cell surface of *P. aeruginosa* ATCC15692 are loaded with iron-free Pvd.[135,136] When the cells are iron starved, the FpvA-Pvd complex seems to be the normal state of this transporter. Based on binding assays, it was shown that iron-free and iron-loaded Pvd bind to a common or overlapping binding site on FpvA, with unexpected similar affinities of 3 and 0.5 nM, respectively.[31,135,136] A larger difference in affinity was expected to explain the efficiency in the iron uptake mechanism observed under iron-limited conditions, where the concentration of iron-free siderophore is higher than the concentration of ferric-siderophore. More recently, the crystal structures of the three different forms of FecA, e.g. FecA, FecA-cit, and FecA-cit-Fe, have shown that this receptor, like FpvA, is able to bind its apo- and ferric-siderophore, dicitrate and diferric-dicitrate, at a common binding site.[168]

Although the binding site is common, the interaction of the apo- and ferric-siderophore with the residues of the binding site in FecA and FpvA is quite different.[53,168] Only the ferric form of the siderophore will be transported inside the cells, since the FecA-cit or FpvA-Pvd complexes are not capable of transport the apo-siderophore. Therefore, the OMT must be able to distinguish between iron-free and ferric-siderophores. In the FecA-cit-Fe complex, the two citrate ions are in planar orientation, which allows their carboxyl groups to coordinate the two ferric ions. In the FecA-cit complex, the iron-free dicitrate are in orthogonal orientation. Comparison of the binding pocket for diferric-dicitrate and iron-free dicitrate shows two Arg residues of the β-barrel involved in ligand binding via hydrogen bonds. In addition, diferric-dicitrate accepts hydrogen bonds from three other residues that do not participate in the binding of iron-free dicitrate by FecA and van der Waals contacts are provided by residues from the β-barrel. However, the most important feature is that only the binding of the diferric-dicitrate to FecA is accompanied by significant conformational changes in two extracellular loops, L7 and L8. The binding of diferric-dicitrate generates a large rearrangement of these loops, which closes the binding pockets, sequestering the ligand in the binding site. When iron-free dicitrate binds to FecA, no binding pocket closure is observed and the iron-free dicitrate remains accessible from the extracellular medium. Although, apo- and Pvd-Fe bind to a common binding site, time resolved fluorescent spectroscopy studies, using the fluorescent properties of Pvd and Pvd-Ga, have shown a different proteic environment for both forms of Pvd.[53] When loaded with metal, the binding site of the siderophore is not as flexible and solvent accessible and its environment is not as polar as in the FpvA-Pvd complex, suggesting that some extracellular

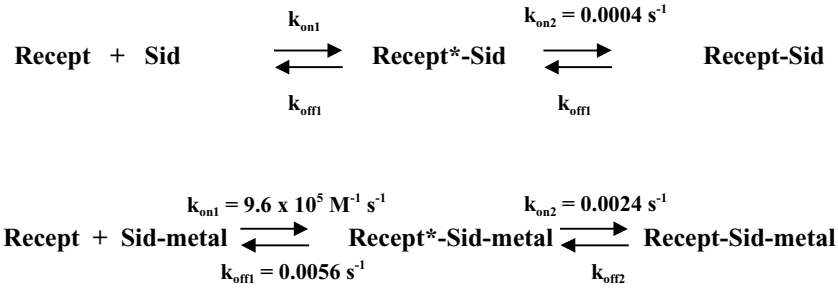


Figure 4. Kinetic model for the binding of Pvd-metal and apo-Pvd to FpvA. The equilibriums represent a minimal kinetic binding model, derived from *in vivo* and *in vitro* binding data.[31] Receptor (recept) stands for FpvA and Sid for Pvd.

loops could form a lid and trap the Pvd-metal in its binding site as observed in FecA-citrate-Fe.

While the FecA structures and FpvA-binding assays both suggest a single binding pocket for the iron-free or ferric-siderophore in the transporter, it seems that the binding of either form of siderophore to the pocket may occur in more than one kinetically distinct step. Kinetic studies using the fluorescent properties of Pvd have shown that both Pvd and Pvd-Ga bind in two kinetic steps to the FpvA transporter[31] (Figure 4). For the first step, the binding process is in the range of seconds for Ga-Pvd ($k_{on1} = 9.6 \times 10^{-5}$ $M^{-1}s^{-1}$, $t_{1/2} = 7$ s for [Pvd-Ga] = 100 nM) and on the same timescale for iron-free Pvd. The second step is slower, in the range of minutes for Ga-Pvd ($k_{on2} = 0.0024$ s^{-1}, $t_{1/2} =$ 5 min) and half an hour for iron-free Pvd ($k_{on2} = 0.0004$ s^{-1}, $t_{1/2} = 29$ min) (Figure 4). A reasonable explanation may be that the two kinetically distinct steps represent first, the recognition and capture of the ligand (iron-free or ferric-siderophore) at the cell surface by the extracellular loops of FpvA and second, a stabilization of the ligand within the binding pocket. Mutagenesis studies[27] and experiments using fluorescent-labeled FepA,[118] the ferric-enterobactin receptor in *E. coli*, suggested the existence of dual ligand-binding sites for this receptor. The initial site located in the extracellular loops contains aromatic residues and probably functions through hydrophobic interactions. The secondary site, located deeper in the protein, contains both charged and aromatic residues.[27] Deletion of loop 3 or 11 of the *E. coli* FhuA receptor inactivated the ferrichrome transport activity, suggesting the existence of a second binding site for this transporter located in the extracellular loops.[48] However, the crystal structure of FpvA, FecA, FhuA, and FepA[24,51,92,168] shows a single binding pocket for the corresponding ferric-siderophore. The Pvd and metal-Pvd kinetic data can be interpreted as follows: iron-free or Fe-Pvd first binds to the single binding site, and then induces a change in the conformation of FpvA. This hypothesis is fully consistent with the time resolved fluorescent spectroscopy data on the FpvA-Pvd

and FpvA-Pvd-Ga complexes.[53] When the siderophore is metal loaded, the Pvd is less solvent accessible and less mobile, like trapped in its binding site. Moreover, cross-linking reactions[141] and spectroscopic observations[28] on FepA suggest loop conformational changes during ligand binding and transport. All these data raise the possibility that such conformational changes in some of the extracellular loops of the transporter are a general facet of siderophore transport dynamics. In the crystal structure of FhuA, these changes were not observed,[52,92] but according molecular dynamics simulations of FhuA, loop L8 appears to be also involved in a mechanism whereby the binding site is gated closed upon ligand binding.[49] Of course, all the data presented here are also consistent with a mechanism in which a binding site at the level of the extracellular loops first recognizes the ferric-siderophore. As a second step, the ligand migrates deeper into the protein and associates with a secondary binding site localized on the plug domain. This induces a conformation change of some extracellular loops and traps the ferric-siderophore in its secondary binding site.

So far, it is not known if the binding of apo- and ferric-siderophore to a common binding site is a feature reserved to the subfamily of signal transducing siderophore receptors or a common feature to all siderophore OMTs in Gram-negative bacteria. Further investigations will be necessary to answer this question, especially on receptors lacking the N-terminus extension.

2.3. Ferric-Siderophore Transport by the Outer Membrane Transporters

The transport of ferric-siderophore across the outer membrane via the OMT consumes energy. There is no known energy source in the outer membrane or in the adjacent periplasmic space. The energy is provided by the electrochemical potential across the cytoplasmic membrane and flows from the cytoplasmic membrane into the OMT via the TonB machinery. This is a protein complex anchored to the cytoplasmic membrane and composed of TonB, ExbB, and ExbD.[85,107,149]

2.3.1. Structure and Stoichiometry of the TonB–ExbB–ExbD Complex

Three *tonB* genes (*tonB1*,[123] *tonB2*,[169] and *tonB3*[68]) have been identified in *P. aeruginosa*. A fourth gene (PA0695)[150] codes for a protein showing weak similarity to the known TonB proteins. Whereas homolog sets of *exbB* and *exbD* have been found linked to *tonB2*[169] and PA0695,[150] in an apparent operon structure, no *exbB* and *exbD* set have been found near the *tonB1* gene. TonB1 works with so far unidentified ExbB and ExbD proteins. Disruption of the *tonB1* gene abrogates siderophore-mediated iron uptake[123] and heme uptake,[169] but

inactivation of *tonB2*, *exbB*, or *exbD* has no adverse effect on iron or heme acquisition in this organism,[169] suggesting that TonB1 is more important for high-affinity iron uptake than TonB2. The involvement of *tonB3* and PA0695 in iron acquisition in *P. aeruginosa* has not yet been tested. Although TonB1 displays sequence significant homology to *E. coli* TonB, the TonB1 protein is the largest TonB protein with an additional 90 amino acid residues at it N-terminus.[123] In *B. cepacia*, *tonB–exbB–exbD* genes have been identified and their involvement in iron acquisition confirmed.[11]

The organization of TonB–ExbB–ExbD complex in the inner membrane and its stoichiometry has been mostly studied in *E. coli* (for review see ref [126]). TonB is anchored to the cytoplasmic membrane by an uncleaved signal peptide and contains a distinctive proline-rich region, which extends the C-terminus of TonB out of the cytoplasmic membrane. ExbD is also N-terminally anchored in the cytoplasmic membrane and extends into the periplasm.[77] ExbB, on the other hand, spans the cytoplasmic membrane three times and is mostly located in the cytoplasm.[76] ExbB and ExbD are therefore topologically partitioned to opposite sides of the cytoplasmic membrane and cross-linking experiments suggest that ExbB and ExbD may exist in a complex with the following ratio: 1 TonB:2 ExbD:7 ExbB.[67] The crystal structure of the C-terminal-soluble domain of TonB of *E. coli* has been solved. The structure shows tight intertwined dimers,[30] suggesting that TonB might be active as a dimer, which is in good agreement with many biochemical data.[25,78,80,115] If it is assumed that TonB is active under a dimeric form, the TonB machinery will be a complex of 520 kDa (2 TonB:4 ExbD:14 ExbB). In this quite huge structure, ExbB and ExbD proteins probably interact through transmembrane segments and together may provide a protein channel for the transmembrane segment of TonB.

The process by which the TonB machinery, composed of several ExbB and ExbD proteins, transfers the energy from the inner membrane to the receptor located in the outer membrane remains unknown. Sequence alignment has shown some homologies between ExbB and MotA and has suggested a similar topology for ExbD and MotB.[29,30] MotA and MotB are believed to form the stator of the motor in bacterial flagellum and the TonB–ExbB–ExbD complex may have common features with molecular machines such as the flagellum or ATP synthetases. In the TonB–ExbB–ExbD complex the pmf might cause as well a torsion motion, which is transduced to the TonB C-terminal domain via the stiff proline-rich region. Then, TonB physically contacts a TonB-dependent OMT, triggering a conformational change in the OMT such that ligand is pumped into the periplasmic space.

Clearly, the mechanism of energy transfer from the cytoplasmic membrane into the outer membrane via the TonB machinery is still unclear at the molecular level, and remains a major question in the field of ferric-siderophore uptake in Gram-negative bacteria.

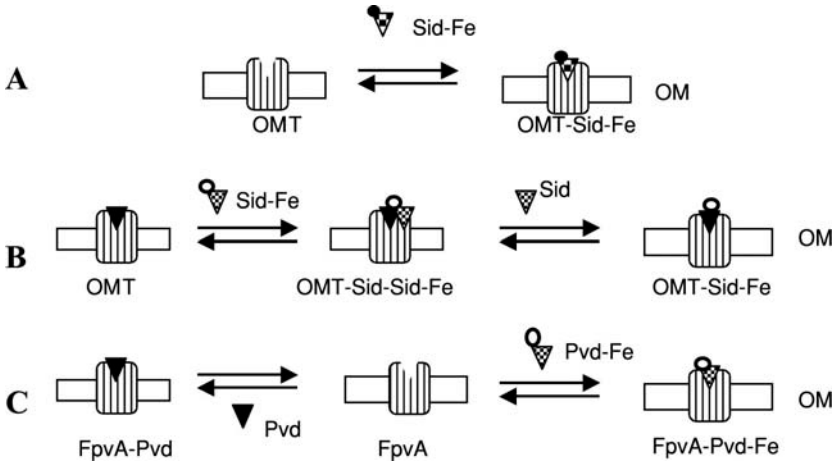


Figure 5. Three different mechanisms describing the formation of the receptor-siderophore-Fe complex in Gram-negative bacteria. In the first mechanism (a), the ferric-siderophore binds to the OMT, which has an empty binding site, and is transported as such across the outer membrane. In the second mechanism (b), the OMT has its binding site loaded with a molecule of iron-free siderophore. The extracellular ferric-siderophore acts as an iron donor for the already bound siderophore on the OMT. In this mechanism, different ferric-siderophores may supply iron to the siderophore bound to the OMT.[146] In the third mechanism (c), the OMT is also loaded with a molecule of iron-free siderophore. However, in this case the extracellular siderophore is able to bind to its binding site on the OMT only after dissociation of the already bound apo-siderophore from the receptor.[31,135]

2.3.2. Formation of the Receptor-Siderophore-Iron Complex

The first step for iron uptake across the outer membrane via siderophores is the formation of an OMT-siderophore-iron complex. Three different mechanisms have been described in the literature for the formation of this complex in Gram-negative bacteria[134] (Figure 5). The first mechanism (Figure 5A) has been described for FhuA and FepA, the ferrichrome and the ferric-enterobactin receptor in *E. coli*, respectively. In this case, the binding site of the receptor is free, and the ferric-siderophore formed in the extracellular medium, binds to it in a TonB and pmf independent way. In a second mechanism (Figure 5B), the extracellular ferric-siderophore acts as an iron donor and brings the iron to the apo-siderophore already bound to the membrane receptor. This mechanism is proposed to work out in *Aeromonas hydrophila*.[146] The third mechanism (Figure 5C) has been described in *P. aeruginosa* for the Pvd pathway and is a mechanism of siderophore displacement on the FpvA receptor, in a two step process: dissociation of the bound Pvd from FpvA followed by subsequent binding of Pvd-Fe.[31,135]

Kinetic studies showed that the FpvA-Pvd complex, which is the normal state of the receptor under iron limitation (see Section 2.2.2), is extremely

stable and needs to be activated by the TonB machinery and the pmf of the inner membrane to get a fast dissociation of the bound iron-free Pvd.[31,135] *In vitro* or in the absence of TonB and pmf, this FpvA-Pvd complex dissociates in the presence of an excess of Pvd-Fe, with a half life time of 25 h.[31,136] *In vivo*, the FpvA receptor is highly activated in the presence of TonB and pmf and dissociates by a so far unknown mechanism with a half life time of 4 min.[31] Such a regulating function of the binding properties of an OMT by the TonB machinery and the pmf has been described only for FpvA and the hemophore receptor HasR in *S. marcescens*[87] (see below in Iron uptake via hemophores). It is possible that TonB and pmf regulate the affinity of FpvA for its apo-siderophore but this regulation cannot be seen under the experimental conditions used. Experiments using the fluorescent properties of Pvd clearly showed that during iron uptake, TonB activates only a few FpvA-Pvd complexes to get a fast release of the apo-siderophore.[31] TonB levels in the cells are indeed likely to be much lower than the number of transport proteins.[124,125,127] Only a few OMTs are activated by the TonB machinery and most likely during a short time.[31,74] Once the FpvA receptor binding site is empty, Pvd-Fe or another molecule of Pvd can bind.

2.3.3. Ligand Transport by the OMT

Binding of the ferric-siderophore to the OMT is not enough to start the translocation process of the iron-loaded ligand across the outer membrane. It is an energy consuming process, which involves the pmf of the inner membrane, the TonB machinery and a conserved N-terminal motif of the OMTs called the TonB box. This region is important for the interaction between the OMT and the TonB protein. Many studies in the literature emphasize the importance of the TonB box in ferric-siderophore uptake and as a mediator of some physical interaction between TonB and TonB-dependent receptors.[26,73,108,142] Full length FhuA or BtuB (*E. coli* vitamin B12 OMT), mutants carrying an inactive TonB box show no iron or vitamin B12 uptake.[14,66,138] Site-directed spin labeling and EPR experiments on BtuB have shown that in the absence of substrate the TonB box exists in a structured helical conformation that contacts the barrel of the transporter.[96] Addition of substrate converts this segment into an extended structure that is highly dynamic, disordered, and probably extended into the periplasm, where it associates with TonB. The TonB box is disordered in the crystals structure of FhuA, but a conformational change occurs when the iron-siderophore binds to the receptor. Moreover, an unfolding of the N-terminal helix (H_1) in the central hatch near the N-terminus was observed in the presence of ferrichrome.[52,92] In FecA bound to ferric-citrate, electron density for the N-terminal residues was not seen.[51,168] The flexibility of the TonB box and the increased flexibility of the N-terminal part in the cork domain upon binding of the ferric-siderophore through an allosteric mechanism, are probably crucial for interactions with the TonB protein.

The mechanism of the siderophore-mediated iron translocation across the outer membrane itself remains unsolved despite the structural information available for FpvA and for four *E. coli* OMTs. It seems unlikely that the cork domain of the OMT is removed during iron uptake and floats freely in the periplasm. The energy needed to break each of the 70 hydrogen bounds (in the case of FecA) that position the globular domain into the lumen of the β-barrel would be too high (>100 kcal/mol). It is more likely that a conformational change in the cork domain allows the formation of a passage for the ferric-siderophore between the barrel and the cork. The only experiment showing conformational changes during iron uptake has been done by Klebba and co-workers on FepA, using site-directed spin labeling and electron spin resonance spectroscopy.[91] They observed conformational changes in the FepA receptor during iron uptake *in vivo*[72] that were TonB, energy and temperature dependent. The crystal structures of FpvA, FhuA, FecA, and FepA show a continuous water filled channel, which is going from the extracellular face to the periplasm, but which is too small to allow the iron-siderophore transport across the outer membrane because of the amino acid side chains. This channel, under the activation of the TonB machinery, may be widened and could be a pathway for the ferric-siderophores to reach the periplasm. Usher *et al*. showed that the over-expressed cork domain of FepA is unfolded in solution but still able to bind the ferric-siderophore.[155] These observations support a structural change of the plug domain for the transport of the ferric-siderophore along the channel. However, more investigation is needed to clarify the translocation mechanism of ferric-siderophore across the outer membrane via these OMTs.

2.4. Proposed Pvd-Fe Uptake Mechanism Via FpvA in *P. aeruginosa*

Among all possible iron uptake pathway in *P. aeruginosa*, the Pvd-Fe uptake mechanism is the best known at the molecular level and is summarized in Figure 6. In this mechanism, the normal state of the FpvA receptor in the outer membrane, under iron-limited conditions, is the FpvA-Pvd complex.[135,136] This complex is extremely stable and activation of the transporter by the pmf and the TonB machinery is needed to get a fast dissociation of the ligand and generate an OMT with a binding site ready for Pvd-Fe binding and uptake.[31] Binding assays have shown that both, Pvd and Pvd-Fe, have similar affinities (a 10-fold difference only) for FpvA and are in competition for a common or overlapping binding site.[31,136] Nonetheless, the binding kinetics for the apo form are appreciably slower than for the ferric form.[31] In the case of binding of apo-Pvd, the TonB machinery will activate again the receptor for a new cycle until Pvd-Fe binding. The binding of Pvd-Fe is a two-step process: the bimolecular step (association of the ligand with the receptor) is followed by a slower step that

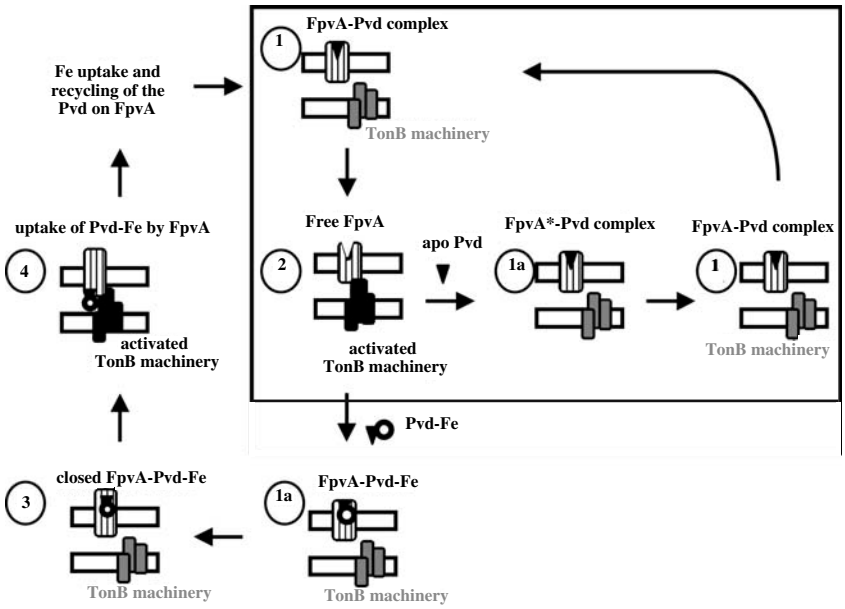


Figure 6. Mechanism of iron uptake by the Pvd pathway in *P. aeruginosa*. In this proposed mechanism (based on the FecA structures and the functional data for FpvA), the FpvA receptors at the cell surface are loaded with iron-free Pvd, under iron-limited conditions (1). Release of the iron-free Pvd occurs only after activation of the FpvA receptor by the TonB machinery. Once the receptor has an empty binding site (2), two events may occur: either an iron-free Pvd binds again to the transporter with formation of an FpvA-Pvd complex (1) or the Pvd-Fe binds to the transporter (3). In the first case, TonB will activate the FpvA receptor until there is binding of ferric-siderophore. In the second case, the binding of ferric-siderophore (3) induces a change of conformation in the receptor, which traps the ferric-siderophore in its binding site (4). The FpvA-Pvd-Fe complex is then ready for transport. The translocation of ferric-siderophore is induced by TonB activation of the transporter (5). Afterward, the ferric-siderophore is in most cases transported into the cytoplasm by an ABC transporter, where iron is released from the siderophore, and the iron-free siderophore is recycled again to the extracellular surface, where it can bind an empty receptor (1) and a new cycle can start.

presumably corresponds to a change of conformation of FpvA.[31] Time resolved fluorescence spectroscopy studies have shown a Pvd, which is less solvent accessible and less mobile in the FpvA-Pvd-Ga complex than in the FpvA-Pvd complex.[53] The binding of the ferric-siderophore may induce a change of conformation of some extracellular loops, as described for FecA,[51,168] which traps the ferric-siderophore in its binding site. This conformation of the OMT can be considered as competent for ferric-siderophore uptake after activation of the transporter by the TonB machinery. The mechanism of translocation of the Pvd-Fe through the FpvA structure and the mechanism of iron release

from Pvd remain unknown. Based on the fluorescent properties of Pvd it could be shown that the iron-free Pvd is recycled on the FpvA receptor and in the extracellular medium after iron release.[133] It is not clear whether iron-free Pvd is recycled first through a step involving the formation of a FpvA-Pvd complex and subsequent released into the medium or whether the Pvd is first released into the medium and then binds to FpvA in the outer membranes.

This proposed mechanism is consistent with the turnovers usually observed for ferric-siderophore uptake in Gram-negative bacteria (1 Pvd-Fe/min/FpvA and 6.4 enterobactin-Fe/min/FepA[140]). Such iron uptake rates are more than sufficient to satisfy the iron requirement of a cell, which is in the order of 10^5 iron ion per generation, through approximately 1000 receptor molecules present in the outer membrane under iron limitation conditions.[20]

The biological function of the binding of apo-Pvd to FpvA remains unknown. At first sight, this binding does not seem to play a key role in the iron uptake mechanism, except in the case the Pvd is recycled in the extracellular medium via FpvA. The binding of iron-free siderophore to an OMT may be involved in the regulation of iron uptake, which must be strictly controlled to avoid the deleterious effects of excessive or insufficient iron levels. In *P. aeruginosa*, the loading status of FpvA (iron-free Pvd vs. Pvd-Fe) depends on the relative concentrations of the two Pvd forms in the medium, and this property may be linked to a regulatory role. This idea is discussed in the paragraph below, describing the iron uptake regulation.

3. HEME ACQUISITION

Heme uptake has also been suggested to play an important role in *P. aeruginosa* infections.[151] Two distinct heme uptake systems, Fur-repressible, encoded by the *phu* and *has* loci, have been described in this bacterium.[113] The *phu* genes include *phuR*, which encodes an outer membrane heme receptor, and *phuSTUVW*, which encodes a typical ABC transporter. The heme uptake in other Gram-negative bacteria involves the same proteic partner as ferric-siderophore uptake, e.g. a specific OMT that is activated by a TonB machinery and an ABC transporter (Figure 1). The second heme uptake system is composed of a heme-OMT gene, *hasR*, in an operon with *hasA* a heme-binding extracellular protein. This uptake system was identified as a homolog of the well characterized Has system in *S. marcescens*.[90]

In *S. marcescens* and several Gram-negative bacteria, heme acquisition involves a secreted heme-binding protein called hemophore, which extracts heme from various hemoproteins and delivers it to a specific TonB-dependent OMT.[164] Extracellular release of HasA in *S. marcescens* and *P. fluorescens* requires a type I secretion apparatus of the ABC family (*hasDEF* operon).[70,89] The well characterized *S. marcescens* hemophore is a monomer, which binds

heme with a stoichiometry of one to one (1:1) and an affinity lower than 10^{-9} M[71]. The crystal structure of the holoprotein has been solved and found to consist of a single module with two residues interacting with the heme.[10] The hemophore receptor HasR in the outer membrane of *S. marcescens* recognizes the heme-free and heme-loaded hemophores with similar affinities (10^{-10} M) and at a unique or overlapping sites. It also recognizes the free or hemoglobin-bound heme.[88] Two *tonB* genes, *tonB* and *hasB,* have been identified in *S. marcescens*.[117] Though TonB and HasB are significantly similar and can replace each other for heme acquisition, only TonB mediates iron acquisition from iron sources other than heme and hemoproteins.[117] For the two *tonB* genes characterized in *P. aeruginosa*, iron and heme acquisition appear to depend on the TonB1 protein.[170] The function of the HasB machinery is to activate the HasR receptor for the uptake of the heme moiety (apo-HasA stays bound to the receptor at the extracellular side) and then subsequently to activate HasR to get dissociation of the hemophore HasA. Over-expression of the HasB machinery (HasB plus corresponding ExbB and ExbD proteins) in *S. marcescens* speeds up the release of empty hemophores from HasR.[87] The mechanism by which HasB complex in the inner membrane causes heme uptake and hemophore release at the cell surface remains unclear.

The heme uptake mechanism of *S. marcescens* clearly shows homology with the Pvd-Fe uptake in *P. aeruginosa* and is summarized in Figure 7. The first major homology is a similar binding affinity of the apo and ferric ligand to a common binding site on the OMT. The second homology between heme uptake via HasA and iron uptake via Pvd is the activation of the release of the apo ligand from the receptor by the TonB machinery. This parallel between these two different iron uptake mechanisms suggests that the mechanism described in Figure 6 for FpvA may be not only specific for FpvA but may be a more general iron uptake mechanism among Gram-negative bacteria.

Sequence alignment studies showed also that HasR of *P. aeruginosa* and *S. marcescens* have both, like FpvA and FecA, an additional N-terminal extension.[54,137,159] Like FpvA and FecA, the HasR receptor belongs to a sub-family of OMTs, with dual function: (i) specific ligand transport across the outer membrane and (ii) sensing of the presence of the loaded ligand to trigger transcription induction.[18] Binding of the heme-loaded hemophore to HasR was shown to be the stimulus for the HasR-mediated signal transduction.[131] In the case of Pvd, it has not yet been determined if apo- or Pvd-Fe were the inducers.[83]

4. TRANSPORT ACROSS THE INNER MEMBRANE

After transport across the outer membrane, the ferric-siderophore complex binds to a periplasmic binding protein that delivers the iron compounds to the integral cytoplasmic membrane proteins of an ABC transporter. This step

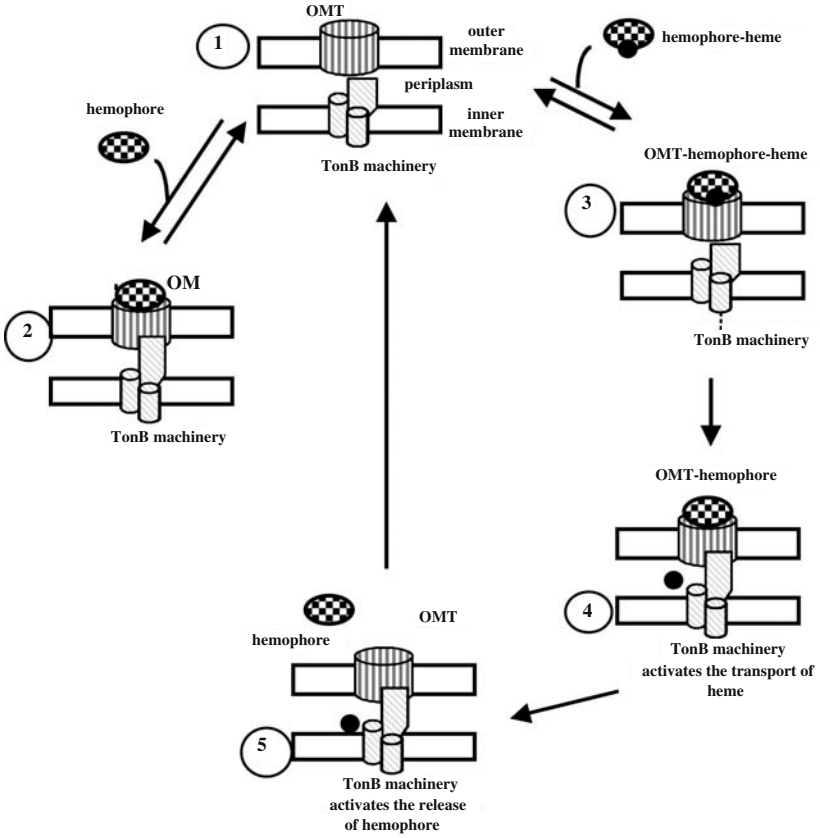


Figure 7. Mechanism of heme uptake in *S. marcescens*. In this mechanism, the HasR receptor binds with the same affinity and to a common binding site both apo- and holo-hemophore (HasA). Two events may occur after step (1): either an heme-free HasA binds to the transporter with formation of a HasR–HasA complex (2), or the holo-HasA binds to the transporter (3). In the case of the binding of holo-HasA (3), the translocation of heme is induced by TonB activation of the transporter (4). Only the heme is transported into the periplasm, HasA stays bound to HasR until the TonB machinery activates HasR to get a fast release of the hemophore.

of the iron acquisition in pseudomonads has received little if any attention and almost no data are available.

In general, ABC transporters involved in iron uptake are composed of: (i) one or several periplasmic substrate binding proteins, (ii) one or two different (homodimer or heterodimer) integral membrane proteins, and (iii) one or two different ATP hydrolases that face the cytoplasm and supply the system with energy.[82] Often there are more different TonB-dependent OMTs in a bacterial cell than corresponding ABC transporters. In *E. coli*, the siderophore receptor

displays higher substrate specificity than the proteins of the ABC transporter.[20] This difference can be most striking, as for instance in *P. aeruginosa*: 23 putative TonB-dependent receptors were identified in the *P. aeruginosa* genome[150] and only four iron-related ABC transporters were found to exist in this organism. The *FepBCDG* homologs of the ABC transporter involved in the uptake of ferric-enterobactin in *E. coli* could be found in the *P. aeruginosa* genome sequence. Homologs of the *E. coli* FeoAB Fe(II) transporter (PA4359–PA4358),[75] of the *S. marcescens* SfuABC[7] (PA5216–PA5217), and of the *Hemophilus influenzae* HitABC[4] (PA4687–PA4688) Fe(III) uptake systems have also been found. But no ABC transporter involved in metal uptake could be localized in the *pch-fptA* locus of pyochelin synthesis and uptake, in the vicinity of *fecA-fecIR* genes involved in the ferric-dicitrate uptake. An ABC transporter can be found in the pvd locus (PA2407–PA2408–PA2409), but the periplasmic binding protein (PA2407) of this transporter seems not to be involved in Pvd-Fe uptake.[114]

According the genome of *P. aeruginosa*, there is clearly a lack of ABC transporter compared to the number of OMTs present. It is unlikely that the ABC transporters involved in the iron uptake in *P. aeruginosa* have a very large specificity and are able to transport ferric-siderophore complexes with different structures. It is probably more reasonable to think that for most of the ferric-siderophore complexes used by *P. aeruginosa*, the dissociation of iron from the siderophore occurs in the periplasm, and only iron is transported in the cytoplasm by one or two of the ABC transporters. Such a mechanism would probably involve the reduction of Fe(III) into Fe(II), in order to facilitate the dissociation step of the iron from its siderophore in the periplasm. The decrease in transport observed in the presence of dipyridyl suggests indeed involvement of siderophore reduction in the process of iron dissociation.[132] Moreover, studies using cells osmotically shocked after incubating with [^{55}Fe]ferri[^{14}C]Pvd[132] and Mössbauer spectroscopy studies[105] suggest a separation of metal and ligand in the periplasmic space for the Pvd iron uptake pathway. These different data are consistent with a mechanism where iron is released from the siderophore by a reduction process in the periplasm, then transported by an ABC transporter into the cytoplasm, but more studies are necessary to demonstrate this hypothesis.

5. IRON TRANSPORT REGULATION

The iron content of the cells must be regulated to conserve energy and substrates, and to avoid iron toxicity. In Gram-negative bacteria, iron regulation is mediated by the Fur protein, which represses the transcription of numerous genes involved in the synthesis of siderophores and in the iron uptake (for a review see refs [64,157]). Fur requires iron (corepressor) in order to bind to a

target sequence ("Fur box") in the promoter region of iron-regulated genes and block their transcription when the level of intracellular iron (Fe^{2+}) reaches a threshold. By contrast, when the cells are iron starved, the apo form of Fur loses its ability to bind DNA and gene transcription occurs. The crystal structure of *P. aeruginosa* Fur protein has been solved.[119] The protein is composed of two domains, the N-terminal domain implicated in DNA binding and the C-terminal domain responsible for homodimerization. The *fur* gene of *B. cepacia* and *P. putida* have been identified and also cloned.[93,158]

The Fur protein is not the only regulator of iron acquisition. Additional regulatory devices, acting positively on the expression of iron uptake genes, have been identified in *Pseudomonas*. Expression of the ferripyochelin receptor FptA in *P. aeruginosa* is pyochelin inducible, with induction involving an AraC-type transcriptional activator (PchR) as well as pyochelin and the FptA transporter itself.[65] The expression of the ferric-enterobactin receptor in *P. aeruginosa* (PfeA) relies on a sensor protein PfeS, belonging to the histidine protein kinase superfamily, and a response regulator PfeR that activates *pfeA* expression following phosphorylation by PfeS.[43] A similar mode of signal transduction has also been reported for PirA, a low affinity receptor for ferric-enterobactin that responds to the iron-regulated PirR–PirS system.[157] At last, a regulatory cascade involving extracytoplasmic function (ECF) sigma factor has been best characterized for the ferric-citrate uptake system of *E*.coli (for a review see ref [23]), the Pvd-Fe uptake system in *P. aeruginosa* (for reviews see refs [157,163]) and the heme uptake system via HasR in *S. marcescens*.[16] Transcription induction is initiated at the cell surface, and a signal is transmitted to the cytoplasm by a signaling mechanism involving three components: an OMT bound with its ferric ligand, an inner membrane regulator protein (also referred to as an anti-sigma factor), and a cytoplasmic sigma factor belonging to the ECF family. The interaction of ferric ligand with its cognate transporter is thought to induce a conformational change in the OMT that is transmitted by an energy-driven, TonB-dependent mechanism and via the N-terminal domain of the OMT, to the inner membrane regulator.[79] Once the OMT has signaled that a ferric-siderophore is bound, the inner membrane regulator modulates the activity of a specific ECF sigma factor, which in turn binds to an RNA polymerase core enzyme and initiates transcription of the iron transport operon. In the *E. coli* Fec system, the anti-sigma factor is FecR and the sigma factor, FecI. In *P. aeruginosa*, the anti-sigma factor FpvR regulates two ECF sigma factors, FpvI and PvdS.[13,163] This is the first example of an anti-sigma factor (FpvR) that directly regulates the activities of two different ECF sigma factors, involving branched signaling system. FpvI binds RNA polymerase and initiates transcription of *fpvA*. PvdS initiates transcription of genes required for the production of Pvd, as well as for the secreted proteins, exotoxin A and PrL endoprotease. This signaling pathway has also clear parallels with the PupB/PupR/PupI system in *P. putida* WCS358.[81]

In the case of the FecA receptor, it seems that ferric-citrate induces the signal transduction via FecR and FecI.[69] In *P. aeruginosa*, the presence of Pvd in the extracellular medium positively regulates the expression of *fpvA*,[59] suggesting that the binding of apo-Pvd to FpvA may be involved in this regulation process. Moreover, the loading status of FpvA (iron-free Pvd vs. Pvd-Fe) depends on the relative concentrations of the two Pvd forms in the medium, and this property may be linked to a regulatory role in *P. aeruginosa*. FpvA may sense iron availability and then either interacts or not with the signal transduction machinery, depending on its loading status with iron-free Pvd or Pvd-Fe. For species like *P. aeruginosa* that possess multiple endogenous siderophore systems, autoinduction of siderophore synthesis and expression of the corresponding OMT, following siderophore interaction with the cognate OMT, could represent a convenient strategy for ensuring selective expression of the most effective iron carrier under particular environmental conditions. Here also further studies are necessary to understand if the binding of an apo-siderophore to its OMT may be involved in the activation of the signal transduction.

6. CONCLUSIONS

The presence in the genome of *P. aeruginosa* of a number of genes encoding putative siderophores OMTs, the ability to use multiple sources of iron for survival and the regulation of known virulence factors by iron highlight the importance of iron for this bacterium and, more in general, for *Pseudomonas*.

Until recently, research on siderophore-mediated iron uptake in *Pseudomonas* was mostly focused on diverse Pvds, and more specifically, on the Pvd uptake pathway in *P. aeruginosa*. The fluorescent properties of Pvd were crucial in providing valuable insights into the understanding of the ferric-siderophore uptake process across the membranes of *P. aeruginosa* (Figure 6). The mechanism proposed from these studies has interesting similarities with the heme uptake via the HasR receptor in *S. marcescens* (Figure 7). In both mechanisms, the ligand (siderophore or hemophore) loaded or not with iron (or heme) binds with close affinities to a common site or to an overlapping binding site on the receptor. Additionally, in both cases the receptor loaded with the apo ligand (apo-siderophore or apo-hemophore) is a very stable complex. An activation of the receptor by the pmf and the TonB machinery is necessary to get a fast dissociation of the apo ligand from the OMT. These similarities in the binding properties of FpvA, HasR, and also FecA and in the iron uptake mechanisms proposed for these receptors suggest that the iron uptake mechanism in Figure 6 is probably not only specific for FpvA but may be a more general mechanism among Gram-negative bacteria.

Despite the determination of the crystal structure of FpvA, FhuA, FepA, and FecA many questions remain unanswered. For instance, the mechanism of

translocation of the ferric-siderophore through a structure composed of a β-barrel domain with the lumen closed by a plug remains unsolved. It is also not yet known how OMTs receive and respond to the pmf and how the TonB machinery transduces energy from the proton gradient of the cytoplasmic membrane to the OMTs.

Another challenge for the future will be to go beyond genomic analysis of all iron putative uptake systems present in the genomes of *Pseudomonas* and characterize them. Transcriptome and proteome approaches are certainly interesting tools to reach these goals. The last 2 years, *P. aeruginosa* proteomic and transcriptome profiling data were reviewed for different environment conditions, among other iron starvation conditions.[61,114,116] Under such conditions, transcriptome analysis confirmed the iron-responsive expression of known genes, but also the identification of many novel iron-regulated genes of unknown function, suggesting that they are involved directly or indirectly in iron metabolism or metabolic adaptation to different iron-availability conditions.[114,116] More recently, the transcription profile of *P. aeruginosa* after interactions with primary normal human airway epithelial cells was determined using Affymetrix GeneChip technology. Surprisingly, the gene expression profiles indicated repression of iron acquisition genes.[55] The number of genes showing these trends increased over time, suggesting that *P. aeruginosa* may be able to acquire ample iron for growth from the epithelial cells during infection.

The recent completion of the *Pseudomonas* Genome Project, in conjunction with the *Pseudomonas* Community Annotation Project (PseudoCAP) has fast-tracked the ability to apply the tools encompassed under the term proteomics or transcriptome to this pathogen. Such global approaches combined with biochemistry, molecular biology, and microbiology studies will allow the research community to answer long-standing questions regarding the ability of *P. aeruginosa* to survive diverse habitats, its pathogenic nature toward humans, and concerning iron metabolism identification of all iron uptake pathways and the regulations involved.

ACKNOWLEDGMENTS

I am very grateful to Dr D. Cobessi (ESBS, Illkirch) for having provided the figure showing the crystal structure of FpvA. I also acknowledge Drs C. Hihi and D. Cobessi for critical reading the manuscript. This work was supported by Vaincre la Mucoviscidose, by the ACI Physique et Chimie du Vivant program from the Ministère de l'Enseignement Supérieur de la Recherche et de la Technologie and by the Center National de la Recherche Scientifique (CNRS).

REFERENCES

1. Abdallah, M.A., and Pattus, F., 2000, Siderophores and iron-transport in microorganisms. *J. Chin. Chem. Soc.*, 47:1–20.
2. Ackerley, D.F., Caradoc-Davies, T.T., and Lamont, I.L., 2003, Substrate specificity of the non-ribosomal peptide synthetase PvdD from *Pseudomonas aeruginosa*. *J. Bacteriol.*, 185:2848–2855.
3. Adams, C., Dowling, D.N., O'Sullivan, D.J., and O'Gara, F., 1994, Isolation of a gene (*pbsC*) required for siderophore biosynthesis in fluorescent *Pseudomonas* sp. strain M114. *Mol. Gen. Genet.*, 243:515–524.
4. Adhikari, P., Kirby, S.D., Nowalk, A.J., Veraldi, K.L., Schryvers, A.B., and Mietzner, T.A., 1995, Biochemical characterization of a *Haemophilus influenzae* periplasmic iron transport operon. *J. Biol. Chem.*, 270:25142–25149.
5. Albrecht-Garry, A.M., Blanc, S., Rochel, N., Ocacktan, A.Z., and Abdallah, M.A., 1994, Bacterial iron transport: coordination properties of pyoverdin PaA, a peptidic siderophore of *Pseudomonas aeruginosa*. *Inorg. Chem.*, 33:6391–6402.
6. Ambrosi, C., Leoni, L., Putignani, L., Orsi, N., and Visca, P., 2000, Pseudobactin biogenesis in the plant growth-promoting rhizobacterium *Pseudomonas* strain B10: identification and functional analysis of the L-ornithine N(5)-oxygenase (*psbA*) gene. *J. Bacteriol.*, 182:6233–6238.
7. Angerer, A., Gaisser, S., and Braun, V., 1990, Nucleotide sequences of the *sfuA*, *sfuB*, and *sfuC* genes of *Serratia marcescens* suggest a periplasmic-binding-protein-dependent iron transport mechanism. *J. Bacteriol.*, 172:572–578.
8. Ankenbauer, R.G., Toyokuni, T., Staley, A., Rinehart, K.L., Jr., and Cox, C.D., 1988, Synthesis and biological activity of pyochelin, a siderophore of *Pseudomonas aeruginosa*. *J. Bacteriol.*, 170:5344–5351.
9. Ankenbauer, R.G., and Quan, H.N., 1994, FptA, the Fe(III)-pyochelin receptor of *Pseudomonas aeruginosa*: a phenolate siderophore receptor homologous to hydroxamate siderophore receptors. *J. Bacteriol.*, 176:307–319.
10. Arnoux, P., Haser, R., Izadi, N., Lecroisey, A., Delepierre, M., Wandersman, C., and Czjzek, M., 1999, The crystal structure of HasA, a hemophore secreted by *Serratia marcescens*. *Nat. Struct. Biol.*, 6:516–520.
11. Asghar, A.H., and Thomas, M.G., 2001, Role of the TonB system in iron uptake by *Burkholderia cepacia*. In Abstracts of *Pseudomonas* 2001, Brussels, Belgium (abstr 46:65).
12. Atkinson, R.A., Salah El Din, A.L., Kieffer, B., Lefevre, J.F., and Abdallah, M.A., 1998, Bacterial iron transport: ^{1}H NMR determination of the three-dimensional structure of the gallium complex of pyoverdin G4R, the peptidic siderophore of *Pseudomonas putida* G4R. *Biochemistry*, 37:15965–15973.
13. Beare, P.A., For, R.J., Martin, L.W., and Lamont, I.L., 2003, Siderophore-mediated cell signalling in *Pseudomonas aeruginosa*: divergent pathways regulate virulence factor production and siderophore receptor synthesis. *Mol. Microbiol.*, 47:195–207.
14. Bell, P.E., Nau, C.D., Brown, J.T., Konisky, J., and Kadner, R.J., 1990, Genetic suppression demonstrates interaction of TonB protein with outer membrane transport proteins in *Escherichia coli*. *J. Bacteriol.*, 172:3826–3829.
15. Bitter, W., Marugg, J.D., de Weger, L.A., Tommassen, J., and Weisbeek, P.J., 1991, The ferric-pseudobactin receptor PupA of *Pseudomonas putida* WCS358: homology to TonB-dependent *Escherichia coli* receptors and specificity of the protein. *Mol. Microbiol.*, 5:647–655.
16. Biville, F., Cwerman, H., Letoffe, S., Rossi, M.S., Drouet, V., Ghigo, J.M., and Wandersman, C., 2004, Haemophore-mediated signalling in *Serratia marcescens*: a new mode of regulation

for an extra cytoplasmic function (ECF) sigma factor involved in haem acquisition. *Mol. Microbiol.*, 53:1267–1277.
17. Boukhalfa, H., and Crumbliss, A.L., 2002, Chemical aspects of siderophore mediated iron transport. *Biometals*, 15:325–339.
18. Braun, V., 1997, Surface signaling: novel transcription initiation mechanism starting from the cell surface. *Arch. Microbiol.*, 167:325–331.
19. Braun, V., 2001, Iron uptake mechanisms and their regulation in pathogenic bacteria. *Int. J. Med. Microbiol.*, 291:67–79.
20. Braun, V., 2003, Iron uptake by *Escherichia coli*. *Front. Biosci*., 8:s1409–s1421.
21. Braun, V., and Braun, M., 2002, Active transport of iron and siderophore antibiotics. *Curr. Opin. Microbiol.*, 5:194–201.
22. Braun, V., and Braun, M., 2002, Iron transport and signaling in *Escherichia coli*. *FEBS Lett.*, 529:78.
23. Braun, V., Mahren, S., and Ogierman, M., 2003, Regulation of the FecI-type ECF sigma factor by transmembrane signalling. *Curr. Opin. Microbiol.*, 6:173–180.
24. Buchanan, S.K., Smith, B.S., Venkatramani, L., Xia, D., Esser, L., Palnitkar, M., Chakraborty, R., van der Helm, D., and Deisenhofer, J., 1999, Crystal structure of the outer membrane active transporter FepA from *Escherichia coli*. *Nat. Struct. Biol.*, 6:56–63.
25. Cadieux, N., Bradbeer, C., and Kadner, R.J., 2000, Sequence changes in the ton box region of BtuB affect its transport activities and interaction with TonB protein. *J. Bacteriol.*, 182:5954–5961.
26. Cadieux, N., and Kadner, R.J., 1999, Site-directed disulfide bonding reveals an interaction site between energy-coupling protein TonB and BtuB, the outer membrane cobalamin transporter. *Proc. Natl. Acad. Sci. U.S.A.*, 96:10673–10678.
27. Cao, Z., Qi, Z., Sprencel, C., Newton, S.M., and Klebba, P.E., 2000, Aromatic components of two ferric enterobactin binding sites in *Escherichia coli* FepA. *Mol. Microbiol.*, 37:1306–1317.
28. Cao, Z., Warfel, P., Newton, S.M., and Klebba, P.E., 2003, Spectroscopic observations of ferric enterobactin transport. *J. Biol. Chem.*, 278:1022–1028.
29. Cascales, E., Lloubes, R., and Sturgis, J.N., 2001, The TolQ–TolR proteins energize TolA and share homologies with the flagellar motor proteins MotA–MotB. *Mol. Microbiol.*, 42:795–807.
30. Chang, C., Mooser, A., Pluckthun, A., and Wlodawer, A., 2001, Crystal structure of the dimeric C-terminal domain of TonB reveals a novel fold. *J. Biol. Chem.*, 276:27535–27540.
31. Clément, E., Mesini, P.J., Pattus, F., Abdallah, M.A., and Schalk, I.J., 2004, The binding mechanism of pyoverdin with the outer membrane receptor FpvA in *Pseudomonas aeruginosa* is dependent on its iron-loaded status. *Biochemistry*, 43:7954–7965.
32. Cobessi, D., Celia, H., Folschweiller, N., Heymann, M., Schalk, I., Abdallah, M., and Pattus, F., 2004, Crystallization and preliminary X-ray analysis of the outer membrane pyoverdine receptor FpvA from *Pseudomonas aeruginosa*. *Acta Crystallogr. D Biol. Crystallogr.*, 60:1467–1469.
33. Cobessi, D., Célia, H., Folschweiller, N., Schalk, I.J., Abdallah, M.A., and Pattus, F., 2005, The crystal structure of the pyoverdin outer membrane receptor FpvA from *Pseudomonas aeruginosa* at 3.6 Å resolution. *J. Mol. Biol.*, 347: 121–34.
34. Cobessi, D., Celia, H., and Pattus, F., 2004, Crystallization and X-ray diffraction analyses of the outer membrane pyochelin receptor FptA from *Pseudomonas aeruginosa*. *Acta Crystallogr. D Biol. Crystallogr.*, 60:1919–1921.
35. Cornelis, P., Hohnadel, D., and Meyer, J.M., 1989, Evidence for different pyoverdine-mediated iron uptake systems among *Pseudomonas aeruginosa* strains. *Infect. Immun.*, 57:3491–3497.

36. Cornelis, P., and Matthijs, S., 2002, Diversity of siderophore-mediated iron uptake systems in fluorescent pseudomonads: not only pyoverdines. *Environ. Microbiol.*, 4:787–798.
37. Cornelis, P., Moguilevsky, N., Jacques, J.F., and Masson, P.L., 1987, Study of the siderophores and receptors in different clinical isolates of *Pseudomonas aeruginosa*. *Antibiot. Chemother.*, 39:290–306.
38. Cornelissen, C.N., and Sparling, P.F., 1994, Identification of receptor-mediated transferrin-iron uptake mechanism in *Neisseria gonorrhoeae*. *Methods Enzymol.*, 235:356–363.
39. Cox, C.D., Rinehart, K.L., Jr., Moore, M.L., and Cook, J.C., Jr., 1981, Pyochelin: novel structure of an iron-chelating growth promoter for *Pseudomonas aeruginosa*. *Proc. Natl. Acad. Sci. U.S.A.*, 78:4256–4260.
40. Crosa, J.H., and Walsh, C.T., 2002, Genetics and assembly line enzymology of siderophore biosynthesis in bacteria. *Microbiol. Mol. Biol. Rev.*, 66:223–249.
41. Cunliffe, H.E., Merriman, T.R., and Lamont, I.L., 1995, Cloning and characterization of *pvdS*, a gene required for pyoverdine synthesis in *Pseudomonas aeruginosa*: PvdS is probably an alternative sigma factor. *J. Bacteriol*., 177:2744–2750.
42. Darling, P., Chan, M., Cox, A.D., and Sokol, P.A., 1998, Siderophore production by cystic fibrosis isolates of *Burkholderia cepacia*. *Infect. Immun.*, 66:874–877.
43. Dean, C.R., and Poole, K., 1993, Expression of the ferric enterobactin receptor (PfeA) of *Pseudomonas aeruginosa*: involvement of a two-component regulatory system. *Mol. Microbiol.*, 8:1095–1103.
44. de Chial, M., Ghysels, B., Beatson, S.A., Geoffroy, V., Meyer, J.M., Pattery, T., Baysse, C., Chablain, P., Parsons, Y.N., Winstanley, C., Cordwell, S.J., and Cornelis, P., 2003, Identification of type II and type III pyoverdine receptors from *Pseudomonas aeruginosa*. *Microbiology*, 149:821–831.
45. De Vos, D., De Chial, M., Cochez, C., Jansen, S., Tummler, B., Meyer, J.M., and Cornelis, P., 2001, Study of pyoverdine type and production by *Pseudomonas aeruginosa* isolated from cystic fibrosis patients: prevalence of type II pyoverdine isolates and accumulation of pyoverdine-negative mutations. *Arch. Microbiol.*, 175:384–388.
46. Devescovi, G., Aguilar, C., Majolini, M.B., Marugg, J., Weisbeek, P., and Venturi, V., 2001, A siderophore peptide synthetase gene from plant-growth-promoting *Pseudomonas putida* WCS358. *Syst. Appl. Microbiol.*, 24:321–330.
47. Dorrestein, P.C., Poole, K., and Begley, T.P., 2003, Formation of the chromophore of the pyoverdine siderophores by an oxidative cascade. *Org. Lett.*, 5:2215–2217.
48. Endriss, F., and Braun, V., 2004, Loop deletions indicate regions important for FhuA transport and receptor functions in *Escherichia coli*. *J. Bacteriol.*, 186:4818–4823.
49. Faraldo-Gomez, J.D., Smith, G.R., and Samson, M.S.P., 2003, Molecular dynamics simulations of the bacterial outer membrane protein FhuA: a comparative study of the ferrichrome-free and bound States. *Biophys. J.*, 85:1406–1420.
50. Ferguson, A.D., Braun, V., Fiedler, H.P., Coulton, J.W., Diederichs, K., and Welte, W., 2000, Crystal structure of the antibiotic albomycin in complex with the outer membrane transporter FhuA. *Protein Sci*., 9:956–963.
51. Ferguson, A.D., Chakraborty, R., Smith, B.S., Esser, L., van der Helm, D., and Deisenhofer, J., 2002, Structural basis of gating by the outer membrane transporter FecA. *Science*, 295:1715–1719.
52. Ferguson, A.D., Hofmann, E., Coulton, J.W., Diederichs, K., and Welte, W., 1998, Siderophore-mediated iron transport: crystal structure of FhuA with bound lipopolysaccharide. *Science*, 282:2215–2220.
53. Folschweiller, N., Gallay, J., Vincent, M., Abdallah, M.A., Pattus, F., and Schalk, I.J., 2002, The interaction between pyoverdin and its outer membrane receptor in *Pseudomonas aeruginosa* leads to different conformers: a time-resolved fluorescence study. *Biochemistry*, 41:14591–14601.

54. Folschweiller, N., Schalk, I.J., Celia, H., Kieffer, B., Abdallah, M.A., and Pattus, F., 2000, The pyoverdin receptor FpvA, a TonB-dependent receptor involved in iron uptake by *Pseudomonas aeruginosa* (review). *Mol. Membr. Biol.*, 17:123–133.
55. Frisk, A., Schurr, J.R., Wang, G., Bertucci, D.C., Marrero, L., Hwang, S.H., Hassett, D.J., and Schurr, M.J., 2004, Transcriptome analysis of *Pseudomonas aeruginosa* after interaction with human airway epithelial cells. *Infect. Immun.*, 72:5433–5438.
56. Fuchs, R., Schafer, M., Geoffroy, V., and Meyer, J.M., 2001, Siderotyping a powerful tool for the characterization of pyoverdines. *Curr. Top. Med. Chem.*, 1:31–57.
57. Gaille, C., Kast, P., and Haas, D., 2002, Salicylate biosynthesis in *Pseudomonas aeruginosa*. Purification and characterization of PchB, a novel bifunctional enzyme displaying isochorismate pyruvate-lyase and chorismate mutase activities. *J. Biol. Chem.*, 277:21768–21775.
58. Genco, C.A., and Desai, P.J., 1996, Iron acquisition in the pathogenic *Neisseria*. *Trends Microbiol.*, 4:179–184.
59. Gensberg, K., Hughes, K., and Smith, A.W., 1992, Siderophore-specific induction of iron uptake in *Pseudomonas aeruginosa*. *J. Gen. Microbiol.*, 138:2381–2387.
60. Ghysels, B., Dieu, B.T., Beatson, S.A., Pirnay, J.P., Ochsner, U.A., Vasil, M.L., and Cornelis, P., 2004, FpvB, an alternative type I ferripyoverdine receptor of *Pseudomonas aeruginosa*. *Microbiology*, 150:1671–1680.
61. Guina, T., Purvine, S.O., Yi, E.C., Eng, J., Goodlett, D.R., Aebersold, R., and Miller, S.I., 2003, Quantitative proteomic analysis indicates increased synthesis of a quinolone by *Pseudomonas aeruginosa* isolates from cystic fibrosis airways. *Proc. Natl. Acad. Sci. U.S.A.*, 100:2771–2776.
62. Haas, B., Kraut, J., Marks, J., Zanker, S.C., and Castignetti, D., 1991, Siderophore presence in sputa of cystic fibrosis patients. *Infect. Immun.*, 59:3997–4000.
63. Handfield, M., Lehoux, D.E., Sanschagrin, F., Mahan, M.J., Woods, D.E., and Levesque, R.C., 2000, In vivo-induced genes in *Pseudomonas aeruginosa*. *Infect. Immun.*, 68:2359–2362.
64. Hantke, K., 2001, Iron and metal regulation in bacteria. *Curr. Opin. Microbiol.*, 4:172–177.
65. Heinrichs, D.E., and Poole, K., 1996, PchR, a regulator of ferripyochelin receptor gene (*fptA*) expression in *Pseudomonas aeruginosa*, functions both as an activator and as a repressor. *J. Bacteriol.*, 178:2586–2592.
66. Heller, K.J., Kadner, R.J., and Gunther, K., 1988, Suppression of the *btuB451* mutation by mutations in the *tonB* gene suggests a direct interaction between TonB and TonB-dependent receptor proteins in the outer membrane of *Escherichia coli*. *Gene*, 64:147–153.
67. Higgs, P.I., Larsen, R.A., and Postle, K., 2002, Quantification of known components of the *Escherichia coli* TonB energy transduction system: TonB, ExbB, ExbD and FepA. *Mol. Microbiol.*, 44:271–281.
68. Huang, B., Ru, K., Yuan, Z., Whitchurch, C.B., and Mattick, J.S., 2004, *TonB3* is required for normal twitching motility and extracellular assembly of type IV pili. *J. Bacteriol.*, 186:4387–4389.
69. Hussein, S., Hantke, K., and Braun, V., 1981, Citrate-dependent iron transport system in *Escherichia coli* K-12. *Eur. J. Biochem.*, 117:431–437.
70. Idei, A., Kawai, E., Akatsuka, H., and Omori, K., 1999, Cloning and characterization of the *Pseudomonas fluorescens* ATP-binding cassette exporter, HasDEF, for the heme acquisition protein HasA. *J. Bacteriol.*, 181:7545–7551.
71. Izadi, N., Henry, Y., Haladjian, J., Goldberg, M.E., Wandersman, C., Delepierre, M., and Lecroisey, A., 1997, Purification and characterization of an extracellular heme-binding protein, HasA, involved in heme iron acquisition. *Biochemistry*, 36:7050–7057.
72. Jiang, X., Payne, M.A., Cao, Z., Foster, S.B., Feix, J.B., Newton, S.M., and Klebba, P.E., 1997, Ligand-specific opening of a gated-porin channel in the outer membrane of living bacteria. *Science*, 276:1261–1264.
73. Kadner, R.J., 1990, Vitamin B12 transport in *Escherichia coli*: energy coupling between membranes. *Mol. Microbiol.*, 4:2027–2033.

74. Kadner, R.J., and Heller, K.J., 1995, Mutual inhibition of cobalamin and siderophore uptake systems suggests their competition for TonB function. *J. Bacteriol*., 177:4829–4835.
75. Kammler, M., Schon, C., and Hantke, K., 1993, Characterization of the ferrous iron uptake system of *Escherichia coli*. *J. Bacteriol.*, 175:6212–6219.
76. Kampfenkel, K., and Braun, V., 1993, Topology of the ExbB protein in the cytoplasmic membrane of *Escherichia coli*. *J. Biol. Chem.*, 268:6050–6057.
77. Kampfenkel, K., Braun, V., Stojiljkovic, I., and Srinivasan, N., 1992, Membrane topology of the *Escherichia coli* ExbD protein. *J. Bacteriol.*, 174:5485–5487.
78. Khursigara, C.M., De Crescenzo, G., Pawelek, P.D., and Coulton, J.W., 2004, Enhanced binding of TonB to a ligand-loaded outer membrane receptor: role of the oligomeric state of TonB in formation of a functional FhuA–TonB complex. *J. Biol. Chem.*, 279:7405–7412.
79. Kim, I., Stiefel, A., Plantor, S., Angerer, A., and Braun, V., 1997, Transcription induction of the ferric citrate transport genes *via* the N-terminus of the FecA outer membrane protein, the Ton system and the electrochemical potential of the cytoplasmic membrane. *Mol. Microbiol.*, 23:333–344.
80. Koedding, J., Howard, P., Kaufmann, L., Polzer, P., Lustig, A., and Welte, W., 2004, Dimerization of TonB is not essential for its binding to the outer membrane siderophore receptor FhuA of *Escherichia coli*. *J. Biol. Chem.*, 279:9978–9986.
81. Koster, M., van Klompenburg, W., Bitter, W., Leong, J., and Weisbeek, P., 1994, Role for the outer membrane ferric siderophore receptor PupB in signal transduction across the bacterial cell envelope. *EMBO J.*, 13:2805–2813.
82. Koster, W., 2001, ABC transporter-mediated uptake of iron, siderophores, heme and vitamin B12. *Res. Microbiol.*, 152:291–301.
83. Lamont, I.L., Beare, P.A., Ochsner, U., Vasil, A.I., and Vasil, M.L., 2002, Siderophore-mediated signaling regulates virulence factor production in *Pseudomonas aeruginosa*. *Proc. Natl. Acad. Sci. U.S.A.*, 99:7072–7077.
84. Lamont, I.L., and Martin, L.W., 2003, Identification and characterization of novel pyoverdine synthesis genes in *Pseudomonas aeruginosa*. *Microbiology*, 149:833–842.
85. Larsen, R.A., Foster-Hartnett, D., McIntosh, M.A., and Postle, K., 1997, Regions of *Escherichia coli* TonB and FepA proteins essential for *in vivo* physical interactions. *J. Bacteriol.*, 179:3213–3221.
86. Lehoux, D.E., Sanschagrin, F., and Levesque, R.C., 2000, Genomics of the 35-kb *pvd* locus and analysis of novel *pvdIJK* genes implicated in pyoverdine biosynthesis in *Pseudomonas aeruginosa*. *FEMS Microbiol. Lett*., 190:141–146.
87. Letoffe, S., Delepelaire, P., and Wandersman, C., 2004, Free and hemophore-bound heme acquisitions through the outer membrane receptor HasR have different requirements for the TonB–ExbB–ExbD complex. *J. Bacteriol.*, 186:4067–4074.
88. Letoffe, S., Deniau, C., Wolff, N., Dassa, E., Delepelaire, P., Lecroisey, A., and Wandersman, C., 2001, Haemophore-mediated bacterial haem transport: evidence for a common or overlapping site for haem-free and haem-loaded haemophore on its specific outer membrane receptor. *Mol. Microbiol.*, 41:439–450.
89. Letoffe, S., Ghigo, J.M., and Wandersman, C., 1994, Secretion of the *Serratia marcescens* HasA protein by an ABC transporter. *J. Bacteriol.*, 176:5372–5377.
90. Letoffe, S., Redeker, V., and Wandersman, C., 1998, Isolation and characterization of an extracellular haem-binding protein from *Pseudomonas aeruginosa* that shares function and sequence similarities with the *Serratia marcescens* HasA haemophore. *Mol. Microbiol.*, 28:1223–1234.
91. Liu, J., Rutz, J.M., Klebba, P.E., and Feix, J.B., 1994, A site-directed spin-labeling study of ligand-induced conformational change in the ferric enterobactin receptor, FepA. *Biochemistry*, 33:13274–13283.

92. Locher, K.P., Rees, B., Koebnik, R., Mitschler, A., Moulinier, L., Rosenbusch, J.P., and Moras, D., 1998, Transmembrane signaling across the ligand-gated FhuA receptor: crystal structures of free and ferrichrome-bound states reveal allosteric changes. *Cell*, 95:771–778.
93. Lowe, C.A., Asghar, A.H., Shalom, G., Shaw, J.G., and Thomas, M.S., 2001, The *Burkholderia cepacia fur* gene: co-localization with *omlA* and absence of regulation by iron. *Microbiology*, 147:1303–1314.
94. McMorran, B.J., Kumara, H.M., Sullivan, K., and Lamont, I.L., 2001, Involvement of a transformylase enzyme in siderophore synthesis in *Pseudomonas aeruginosa*. *Microbiology*, 147:1517–1524.
95. McMorran, B.J., Merriman, M.E., Rombel, I.T., and Lamont, I.L., 1996, Characterisation of the *pvdE* gene which is required for pyoverdine synthesis in *Pseudomonas aeruginosa*. *Gene*, 176:55–59.
96. Merianos, H.J., Cadieux, N., Lin, C.H., Kadner, R.J., and Cafiso, D.S., 2000, Substrate-induced exposure of an energy-coupling motif of a membrane transporter. *Nat. Struct. Biol.*, 7:205–209.
97. Merriman, T.R., Merriman, M.E., and Lamont, I.L., 1995, Nucleotide sequence of *pvdD*, a pyoverdine biosynthetic gene from *Pseudomonas aeruginosa*: PvdD has similarity to peptide synthetases. *J. Bacteriol.*, 177:252–258.
98. Meyer, J.M., 1992, Exogenous siderophore-mediated iron uptake in *Pseudomonas aeruginosa*: possible involvement of porin OprF in iron translocation. *J. Gen. Microbiol.*, 138:951–958.
99. Meyer, J.M., Azelvandre, P., and Georges, C., 1992, Iron metabolism in *Pseudomonas*: salicylic acid, a siderophore of *Pseudomonas fluorescens* CHAO. *Biofactors*, 4:23–27.
100. Meyer, J.M., Geoffroy, V.A., Baida, N., Gardan, L., Izard, D., Lemanceau, P., Achouak, W., and Palleroni, N.J., 2002, Siderophore typing, a powerful tool for the identification of fluorescent and nonfluorescent pseudomonads. *Appl. Environ. Microbiol.*, 68:2745–2753.
101. Meyer, J.M., Hohnadel, D., and Halle, F., 1989, Cepabactin from *Pseudomonas cepacia*, a new type of siderophore. *J. Gen. Microbiol.*, 135:1479–1487.
102. Meyer, J.M., Neely, A., Stintzi, A., Georges, C., and Holder, I.A., 1996, Pyoverdin is essential for virulence of *Pseudomonas aeruginosa*. *Infect. Immun.*, 64:518–523.
103. Meyer, J.M., Stintzi, A., De Vos, D., Cornelis, P., Tappe, R., Taraz, K., and Budzikiewicz, H., 1997, Use of siderophores to type pseudomonads: the three *Pseudomonas aeruginosa* pyoverdine systems. *Microbiology*, 143:35–43.
104. Meyer, J.M., Van, V.T., Stintzi, A., Berge, O., and Winkelmann, G., 1995, Ornibactin production and transport properties in strains of *Burkholderia vietnamiensis* and *Burkholderia cepacia* (formerly *Pseudomonas cepacia*). *Biometals*, 8:309–317.
105. Mielczarek, E.V., Royt, P.W., and Toth-Allen, J., 1990, A Mössbauer spectroscopy study of cellular acquisition of iron from pyoverdine by *Pseudomonas aeruginosa*. *Biol. Met*. 3:34–38.
106. Miyazaki, H., Kato, H., Nakazawa, T., and Tsuda, M., 1995, A positive regulatory gene, *pvdS*, for expression of pyoverdin biosynthetic genes in *Pseudomonas aeruginosa* PAO. *Mol. Gen. Genet.*, 248:17–24.
107. Moeck, G.S., and Coulton, J.W., 1998, TonB-dependent iron acquisition: mechanisms of siderophore-mediated active transport. *Mol. Microbiol.*, 28:675–681.
108. Moeck, G.S., Coulton, J.W., and Postle, K., 1997, Cell envelope signaling in *Escherichia coli*. Ligand binding to the ferrichrome-iron receptor *fhuA* promotes interaction with the energy-transducing protein TonB. *J. Biol. Chem.*, 272:28391–28397.
109. Mohn, G., Koehl, P., Budzikiewicz, H., Lefevre, J.F., Atkinson, R.A., Salah El Din, A.L., Kieffer, B., and Abdallah, M.A., 1994, Solution structure of pyoverdin GM-II. *Biochemistry*, 33:2843–2851.

110. Morris, J., Donnelly, D.F., O'Neill, E., McConnell, F., and O'Gara, F., 1994, Nucleotide sequence analysis and potential environmental distribution of a ferric pseudobactin receptor gene of *Pseudomonas* sp. strain M114. *Mol. Gen. Genet.*, 242:9–16.
111. Mossialos, D., Meyer, J.M., Budzikiewicz, H., Wolff, U., Koedam, N., Baysse, C., Anjaiah, V., and Cornelis, P., 2000, Quinolobactin, a new siderophore of *Pseudomonas fluorescens* ATCC 17400, the production of which is repressed by the cognate pyoverdine. *Appl. Environ. Microbiol.*, 66:487–492.
112. Mossialos, D., Ochsner, U., Baysse, C., Chablain, P., Pirnay, J.P., Koedam, N., Budzikiewicz, H., Fernandez, D.U., Schafer, M., Ravel, J., and Cornelis, P., 2002, Identification of new, conserved, non-ribosomal peptide synthetases from fluorescent pseudomonads involved in the biosynthesis of the siderophore pyoverdine. *Mol. Microbiol.*, 45:1673–1685.
113. Ochsner, U.A., Johnson, Z., and Vasil, M.L., 2000, Genetics and regulation of two distinct haem-uptake systems, *phu* and *has*, in *Pseudomonas aeruginosa*. *Microbiology*, 146:185–198.
114. Ochsner, U.A., Wilderman, P.J., Vasil, A.I., and Vasil, M.L., 2002, GeneChip expression analysis of the iron starvation response in *Pseudomonas aeruginosa*: identification of novel pyoverdine biosynthesis genes. *Mol. Microbiol.*, 45:1277–1287.
115. Ogierman, M., and Braun, V., 2003, Interactions between the outer membrane ferric citrate transporter FecA and TonB: studies of the FecA TonB box. *J. Bacteriol*., 185:1870–1885.
116. Palma, M., DeLuca, D., Worgall, S., and Quadri, L.E., 2004, Transcriptome analysis of the response of *Pseudomonas aeruginosa* to hydrogen peroxide. *J. Bacteriol*., 186:248–252.
117. Paquelin, A., Ghigo, J.M., Bertin, S., and Wandersman, C., 2001, Characterization of HasB, a *Serratia marcescens* TonB-like protein specifically involved in the haemophore-dependent haem acquisition system. *Mol. Microbiol.*, 42:995–1005.
118. Payne, M.A., Igo, J.D., Cao, Z., Foster, S.B., Newton, S.M., and Klebba, P.E., 1997, Biphasic binding kinetics between FepA and its ligands. *J. Biol. Chem.*, 272:21950–21955.
119. Pohl, E., Haller, J.C., Mijovilovich, A., Meyer-Klaucke, W., Garman, E., and Vasil, M.L., 2003, Architecture of a protein central to iron homeostasis: crystal structure and spectroscopic analysis of the ferric uptake regulator. *Mol. Microbiol.*, 47:903–915.
120. Poole, K., and McKay, G.A., 2003, Iron acquisition and its control in *Pseudomonas aeruginosa*: many roads lead to Rome. *Front. Biosci.*, 8:661–686.
121. Poole, K., Neshat, S., Krebes, K., and Heinrichs, D.E., 1993, Cloning and nucleotide sequence analysis of the ferripyoverdine receptor gene *fpvA* of *Pseudomonas aeruginosa*. *J. Bacteriol.*, 175:4597–4604.
122. Poole, K., Young, L., and Neshat, S., 1990, Enterobactin-mediated iron transport in *Pseudomonas aeruginosa*. *J. Bacteriol.*, 172:6991–6996.
123. Poole, K., Zhao, Q., Neshat, S., Heinrichs, D.E., and Dean, C.R., 1996, The *Pseudomonas aeruginosa tonB* gene encodes a novel TonB protein. *Microbiology*, 142:1449–1458.
124. Postle, K., 1990, Aerobic regulation of the *Escherichia coli tonB* gene by changes in iron availability and the *fur* locus. *J. Bacteriol.*, 172:2287–2293.
125. Postle, K., 1990, TonB and the gram-negative dilemma. *Mol. Microbiol.*, 4:2019–2025.
126. Postle, K., and Kadner, R.J., 2003, Touch and go: tying TonB to transport. *Mol. Microbiol.*, 49:869–882.
127. Postle, K., Reznikoff, W.S., Stojiljkovic, I., and Srinivasan, N., 1979, Identification of the *Escherichia coli tonB* gene product in minicells containing *tonB* hybrid plasmids. *J. Mol. Biol.*, 131:619–636.
128. Ratledge, C., and Dover, L.G., 2000, Iron metabolism in pathogenic bacteria. *Annu. Rev. Microbiol.*, 54:881–941.
129. Reimmann, C., Patel, H.M., Serino, L., Barone, M., Walsh, C.T., and Haas, D., 2001, Essential PchG-dependent reduction in pyochelin biosynthesis of *Pseudomonas aeruginosa*. *J. Bacteriol.*, 183:813–820.

130. Reimmann, C., Serino, L., Beyeler, M., and Haas, D., 1998, Dihydroaeruginoic acid synthetase and pyochelin synthetase, products of the *pchEF* genes, are induced by extracellular pyochelin in *Pseudomonas aeruginosa*. *Microbiology*, 144:3135–3148.
131. Rossi, M.S., Paquelin, A., Ghigo, J.M., and Wandersman, C., 2003, Haemophore-mediated signal transduction across the bacterial cell envelope in *Serratia marcescens*: the inducer and the transported substrate are different molecules. *Mol. Microbiol.*, 48:1467–1480.
132. Royt, P.W., 1990, Pyoverdine-mediated iron transport. Fate of iron and ligand in *Pseudomonas aeruginosa*. *Biol. Met*. 3:28–33.
133. Schalk, I.J., Abdallah, M.A., and Pattus, F., 2002, Recycling of pyoverdin on the FpvA receptor after ferric pyoverdin uptake and dissociation in *Pseudomonas aeruginosa*. *Biochemistry*, 41:1663–1671.
134. Schalk, I.J., Abdallah, M.A., Pattus, F., Held, K.G., and Postle, K., 2002, A new mechanism for membrane iron transport in *Pseudomonas aeruginosa*. *Biochem. Soc. Trans.*, 30:702–705.
135. Schalk, I.J., Hennard, C., Dugave, C., Poole, K., Abdallah, M.A., and Pattus, F., 2001, Iron-free pyoverdin binds to its outer membrane receptor FpvA in *Pseudomonas aeruginosa*: a new mechanism for membrane iron transport. *Mol. Microbiol.*, 39:351–360.
136. Schalk, I.J., Kyslik, P., Prome, D., van Dorsselaer, A., Poole, K., Abdallah, M.A., and Pattus, F., 1999, Copurification of the FpvA ferric pyoverdin receptor of *Pseudomonas aeruginosa* with its iron-free ligand: implications for siderophore-mediated iron transport. *Biochemistry*, 38:9357–9365.
137. Schalk, I.J., Yue, W.W., and Buchanan, S.K., 2004, Recognition of iron-free siderophores by TonB-dependent iron transports. *Mol. Microbiol.*, i54:14–22.
138. Schoffler, H., and Braun, V., 1989, Transport across the outer membrane of *Escherichia coli* K12 *via* the FhuA receptor is regulated by the TonB protein of the cytoplasmic membrane. *Mol. Gen. Genet.*, 217:378–383.
139. Schryvers, A.B., Stojiljkovic, I., Vasil, M.L., and Ochsner, U.A., 1999, Iron acquisition systems in the pathogenic *Neisseria*. *Mol. Microbiol.*, 32:1117–1123.
140. Scott, D.C., Cao, Z., Qi, Z., Bauler, M., Igo, J.D., Newton, S.M., and Klebba, P.E., 2001, Exchangeability of N termini in the ligand-gated porins of *Escherichia coli*. *J. Biol. Chem.*, 276:13025–13033.
141. Scott, D.C., Newton, S.M., and Klebba, P.E., 2002, Surface loop motion in FepA. *J. Bacteriol.*, 184:4906–4911.
142. Skare, J.T., Ahmer, B.M., Seachord, C.L., Darveau, R.P., and Postle, K., 1993, Energy transduction between membranes. TonB, a cytoplasmic membrane protein, can be chemically cross-linked *in vivo* to the outer membrane receptor FepA. *J. Biol. Chem.*, 268:16302–16308.
143. Sokol, P.A., Darling, P., Lewenza, S., Corbett, C.R., and Kooi, C.D., 2000, Identification of a siderophore receptor required for ferric ornibactin uptake in *Burkholderia cepacia*. *Infect. Immun.*, 68:6554–6560.
144. Sokol, P.A., Lewis, C.J., and Dennis, J.J., 1992, Isolation of a novel siderophore from *Pseudomonas cepacia*. *J. Med. Microbiol.*, 36:184–189.
145. Stephan, H., Freund, S., Beck, W., Jung, G., Meyer, J.M., and Winkelmann, G., 1993, Ornibactins a new family of siderophores from *Pseudomonas*. *Biometals*, 6:93–100.
146. Stintzi, A., Barnes, C., Xu, J., and Raymond, K.N., 2000, Microbial iron transport *via* a siderophore shuttle: a membrane ion transport paradigm. *Proc. Natl. Acad. Sci. U.S.A.*, 97:10691–10696.
147. Stintzi, A., Cornelis, P., Hohnadel, D., Meyer, J.M., Dean, C., Poole, K., Kourambas, S., and Krishnapillai, V., 1996, Novel pyoverdine biosynthesis gene(s) of *Pseudomonas aeruginosa* PAO. *Microbiology*, 142:1181–1190.
148. Stintzi, A., Johnson, Z., Stonehouse, M., Ochsner, U., Meyer, J.M., Vasil, M.L., and Poole, K., 1999, The *pvc* gene cluster of *Pseudomonas aeruginosa*: role in synthesis of the pyoverdine chromophore and regulation by PtxR and PvdS. *J. Bacteriol.*, 181:4118–4124.

149. Stojiljkovic, I., and Srinivasan, N., 1997, *Neisseria meningitidis* tonB, exbB, and exbD genes: Ton-dependent utilization of protein-bound iron in *Neisseriae*. *J. Bacteriol.*, 179:805–812.
150. Stover, C.K., Pham, X.Q., Erwin, A.L., Mizoguchi, S.D., Warrener, P., Hickey, M.J., Brinkman, F.S., Hufnagle, W.O., Kowalik, D.J., Lagrou, M., Garber, R.L., Goltry, L., Tolentino, E., Westbrock-Wadman, S., Yuan, Y., Brody, L.L., Coulter, S.N., Folger, K.R., Kas, A., Larbig, K., Lim, R., Smith, K., Spencer, D., Wong, G.K., Wu, Z., Paulsen, I.T., Reizer, J., Saier, M.H., Hancock, R.E., Lory, S., and Olson, M.V., 2000, Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen. *Nature*, 406:959–964.
151. Takase, H., Nitanai, H., Hoshino, K., and Otani, T., 2000, Impact of siderophore production on *Pseudomonas aeruginosa* infections in immunosuppressed mice. *Infect. Immun.*, 68:1834–1839.
152. Teintze, M., Hossain, M.B., Barnes, C.L., Leong, J., van der Helm, D., Yun, C.W., Bauler, M., Moore, R.E., Klebba, P.E., and Philpott, C.C., 1981, Structure of ferric pseudobactin, a siderophore from a plant growth promoting *Pseudomonas*. *Biochemistry*, 20:6446–6457.
153. Teintze, M., and Leong, J., 1981, Structure of pseudobactin A, a second siderophore from plant growth promoting *Pseudomonas* B10. *Biochemistry*, 20:6457–6462.
154. Thupvong, T., Wiideman, A., Dunn, D., Oreschak, K., Jankowicz, B., Doering, J., and Castignetti, D., 1999, Sequence heterogeneity of the ferripyoverdine uptake (*fpvA*), but not the ferric uptake regulator (*fur*), genes among strains of the fluorescent pseudomonads *Pseudomonas aeruginosa*, *Pseudomonas aureofaciens*, *Pseudomonas fluorescens* and *Pseudomonas putida*. *Biometals*, 12:265–274.
155. Usher, K.C., Ozkan, E., Gardner, K.H., and Deisenhofer, J., 2001, The plug domain of FepA, a TonB-dependent transport protein from *Escherichia coli*, binds its siderophore in the absence of the transmembrane barrel domain. *Proc. Natl. Acad. Sci. U.S.A.*, 98:10676–10681.
156. Vandenende, C.S., Vlasschaert, M., and Seah, S.Y., 2004, Functional characterization of an aminotransferase required for pyoverdine siderophore biosynthesis in *Pseudomonas aeruginosa* PAO1. *J. Bacteriol.*, 186:5596–5602.
157. Vasil, M.L., and Ochsner, U.A., 1999, The response of *Pseudomonas aeruginosa* to iron: genetics, biochemistry and virulence. *Mol. Microbiol.*, 34:399–413.
158. Venturi, V., Ottevanger, C., Bracke, M., and Weisbeek, P., 1995, Iron regulation of siderophore biosynthesis and transport in *Pseudomonas putida* WCS358: involvement of a transcriptional activator and of the Fur protein. *Mol. Microbiol.*, 15:1081–1093.
159. Visca, P., 2004, Iron regulation and siderophore signalling in virulence by *Pseudomonas aeruginosa*, pp. 69–123. In J.-L. Ramos (ed.), *Pseudomonas*, Vol. 2. Kluwer Academic/Plenum Publishers, New York.
160. Visca, P., Ciervo, A., and Orsi, N., 1994, Cloning and nucleotide sequence of the *pvdA* gene encoding the pyoverdin biosynthetic enzyme L-ornithine N5-oxygenase in *Pseudomonas aeruginosa*. *J. Bacteriol.*, 176:1128–1140.
161. Visca, P., Ciervo, A., Sanfilippo, V., and Orsi, N., 1993, Iron-regulated salicylate synthesis by *Pseudomonas* spp. *J. Gen. Microbiol.*, 139:1995–2001.
162. Visca, P., Colotti, G., Serino, L., Verzili, D., Orsi, N., and Chiancone, E., 1992, Metal regulation of siderophore synthesis in *Pseudomonas aeruginosa* and functional effects of siderophore-metal complexes. *Appl. Environ. Microbiol.*, 58:2886–2893.
163. Visca, P., Leoni, L., Wilson, M.J., and Lamont, I.L., 2002, Iron transport and regulation, cell signalling and genomics: lessons from *Escherichia coli* and *Pseudomonas*. *Mol. Microbiol.*, 45:1177–1190.
164. Wandersman, C., and Stojiljkovic, I., 2000, Bacterial heme sources: the role of heme, hemoprotein receptors and hemophores. *Curr. Opin. Microbiol.*, 3:215–220.
165. Winkelmann, G., 2002, Microbial siderophore-mediated transport. *Biochem. Soc. Trans.*, 30:691–696.

166. Wolz, C., Hohloch, K., Ocaktan, A., Poole, K., Evans, R.W., Rochel, N., Albrecht-Gary, A.M., Abdallah, M.A., and Doring, G., 1994, Iron release from transferrin by pyoverdin and elastase from *Pseudomonas aeruginosa*. *Infect. Immun.*, 62:4021–4027.
167. Xiao, R., and Kisaalita, W.S., 1997, Iron acquisition from transferrin and lactoferrin by *Pseudomonas aeruginosa* pyoverdin. *Microbiology*, 143:2509–2515.
168. Yue, W.W., Grizot, S., and Buchanan, S.K., 2003, Structural evidence for iron-free citrate and ferric citrate binding to the TonB-dependent outer membrane transporter FecA. *J. Mol. Biol.*, 332:353–368.
169. Zhao, Q., and Poole, K., 2000, A second *tonB* gene in *Pseudomonas aeruginosa* is linked to the *exbB* and *exbD* genes. *FEMS Microbiol. Lett*., 184:127–132.

2

CLONAL VARIATIONS IN *PSEUDOMONAS AERUGINOSA*

Burkhard Tümmler

Klinische Forschergruppe, OE 6711
Medizinische Hochschule Hannover
D-30625 Hannover, Germany

1. INTRODUCTION

The genetic diversity within a bacterial species is determined by the number and size of chromosomal and extrachromosomal elements, rates of nucleotide substitution, recombination, genome rearrangements and gene flow, and both the size and growth of the bacterial population. Most species of bacteria that were initially analyzed, were of a clonal nature.[77] The structural characteristics of a clonal population are the paucity of genotypes, linkage disequilibrium among gene loci, and recovery of closely related genotypes over large geographic areas and/or over long periods of time. The accumulation of molecular data during the last 15 years and the growing evidence of the occurrence of horizontal gene transfer among bacteria in nature, however, have led to consideration that bacterial populations are not invariably clonal but range from the highly sexual *Neisseria gonorrhoeae* to the almost strictly clonal *Salmonella*.[80]

The metabolically versatile *Pseudomonas aeruginosa* is present in soil and aquatic habitats, but it is also an important opportunistic pathogen for humans, animals, and plants. Typing of strain collections in single nucleotide polymorphisms (SNPs), DNA fragment length polymorphisms and phenotypic traits indicated that the current *P. aeruginosa* population is in linkage equilibrium and consists of a net of equivalent genotypes (termed clones), whereby a subset of clones is overrepresented due to epidemic spread.[36,60] Isolates from the

Pseudomonas, Volume 4, edited by Juan-Luis Ramos and Roger C. Levesque

inanimate environment and clinical habitats have been shown to share the same chemotaxonomic profile[23] and repertoire of metabolic and virulence traits.[1] Irrespective of their origin, isolates from disease and environment were similarly proficient in the degradation of environmental pollutants and secretion of virulence factors.[1] In other words, there are no disease- or habitat-associated clones. However, we do observe adaptation of *P. aeruginosa* to a particular niche. Most data exists of how *P. aeruginosa* colonizes and persists in the atypical habitat of the cystic fibrosis (CF) lung where independent of the genetic background of the clone a convergent evolution towards common phenotypes takes place.[89]

This chapter summarizes our current knowledge about the inter- and intraclonal diversity of genotype and phenotype of *P. aeruginosa.*

2. INTRA- AND INTERCLONAL GENOME DIVERSITY

Physical mapping and sequencing and Southern hybridization data indicate that the *P. aeruginosa* genome is made up of a mosaic of a conserved core and variable accessory segments.[20,31,36,66,84] The core genome is characterized by a conserved synteny of genes and a low average nucleotide substitution rate. Clone- or strain-specific genome islands and genome islets define the accessory part of the chromosome and lead to fluctuations in the genome size, which can range from 5.2 to 7 Mbp.[73]

2.1. Clonal Variation of the Core Genome

The complete genome sequence of strain PAO1[85] is the genetic blueprint for *P. aeruginosa*. Genomic DNA hybridization of in total 39 *P. aeruginosa* strains of diverse origin onto PAO1 microarrays[20,95] detected the presence of almost 90% of the 5570 predicted PAO1 protein coding sequences in all strains. Hence, the core genome is made up of about 5000 highly conserved genes.

Interclonal sequence variation is low in the *P. aeruginosa* core genome. Comparative sequencing of housekeeping genes in strain collections revealed an average rate of sequence polymorphism of 0.3%, which is about one order of magnitude lower than in comparable housekeeping genes of *Salmonella enterica.*[36] The ratio of non-synonymous to synonymous nucleotide substitutions is about 1:6. Sequence variation within clones is substantially lower than the already low sequence diversity amongst unrelated clones: Within 300 kb of bulk sequence, just a single synonymous nucleotide substitution was detected in one of four analyzed strains.[36,43] In other words, members of a clone are characterized by virtually identical core genome sequence in all segments with low sequence diversity.

Figure 1 shows the comparison of 49 single nucleotide substitutions (SNPs) of *P. aeruginosa* detected in *oriC*, *ampC, citS, fliC, oprI* with 500 SNPs of *S. enterica* detected in *gapA*, *putP*, and *mdh.*[37] In contrast to the high GC

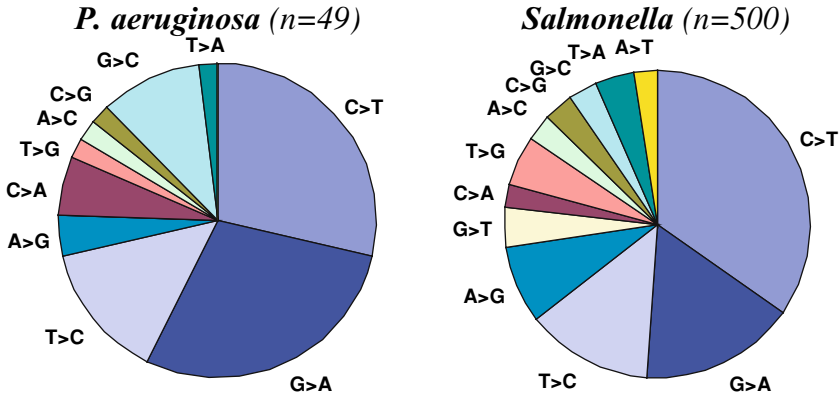


Figure 1. Pie charts showing all single base substitutions detected in *oriC*, *citS*, *ampC*, *oprI,* and *fliC* sequences of 18 *P. aeruginosa* strains and in *gapA*, *putP*, and *mdh* of 16 *Salmonella* strains.

content of the bulk *P. aeruginosa* chromosome (67%), the phylogenetically closely related but ecologically distinct *S. enterica* exhibits a much lower GC content (50–53% GC) and a less pronounced codon usage bias.[93] The quantitative distribution of nucleotide substitution types is similar in both bacterial species except for the more frequent G→C transversion in the GC rich *P. aeruginosa*. In spite of their dissimilar GC content and codon usage bias, about 75% of nucleotide substitutions are transitions. C→T is the most abundant substitution, followed by G→A, T→C, and A→G (Figure 1). Both nucleotide substitution profile and transition-to-transversion ratio are non-randomly distributed. The observed bias probably reflects first, the thermodynamic stability and geometric selection of the mismatches; and second, evolutionarily conserved error-correction mechanisms, i.e. the preferential repair of lesions on the transcribed strand.[37] This argument that nucleotide sequence variation is not governed by chromosomal GC content and codon usage but by DNA structure and repair is substantiated by the finding that the same profiles seen in bacteria were also observed for sequence variants and disease-causing mutations in phylogenetically distant mammalian genomes.[37]

Twenty-five regions of significantly elevated sequence variation were uncovered in pairwise comparisons between PAO1 and three other partially sequenced strains (Table 1).[84] There were no obvious sequence features unique to these regions that flag them relative to the rest of the genomic sequence, albeit a lower average GC contents and an underutilization of the most frequently used codons were noted.

The genome segments with the highest level of sequence diversity are the genes whose products are involved in flagellar biosynthesis and genes whose products are involved in the biosynthesis of the siderophore pyoverdine and the receptor for ferripyoverdines.[84] Pyoverdine is the primary siderophore of

Table 1. PAO1 sequence coordinates for regions of high sequence diversity[a] (modified from Table 3 in ref [84] by permission of the authors and the American Society for Microbiology).

Coordinates[b] of PAO1		Length PAO1	No. of	No. of bases		PAO1
Start	End	(bp)[c]	SNPs[d]	available[e]	%SNPs[f]	ORFs[g]
290,000	294,999	5000	83	1488	5.58	PA0259–PA0262
320,000	324,999	5000	97	1754	5.53	PA0285–PA0289
515,000	519,999	5000	143	2103	6.80	PA0457–PA0459
530,000	534,999	5000	111	2288	4.85	PA0470–PA0473
555,000	559,999	5000	66	1252	5.27	PA0495–PA0500
652,500	657,499	5000	76	1171	6.49	PA0594–PA0596
685,000	689,999	5000	44	789	5.58	PA0625–PA0633
695,000	699,999	5000	84	1436	5.85	PA0640–PA0643
790,000	799,999	10,000	274	2535	10.81	PA0719–PA0731
1,060,000	1,064,999	5000	57	926	6.16	PA0976–PA0982
1,170,000	1,174,999	5000	224	1881	11.91	PA1084–PA1087[h]
2,150,000	2,154,999	5000	88	842	10.45	PA1967–PA1972
2,550,000	2,554,999	5000	59	1346	4.38	PA2312–PA2317
2,635,000	2,654,999	20,000	451	5787	7.79	PA2383–PA2397[i]
2,660,000	2,664,999	5000	115	551	20.87	PA2399[i]
2,672,500	2,677,499	5000	186	1132	16.43	PA2402[i]
2,682,500	2,692,499	10,000	139	1723	8.07	PA2402–PA2409[i]
3,647,500	3,652,499	5000	81	1417	5.72	PA3260–PA3264
4,060,000	4,064,999	5000	85	1695	5.01	PA3624–PA3629
5,037,500	5,042,499	5000	98	1923	5.10	PA4500–PA4503
5,070,000	5,074,999	5000	162	2030	7.98	PA4526–PA4532
5,097,500	5,102,499	5000	119	1825	6.52	PA4549–PA4554[j]
5,187,500	5,192,499	5000	91	1561	5.83	PA4625
5,722,500	5,727,499	5000	162	1966	8.24	PA5084–PA5089
6,090,000	6,094,999	5000	93	618	15.05	PA5412–PA5415

[a] Sequencing data for all three strains that had been subjected to whole genome shot-gun sequencing was compared to the PAO1 reference in 5000-bp sliding windows that were sequentially offset by 2500 bp. Regions with at least 500 bp of alignable sequence that exhibited nucleotide diversity values greater than three standard deviations (>4.35% sequence differences) from the mean value of 0.5% are shown.

[b] Sequence coordinates of the PAO1 reference sequence, accessible at www.pseudomonas.com.

[c] Overall span of the PAO1 region encompassed by high sequence variation.

[d] Number of SNPs detected.

[e] Total number of alignable bases in which SNPs were detected.

[f] Percent SNPs among alignable bases.

[g] Annotated ORFs within high-diversity regions. A more comprehensive description of these genes is available at www.genome.washington.edu/UWGC and at www.pseudomonas.com.

[h] Flagellar biogenesis genes.

[i] Pyoverdine locus.

[j] Minor type IV pili prepilin.

P. aeruginosa. Each strain makes one of three pyoverdine types, each type with a distinct peptide chain that is synthesized non-ribosomally. The pyoverdine region spans an interval of approximately 50 kb in PAO1. The three divergent sequence types correspond to the three structural types of pyoverdines.[79] The outer-membrane pyoverdine receptor, FvpA (PA2398), is also type-specific, transporting only its corresponding pyoverdine. FvpA exhibits the largest variations with about 50% mismatch of amino acid pairwise alignment between genes of each pyoverdine type.[15,42,63] FpvA moreover shows substantial intratype variation and apparently accumulated non-synonymous changes at a elevated rate which has been interpreted as strong evidence of positive selection.[25,79] The next most divergent genes with 15–40% mismatch of amino acid pairwise alignment are immediately adjacent to *fpvA*, and include the ABC transporter *pvdE* (PA2397) and the non-ribosomal peptide synthetase genes *pvdD, pvdJ,* and *pvdI* (PA2399–PA2402). Besides *fvpA* 10 further hotspots of elevated intratype sequence divergence were identified. Since these islands of about 100 bp in length are located within regions that are divergent between pyoverdine types and since intratype differences are very similar to those between pyoverdine types, the sequence divergence probably arose from recombinations.[79]

The flagellin biosynthesis genes encode the elements for the serologically distinct a- and b-type flagellae. The flagellum confers motility and chemotaxis, facilitates adherence to cells and inanimate surfaces and contributes to the colonization and invasion of hosts during infection. Flagellins, a- and b-type, are 74% identical in the nucleotide sequence and 63–65% identical in the amino acid sequence.[7,82,94] They share nearly identical N- and C-terminal sequences, whereas the central region is variable in size and primary structure. This central part is also the major region of intratype sequence variation among a-type *fliC* genes. Based on the amino acid sequences of flagellins from 24 a-type *P. aeruginosa* strains, two subtypes, A1 and A2 were recognized that differ in the central regions by 13 amino acid substitutions and two small deletions of three- and four- amino acids.[7] Although a-type and b-type flagellins differ by 37–38% in their primary structure, the impact of sequence diversity on secondary and tertiary structure is low. A1, A2, and b-type flagellins match perfectly in their profiles for hydrophobicity, flexibility of the peptide backbone, antigenic index, and probability of surface exposure.[83] The constraints for the efficient multimerization of subunits to a functional flagellum are probably so tight that the polymorphic proteins fold into a similar three-dimensional structure. a-type flagellins are glycosylated.[12] a-type strains carry a polymorphic genomic island that is essential for glycosylation of flagellin.[4] An a-type strain either harbors the long version of the island of 14 open reading frames (*orfA* to *orfN*) or an abbreviated version (short island) in which *orfD*, -*E*, and -*H* are polymorphic and *orfI*, -*J*, -*K*, -*L*, and -*M* are absent.[7] The glycosylation island is located upstream of *fliC*. Comparative sequencing between strains PAK and PAO1 as representatives for a- and b-type flagella revealed that the polymorphic region

of the flagellar regulon encompasses the region from *flgK* (PA1086) at the 5′ end up to amino acid 88 of *fleP* (PA1096) at the 3′ end: 5′-*flgKL* – glycosylation island – *fliC* – *fleL* – *fliDSS*′– *fleP*.[4,84] Correspondingly, there are two types of flagellar cap proteins, FliD, which are only 58% identical at the nucleotide level and 43% identical at the amino acid level.[5] These genes are co-inherited with their cognate flagellin gene types, a or b.

The most substantial interclonal sequence variation for a single gene common to all *P. aeruginosa* is observed for the *pilA* gene (PA4525) encoding the type IV pili that play a major role in mediating the adhesion of the bacterial cell to host tissue. All classic pilin subunits share characteristic features, including a six- or seven-amino acid leader peptide, an *N*-methylated phenylalanine as the first residue of the mature protein and a highly conserved N-terminus with 25–30 hydrophobic amino acids, but otherwise the primary sequence is highly variable. The published *pilA* sequences segregate into five groups exhibiting less than 30% nucleotide identity that provide fewer homologies between themselves than with pilins of different species.[36,81] Each group carries a specific sequence insertion downstream of *pilA*. Group I members share about 85%, group II members about 65% nucleotide identity amongst themselves.[13] The type IV pili of *P. aeruginosa* are no more closely related to each other or to other γ-Proteobacteria genera *Escherichia, Aeromonas, Vibrio,* and *Moraxella* than they are to the pili of the β-Proteobacteria *Neisseria* and *Eikenella*, the genus *Dichelobacter*, representative of the deepest branching γ-Proteobacteria, or the phylogenetically distant δ-Proteobacterium *Myxococcus*.[81] *P. aeruginosa* probably acquired its pilin genes from the *Moraxella* lineage, because the *pilA* genes still retain the GC and codon usage characteristics of *Moraxella* pilin genes.[48]

The type III secretion system as one of the major virulence determinants of *P. aeruginosa* transports four known effector proteins: ExoS, ExoT, ExoU, and ExoY. The bifunctional ExoS exerts its cytotoxic activities by a GTPase-activating domain and a ADP-ribosyltransferase activity.[8] ExoT is also an ADP-ribosyltransferase but has only 0.2% of the catalytic activity of ExoS. Like ExoS, it is a GTPase-activating protein for Rho GTPases. ExoU is a potent patatin-like phospholipase that causes rapid cell death following its injection into host cells.[71] ExoY is an adenylate cyclase that elevates the intracellular cAMP levels in eukaryotic cells and causes rounding of certain cell types.[90] The genes encoding the secretion, translocation, and regulatory machinery of the type III secretion system are clustered together in the *P. aeruginosa* chromosome. The genes encoding the type III effector proteins, however, are scattered throughout the chromosome. In an epidemiological study on 115 *P. aeruginosa* isolates[22] the large chromosomal locus and *exoT* (PA0044) were present in all isolates. In contrast, the *exoS* (PA3841), *exoU,* and *exoY* (PA2191) genes were variable traits. Overall, 72% of examined isolates contained the *exoS* gene, 28% contained the *exoU* gene, and 89% contained the *exoY* gene. An inverse correlation was noted between the presence of the *exoS* and *exoU* genes in that all isolates except two, one containing both genes and another containing neither

of them, contained either *exoS* or *exoU* but not both. No significant difference in *exoS, exoU,* or *exoY* prevalence was observed between clinical and environmental isolates or between isolates cultured from different disease sites except for respiratory isolates from patients with CF. CF isolates harbored the *exoU* gene less frequently and the *exoS* gene more frequently than did isolates from some of the other sites of infection, including the respiratory tract of patients without CF. These results suggest that the *P. aeruginosa* type III secretion system is present in nearly all clinical and environmental isolates but that individual isolates differ in their effector genotypes.

P. aeruginosa lipopolysaccharide (LPS) is composed of lipid A, the core oligosaccharide, and the long chain polysaccharides (O-antigen) (detailed information in the article by Lam *et al*., volume 3, Chapter 1 of this monograph series). In short, the majority of *P. aeruginosa* produces two distinct forms of O-antigens called A-band and B-band. Differences in the chemical structure of the B-band LPS are responsible for the serogroup specificity of the respective strains and has been employed for many years for serotyping of *P. aeruginosa* isolates. The major set of enzymes responsible for O-antigen B-band synthesis and assembly are encoded in a single, large gene cluster (PA3160–PA3141 in strain PAO1). Raymond *et al*.[64] sequenced this B-band gene island in all 20 IATS reference serotype strains. Eleven groups of gene clusters were identified that are highly divergent from one another at the DNA sequence level. Within each group a high degree of sequence conservation was observed. The B-band gene islands of serotypes O1, O4, O6, O9, and O12 constitute each a distinct gene cluster. Groups with two members include O3 and O15 (Lory), O7 and O8, O10 and O19, O11 and O17, and O13 and O14. The largest group with 98% sequence identity contains strains of serotypes O2, O5, O16, O18, and O20, consequently the variations in the structures among these serotypes are not conferred by the B-band gene island. This argument also applies to the O10–O19 group, for which no DNA sequence differences were found in 16 kbp of sequence, and to the O7–O8 group, in which only two conservative amino acid changes were identified.

In summary, to date the following genes and gene clusters exhibit the largest interclonal genetic diversity in the core genome: the pyoverdine locus, the flagellar regulon, *pilA*, the type III secretion effector proteins and the O-antigen biosynthesis locus. Each locus is present in all strains, but the genes in each locus are highly divergent between strains. This "replacement island" phenomenon presumably results from diversifying selection, a type of selection that maintains multiple alleles in the population.[79]

Mosaic genes are a further source of genetic diversity. Evidence for a mosaic gene structure is drawn from SNP haplotype[36,83] or the detection of cassettes.[82] Our current knowledge about mosaic genes in *P. aeruginosa* is restricted to *ampC* (PA4110),[83] *fleP* (PA1096),[4] *fliC* (PA1092),[36,83] *mucABCD* (PA0763–PA0766),[11] and *oprD* (PA0958).[61] The *ampC* sequences of 18 strains were compiled into 12 groups by their diagnostic SNP patterns.[83] No

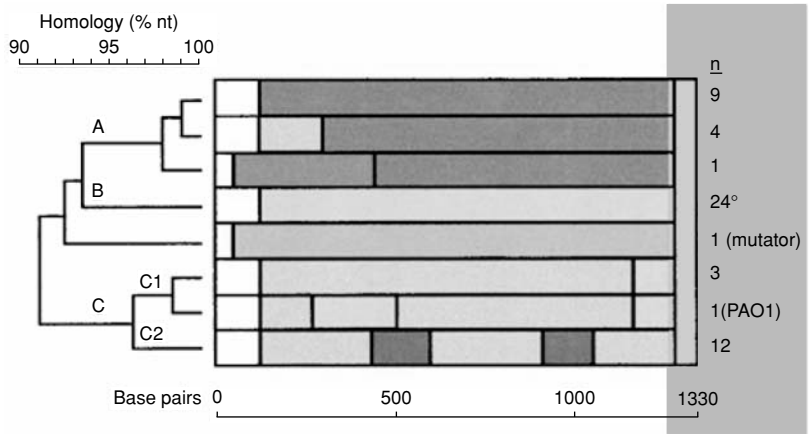


Figure 2. Graphical representation of the mosaic structure of the *oprD* gene in 55 unbiased *P. aeruginosa* strains. Including 10 clone C isolates; *n*, number of isolates belonging to a subgroup. Reproduced from the article by Pirnay *et al.*[61] by permission of the authors and Blackwell Publishing.

linkage of dimorphisms was observed which indicated repetitive intragenic recombination events. However, the low nucleotide substitution rate and the low number of analyzed strains did not allow statistical evidence of a putative mosaic structure of *ampC*. *fleP* is a mosaic gene because the 3′ end of the polymorphic region of the flagellar regulon is located within *fleP*.[4] The first 264 nucleotide of *fleP* were only 56% identical between strains PAK and PAO1 but the last 31 nucleotides of *fleP* were 100% identical. a-Type *fliC* genes contain a variable 141-bp central cassette showing 28% nucleotide and 40% amino acid diversity.[82] Significant non-random clustering of polymorphic sites within this cassette indicated an intragenic recombination event and a mosaic gene structure.[36] Sequencing of 37 CF *P. aeruginosa* isolates in the *mucABD* operon uncovered 16 SNP genotypes. The non-random distribution of conserved SNP blocks visualized the mosaic structure of the *muc* operon.[11] Sequence analysis of *oprD* in 55 *P. aeruginosa* isolates, collected over a period of 15 years from various, spatially separated, clinical and environmental habitats, uncovered a microscale mosaic structure of *oprD*.[61] All sequences fell into three main groups, which differ by 7–9% of nucleotides. Several recombinational exchanges of DNA blocks of 100–300 bp led to a mosaic gene structure and caused a further divergence into subgroups (Figure 2).

Our knowledge about sequence variation resides on the complete genome sequence of two strains, whole genome shot-gun sequencing in another three strains and comparative sequencing of strain collections in 11 loci. Considering this rather limited body of comparative sequence data the proportion of five genes with intragenic mosaicism is substantial. In other words, intragenic recombination may be a major driving force for genetic diversity of *P. aeruginosa.*

2.2. Clonal Variation of the Accessory Genome

Genome diversity is accomplished by sequence variation in coding and non-coding regions of the core genome and by a differential repertoire of the accessory genome the latter being made up of genome islands and genome islets and of mobile genetic elements such as phages, plasmids, and transposons. The reader is referred to volume 1, Chapters 6–8 to get comprehensive information about the features of phages, plasmids, and transposons in *Pseudomonas*. This chapter focuses on the variation of chromosomal contents.

Diversity of the *P. aeruginosa* chromosome was first studied by Southern hybridization analysis.[31,66] *Spe*I macrorestriction fragment length diversity was scanned in 60 unrelated clones for using probes of known map position of the PAO1 chromosome. The *oriC*-containing *Spe*I fragment was the most conserved *Spe*I fragment on the chromosome. Small insertions or deletions lead to a variation of ±10% of chromosomal contents in this region of the origin of replication (Figure 3). Few fragment length classes were seen for most analyzed segments indicating an intermediate range of diversity. In contrast, extensive genomic diversity was detected around the *pilA* and *lipH* loci that later turned out to be hotspots for the integration of genome islands (see below). In other words, the gene contig of the core genome is interrupted by few islets around *oriC* as one extreme and by large segments around *pilA* and *lipA*. Heuer *et al*.[31] studied the same strain collection by probing the chromosome in four regions with 40–114 kb large PAO1 *Spe*I fragments cloned into yeast artificial chromosomes (YACs). In one region the broad distribution of hybridizing *Spe*I fragment size indicated substantial genome plasticity, but otherwise only few bands within narrow fragment length classes reacted with the probe. The low complexity of the hybridization pattern indicates that conserved PAO1 coding and non-coding sequence is maintained as contigs in *P. aeruginosa*. Intrachromosomal shuffling of sequence is rare. In other words, gene order established for strain PAO1 should be valid for most *P. aeruginosa*. YAC hybridizations compare genomes at low resolution so that the disruption of the sequence contig by small genome islets is not resolved. Indirect evidence for the presence of such small insertions and deletions was provided by the strain-to-strain variation of up to 10% in macrorestriction fragment size.

Information about the PAO1 accessory genome in terms of genome islets and genome islands has meanwhile been obtained by hybridization of genomic DNA from strain collections onto PAO1 microarrays. Wolfgang *et al*.[95] analyzed 18 strains of diverse origin. Strain-specific genes were localized to 90 discrete regions relative to the PAO1 genome. Many of these regions are composed of small gene blocks (one to four genes) that showed variability in one or more strains. These variable blocks likely contain genes that are highly polymorphic at the nucleotide sequence level or are gained or lost through local recombination events. A second pattern, which was more readily apparent, is characterized by large clusters of tandem genes that show varying levels of

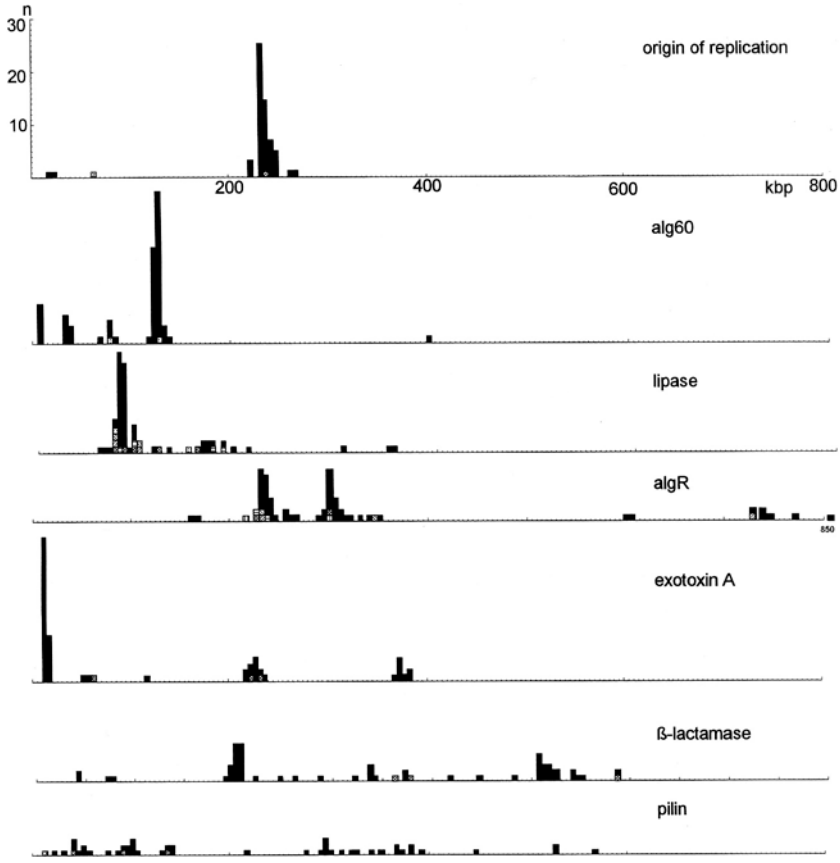


Figure 3. Southern analysis of *P. aeruginosa* clones: size variation of hybridizing *Spe*I fragments. The number of fragments detected by one probe is given as a function of size. Fragment length was counted in 5 kbp increments. Fragment size variation within a clone is indicated by identical open symbols.

polymorphism between strains. Twenty-four of these regions (termed variable segments) were identified. These variable segments are scattered throughout the genome; however, nine segments are immediately adjacent to tRNA or tmRNA genes. Ernst *et al*.[20] detected 38 PAO1 gene islands to be absent or divergent in at least 2 out of 14 examined clones. Besides the pyoverdine cluster, the flagellar regulon and the O-antigen biosynthetic gene cluster described above, a further not yet characterized exopolysaccharide is encoded by a gene island (PA1383–PA1393). Numerous islands in *P. aeruginosa* encode pyocins or phage proteins.[20] Six further islands were each adjacent to 1 of the 10 members of the *vgr* gene family, genes associated with rearrangement hotspots in the *E. coli*

chromosome. Seven of these 38 islands belong to the subset of 10 chromosomal regions whose low G + C content suggested that they were sites of recent horizontal transfer in PAO1.[85,95]

The hotspots for gene island replacement are apparently the regions where most intra- and interclonal genome diversity takes place. Intraclonal genome diversity has so far been studied in 21 *P. aeruginosa* isolates of clone C.[67] Clone C is one of the major clones in the *P. aeruginosa* population and has frequently been isolated from inanimate and disease habitats.[18,68] Clone C consists of closely related genotypes (also called clonal variants), each of which is characterized by a unique macrorestriction fragment pattern. Within clone C the total genome size varies at maximum by 300 kb. In total 34 different insertions or deletions were mapped that each were present in 1 to 13 strains. The acquisition and loss of DNA occurred preferentially around the terminus of replication but was not observed around the origin of replication, from about *rrnC* to *rrnA* (Figure 4). Three regions close to the *phnAB, pilA*, and *lipH* loci were subject to extensive variation processes. These hypervariable regions of the clone C chromosomes match with the hotspots of variation in the Southern and PAO1 microarray hybridization experiments. Subsequent sequencing revealed that most larger genome islands are located in these regions.

The ca. 110 kb large hypervariable region located near the *lipH* gene was sequenced in two clone C strains, strain C and strain SG17M.[43] In both strains the region consists of an individual strain-specific genome island of 111 (strain C) or 106 (SG17M) open reading frames (ORFs) and of a 7 kb stretch of clone C-specific sequence of nine ORFs. The left boundary of the islands is a cluster of tRNA genes comprising one $tRNA^{Glu}$ gene followed by two identical $tRNA^{Gly}$ genes separated by 84 bp, one serving as the integration site for the *P. aeruginosa* genome island PAGI-2 in strain C, the other for PAGI-3 in SG17M. PAGI-2 and PAGI-3 terminate at the right end with the terminal 16 and 24 nucleotides of the 3′ end of the $tRNA^{Gly}$ gene, respectively. The same organization is seen for the *Pseudomonas clc* genome island that contains the genes encoding the degradation of 3-chlorobenzoate (see Chapter 16 by J.R. van der Meer in this volume for more information). In all three islands the first ORF adjacent to the $tRNA^{Gly}$ gene encodes a bacteriophage P4-related multidomain integrase with an unusual transposase-like C-terminus. PAGI-2 and PAGI-3 have a bipartite structure. The first part adjacent to the tRNA gene consists of strain-specific ORFs encoding metabolic functions and transporters, the majority of which has homologs of known function in other eubacteria. The second part is made up of a syntenic set of ORFs the majority of which is classified as conserved hypotheticals. Forty-seven of these ORFs are arranged in the same order in both islands with a pairwise amino acid identity of 35–88% (Figure 5). Interestingly, PAGI-2 is also found with 100% sequence identity in the *Ralstonia metallidurans* CH34 chromosome[43] indicating that first, the genome island is also present in other phylogenetically distant taxa,

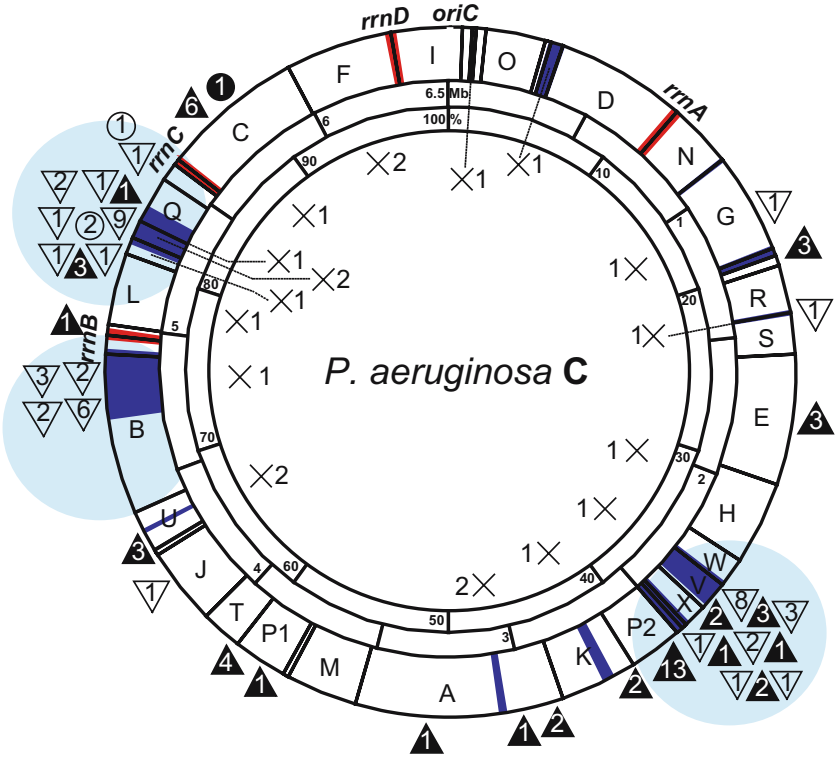


Figure 4. *Spe*I restriction map of *P. aeruginosa* C summarizing all chromosomal changes that occurred within the 21 analyzed clone C strains. Open triangles represent deletions, filled triangles represent insertions, open circles indicate deletion of a *Spe*I site, and filled circles indicate additional *Spe*I sites. Crosses indicate endpoints of recombination. Numbers in the symbols refer to the frequency of an additional genome alteration. Shaded regions indicate additional genetic material in strain C in comparison to PAO1. The hotspots of gene replacement around the *pilA* and *lipH* loci are indicated by the large circles.

and second, this type of island may have closer homologs in other clones and taxa than within the same clone. Subsequent hybridization analyses revealed that additional copies of PAGI-2 that first had been sequenced in a *P. aeruginosa* isolate from a German CF patient's lung, were present in the majority of tested *R. metallidurans* and *R. campiniensis* isolates from wastewater and polluted habitats in Europe and North America.[38]

PAGI-2 and PAGI-3 are prototypes for tRNA-associated gene islands that are causative for the genetic make-up of one of the hypervariable areas of the *P. aeruginosa* chromosome. The other two hypervariable regions in the *P. aeruginosa* chromosome with pronounced genomic variability reside in the

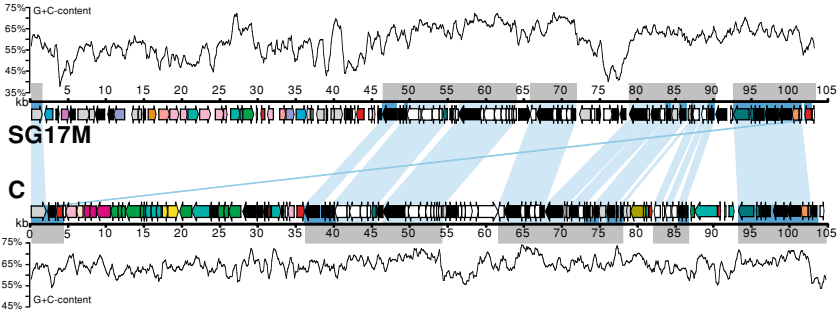


Figure 5. Comparison of the strain-specific gene islands in the *P. aeruginosa* clone C strains SG17M (upper line) and C (lower line). Gene are represented by arrows. Homologous ORFs are linked by light bars. Genes with homologs in the *Xylella fastidiosa* genome island[78] are highlighted with a dark background. Gray boxes above and below the gene maps mark the syntenic set of core genes that are characteristic for this type of island.[51] Reproduced from the article by Larbig *et al*.[43] with permission by the authors and the American Society for Microbiology.

vicinity to the *pilA* and *oprL – phnAB* loci. The duplicated copies of a $tRNA^{Lys}$ gene were identified as the hotspots for the integration and excision of DNA in these regions. The large plasmid pKLK106 sequentially recombined with either of the two $tRNA^{Lys}$ genes in *P. aeruginosa* clone K strains, giving rise to reversible rearrangement of a 106 kb genome island in sequential isolates from CF patients[35] (Figure 9). In all investigated clone K strains, both episomal and chromosomal copies were detected. During the propagation of single colonies on agar plates *in vitro*, progeny that had retargeted pKLK106 into the other $tRNA^{Lys}$ locus were regularly observed, indicating that pKLK106 is mobilized and reintegrated into the clone K chromosomes at high frequency.

In strain PAO1 the $tRNA^{Lys}(1)$ gene close to *oprL–phnAB* is located between coding sequences PA0976 and PA0977. The 8.9 kb DNA block 3′ of $tRNA^{Lys}$ from PA0977 to PA0987 represents a non-conserved insertion that terminates with duplicated 22 bp of the 3′ end of the $tRNA^{Lys}(1)$ gene, presumably the former attP-site of the integrated element. This 8.9 kb block of PAO1-specific DNA is absent in clone K strains, harboring PA0988 as their first PAO1 homolog downstream of $tRNA^{Lys}(1)$.[35]

In strain C a 23.4 kb large gene island termed PAGI-4 is integrated at this $tRNA^{Lys}(1)$ site.[39] PAGI-4 substitutes PA0977 to PA0994 and consists of two blocks of non-PAO1 sequence that each are flanked by short stretches of PAO1-homologous sequence. The first block of 9.5 kb of non-PAO1 sequence flanked by truncated versions of PA0977 and PA0980, shares conserved synteny and 87–99% amino acid sequence with ORFs of PAGI-2, PAGI-3, and pKLC102 (see below). The second 12.7 kb DNA segment flanked by truncated versions of PA0981 and PA0994 encodes the typical elements of a transposon similar to Tn*4652* from *Pseudomonas putida.*

Strain PA14 carries the 10.7 kb island PAPI-2 at this location that shares substantial sequence similarity with the PAO1 genome island.[30] The PAO1 pyocin genes PA0984-85 are replaced in PAPI-2 by the cytotoxin *exoU* gene and its chaperone *spcU*, and accordingly PAPI-2 has termed a pathogenicity island. In two clinical isolates another 81 kb island that also contains the *exoU* gene has been identified to again reside at the very same genomic position.[95] In summary, five different genome islands varying of 8.9–106 kb in size have yet been identified to integrate into the $tRNA^{Lys}$(1) gene close to *oprL–phnAB*.

Three different genome islands PAPI-1, pKLK106, and pKLC102 are known to insert into the $tRNA^{Lys}$(2) gene close to *pilA*. pKLK106 and pKLC102 are highly homologous plasmids. Clone K and clone C strains from the environment harbored chromosomal and episomal copies of this mobile genetic element.[39] PAPI-1, pKLK106, and pKLC102 share numerous features: approximate size (108, 106, 102 kb), a $tRNA^{Asp}$, $tRNA^{Pro}$, and $tRNA^{Lys}$ gene cluster at their leftward PAO1 junction, and a direct repeat of the 3′ half of the $tRNA^{Lys}$ gene at their right border, and the integrase and the chromosome partitioning genes at the ends of the island, similar to PAGI-2 and PAGI-3.[30,39] PAPI-1 is a pathogenicity island because it carries at least 19 virulence factors that occur on genomic islands found in a wide spectrum of other pathogenic bacteria.[30] pKLC102 contains the 8.5 kb *chvB* gene homologs of which are known to confer host tropism and virulence and to be essential for the interaction of the bacterium with its eukaryotic host. PAPI-1 and pKLC102 encode type IV group B pili and type IV thin sex pili, respectively, and share a set of homologs found as island-specific genes in PAGI-2, PAGI-3 (see above), and numerous genome islands in other proteobacteria (Figure 6). Fifteen of 33 core genes common to 15 genome islands from β- and γ-Proteobacteria were congruent with the phylogenetic relationships of each of the individual genes indicating that all five large genome islands known so far in *P. aeruginosa* belong to one family of related syntenic genomic islands with a deep evolutionary origin.[51] The mobile pKLC102 shares with PAPI-1 the phage module that conferred integrase, the *att* element and the syntenic set of genes, but it differs from PAPI-1 in carrying a plasmid module that conferred *oriV* and genes for replication, partitioning, and conjugation.[39]

The only large genome island known so far that is not associated with a tRNA gene is the 49 kbp PAGI-1. This first described genome island in *P. aeruginosa* is widely distributed in the population.[45] The island was probably assembled from two ancestral components of different G + C content. 35 kb of the higher G + C content portion is also found in the *P. putida* KT2440 genome.[53] PAGI-1 contains genes potentially involved in oxidative stress resistance, and replaced PAO1 genes PA2218 to PA2222. Furthermore, in other *P. aeruginosa* strains, this region contains an insertion of 3 kbp of DNA unrelated to PAGI-1.

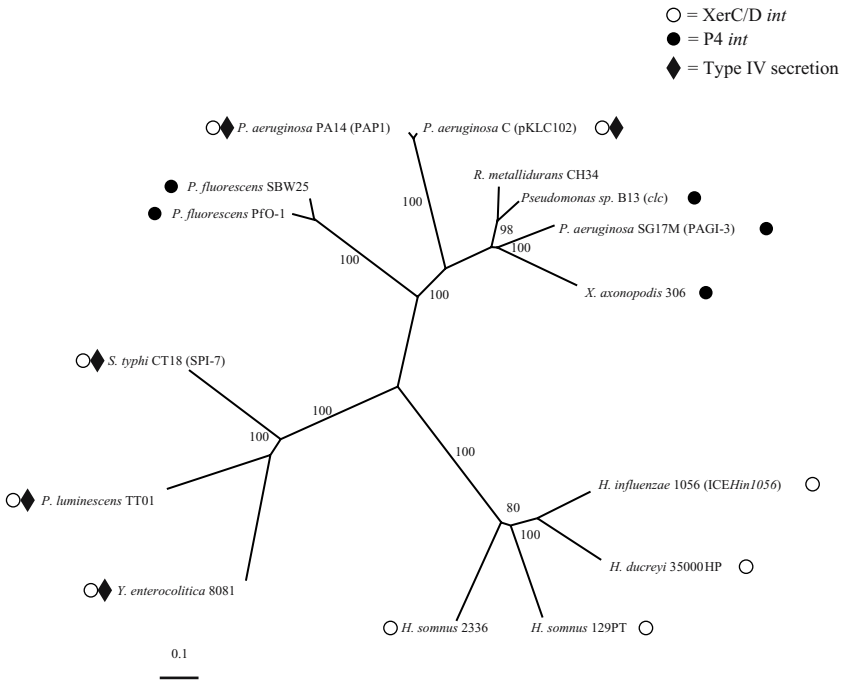


Figure 6. The amino acid sequences of the 15 predicted sequences common to 15 genome islands were concatenated and aligned by ClustalX. The alignment of each of the genes alone was consistent with the alignment illustrated. Reproduced from the article by Mohd-Zain *et al*.[51] with permission by the authors and the American Society for Microbiology.

In summary, the accessory genome of *P. aeruginosa* is made up of numerous genome islets and islands. Most genome islands analyzed so far are integrated into tRNA genes and share a signature of syntenic genes that are widespread among proteobacteria. Accessory genes are nestled among these core genes and confer a diverse repertoire of strain-specific features.

2.3. Population Biology of *P. aeruginosa*

Bacteria can have population structures ranging from the fully sexual to the highly clonal. Several independent studies in the last years demonstrated that *P. aeruginosa* has a nonclonal population structure punctuated by highly successful epidemic clones or clonal complexes.[14,36,52,60,76]

By applying multilocus SNP typing on two unrelated strain collections, the index of association was consistently calculated in two independent studies to be 0.29[14] and 0.31[36] indicating that *P. aeruginosa* has a nonclonal population

structure. The index of association is a measure of the extent of linkage equilibrium within a population by quantifying the amount of recombination among a set of sequences and detecting associations between alleles at different loci. Comparisons of the topologies of neighbor-joining trees for the nucleotide sequences of individual loci revealed in both studies that there was little, if any, congruence between the trees. Strains, which belong to the same genotype, are characterized by non-random association of alleles that is not disrupted by recombination. In contrast, the recombination frequency of large chromosomal segments between genotypes is high enough to break up clonal associations and have all genotypes in linkage equilibrium to each other. Hence, the *P. aeruginosa* genotypes are equivalent biovars that form a net-like population structure. Each genotype represents a cluster of closely related strains (clonal variants) that share identical alleles.[36]

When typing a strain, the core genome can be represented by the multilocus SNP genotype of conserved genes whereas both core and accessory genome can be represented by the PFGE-separated macrorestriction fragment profile. By comparative SNP and *Spe*I PFGE genotyping[52] numerous cases were resolved whereby strains shared the SNP genotype but had different *Spe*I macrorestriction profiles. Interestingly, this finding applied to the most abundant *Spe*I genotypes. Further examples were the completely sequenced strains PAO1 and PA14 isolated in Australia and the US, respectively, which shared their SNP genotype with numerous clinical and environmental isolates from Europe. This data indicate the high proportion of dominant epidemic clones in the *P. aeruginosa* population. These epidemic clones such as the European clone C, the Australian, and the UK epidemic clones have unrelated genotypes, suggesting that they have evolved independently.[14,76]

RFLP analysis of the chromosome and SNP analysis of individual genes measure different evolutionary forces. The conservation of the SNP genotypes and the divergence of *Spe*I macrorestriction patterns in strains sharing the same SNP profile agree with the idea that the core genome of *P. aeruginosa* is highly conserved and that its evolution and structure rely more on acquisition, loss, and rearrangements of genome islands and genome islets than on point mutations. In other words, horizontal gene transfer has a more important role than point mutations on the evolution of *P. aeruginosa* in most habitats. The only known exception seems to be the uncommon habitat of the human respiratory tract where a high proportion of hypermutable *P. aeruginosa* strains emerges over time.[56]

In enterobacteria a single genotype predominates one habitat (Figure 7). Genotypes are associated with particular pathogenicity islands which result in disease-associated clones. In contrast, there is no correlation between *P. aeruginosa* clones and habitats (Figure 7). Dominant clones are ubiquitously distributed in both disease and environmental habitats: for example, members of the same clone were recovered from oil shale and from the lungs of patients

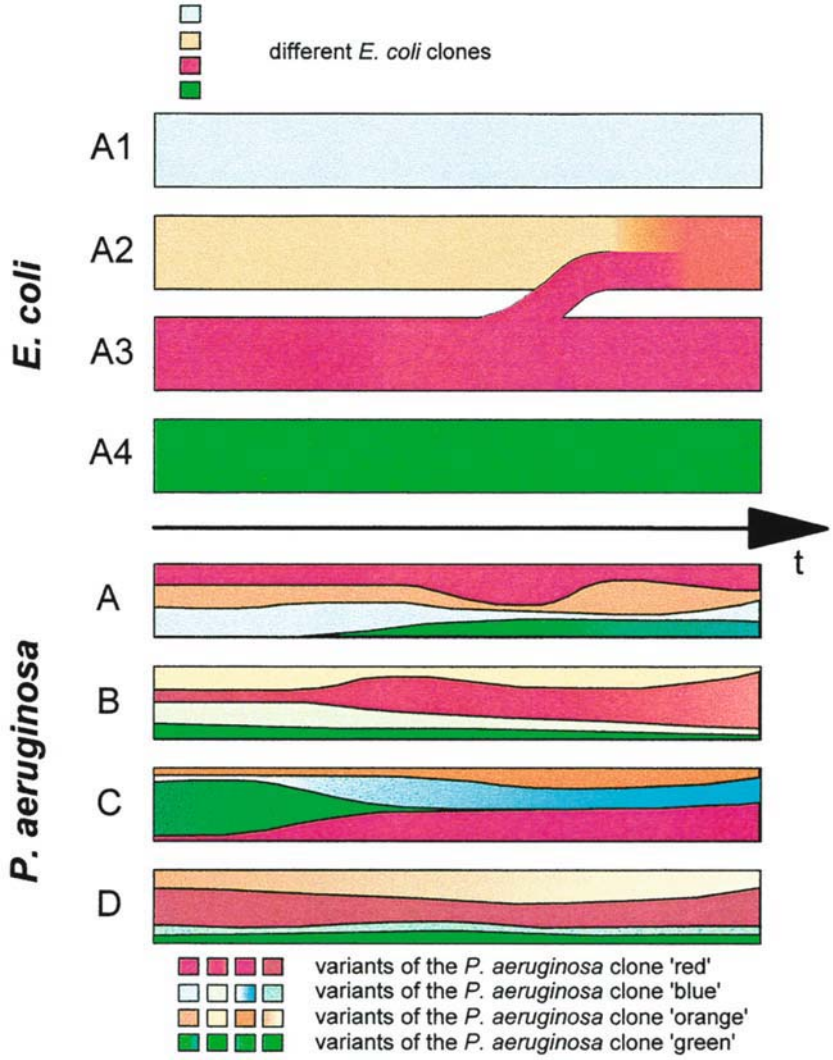


Figure 7. Clonal population structure of *E. coli* and *P. aeruginosa* differing in spatio-temporal distribution. The related disease habitats of *E. coli* are designated A1, A2, A3, and A4; the diverse disease and environmental habitats of *P. aeruginosa* are symbolized by A, B, C, and D. Whereas *E. coli* shows a clear correlation between clone and habitat (disease-associated clones), being only occasionally interrupted by horizontal gene transfer, the same spectrum of *P. aeruginosa* clones colonizes even unrelated habitats. Individual variants of a certain clone may predominate in several niches. Variants of *P. aeruginosa* clones undergo adaptive genetic changes, suggested by the shading. Reproduced from the article by Kiewitz and Tümmler[36] with permission by the authors and the American Society for Microbiology.

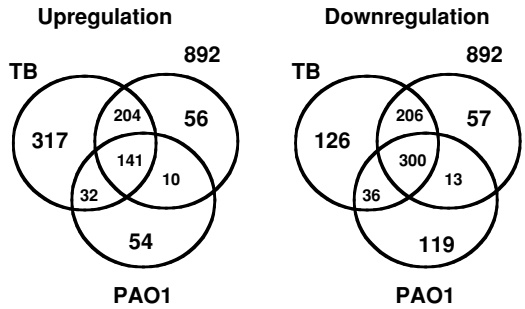


Figure 8. Transcriptome analyses in strain PAO1 and the two clone TB strains TB and 892: PAO1 genes with significantly changed expression levels in hydrogen peroxide treated *P. aeruginosa* TB, 892, and PAO1. Genes that were found to be up- or down-regulated in two or three strains are indicated in the intersections.

with CF.[23,52] Disease and environmental isolates of *P. aeruginosa* clones are indistinguishable in their genotypic and chemotaxonomic properties[14,23,52] and are functionally equivalent in several traits relevant for their virulence and environmental properties.[1] In summary, *P. aeruginosa* appears to be so versatile that it can colonize a variety of different ecological niches without specialization (Figure 7).

3. CLONAL VARIATIONS OF PHENOTYPE

The polymorphic loci in the *P. aeruginosa* core genome lead to a clone-specific repertoire of pyoverdines, LPS, pili, and flagella, whereby the latter two do not only vary in the primary amino acid sequence, but also in the posttranslational glycosylation pattern. Most interclonal variations are conferred by the genome islets and islands of the accessory genome, but their impact on phenotype has yet not been resolved. The differential sets of O-antigens, pyocins, and phage receptors have been exploited for decades for strain typing, but the first systematic comparisons of global mRNA transcript and protein profiles in genetically typed strains have only reported recently.

Salunkhe *et al.*[69,70] compared the inter- and intraclonal diversity of the PAO1 transcriptome between strain PAO1 and the two clone TB strains TB and 892 under different environmental conditions. The number of expressed genes differed by clone. The PAO1 strain expressed 64% of its genes in LB medium, whereas the TB and 892 strains expressed more than 70% of genes. When the strains were exposed to the stressor hydrogen peroxide, 24, 17, 12% of the 5900 ORFs were significantly differentially regulated in *P. aeruginosa* TB, 892 and PAO1, respectively. Only 441 genes were consistently differentially expressed in the three strains. 729 genes showed strain-specific responses and 501 genes

were similarly regulated in two strains (Figure 8). The expression profile was significantly more related between the two members of the TB clone than between either strain with PAO1. The 892 strain shared 84% and 88% of up- and down-regulated genes with its more virulent clonal variant TB that exhibited a more strain-specific expression profile with 50% and 76% of up- and down-regulated genes being in common with 892. It is noteworthy that strains TB and 892 exhibit identical SNP and *Spe*I genotypes[52] and that the DNA sequence was 100% identical in more than 100 kb of randomly selected loci of the core genome.[36] Nevertheless the global mRNA expression profiles of the two clonal variants were divergent when the strains were cultivated simultaneously in LB medium in the presence or absence of oxidative stress. Apparently there are a few sequence variations in some key genes that account for the divergent phenotypes. In other words, genetically very closely related *P. aeruginosa* strains present a strain-specific mRNA expression profile that under carefully controlled identical conditions is highly reproducible, but distinct from members of the same clone.

The proteome of the two clonal variants TB and 892 has also been compared by two-dimensional polyacrylamide gel electrophoresis coupled to mass spectrometry to map the extracellular, intracellular, and surface sub-proteomes and to identify differentially expressed proteins.[3] About 4% of all detected protein spots were differentially expressed between both strains including absent or present spots and spots with a more than two-fold changed intensity. Nineteen of 78 differentially expressed spots were identified by mass spectrometry on the basis of a predicted gene product in the genome database for *P. aeruginosa* PAO1, for 13 additional spots mass fingerprints were obtained which most likely represent clone-specific proteins of the TB lineage. Many of the protein spots in TB that were missing or expressed at lower levels in the less virulent 892 strain were identified as quorum-sensing regulated virulence factors.

Strains of *P. aeruginosa* can be phenotypically classified by their mode of pathogenicity as either invasive, where the bacterium is internalized by host cells, or cytotoxic, where the host cell is killed without internalization through the expression of cytotoxicity factors. These phenotypes are thought to depend primarily on the interactions of pseudomonal membrane and secreted proteins with host cells. Nouwens *et al*.[54] compared the proteomes of the outer membrane and extracellular protein-enriched fractions from the invasive strain PAO1 and the cytotoxic strains 6206. Membrane protein strain differences were typically the result of minor amino acid sequence variations resulting in small mass and isoelectric point shifts visible on two-dimensional gels. Analysis of extracellular proteins from stationary phase growth, however, revealed significantly different protein profiles between the two unrelated clones. Extracellular fractions from the invasive PAO1 strain were dominated by extracellular proteases including elastase (LasB), LasA protease, and chitin-binding protein, as well as several previously designated ‚conserved hypotheticals' of unknown function. Conversely, extracellular fractions from strain 6206 consisted mainly of

cellular and membrane exposed proteins including GroEL, DnaK, and flagellar subunits. These are thought to result from cellular turnover during growth and the reliance on the secretory mechanisms of this strain to produce high levels of cytotoxicity factors, such as ExoU, which may be produced only upon specific interactions with host cells.

Wehmhöner *et al*.[92] compared the proteome profiles of six *P. aeruginosa* clones grown in modified minimal Vogel-Bonner medium. The proteome analysis revealed almost identical patterns for the cellular extracts, whereas interclonal diversity were demonstrated for the secretomes of cultured *P. aeruginosa*. The diversity was even greater for the immunogenic protein patterns expressed *in vivo*. The observed interclonal variability of the secretome may reflect the differential, clone-specific regulation of gene expression and/or the utilization of genes that are not encoded by the core genome but are encoded by the highly dynamic accessory genome. To differentiate between the two mechanisms, Wehmhöner *et al*.[92] also analyzed the proteomes of sequential clonal variants with diverse morphotypes. A *P. aeruginosa* isolate that formed irregularly shaped colonies was compared with a hyperpiliated and autoaggregative *P. aeruginosa* small colony variant. The expression profiles of cellular extracts of the two morphotypes exhibited only minor differences, in contrast to the marked differences in the expression profiles of the extracellular fractions. Mass spectrometry revealed that the small colony variant overexpressed proteins secreted by the type I and type III secretion systems. This finding implies that the variability of the secretome is due to differential regulation of protein expression, possibly as a consequence of small adaptational mutations. These observations were backed up by genome-wide transcriptional profiles of the two clonal morphotypes.[91] Of the more than 300 differentially expressed genes, the upregulation of the type III secretion system and the respective effector proteins in the small colony variant was the most striking finding. The conserved intracellular proteome of strains grown *in vitro* probably reflects the fact that the need for adaptation under these conditions is low, and inter- and intraclonal differences that reflect the versatility of niche specialists are not likely to be detected. Moreover, the cellular proteome comprises mostly proteinaceous cell constituents that are expected to be species-specific but not clone- or strain-specific. However, the secretome expression is strongly strain and morphotype-specific. Since the secreted *P. aeruginosa* proteins come into direct contact with their environment, they could be especially important and thus be essential for bacterial adaptation. Moreover, the secretome includes important virulence factors essential for establishment of an infection within the human host. In summary, the secretome is a sensitive measure of *P. aeruginosa* strain variation.

A special case of intraclonal diversity are the strain variations that occur during subculturing *in vitro*. A timely and important example is the completely sequenced reference strain PAO1. The sequenced strain[85] differs from the ancestor strain that had been independently physically mapped in Australia and Germany,[33] by a 1.7 Mbp inversion between the *rrnA* and *rrnB* loci and an about

20 kbp deletion close to the *rrnC* locus. Moreover, PAO1 stocks maintained at different laboratories are not identical in phenotypic traits, a spectacular example being the differential virulence in infection models even though the strain had been originally obtained from the same public collection (own unpublished data). During storage and subculturing the PAO1-derived isolates apparently diversified by inversion, deletion, and point mutation.

4. INTRACLONAL EVOLUTION AND DIVERSITY IN CLINICAL HABITATS

4.1. Hospital-Acquired Infections

P. aeruginosa is resistant to many antimicrobial agents and a major source for nosocomial infections in predisposed individuals. Hence, the major practical issue to assess clonality and intraclonal evolution are infection control measures to determine epidemic clonality amongst multidrug-resistant strains or to document outbreaks of (drug-resistant) clones (as examples see refs [34,50,58,59]). A few groups combined the molecular epidemiology of hospital-acquired infection with the characterization of intraclonal diversity.

Hocquet *et al*.[32] retrospectively analyzed the intraclonal variation of drug resistance of a serotype O:6 multidrug-resistant *P*. *aeruginosa* clone during a 4-year long outbreak at a French University Hospital. This clone was initially recognized because of its particular susceptibility profile to aminoglycosides [conferred by an ANT (2")-I enzyme] and fluoroquinolones (caused by mutations in the QRDR of *gyrA* and *parC*) and because of its elevated resistance to many β-lactams. The susceptibility profile of this epidemic clone to fluoroquinolones and aminoglycosides was relatively stable during the outbreak but showed important isolate-to-isolate variations in the susceptibility to β-lactams. Analysis of 18 genotypically related isolates selected on a quarterly basis demonstrated alterations in DNA topoisomerases, constitutive overexpression of the MexXY efflux system, derepression of intrinsic AmpC β-lactamase and sporadic deficiency in the carbapenem-selective porin OprD. Of the 18 isolates, 14 were also found to overproduce the efflux system MexAB-OprM as a result of alteration of the repressor protein MexR. Of the four isolates exhibiting wild-type MexAB-OprM expression despite the MexR alteration, two appeared to harbor secondary mutations in the *mexA-mexR* intergenic region and one harbored secondary mutations in the putative ribosome binding site located upstream of the *mexAB-oprM* operon. In conclusion, many mechanisms were involved in the multiresistance phenotype and the clone sporadically underwent substantial genetic and phenotypic variations during the course of the outbreak.

P. aeruginosa is responsible for severe nosocomial pneumonia in mechanically ventilated patients. Denervaud *et al*.[16] collected 442 *P. aeruginosa* isolates during the first 3 days of documented colonization of 13 intubated

patients in order to study quorum-sensing dependent phenotypic traits. The 442 isolates belonged to nine different clones. Eighty-one percent of the isolates produced homoserine lactones and quorum-sensing dependent extracellular virulence factors, including total exoprotease, elastase, HCN, pyocyanin, and rhamnolipids, at levels equivalent to those of the reference strain PAO1, but 19% of the isolates were deficient in cell-to-cell signaling, because the *lasR* gene encoding the LasR transcriptional regulator was inactivated by various mutations. A subset of these isolates also had mutations in the *rhlR* gene, probably explaining the defect in both homoserine lactone and extracellular virulence factor production. Since the homoserine lactone production of these strains was complemented by the chromosomal insertion of the wild-type *lasR* and *rhlR* genes, additional mutations are unlikely. Three of the 13 patients presented a *P. aeruginosa* pneumonia as a complication of their respiratory colonization of whom two subsequently developed a *P. aeruginosa* bacteremia. These bacteremic isolates were clonal variants carrying the lasR or lasR/rhlR mutants. This is the first report on clinical isolates that are unable to produce cell-to-cell signals as a result of both *lasR* and *rhlR* mutations, and it is interesting to note that the intraclonal evolution toward loss of quorum-sensing was associated with the gain-of-invasiveness to breach the airway epithelial barrier.

4.2. Cystic Fibrosis

Most information about the evolution of intraclonal diversity of *P. aeruginosa* was obtained from retrospective cross-sectional and longitudinal analyses of isolates recovered from the atypical habitat of the CF lung. CF is a severe monogenic disorder of ion transport in exocrine glands that is caused by mutations in the CF transmembrane conductance regulator *(CFTR)* gene. The basic defect predisposes to chronic bacterial airway infections, particularly with *P. aeruginosa.* The *P. aeruginosa* infections in CF are a paradigm of how environmental bacteria can conquer, adapt, and persist in an atypical habitat and successfully evade defense mechanisms and chemotherapy in a susceptible host. Airway infections with *P. aeruginosa* in individuals with CF are unique in that they chronically affect a host who is immunocompetent in terms of cellular and humoral responses but is immunocompromised by impaired airway clearance. Once *P. aeruginosa* has taken residence in the CF lungs, the organism is notoriously resistant to eradication by chemotherapy. The pseudomonads chronically colonize the bronchiolar lumen and virtually never breach the epithelial barrier (reviews: refs [26,89]).

4.2.1. Clonal Variations of Genotype in CF Lungs

Most individuals with CF become chronically colonized with a single clone of *P. aeruginosa* that stays in the lungs for many years. Turnover of clones was seen in the author's laboratory after 5–15 years in about half of the

investigated patients. Transient or permanent co-colonization with more than one clone was seen in 20–30% of patients. These conditions that CF lungs are chronically infected for years by one or a few *P. aeruginosa* clones are precisely those that theoretical studies predict for the evolution of mechanisms that augment the rate of variation. Determination of spontaneous mutation rates in 128 *P. aeruginosa* isolates from 30 CF patients revealed a high proportion (20%) with an increased mutation frequency (mutators).[55,56] Seven out of 11 analyzed CF mutator strains were found to be defective in the mismatch repair system. The alterations in the *mutS*, *mutL,* and *uvrD* genes were found to be responsible for the mutator phenotype. In four cases (three *mutS* and one *mutL*), the genes contained frameshift mutations. The fourth *mutS* strain showed a 3.3 kb insertion after the 10th nucleotide of the *mutS* gene, and a 54 nucleotide deletion between two eight nucleotide direct repeats. This deletion, involving domain II of MutS, was found to be the main one responsible for *mutS* inactivation. The second *mutL* strain presented a K310M mutation, equivalent to K307 in *E. coli* MutL, a residue known to be essential for its ATPase activity. Finally, the uvrD strain had three amino acid substitutions within the conserved ATP binding site of the deduced UvrD polypeptide, showing defective mismatch repair activity. In summary, intraclonal evolution of *P. aeruginosa* in CF lungs can be driven by hypermutable clonal variants. Since the proportion of mutators in the population increases over time, point mutations preferentially accumulate during the late stages of the infection.

Mutator strains were not found in 75 non-CF patients acutely infected with *P. aeruginosa*,[56] but were also seen in 30 patients with non-CF underlying chronic respiratory diseases (22 with bronchiectasis and 8 with chronic obstructive pulmonary disease).[46] Seventeen of the 30 patients were colonized with hypermutable strains. The *mutS* gene was inactivated in isolates from 11 patients. Multiple antimicrobial resistance was documented in 42% of the hypermutable strains in contrast to 0% of the non-hypermutable strains. This study demonstrates that first, hypermutation is a key factor for the emergence of the multidrug resistance phenotype; and that second, in contrast to what has been described in acute processes, hypermutable *P. aeruginosa* strains are highly prevalent in chronic infections of the human respiratory tract.[46,56]

Besides the mismatch repair system and the targets that confer multiple antimicrobial resistance, further known hotspots for mutation in CF isolates are the *mucA* and *lasR* genes. Inactivation of *lasR* is often causative for the loss of production of *N*-acylhomoserine lactones (AHL) which is not rare in CF, particularly if the lung is co-colonized with *Burkholderia cepacia*.[19,24] The mucoid phenotype of *P. aeruginosa* has been linked to mutations in a gene cluster designated as the *mucABCD* genes that encode proteins that inhibit the activity of the alternative σ-factor AlgU.[49] When alginate production is minimal, AlgU (also referred to as AlgT) is bound in a complex along with MucA and MucB, but under environmental stress conditions this complex is disrupted,

leading to the release of AlgU into the cytosol. AlgU acts on the key alginate biosynthesis gene *alg*D, which encodes a GDP mannose dehydrogenase, and on *alg*R, a response regulator genes that increases alginate synthesis.[74] Mutations in *mucA, B,* and *D* are held responsible for alginate overproduction and conversion to a stable mucoid phenotype in *P. aeruginosa*, while mutations in *mucC* do not cause any overt effects on alginate synthesis.[9,49] Consistent with this hypothesis, mutations in *mucA* have been detected in mucoid *P. aeruginosa* strains isolated from chronically infected CF patients.[2,9,10,75,84] Alginate production is also dependent on a second alternative σ-factor, RpoN, and likely on other mutations in genes which are not known at present. However, according to more recent studies mutations in the *mucABD* cluster are not exclusively correlated to overexpression of alginate in *P. aeruginosa* CF isolates.[11,84] The combined analysis of quantitative alginate expression and *mucA*, *mucB, mucD,* and *algU* sequencing in 37 *P. aeruginosa* strains revealed that a distinct proportion of phenotypically non-mucoid *P. aeruginosa* strains carried *mucA* stop mutations which were also present in alginate-overexpressing, mucoid *P. aeruginosa* strains.[11] Since sequence analysis of *algU* did not reveal any mutational genetic changes, other, unknown, mechanisms are presumably regulating alginate expression in these *mucA* mutated strains.

P. aeruginosa in CF lungs is not only prone to point mutations, but also to gross changes of the chromosomal frame. When Ernst *et al*.[20] analyzed 13 isolates from seven young CF children by hybridization on PAO1 whole genome DNA microarrays, they detected 2 strains with large deletions (strain CF250: 119 kb (PA1909–PA2010); strain CF5296: 189 kbp (PA2273–PA2409). The latter deletion eliminates the hypervariable pyoverdine locus.

Reversible genome rearrangements were seen in CF isolates which were carrying the mobile genetic element pKLK106. pKLK106 reversibly recombined with sequential clone K chromosomes at one of the two $tRNA^{Lys}$ genes[35] (Figure 9). In all investigated sequential clone K CF strains both episomal and chromosomal copies were detected.

Physical mapping of 18 CF clone C isolates revealed inversions in eight strains, two of which were harboring two nested inversions.[67] Besides one small scale inversion of 40 kb, the inversions ranged from 1 to 5 Mbp whereby their recombination endpoints scattered on the chromosome. In six cases the region of the terminus of replication was included in the recombinational exchange and was shifted by maximal 17% of genome size (Figure 4). A hotspot of recombination was mapped to the pKLC102 locus:[40] All investigated clone C isolates from aquatic habitats and the hospital environment harbored chromosomal and episomal copies of pKLC102. However, many isolates from CF lungs contained either no (C5) or only chromosomally integrated pKLC102 (C2) (Figure 9). Of the four subgroups of clone C,[67] subgroup C was exclusively represented by CF lung isolates and differed from the other three groups by the insertion of the class 1 composite transposon TNCP23 into chromosomally

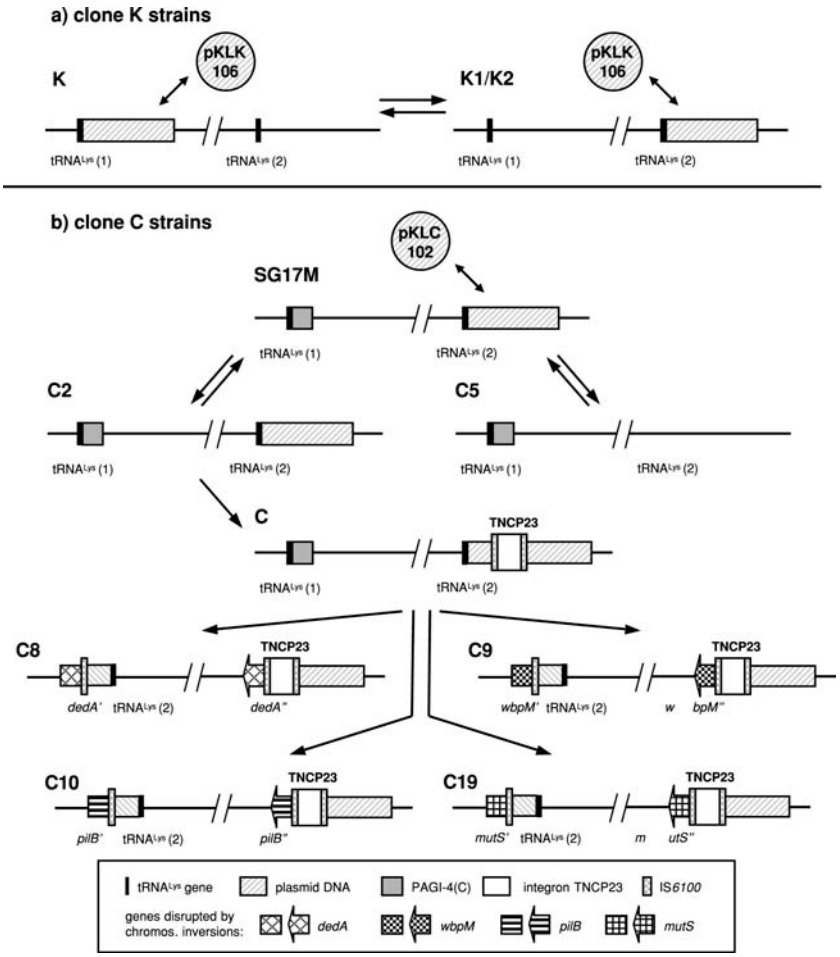


Figure 9. Evolution of *P. aeruginosa* strains linked to plasmid DNA. (a) Reversible integration of plasmid DNA into two possible sites of clone K strains. (b) Different forms of plasmid DNA in clone C strains. In subgroup SG17M pKLC102 is found episomally and integrated into the genome at $tRNA^{Lys}(2)$. Strain C5 apparently lost the pKLC102 DNA, while strain C2 only harbors the integrated form. In subgroup C the integron TNCP23 inserted into chromosomally integrated pKLC102. Free plasmid is not detectable in subgroup C strains indicating that TNCP23 prevented mobilization. TNCP23 is flanked by copies of IS*6100*. Intramolecular transposition of the left copy of IS*6100*-L is coupled with an inversion of the chromosomal region between the transposed copy and IS*6100*-L in some strains of subgroup C. For these strains C8, C9, C10, and C19, the $tRNA^{Lys}(1)$ area is not shown. Reproduced from the article by Klockgether *et al.*[39] by permission of the authors and the American Society for Microbiology.

integrated pKLC102, which may have been acquired because of the encoded *aadB* gene conferring gentamicin resistance (Figure 9). *P. aeruginosa* converges in CF lungs to a common phenotype characterized by the decreased production of membrane components, cellular appendages and secreted factors (see below). This phenotypic signature was partially gained in subgroup C strains by TNCP23-mediated chromosome remodeling. Intramolecular transposition of the active IS-*6100* element of TNCP23 led to large chromosomal inversions which disrupted genes that are typically inactivated during the adaptation of *P. aeruginosa* to the atypical habitat of CF lungs (Figure 9). In parallel the integrity of pKLC102 was destroyed. The two attachment sites were separated so that the genetic contents of pKLC102 was irreversibly fixed in the chromosome. In summary, Figure 9 visualizes the evolution of plasmid pKLC102 from a mobile genetic element to an irreversibly fixed genome island that finally was disrupted and distributed to separate chromosomal regions. It should be noted that the increasing complexity of genome organization caused by insertion, transposition, and inversion was accompanied by mutation, deletion, and/or duplication of sequence close to the breakpoint.

In summary, inversions, deletions, and point mutations are common for the intraclonal evolution of *P. aeruginosa* in CF lungs.

4.2.2. Clonal Variations of Phenotype in CF Lungs

Apart from the evolution of genome organization, *P. aeruginosa* develops common phenotypic features irrespective of clonal descent. This phenotypic conversion is characteristic for the CF lung habitat and, on the whole, is genetically fixed and irreversible. In other words, some aspects of intraclonal evolution are similar in all clones. Strains become LPS deficient (non-typable or polyagglutinating rough strains) and sensitive to lysis by complement.[27] On the other hand, CF isolates produce modified lipid A forms containing palmitate and aminoarabinose that are associated with resistance to cationic antimicrobial peptides and stronger induction of inflammatory responses such as interleukin 8 expression.[21] *P. aeruginosa* strains vary in their differential repertoire of bacteriophage receptors and the production of pyocins, which lyse susceptible *P. aeruginosa* strains. These differential properties are gradually lost in most *P. aeruginosa* during chronic colonization of the CF lung. Susceptibility to phages and secretion of pyocins decrease within a few years time.[65] The production of the major siderophore pyoverdine also changes over time. Pyoverdine-negative mutants emerge, but retain the capacity to take up pyoverdines.[17] CF strains become immotile owing to the loss of their flagella.[47] CF isolates from early colonization were highly motile and expressed both flagellin and pilin. However, about 40% of more than a 1000 examined isolates from chronically colonized CF patients lacked flagellin expression and were nonmotile. Sequential isolates remained consistently nonmotile. Lack of motility was rare among environmental isolates (1.4%) and other clinical isolates (3.7%) of *P. aeruginosa* examined.[47]

Moreover, during chronic colonization of the CF lung most *P. aeruginosa* strains reduce or even abolish the production of type II and III secretion effector proteins and thus reduce cytotoxicity. While the killing of epithelial and phagocytic cells may be an important feature of acute infections, the same virulence mechanisms appear to be incompatible with chronic colonization of CF patients.[26] *P. aeruginosa* isolates were examined that had been obtained from 7 patients soon after their initial colonization and then again more than a decade later, after the establishment of chronic lung infections.[44] Early isolates were typically cytotoxic, the exception being the highly mucoid strains. Variants of the same clone, isolated years later from the same patients, were found to be nontoxic, suggesting that there has been a selection for loss of cytotoxicity. In many cases restoration of type III regulation, through the expression of the ExsA regulator, was able to reestablish ExoS secretion and cytotoxicity. However, this was not the only mechanism of attenuation since expression of ExsA was not able to restore ExoS secretion in all of the clinical isolates. Moreover, some strains accumulated more than one mutation in the type III secretion system so that ExoS secretion could be restored with ExsA expression; but cytotoxicity was still attenuated. What was also apparent from the pairwise comparisons of early and late isolates was that the phenomenon of delayed cytotoxicity associated with type II secretion effector proteins was also lost in later isolates. The respiratory tracts of CF patients therefore provide a strongly selective environment for the accumulation of pathoadaptive mutations, which favor a chronic existence that often necessitates the elimination of a potent cytotoxic mechanism.

Another common feature of intraclonal evolution in the CF lung is the diversification of morphotype, the hallmarks being the emergence of small colony variants[28,29] and mucoid colonies.[26,57] The alginate-overexpressing mucoid phenotype is typical for CF isolates and very uncommon in other habitats.[26] A subgroup of hyperpiliated small colony variants is prone to biofilm formation and induction of type II and type III secretion[29,91] which leads to increased virulence in infection models in contrast to the notion that most *P. aeruginosa* isolates from chronically colonized CF lungs are typically attenuated in virulence (see above). Auxotrophy is common in CF, particularly in those with severe underlying pulmonary disease.[87] At this late stage the auxotroph count exceeds more than 50% total CFU. The majority of auxotrophs required methionine as the sole factor.[86] In summary, the common conversion of phenotype of *P. aeruginosa* in CF lungs starts with morphotype diversification and loss of outer membrane constituents and cell appendages and ends with a progressive change from proto- to auxotrophy. At this final stage of adaptation to the CF lung, the accumulation of pathoadaptive mutations will probably impair the fitness of the bacteria to grow in other habitats.

Besides this convergence in phenotype that progressively happens over time, numerous phenotypic features are fluctuating in the *P. aeruginosa* CF

lung communities. These periodic changes in phenotype probably result from the emergence and disappearance of clonal variants with differential fitness. The repeated courses of regular antipseudomonal intravenous chemotherapy and the long-term administration of aerosolized aminoglycosides inadvertantly will select for resistant variants, and correspondingly fluctuations in the susceptibility patterns to antipseudomonal agents are a common finding in CF sequential isolates.

The adhesion phenotypes also vary strongly over time. *P. aeruginosa* mainly resides in the bronchiolar lumen of the CF airways embedded into a matrix of DNA, bacterial exopolysaccharides and human mucins.[88] *P. aeruginosa* uses chiefly proteins of its flagellar apparatus to initiate this binding and recognizes a variety of oligosaccharides that have been identified in mucins.[6,62] Among these are both neutral oligosaccharides and several forms of acidic oligosaccharides derived from the Lewis antigens.[72] Serial *P. aeruginosa* isolates from CF patients with advanced lung disease were characterized in their binding to CF human tracheobronchial mucins from three of these patients.[88] The strains differed strongly in their specificity for and affinity to mucin carbohydrate. Intra- and interclonal variation was equally pronounced indicating that the mucin-binding phenotype is not conserved within a particular clone.

Binding capacity to the airway epithelium is another trait subject to intraclonal variation. *P. aeruginosa* binding capacity to respiratory epithelial cells was studied in a representative panel of 634 sequential *P. aeruginosa* strains isolated from 26 CF patients, from the onset of colonization for up to 15 years of infection.[41] Adherence was strongly varying between clonal variants sampled at the same or different times, albeit three types divergent in the temporal evolution were noted: predominantly high binders, predominantly low binders at all times, or a shift from high binders at early colonization to low binders later on. Patients chronically harboring high binders had a worse prognosis than the others indicating that adhesion to the airway epithelium is a relevant pathogenic trait for *P. aeruginosa* to colonize and to persist in the CF lung.

Intraclonal variation may also be caused by AHL-dependent signaling between *B. cepacia* and *P. aeruginosa*.[24] When patients became transiently co-infected with an AHL-producing *B. cepacia* strain, AHL production by the co-residing *P. aeruginosa* isolates was switched off. However, months after the last *B. cepacia*-positive sputum the initial *P. aeruginosa* AHL profile was regained suggesting that AHL-mediated cross-talk between the two pathogens may affect the virulence of the mixed consortium which, in turn, may have selected for *P. aeruginosa* mutants producing lowered amounts of AHLs.

In summary, the CF lung habitat triggers a conversion of bacterial phenotype, but *P. aeruginosa* retains enough flexibility to recognize its environment of host cells and polymers and to respond to selective pressures such as antimicrobial chemotherapy or co-colonization with other taxa.

REFERENCES

1. Alonso, A., Rojo, F., and Martínez, J.L., 1999, Environmental and clinical isolates of *Pseudomonas aeruginosa* show pathogenic and biodegradative properties irrespective of their origin. *Environ. Microbiol.*, 1:421–443.
2. Anthony, M., Rose, B., Pegler, M.B., Elkins, M., Service, H., Thamotharampillai, K., Watson, J., Robinson, M., Bye, P., Merlino, J., and Harbour, C., 2002, Genetic analysis of *Pseudomonas aeruginosa* isolates from the sputa of Australian adult cystic fibrosis patients. *J. Clin. Microbiol.*, 40:2772–2778.
3. Arevalo-Ferro, C., Buschmann, J., Reil, G., Görg, A., Wiehlmann, L., Tümmler, B., Eberl, L., and Riedel, K., 2004, Proteome analysis of intraclonal diversity of two *Pseudomonas aeruginosa* TB clone isolates. *Proteomics*, 4:1241–1246.
4. Arora, S., Bangera, S., Lory, S., and Ramphal, R., 2001, A genomic island in *Pseudomonas aeruginosa* carries the determinants of flagellin glycosylation. *Proc. Natl. Acad. Sci. USA*, 98:9342–9347.
5. Arora, S.K., Dasgupta, N., Lory, S., and Ramphal, R., 2000, Identification of two distinct types of flagellar cap proteins, FliD, in *Pseudomonas aeruginosa*. *Infect. Immun.*, 68:1474–1479.
6. Arora, S.K., Ritchings, B.W., Almira, E.C., Lory, S., and Ramphal, R., 1998, The *Pseudomonas aeruginosa* flagellar cap protein, FliD, is responsible for mucin adhesion. *Infect. Immun.*, 66:1000–1007.
7. Arora, S.K., Wolfgang, M.C., Lory, S., and Ramphal, R., 2004, Sequence polymorphism in the glycosylation island and flagellins of *Pseudomonas aeruginosa*. *J. Bacteriol.*, 186:2115–2122.
8. Barbieri J.T., 2000, *Pseudomonas aeruginosa* exoenzyme S, a bifunctional type-III secreted cytotoxin. *Int. J. Med. Microbiol.*, 290:381–387.
9. Boucher, J.C., Martinez-Salaza, J., Schurr, M.J., Mudd, M., Yu, H., and Deretic, V., 1996, Two distinct loci affecting conversion to mucoidy in *Pseudomonas aeruginosa* in cystic fibrosis encode homologs of the serine protease HtlA. *J. Bacteriol.*, 178:511–523.
10. Boucher, J.C., Yu, H., Mudd, M.H., and Deretic, V., 1997, Mucoid *Pseudomonas aeruginosa* in cystic fibrosis: characterization of muc mutations in clinical isolates and analysis of clearance in a mouse model of respiratory infection. *Infect. Immun.*, 65:3838–3846.
11. Bragonzi, A., Worlitzsch, D., Tümmler, B., Salerno, P., Pier, G.P., and Döring, G., 2006, Impact of mutational changes in genetic regulatory loci on the expression of *Pseudomonas aeruginosa* alginate in clinical isolates from patients with cystic fibrosis. Submitted.
12. Brimer, C.D., and Montie, T.C., 1998, Cloning and comparison of *fliC* genes and identification of glycosylation in the flagellin of *Pseudomonas aeruginosa* a-type strains. *J. Bacteriol.*, 180:3209–3217.
13. Castric, P.A., and Deal, C.D., 1994, Differentiation of *Pseudomonas aeruginosa* pili based on sequence and B-cell epitope analyses. *Infect. Immun.*, 62:371–376.
14. Curran, B., Jonas, D., Grundmann, H., Pitt, T., and Dowson, C.G., 2004, Development of a multilocus sequence typing scheme for the opportunistic pathogen *Pseudomonas aeruginosa*. *J. Clin. Microbiol.*, 42:5644–5649.
15. De Chial, M., Ghysels, B., Beatson, S.A., Geoffroy, V., Meyer, J.M., Pattery, T., Baysse, C., Chablain, P., Parsons, Y.N., Winstanley, C., Cordwell, S.J., and Cornelis, P., 2003, Identification of type II and type III pyoverdine receptors from *Pseudomonas aeruginosa*. *Microbiology*, 149:821–831.
16. Denervaud, V., TuQuoc, P., Blanc, D., Favre-Bonte, S., Krishnapillai, V., Reimmann, C., Haas, D., and van Delden, C., 2004, Characterization of cell-to-cell signaling-deficient *Pseudomonas aeruginosa* strains colonizing intubated patients. *J. Clin. Microbiol.*, 42:554–562.
17. De Vos, D., De Chial, M., Cochez, C., Jansen, S., Tümmler, B., Meyer, J.M., and Cornelis, P., 2001, Study of pyoverdine type and production by *Pseudomonas aeruginosa* isolated

from cystic fibrosis patients: prevalence of type II pyoverdine isolates and accumulation of pyoverdine-negative mutations. *Arch. Microbiol*., 175:384–388.
18. Dinesh, S.D., Grundmann, H., Pitt, T.L., and Römling, U., 2003, European-wide distribution of *Pseudomonas aeruginosa* clone C. *Clin. Microbiol. Infect*., 9:1228–1233.
19. Eberl, L., and Tümmler, B., 2004, *Pseudomonas aeruginosa* and *Burkholderia cepacia* in cystic fibrosis: genome evolution, interactions and adaptation. *Int. J. Med. Microbiol*., 294:123–131.
20. Ernst, R.K., D'Argenio, D.A., Ichikawa, J.K., Bangera, M.G., Selgrade, S., Burns, J.L., Hiatt, P., McCoy, K., Brittnacher, M., Kas, A., Spencer, D.H., Olson, M.V., Ramsey, B.W., Lory, S., and Miller, S.I., 2003, Genome mosaicism is conserved but not unique in *Pseudomonas aeruginosa* isolates from the airways of young children with cystic fibrosis. *Environ. Microbiol*., 5:1341–1349.
21. Ernst, R.K., Yi, E.C., Guo, L., Lim, K.B., Burns, J.L., Hackett, M., and Miller, S.I., 1999, Specific lipopolysaccharide found in cystic fibrosis airway *Pseudomonas aeruginosa*. *Science*, 286:1561–1565.
22. Feltman, H., Schulert, G., Khan, S., Jain, M., Peterson, L., and Hauser, A.R., 2001, Prevalence of type III secretion genes in clinical and environmental isolates of *Pseudomonas aeruginosa*. *Microbiology*, 147:2659–2669.
23. Foght, J.M., Westlake, D.W.S., Johnson, W.M., and Ridgway, H.F., 1996, Environmental gasoline-utilizing isolates and clinical isolates of *Pseudomonas aeruginosa* are taxonomically indistinguishable by chemotaxonomic and molecular techniques. *Microbiology*, 142:2333–2340.
24. Geisenberger, O., Givskov, M., Riedel, K., Hoiby, N., Tümmler, B., and Eberl, L., 2000, Production of *N*-acyl-L-homoserine lactones by *P. aeruginosa* isolates from chronic lung infections associated with cystic fibrosis. *FEMS Microbiol. Lett*., 184:273–278.
25. Ghysels, B., Dieu, B.T., Beatson, S.A., Pirnay, J.P., Ochsner, U.A., Vasil, M.L., and Cornelis, P., 2004, FpvB, an alternative type I ferripyoverdine receptor of *Pseudomonas aeruginosa*. *Microbiology*, 150:1671–1680.
26. Govan, J.R., and Deretic, V., 1996, Microbial pathogenesis in cystic fibrosis: mucoid *Pseudomonas aeruginosa* and *Burkholderia cepacia*. *Microbiol. Rev*., 60:539–574.
27. Hancock, R.E., Mutharia, L.M., Chan, L., Darveau, R.P., Speert, D.P., and Pier, G.B., 1983, *Pseudomonas aeruginosa* isolates from patients with cystic fibrosis: a class of serum-sensitive, nontypable strains deficient in lipopolysaccharide O side chains. *Infect. Immun*., 42:170–177.
28. Häussler, S., Tümmler, B., Weissbrodt, H., Rohde, M., and Steinmetz, I., 1999, Small-colony variants of *Pseudomonas aeruginosa* in cystic fibrosis. *Clin. Infect. Dis*., 29:621–625.
29. Häussler, S., Ziegler, I., Lottel, A., von Götz, F., Rohde, M., Wehmhöner, D., Saravanamuthu, S., Tümmler, B., and Steinmetz, I., 2003, Highly adherent small-colony variants of *Pseudomonas aeruginosa* in cystic fibrosis lung infection. *J. Med. Microbiol*., 52:295–301.
30. He, J., Baldini, R.L., Deziel, E., Saucier, M., Zhang, Q., Liberati, N.T., Lee, D., Urbach, J., Goodman, H.M., and Rahme, L.G., 2004, The broad host range pathogen *Pseudomonas aeruginosa* strain PA14 carries two pathogenicity islands harboring plant and animal virulence genes. *Proc. Natl. Acad. Sci. USA*, 101:2530–2535.
31. Heuer, T., Bürger, C., Maaß, G., and Tümmler, B., 1998, Cloning of prokaryotic genomes in yeast artificial chromosomes: application to the population genetics of *Pseudomonas aeruginosa*. *Electrophoresis*, 19:486–494.
32. Hocquet, D., Bertrand, X., Köhler, T., Talon, D., and Plesiat, P., 2003, Genetic and phenotypic variations of a resistant *Pseudomonas aeruginosa* epidemic clone. *Antimicrob. Agents Chemother*., 47:1887–1894.
33. Holloway, B.W., Römling, U., and Tümmler, B., 1994, Genomic mapping of *Pseudomonas aeruginosa* PAO. *Microbiology*, 140:2907–2929.
34. Jones, R.N., Deshpande, L., Fritsche, T.R., and Sader H.S., 2004, Determination of epidemic clonality among multidrug-resistant strains of *Acinetobacter* spp. and *Pseudomonas*

aeruginosa in the MYSTIC Programme (USA, 1999–2003). *Diagn. Microbiol. Infect. Dis*., 49:211–216.

35. Kiewitz, C., Larbig, K., Klockgether, J., Weinel, C., and Tümmler, B., 2000, Monitoring genome evolution ex vivo: reversible chromosomal integration of a 106 kb plasmid at two tRNA(Lys) gene loci in sequential *Pseudomonas aeruginosa* airway isolates. *Microbiology*, 146:2365–2373.
36. Kiewitz, C., and Tümmler, B., 2000, Sequence diversity of *Pseudomonas aeruginosa*: impact on population structure and genome evolution. *J. Bacteriol.*, 182:3125–3135.
37. Kiewitz, C., and Tümmler, B., 2002, Similar profile of single nucleotide substitution types in bacteria and human genetic disease. *Genome Lett.*, 1:111–114.
38. Klockgether, J., 2004, Geninseln als Quelle der Genomdiversität von *Pseudomonas aeruginosa*. Dissertation. University of Hannover.
39. Klockgether, J., Reva, O., Larbig, K., and Tümmler, B., 2004, Sequence analysis of the mobile genome island pKLC102 of *Pseudomonas aeruginosa* C. *J. Bacteriol*., 186:518–534.
40. Kresse, A.U., Dinesh, S.D., Larbig, K., and Römling, U., 2003, Impact of large chromosomal inversions on the adaptation and evolution of *Pseudomonas aeruginosa* chronically colonizing cystic fibrosis lungs. *Mol. Microbiol*., 47:145–158.
41. Laabs, U., Gudowius, P., Wiehlmann, L., Limpert, A.S., Filloux, A., Tümmler, B., and de Bentzmann, S., 2006, High binding capacity of *Pseudomonas aeruginosa* cystic fibrosis isolates during chronic infection is associated with worse clinical outcome. (submitted).
42. Lamont, I.L., and Martin, L.W., 2003, Identification and characterization of novel pyoverdine synthesis genes in *Pseudomonas aeruginosa*. *Microbiology*, 149:833–842.
43. Larbig, K.D., Christmann, A., Johann, A., Klockgether, J., Hartsch, T., Merkl, R., Wiehlmann, L., Fritz, H.J., and Tümmler, B., 2002, Gene islands integrated into tRNA(Gly) genes confer genome diversity on a *Pseudomonas aeruginosa* clone. *J. Bacteriol*., 184:6665–6680.
44. Lee, V.T., Smith, R.S., Tümmler, B., and Lory, S., 2005, Activities of *Pseudomonas aeruginosa* effectors secreted by the Type III secretion system in vitro and during infection. *Infect. Immun*., 73:1695–1705.
45. Liang, X., Pham, X.Q., Olson, M.V., and Lory, S., 2001, Identification of a genomic island present in the majority of pathogenic isolates of *Pseudomonas aeruginosa*. *J. Bacteriol.*, 183:843–853.
46. Macia, M.D., Blanquer, D., Togores, B., Sauleda, J., Perez, J.L., and Oliver, A., 2005, Hypermutation is a key factor for multiple antimicrobial resistance development in *Pseudomonas aeruginosa* chronic lung infections. *Antimicrob. Agents Chemother.*, 49: 3382–3386.
47. Mahenthiralingam, E., Campbell, M.E., and Speert, D.P., 1994, Nonmotility and phagocytic resistance of *Pseudomonas aeruginosa* isolates from chronically colonized patients with cystic fibrosis. *Infect. Immun*., 62:596–605.
48. Marrs, C.F., Schoolnik, G., Koomey, J.M., Hardy, J., Rothbard, J., and Falkow, S., 1985, Cloning and sequencing of a *Moraxella bovis* pilin gene. *J. Bacteriol*., 163:132–139.
49. Martin, D.W., Schurr, M.J., Mudd, M.H., Govan, J.R.W., Holloway, B.W., and Deretic, V., 1993, Mechanisms of conversion to mucoidy in *Pseudomonas aeruginosa* infecting cystic fibrosis patients. *Proc. Natl. Acad. Sci. USA*, 90:8377–8381.
50. Miranda, G., Leanos, B., Marquez, L., Valenzuela, A., Silva, J., Carrillo, B., Munoz, O., and Solorzano, F., 2001, Molecular epidemiology of a multiresistant *Pseudomonas aeruginosa* outbreak in a paediatric intensive care unit. *Scand. J. Infect. Dis*., 33:738–743.
51. Mohd-Zain, Z., Turner, S.L., Cerdeno-Tarraga, A.M., Lilley, A.K., Inzana, T.J., Duncan, A.J., Harding, R.M., Hood, D.W., Peto, T.E., and Crook, D.W., 2004, Transferable antibiotic resistance elements in *Haemophilus influenzae* share a common evolutionary origin with a diverse family of syntenic genomic islands. *J. Bacteriol*., 186:8114–8122.
52. Morales, G., Wiehlmann, L., Gudowius, P., van Delden, C., Tümmler, B., Martínez, J.L., and Rojo, F., 2004, Structure of *Pseudomonas aeruginosa* populations analyzed by single nucleotide

polymorphism and pulsed-field gel electrophoresis genotyping. *J. Bacteriol*., 186:4228–4237.

53. Nelson, K.E., Weinel, C., Paulsen, I.T., Dodson, R.J., Hilbert, H., Martins dos Santos, V.A., Fouts, D.E., Gill, S.R., Pop, M., Holmes, M., Brinkac, L., Beanan, M., DeBoy, R.T., Daugherty, S., Kolonay, J., Madupu, R., Nelson, W., White, O., Peterson, J., Khouri, H., Hance, I., Chris Lee, P., Holtzapple, E., Scanlan, D., Tran, K., Moazzez, A., Utterback, T., Rizzo, M., Lee, K., Kosack, D., Moestl, D., Wedler, H., Lauber, J., Stjepandic, D., Hoheisel, J., Straetz, M., Heim, S., Kiewitz, C., Eisen, J.A., Timmis, K.N., Düsterhöft, A., Tümmler, B., and Fraser, C.M., 2002, Complete genome sequence and comparative analysis of the metabolically versatile *Pseudomonas putida* KT2440. *Environ. Microbiol*., 4:799–808.
54. Nouwens, A.S., Willcox, M.D., Walsh, B.J., and Cordwell, S.J., 2002, Proteomic comparison of membrane and extracellular proteins from invasive (PAO1) and cytotoxic (6206) strains of *Pseudomonas aeruginosa*. *Proteomics*, 2:1325–1346.
55. Oliver, A., Baquero, F., and Blazquez, J., 2002, The mismatch repair system (*mutS, mutL* and *uvrD* genes) in *Pseudomonas aeruginosa*: molecular characterization of naturally occurring mutants. *Mol. Microbiol*., 43:1641–1650.
56. Oliver, A., Canton, R., Campo, P., Baquero, F., and Blaźquez, J., 2000, High frequency of hypermutable *Pseudomonas aeruginosa* in cystic fibrosis lung infection. *Science*, 288:1251–1254.
57. Pedersen, S.S., Hoiby, N., Espersen, F., and Koch, C., 1992, Role of alginate in infection with mucoid *Pseudomonas aeruginosa* in cystic fibrosis. *Thorax*, 47:6–13.
58. Pellegrino, F.L., Teixeira, L.M., Carvalho, Md., Mda, G., Aranha Nouer, S., Pinto De Oliveira, M., Mello Sampaio, J.L., D'Avila Freitas, A., Ferreira, A.L., Amorim, Ed., Ede, L., Riley, L.W., and Moreira, B.M., 2002, Occurrence of a multidrug-resistant *Pseudomonas aeruginosa* clone in different hospitals in Rio de Janeiro, Brazil. *J. Clin. Microbiol*., 40:2420–2424.
59. Pirnay, J.P., De Vos, D., Cochez, C., Bilocq, F., Pirson, J., Struelens, M., Duinslaeger, L., Cornelis, P., Zizi, M., and Vanderkelen, A., 2003, Molecular epidemiology of *Pseudomonas aeruginosa* colonization in a burn unit: persistence of a multidrug-resistant clone and a silver sulfadiazine-resistant clone. *J. Clin. Microbiol*., 41:1192–1202.
60. Pirnay, J. P., De Vos, D., Cochez, C., Bilocq, F., Vanderkelen, A., Zizi, M., Ghysels, B., and Cornelis, P., 2002, *Pseudomonas aeruginosa* displays an epidemic population structure. *Environ. Microbiol.*, 4:898–911.
61. Pirnay, J.P., De Vos, D., Mossialos, D., Vanderkelen, A., Cornelis, P., and Zizi, M., 2002, Analysis of the *Pseudomonas aeruginosa oprD* gene from clinical and environmental isolates. *Environ. Microbiol*., 4:872–882.
62. Ramphal, R., and Arora, S.K., 2001, Recognition of mucin components by *Pseudomonas aeruginosa. Glycoconj. J.*, 18:709–713.
63. Ravel, J., and Cornelis, P., 2003, Genomics of pyoverdine-mediated iron uptake in pseudomonads. *Trends Microbiol.*, 11:195–200.
64. Raymond, C.K., Sims, E.H., Kas, A., Spencer, D.H., Kutyavin, T.V., Ivey, R.G., Zhou, Y., Kaul, R., Clendenning, J.B., and Olson, M.V., 2002, Genetic variation at the O-antigen biosynthetic locus in *Pseudomonas aeruginosa*. *J. Bacteriol*., 184:3614–3622.
65. Römling, U., Fiedler, B., Bosshammer, J., Grothues, D., Greipel, J., von der Hardt, H., and Tümmler, B., 1994, Epidemiology of chronic *Pseudomonas aeruginosa* infections in cystic fibrosis. *J. Infect. Dis*., 170:1616–1621.
66. Römling, U., Greipel, J., and Tümmler, B., 1995, Gradient of genomic diversity in the *Pseudomonas aeruginosa* chromosome. *Mol. Microbiol*., 17:323–332.
67. Römling, U., Schmidt, K.D., and Tümmler, B., 1997, Large genome rearrangements discovered by the detailed analysis of 21 *Pseudomonas aeruginosa* clone C isolates found in environment and disease habitats. *J. Mol. Biol*., 271:386–404.
68. Römling, U., Wingender, J., Müller, H., and Tümmler, B., 1994, A major *Pseudomonas aeruginosa* clone common to patients and aquatic habitats. *Appl. Environ. Microbiol*., 60:1734–1738.

69. Salunkhe, P., Götz, F., Wiehlmann, L., Lauber, J., Buer, J., and Tümmler B., 2002, GeneChip expression analysis of the response of *Pseudomonas aeruginosa* to paraquat-induced superoxide stress. *Genome Lett*., 1:165–174.
70. Salunkhe, P., Töpfer, T., Buer, J., and Tümmler, B., 2005, Genome-wide transcriptional profiling of the steady state response of Pseudomonas aeruginosa to hydrogen peroxide. J. Bacteriol., 187: 2565–2572.
71. Sato, H., Frank, D.W., Hillard, C.J., Feix, J.B., Pankhaniya, R.R., Moriyama, K., Finck-Barbancon, V., Buchaklian, A., Lei, M., Long, R.M., Wiener-Kronish, J., and Sawa, T., 2003, The mechanism of action of the *Pseudomonas aeruginosa*-encoded type III cytotoxin, ExoU. *EMBO J.*, 22:2959–2969.
72. Scharfman, A., Arora, S.K., Delmotte, P., Van Brussel, E., Mazurier, J., Ramphal, R., and Roussel, P., 2001, Recognition of Lewis x derivatives present on mucins by flagellar components of *Pseudomonas aeruginosa. Infect. Immun*., 69:5243–5248.
73. Schmidt, K.D., Tümmler, B., and Römling, U., 1996, Comparative genome mapping of *Pseudomonas aeruginosa* PAO with *P. aeruginosa* C, which belongs to a major clone in cystic fibrosis patients and aquatic habitats. *J. Bacteriol*., 178:85–93.
74. Schurr, M.J., Martin, D.W., Mudd, M.H., and Deretic, V., 1994, Gene cluster controlling conversion to alginate-overproducing phenotype in *Pseudomonas aeruginosa*: functional analysis in a heterologous host and role in the instability of mucoidy. *J. Bacteriol*., 176:3375–3382.
75. Schurr, M.J., Yu, H., Martinez-Salazar, J.M., Boucher, J.C., and Deretic, V., 1996, Control of AlgU, a member of the σ^E-like family of stress sigma factors, by the negative regulators MucA and MucB and *Pseudomonas aeruginosa* conversion to mucoidy in cystic fibrosis. *J. Bacteriol*., 178:4997–5004.
76. Scott, F.W., and Pitt, T.L., 2004, Identification and characterization of transmissible *Pseudomonas aeruginosa* strains in cystic fibrosis patients in England and Wales. *J. Med. Microbiol*., 53:609–615.
77. Selander, R.K., and Musser, J.M., 1990, Population genetics of bacterial pathogenesis, pp. 11–36. In B.H. Iglewski, and V.L. Clark (eds.), *Molecular Basis of Bacterial Pathogenesis,* Vol. X. The Bacteria. Academic Press, San Diego, CA.
78. Simpson, A.J., Reinach, F.C., Arruda, P., Abreu, F.A., Acencio, M., Alvarenga, R., Alves, L.M., Araya, J.E., Baia, G.S., Baptista, C.S., Barros, M.H., Bonaccorsi, E.D., Bordin, S., Bove, J.M., Briones, M.R., Bueno, M.R., Camargo, A.A., Camargo, L.E., Carraro, D.M., Carrer, H., Colauto, N.B., Colombo, C, Costam F.F., Costa, M.C., Costa-Neto, C.M., Coutinho, L.L., Cristofani, M., Dias-Neto, E., Docena, C., El-Dorry, H., Facincani, A.P., Ferreira, A.J., Ferreira, V.C., Ferro, J.A., Fraga, J.S., Franca, S.C., Franco, M.C., Frohme, M., Furlan, L.R., Garnier, M., Goldman, G.H., Goldman, M.H., Gomes, S.L., Gruber, A., Ho, P.L., Hoheisel, J.D., Junqueira, M.L., Kemper, E.L., Kitajima, J.P., Krieger, J.E., Kuramae, E.E., Laigret, F., Lambais, M.R., Leite, L.C., Lemos, E.G., Lemos, M.V., Lopes, S.A., Lopes, C.R., Machado, J.A., Machado, M.A., Madeira, A.M., Madeira, H.M., Marino, C.L., Marques, M.V., Martins, E.A., Martins, E.M., Matsukuma, A.Y., Menck, C.F., Miracca, E.C., Miyaki, C.Y., Monteriro-Vitorello, C.B., Moon, D.H., Nagai, M.A., Nascimento, A.L., Netto, L.E., Nhani, A. Jr., Nobrega, F.G., Nunes, L.R., Oliveira, M.A., de Oliveira, M.C., de Oliveira, R.C., Palmieri, D.A., Paris, A., Peixoto, B.R., Pereira, G.A., Pereira, H.A. Jr., Pesquero, J.B., Quaggio, R.B., Roberto, P.G., Rodrigues, V., de M Rosa, A.J., de Rosa, V.E. Jr., de Sa, R.G., Santelli, R.V., Sawasaki, H.E., da Silva, A.C., da Silva, A.M., da Silva, F.R., da Silva, W.A. Jr., da Silveira, J.F., Silvestri, M.L., Siqueira, W.J., de Souza, A.A., de Souza, A.P., Terenzi, M.F., Truffi, D., Tsai, S.M., Tsuhako, M.H., Vallada, H., Van Sluys, M.A., Verjovski-Almeida, S., Vettore, A.L., Zago, M.A., Zatz, M., Meidanis, J., and Setubal, J.C., 2000, The genome sequence of the plant pathogen *Xylella fastidiosa*. The *Xylella fastidiosa* consortium of the organization for nucleotide sequencing and analysis. *Nature*, 406:151–157.

79. Smith, E.E., Sims, E.H., Spencer, D.H., Kaul, R., and Olson, M.V., 2005, Evidence for diversifying selection at the pyoverdine locus of *Pseudomonas aeruginosa*. *J. Bacteriol.*, 187:2138–2147.
80. Smith, J.M., Smith, N.H., O'Rourke, M., and Spratt, B.G., 1993, How clonal are bacteria? *Proc. Natl. Acad. Sci. USA*, 90:4384–4388.
81. Spangenberg, C., Fislage, R., Römling, U., and Tümmler B., 1997, Disrespectful type IV pilins. *Mol. Microbiol.*, 25:203–204.
82. Spangenberg, C., Heuer, T., Bürger, C., and Tümmler, B., 1996, Genetic diversity of flagellins of *Pseudomonas aeruginosa*. *FEBS Lett*., 396:213–217.
83. Spangenberg, C., Montie, T.C., and Tümmler, B., 1998, Structural and functional implications of sequence diversity of *Pseudomonas aeruginosa* genes *oriC, ampC* and *fliC*. *Electrophoresis*, 19:545–550.
84. Spencer, D.H., Kas, A., Smith, E.E., Raymond, C.K., Sims, E.H., Hastings, M., Burns, J.L., Kaul, R., and Olson M.V., 2003, Whole-genome sequence variation among multiple isolates of *Pseudomonas aeruginosa*. *J. Bacteriol*., 185:1316–1325.
85. Stover, C.K., Pham, X.Q., Erwin, A.L., Mizoguchi, S.D., Warrener, P., Hickey, M.J., Brinkman, F.S., Hufnagle, W.O., Kowalik, D.J., Lagrou, M., Garber, R.L., Goltry, L., Tolentino, E., Westbrock-Wadman, S., Yuan, Y., Brody, L.L., Coulter, S.N., Folger, K.R., Kas, A., Larbig, K., Lim, R., Smith, K., Spencer, D., Wong, GK., Wu, Z., Paulsen, I.T., Reizer, J., Saier, M.H., Hancock, R.E., Lory, S., and Olson, M.V., 2000, Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen. *Nature*, 406:959–964.
86. Taylor, R.F., Hodson, M.E., and Pitt, T.L., 1992, Auxotrophy of *Pseudomonas aeruginosa* in cystic fibrosis. *FEMS Microbiol. Lett*., 71:243–246.
87. Thomas, S.R., Ray, A., Hodson, M.E., and Pitt, T.L., 2000, Increased sputum amino acid concentrations and auxotrophy of *Pseudomonas aeruginosa* in severe cystic fibrosis lung disease. *Thorax*, 55:795–797.
88. Tümmler, B., Bosshammer, J., Breitenstein, S., Brockhausen, I., Gudowius, P., Herrmann, C., Herrmann, S., Heuer, T., Kubesch, P., Mekus, F., Römling, U., Schmidt, K.D., Spangenberg, C., and Walter, S., 1997, Infections with *Pseudomonas aeruginosa* in patients with cystic fibrosis. *Behring Inst. Mitt.*, 98:249–255.
89. Tümmler, B., and Kiewitz C., 1999, Cystic fibrosis: an inherited susceptibility to bacterial respiratory infections. *Mol. Med. Today*, 5:351–358.
90. Vallis, A.J., Finck-Barbancon, V., Yahr, T.L., and Frank, D.W., 1999, Biological effects of *Pseudomonas aeruginosa* type III-secreted proteins on CHO cells. *Infect. Immun*., 67:2040–2044.
91. von Götz, F., Häussler, S., Jordan, D., Saravanamuthu, S.S., Wehmhöner, D., Strüssmann, A., Lauber, J., Attree, I., Buer, J., Tümmler, B., and Steinmetz, I., 2004, Expression analysis of a highly adherent and cytotoxic small colony variant of *Pseudomonas aeruginosa* isolated from a lung of a patient with cystic fibrosis. *J. Bacteriol*., 186:3837–3847.
92. Wehmhöner, D., Häussler, S., Tümmler, B., Jansch, L., Bredenbruch, F., Wehland, J., and Steinmetz, I., 2003, Inter- and intraclonal diversity of the *Pseudomonas aeruginosa* proteome manifests within the secretome. *J. Bacteriol*., 185:5807–5814.
93. West, S.E., and Iglewski B.H., 1988, Codon usage in *Pseudomonas aeruginosa*. *Nucleic Acids Res*., 16:9323–9335.
94. Winstanley, C., Coulson, M.A., Wepner, B., Morgan, J.A., and Hart, C.A., 1996, Flagellin gene and protein variation amongst clinical isolates of *Pseudomonas aeruginosa*. *Microbiology*, 142:2145–2151.
95. Wolfgang, M.C., Kulasekara, B.R., Liang, X., Boyd, D., Wu, K., Yang, Q., Miyada, C.G., and Lory S., 2003, Conservation of genome content and virulence determinants among clinical and environmental isolates of *Pseudomonas aeruginosa*. *Proc. Natl. Acad. Sci. USA*, 100:8484–8489.

3

PSEUDOMONAS AERUGINOSA PHOSPHOLIPASES AND PHOSPHOLIPIDS

Michael L. Vasil

University of Colorado at Denver and Health Sciences Center
Aurora, CO 80045, USA

The diversity of roles that phospholipases play in biology and medicine is exceptional. In the past decade, this class of enzymes has proven to be considerably more complex than initially perceived and their impact on an assortment of basic cellular processes in eukaryotes, including oncogenesis and inflammation has become widely appreciated. Likewise, there are sundry functions for phospholipases in prokaryotic biology, including their noteworthy contributions to microbial virulence. For example, individual members of a homologous class of phospholipases C (PLCs), usually produced by gram-positive bacteria (GP-PLCs), serve vastly different functions in pathogenesis. One member of this class is an extremely potent extracellular toxin (e.g. *Clostridium perfringens* α toxin), while another contributes to the intricate mechanisms of the intracellular and intercellular trafficking in a facultative intracellular pathogen (e.g. PlcA and PlcB of *Listeria monocytogenes*).

There are several major distinct classes of phospholipases (Figure 1, Table 1). Even within a specific class (e.g. PLCs), there are subclassifications based on biochemical parameters (e.g. need for metal cofactors such as zinc) and differences in substrate preference (phosphatidylcholine vs. phosphatidylinositol). These differences can ultimately have a profound effect on whether a particular phospholipase will or will not have an impact in a particular environment, such as different locations in an infected host. In addition, it is important to note that the products created by phospholipases play a fundamental

Pseudomonas, Volume 4, edited by Juan-Luis Ramos and Roger C. Levesque

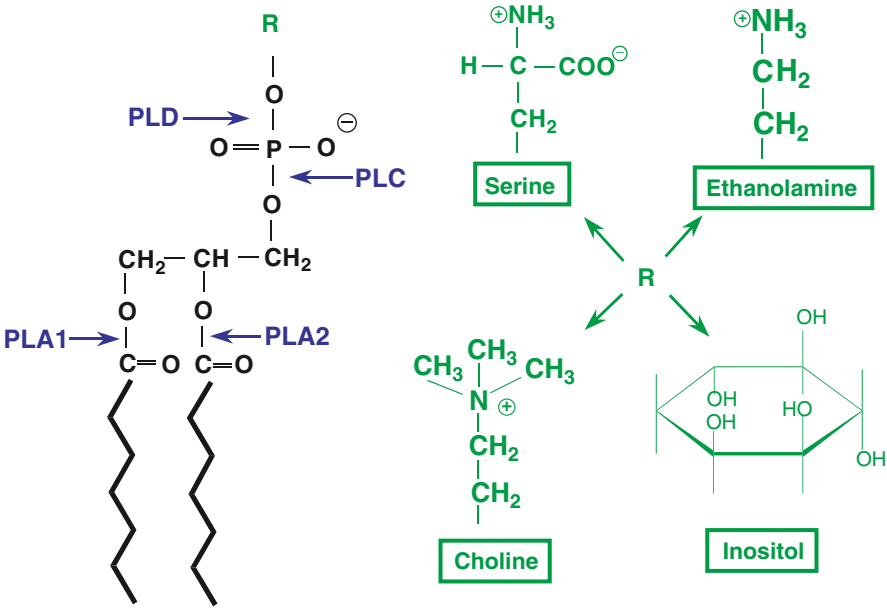


Figure 1. Structure of a typical glycerophospholipid (diacyl) with the cleavage sites for phospholipase C (PLC), phospholipase D (PLD), phospholipase A1 (PLA1), and phospholipase A2 (PLA2) designated by arrows. The structures of some of the possible head groups substituting for R are shown.

role in the biological effects ultimately attributed to phospholipases (Figure 2). Diacylglycerol (DAG) generated by a PLC can induce proliferative effects in eukaryotic cells, while ceramide generated by a sphingomyelinase (SMase) can cause cell death through apoptosis (Figure 2).[20,21] In some cases, the same protein can have both PLC and SMase activity (see Section 4.1 below).[41,76]

1. PHOSPHOLIPASE C

A PLC was the first enzymatic activity to be associated with the mode of action of a bacterial toxin. In 1941, Macfarlane and Knight[42] demonstrated that the highly cytotoxic α toxin of *C. perfringens* has PLC activity. Since then, PLC activity has been demonstrated in a variety of other pathogenic bacteria, including *Staphylococcus aureus,*[44] *Legionella pneumophila,*[3] *Helicobacter pylori,*[6] *Mycoplasma* spp.,[12] *L. monocytogenes*,[7,45] *Mycobacterium tuberculosis*,[30] *Francisella tularensis,*[62] *Burkholderia pseudomallei*,[35] and *Pseudomonas aeruginosa*.[5,57,76] The biochemistry of some bacterial PLCs, particularly the α toxin of *C. perfringens*, was studied in detail early on because they were

Table 1. Known phospholipases of *Pseudomonas aeruginosa*.

Type	Gene designation[a]	MW mature protein (kDa)	Secretory pathway		Chaperone	Substrate specificity
			CM	OM		
Phospholipase C	*plcH* (PA0844)	78.4	TAT	Xcp	PlcR	PC, SM
Sphingomyel in synthase	*plcH* (PA0844)	78.4	TAT	Xcp	PlcR	*erythro*-ceramide and monoacyl or diacyl phosphatidyl choline
Phospholipase C	*plcN* (PA3319)	73.5	TAT	Xcp	?	PC, PS
Phospholipase C	*plcB* (PA0026)	34.5	Sec	Xcp	PpiP (PA0027) and PA0028	PC, PE, PS, SM
Phospholipase C	*plcA* (PA3464)	46.6	Sec	?	?	PC, PS, PE
Phospholipase A	*exoU*	74	Type III		SpcU	LysoPC, PC (Others ?)
Phospholipase D	*pldA* (PA3487)	116	Unknown – periplasmic		?	PC (Others ?)

[a] Annotation number given in parentheses (http//www.pseudomonas.com). *exoU* and *spcU* are not present in sequenced strain PAO1; CM, cytoplasmic membrane; OM, outer membrane.

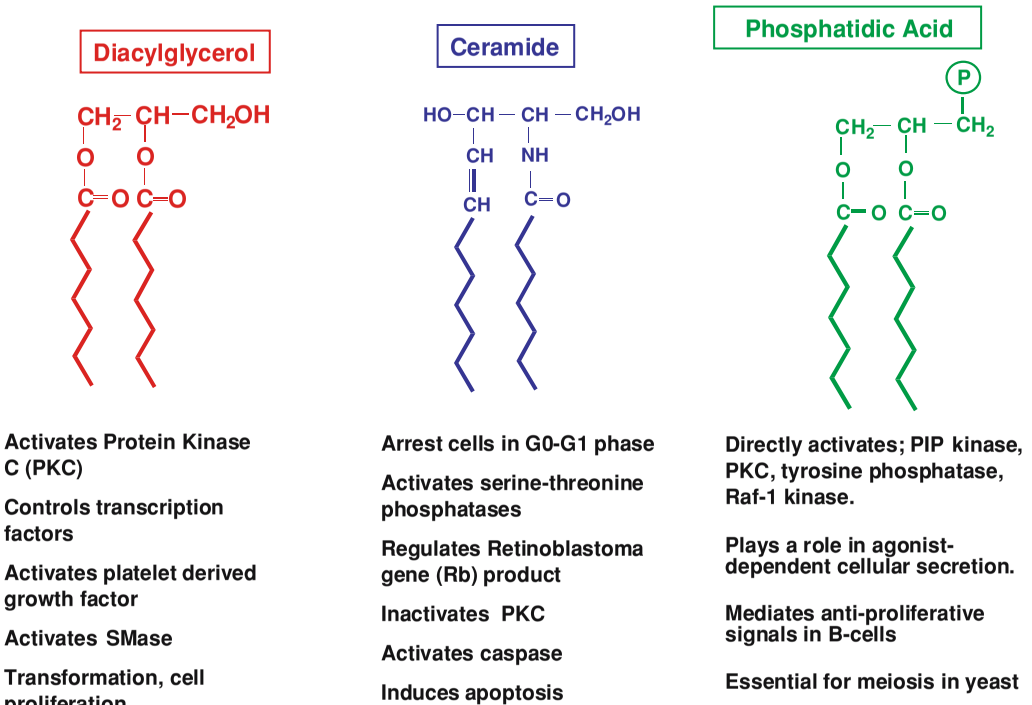


Figure 2. Eukaryotic signaling effects from the products of phospholipases described in this chapter. Diacylglycerol is generated by PLC, Ceramide is generated by SMase, and Phosphatidic acid is generated by PLD.

useful tools for membranologists to probe the composition and organization of membrane phospholipids. In contrast, the role of PLCs in bacterial pathogenesis was largely ignored until a little over a decade ago there was a resurgence of interest in the potential role of PLC in prokaryotic pathogenesis. Camilli *et al*.,[7] Geoffroy *et al*.[17], and Marquis *et al*.[45] demonstrated that two PLCs, a phosphatidylinositol-specific PLC (PI-PLC) and a phosphatidylcholine-specific PLC (PC-PLC) of *L. monocytogenes* (PlcA and PlcB, respectively) contribute to the cell-to-cell spread of this opportunistic pathogen. Marquis *et al*.[45] also provided evidence that activation of the PC-PLC precursor (proPC-PLC) is specific for certain compartments of eukaryotic cells and that this is controlled by a combination of bacterial and host factors. Wadsworth and Goldfine[87] provided evidence that the DAG, generated by the *L. monocytogenes* PC-PLC, along with Listeriolysin O, resulted in increased intracellular Ca^{++} mediated by the translocation of protein kinase C (PKC) delta in murine macrophage. Contemporary studies of the *C. perfringens* α toxin provided a more detailed view of the molecular architecture of the GP-PLCs, as well as provided a better appreciation of their diverse functions.[9] They also afforded a fresh perspective about the role of α toxin in the pathogenesis of clostridial myonecrosis (i.e. gas gangrene). More specifically the α toxin, in synergy with the pore-forming toxin, perfringolysin O (PFO), is required for necrosis of muscle tissue, inhibition of the influx of polymorphonuclear leukocytes (PMN) into the tissue lesions, and thrombosis formation, all of which are pathognomonic of this severe condition.[2] These data exemplify how a single class of zinc-dependent PLCs produced by distinct gram-positive pathogens can serve vastly dissimilar roles in pathogenic processes.

Until the discovery of the PLCs of *P. aeruginosa* (see below), the GP-PLCs were the only bacterial PLCs that were well characterized, particularly with regard to their possible role in virulence. In the 1960s, Liu[39] identified an extracellular PLC of *P. aeruginosa* and designated it the "Heat-labile Hemolysin" because heat inactivation of the PLC in extracellular fractions, simultaneously lead to inactivation of their hemolytic activity on human and sheep erythrocytes. It was not until nearly two decades later that a gene encoding a specific protein (hemolytic phospholipase C – PlcH) with both hemolytic and PLC activities was identified.[60,84] Once the sequence of PlcH was determined, it became clear that it did not belong to any previously known class of PLCs, including the GP-PLCs.[60] Although, it hydrolyses both phosphatidylcholine (PC) and sphingomyelin (SM), as do the PLCs from *C. perfringens* and *L. monocytogenes* basically, similarity ends there. The size of PlcH (78 kDa) is nearly twice as large as the largest previously known bacterial PLC (48 kDa) (i.e. *C. perfringens* α toxin). It has no requirement for zinc as a cofactor, as do GP-PLCs, nor is its sequence in any way even remotely similar to the GP-PLCs. PlcH consequently became the founding member of a novel class of prokaryotic PLCs.[76] PlcH and another member (PlcN) of this new class of enzymes produced by *P. aeruginosa*

will be described in detail below. What is more, we have now identified two additional, previously unrecognized, PLCs of *P. aeruginosa* that belong to the GP-PLCs class of PLCs.[5,77] These will also be discussed below.

2. PHOSPHOLIPASE D

Phospholipases D (PLDs) are virtually ubiquitous in both plants and animals, but they are relatively scarce in prokaryotes.[78] Most of the well-studied PLDs specifically hydrolyse PC to generate choline and phosphatidic acid (PA) (Figure 2), although phosphatidylinositol-specific PLDs have also been identified. PLDs, as well as their products (e.g. PA), have become the focus of intensive efforts to understand their wide-ranging functions in eukaryotic cell biology. Eukaryotic PLDs participate in such assorted processes as meiosis in yeast, phagocytosis, and intracellular killing of *M. tuberculosis* by macrophage and signal transduction in response to lipopolysaccharide (LPS).[48] More recently, de Torres Zabela *et al*.[13] reported that *Arabidopsis* PLDs are significantly upregulated in response to *Pseudomonas. syringae*, leading to a hypersensitive response in this plant. When PLDs have been found in bacteria, including *P. aeruginosa,* they often play a role in virulence. The most notable of these is the PLD of *Yersinia pestis*. While it was initially called the "murine toxin" of plague and thought to be responsible for the death of certain mammalian hosts (e.g. rats), more recent studies directly implicated this PLD in, by far, more intricate pathogenic processes. First, contrary to its early designation as an extracellular toxin, in reality it is strictly a cytoplasmic protein of *Y. pestis*. Moreover, it does not participate in virulence in the mammalian host but is, in fact, required for the survival of *Y. pestis* in the midgut of the insect vector (i.e. flea).[25,26] With regard to other prokaryotic PLDs, Wilderman *et al*.[89] recently identified and extensively characterized a PLD from *P. aeruginosa* that is homologous to the mammalian subclass of the PLD Superfamily of enzymes. Its genetics, biochemistry, and role in virulence will be further discussed below.

3. PHOSPHOLIPASE A

This class of phospholipases constitutes the most diverse group of phospholipid modifying enzymes and accordingly, display a remarkable range of biological functions. Phospholipases A (PLAs) have been mostly studied in terms of their roles in the generation of powerful lipid signaling molecules such as arachidonic acid (Figure 3).[1,75] However, more recently they have been progressively better appreciated for their roles in intracellular membrane trafficking events and their ability to directly affect the structure and function

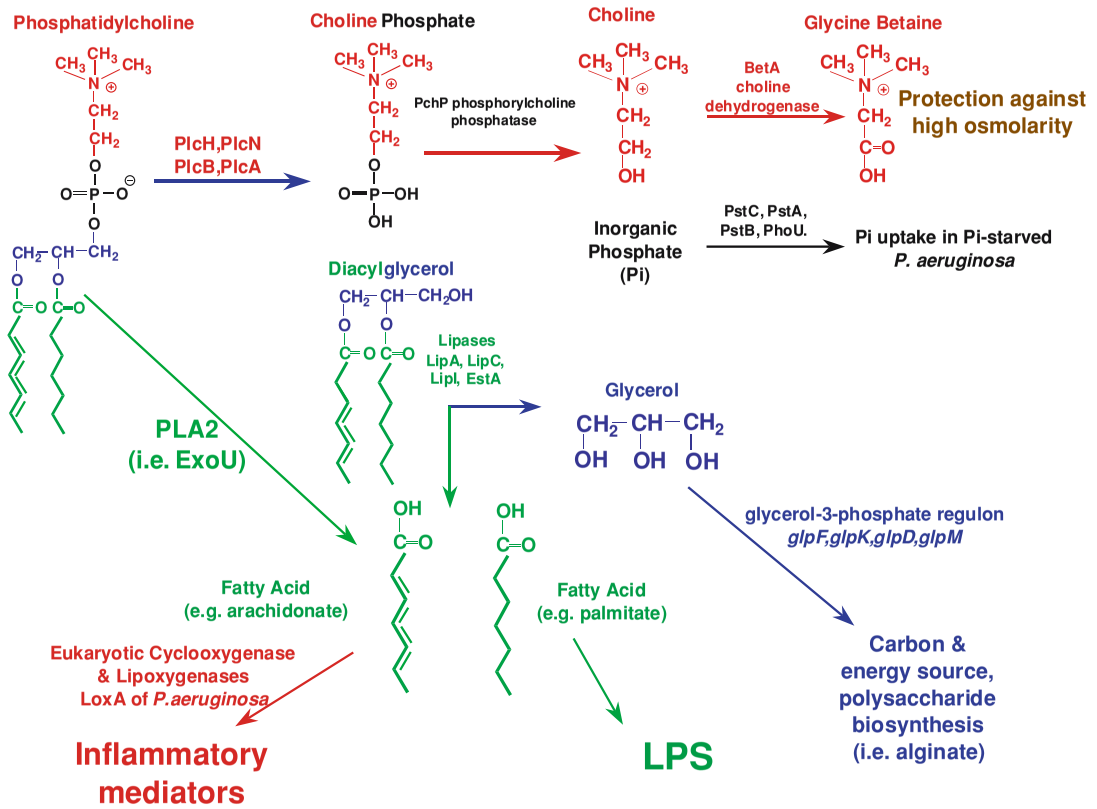


Figure 3. Potential products and their functions that can be generated by known enzymes of *Pseudomonas aeruginosa* from phosphatidylcholine starting with any of its four PLCs or ExoU (PLA2).

of eukaryotic membranes.[75] While there are two major subclasses of PLAs (i.e. PLA1 and PLA2; see Figure 1), PLA2s are by far the most widely and thoroughly characterized group. Belonging to the PLA2 class, there is also a newly classified group of proteins, now known as "patatins," some of which are known to have PLA activity.[27] They were initially identified in plants (i.e. potatoes), but proteins with homologous regions (i.e. patatin domains) are also found in animals and bacteria. PLAs had been previously associated with virulence in *H. pylori*,[36] *L pneumophila,*[16] *Yersinia enterocolitica*,[67] *Rickettsia prowazekii*,[91] and *Streptococcus pyogenes,*[51] but no protein with PLA activity had been formerly associated with *P. aeruginosa.* In the past few years, however, a domain of a well-established virulence factor of *P. aeruginosa* (ExoU) was found to align with patatin domains and with regions of human cytosolic and calcium independent PLA2s.[61,66,79] Subsequently, two independent groups confirmed that ExoU, which is secreted into the cytosol of eukaryotic cells via the Type III secretory system, has PLA activity.[58,66] This novel toxin-enzyme and its potential role in the pathogenesis of *P. aeruginosa* infections will be examined below.

4. PLCS OF *P. AERUGINOSA*

Until quite recently, there were only two known phospholipases of this opportunist, PlcH and PlcN.[76] The sequences of both are notably similar and both belong to a growing number of homologous proteins, some with known PLC activity, most frequently found in prokaryotes, which tend to have high G+C content (e.g. *Pseudomonas* spp. *Mycobacterium* spp. *Streptomyces spp. Burkholderia* spp. *Bordetella* spp., *Xanthamonas* spp. *Caulobacter* spp.).[76] Some of these include PLCs that are known virulence determinants. Proteins belonging to this novel class of PLCs have been reported to exist in fungi (e.g. *Aspergillus*) and plants (e.g. *Arabidopsis*), as well.[94] As stated earlier, these PLCs constitute a novel class of PLCs. They are not similar at the sequence or biochemical level to the GP-PLCs described above. Yet, a couple of years ago, we detected an unanticipated PLC activity in culture supernatants of a mutant with the inability to secrete either PlcH or PlcN (i.e. a twin arginine translocase mutant—TAT). Further analysis of this mutant (ΔTatC) and a mutant with *plcH* and *plcN* deleted (ΔPlcHΔPlcN) led to the discovery of a third PLC of *P. aeruginosa* that belongs to the GP-PLC class of PLCs described above.[5] The nuances relating of the identification of the gene encoding this third PLC of *P. aeruginosa* are described in greater detail elsewhere.[5] This new PLC (PlcB) has sequence similarity, especially in the region of the active site, to the GP-PLC class and like the GP-PLCs its activity is zinc dependent. Subsequently, within the past year we discovered a fourth PLC that shares limited similarity in the active site region to PlcB and the GP-PLC class.[77] We designated this PLC,

PlcA. The genetics, biochemistry, and biology of these four PLCs are reviewed below.

4.1. PlcH

Early on, Liu[39] reported that the hemolytic phospholipase (i.e. PlcH) is preferentially expressed under phosphate (Pi)-limiting conditions. This has led more than a few authors to surmise that the major purpose of PlcH, along with phosphatases, is to scavenge Pi by degrading host phospholipids (Figure 3). While this may indeed be of benefit to *P. aeruginosa* in a Pi-limiting environment, it is a very limited perspective, *vis-à-vis* the real potential of this virulence determinant. The Pi-starvation inducible expression of PlcH was verified at the transcriptional level and it was determined that PlcH is expressed from a three-gene operon, which contains two overlapping genes encoding calcium-binding proteins (PlcR1 and PlcR2) that are chaperones for PlcH.[10,60,76] However, further analyses revealed that Pi-starvation is not the sole condition under which the PlcHR operon is expressed. Shortridge *et al*.[74] found that the PlcHR operon is still expressed even under Pi sufficient conditions, provided choline is present in the environment (Table 2). Actually, there are two promoters for the PlcHR operon, one is a Pi-starvation inducible promoter and the other is induced by choline and independent of Pi-starvation. These and other data led to a new hypothesis that is parallel to the one in which Pi scavenging is the sole objective of PlcH. We put forward a scenario whereby this virulence determinant, through its ability to generate another moiety from phospholipids (choline phosphate), would provide much more than just a single ion (i.e. Pi) for the survival of *P. aeruginosa*. First, *P. aeruginosa* can utilize choline phosphate as a sole source of carbon, nitrogen and Pi (Figure 3).[72] Consequently, it would be possible for this organism to survive solely on phospholipids for these needs if it is expressing PlcH or other PLCs. A further benefit of choline phosphate is that it will enable *P. aeruginosa* to survive and grow in high osmolarity environments.[74,85] Choline, once it is generated from choline phosphate by a phosphatase, in *P. aeruginosa,* as well as in other bacteria, can be readily taken up and converted to the osmolyte, glycine betaine (Figure 3). Under conditions of high osmolarity (e.g. lungs of CF patients) the ability of *P. aeruginosa* to generate choline, and convert it to glycine betaine, through the hydrolysis of phospholipids could allow it survive in this hostile environment.[95] Other organisms that do not produce PLCs would be unable to survive in such a harsh milieu where phospholipids might be the only source of choline. It is worth mentioning that a significant fraction of the phospholipids in mammalian membranes are choline phospholipids. Furthermore, only PlcH, but not any other *P. aeruginosa* PLCs, has a strict substrate requirement for choline containing phospholipids, principally PC and SM. Based on these data, it is more than likely that the function of PlcH

Table 2. Genetics and role in virulence.

Name	Incidence (%)	Genetic regulation	Cytotoxic properties	Virulence in model infections	Putative function
PlcH	100	Pi-starvation inducible, choline inducible	Hemolytic; highly toxic to endothelial cells (pM), moderate toxicity to macrophage (nM), weakly cytotoxic (μM) to epithelial cells or fibroblasts	Chronic pulmonary – rats; thermal injury – mice, acute; pulmonary – rabbits; endocarditis – rabbits	Nutrient acquisition; osmoprotection; eukaryotic lipid signaling, cytotoxin, synthesis of sphingomyelin
PlcB	100	Pi-starvation, Vfr, adenyl cyclase (CyaAB), quorum sensing	Unknown	NT	Chemotaxis toward phospholipids
ExoU	~30 (associated with PAI-like element)	Low calcium	Broad cytotoxicity to eukaryotic cells including yeast; requires eukaryotic cytosolic factor for cytotoxicity	Acute pneumonia – mice	Overt cytotoxicity; dissemination
PldA	~30 (associated with mobile element *vgr*)	Unknown	Unknown	Chronic pulmonary – rats	Prokaryotic lipid signaling

NT, not tested; Incidence, frequency gene is found in *P. aeruginosa*.

is not merely Pi scavenging; it also affords *P. aeruginosa* with other nutrients (C and N), as well as an osmoprotectant (choline).

Although the above scenario presents a more expansive view of the function of PlcH, it is likewise probably too limited. *P. aeruginosa* also produces several extracellular lipases that can remove the fatty acids from DAG, but they cannot attack the fatty acid ester bonds when they are present in phospholipids (Figure 3). Consequently, once PlcH removes the choline phosphate head group from PC, free fatty acids can now be generated by the extracellular lipases of *P. aeruginosa*.[63] Depending on the fatty acids that would be present on a particular kind of PC (e.g. arachidonyl, palmitoyl), there could be a vastly distinct biological outcomes from the action of PlcH and the extracellular DAG-lipases of *P. aeruginosa*. The fatty acids so generated by the action of these enzymes can be incorporated into its LPS (e.g. palmitoyl) or they (e.g. arachidonyl) may be further modified, into extremely potent inflammatory mediators, usually by eukaryotic enzymes (Figure 3).[14,82] However, Vance *et al*.[83] recently identified and characterized an extracellular lipoxygenase of *P. aeruginosa* that converts arachidonic acid into 15-hydroxyeicosatetraenoic acid (15-HETE), which has regulatory effects on immune and nonimmune cells. Finally, the glycerol moiety that would be left from the action of PlcH and DAG-lipases could be shunted into energy production or polysaccharide biosynthesis (e.g. alginate) (Figure 3)[70,71]. Admittedly such scenarios are hypothetical, but they are supported by experimental data. Thus, it is exceedingly probable that the function of PlcH is, by far, more substantial than its ability to merely increase the availability Pi.

Even though the above examples provide a compelling rationale that PlcH ultimately enables *P. aeruginosa* to utilize the entire phospholipid rather than just certain moieties (e.g. Pi), there are other properties of PlcH, which argue that its contribution to virulence is even more profound than nutrient acquisition. As noted above, PlcH is hemolytic for human erythrocytes, but not all mammalian red blood cells. Furthermore, some but not all, PLCs are hemolytic.[85] For example, even the highly similar homolog of PlcH (i.e. PlcN) is not hemolytic for human erythrocytes, nor is the PlcB of *P. aeruginosa* despite the fact that all of these PLCs are able to hydrolyze, at least, some of the phospholipids in the membranes of these cells. More recently, we obtained data indicating that, not only is PlcH cytotoxic to live eukaryotic cells, it is selectively toxic (i.e. highly toxic to some kinds of cells and minimally toxic or nontoxic to others), not unlike the ADP-ribosyltransferase A-B type toxins of many bacteria (e.g. exotoxin A, diphtheria toxin, cholera toxin).[76,77] We determined that highly purified preparations of PlcH induce programmed cell death (apoptosis) in an assortment of eukaryotic cell types (e.g. human monocytes, human endothelial cells). More exactly, even though it is highly cytotoxic to some cell types (e.g. endothelial) PlcH exhibits minimal cytotoxic effects to other cell types (e.g.

human epithelial, mouse fibroblasts). These differences are vast. That is, human vascular endothelial cells (HUVEC) are highly susceptible to picomolar concentrations of PlcH, while micromolar concentrations of PlcH have little if any cytotoxic effect on A549 cells (human airway epithelial cells). Because both susceptible and resistant cell lines contain PC, as well as SM, in the outer leaflet of their cytoplasmic membranes, it is extremely unlikely that the cytotoxicity of PlcH is solely due to its ability to hydrolyze these phospholipids. We have also just obtained data indicating that PlcH interacts with a specific class of calcium-dependent eukaryotic cell receptors (i.e. integrins) through an RGD (Asp-Gly-Glu) motif on in PlcH.[76,77] This highly specific cytotoxic nature of PlcH adds a novel dimension to its role in *P. aeruginosa* virulence. Although, the mechanism by which PlcH induces an apoptotic cell death in endothelial cells is not known at this time, this attribute may serve key functions during specific kinds of *P. aeruginosa* infections. The extreme toxicity of PlcH for endothelial cells is pertinent to the high mortality, blood-borne infections caused by *P. aeruginosa* in immunocompromised individuals with blood dyscrasias. In order for *P. aeruginosa* to enter tissues from blood vessels, they must adhere to and penetrate the endothelial lining of the vasculature. During septicemia, *P. aeruginosa* has the propensity to establish a nidus near the endothelial cell lining of blood vessels, thereby facilitating recurrent seeding of the bloodstream.[80] These foci are often complicated by vasculitis and thrombosis. While endothelial cells may provide protection against *P. aeruginosa* through their increasingly recognized role in innate immune defenses (e.g. production of cytokines), it is also possible that the release of cytokine by PlcH stressed endothelial cells may be detrimental to the host.[43] Only recently has it been recognized that repair of damage to the endothelial lining of the vasculature occurs not by localized proliferation of the undamaged endothelial cells, but that it takes place through circulating, bone marrow derived, progenitor endothelial cells.[64] If PlcH is even moderately cytotoxic to these progenitor endothelial cells, then the healing of these lesions could also be in jeopardy. Moreover, PlcH has been shown to induce platelet aggregation and activation through a novel mechanism of action.[11] Consequently it is not unrealistic to envision a scenario by which PlcH induces damage to the endothelial lining of blood vessels that cannot be repaired by progenitor endothelial cells, causing an influx of platelets, that are then activated by PlcH, and contribute to blood clotting and the thrombotic lesions (i.e. ecthyma gangrenosum) pathognomonic of *P. aeruginosa* sepsis.

Ultimately, it is probably unrealistic to entirely fathom the definitive contribution that a particular extracellular enzyme or virulence factor (i.e. PlcH) makes to the survival of a pathogen. Nevertheless, data about PlcH clearly indicate that it has a noteworthy repertoire of purposes. A corollary to this view is that the distinct biological role that PlcH may play, at a given time, may depend on a particular set of circumstances (e.g. a mammalian host or decaying plant

material) or a certain environment (e.g. low Pi or the presence of choline) in which *P. aeruginosa* finds itself. In this regard, there is yet another dimension of PlcH that supports these notions. It will be explained in Section 7.

4.2. PlcN

The second extracellular PLC of *P. aeruginosa* was discovered once a PlcH deletion (ΔPlcH) mutant was constructed in strain PAO1.[57] A ΔPlcH mutant still expressed an extracellular PLC activity that was detected when *P. aeruginosa* was grown under Pi-limiting conditions, but not when it was grown under Pi-sufficient conditions (Table 2). However, such supernatants had no hemolytic activity on human or sheep erythrocytes. The cloning of the gene encoding this phospholipase (i.e. PlcN) revealed that it encodes a protein (73 kDa), with significant similarity (55%) and size (73 kDa) to PlcH. Additional characterization of PlcN confirmed that it is not hemolytic and that it has similar, but distinct, substrate preferences in comparison to PlcH.[56] Although PlcN is active on PC, it has no detectable activity on SM. However, it is active on the non-choline containing phospholipid phosphatidylserine (PS). There are other features of PlcN, which further convincingly indicate that it is not simply redundant to PlcH in terms of its contribution to the biology to *P. aeruginosa*. Besides the differences already mentioned, the predicted isoelectric point of PlcN is basic (8.8), while the overall p*I* of PlcH is acidic (5.5). However, the deduced pI of each PLC is not uniform over their entire sequences. That is, while the pI of the N-terminal two-thirds of PlcH and PlcN are very similar (5.5 and 6.3, respectively), the pI of the remaining C-terminal portions of PlcH and PlcN (5.7 vs. 10.2) are very different.[56,76] These and other data suggest that these extracellular PLCs are composed of distinct domains. Actually, the C-terminal portions (i.e. outside their conserved PLC domains) of PlcH and PlcN are now classified as separate domains of unknown functions (DUF756) under the Conserved Domain Search algorithm at the NCBI web site. Such a composite structure is quite common in eukaryotic PLCs where there is a core PLC domain, and separate domains that bestow additional functions (e.g. calcium binding, membrane binding) to the entire protein. These facets of PlcH and PlcN, along with others, plainly support the hypothesis that they provide distinct roles in the lifestyle of *P. aeruginosa*. Still, currently PlcN has not been nearly as well studied in terms of its participation in microbial virulence as PlcH. Thus, its role in pathogenesis is considerably less obvious than that of PlcH. In any event, it is salient to point out that eukaryotic cells undergoing apoptosis (e.g. PlcH induced) flip the more PS rich cytoplasmic side of their membranes, to the extra cytoplasmic side.[47] Perhaps, PlcN is able to hydrolyze this newly exposed PS, thereby accelerating eukaryotic cell death initiated by PlcH. Finally, as described below in Section 7, we propose that PlcN, along with PlcH, may under certain circumstances actually participate in phospholipid biosynthesis in *P. aeruginosa*.

4.3. PlcB and PlcA

As recounted above, even homologous phospholipases produced by the same organism may have distinct biochemical or biophysical properties that can determine whether they will or will not be active under a given set of circumstances. These and other factors undoubtedly contributed to the delayed discovery of the third and fourth extracellular PLCs of *P. aeruginosa*. PlcB was first encountered in culture supernatants of a mutant that is deficient in the secretion of PlcH and PlcN. The TAT secretory system is required for the secretion of these PLCs through the inner membrane of *P. aeruginosa*.[53,86] However, the secretion of other extracellular proteins via the Sec-translocase is generally unaffected in a ΔTatC mutant. We detected an unanticipated PLC activity in culture supernatants of a *P. aeruginosa* ΔTatC mutant.[5] Further characterization of this activity led to the identification of the gene encoding PlcB. Prior to that, the *plcB* gene was annotated as PA0026, encoding a 36 kDa hypothetical protein with unknown, unclassified function (HUU). When the sequence of PlcB was further scrutinized using the Conserved Domain Search algorithm at the NCBI web site, we discovered that a limited region (∼60 amino acids) of PlcB shares some degree of similarity to the zinc-dependent GP-PLCs. Moreover, the conserved His residues in this homologous region are required for coordination of three zinc ions in the active site of this class of PLCs.[22,29] Mutagenesis of one of these His residues in PlcB resulted in abrogation of its PLC activity.[5] The PlcB gene is part of a three-gene operon, which includes a gene (PA0027) encoding a peptidyl prolyl *cis–trans* isomerase (PpiP) and a gene (PA0028) encoding a protein predicted to be a proline rich lipoprotein.[4] Although, the functions of these other proteins are not entirely clear at the present time, preliminary data suggest that one (PpiP) is required for the proper folding of PlcB, while the other (PA0028) is required for the secretion (e.g. chaperone) of PlcB through the membrane of *P. aeruginosa*.[4] It is also of interest to point out that PpiP is secreted to the periplasm via the TAT secretory pathway. This feature exemplifies the cooperative nature of the Sec (PlcB) and the TAT translocation systems (PpiP) in terms of exporting proteins that may ultimately interact with each other. Such is the case with PlcH and PlcR1, which ultimately form an extracellular heterodimer. PlcH is exported to the periplasm via the TAT system, while PlcR is exported through Sec.[76] The PlcHR1 heterodimer is then secreted through the outer membrane via the Xcp apparatus of *P. aeruginosa*.

The substrate specificity of PlcB provided some hints with regard to its potential function in the biology of *P. aeruginosa.* PlcB is active on PC, SM, and PS; however in contrast to either PlcH or PlcN, it is also highly active on phosphatidylethanolamine (PE) (Figure 1, Table 1). Since, it had been previously reported that *P. aeruginosa* is able to exhibit twitching motility-mediated

chemotaxis up a gradient of this particular phospholipid, it was reasoned that PlcB might somehow play a role in this process.[33] This supposition was confirmed when it was demonstrated that PlcB is required for the chemotaxis of *P. aeruginosa* up a gradient of diolyl-PE or diolyl-PC, but it is not necessary for twitching-mediated chemotaxis toward the dilauryl forms of these phospholipids.[5] It is probable that fatty acids (diolyl), released by the action of extracellular lipases on the DAG generated from the action of PlcB on PE or PC, ultimately contribute the specificity of this chemotactic response.

From the perspective of the would-be role of PlcB in *P. aeruginosa* biology, it is germane that the major lipid constituent of mammalian lung surfactant is PC. In fact, both PC and PE, along with SM levels, are increased in the bronchoalveolar lavage fluid (BAL) of young adults with cystic fibrosis (CF) by comparison with BAL from age matched controls, without CF.[49] It is possible that the increased levels of these phospholipids serve as a chemoattractants to *P. aeruginosa*, which initially colonizes the upper airways in a CF patient. The choline that would be generated by any of the *P. aeruginosa* PLCs from this PC rich resource could also protect it against the increased osmolarity it will encounter in the lower airways of the CF lung.

The regulation of PlcB expression offers insights into the *raison d'être* for existence of yet another extracellular PLC in *P. aeruginosa* (Table 2). In spite of the fact that the expression of PlcH, PlcN, and PlcB is Pi-starvation inducible, there are several reports from independent investigators, based on microarray data, indicating that the expression of *plcB* (i.e PA0026), but not *plcH* or *plcN*, is dependent upon homoserine lactone-mediated quorum sensing.[69,88] Additional microarray data support the view that other environmental factors differentially influence the expression of PlcH, PlcN, and PlcB, as well. Wolfgang *et al*.[92] reported that the expression of PlcN and PlcB, but not PlcH, were increased when *P. aeruginosa* was exposed to muco-purulent airway liquids from chronically infected CF patients, as compared to when it was grown minimal media alone. In a separate study, these investigators provided experimental results indicating that expression of PlcB and PlcN are also affected by the *P. aeruginosa* cAMP-binding protein, Vfr, but regulation of PlcH expression was not.[93] In contrast, an adenyl cyclase mutant (Δ*cyaAB*) expressed significantly reduced levels of PlcB transcript, but transcription of the PlcN and PlcH genes were unaffected or slightly increased in this mutant.[93] The intricate differences in the regulation of *plcB, plcH*, and *plcN* further typify the distinct purpose for the proteins they encode, despite the fact that all are classified as PLCs.

PlcA is the most recently identified PLC of *P. aeruginosa,* and therefore the least understood in terms of its utility to *P. aeruginosa*, much less its contribution to virulence. Once we had recognized that a portion of the PlcB sequence shares similarity with a region of the active site of the GP-PLCs, we were further analyzing its properties through an algorithm using Hidden Markov

Models that represent all proteins of known structure (http://supfam.mrc-lmb.cam.ac.uk/SUPERFAMILY/hmm.html).[8] The results of this analysis revealed that the genome of *P. aeruginosa* actually encodes two separate proteins, containing the zinc-dependent PLC active site motif that is present in the PlcB of *P. aeruginosa* and the GP-PLCs. Of course one is PlcB (PA0026), but the other is encoded by a gene in the *P. aeruginosa* genome annotated as PA3464. PA3464, like PA0026 (i.e. PlcB), was annotated in *P. aeruginosa* Genome Project web site (http://www.pseudomonas.com) as encoding a hypothetical, unclassified, unknown protein. Paradoxically, the Superfamily site suggested that it is a PLC. We have now expressed the protein encoded by PA3464 and demonstrated that it has PLC activity on PC, PS, and PE, but it has no detectable activity on SM, as PlcB does.[76] PlcA is predicted to be a 46.7 kDa acidic (pI 5.2) protein (Table 1). The *P. aeruginosa* Genome Project web site does indicate that, based on the PSORT algorithm, PlcA has a Sec-type signal peptide. Nevertheless, at the present its ultimate location in *P. aeruginosa* is unknown. Recently, we were able to detect the expression of PlcA in *P. aeruginosa*. Preliminary data indicate that its expression does not appear to be induced under Pi-limiting conditions and does appear to be influenced by choline. However, in contrast to the choline induced expression of PlcH, the expression of PlcA is inhibited by the presence of choline in the media. An even more concerted effort is going to be needed in order to tease out further information about the functionality of PlcA to *P. aeruginosa*.

5. PLDS OF *P. AERUGINOSA*

As related above, this type of phospholipase is quite unusual in prokaryotes. However, it should be pointed out that PLDs belong to a rather large class of enzymes referred to as the "PLD Superfamily."[48,78] It includes mammalian and plant PLDs, a Vaccina virus protein, cardiolipin synthase and endonucleases, and phosphatidylserine synthetase. This classification is based on the presence of conserved active site motif, HXKX4DX6G(G/S), which is usually present in two copies in most members (e.g. PLDs) and a single copy in others (e.g. endonucleases). Among the PLD members, there are two major subclassifications based on differences in homology. Plant PLDs comprise one class, while the fungal (yeast) and animal (mammalian and nematode) comprise the other. Bacterial PLDs usually belong to the plant class.

There is a wealth of information regarding critical roles of PLDs in notably various physiological and genetic processes in eukaryotic cells. In contrast, little is known about the role of PLDs in prokaryotes, which is largely due to their relative rare occurrence in bacteria. As mentioned above, the best-studied one is the so-called "murine toxin" of *Y. pestis*.[25,26] However, there are members of the "PLD Superfamily" that are actually quite common in

prokaryotes. Cardiolipin, along with phosphatidylglycerol (PG) and PE, is a major phospholipid of prokaryotic membranes and it is formed in both gram-positive and gram-negative bacteria by the transfer of a phosphatidyl group from one PG molecule to another via cardiolipin synthetase (CLS).[81] In contrast, *bona fide* PLDs, which remove the choline head group from PC are few and far between in these organisms. Moreover, CLSs are almost exclusively membrane bound while bacterial PLDs are more likely to be located in soluble fractions (e.g. cytoplasm, periplasm). Another significant issue is that, while the substrate (e.g. PG) for CLSs is plentiful in bacteria, the preferred substrate for PLDs (i.e. PC) is usually very limited or absent in most bacteria.[46] These differences illustrate the distinct functions that each of these subclasses of the PLD Superfamily plays in prokaryotic biology.

Wilderman *et al*.[89] initially identified a true PLD (i.e. not a CLS) activity in supernatants of late stationary grown *P. aeruginosa* strain PAO1. Examination of the annotated genomic database (http://www.pseudomonas.com) revealed that, in addition to genes encoding CLSs, there are two additional genes encoding proteins, each with two copies of the HXKX4DX6G(G/S) PLD motif. Further characterization of one of these putative PLDs (i.e. PldA) revealed that its primary location is actually periplasmic. The earlier detection of PLD activity in culture supernatants was likely due to release of PLD activity through lysis during stationary phase, not unlike that observed early on with the *Y. pestis* PLD. PldA has extended regions of homology with animal PLDs, and very limited homology to plant PLDs.[89] There are also unique cassette-like regions with no homology with any sequences in all the databases examined. These investigators proposed that the unique regions play a role in regulating PLD activity in *P. aeruginosa* and that they might roughly correspond to the domains in PLDs that interact with eukaryotic regulatory factors, such as ARF, Ras, and Rho. Since PldA uses PC as a preferred substrate, and it is located in the periplasm, it was difficult to envisage how it might have access to this substrate because PC was not known to be present in either the inner or outer membranes of *P. aeruginosa*. This apparent paradox led Wilderman *et al*.[90] to investigate whether *P. aeruginosa* is one of the relatively few bacteria capable of PC synthesis. Indeed, they reported that *P. aeruginosa* contains a small, but significant amount of PC in both membranes, where it would be accessible to PldA. They also discovered that *P. aeruginosa* is able to synthesize PC, *de novo*, through the condensation of choline with CDP-DAG by a PC synthase (Pcs). Consequently, the presence of PC and the periplasmic location of PldA provided a new perspective into the possible role of PldA in *P. aeruginosa*. That is, in contrast to the extracellular nature of bacterial PLCs (e.g. PlcH and α toxin) and their propensity to affect host phospholipid metabolism, it is more likely that PldA plays a direct role in phospholipid-mediated signaling events in *P. aeruginosa*. This scenario, however, does not mitigate against PldA playing a role in *P. aeruginosa* pathogenesis. Wilderman *et al*.[89] demonstrated that a PldA

deletion mutant (ΔPldA) was deficient in its ability to compete with the wild-type parental strain in a chronic pulmonary infection model in rats (Table 1). The ΔPldA mutant had no apparent *in vitro* growth defect compared to the parental strain. Although these data do not shed any light on the specific role of PldA in virulence, they certainly are indicative of a more expansive role for bacterial phospholipases in pathogenesis. As a final point about *P. aeruginosa* PLDs, Wilderman *et al*. reported that *pldA* is immediately adjacent to a highly variable genetic element (i.e. *vgr*). The *vgr* genes are components of multicopy rearrangement hot-spots (*Rhs*) first described in *Escherichia coli*. Both *E. coli* and *P. aeruginosa* carry multiple copies of homologs of *vgr* genes, but the actual number in any given strain in these bacteria is highly variable, ranging from 1 to 10 copies.[24,89] The function of the proteins encoded by the *vgr* genes is unknown, but it is thought that they are located on the bacterial surface. In any case, Wilderman *et al*. found that approximately 30% of *P. aeruginosa* strains including those from patients with CF, as well as environmental strains, carry a gene encoding PldA. What is more, they noted that the *pldA* gene is invariably associated with a specific copy of a *vgr* gene and a gene encoding a hydrophobic lipoprotein of unknown function (Table 2). These three genes appear to comprise a pathogenicity island (PAI). They are either, entirely present, or wholly absent in all strains (57 isolates) of *P. aeruginosa* examined and, the regions flanking this PAI are highly conserved whether the PAI is present or not. Perhaps this PAI provides a selective advantage to this *P. aeruginosa* in a niche, yet to be defined. Although *pldA* encodes a protein most similar to eukaryotic PLDs, it may have been acquired from an unidentified prokaryotic intermediate along with a *vgr*.[23,24,89] The possible involvement of the multiple copies of the highly variable *vgr* genes in the evolution of *P. aeruginosa* through recombination and horizontal gene transfer also has powerful implications in the ongoing evolution of this opportunistic pathogen. Notably, there is yet another gene in *P. aeruginosa* encoding a protein with significant similarity to plant PLDs. This gene is likewise associated with a *vgr* gene and is present in a portion (∼90%) of *P. aeruginosa* strains examined.[89]

6. PLA2 (EXOU) OF *P. AERUGINOSA*

Although this phospholipase (i.e. ExoU) has been recently competently reviewed elsewhere, there are a few additional comments apropos to the topics in this chapter.[65] In contrast to the other phospholipases discussed herein that were initially identified by their enzymatic activities, ExoU was first identified through its association with the Type III secretion system and by its association with certain strains of *P. aeruginosa,* which are highly cytotoxic to epithelial cells.[15] It was eventually determined that the gene encoding ExoU is associated

with a PAI-like genome fragment and that it is present in approximately only 30% of *P. aeruginosa* strains thus far examined, including those from clinical and environmental sources.[73] This frequency is reminiscent of that of the *pldA* gene, which is likewise associated with a PAI (see above). In contrast, the genes encoding PlcH, PlcN, PlcB, and PlcA appear to be associated with a more stable region of the *P. aeruginosa* core chromosome and we have not encountered any strain of *P. aeruginosa* that lacks any of these phospholipase genes.[5,76,77,85] Perhaps, ExoU and PldA provide a selective advantage in only certain niches where *P. aeruginosa* resides. For example, ExoU does not appear to play a determinative role in the pathogenesis of *P. aeruginosa* to *C. elegans* or *Arabidopsis,* but data from several studies support the notion that ExoU contributes to specific kinds of human infections (e.g. hospital acquired pneumonias).[58,68]

The enzymatic activity of ExoU was only recently uncovered.[65,66] Probably contributing to this delay is the fact that the PLA activity of this virulence factor requires an, as yet, unidentified eukaryotic cytosolic factor (Table 2). For this, and for the same reasons that PlcB was previously unrecognized as a PLC, it would not be too surprising if other proteins of *P. aeruginosa*, particularly those with the classification HUU, are ultimately established to have phospholipase activity. The first clue that ExoU might have phospholipase activity emerged from several independent observations that inhibitors of eukaryotic PLA2s also inhibited the cytotoxicity of ExoU.[59] While data using enzymatic inhibitors can certainly be enticing, there are also pitfalls associated with their use. That is, they could be blocking cellular processes downstream of the primary effect of toxin (e.g. ExoU). The breakthrough that was perhaps more indicative of ExoU phospholipase activity came when it was discovered that a region of ExoU aligned to the patatin motif, and two human PLA2s. Mutagenesis of two amino acids of ExoU in its conserved PLA region reduced its toxicity.[65,66] Similar to the other phospholipase described in this chapter, ExoU appears to have separate domains with distinct functions. Only the N-terminal half of ExoU contains the conserved patatin domain. The function of the other portion is unknown, but it is rational to suppose that it plays a role in interaction with the eukaryotic cytosolic factor required for its toxicity.

The precise mechanism of the enzymatic activity of ExoU and its substrate preferences are still not well defined. Furthermore, it is not at all certain how its phospholipase activity imparts cytotoxicity to ExoU. Based on its similarity to patatins and human PLA2 and some preliminary biochemical experiments with eukaryotic cell extracts, some investigators originally suggested that ExoU is a PLA2 (Figures 1 and 3). However, the latest experimental results suggest that ExoU should more specifically be classified as a lysophospholipase A.[58,79] That is because ExoU is considerably more (10X) active on lysophospholipids (i.e. those that only have a single fatty acid chain in either the *sn*-1 or *sn*-2 position) than it is on phospholipids with two fatty acid moieties (e.g. dipalmitoyl-PC).

Part of the problem in identifying the specific nature of the enzymatic activity of ExoU emanates from its requirement for a eukaryotic cytosolic factor. In any case, at present, it is difficult to envision how a lysophospholipase evokes such an extraordinary cytotoxic phenotype in a broad range of eukaryotic cells, including yeast. It does not seem likely that such an enzymatic activity could physically disrupt the integrity of eukaryotic cell membranes and lead to their rapid lysis. This is particularly true for live eukaryotic cells that are continually synthesizing new membrane phospholipids. Perhaps, downstream signaling events from the primary target of ExoU are responsible for the exuberant cytotoxicity attributed to this virulence determinant. A striking parallel to the conundrum of how the enzymatic activity of ExoU relates to its highly toxic (cytolytic) nature for eukaryotic cells, is also observed with PlcH. Although it is able to hydrolyze membrane phospholipids, it is rather difficult to reconcile how the enzymatic activity of extremely small amounts of PlcH (see above) is connected to the induction of complex eukaryotic cell signaling process (i.e. apoptosis). It is possible that the biochemical information we currently have about all of the phospholipases discussed in this chapter is inadequate and that the enzymatic activities that actually are the cause of the observed biological phenotypes have yet to be clearly elucidated. The following section addresses some of these matters.

7. PHOSPHOLIPASES: HYDROLASE OR SYNTHASE?

Basically, the names of enzymes are defined by their products, which are usually detected in an *in vitro* biochemical reaction. However, the detection of particular product in such experiments may not actually reflect biological reality. For example, the ADP ribosyltransferase toxins (e.g. exotoxin A, diphtheria toxin) will hydrolyze the ADP ribose moiety from NAD and transfer it to a water molecule (i.e. NAD glycohydrolase).[19,40] Cholera toxin can also ADP-ribosylate small artificial substrates such as agmatine.[55] However, the biologically relevant acceptor for the ADP ribose moiety is actually the eukaryotic protein, EF2, or a subunit of an adenyl cyclase regulator. Moreover, some proteins may have multiple enzymatic activities and preferred substrates (e.g. PC and SM for PlcH) further confounding the chance of identifying all the biologically significant substrates and products. Such could be the case for any, or all, of phospholipases described herein.

In terms of one of the phospholipases discussed in this chapter, Tamura *et al*.[79] reported that glycerol, rather than, for instance water, might serve as an acceptor for the cleaved fatty acid in the ExoU-catalyzed lysophospholipase reaction. Provided this acyl transferase activity of ExoU occurs in the cytoplasm of eukaryotic cells then, this product (a monoacylglycerol) or its derivatives

(diacyl or triacyl glycerol), rather than fatty acids *per se* could ultimately be significant to biological activity (e.g. cytotoxicity) of this protein.

Also in relation to this issue, Luberto *et al*.[41] verified that, although PlcH has PLC and SMase, they found that depending on whether ceramide (CM) is available or not, PlcH also has SM synthase activity (Figure 4) That is, it is able to transfer the choline phosphate head group from PC to CM resulting in the synthesis of SM and the production of DAG. The specificity of this activity is demonstrated by the fact that the choline phosphate must come from a phospholipid. PlcH will not use nonlipid substrates (e.g. nitrophenyl-phosphorylcholine) as a donor for the choline phosphate this reaction. However, if SM is present, without CM, PlcH will remove the choline phosphate head group from SM thereby generating CM. Transfer of a phosphorylcholine head group from SM to DAG is not observed. Hence, it is the relative concentrations of PC, SM, and CM, which determine whether PlcH has SMase, PLC (i.e hydrolase), or SM synthase activity. These three distinct enzymatic activities attributable to a single protein (i.e PlcH) beg the question; what are its actual biologically relevant substrates and products? More than likely, the answer depends upon the environment in which *P. aeruginosa* may be found at a given time. For example in lung surfactant, while there is plenty of PC and SM, there is little if any CM available. Consequently, one might imagine in such a situation PlcH would act as a PLC or as a SMase. On the other hand, in terms of the cytotoxicity of PlcH, because it induces apoptosis in endothelial cells, rather than a proliferative response, it is logical to suppose that the cytotoxicity of PlcH is primarily associated with its SMase activity rather than its PLC activity. The rationale for this hypothesis is that, as cited earlier, CM induces an apoptotic response in eukaryotic cells, while DAG, causes cell proliferation and transformation of eukaryotic cells (Figure 2).[20]

An alternative point of view that could change the perceived function of PlcH in *P. aeruginosa* is that, it is possible this organism could actually synthesize CM *de novo* and thereby provide a precursor for SM synthesis by PlcH, if it was otherwise not present (e.g. lung surfactant). However, this hypothesis would seem to challenge conventional wisdom because so far bacteria, with the possible but unproven exception of *Mycoplasma*, are not known to contain SM, much less synthesize it.[52] Actually, there are some bacteria, not too distantly related to *P. aeruginosa* (i.e. *Sphingomonas* spp.), which contain glycosphingolids as part of their outer membranes, and *Bacteriodes* has been shown to contain CM in its membranes.[38,52] In *Sphingomonas* spp., glycosphingolipids actually substitute for the ubiquitous LPS present in other gram-negative bacteria.[32] While there are now ample data indicating that bacteria contain CM, nothing is known about the pathway for the synthesis of sphingolipids in these organisms. Even so, data are available indicating that the sphingosine precursor of CM, sphinganine, is also present in these bacteria (Figure 4).[31] Concerning the possibility that *P. aeruginosa* is able to synthesize CM, it is interesting to point out that an

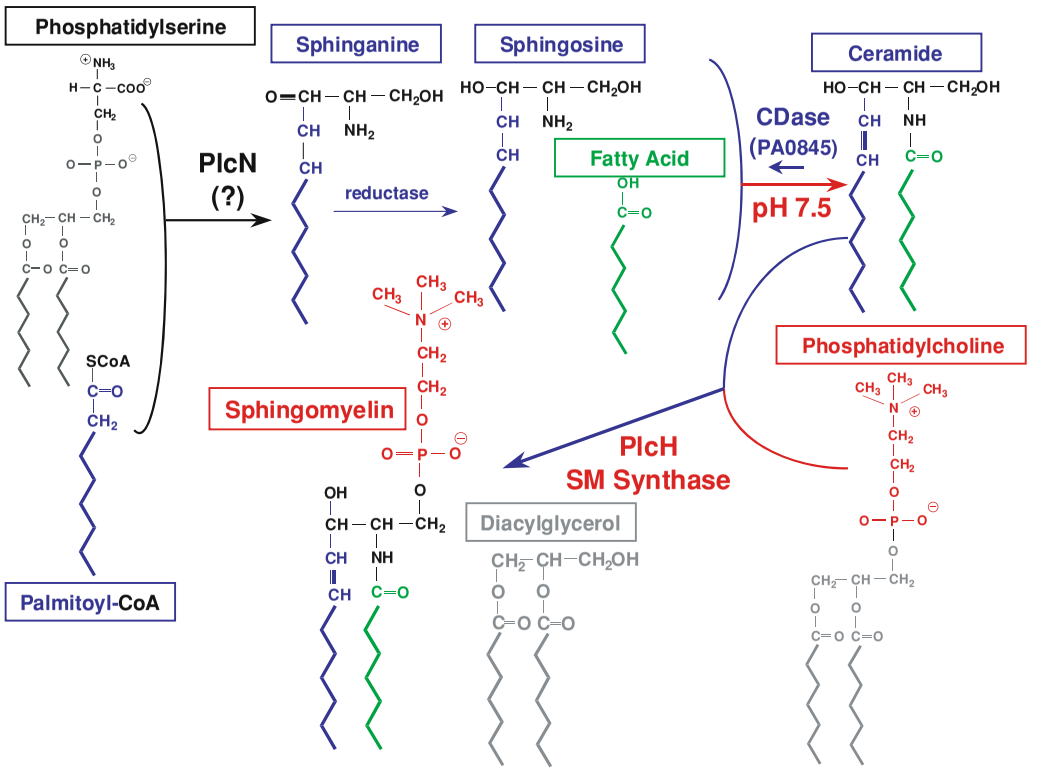


Figure 4. Proposed pathway for the synthesis of sphingomyelin (SM) and its precursors (e.g. ceramide and sphinganine) by extracellular enzymes (i.e. PlcH, PlcN, CDase) of *P. aeruginosa*.

extracellular ceramidase (CMase) of this organism has been isolated and well characterized.[54] It is likewise notable that the gene encoding this enzyme is immediately adjacent (5′) to the PlcH gene in *P. aeruginosa*. However, as shown in Figure 4, such an enzyme would actually be counterproductive in terms of SM synthesis because it would degrade CM, not synthesize it. At variance with this view is the fact that its designated enzymatic activity (i.e. CDase) is dependent upon certain specific conditions (see below). If different ones are present, this protein readily catalyzes the reverse reaction leading to CM synthesis.[34] This *P. aeruginosa* protein was initially classified as an alkaline CMase because it is most active at an alkaline pH (pH optimum 8.5). Nevertheless, Kita *et al.*[34] reported that the reverse reaction occurs very efficiently at near neutral pH (pH optimum 7.5). That is, this protein can catalyze the condensation of free fatty acids to sphingosine resulting in the synthesis of CM, rather than in its degradation (Figure 4). Consequently, it is now tenable to take the view that *P. aeruginosa* has the enzymes to actually synthesize SM when PC, sphingosine, and fatty acids are present in the environment. What is more, it is possible that PlcN also contributes to the *de novo* biosynthesis of SM in *P. aeruginosa* (Figure 4) through the production of a CM precursor.[50] As mentioned earlier, PlcN, but not PlcH, is able to hydrolyze PS to phosphoryl serine and DAG. If PlcN like PlcH can transfer the head group of this phospholipid (i.e. serine) to something other than water, such as palmitoyl-CoA, then this reaction would actually yield sphinganine precursor (Figure 4).[50] Sphinganine would then need to be reduced by an NADPH-dependent reductase (Figure 4). These are relatively minor modifications to convert sphinganine to sphingosine, but it is not improbable that either PlcN itself makes these minor alterations during the course of the reaction or that other *P. aeruginosa* enzymes can make these conversions (Figure 4). In any case, such a scenario would enable *P. aeruginosa* to completely synthesize SM *de novo* through PlcH, its so-called alkaline CDase (i.e a ceramide synthase), and PlcN. Perhaps one or two other as yet unidentified enzymes are also necessary.

Even though *P. aeruginosa* may be able to synthesize SM the benefit of such a phospholipid to its survival is not immediately obvious. *P. aeruginosa* can also synthesize the eukaryotic-like phospholipid PC and Wilderman *et al*.[90] determined that PC deficient mutants of *P. aeruginosa* are less fit after being under certain environmental stresses (i.e. freezing). With regard to SM, this phospholipid contains long, largely saturated acyl chains thereby causing tighter packing in membranes and resulting in their concentration in lipid rafts in eukaryotic cells. SM also has a much higher melting temperature than glycerol containing phospholipids. Along with cholesterol and glycosphingolipids SM is the lipid that is most correlated with an increased viscosity of airway secretions. Of note, several independent studies reported increased levels of SM in the bronchiolar alveolar lavage fluid of CF patients in comparison to non-CF patients, including those with other pulmonary diseases.[18,49] A more

recent report also found that there is a significant correlation between increased levels of PLC production by *P. aeruginosa* and poor pulmonary function in CF patients.[37] While such data do not directly connect the synthesis of SM by *P. aeruginosa* with virulence, they offer novel perspectives into the potential capabilities of these enzymes in pathogenesis, whether they have hydrolase (i.e. PLC) or synthase activities.

As a final point, the paradigms regarding the role of all the phospholipases discussed in this chapter are admittedly biased toward their possible contributions to mammalian pathogenesis. It is just as likely that they also have utility in other aspects of the lifestyle of this organism, including its survival in riparian soil, in plants or even to its ability to compete with other microorganisms. PlcH was recently shown to be, in some way, involved in killing the opportunistic fungal pathogen, *Candida albicans.*[28] Perhaps their multifunctional nature is merely a reflection of and contributes to the remarkable adaptability of *P. aeruginosa*.

ACKNOWLEDGMENTS

The author gratefully acknowledges the thoughtful and friendly interactions with the remarkable colleagues: Howard Goldfine, Yusuf Hannun, Chiara Luberto who have taught him to better appreciate the more intricate, but critical, points of the biochemistry and enzymology of phospholipids and phospholipases. He very much appreciates the exceptionally hard work and highly valued intellectual input of the following postdoctoral fellows, technical personnel, and graduate students from his laboratory (alphabetically arranged) including: Adam Barker, Randy Berka, Adela Cota-Gómez, Gregory Gray, Kristine Johansen, Zaiga Johnson, Urs Ochsner, Rachel Ostroff, Sarah Parker, Arthur Pritchard, Andy Sage, Virginia Shortridge, Martin Stonehouse, Adriana Vasil, and P.J Wilderman. All made notable contributions to the information described in this chapter. Finally, he sincerely thanks all other exceedingly capable investigators, too numerous to mention herein, for their collaborative efforts that made key contributions as well.

REFERENCES

1. Arni, R.K., and Ward, R.J., 1996, Phospholipase A2 – a structural review. *Toxicon*, 34:827–841.
2. Awad, M.M., Ellemor, D.M., Boyd, R.L., Emmins, J.J., and Rood, J.I., 2001, Synergistic effects of alpha-toxin and perfringolysin O in *Clostridium perfringens*-mediated gas gangrene. *Infect. Immun.*, 69:7904–7910.
3. Baine, W.B., 1985, Cytolytic and phospholipase C activity in *Legionella* species. *J. Gen. Microbiol.*, 131:1383–1391.

4. Barker, A.P., and Vasil, M.L., 2005, Unpublished observations.
5. Barker, A.P., Vasil, A.I., Filloux, A., Ball, G., Wilderman, P.J., and Vasil, M.L., 2004, A novel extracellular phospholipase C of *Pseudomonas aeruginosa* is required for phospholipid chemotaxis. *Mol. Microbiol.*, 53:1089–1098.
6. Bode, G., Barth, R., Song, Q., and Adler, G., 2001, Phospholipase C activity of *Helicobacter pylori* is not associated with the presence of the *cagA* gene. *Eur. J. Clin. Invest.*, 31:344–348.
7. Camilli, A., Goldfine, H., and Portnoy, D.A., 1991, *Listeria monocytogenes* mutants lacking phosphatidylinositol-specific phospholipase C are avirulent. *J. Exp. Med.*, 173:751–754.
8. Choo, K.H., Tong, J.C., and Zhang, L., 2004, Recent applications of Hidden Markov models in computational biology. *Genomics Proteomics Bioinformatics*, 2:84–96.
9. Clark, G.C., Briggs, D.C., Karasawa, T., Wang, X., Cole, A.R., Maegawa, T., Jayasekera, P.N., Naylor, C.E., Miller, J., Moss, D.S., Nakamura, S., Basak, A.K., and Titball, R.W., 2003, *Clostridium absonum* alpha-toxin: new insights into clostridial phospholipase C substrate binding and specificity. *J. Mol. Biol.*, 333:759–769.
10. Cota-Gomez, A., Vasil, A.I., Kadurugamuwa, J., Beveridge, T.J., Schweizer, H.P., and Vasil, M.L., 1997, PlcR1 and PlcR2 are putative calcium-binding proteins required for secretion of the hemolytic phospholipase C of *Pseudomonas aeruginosa*. *Infect. Immun.*, 65:2904–2913.
11. Coutinho, I.R., Berk, R.S., and Mammen, E., 1988, Platelet aggregation by a phospholipase C from *Pseudomonas aeruginosa*. *Thromb. Res.*, 51:495–505.
12. de Silva, N.S., and Quinn, P.A., 1987, Rapid screening assay for phospholipase C activity in mycoplasmas. *J. Clin. Microbiol*., 25:729–731.
13. de Torres Zabela, M., Fernandez-Delmond, I., Niittyla, T., Sanchez, P., and Grant, M., 2002, Differential expression of genes encoding *Arabidopsis* phospholipases after challenge with virulent or avirulent *Pseudomonas* isolates. *Mol. Plant. Microbe. Interact.*, 15:808–816.
14. Ernst, R.K., Yi, E.C., Guo, L., Lim, K.B., Burns, J.L., Hackett, M., and Miller, S.I., 1999, Specific lipopolysaccharide found in cystic fibrosis airway *Pseudomonas aeruginosa*. *Science*, 286:1561–1565.
15. Finck-Barbancon, V., Goranson, J., Zhu, L., Sawa, T., Wiener-Kronish, J.P., Fleiszig, S.M., Wu, C., Mende-Mueller, L., and Frank, D.W., 1997, ExoU expression by *Pseudomonas aeruginosa* correlates with acute cytotoxicity and epithelial injury. *Mol. Microbiol.*, 25:547–557.
16. Flieger, A., Rydzewski, K., Banerji, S., Broich, M., and Heuner, K., 2004, Cloning and characterization of the gene encoding the major cell-associated phospholipase A of *Legionella pneumophila*, plaB, exhibiting hemolytic activity. *Infect. Immun.*, 72:2648–2658.
17. Geoffroy, C., Raveneau, J., Beretti, J.L., Lecroisey, A., Vázquez-Boland, J.A., Alouf, J.E., and Berche, P., 1991, Purification and characterization of an extracellular 29-kilodalton phospholipase C from *Listeria monocytogenes*. *Infect. Immun.*, 59:2382–2388.
18. Girod, S., Galabert, C., Lecuire, A., Zahm, J.M., and Puchelle, E., 1992, Phospholipid composition and surface-active properties of tracheobronchial secretions from patients with cystic fibrosis and chronic obstructive pulmonary diseases. *Pediatr. Pulmonol.*, 13:22–27.
19. Han, X.Y., and Galloway, D.R., 1995, Active site mutations of *Pseudomonas aeruginosa* exotoxin A. Analysis of the His440 residue. *J. Biol. Chem.*, 270:679–684.
20. Hannun, Y.A., and Luberto, C., 2000, Ceramide in the eukaryotic stress response. *Trends Cell Biol.*, 10:73-80.
21. Hannun, Y.A., Luberto, C., and Argraves, K.M., 2001, Enzymes of sphingolipid metabolism: from modular to integrative signaling. *Biochemistry*, 40:4893–4903.
22. Hansen, S., Hansen, L.K., and Hough, E., 1993, The crystal structure of tris-inhibited phospholipase C from *Bacillus cereus* at 1.9 A resolution. The nature of the metal ion in site 2. *J. Mol. Biol.*, 231:870–876.
23. Hill, C.W., Feulner, G., Brody, M.S., Zhao, S., Sadosky, A.B., and Sandt, C.H., 1995, Correlation of Rhs elements with *Escherichia coli* population structure. *Genetics*, 141:15–24.

24. Hill, C.W., Sandt, C.H., and Vlazny, D.A., 1994, Rhs elements of *Escherichia coli*: a family of genetic composites each encoding a large mosaic protein. *Mol. Microbiol.*, 12:865–871.
25. Hinnebusch, B.J., Rudolph, A.E., Cherepanov, P., Dixon, J.E., Schwan, T.G., and Forsberg, A., 2002, Role of *Yersinia* murine toxin in survival of *Yersinia pestis* in the midgut of the flea vector. *Science*, 296:733–735.
26. Hinnebusch, J., Cherepanov, P., Du, Y., Rudolph, A., Dixon, J.D., Schwan, T., and Forsberg, A., 2000, Murine toxin of *Yersinia pestis* shows phospholipase D activity but is not required for virulence in mice. *Int. J. Med. Microbiol.*, 290:483–487.
27. Hirschberg, H.J., Simons, J.W., Dekker, N., and Egmond, M.R., 2001, Cloning, expression, purification and characterization of patatin, a novel phospholipase A. *Eur. J. Biochem.*, 268:5037–5044.
28. Hogan, D.A., and Kolter, R., 2002, Pseudomonas–Candida interactions: an ecological role for virulence factors. *Science*, 296:2229–2232.
29. Hough, E., Hansen, L.K., Birknes, B., Jynge, K., Hansen, S., Hordvik, A., Little, C., Dodson, E., and Derewenda, Z., 1989, High-resolution (1.5 A) crystal structure of phospholipase C from *Bacillus cereus*. *Nature*, 338:357–360.
30. Johansen, K.A., Gill, R.E., and Vasil, M.L., 1996, Biochemical and molecular analysis of phospholipase C and phospholipase D activity in mycobacteria. *Infect. Immun.*, 64:3259–3266.
31. Kawahara, K., Lindner, B., Isshiki, Y., Jakob, K., Knirel, Y.A., and Zahringer, U., 2001, Structural analysis of a new glycosphingolipid from the lipopolysaccharide-lacking bacterium *Sphingomonas adhaesiva*. *Carbohydr. Res.*, 333:87–93.
32. Kawahara, K., Moll, H., Knirel, Y.A., Seydel, U., and Zahringer, U., 2000, Structural analysis of two glycosphingolipids from the lipopolysaccharide-lacking bacterium *Sphingomonas capsulata*. *Eur. J. Biochem.*, 267:1837–1846.
33. Kearns, D.B., Robinson, J., and Shimkets, L.J., 2001, *Pseudomonas aeruginosa* exhibits directed twitching motility up phosphatidylethanolamine gradients. *J. Bacteriol.*, 183:763–767.
34. Kita, K., Okino, N., and Ito, M., 2000, Reverse hydrolysis reaction of a recombinant alkaline ceramidase of *Pseudomonas aeruginosa*. *Biochim. Biophys. Acta.*, 1485:111–120.
35. Korbsrisate, S., Suwanasai, N., Leelaporn, A., Ezaki, T., Kawamura, Y., and Sarasombath, S., 1999, Cloning and characterization of a nonhemolytic phospholipase C gene from *Burkholderia pseudomallei. J. Clin. Microbiol.*, 37:3742–3745.
36. Langton, S.R., and Cesareo, S.D., 1992, *Helicobacter pylori* associated phospholipase A2 activity: a factor in peptic ulcer production? *J. Clin. Pathol.*, 45:221–224.
37. Lanotte, P., Mereghetti, L., Lejeune, B., Massicot, P., and Quentin, R., 2003, *Pseudomonas aeruginosa* and cystic fibrosis: correlation between exoenzyme production and patient's clinical state. Pediatr. *Pulmonol.*, 36:405–412.
38. Lev, M., 1979, Sphingolipid biosynthesis and vitamin K metabolism in *Bacteroides melaninogenicus*. *Am. J. Clin. Nutr.*, 32:179–186.
39. Liu, P.V., 1966, The roles of various fractions of *Pseudomonas aeruginosa* in its pathogenesis. II. Effects of lecithinase and protease. *J. Infect. Dis.*, 116:112–116.
40. Lory, S., Carroll, S.F., Bernard, P.D., and Collier, R.J., 1980, Ligand interactions of diphtheria toxin. I. Binding and hydrolysis of NAD. *J. Biol. Chem.*, 255:12011–2015.
41. Luberto, C., Stonehouse, M.J., Collins, E.A., Marchesini, N., El-Bawab, S., Vasil, A.I., Vasil, M.L., and Hannun, Y.A., 2003, Purification, characterization, and identification of a sphingomyelin synthase from *Pseudomonas aeruginosa*. PlcH is a multifunctional enzyme. *J. Biol. Chem.*, 278:32733–32743.
42. Macfarlane, M.G., and Knight, B.C., 1941, The biochemistry of bacterial toxins. I The lecithinase activity of *Clostridium welchii* toxin. *Biochem. J.*, 35:884–902.
43. Mantovani, A., Sozzani, S., Vecchi, A., Introna, M., and Allavena, P., 1997, Cytokine activation of endothelial cells: new molecules for an old paradigm. *Thromb. Haemost.*, 78:406–414.

44. Marques, M.B., Weller, P.F., Parsonnet, J., Ransil, B.J., and Nicholson-Weller, A., 1989, Phosphatidylinositol-specific phospholipase C, a possible virulence factor of *Staphylococcus aureus*. *J. Clin. Microbiol.*, 27:2451–2454.
45. Marquis, H., Goldfine, H., and Portnoy, D.A., 1997, Proteolytic pathways of activation and degradation of a bacterial phospholipase C during intracellular infection by *Listeria monocytogenes*. *J. Cell. Biol.*, 137:1381–1392.
46. Martínez-Morales, F., Schobert, M., Lopez-Lara, I.M., and Geiger, O., 2003, Pathways for phosphatidylcholine biosynthesis in bacteria. *Microbiology*, 149:3461–3471.
47. Maulik, N., Kagan, V.E., Tyurin, V.A., and Das, D.K., 1998, Redistribution of phosphatidylethanolamine and phosphatidylserine precedes reperfusion-induced apoptosis. *Am. J. Physiol.*, 274:H242–H248.
48. McDermott, M., Wakelam, M.J., and Morris, A.J., 2004, Phospholipase D. *Biochem. Cell. Biol.*, 82:225–253.
49. Meyer, K.C., Sharma, A., Brown, R., Weatherly, M., Moya, F.R., Lewandoski, J., and Zimmerman, J.J., 2000, Function and composition of pulmonary surfactant and surfactant-derived fatty acid profiles are altered in young adults with cystic fibrosis. *Chest*, 118:164–174.
50. Meyer, S.G., and de Groot, H., 2003, [14C] serine from phosphatidylserine labels ceramide and sphingomyelin in L929 cells: evidence for a new metabolic relationship between glycerophospholipids and sphingolipids. *Arch. Biochem. Biophys.*, 410:107–111.
51. Nagiec, M.J., Lei, B., Parker, S.K., Vasil, M.L., Matsumoto, M., Ireland, R.M., Beres, S.B., Hoe, N.P., and Musser, J.M., 2004, Analysis of a novel prophage-encoded group A Streptococcus extracellular phospholipase A(2). *J. Biol. Chem.*, 279:45909–45918.
52. Naka, T., Fujiwara, N., Yabuuchi, E., Doe, M., Kobayashi, K., Kato, Y., and Yano, I., 2000, A novel sphingoglycolipid containing galacturonic acid and 2-hydroxy fatty acid in cellular lipids of Sphingomonas yanoikuyae. *J. Bacteriol.*, 182:2660–2663.
53. Ochsner, U.A., Snyder, A., Vasil, A.I., and Vasil, M.L., 2002, Effects of the twin-arginine translocase on secretion of virulence factors, stress response, and pathogenesis. *Proc. Natl. Acad. Sci. U.S.A.*, 99:8312–8317.
54. Okino, N., Ichinose, S., Omori, A., Imayama, S., Nakamura, T., and Ito, M., 1999, Molecular cloning, sequencing, and expression of the gene encoding alkaline ceramidase from *Pseudomonas aeruginosa*. Cloning of a ceramidase homologue from Mycobacterium tuberculosis. *J. Biol. Chem.*, 274:36616–36622.
55. Osborne, J.C., Jr., Stanley, S.J., and Moss, J., 1985, Kinetic mechanisms of two NAD:arginine ADP-ribosyltransferases: the soluble, salt-stimulated transferase from turkey erythrocytes and choleragen, a toxin from *Vibrio cholerae*. *Biochemistry*, 24:5235–5240.
56. Ostroff, R.M., Vasil, A.I., and Vasil, M.L., 1990, Molecular comparison of a nonhemolytic and a hemolytic phospholipase C from *Pseudomonas aeruginosa*. *J. Bacteriol.*, 172:5915–5923.
57. Ostroff, R.M., and Vasil, M.L., 1987, Identification of a new phospholipase C activity by analysis of an insertional mutation in the hemolytic phospholipase C structural gene of *Pseudomonas aeruginosa*. *J. Bacteriol.*, 169:4597–601.
58. Pankhaniya, R.R., Tamura, M., Allmond, L.R., Moriyama, K., Ajayi, T., Wiener-Kronish, J.P., and Sawa, T., 2004, *Pseudomonas aeruginosa* causes acute lung injury via the catalytic activity of the patatin-like phospholipase domain of ExoU. *Crit. Care Med.*, 32:2293–2299.
59. Phillips, R.M., Six, D.A., Dennis, E.A., and Ghosh, P., 2003, In vivo phospholipase activity of the *Pseudomonas aeruginosa* cytotoxin ExoU and protection of mammalian cells with phospholipase A2 inhibitors. *J. Biol. Chem.*, 278:41326–41332.
60. Pritchard, A.E., and Vasil, M.L., 1986, Nucleotide sequence and expression of a phosphate-regulated gene encoding a secreted hemolysin of *Pseudomonas aeruginosa*. *J. Bacteriol.*, 167:291–298.
61. Rabin, S.D., and Hauser, A.R., 2005, Functional regions of the *Pseudomonas aeruginosa* cytotoxin ExoU. Infect. *Immun.*, 73:573–582.

62. Reilly, T.J., Baron, G.S., Nano, F.E., and Kuhlenschmidt, M.S., 1996, Characterization and sequencing of a respiratory burst-inhibiting acid phosphatase from *Francisella tularensis*. *J. Biol. Chem.*, 271:10973–10983.
63. Rosenau, F., and Jaeger, K., 2000, Bacterial lipases from *Pseudomonas*: regulation of gene expression and mechanisms of secretion. *Biochimie*, 82:1023–1032.
64. Rosenzweig, A., 2003, Endothelial progenitor cells. *N. Engl. J. Med.*, 348:581–582.
65. Sato, H., and Frank, D.W., 2004, ExoU is a potent intracellular phospholipase. *Mol. Microbiol.*, 53:1279–1290.
66. Sato, H., Frank, D.W., Hillard, C.J., Feix, J.B., Pankhaniya, R.R., Moriyama, K., Finck-Barbancon, V., Buchaklian, A., Lei, M., Long, R.M., Wiener-Kronish, J., and Sawa, T., 2003, The mechanism of action of the *Pseudomonas aeruginosa*-encoded type III cytotoxin, ExoU. *EMBO. J.*, 22:2959–2969.
67. Schmiel, D.H., Wagar, E., Karamanou, L., Weeks, D., and Miller, V.L., 1998, Phospholipase A of *Yersinia enterocolitica* contributes to pathogenesis in a mouse model. *Infect. Immun.*, 66:3941–3951.
68. Schulert, G.S., Feltman, H., Rabin, S.D., Martin, C.G., Battle, S.E., Rello, J., and Hauser, A.R., 2003, Secretion of the toxin ExoU is a marker for highly virulent *Pseudomonas aeruginosa* isolates obtained from patients with hospital-acquired pneumonia. *J. Infect. Dis.*, 188:1695–1706.
69. Schuster, M., Lostroh, C.P., Ogi, T., and Greenberg, E.P., 2003, Identification, timing, and signal specificity of *Pseudomonas aeruginosa* quorum-controlled genes: a transcriptome analysis. *J. Bacteriol.*, 185:2066–2079.
70. Schweizer, H.P., Jump, R., and Po, C., 1997, Structure and gene-polypeptide relationships of the region encoding glycerol diffusion facilitator (*glpF*) and glycerol kinase (*glpK*) of *Pseudomonas aeruginosa*. *Microbiology*, 143:1287–1297.
71. Schweizer, H.P., and Po, C., 1994, Cloning and nucleotide sequence of the glpD gene encoding sn-glycerol-3-phosphate dehydrogenase of *Pseudomonas aeruginosa*. *J. Bacteriol.*, 176:2184–2193.
72. Serra, A.L., Mariscotti, J.F., Barra, J.L., Lucchesi, G.I., Domenech, C.E., and Lisa, A.T., 2002, Glycine betaine transmethylase mutant of *Pseudomonas aeruginosa*. *J. Bacteriol.*, 184:4301–4303.
73. Shaver, C.M., and Hauser, A.R., 2004, Relative contributions of *Pseudomonas aeruginosa* ExoU, ExoS, and ExoT to virulence in the lung. *Infect. Immun.*, 72:6969–6977.
74. Shortridge, V.D., Lazdunski, A., and Vasil, M.L., 1992, Osmoprotectants and phosphate regulate expression of phospholipase C in *Pseudomonas aeruginosa*. *Mol. Microbiol.*, 6:863–871.
75. Six, D.A., and Dennis, E.A., 2000, The expanding superfamily of phospholipase A(2) enzymes: classification and characterization. *Biochim. Biophys. Acta.*, 1488:1–19.
76. Stonehouse, M.J., Cota-Gomez, A., Parker, S.K., Martin, W.E., Hankin, J.A., Murphy, R.C., Chen, W., Lim, K.B., Hackett, M., Vasil, A.I., and Vasil, M.L., 2002, A novel class of microbial phosphocholine-specific phospholipases C. *Mol. Microbiol.*, 46:661–676.
77. Stonehouse, M.J., and Vasil, M.L., 2005, unpublished observations.
78. Stuckey, J.A., and Dixon, J.E., 1999, Crystal structure of a phospholipase D family member. *Nat. Struct. Biol.*, 6:278–284.
79. Tamura, M., Ajayi, T., Allmond, L.R., Moriyama, K., Wiener-Kronish, J.P., and Sawa, T., 2004, Lysophospholipase A activity of *Pseudomonas aeruginosa* type III secretory toxin ExoU. *Biochem. Biophys. Res. Commun.*, 316:323–331.
80. Teplitz, C., 1965, Pathogenesis of *Pseudomonas vasculitis* and septic legions. *Arch. Pathol.*, 80:297–307.
81. Tropp, B.E., 1997, Cardiolipin synthase from *Escherichia coli*. *Biochim. Biophys. Acta.*, 1348:192–200.

82. Vance, D.E., and J.E., V., 2002, Biochemistry of Lipids, Lipoproteins and Membranes, 4th ed., Vol. 36. Elsevier, Boston.
83. Vance, R.E., Hong, S., Gronert, K., Serhan, C.N., and Mekalanos, J.J., 2004, The opportunistic pathogen *Pseudomonas aeruginosa* carries a secretable arachidonate 15-lipoxygenase. *Proc. Natl. Acad. Sci. U.S.A.*, 101:2135–2139.
84. Vasil, M.L., Berka, R.M., Gray, G.L., and Nakai, H., 1982, Cloning of a phosphate-regulated hemolysin gene (phospholipase C) from *Pseudomonas aeruginosa*. *J. Bacteriol.*, 152:431–440.
85. Vasil, M.L., Graham, L.M., Ostroff, R.M., Shortridge, V.D., and Vasil, A.I., 1991, Phospholipase C: molecular biology and contribution to the pathogenesis of *Pseudomonas aeruginosa*. *Antibiot. Chemother.*, 44:34–47.
86. Voulhoux, R., Ball, G., Ize, B., Vasil, M.L., Lazdunski, A., Wu, L.F., and Filloux, A., 2001, Involvement of the twin-arginine translocation system in protein secretion via the type II pathway. *EMBO J.*, 20:6735–6741.
87. Wadsworth, S.J., and Goldfine, H., 2002, Mobilization of protein kinase C in macrophages induced by *Listeria monocytogenes* affects its internalization and escape from the phagosome. *Infect. Immun.*, 70:4650–4660.
88. Wagner, V.E., Bushnell, D., Passador, L., Brooks, A.I., and Iglewski, B.H., 2003, Microarray analysis of *Pseudomonas aeruginosa* quorum-sensing regulons: effects of growth phase and environment. *J. Bacteriol.*, 185:2080–2095.
89. Wilderman, P.J., Vasil, A.I., Johnson, Z., and Vasil, M.L., 2001, Genetic and biochemical analyses of a eukaryotic-like phospholipase D of *Pseudomonas aeruginosa* suggest horizontal acquisition and a role for persistence in a chronic pulmonary infection model. *Mol. Microbiol.*, 39:291–303.
90. Wilderman, P.J., Vasil, A.I., Martin, W.E., Murphy, R.C., and Vasil, M.L., 2002, *Pseudomonas aeruginosa* synthesizes phosphatidylcholine by use of the phosphatidylcholine synthase pathway. *J. Bacteriol.*, 184:4792–4799.
91. Winkler, H.H., and Daugherty, R.M., 1989, Phospholipase A activity associated with the growth of *Rickettsia prowazekii* in L929 cells. *Infect. Immun.*, 57:36–40.
92. Wolfgang, M.C., Jyot, J., Goodman, A.L., Ramphal, R., and Lory, S., 2004, *Pseudomonas aeruginosa* regulates flagellin expression as part of a global response to airway fluid from cystic fibrosis patients. *Proc. Natl. Acad. Sci. U.S.A.*, 101:6664–6668.
93. Wolfgang, M.C., Lee, V.T., Gilmore, M.E., and Lory, S., 2003, Coordinate regulation of bacterial virulence genes by a novel adenylate cyclase-dependent signaling pathway. *Dev. Cell.*, 4:253–263.
94. Yuki, N., Awai, K., Matsuda, T., Yoshioka, Y., Takamiya, K., and Ohata, H., 2004, A novel phosphatidylcholine-hydrolyzing phospholipase C induced by phosphate starvation in *Arabidopsis*. *J. Biol. Chem.*, 280:7469–7476.
95. Zabner, J., Smith, J.J., Karp, P.H., Widdicombe, J.H., and Welsh, M.J., 1998, Loss of CFTR chloride channels alters salt absorption by cystic fibrosis airway epithelia in vitro. *Mol. Cell.*, 2:397–403.

4

IN VIVO FUNCTIONAL GENOMICS OF *PSEUDOMONAS*: PCR-BASED SIGNATURE-TAGGED MUTAGENESIS

Roger C. Levesque

Microbiologie Moléculaire et Génie des Protéines
Département de Biologie Médicale et Pavillon
Charles-Eugène Marchand
Faculté de médecine Université Laval
Sainte-Foy, Québec
G1K 7P4, Canada

1. INTRODUCTION

The genus *Pseudomonas* encompasses closely related bacterial species that behave as pathogens in specific situations and can cross a large evolutionary gulf to naturally or experimentally infect organisms from other phyla or kingdoms. Hence, *Pseudomonas* can fill practically any ecological niche including survival in inhospitable Mars-like environments, in livings hosts such as unicellular life-forms, in plants, in insects, in animals, and in humans.[74,89] In humans, *Pseudomonas aeruginosa* has been considered an opportunistic pathogen; in most cases the opportunity being a conditional trauma such as immunosuppressed, cancer, and burn patients.[25–27] Unusual situations in humans are also excellent opportunities for *P. aeruginosa* to initiate and establish a nosocomial infection such as in the hospital intensive care units where catheters, respiratory devices, and other medical implants are used in combination with an arsenal of immunomodulators and antibiotics.[31,56] In addition, there is a specific and unusual human medical condition in Cystic Fibrosis where *Pseudomonas* species,

Pseudomonas, Volume 4, edited by Juan-Luis Ramos and Roger C. Levesque

mostly *P. aeruginosa* will cause a permanent chronic lung infection leading to irreversible tissue damage and death.[71,93,95]

Recent analysis of the 67% G + C 6.3 Mb genome content of *P. aeruginosa* strain PAO1 has revealed a repertoire of more than 468 open reading frames (ORFs) containing 543 regulatory motifs characteristic of transcriptional regulators, 55 sensors, 89 response regulators, and 14 sensor–response regulatory hybrids of two component systems and at least 12 potential RND efflux systems. These proteins represent approximately 8.4% of the genome coding capacity acting as sensors interacting with the environment or with the host permitting an adaptive response to aggression and to antibiotic resistance.[36,83] Of the 5570 ORFs, *P. aeruginosa* strain PAO1 encodes 1780 (32%) genes having no homology to any previously reported sequences, 1590 (28.5%) genes having a function proposed based on the presence of conserved amino acid motif, structural feature or limited homology and 769 (13.8%) homologues of previously reported genes of unknown function. In terms of genes characterized, 1059 (19%) have a function based upon a strongly homologous gene experimentally demonstrated in another organism; whereas only 372 genes (6.7%) have a function experimentally demonstrated in *P. aeruginosa*. Hence, a massive amount of information implicated in virulence and for *in vivo* maintenance of *P. aeruginosa* is hiding in more than 1780 hypothetical and unknown proteins, in 1590 genes having conserved motifs and in 769 homologues of previously reported genes of unknown function. This intractable problem can be addressed by signature-tagged mutagenesis (STM).

A clear understanding how *Pseudomonas* has the capability to initiate, invade, maintain, and colonize a mammalian host to assure its survival is a prerequisite for understanding the virulence behavior of this organism in particular, and all opportunistic pathogens in general.[3,11,21,28,34] Indeed, novel strategies are essential for identifying genes expressed solely in the host during the infection process and also genes which are absolutely essential for initiating and maintaining *Pseudomonas* in the host.[19–21,42] Unfortunately, traditional screening in animal models of infection for mutants covering a complete genome and based upon a gene by gene mutational approach is not feasible *in vivo*, even with today's capabilities in genomics and in proteomics. For example, a significant analysis of virulence determinants for the *P. aeruginosa* 6.3 Mb genome encoding 5570 ORFs would require in a model of infection a minimum of 5570 animals; statistical validity would recommend groups of at least five individuals giving a total of 27,850 animals, an impossible and unjustifiable task.

Identification of function loss and hence of genes essential in the infection process can now be addressed by STM. This is an extremely powerful and elegant bacterial genetics approach for *in vivo* functional genomics, particularly when used in combination with bioinformatics, proteomics, and transcriptome analysis to identity genes and their products essential for *in vivo* maintenance.[20,33,34,40,87,88]

The technology has become accessible to a larger number of small laboratories, especially as screening can now be done rapidly by using PCR instead of hybridization. The design of tags has been simplified, and several different mini-Tn5s, each with a unique phenotypic selection, can now be utilized.[50,51] We will not discuss all the technical details here but will focus on the concept and results available to date with *P. aeruginosa*. The reader is referred to more technical publications for the PCR-based STM.[50,51] These modifications of STM, developed in our laboratory, have been applied to the study of *P. aeruginosa* strain PAO1.

In this chapter, we focus on the biological significance of the data obtained in comparisons with other *in vivo* methods such as IVET, transcriptome, and proteome analysis.[33,67] Application of IVET to *Pseudomonas* has been described in Volume I, Chapter 11. In studying *in vivo* maintenance and gene expression of *P. aeruginosa*, STM seemed more appropriate than IVET because the negative selection identifies unique genes essential for *in vivo* survival. In contrast, IVET identifies a large portion of housekeeping genes upregulated in these conditions. These genes certainly play significant role for survival *in vivo* and IVET is a complementary approach and we, and others, have applied both approaches to *Pseudomonas*; in every case there seemed to be little overlaps in the results. Unfortunately, a highly versatile organism like *Pseudomonas* has alternative metabolic pathways to compensate IVET mutant deficiencies.

2. SIGNATURE-TAGGED MUTAGENESIS

Transposable elements have been traditional tools in bacterial genetics for creating insertional mutations. Over the years, the limitations of transposons use in *Pseudomonas* such as bacterial host specificity and transfer and their caveats including problems of multiple insertions, chromosomal rearrangements, and in many cases, polar effects were eliminated by the development of mini-transposons, mobilizable suicide plasmid vectors, and optimized protocols in transfer using bacterial conjugation and electroporation.[23,79,80,82] In addition to the development of mini-transposons, completely *in vitro* transposition methods are also available.[23,32,77]

STM, developed 10 years ago by David Holden and colleagues, represents a major application of an "en masse" transposon mutagenesis and high-throughput screening technique.[38,81] STM is a mutation-based screening method that uses a population of bacterial mutants for the identification of multiple virulence genes of microbial pathogens by negative selection. The technique depends upon *in vivo* selection of virulent organisms; while those mutants whose virulence genes are inactivated will not persist *in vivo*. This

allows a relatively rapid, unbiased search for virulence genes, using experimental infection models as hosts to select against strains carrying mutations in genes affecting virulence and bacterial maintenance. The beauty of the technology is an *in vivo* selection process done by the host among a mixed population of mutants. Hence, to avoid the typical labor-intensive screening of individual mutants, each mutant is tagged with a different and unique DNA signature; the power of STM is that it allows large numbers of different mutants to be screened simultaneously in the same host. There is now enough data, experience, and knowledge from more than 10 years of work and from 40 distinct reports in different bacterial systems to indicate that STM has numerous applications.[4] The power of the technique relies mostly on the rounds of selection done in an appropriate model being an animal, plant, insect, cell line, or environmental condition selected. In several cases, different host models can be tested with the same collection of mutants.[49]

In the classical STM, a comparative hybridization technique was used with a collection of transposons each modified by the incorporation of a distinct DNA tag sequence. The concept was that when the tagged transposon integrates into the bacterial chromosome, each individual mutant can be distinguished from every other by hybridization. The first tag collections were designed as short DNA segments containing 40-bp variable regions giving probe specificity flanked by invariant arms that facilitated the co-amplification and radioactive labeling of the central portions by PCR, and which were subsequently used as hybridization probes.[38] Colony or DNA dot blots were prepared from these mutants, and compared by hybridization to DNA prepared when the same pool of strains passaged in a model of infection. PCR was then used to prepare labeled probes, representing the tags present in a defined inoculum (*in vitro* input) and recovered from the host (*in vivo* output). Hybridization of the tags from the input and the output pools were compared. This permitted the identification of mutants that failed to grow *in vivo*, because these tags from mutants effective for *in vivo* maintenance will not be present in the output pools. These mutants can then be identified and recovered from the original bacterial arrays. The nucleotide sequence of genomic DNA flanking the transposon insertion site can be determined and the inactivated gene identified.

Because of inherent cross-hybridization signals between tags incorporated into mini-Tn5 transposons, their suitability was checked prior to use by hybridization of amplified labeled tags to DNA colony blots of mutants used to generate the probes. Mutants from the *in vitro* pool whose tags failed to yield clear signals were discarded; while those that gave good signals were assembled into new pools for screening into animals. This careful analysis was done prior to STM, so as to diminish the inherent problems of hybridization caused by problematic tags.[64]

2.1. Description of the PCR-Based STM

The PCR-based STM is an important variant of the hybridization method, easier to use and can be divided into two major steps. In contrast to the STM done by hybridization, the first *in vitro* step involved the construction of defined complementary oligonucleotides assembled to form double-stranded DNA tags which are specific for PCR amplification, do not cross-amplify, and are optimized for PCR multiplex and polymerase priming.[51] These tags are cloned as complementary oligonucleotides into a mini-transposable element; transposons are selected with different antibiotic markers as a reliable method to select for the recipients; and mutants are carefully assembled in an array of tagged mutants.[50] The PCR-based STM method developed uses 24 pairs of complementary 21-bp oligonucleotides. These pairs are annealed to generate 24 double-stranded DNA tags, which are then cloned into each of the four plasmid-located mini-Tn5 derivatives, mini-Tn5Km2, mini-Tn5Tc, mini-Tn5Cm, and mini-Tn5TcGFP.[50] These Tn5-derived mini-transposons contain genes encoding kanamycin resistance (Km), tetracycline resistance (Tc), chloramphenicol (Cm) resistance, and green fluorescent protein (GFP), respectively. All plasmids have unique cloning sites for tag insertion, and are flanked by 19-bp terminal repeats encoding the I and the O ends for transposase action and are capable of only a single transposition event. The transposons are located on the pUT plasmid vector, an R6K-based suicide delivery plasmid where the Pi protein is furnished in *trans* by the donor bacteria. The PCR-based STM comprises a collection of 96 (24 tags $\times$ 4 mini-Tn5 plasmid vectors) tagged transposons (available upon request).

The second step requires an experimental infection host system such as an animal or cell for *in vivo* screening of the library.[70] We strongly believe that the model selected for screening the "output pool" of STM mutants for negative selection is the crux of STM. In this second phase, the experimental design needs to take into account: (1) the power and limitations of the model of infection used, (2) the number of rounds of STM screening to be done, and (3) the number of different hosts to be utilized. When performing analysis by STM, it is crucial to take into account crucial parameters including: (4) the inoculum size (from 1 to 10^6bacteria/ml), (5) the number of STM mutants to be utilized in each experimental host (from 10 to 96), (6) the route for initiation of the infection, and (7) the duration of selection of STM mutant which may vary from a few hours to several weeks.[6,8,35,60,65] Once this is determined and the selection process completed, crucial data must be obtained by PCR, negative clones selected by multiplex and single PCR analysis, determination of the DNA sequence around the site of the insertion mutation, which may be obtained by cloning the transposon marker or by RT-PCR.

The final phase, and the ultimate goal of STM, is to characterize genes shown to play a significant role in virulence, and to define their function and

role in pathogenicity. Hence, STM also involves a systematic characterization of single mutants selected *in vivo* by construction of gene knockouts giving a clear genetic background so as to eliminate polar effects. Additional analysis involves bacterial cell growth and maintenance *in vitro* for auxotrophy but more so *in vivo* by measuring the growth index and or the competitive index.[8] To make fast progress in analysis of selected *Pseudomonas* STM mutants, several transposon mutant libraries of all PAO1 genes are now available upon requests as *Pseudomonas* or as BAC clones for analysis of specific genomic regions of interest.[24,41,53] The general scheme for PCR-based STM is presented in Fig. 1.

An alternative to the clone by clone analysis of STM mutants selected after screening, studies can now be combined with transcriptomics using the Affymetrix *P. aeruginosa* strains PAO1 DNA oligonucleotide chip; proteomics and IVET studies are also complementary prospects. Metabolomics can also be attempted with *Pseudomonas* STM mutants using the Phenotype MicroArrays system, a high-throughput technology for simultaneous testing of 700 cellular phenotypes by merely pipetting a cell suspension into microplate arrays.[10,58,67,78]

3. *IN VIVO* STM OF *P. AERUGINOSA*

Our interest in *P. aeruginosa* STM has been prompted by several questions: what are the patterns of initiation, growth, establishment, and maintenance for progression of a chronic infection? What are the bacterial growth rates and population sizes at these different stages? What nutrients and substrates become available as infection progresses? How do bacteria metabolize them? What virulence determinants are produced at these different and specific periods during infection? How is the production of virulence determinants regulated? What are the effects of environmental parameters? Again one must remember that studies, including those done by STM must cover part of Koch's postulates. This becomes a difficult case when studying opportunistic pathogens including *Pseudomonas* because no adequate experimental models exist to represent infected host.[59]

To date, only two studies have been done by STM using *P. aeruginosa* strains PAO1 and TBFC10389. PAO1 was screened in the rat lung model of chronic infection and in *Drosophila*; while TBFC10389 was screened in tissue culture using polymorphonuclear (PMN) granulocytes.[40,49,50,51,52,72,91]

3.1. Global Analysis and Distribution of STM Mutants

PCR-based STM was used for high-throughput screening of a collection of 7968 *P. aeruginosa* mutants in a rat model of chronic respiratory infection and is summarized in Figure 1 and in Figure 2.[18,52] As depicted in Figure 2

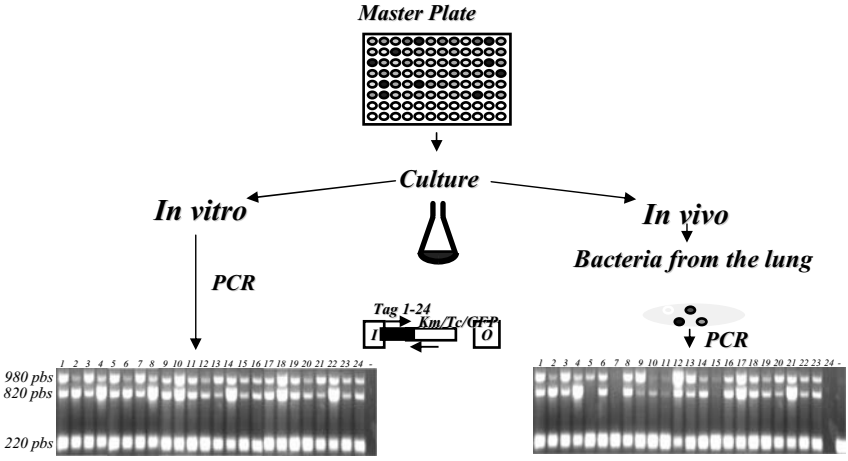


Figure 1. General scheme for the PCR-based STM using master plates assembled from a collection of 7968 *P. aeruginosa* clones carrying insertions of the mini-Tn5 Km, Tc or GFP markers.[49,50,51,52,72]

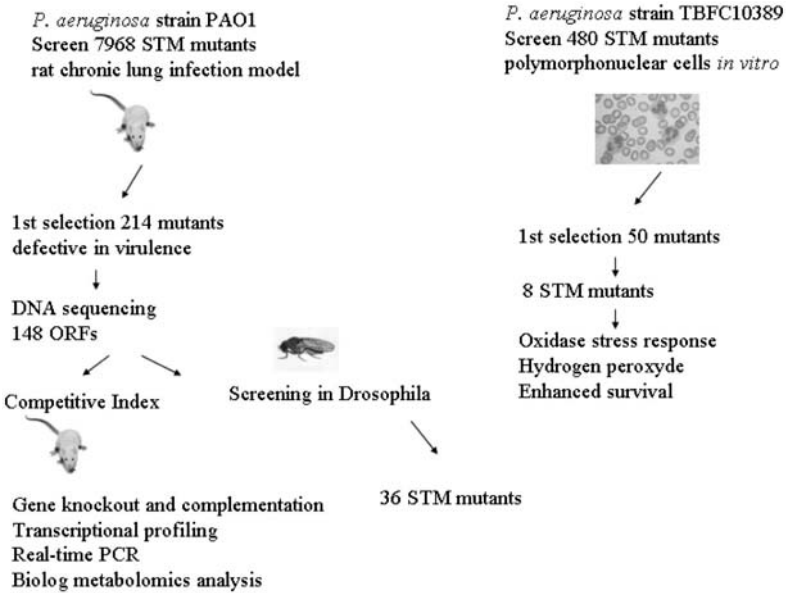


Figure 2. General scheme used for screening *P. aeruginosa* STM mutants using different models of infection.[49,50,51,52,72,91]

and after three rounds of screening, a total of 214 mutants were shown to be attenuated in lung infection and were retained for further studies.

Analysis of the genome distribution of 160 unique STM mutants represented transposition event into 148 distinct ORFs. Of significance in virulence studies and indicating the power of STM for studying *in vivo* functional genomics was that many *P. aeruginosa* STM insertions were found into genes expressing hypothetical, unknown, unclassified proteins. Indeed, this first group represented 67 of 148 insertion events (42.6%). The second interesting group included 26 proteins implicated in the transport of small molecules (9.5%), 7 proteins were secreted factors (4.8%), and 7 transcriptional regulators (4.8%). Few insertions were found in genes coding for classes of protein involved in adaptation and protection (2), amino acid biosynthesis and metabolism (4), carbon compound catabolism (2), cell division (1), cell wall and lipopolysaccharide biosynthesis (1), central intermediary metabolism (3), chaperone and heat shock (1). Bioinformatics analysis for protein localization indicated that 40% of *P. aeruginosa* STM mutants defective *in vivo* had defects in proteins localized at or near the surface of the bacterial cell such as the inner membrane, the periplasm, and the outer membrane; while 15% of STM mutants identified proteins whose localizations were unknown.[53] Overall, the number of *in vivo* attenuated STM mutants identified represents ≈2% of the total 7968 *P. aeruginosa* STM mutant strains screened; this value is lower than the 4%–10% previously reported in different bacterial systems and may be due to the highly stringent conditions used when selecting STM clones by PCR screening.[35,38,65] The complete list of STM mutants identified to date as essential for lung infection is available at: http://rclevesque.rsvs.ulaval.ca/TablewebsiteTMlist.pdf

3.2. Detailed Analysis of *P. aeruginosa* STM mutants

One of the gold standards of STM is to clearly demonstrate the capability of the method in identifying previously characterized virulence factors in a known bacterial system such as *P. aeruginosa*. Bacterial cellular processes essential for *in vivo* survival and known to be crucial for *P. aeruginosa* pathogenesis are intimately involved in motility and attachment such as pili and type IV fimbriae.[2,62] Bacterial flagellins are mediators of pathogenicity and host immune responses in mucosa.[76] The flagellum is an exquisitely engineered chemi-osmotic nanomachine; nature's most powerful rotary motor, harnessing a trans-membrane ion-motive force to drive a filamentous propeller.[68] Flagella contribute to the virulence of pathogenic bacteria through chemotaxis, adhesion to and invasion of host surfaces. Flagellin is the structural protein that forms the major portion of flagellar filaments. Thus, flagellin consists of a

Table 1. *P. aeruinosa* genes essential for lung infection identified by STM and previously reported as essential for virulence[72]

Strain name	PA position	Gene name[a]	Description	Homologue[b]	Localization[c]
STM4528	PA4528	*pilD*	Motility and attachment	100% PilD *P. aeruginosa*	IM
STM0410	PA0410	*pilI*	Motility and attachment	100% PilI *P. aeruginosa*	C
STM4554	PA4554	*pilY1*	Motility and attachment	100% PilY1 *P. aeruginosa*	IM
STM0762	PA0762	*algU*	Sigma factor	100% AlgU *P. aeruginosa*	C
STM4446	PA4446	*algW*	Secretion of alginate	100% AlgW *P. aeruginosa*	P
STM0765	PA0765	*mucC*	Transcription regulator	100% MucC *P. aeruginosa*	U
STM1248	PA1248	*aprf*	Alkaline protease	100% AprF *P. aeruginosa*	OM
STM3478	PA3478	*rhlB*	Rhamnosyl transferase	100% RhlB *P. aeruginosa*	C
STM3831	PA3 831	*pepA*	Leucine amino peptidase	100% PhpA *P. aeruginosa*	C
STM5112	PA5112	*estA*	Esterase	69% YtrP *P. putida*	OM
STM5449	PA5449	*wbpX*	Glycosyl transferase	100% WpbX *P. aeruginosa*	C

C, cytoplasmic; IM, inner membrane; P, periplasm; OM, outer membrane; U, unknown.

[a] The gene name assigned by the annotation and sequencing group at the http://www.pseudomonas.com internet site.

[b] The name of the protein homologue, when available with recent BLAST analysis and updating data from the http://www.pseudomonas.com site. There are four classes (1) Function experimentally demonstrated in *P. aeruginosa*; (2) Function of highly similar gene experimentally demonstrated in another organism, and gene context consistent of pathways it is involved in, if known; (3) Function proposed based on presence of conserved amino acid motif, structural feature, or limited sequence similarity to an experimentally studied gene; (4) Homologues of previously reported genes of unknown function, or no similarity to any previously reported sequences.

[c] Protein localization was determined using PSORTB.

conserved domain that is widespread in bacterial species and is dedicated to filament polymerization. Conversely, mammalian hosts detect the conserved domain on flagellin monomers through Toll-like receptor (TLR) 5, which triggers proinflammatory and adaptive immune responses.[76] In agreement with these observations and as shown in Table 1, *P. aeruginosa* STM mutants screened in the rat lung had insertions in *pilD* (PA4528), *pilI* (PA0410), and *pilY1* (PA4554), respectively.

Significant for *in vivo* maintenance and for the production of biofilms, *P. aeruginosa* is notorious for the copious amounts of alginate and mucus it

produces in the lungs of cystic fibrosis patients.[17] Alginate forms a physical/chemical barrier protecting the bacterium from its environment, behaves as a virulence factor by acting as an antiphagocytic agent, impairs the efficacy of aminoglycoside antibiotics, and it suppresses neutrophil and lymphocyte function. Environmental stresses such as oxidative stress triggers dramatic increases in alginate biosynthesis and secretion, favors the appearance of the mucoid phenotype which has devastating effects on antibiotic treatment in CF individuals. Thus, the recovery of attenuated *P. aeruginosa* STM strains carrying mutations in *algU* (PA0762), *algW* (PA4446), and in *mucC* (PA0765), the functions of which are closely associated with alginate biosynthesis, its secretion and transcriptional regulation are highly relevant to validate STM.[12,54,94]

P. aeruginosa in a repertoire of extracytoplasmic factors implicated in virulence and *in vivo* maintenance. The need for iron *in vivo* is critical and mediated by siderophores or iron (III) chelators produced by *P. aeruginosa* in response to iron limitation, and their structure can be extremely diverse.[85] The production of PVDs and their cognate receptors depends on two extracytoplasmic sigma factors (ECF-σ), PvdS, and FpvI. PvdS also controls the transcription of two *P. aeruginosa* extracellular virulence factors, the endoprotease PrpL and exotoxin.[86] Recently, a microarray analysis of the *P. aeruginosa* genes transcribed under conditions of iron limitation in wild-type and in a Δ*pvdS* mutant led to the discovery of new PvdS-regulated genes for the biosynthesis of PVDs.[69] It has been suggested that extracellular virulence determinants might not be identified by STM because they can be complemented by other mutants in the pool.[66] We have shown that *P. aeruginosa* STM mutants carried insertions in a plethora of secreted factors and toxins known to be critical in virulence including an alkaline protease (PA1248), a rhamnosyl surfactant (PA3478), an amino peptidase (PA3831), an esterase (PA5112), and a glycosyl transferase (PA5449) involved in LPS biosynthesis.[7,39,55,61,72]

After confirming the attenuation of virulence with several rounds of screening in a selected host, one has to find a strategy to evaluate the significance of STM mutants isolated. One simple approach is to simply determine the expression of known virulence factors from STM mutants isolated. For *P. aeruginosa* STM mutants, this was done to determine whether the 148 STM strains had defects in the expression of known virulence factors such as proteases, lipases, phospholipases, exopolysaccharides, pyoverdine, pyocyanin, and motility.[5,9,13–16,43] From the 148 STM strains tested, 36 were defective in the production of known virulence factors such as the proteases LasA and LasB (PA2895 and PA3498), pyocyanine (PA1157, PA2639, PA4887, PA5312), and pyoverdin (PA5441). Of those, the majority were defective in the swimming and swarming phenotype, and 14 STM mutants were found to have insertions in hypothetical or unknown proteins.[44,45,63,72]

3.3. Screening in Drosophila and in PMN Cells

Because of the repertoire of host models of infection available, *P. aeruginosa* STM mutants can be categorized further by *in vivo* screening in alternative hosts.[1,37,42,47] Selected *P. aeruginosa* mutants obtained by various genetic methods have been screened in a repertoire of hosts including plants such as *Arabidopsis thaliana* and lettuce, the amoeba *Dictyostelium discoideum*, the nematode *Caenorhabditis elegans*, the wax moth *Bombix mori*, the fly *Drosophila melanogaster* and mammalian hosts such as various strains of mouse and rat.[22,47,57,75] Certain gene products may be directly or indirectly important for initiation or maintenance of the infection and would presumably be niche-dependent or may be expressed specifically in certain host tissues only. In these various models, if the duration of the infection for STM screening *in vivo* is short presumably representing an acute infection, genes important for establishment of the infection will be found or, if the duration of infection is long or chronic, genes important for maintenance of infection could be identified.

For example, the collection of 148 *P. aeruginosa* STM mutants defective for maintenance in the rat lung were screened in *Drosophila* and results are shown in Table 2; 8 mutants were found to be highly attenuated giving less than 15% lethality.[72] In *Drosophila*, these *P. aeruginosa* STM mutants were found to have insertions mostly in genes encoding amino acid production, nucleotide biosynthesis, and in enzymes of central metabolism. Two STM mutants had insertions in *pilI* (PA0410) and one in a hypothetical gene (PA5441), but affecting pyoverdine overproduction.

There is a precedent for particular interest of STM mutants defective in amino acid biosynthesis and a relationship to *in vivo* maintenance. An *aroA* deletion mutant of *P. aeruginosa* strain PAO1 showing auxotrophy for aromatic amino acids (as verified by the inability to grow on minimal media unless supplemented with aromatic amino acids) has been defined as an excellent candidate for intranasal vaccine. When evaluated for safety and immunogenicity in mice, the PAO1Δ*aroA* strain could be applied either intranasally or intraperitoneally at doses up to 5×10^9 CFU per mouse without adverse effects.[73] Pyoverdine but more so pyocyanin is a blue redox-active secondary metabolite that is produced in large quantities in sputum from infected patients with CF.[48] Pyocyanin is an evolutionary conserved virulence factor of *P. aeruginosa* hosts of multiple phyla and is crucial for lung infection in mice. Pyocyanin targets in yeasts have 60% mammalian orthologs which include the V-ATPase; their inhibition would presumably affect vesicular transport and protein trafficking, ATP availability, cellular respiration via electron transport, and oxidative stress leading to injury and killing of lung epithelial cells.[48]

A distinct but complementary approach to PCR-based STM is a hybridization-based STM using a 40-mers tag adapted to the 67% G + C content

Table 2. *P. aeruginosa* genes essential for lung infection and for killing of *Drosophila*[72]

STM clone PA position	Gene name	Description	Relevant characteristics
STMPA0410	*pilI*	Motility and attachment	Twitching, swarming, *Drosophila*
STMPA3831	*PepA*	Leucine amino peptidase	Swarming
STMPA2876	*PyrF*	Orotidine decarboxylase	*Drosophila*
STMPA5131	*Pgm*	Phosphoglycerate mutase	*Drosophila*
STMPA1927	*MetE*	Homocysteine methyl transferase	Swarming
STMPA0552	*Pgk*	Phosphoglycerate kinase	Lipase, pyocyanin, pyoverdine, *Drosophila*
STMPA2639	*NuoD*	NADH dehydrogenase	Pyocyanin, motility
STMPA2998	*NrqB*	Translocating oxido reductase	Swarming
STMPA3286	C, hypothetical	Hypothetical, unclassified, unknown	Swarming
STMPA4115	C, hypothetical	Hypothetical, unclassified, unknown	Swarming
STMPA4488	C, hypothetical	Hypothetical, unclassified, unknown	Swarming
STMPA4489	C, hypothetical	Hypothetical, unclassified, unknown	*Drosophila*
STMPA4491	C, hypothetical	Hypothetical, unclassified, unknown	Swarming
STMPA5441	C, hypothetical	Hypothetical, unclassified, unknown	Pyoverdine (++), *Drosophila*
STMPA4564	C, hypothetical	Hypothetical, unclassified, unknown	Swarming
STMPA2972	C, hypothetical	Hypothetical, unclassified, unknown	Swarming
STMPA3756	C, hypothetical	Hypothetical, unclassified, unknown	Swarming
STMPA4692	C, hypothetical	Hypothetical, unclassified, unknown	Swarming
STMPA5078	C, hypothetical	Hypothetical, unclassified, unknown	Swarming
STMPA3826	C, hypothetical	Hypothetical, unclassified, unknown	Swarming
STMPA1009	Hypothetical	Hypothetical, unclassified, unknown	Swarming
STMPA2895	Hypothetical	Hypothetical, unclassified, unknown	Protease
STMPA3173		Probable short-chain dehydrogenase	Swarming
STMPA3001		Probable phosphate dehydrogenase	*Drosophila*

(*Continued*)

Table 2. (*Continued*)

STM clone PA position	Gene name	Description	Relevant characteristics
STMPA3498		Probable oxido reductase	Protease
STMPA5312		Probable aldehyde dehydrogenase	Pyocyanin
STMPA5327		Probable oxido reductase	Twitching
STMPA0041		Probable hemagglutinin	Swarming
STMPA0584	*cca*	tRNA nucleotidyl transferase	Twitching, *Drosophila*
STMPA0090		Probable ClpA/B-type chaperone	Swarming
STMPA3876	*narK2*	Nitrite extrusion protein 2	Motility
STMPA4887		Probable MFS transporter	Pyocyanin (++)
STMPA0073		Probable ABC transporter	Swarming
STMPA0151		Probable TonB-dependent receptor	Biofilm
STMPA1863	*ModA*	Molybdate-binding precursor	Swarming, twitching motility
STMPA1157		Probable two-component regulator	Pyocyanin (++)

C, conserved; ++ ,indicate higher levels of pyocyanin and of pyoverdine when compared to the wild-type.

of *Pseudomonas* and cloned into the plasmoson pTnModOGm.[91] Even though the genomic sequence is unknown, the *P. aeruginosa* strain TBFC10389 was used instead of PAO1 because of its persistence in professional phagocytes.[84] PMNs were infected with 480 *P. aeruginosa* STM mutants; 50 did not survive the negative selection in PMNs and were identified as defective in the expression of a gene or operon essential for the bacterium's intracellular survival.[91] As shown in Table 3, Several rounds of screening identified 6 clones defective in PMN maintenance with insertions in PA2613 (YcaJ, conserved hypothetical), PA3344 (RecQ, helicase), PA3953 (YrcD, conserved hypothetical), PA4621 (oxidoreductase), PA5252 (YheS, ABC transporter), and PA5349 (rubredoxin reductase); two mutants showed enhanced survival in granulocytes, namely mutants with insertions in PA1992 (two component sensor) and PA5040 (PilQ Type IV fimbriae biogenesis) (Table 3). Oxidase stress response genes are essential for intracellular PMN survival; helicase is involved in the inactivation of peroxide or DNA repair; while rubredoxin reductase can substitute for superoxide dismutase and oxidates aliphatic hydrocarbons. The mutants that harbored transposons in the helicase or rubredoxin reductase genes did not grow in the presence of 0.8 mM hydrogen peroxide.

Even with such minimal screening, a comparison between the 8 mutants (Table 3) obtained from the hybridization-based STM (480 mutants) screened in PMNs, the 148 STM mutants obtained from the PCR-based STM mutants (7968) screened in the chronic rat lung model of infection and comparisons

Table 3. *P. aeruginosa* genes essential for maintenance in polymorphonuclear cells[91]

Strain name	PA position	Gene name[a]	Description	Homologue[b]	Localization[c]
STM2613	PA2613	*YcaJ*	Conserved hypothetical	74% Putative polynucleotide	U
STM3344	PA3344	*RecQ*	ATP DNA helicase	66% RecQ *E.coli*	C
STM3953	PA3953	*YrdC*	Conserved hypothetical	55% Hypothtical *B. subtilis*	U
STM4621	PA4621		Probable oxidoreductase		U
STM5252	PA5252	*YheS*	ABC transporter	70% hypothetical *E.coli*	U
STM5349	PA5349		Rubredoxin reductase	59% Rubredoxin *A. calcoaceticus*	C
STM1992	PA1992	*FlhS*	Two-component sensor	56% FlhS *P. denitrificans*	C
STM5040	PA5040	*PilQ*	Type 4 fimbrial precursor	98% PilQ *P. aeruginosa*	OM

with typical mutants defective in virulence, identified previously in various laboratories, revealed exciting features. For example, the *P. aeruginosa* 34 kDa H_2O_2-responsive transactivator OxyR encoded by PA5344 is required for full virulence in rodent and insect models of infection and for resistance to human neutrophils (Figure 3).[46] *P. aeruginosa* lacking OxyR is exquisitely susceptible to H_2O_2, even with wild-type catalase activity. In addition, the OxyR mutant cannot grow on LB agar because the autoxidizable components in the medium generate H_2O_2 at 1 μM which represents a concentration detected in peripheral blood from human donors and is sufficient to kill these organisms.[46] In STM screening by both methods, the hybridization-based STM has identified a mutant having a defect in PA5349 (rubredoxin reductase) essential for maintenance in PMNs; while the PCR-based STM has identified an insertion in PA5347 (hypothetical protein, unknown) essential for maintenance in the rat lung.[72,91] It remains to be demonstrated if both genes are part of the same operon and are regulated by the PA5344 OxyR transactivator. Additional comparisons between the STM data obtained from both methods (Figure 3) identified a crucial operon in infection; an insertion in PA4621 (probable oxidoreductase) was found essential for PMN maintenance and an insertion into PA4620 having 52% identity with a 4-hydroxybenzoyl CoA reductase was crucial for maintenance in the rat lung.[72,91] Similar comparative analysis of *P. aeruginosa* genes identified using STM, IVET and with mutants used for identification of essential virulence genes has confirmed decisively the impact of STM in virulence studies of *Pseudomonas*.

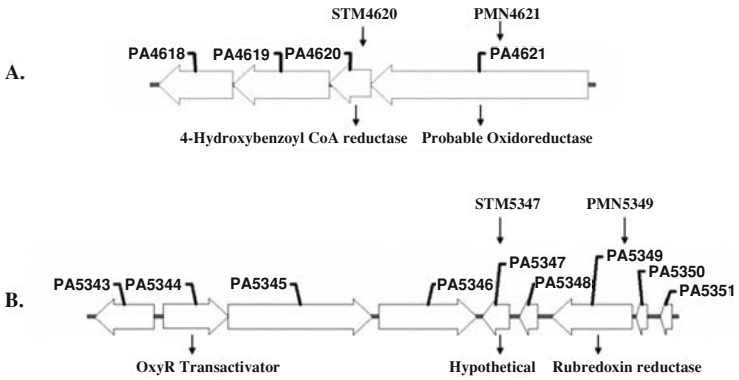


Figure 3. Genomic organization of two regions in the chromosome of *P. aeruginosa* identified as essential for maintenance in the rat lung and in polymorphonuclear cells. (A) Structure of the PA4618–PA46221 operon; (B) Gene organization of the PA5343–PA5351 region encoding the OxyR (PA5344) transactivator essential for full virulence in rodent, in Drosophila, and in PMNs. Abbreviations: STM, signature-tagged mutagenesis; PMN, Polymorphonuclear cells. Details are given in the text.

4. STM AND CHIP ANALYSIS

Since several genes found to be essential for lung infection have also been identified by other high-throughput screening methods such as microarrays and transcriptome analysis, it is crucial to compare the chip data with STM mutants found to date.[30,78,87,90,92] However, such comparisons should be done with caution. The comparisons of chip data between different laboratories, where experiments have been done in different conditions had raised initially a certain amount of controversies which can now be resolved with the experience of using *Pseudomonas* chips and additional data available.

Since the probes on the GeneChip from Affymetrix were designed in a sequence specific way, the small sequence differences between *P. aeruginosa* strains may result in several probes hybridizing poorly or not at all. When using the PA14 strain to hybridize the Affymetrix GeneChip, which was designed from the PAO1 sequence data, one should remember that the genomic sequence differences between PAO1 and PA14 are at least 5%. A BLAST result indicated that only about 75% of the probes on the Affymetrix chip are actually targeting a location in the PA14 genome, which means that 25% of probes will not hybridize properly and must be masked out before subsequent analysis. One additional caution is to link probes to the corresponding PA14 genes, which might be different in the PAO1 annotation. The hybridization of the Affymetrix GeneChip with a strain that was not designed for it represents difficult work and there are even more difficulties in data analysis. If the *P. aeruginosa* strain used

is not sequenced, it will be difficult to find out which probes are not working and which probes are targeting a different location in the genome that they were designed for. Hence, it will be difficult to interpret whether what you have observed in the subsequent data analysis is real.

Transcriptome analysis for identification, timing, and signal specificity of *P. aeruginosa* quorum-controlled genes was compared to the PCR-based STM study.[72,78] Even though the quorum sensing activity was not monitored in STM mutants essential for maintenance in the rat lung, the genes identified in both cases were PA158 (probable RND efflux transporter), PA3734 (hypothetical), PA4172 (probable nuclease), PA3284 (hypothetical), and PA4692 (conserved hypothetical). Comparisons between the data from the microarray analysis of ECF sigma factor AlgU (*P. aeruginosa* sigma E)-dependent gene expression in *P. aeruginosa* with the PCR-based STM data identified PA1592 (hypothetical, *lptA*).[29,30] In addition to several of the typical known virulence factors, PCR-based STM identified a novel gene previously identified in the co-ordinate regulation of bacterial virulence genes by a novel calcium-dependent adenylate cyclase-dependent signaling pathway, PA4983, a two component response regulator.[92] It was apparent that comparisons between the *P. aeruginosa* STM mutants and the transcriptome of cAMp- and Vfr-deficient mutants indicated that numerous host-directed virulence determinants, including motility systems, attachment organelles, and the type II secretory pathway proteins of which many were identified by STM, are co-ordinately regulated under conditions that control expression of the type III secretion systems, which represents one of the more specific host-directed bacterial virulence determinants.[72,92]

5. CONCLUSION

Obviously, functional genomics of *Pseudomonas* and particularly *P. aeruginosa* has benefited immensely from the availability of the complete and annotated sequence of strain PAO1 and more recently of the completed sequence of PA14. Without this crucial information, it would be quite inconceivable to attempt functional genomics *in vivo* using STM, IVET or any other large-scale genomics method. Obviously, the field of *Pseudomonas* will benefit even more from the sequencing of additional species and strains of *Pseudomonas*. There is now clear evidence that *P. aeruginosa* strains can vary significantly in their genome content from a few hundred base pairs to several megabases. What is available today as the annotated sequence of PAO1 may actually represent some core genomic sequences, and metagenomics analysis will be pertinent to understanding virulence in an opportunistic pathogen such as *P. aeruginosa*. Hence, what have we learned from STM? One of the first significance of STM in *P. aeruginosa* is the clean correlation between virulence factors identified

with those previously known virulence factors reported by many laboratories around the world. In terms of functional genomics and significance, STM has pinpointed and identified several hypothetical and unknown proteins whose function *in vivo* is crucial for maintenance of *P. aeruginosa*. An interesting future prospect will be to analyze STM mutants of the PAO1 strain and the PA14 strain in similar models of infection.

Finally, re-analysis of the data from STM using various Pseudomonads in different models of infection coupled to transcriptomics and proteomics and its integration using a biological systems approach should give a better understanding of how an opportunistic pathogen competes and survives in any environment.

ACKNOWLEDGMENTS

Work in R.C.L. laboratory is funded by the Canadian Institute for Health Research and a team grant from Le Fonds de la Recherche en Santé du Québec. R.C.L. is a scholar of exceptional merit from the FRSQ.

REFERENCES

1. Aballay, A. and Ausubel, F.M., 2002, *Caenorhabditis elegans* as a host for the study of host-pathogen interactions. *Curr. Opin. Microbiol*., 5:97–101.
2. Alm, R.A. and Mattick, J.S., 1997, Genes involved in the biogenesis and function of type-4 fimbriae in *Pseudomonas aeruginosa*. *Gene*, 192:89–98.
3. Apidianakis, Y., Mindrinos, M.N., Xiao, W., Lau, G.W., Baldini, R.L., Davis, R.W., and Rahme, L.G., 2005, Profiling early infection responses: *Pseudomonas aeruginosa* eludes host defenses by suppressing antimicrobial peptide gene expression. *Proc. Natl. Acad. Sci. U.S.A.*, 102:2573–2578.
4. Autret, N. and Charbit, A., 2005, Lessons from signature tagged-mutagenesis on the infectious mechanisms of pathogenic bacteria. *FEMS Microbiol. Rev.*, 29:703–717.
5. Beatson, S.A., Whitchurch, C.B., Sargent, J.L., Levesque, R.C., and Mattick, J.S., 2002, Differential regulation of twitching motility and elastase production by Vfr in *Pseudomonas aeruginosa*. *J. Bacteriol.*, 184:3605–3613.
6. Benton, B.M., Zhang, J.P., Bond, S., Pope, C., Christian, T., Lee, L., Winterberg, K.M., Schmid, M.B., and Buysse, J.M., 2004, Large-scale identification of genes required for full virulence of *Staphylococcus aureus*. *J. Bacteriol.*, 186:8478–8489.
7. Berk, R.S., Brown, D., Coutinho, I., and Meyers, D., 1987, *In vivo* studies with two phospholipase C fractions from *Pseudomonas aeruginosa*. *Infect. Immun.*, 55:1728–1730.
8. Beuzon, C.R. and Holden, D.W., 2001, Use of mixed infections with *Salmonella* strains to study virulence genes and their interactions*in vivo*. *Microbes Infect.*, 3:1345–1352.
9. Bever, R.A. and Iglewski, B.H., 1988, Molecular characterization and nucleotide sequence of the *Pseudomonas aeruginosa* elastase structural gene. *J. Bacteriol.*, 170:4309–4314.
10. Bochner, B.R., Gadzinski, P., and Panomitros, E., 2001, Phenotype microarrays for high-throughput phenotypic testing and assay of gene function. *Genome. Res.*, 11:1246–1255.
11. Botes, J., Williamson, G., Sinickas, V., and Gurtler, V., 2003, Genomic typing of *Pseudomonas aeruginosa* isolates by comparison of riboprinting and PFGE: correlation of experimental

results with those predicted from the complete genome sequence of isolate PAO1. *J. Microbiol. Methods.*, 55:231–240.

12. Boucher, J.C., Schurr, M.J., Yu, H., Rowen, D.W., and Deretic, V., 1997, *Pseudomonas aeruginosa* in cystic fibrosis: role of MucC in the regulation of alginate production and stress sensitivity. *Microbiology*, 143:3473–3480.
13. Braun, P., Ockhuijsen, C., Eppens, E., Koster, M., Bitter, W., and Tommassen, J., 2001, Maturation of *Pseudomonas aeruginosa* elastase. Formation of the disulfide bonds. *J. Biol. Chem.*, 276:26030–26035.
14. Britigan, B.E., Roeder, T.L., Rasmussen, G.T., Shasby, D.M., McCormick, M.L., and Cox, C.D., 1992, Interaction of the *Pseudomonas aeruginosa* secretory products pyocyanin and pyochelin generates hydroxyl radical and causes synergistic damage to endothelial cells. Implications for *Pseudomonas*-associated tissue injury. *J. Clin. Invest.*, 90:2187–2196.
15. Budzikiewicz, H., 2001, Siderophores of the human pathogenic fluorescent pseudomonads. *Curr. Top. Med. Chem.*, 1:1–6.
16. Caballero, A.R., Moreau, J.M., Engel, L.S., Marquart, M.E., Hill, J.M., and O'Callaghan, R.J., 2001, *Pseudomonas aeruginosa* protease IV enzyme assays and comparison to other *Pseudomonas* proteases. *Anal. Biochem.*, 290:330–337.
17. Caiazza, N.C. and O'Toole, G.A., 2004, SadB is required for the transition from reversible to irreversible attachment during biofilm formation by *Pseudomonas aeruginosa*. *J. Bacteriol.*, 186:4476–4485.
18. Cash, H.A., Woods, D.E., McCullough, B., Johanson, W.G., Jr., and Bass, J.A., 1979, A rat model of chronic respiratory infection with *Pseudomonas aeruginosa*. *Am. Rev. Respir. Dis.*, 119:453–459.
19. Chiang, S.L., Mekalanos, J.J., and Holden, D.W., 1999, *In vivo* genetic analysis of bacterial virulence. *Annu. Rev. Microbiol.*, 53:129–154.
20. Choi, J.Y., Sifri, C.D., Goumnerov, B.C., Rahme, L.G., Ausubel, F.M., and Calderwood, S.B., 2002, Identification of virulence genes in a pathogenic strain of *Pseudomonas aeruginosa* by representational difference analysis. *J. Bacteriol.*, 184:952–961.
21. Cobb, J.P., Mindrinos, M.N., Miller-Graziano, C., Calvano, S.E., Baker, H.V., Xiao, W., Laudanski, K., Brownstein, B.H., Elson, C.M., Hayden, D.L., Herndon, D.N., Lowry, S.F., Maier, R.V., Schoenfeld, D.A., Moldawer, L.L., Davis, R.W., Tompkins, R.G., Bankey, P., Billiar, T., Camp, D., Chaudry, I., Freeman, B., Gamelli, R., Gibran, N., Harbrecht, B., Heagy, W., Heimbach, D., Horton, J., Hunt, J., Lederer, J., Mannick, J., McKinley, B., Minei, J., Moore, E., Moore, F., Munford, R., Nathens, A., O'Keefe, G., Purdue, G., Rahme, L., Remick, D., Sailors, M., Shapiro, M., Silver, G., Smith, R., Stephanopoulos, G., Stormo, G., Toner, M., Warren, S., West, M., Wolfe, S., and Young, V., 2005, Application of genome-wide expression analysis to human health and disease. *Proc. Natl. Acad. Sci. U.S.A.*, 103:4801–4806.
22. Cosson, P., Zulianello, L., Join-Lambert, O., Faurisson, F., Gebbie, L., Benghezal, M., van Delden, C., Curty, L.K., and Kohler, T., 2002, *Pseudomonas aeruginosa* virulence analyzed in a *Dictyostelium discoideum* host system. *J. Bacteriol.*, 184:3027–3033.
23. de Lorenzo, V., Herrero, M., Jakubzik, U., and Timmis, K.N., 1990, Mini-Tn5 transposon derivatives for insertion mutagenesis, promoter probing, and chromosomal insertion of cloned DNA in gram-negative eubacteria. *J. Bacteriol.*, 172:6568–6572.
24. Dewar, K., Sabbagh, L., Cardinal, G., Veilleux, F., Sanschagrin, F., Birren, B., and Levesque, R.C., 1998, *Pseudomonas aeruginosa* PAO1 bacterial artificial chromosomes: strategies for mapping, screening, and sequencing 100 kb loci of the 5.9 Mb genome. *Microb. Comp. Genomics.*, 3:105–117.
25. Dos Santos, V.A., Heim, S., Moore, E.R., Stratz, M,. and Timmis, K.N., 2004, Insights into the genomic basis of niche specificity of *Pseudomonas putida* KT2440. *Environ. Microbiol.*, 6:1264–1286.

26. Eberl, L. and Tummler, B., 2004, *Pseudomonas aeruginosa* and *Burkholderia cepacia* in cystic fibrosis: genome evolution, interactions and adaptation. *Int. J. Med. Microbiol.*, 294:123–131.
27. Ernst, R.K., D'Argenio, D.A., Ichikawa, J.K., Bangera, M.G., Selgrade, S., Burns, J.L., Hiatt, P., McCoy, K., Brittnacher, M., Kas, A., Spencer, D.H., Olson, M.V., Ramsey, B.W., Lory, S., and Miller, S.I., 2003, Genome mosaicism is conserved but not unique in *Pseudomonas aeruginosa* isolates from the airways of young children with cystic fibrosis. *Environ. Microbiol.*, 5:1341–1349.
28. Finnan, S., Morrissey, J.P., O'Gara, F., and Boyd, E.F., 2004, Genome diversity of *Pseudomonas aeruginosa* isolates from cystic fibrosis patients and the hospital environment. *J. Clin. Microbiol.*, 42:5783–5792.
29. Firoved, A.M., Boucher, J.C., and Deretic, V., 2002, Global genomic analysis of AlgU (sigma (E))-dependent promoters (sigmulon) in *Pseudomonas aeruginosa* and implications for inflammatory processes in cystic fibrosis. *J. Bacteriol.*, 184:1057–1064.
30. Firoved, A.M. and Deretic, V., 2003, Microarray analysis of global gene expression in mucoid *Pseudomonas aeruginosa*. *J. Bacteriol.*, 185:1071–1081.
31. Gibson, R.L., Burns, J.L., and Ramsey, B.W., 2003, Pathophysiology and management of pulmonary infections in cystic fibrosis. *Am. J. Respir. Crit. Care Med.*, 168:918–951.
32. Goryshin, I.Y. and Reznikoff, W.S., 1998, Tn5 *in vitro* transposition. *J. Biol. Chem.*, 273:7367–7374.
33. Handfield, M., Lehoux, D.E., Sanschagrin, F., Mahan, M.J., Woods, D.E., and Levesque, R.C., 2000, *In vivo*-induced genes in *Pseudomonas aeruginosa*. *Infect. Immun.*, 68:2359–2362.
34. Handfield, M. and Levesque, R.C., 1999, Strategies for isolation of *in vivo* expressed genes from bacteria. *FEMS Microbiol Rev.*, 23:69–91.
35. Hava, D.L. and Camilli, A., 2002, Large-scale identification of serotype 4 *Streptococcus pneumoniae* virulence factors. *Mol. Microbiol.*, 45:1389–1406.
36. He, J., Baldini, R.L., Deziel, E., Saucier, M., Zhang, Q., Liberati, N.T., Lee, D., Urbach, J., Goodman, H.M., and Rahme, L.G., 2004, The broad host range pathogen *Pseudomonas aeruginosa* strain PA14 carries two pathogenicity islands harboring plant and animal virulence genes. *Proc. Natl. Acad. Sci. U.S.A.*, 101:2530–2535.
37. Hendrickson, E.L., Plotnikova, J., Mahajan-Miklos, S., Rahme, L.G., and Ausubel, F.M., 2001, Differential roles of the *Pseudomonas aeruginosa* PA14 *rpoN* gene in pathogenicity in plants, nematodes, insects, and mice. *J. Bacteriol.*, 183:7126–7134.
38. Hensel, M., Shea, J.E., Gleeson, C., Jones, M.D., Dalton, E., and Holden, D.W., 1995, Simultaneous identification of bacterial virulence genes by negative selection. *Science*, 269:400–403.
39. Horvat, R.T. and Parmely, M.J., 1988, *Pseudomonas aeruginosa* alkaline protease degrades human gamma interferon and inhibits its bioactivity. *Infect. Immun.*, 56:2925–2932.
40. Hunt, T.A., Kooi, C., Sokol, P.A., and Valvano, M.A., 2004, Identification of *Burkholderia cenocepacia* genes required for bacterial survival *in vivo*. *Infect. Immun.*, 72:4010–4022.
41. Jacobs, M.A., Alwood, A., Thaipisuttikul, I., Spencer, D., Haugen, E., Ernst, S., Will, O., Kaul, R., Raymond, C., Levy, R., Chun-Rong, L., Guenthner, D., Bovee, D., Olson, M.V., and Manoil, C., 2003, Comprehensive transposon mutant library of *Pseudomonas aeruginosa*. *Proc. Natl. Acad. Sci. U.S.A.*, 100:14339–14344.
42. Jander, G., Rahme, L.G., and Ausubel, F.M., 2000, Positive correlation between virulence of *Pseudomonas aeruginosa* mutants in mice and insects. *J. Bacteriol.*, 182:3843–3845.
43. Kessler, E., Safrin, M., Gustin, J.K., and Ohman, D.E., 1998, Elastase and the LasA protease of *Pseudomonas aeruginosa* are secreted with their propeptides. *J. Biol. Chem.*, 273:30225–30231.
44. Kohler, T., Curty, L.K., Barja, F., van Delden, C., and Pechere, J.C., 2000, Swarming of *Pseudomonas aeruginosa* is dependent on cell-to-cell signaling and requires flagella and pili. *J. Bacteriol.*, 182:5990–5996.

45. Konig, B., Jaeger, K.E., Sage, A.E., Vasil, M.L., and Konig, W., 1996, Role of *Pseudomonas aeruginosa* lipase in inflammatory mediator release from human inflammatory effector cells (platelets, granulocytes, and monocytes. *Infect. Immun.*, 64:3252–3258.
46. Lau, G.W., Britigan, B.E., and Hassett, D.J., 2005, *Pseudomonas aeruginosa* OxyR is required for full virulence in rodent and insect models of infection and for resistance to human neutrophils. *Infect. Immun.*, 73:2550–2553.
47. Lau, G.W., Goumnerov, B.C., Walendziewicz, C.L., Hewitson, J., Xiao, W., Mahajan-Miklos, S., Tompkins, R.G., Perkins, L.A., and Rahme, L.G., 2003, The *Drosophila melanogaster* toll pathway participates in resistance to infection by the gram-negative human pathogen *Pseudomonas aeruginosa*. *Infect. Immun.*, 71:4059–4066.
48. Lau, G.W., Hassett, D.J., Ran, H., and Kong, F., 2004, The role of pyocyanin in *Pseudomonas aeruginosa* infection. *Trends Mol. Med.*, 10:599–606.
49. Lehoux, D.E. and Levesque, R.C., 2000, Detection of genes essential in specific niches by signature-tagged mutagenesis. *Curr. Opin. Biotechnol.*, 11:434–439.
50. Lehoux, D.E., Sanschagrin, F., Kukavica-Ibrulj, I., Potvin, E., and Levesque, R.C., 2004, Identification of novel pathogenicity genes by PCR signature-tagged mutagenesis and related technologies. *Methods Mol. Biol.*, 266:289–304.
51. Lehoux, D.E., Sanschagrin, F., and Levesque, R.C., 1999, Defined oligonucleotide tag pools and PCR screening in signature-tagged mutagenesis of essential genes from bacteria. *Biotechniques*, 26:473–478, 480.
52. Lehoux, D.E., Sanschagrin, F., and Levesque, R.C., 2002, Identification of *in vivo* essential genes from *Pseudomonas aeruginosa* by PCR-based signature-tagged mutagenesis. *FEMS Microbiol. Lett.*, 210:73–80.
53. Lewenza, S., Gardy, J.L., Brinkman, F.S., and Hancock, R.E., 2005, Genome-wide identification of *Pseudomonas aeruginosa* exported proteins using a consensus computational strategy combined with a laboratory-based PhoA fusion screen. *Genome. Res.*, 15:321–329.
54. Lizewski, S.E., Schurr, J.R., Jackson, D.W., Frisk, A., Carterson, A.J., and Schurr, M.J., 2004, Identification of AlgR-regulated genes in *Pseudomonas aeruginosa* by use of microarray analysis. *J. Bacteriol.*, 186:5672–5684.
55. Louis, D., Sorlier, P., and Wallach, J., 1998, Quantitation and enzymatic activity of the alkaline protease from *Pseudomonas aeruginosa* in culture supernatants from clinical strains. *Clin. Chem. Lab. Med.*, 36:295–298.
56. Lyczak, J.B., Cannon, C.L., and Pier, G.B., 2002, Lung infections associated with cystic fibrosis. *Clin. Microbiol. Rev.*, 15:194–222.
57. Mahajan-Miklos, S., Rahme, L.G., and Ausubel, F.M., 2000, Elucidating the molecular mechanisms of bacterial virulence using non-mammalian hosts. *Mol. Microbiol.*, 37:981–988.
58. Malhotra, S., Silo-Suh, L.A., Mathee, K., and Ohman, D.E., 2000, Proteome analysis of the effect of mucoid conversion on global protein expression in *Pseudomonas aeruginosa* strain PAO1 shows induction of the disulfide bond isomerase, dsbA. *J. Bacteriol.*, 182:6999–7006.
59. Manger, I.D. and Relman, D.A., 2000, How the host 'sees' pathogens: global gene expression responses to infection. *Curr. Opin. Immunol.*, 12:215–218.
60. Maroncle, N., Balestrino, D., Rich, C., and Forestier, C., 2002, Identification of *Klebsiella pneumoniae* genes involved in intestinal colonization and adhesion using signature-tagged mutagenesis. *Infect. Immun.*, 70:4729–4734.
61. Marty, N., Pasquier, C., Dournes, J.L., Chemin, K., Chavagnat, F., Guinand, M., Chabanon, G., Pipy, B., and Montrozier, H., 1998, Effects of characterised *Pseudomonas aeruginosa* exopolysaccharides on adherence to human tracheal cells. *J. Med. Microbiol.*, 47:129–134.
62. Mattick, J.S., Whitchurch, C.B., and Alm, R.A., 1996, The molecular genetics of type-4 fimbriae in *Pseudomonas aeruginosa*-a review. *Gene*, 179:147–155.

63. Mavrodi, D.V., Bonsall, R.F., Delaney, S.M., Soule, M.J., Phillips, G., and Thomashow, L.S., 2001, Functional analysis of genes for biosynthesis of pyocyanin and phenazine-1-carboxamide from *Pseudomonas aeruginosa* PAO1. *J. Bacteriol.*, 183:6454–6465.
64. Mecsas, J., 2002, Use of signature-tagged mutagenesis in pathogenesis studies. *Curr. Opin. Microbiol.*, 5:33–37.
65. Merrell, D.S., Hava, D.L., and Camilli, A., 2002, Identification of novel factors involved in colonization and acid tolerance of *Vibrio cholerae*. *Mol. Microbiol.*, 43:1471–1491.
66. Miller, V.L., 1999, Signature-tagged mutagenesis and the hunt for virulence factors: response. *Trends Microbiol.*, 7:388.
67. Nouwens, A.S., Walsh, B.J., and Cordwell, S.J., 2003, Application of proteomics to *Pseudomonas aeruginosa*. *Adv. Biochem. Eng. Biotechnol.*, 83:117–140.
68. Pallen, M.J., Penn, C.W., and Chaudhuri, R.R., 2005, Bacterial flagellar diversity in the post-genomic era. *Trends Microbiol.*, 13:143–149.
69. Palma, M., Worgall, S., and Quadri, L.E., 2003, Transcriptome analysis of the *Pseudomonas aeruginosa* response to iron. *Arch. Microbiol.*, 180:374–379.
70. Pedersen, S.S., Shand, G.H., Hansen, B.L., and Hansen, G.N., 1990, Induction of experimental chronic *Pseudomonas aeruginosa* lung infection with *P. aeruginosa* entrapped in alginate microspheres. *Apmis*, 98:203–211.
71. Pier, G.B., 2002, CFTR mutations and host susceptibility to *Pseudomonas aeruginosa* lung infection. *Curr. Opin. Microbiol.*, 5:81–86.
72. Potvin, E., Lehoux, D.E., Kukavica-Ibrulj, I., Richard, K.L., Sanschagrin, F., Lau, G.W., and Levesque, R.C., 2003, *In vivo* functional genomics of *Pseudomonas aeruginosa* for high-throughput screening of new virulence factors and antibacterial targets. *Environ. Microbiol.*, 5:1294–1308.
73. Priebe, G.P., Brinig, M.M., Hatano, K., Grout, M., Coleman, F.T., Pier, G.B., and Goldberg, J.B., 2002, Construction and characterization of a live, attenuated *aroA* deletion mutant of *Pseudomonas aeruginosa* as a candidate intranasal vaccine. *Infect. Immun.*, 70:1507–1517.
74. Rahme, L.G., Ausubel, F.M., Cao, H., Drenkard, E., Goumnerov, B.C., Lau, G.W., Mahajan-Miklos, S., Plotnikova, J., Tan, M.W., Tsongalis, J., Walendziewicz, C.L., and Tompkins, R.G., 2000, Plants and animals share functionally common bacterial virulence factors. *Proc. Natl. Acad. Sci. U.S.A.*, 97:8815–8821.
75. Rahme, L.G., Tan, M.W., Le, L., Wong, S.M., Tompkins, R.G., Calderwood, S.B., and Ausubel, F.M., 1997, Use of model plant hosts to identify *Pseudomonas aeruginosa* virulence factors. *Proc. Natl. Acad. Sci. U.S.A.*, 94:13245–13250.
76. Ramos, H.C., Rumbo, M., and Sirard, J.C., 2004, Bacterial flagellins: mediators of pathogenicity and host immune responses in mucosa. *Trends Microbiol.*, 12:509–517.
77. Reznikoff, W.S., Goryshin, I.Y., and Jendrisak, J.J., 2004, Tn5 as a molecular genetics tool: *in vitro* transposition and the coupling of *in vitro* technologies with *in vivo* transposition. *Methods Mol. Biol.*, 260:83–96.
78. Schuster, M., Lostroh, C.P., Ogi, T., and Greenberg, E.P., 2003, Identification, timing, and signal specificity of *Pseudomonas aeruginosa* quorum-controlled genes: a transcriptome analysis. *J. Bacteriol.*, 185:2066–2079.
79. Schweizer, H.P., 1991, *Escherichia-Pseudomonas* shuttle vectors derived from pUC18/19. *Gene*, 97:109–121.
80. Schweizer, H.P., Klassen, T., and Hoang, T., 1996, In K.F. Teruko Nakazawa, Dieter Haas, Simon Silvert (ed.), *Molecular Biology of Pseudomonads*. pp. 229–237, ASM Press, Washington, DC.
81. Shea, J.E., Santangelo, J.D., and Feldman, R.G., 2000, Signature-tagged mutagenesis in the identification of virulence genes in pathogens. *Curr. Opin. Microbiol.*, 3:451–458.
82. Simon, R., Priefer, U., and Pühler, A., 1983, A broad host range mobilization system for *in vivo* genetic engineering: transposon mutagenesis in Gram negative bacteria. *BioTechnology*, 1:784– 791.

83. Stover, C.K., Pham, X.Q., Erwin, A.L., Mizoguchi, S.D., Warrener, P., Hickey, M.J., Brinkman, F.S., Hufnagle, W.O., Kowalik, D.J., Lagrou, M., Garber, R.L., Goltry, L., Tolentino, E., Westbrock-Wadman, S., Yuan, Y., Brody, L.L., Coulter, S.N., Folger, K.R., Kas, A., Larbig, K., Lim, R., Smith, K., Spencer, D., Wong, G.K., Wu, Z., Paulsen, I.T., Reizer, J., Saier, M.H., Hancock, R.E., Lory, S., and Olson, M.V., 2000, Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen. *Nature*, 406:959–964.
84. Tummler, B., 1987, Unusual mechanism of pathogenicity of *Pseudomonas aeruginosa* isolates from patients with cystic fibrosis. *Infection*, 15:311–312.
85. Vasil, M.L. and Ochsner, U.A., 1999, The response of *Pseudomonas aeruginosa* to iron: genetics, biochemistry and virulence. *Mol. Microbiol.*, 34:399–413.
86. Visca, P., Leoni, L., Wilson, M.J., and Lamont, I.L., 2002, Iron transport and regulation, cell signalling and genomics: lessons from *Escherichia coli* and *Pseudomonas*. *Mol. Microbiol.*, 45:1177–1190.
87. Wagner, V.E., Bushnell, D., Passador, L., Brooks, A.I., and Iglewski, B.H., 2003, Microarray analysis of *Pseudomonas aeruginosa* quorum-sensing regulons: effects of growth phase and environment. *J. Bacteriol.*, 185:2080–2095.
88. Wagner, V.E., Gillis, R.J., and Iglewski, B.H., 2004, Transcriptome analysis of quorum-sensing regulation and virulence factor expression in *Pseudomonas aeruginosa*. *Vaccine*, 22(Suppl 1):S15–S20.
89. Walker, T.S., Bais, H.P., Deziel, E., Schweizer, H.P., Rahme, L.G., Fall, R., and Vivanco, J.M., 2004, *Pseudomonas aeruginosa*-plant root interactions. Pathogenicity, biofilm formation, and root exudation. *Plant Physiol.*, 134:320–331.
90. Whiteley, M., Bangera, M.G., Bumgarner, R.E., Parsek, M.R., Teitzel, G.M., Lory, S., and Greenberg, E.P., 2001, Gene expression in *Pseudomonas aeruginosa* biofilms. *Nature*, 413:860–864.
91. Wiehlmann, L., Salunkhe, P., Larbig, K., Ritzka, R., and Tummler, B., 2002, Signature-tagged mutagenesis of *Pseudomonas aeruginosa*. *Genome Lett.*, 3:131–139.
92. Wolfgang, M.C., Lee, V.T., Gilmore, M.E., and Lory, S., 2003, Coordinate regulation of bacterial virulence genes by a novel adenylate cyclase-dependent signaling pathway. *Dev. Cell.*, 4:253–263.
93. Worlitzsch, D., Tarran, R., Ulrich, M., Schwab, U., Cekici, A., Meyer, K.C., Birrer, P., Bellon, G., Berger, J., Weiss, T., Botzenhart, K., Yankaskas, J.R., Randell, S., Boucher, R.C., and Doring, G., 2002, Effects of reduced mucus oxygen concentration in airway *Pseudomonas* infections of cystic fibrosis patients. *J. Clin. Invest.*, 109:317–325.
94. Wu, W., Badrane, H., Arora, S., Baker, H.V., and Jin, S., 2004, MucA-mediated coordination of type III secretion and alginate synthesis in *Pseudomonas aeruginosa*. *J. Bacteriol.*, 186:7575–7585.
95. Yu, H. and Head, N.E., 2002, Persistent infections and immunity in cystic fibrosis. *Front. Biosci.*, 7:D442–D457.

5

A GENOME-WIDE MUTANT LIBRARY OF *PSEUDOMONAS AERUGINOSA*

Michael A. Jacobs[1] and Colin Manoil[2]

[1]Genome Center, Department of Medicine
University of Washington
Seattle, WA 98105
[2]Department of Genome Sciences
University of Washington
Seattle, WA 98195, USA

Key Words: transposon, mutant library, gene fusion, essential gene

1. INTRODUCTION

Although transcription microarray and proteomic technologies can identify large numbers of genes expressed under a particular condition, the biological meaning of such correlations is generally unclear without further analysis. Mutant studies can help clarify such a gene's function in the process being studied, but the generation large numbers of mutations inactivating candidate genes is time-consuming and commonly limits the number of genes examined. This chapter summarizes the construction and attributes of a large transposon mutant library of *Pseudomonas aeruginosa* designed to help overcome this limitation. The library includes multiple insertions in most nonessential genes of the bacterium, and can be used systematically to examine the phenotypes of mutations in candidate genes that have been associated with a biological process of interest using other approaches. In addition, the library can be screened directly to provide virtually complete identification of genes required for any

Pseudomonas, Volume 4, edited by Juan-Luis Ramos and Roger C. Levesque

process for which a suitable screen can be devised. In addition to the generation of loss-of-function insertion mutations, the transposons used to generate the strain library have attributes which facilitate additional downstream genetic studies.

2. CONSTRUCTION OF THE TRANSPOSON MUTANT COLLECTION

Two transposon Tn*5* derivatives, IS*phoA*/hah and IS*lacZ*/hah, were used to generate the mutant library (Figure 1).[1,6] The transposons generate alkaline phosphatase or β-galactosidase translational gene fusions if inserted in target genes in the appropriate orientation and reading frame (Figure 2). Such in-frame insertions may be converted into 63 codon insertions ("i63") by *loxP* X *loxP* recombination catalyzed by Cre recombinase (Figures 1 and 2). The resulting sequence encodes an internal tag which includes an influenza hemagglutinin epitope and a hexahistidine metal affinity purification motif, which may facilitate further studies of individual gene products.[1] The recombination event also eliminates the tetracycline resistant determinant of the transposon, making it possible to reutilize the marker in constructing multiple mutants.[4]

As summarized in the process schematic (Figure 3), transposon insertions were generated in *P. aeruginosa* PAO1 (from L. Passador and B. Iglewski, University of Rochester) by mating the strain with either of two *Escherichia coli* donors carrying conjugation-proficient transposon delivery suicide plasmids derived from pUT.[5] The mutagenized cultures were plated on large bioassay-scale L agar plates containing tetracycline (to select for transposon insertions), chloramphenicol (to select against donor cells), and chromogenic indicator (5-bromo-4-chloro-3-indolyl galactoside [Xgal] or 5-bromo-4-chloro-3-indolyl phosphate [XP]) to detect hybrid proteins.[6] Colonies were arrayed into 384-well plates using a Qpix robot (Genetix Ltd.). The robot arrays a 384-well plate in about 15 min, and can array colonies into several 384-well plates from a single bioassay plate while running unattended. In our experience, an expert technician requires at least an hour to array a single 384-well plate, and the risk of cross-contamination and other errors is considerably greater than with robotic arraying. In the arraying step, we generally used colony selection criteria that were as conservative as possible to help ensure that unique colonies were picked and that "fake" colonies were avoided. For plates on which colonies were scarce, it was necessary to relax the criteria in order to pick as many colonies as possible, a practice which typically resulted in a greater number of duplicate picks (due to overlapping images guiding the picking) and to a larger number of blank wells due to recognition errors. Even given these complications, automated picking is

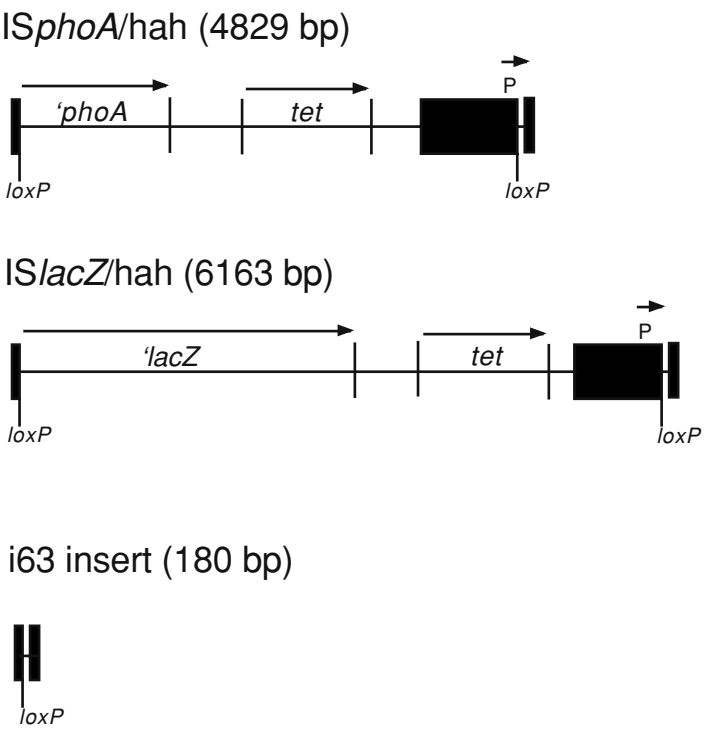


Figure 1. Transposon derivatives. The structures of transposons IS*phoA*/hah and IS*lacZ*/hah are diagrammed, as is the structure of the i63 insertion element generated by *loxP* X *loxP* recombination from either transposon. Sequences derived from IS*50* are represented as filled rectangles. The sequences corresponding to different parts of IS*phoA*/hah are (in bp): IS*50* outside end 1–51; *loxP*, 66–99; '*phoA*, 116–1448; *tet*, 2123–3310; *loxP*, 4714–4747; IS50 inside end, 4810–4829. The sequences in bp corresponding to different parts of IS*lacZ*/hah are: IS*50* outside end,1–51; *loxP*, 66–99; "*lacZ*, 126–3170; *tet*, 3778–4965; *loxP*, 6049–6082; IS50 inside end, 6144–6163. The i63 insert carries a single *loxP* sequence at bp 66–99. Insertions of all three elements have nine bp duplications of target DNA at the insertion sites. When inserted in-frame, the i63 insertion encodes the following sequence: (SPTA)DSYTQVASWTEPFPFSIQGDLITSYNVCYTKLLIKHHHHHH-YPYDVPDYARDPRSDQET(VADEG)XX, where the residues in parentheses are alternative possibilities encoded in part by target gene sequences, "X" refers to residues encoded entirely by tare gene sequences duplicated at the insertion site, and the hemagglutinin epitope is underlined. The DNA sequences of the three elements are available at: http://www.gs.washington.edu/labs/manoil/index.htm.

essential in constructing a collection such as this consisting of tens of thousands of strains.

During picking, we attempted to select only white colonies for some plates, and only blue colonies for other plates, although the robot occasionally made "incorrect" color assignments. When picking blue colonies, we grew

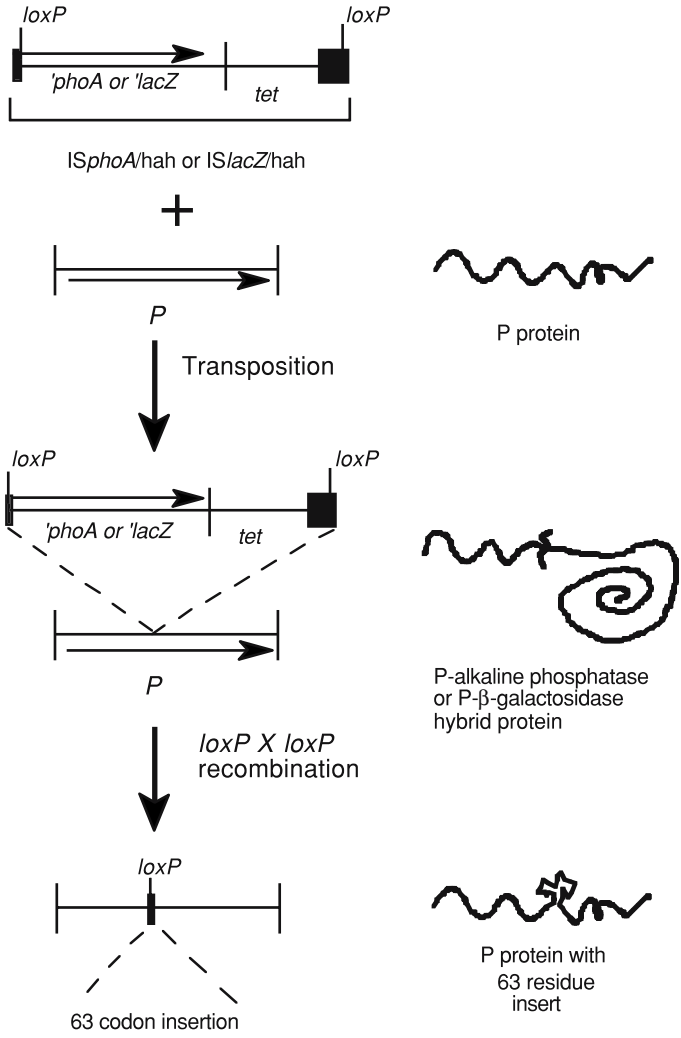


Figure 2. Generation of gene fusions and internal tags by transposon insertion. The diagram represents the generation of a β-galactosidase and alkaline phosphatase translational gene fusions by IS*phoA*/hah and IS*lacZ*/hah insertion and the conversion of such insertions into i63 insertions by Cre recombination. Transposon mutagenesis of *P. aeruginosa* was carried out through conjugal transfer of a suicide plasmid (pCM639 or pIT2) using an *E. coli* donor.[6] *LoxP* X *loxP* recombination may be induced through the conjugal introduction of a nonreplicating plasmid carrying *cre.*[1]

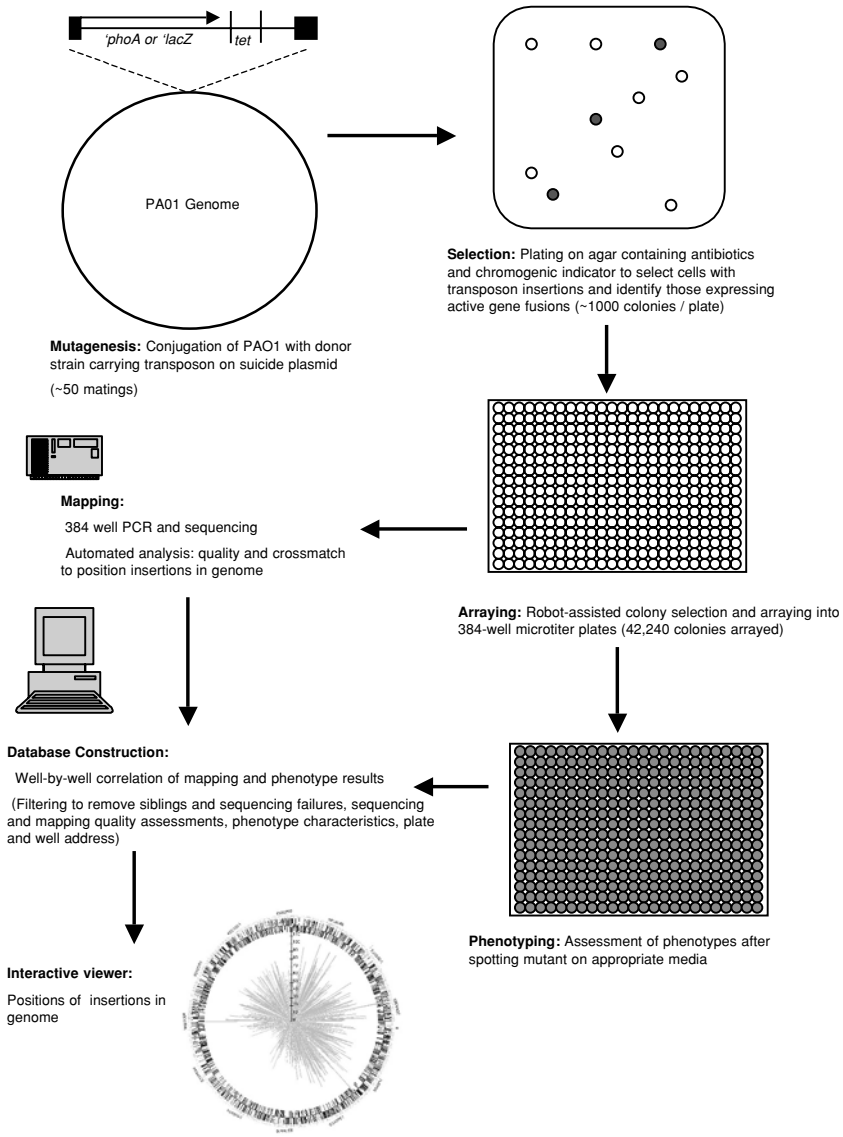


Figure 3. Generation, analysis, and maintenance of the *P. aeruginosa* PAO1 mutant library. The steps used to generate transposon insertions, amplify insertion junction fragments, sequence the fragments, and assign insertion sites to the *P. aeruginosa* genome are summarized.

plates to near-confluence, in order that blue colonies stood out against a white colony background and were easier for the robot to recognize. Although this tactic probably increased the frequency at which arrayed mutants were contaminated by adjacent cells, streak tests of representative samples on indicator indicated this contamination was not a major problem.

High throughput methods were essential for sequence-mapping and storing the mutant collection.[6] The methods utilized were originally developed for sequencing genomes, in which low quality reads were efficiently screened out, and individual strains stored in glycerol plates were rarely re-used.

By utilizing a 384-well format, the volumes of reagents and associated costs of PCR and sequencing were minimized. In addition, the 384-well format allows for efficient storage of all of the strains. However, the spatial constriction of this format also presents difficulties for *P. aeruginosa*, since the cells often stick to each other, forming films at the surfaces of some wells. When using a 384-pin replicator to distribute cells from such a culture, large clumps of cells may stick to a pin, forming a globule that is large enough to contaminate adjacent wells upon inoculation. This property also added to the difficulty of carrying out PCR.[6] We found that the problem was reduced if partially thawed glycerol stocks were used as a source of template in PCR reactions and for inoculating replica plates. The mutant collection consists of 111 plates (384 strains/plate) and may be stored comfortably on one shelf of a standard -80° freezer. Optimizing conditions for long-term storage of the strains and replication if the library is continuing. It is our experience that *Pseudomonas* strains stored in glycerol at 80°C for long periods lose viability, and it has been suggested that DMSO cultures survive better (R. Hancock, personal communication). We maintain several copies of the collection in separate locations. The original strains are kept as pristine as possible and a rotating stock of replica plates is produced for purposes of distributing strains, which requires repeated free-thaw cycles.

Mapping of the strains was made possible by using the reference genome sequence for *P. aeruginosa* PAO1.[8] DNA fragments which included transposon insertion junctions were amplified and sequenced using a semi-degenerate PCR scheme,[6] as described in detail at http://www.genome.washington.edu/UWGC/pseudomonas/pdf/Supplementary_Methods.pdf

A PERL script, illustrated as a flow chart in Figure 4, was used to cross-match the junction sequences against the transposon sequence and the *P. aeruginosa* genome, and then to determine the orientation and position of the insertions relative to annotated open reading frames. Data from the collection were stored in a Microsoft Access database. Visual Basic code was written to determine the number of unique hits (and to screen out duplicate hits), and to determine the number of cases of discrepancies (for strains that had been re-sequenced). A subset of these data was incorporated into an Oracle database for access by outside investigators (see below).

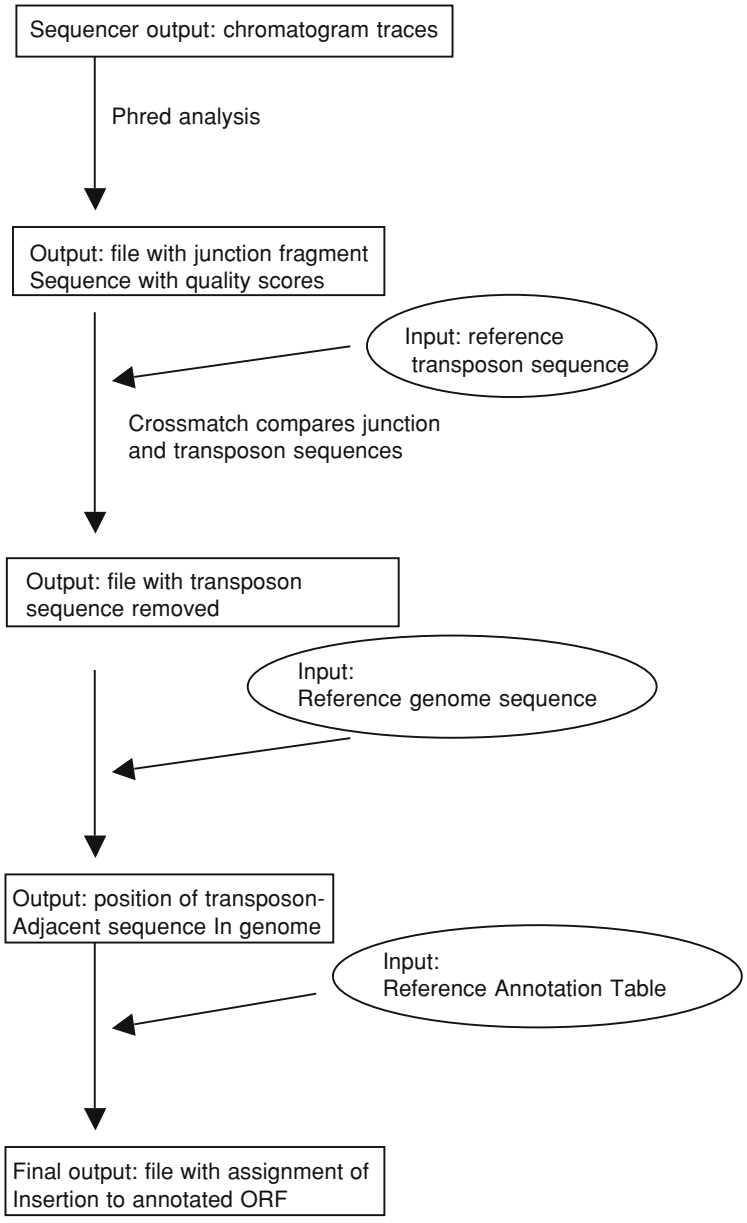


Figure 4. Assignment of transposon insertions to the *P. aeruginosa* genome sequence. The flow chart summarizes the computational steps used to position transposon insertions in the *P. aeruginosa* genome. In many cases, the quality of the junction fragment sequence was insufficient to allow the transposon end to be recognized, and other parts of the sequence were used to make an (approximate) assignment. The Phred and Crossmatch algorithms have been described.[2,3]

Table 1. Makeup of the *P. aeruginosa* mutant library.

Mutants Arrayed	42,240
Insertions Mapped	36,154
Mapping Success Rate	80%
Identical insertion locations	4423
Unique insertion locations	30,100
Insertions within ORFs	27,263
Insertions between ORFs	2837
ORFs hit internally	4892
ORFs never hit internally	678
Average hits per ORF	5.05X
In-frame insertion locations	4823
Reporter-active (blue colony) in-frame insertions	2546
ORFs with in-frame insertions	2582
Mutants scored for colony phenotype	42,240
Twitching (swarming) – defective mutants	709
ORFs with twitching (swarming) – defective mutations	360
Auxotrophic mutants	813
ORFs with auxotrophic mutations	546

3. COMPOSITION OF THE MUTANT COLLECTION

The collection is made up of 42,240 strains. Of these, 34,837 have insertions mapped to a location in the *P. aeruginosa* genome, a success rate of 82%. Once duplicate records and siblings were eliminated, 30,100 unique insertion locations were identified (Table 1). Approximately 90% of insertions were within ORFs, corresponding well to the fraction of the genome predicted to be coding.[8] Reporter gene activities of the strains based on colony color on chromogenic indicator media are summarized in Table 2, and indicate that

Table 2. Results of reporter active–inactive (blue–white) scoring.

	IS*phoA*/hah insertions	IS*lacZ*/hah insertions
Unique Insertions	15,063	15,037
Reporter active (blue colony)	1973	2416
Reporter inactive (white colony)	13,090	12,621
Total ORFs with $\geq$1 insertion	4313	4430
ORFs with $\geq$1 active insertion	954	1436
ORFs with $\geq$1 inactive insertion	4206	4221
ORFs with both active and inactive insertions	847	1227
Unique in-frame insertions	2821	2002
Active in-frame insertions	1387	1159
Inactive in-frame insertions	1434	843
ORFS with in-frame insertions	1761	1432
ORFS with active in-frame insertions	738	768
ORFs with inactive in-frame insertions	1167	729
ORFS with both active and inactive in-frame insertions	144	65

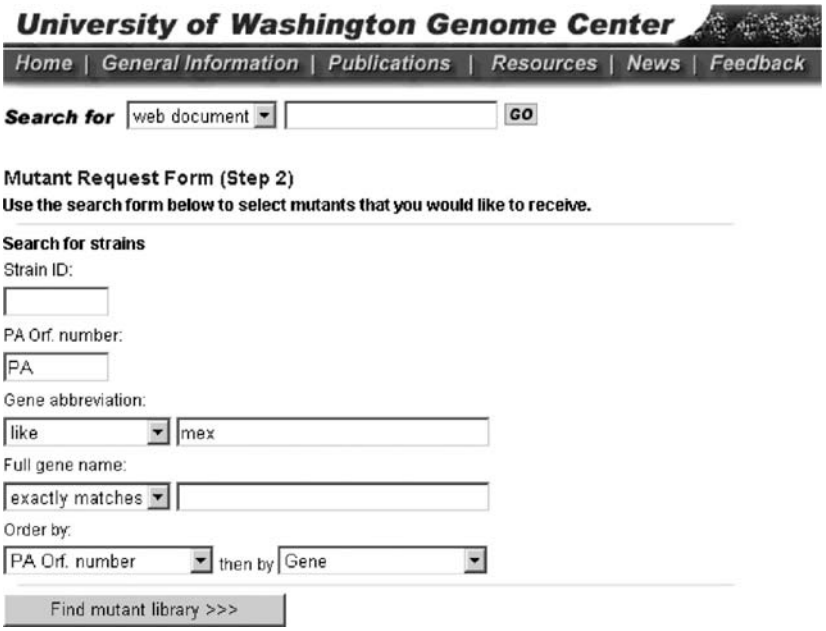


Figure 5. Searching the *P. aeruginosa* mutant collection. The window shown (at http://www.genome.washington.edu/UWGC/index.cfm) is used to search the mutant library for strains of interest. The example shows the input for a search for mutants in *mex* family genes. Clicking the "Find mutant library" button leads to the window shown in Figure 6.

roughly half of the in-frame insertions generated are expressed under the isolation conditions. 678 ORFs were never hit by a transposon insertion, and are called "candidate essential genes."[6]

4. MUTANT DISTRIBUTION

A primary motivation for assembling the *P. aeruginosa* library was to create a resource of mutants for researchers worldwide. Researchers may use reverse genetic techniques, such as bioinformatic searches, microarray data, or proteomics data to identify genes of interest in *P. aeruginosa*. Once a list of interesting genes is generated, the corresponding mutant strains may be obtained, saving time and resources that would otherwise be required for constructing the mutants.

To facilitate the distribution of mutant strains, we created a publicly accessible website at http://www.genome.washington.edu/UWGC/index.cfm. Researchers may search for strains by ORF number, gene name or gene abbreviation. (The transposon insertion positions are also accessible though the Pseudomonas.com website in the Gbrowse viewer.) An example of a search for

mutants affecting the *mex* (Multidrug efflux) family of genes[7] is illustrated in Figures 2 and 6 and summarized in Table 6. The requestor can search for mutants affecting a family of genes by selecting the appropriate gene abbreviation ("*mex*") and the "like" function instead of the "exactly matches" function. (In searching for insertions in specific genes, the most accurate way to search is by PA ORF number, since gene names may change between updates.) Requestors can register their information, and then activate a search in which mutants of interest may be saved for an order to be placed (Figure 6).

We prepare mutants for distribution by streaking the appropriate strain onto an L agar plate, and after growth making a stab using cells from a heavy part of the streak (rather than from a single colony). Immediately after receipt, we recommend that each strain be streaked out on tetracycline selective, and that a representative sample from the thickest part of the streak be stored frozen. This is to preserve all strains in a streak in the event of multiple strains are pretent. We also strongly recommend that the identity of all strains be confirmed prior to use (see below).

In order to distribute mutant strains without restriction, we require shipping expenses be paid by individuals making requests. All other costs are carried by the University of Washington Genome Center. We ask that requestors provide an account number for Federal Express (the only authorized shipper for transporting biological hazards in the US) when submitting their requests. For international shipments, we also require that requestors determine the necessary import permits required by their countries for dangerous goods. Up to 50 strains may be placed in a single shipping container, and shipping costs at present (October, 2004) are approximately $40 per order for domestic orders and $150 for international orders.

5. CONFIRMATION OF MUTANT IDENTITY

Some wells of the mutant library plates contain more than one strain, a consequence of the high throughput methodology used to assemble and replicate the collection. Accordingly, it is essential that the strains obtained from the collection be confirmed prior to experimental use. After receiving a strain we recommend that cells from the stab be streaked to L agar, followed by assay of several individual colonies using PCR. Two tests are necessary, one to show that the intact gene corresponding to the insertion is absent, and a second to show that the transposon insertion location is approximately correct. For each gene of interest, primers flanking the gene are designed ("f" and "r" in Figure 7), and used to test amplification of the wild-type fragment using the mutant strains and the wild-type parent. Depending on PCR conditions, no fragment or a fragment corresponding to a very large product will result from a correctly assigned mutant strain. Assigning the position of a transposon insertion requires a transposon-specific primer (either "Hah minus 138" (5′-cgggtgcagtaatatcgccct-3′) for

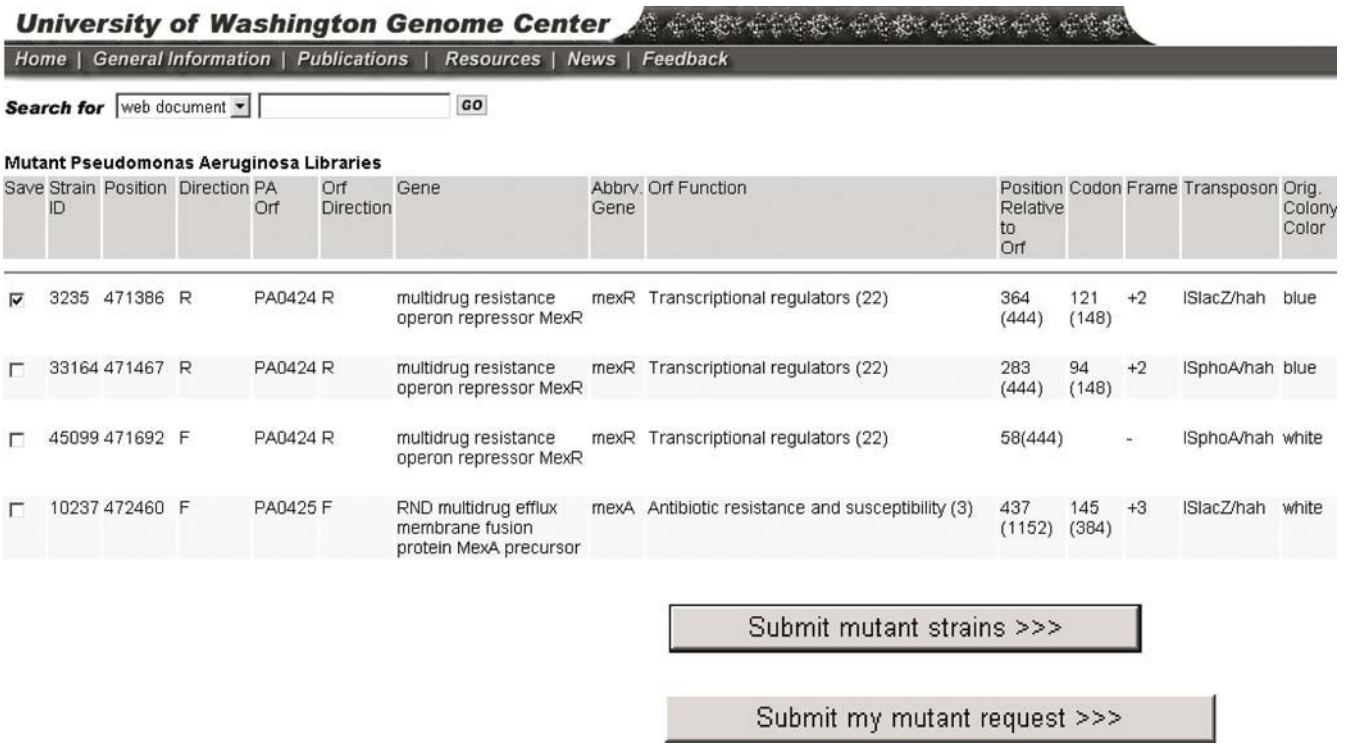


Figure 6. Requesting strains. The first four entries from the list of mutants returned in response to the query illustrated in Figure 5 are shown. A summary of the total list (97 hits in 10 genes) is presented in Table 6. To request a particular strain, the corresponding box is checked (strain 3235 in the example shown). As indicated, strain 3235 is an in-frame (+2) insertion of IS*lacZ*/hah at bp 364 out of 444 in *mexR*. The mutant forms $LacZ^+$ (blue) colonies on L agar containing X-Gal indicator. After strains of interest from the list have been selected, the "Submit mutant strains" button is clicked, and the next screen will show "Your mutant request has been saved". Additional searches and requests may then be carried out, and when the process has been completed, the "Submit my mutant request" button is clicked and the compete request is submitted.

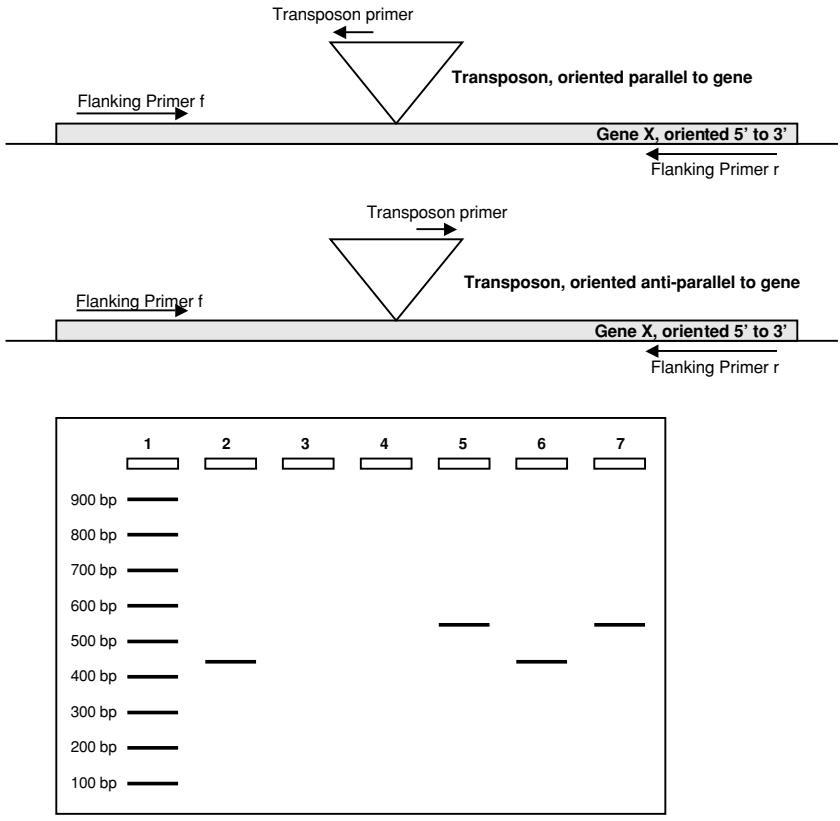


Figure 7. Verifying mutant identity. A summary of the procedure used to verify the identify of mutants obtained from the *P. aeruginosa* mutant collection is presented. The top of the figure represents the two possible orientations of an IS*lacZ*/hah transposon in a gene and the primers (two gene-specific and one transposon-specific) used for the analysis. The gel represented at the bottom corresponds to a (hypothetical) PCR analysis of two single colonies derived from a streak from a stab of strain 3235 which has a *lacZ* transposon insertion orientel parallel to ORF PA 0424 (Figure 6). Primers amplifying the entire *mexR* gene (444 bp) plus 40 bp (20 bp for each primer) would generate a PCR product of 464 bp (lane 2). If the colony carried the expected insertion, amplification using primes f and r would not yield a wild-type sized product (lane 3) nor would a reaction using flanking primer r and the transposon primer. However, amplification with the f primer and *lacZ* 148 would generate a 552 bp fragment (364 bp between the start of the gene, 148 bp between *lacZ* 148 and the junction, and 2 primers of 20 bp each). Lanes 6 and 7 represent a pattern sometimes observed in which both wild-type and insertion fragments are observed, perhaps due to insertion of the transposon in one copy of a genomic tandem duplication.

IS*phoA*/hah insertions or "*lacZ* 148" (5′-gggtaacgccagggttttcc-3′) for IS*lacZ*/hah insertions). The appropriate transposon primer is used in conjunction with one of the flanking primers used in the first PCR test, depending on the orientation of the transposon relative to the gene (Figure 7).

Table 3. Mutants requested (by functional category).

ORF function	Strains available	Strains requested	Proportion (%)
Chemotaxis	411	186	45
Motility and attachment	615	264	43
Secreted factors	493	167	34
Protein secretion/export apparatus	372	117	31
DNA replication, recombination, modification, and repair	418	121	29
Cell wall/LPS/capsule	340	95	28
Antibiotic resistance and susceptibility	165	45	27
Adaptation, protection	430	107	25
Transcriptional regulators	1949	482	25
Two-component regulatory systems	1030	213	21
Related to phage, transposon, or plasmid	443	78	18
Chaperones and heat shock proteins	275	41	15
Membrane proteins	259	36	14
Fatty acid and phospholipid metabolism	369	46	12
Transport of small molecules	4851	541	11
Translation, post-translational modification, degradation	503	55	11
Hypothetical, unclassified, unknown	12,547	1297	10
Energy metabolism	1061	107	10
Carbon compound catabolism	735	68	9
Central intermediary metabolism	458	37	8
Amino acid biosynthesis and metabolism	1280	96	8
Putative enzymes	2905	215	7
Cell division	62	4	6
Biosynthesis of cofactors, prosthetic groups, and carriers	412	16	4
Transcription, RNA processing, and degradation	226	3	1
Nucleotide biosynthesis and metabolism	236	1	<1

6. MUTANTS DISTRIBUTED FROM SEATTLE

As of October 15th, 2004, 4575 strains have been shipped to 127 researchers corresponding to 79 cities in 15 countries. When the strains are sorted by functional category, independent of ORFs, it is clear that some categories are far more popular than others. Chemotaxis, motility and attachment, and secreted factors are the most frequently requested (Table 3), likely due to their known involvement in virulence and recognition by host immune responses. When the collection is viewed in the context of individual genes, the number of genes with insertions is roughly proportional to their representation by functional category in the genome (Table 4). In the most popular categories, mutations in all of the genes with insertions ("available" genes). The ten "most popular" genes include three regulators and one gene needed for the synthesis of an extracellular regulatory molecule (Table 5).

Table 4. Mutants requested (by fraction of genes in different functional categories).

Gene function	Number of genes	% of genes	Genes with insertions	Genes with insertions requested	% genes with insertions requested
Antibiotic resistance and susceptibility	19	0.3	18	18	100
Chemotaxis	45	0.8	45	45	100
Motility and attachment	67	1.2	65	55	85
Related to phage, transposon, or plasmid	62	1.1	54	41	76
Cell wall/LPS/capsule	86	1.5	56	39	70
DNA replication, recombination, modification, and repair	81	1.5	66	45	68
Two-component regulatory systems	116	2.1	113	71	63
Protein secretion/export apparatus	84	1.5	68	42	62
Secreted factors	60	1.1	55	33	60
Adaptation, protection	66	1.2	57	30	53
Transport of small molecules	559	10.0	533	263	49
Membrane proteins	43	0.8	39	18	46
Chaperones and heat shock proteins	52	0.9	44	20	46
Fatty acid and phospholipid metabolism	57	1.0	42	18	43
Transcriptional regulators	403	7.2	376	161	43
Central intermediary metabolism	65	1.2	58	19	33
Energy metabolism	170	3.1	135	42	31
Hypothetical, unclassified, unknown	2381	42.7	2120	592	28
Translation, post-translational modification, degradation	149	2.7	74	18	24
Putative enzymes	457	8.2	432	104	24
Carbon compound catabolism	134	2.4	123	28	23
Amino acid biosynthesis and metabolism	151	2.7	138	31	23
Cell division	26	0.5	15	3	20
Biosynthesis of cofactors, prosthetic groups, and carriers	132	2.4	88	11	13
Transcription, RNA processing, and degradation	45	0.8	32	3	9
Nucleotide biosynthesis and metabolism	60	1.1	46	1	2
Totals	5570		4892	1751	

Table 5. The most popular genes in *P. aeruginosa.*

PA ORF number	Times requested	Gene name	Function
PA4525	30	*pilA*	Motility and attachment
PA1092	30	*fliC*	Motility and attachment
PA1003	26	*pqsR (mvfR)*	Transcriptional regulators; Quinolone signal response
PA3477	25	*rhlR*	Transcriptional regulators; Quorum sensing
PA3790	23	*oprC*	Transport of small molecules
PA3622	23	*rpoS*	Transcriptional regulators
PA0652	23	*vfr*	Transcriptional regulators
PA3841	21	*exoS*	Secreted factors (toxins, enzymes, alginate)
PA3930	18	*cioA*	Energy metabolism
PA3724	17	*lasB*	Secreted factors (toxins, enzymes, alginate)
PA0996	17	*pqsA*	Quinolone signal response

Table 6. Mex mutants in the *P. aeruginosa* mutant library.

Gene	PA ORF	Hits
mexR	PA0424	3
mexA	PA0425	9
mexB	PA0426	19
mexH, mexY, amrB	PA2018	10
mexG, mexX, amrA	PA2019	1
mexZ amrR	PA2020	2
mexE	PA2493	5
mexF	PA2494	15
mexD	PA4598	24
mexC	PA4599	9

7. WHOLE LIBRARY PHENOTYPIC SCREENS

We found that it is feasible to carry out phenotypic screens of the entire mutant library. On the basis of comparisons with earlier studies of two well-defined processes (twitching or swarming motility and growth on minimal medium), the approach appears to provide virtually complete lists of genes in which loss-of-function mutations confer the phenotypes of interest.[6] The collection has more recently also been screened for mutants showing reduced growth under anaerobic conditions and altered sensitivity to the antibiotic tobramycin (I. Thaipittsukul; S. Lee, unpublished results). In each case, the protocol followed was to spot small aliquots of each mutant onto appropriate medium using a 384-pin pronging device, followed by incubation, digital imaging, and one or more rounds of scoring by visual inspection.

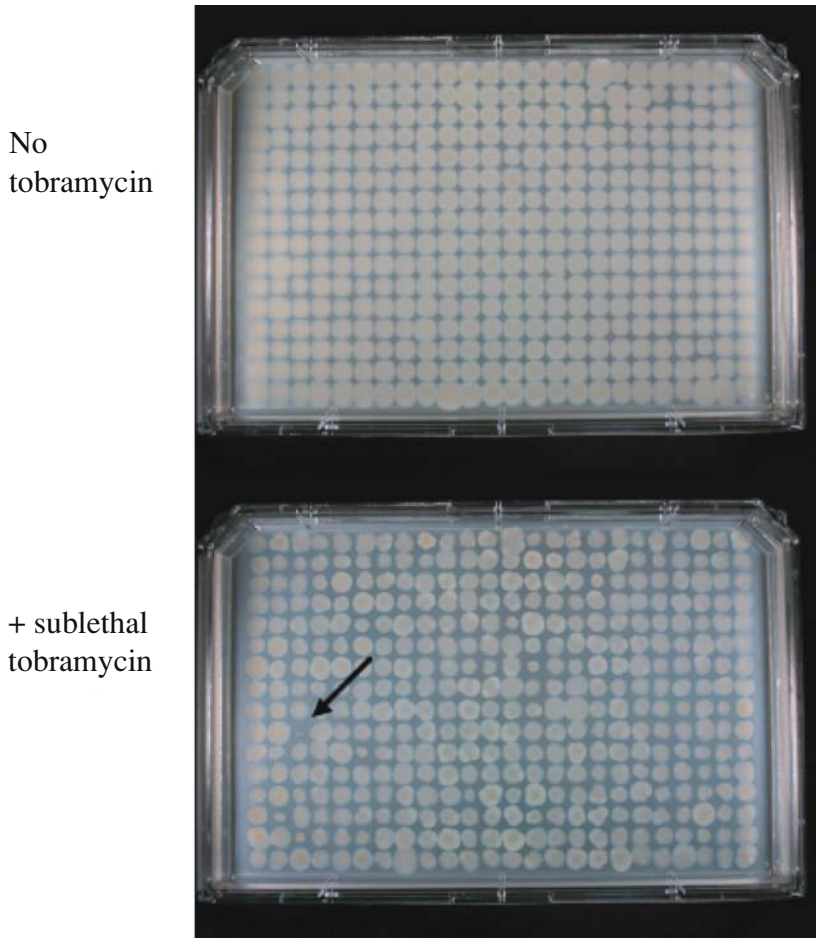


Figure 8. Screen for increased tobramycin sensitivity. The growth of mutants from one of the 384-well mutant library plates is shown on L agar devoid of tobramycin (top) or containing a sublethal concentration of tobramycin (1 μg/ml) (bottom). The position of one mutant appearing to be highly tobramycin sensitive is indicated. Note that several other mutant spots also show reduced growth in the presence of tobramycin.

An example of the results from the tobramycin resistance screen are shown in Figure 8. Although cells in all 384 positions grew on the no tobramycin plate, one failed to grow and several show reduced growth on agar containing a sublethal concentration of the antibiotic, indicating hypersensitivity. Additional studies are required to determine which of the gene-phenotype associations indicated by results such as these are valid, including examination of the other insertion alleles of the gene.

Our experience with whole library screening has led to several general conclusions:

- The existence of multiple alleles for most genes largely compensates for the low-level contamination of some wells since alleles of the same genes from uncontaminated wells can be phenotypically identified even if the contaminated mutants cannot.
- The fraction of mutations in a particular gene which are identified in a given screen is a function of the strength of the phenotype and the extent to which contaminating cells interfere with recognition. For example, mutations leading to tobramycin resistance were very efficiently identified, presumably because the existence of sensitive contaminants did not interfere with resistant cell growth. In contrast, twitching (swarming)-deficient mutants were less efficiently detected, presumably because the presence of small numbers of motile bacteria obscured the sharp colony edge of nonmotile mutants.
- Setting a low threshold for including potential mutants showing to a phenotype helps associate the largest number of genes with the property of interest. The corresponding increased number of false identifications tends to be compensated for by the redundancy of mutants in the collection, since only the genes for which a significant fraction of alleles are detected are typically considered promising for further analysis.
- Saving permanent images of phenotyping plates makes it possible to re-examine results after an initial screen has been completed. Our general experience has been that as a first screen has progressed, the ability to recognize more subtle phenotypic variants has increased, so that re-screening has identified mutants missed in the first screen.
- For some screens, it is helpful to spot duplicate plates to provide two independent tests of behavior under a particular condition. This practice allows an immediate check on the reproducibility of the phenotype observed for a given mutant, and is particularly valuable for phenotypes for which variability in the number of cells transferred in the original spotting step significantly affects scoring.

8. CONCLUSIONS

The *P. aeruginosa* mutant library appears to be a useful resource both as a source of mutants to analyze gene associations revealed using other techniques and directly for comprehensive screens for particular phenotypes. Increased automation in such phenotyping should substantially increase the utility of such screens in the future.

ACKNOWLEDGMENTS

We gratefully acknowledge Sam Lee for providing the data shown in Figure 8, Elizabeth Sims for shipping and maintaining the strain collection, Gregory Alexander for construction and management of the web site, Stephen Ernst for database utility design, and Iyarit Thaipisuttikul, Sam Lee, Rajinder Kaul, and Maynard Olson for helpful discussions. This work was supported by grants from the Cystic Fibrosis Foundation.

REFERENCES

1. Bailey, J. and Manoil, C., 2002, Genome-wide internal tagging of bacterial exported proteins. *Nature Biotechnol*., 20:839–841.
2. Ewing, B., Hillier, L. et al., 1998, Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res.*, 8:175–185.
3. Ewing, B. and Green, P., 1998, Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res*., 8:186–194.
4. Gallagher, L., McKnight, S., Kuznetsova, M., Pesci, E., and Manoil, C., 2002, Functions required for extracellular quinolone signaling by *Pseudomonas aeruginosa*. *J. Bacteriol.*, 184:6472–6480.
5. Herrero, M., de Lorenzo, V., and Timmis, K.N., 1990, Transposon vectors containing non-antibiotic resistance selection markers for cloning and stable chromosomal insertion of foreign genes in gram-negative bacteria. *J. Bacteriol.,* 172:6557–6567.
6. Jacobs, M.A., Alwood, A., Thaipisuttikul, I., Spencer, D., Haugen, E., Ernst, S., Will, O., Kaul, R., Raymond, C., Levy, R., Chun-Rong, L., Guenthner, D., Bovee, D., Olson, M.V., Manoil, C., 2003, Comprehensive transposon mutant library of *Pseudomonas aeruginosa*. Proc. Natl. Acad. Sci. USA, 100:14339–14344.
7. Poole, K., 2004, Efflux pumps, pp. 635–674. In J.-L. Ramos (ed.), *Pseudomonas* vol. 1, Chapter 21. Plenum Publishers, New York.
8. Stover, C.K. et al., 2000, Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen. *Nature*, 406:959–964.

6

BIOGENESIS AND FUNCTION OF TYPE IV PILI IN *PSEUDOMONAS* SPECIES

Cynthia B. Whitchurch

Department of Microbiology
Monash University, Clayton
Victoria 3800, Australia

Key Words: fimbriae, twitching motility, natural transformation, infection, adhesion, biofilm

1. INTRODUCTION

Type IV pili or fimbriae are non-flagellar, filamentous surface appendages that are associated with a number of biological activities in bacteria. These processes include a form of surface translocation termed twitching motility; bacteriophage sensitivity; attachment to biotic (bacteria, plant, animal) and abiotic surfaces; biofilm development; and the uptake of naked DNA by natural transformation. Many of these biological functions are reliant on the ability of these structures to extend and retract.

Type IV pili/fimbriae are found throughout the eubacteria. They are produced by many species of Gram-negative bacteria and have been most extensively studied in *Pseudomonas aeruginosa*, *Neisseria gonorrhoeae*, *N. meningitidis*, *Dichelobacter nodosus*, *Moraxella bovis*, *Myxococcus xanthus*, and *Synechocystis* species PCC 6803, enteropathogenic *Eschericia coli* (EPEC) and *Vibrio cholerae*.[98,99,143] Type IV pili/fimbriae are also produced by some Gram-positive bacteria including *Ruminococcus albus*[163,168] and probably

Pseudomonas, Volume 4, edited by Juan-Luis Ramos and Roger C. Levesque

Streptococcus sanguis[100] though evidence for the latter awaits confirmation *via* molecular analyses.

The nomenclature "type IV" is derived from the classification scheme outlined by Ottow in 1975 in which he grouped non-flagellar filamentous structures into six types based largely on morphological characteristics.[156] Interestingly, throughout this nomenclature system Ottow reserved the use of the term "pili" to specifically refer to conjugative sex pili and used the term "fimbriae" for all other non-flagellar filamentous surface organelles. The term "fimbriae" was also favored by other investigators[72,82] however, historically, the terms "pili" and "fimbriae" have been used interchangeably. In reference to the non-flagellar filamentous structures that were classified by Ottow as "Group IV," the terminology "type IV pili" and its abbreviation "tfp" have become very much the vogue in recent years and will be adopted in this review.

As defined by Ottow (1975) type IV pili are ". . . flexible, rod-like, polarly inserted fimbriae." The tfp of *P. aeruginosa* were amongst the first of this class described in bacteria.[20,21,25,82,104,226] *P. aeruginosa* tfp have a diameter of 5–6 nm[25,50,77,226] and a hollow core of 1.2 nm.[77] Due to their retractile nature, the length of these structures varies significantly, but on average are about 1–4 μm in length[22,24,27] although they can range in length up to 10 μm.[25,226] A number of studies have examined tfp production by *P. aeruginosa* under different culture conditions and have shown that *P. aeruginosa* tfp are produced throughout plate culture;[22,119,209,226] in logarithmic broth culture with gentle agitation;[22,25,226] and by stationary phase bacteria when cultured without aeration or with only gentle agitation[22,119,209,226] whereas vigorously shaken stationary phase *P. aeruginosa* cultures do not produce detectable tfp.[22,119,209,226]

Type IV pili have also been identified in other *Pseudomonas* species including *P. syringae*,[180–183] *P. stutzeri*[84,98] and *P. fluorescens*[138] (see Section 2.6). Historically, however, *P. aeruginosa* has served as one of the primary model organisms for the investigation of tfp and its associated functions in bacteria. It is for this reason that this chapter will largely focus on the tfp of *P. aeruginosa*. This chapter will review the current state of knowledge relating to the structure, biogenesis, and function of the tfp of the genus *Pseudomonas* and will also highlight advances made in the study of these structures in other type IV piliated bacteria.

2. STRUCTURE AND BIOGENESIS OF TYPE IV PILI

By definition, type IV pili are long, flexible filaments attached at the poles of the bacterial cell.[156] In stark contrast to the relative simplicity of the external pilus fiber which is composed of several thousand copies of a single protein subunit,[81,159] the type IV pilus structure located within the bacterial cell wall

is a sophisticated multimeric protein machinery that mediates assembly and retraction of the extracellular tfp filaments and which shares many similarities to the type II protein secretion systems and archeal flagella.[162] The regulation of tfp biogenesis and function is also extremely complex involving the integration of numerous regulatory systems. To date, nearly 40 proteins have been identified that are involved in the biogenesis, function, and regulation of tfp in *P. aeruginosa*.

2.1. The Major Subunit PilA

The extracellular type IV pilus fiber is comprised of several thousand copies of a single subunit protein termed pilin or PilA in *P. aeruginosa*. *P. aeruginosa* PilA and related pilins are synthesized as a precursor with a short, basic, highly conserved N-terminal leader sequence that is removed by endoproteolytic cleavage between an invariant glycine residue and a phenylalanine that is then N-methylated in the mature protein[80,152,157,161,192,213] (Figure 1). Type IV pilin proteins from diverse species are highly conserved across the leader peptide and approximately the first 30 residues of the mature protein (Figure 1). This domain is highly hydrophobic and is thought to be involved in subunit–subunit interactions in the assembled filament.[50] This hydrophobic domain may also be involved in directing the pilin subunits to the inner membrane for processing by the endopeptidase PilD, and for accumulation of a subunit pool for filament assembly and pilin recycling during disassembly (see Sections 2.3 and 2.4). The remainder of the pilin molecule contains moderately conserved and hypervariable regions (Figure 1).

Type IV pilin proteins can be grouped into two subfamilies, the classic type IVa pilins include those from *P. aeruginosa*, *N. gonorrhoeae*, *D. nodosus*, and *M. bovis* whereas the pilins from second subgroup include those of the toxin-coregulated pili (TCP) of *V. cholerae* and the bundle-forming pili (BFP) of enteropathogenic *E. coli* are classified as type IVb pilins (Figure 1a). The Type IVb pilin precursors are closely related to the type IVa pilins but are synthesized with longer leader peptides, have either methionine, valine or leucine rather than phenylalanine as the first amino acid of the mature protein and are generally larger than the type IVa pilins. As shown in Figure 1b, the pilins of *P. syringae*, *P. fluorescens*, *P. stutzeri*, and *P. putida* are closely related to that of *P. aeruginosa* and would be classified as type IVa. A conserved feature of the *Pseudomonas* pilin proteins is the presence of an intrachain disulphide bond loop (DSL) located at the far C-terminus of the molecule (Figure 1b). In *P. aeruginosa* the DSL has been shown to contain the epithelial cell-receptor-binding domain which, despite a high degree of primary sequence divergence between *P. aeruginosa* strains, is antigenically conserved and demonstrates conservation of receptor specificity (see Section 4.1).

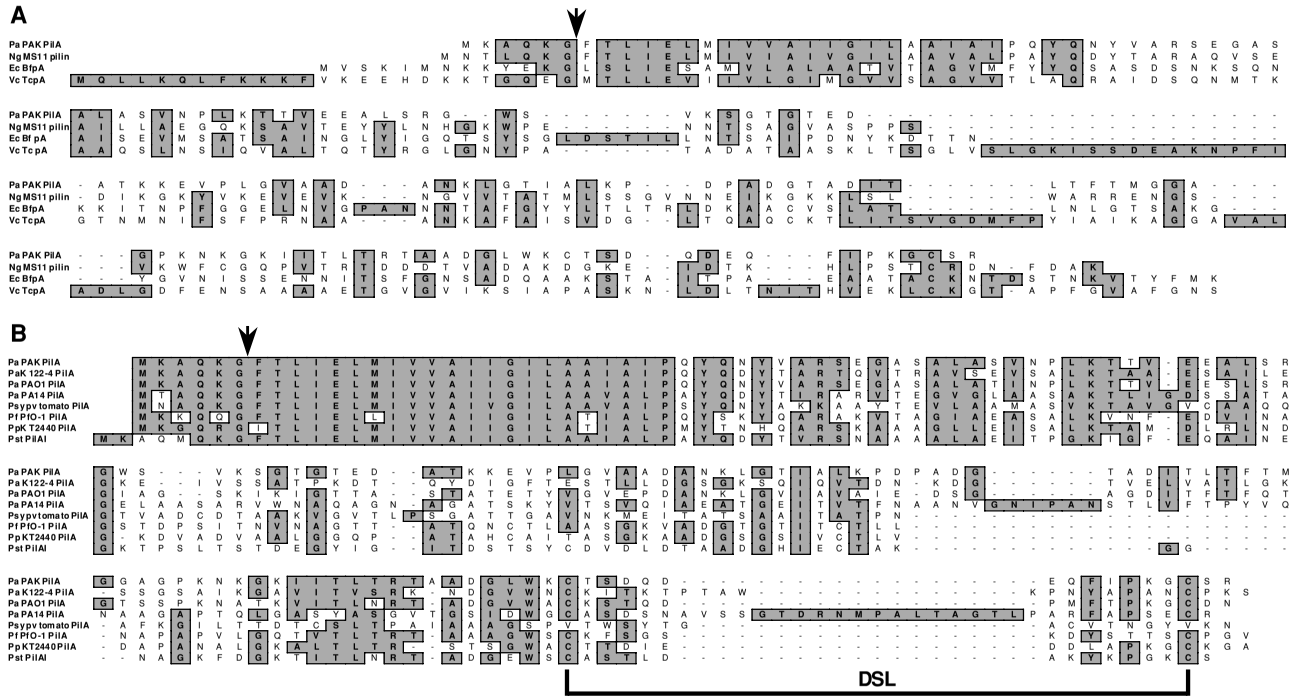


Figure 1. Alignment of type IV pilin sequences. (A). Alignment of type IVa pilins of *P. aeruginosa* (Pa) and *N. gonorrhoeae* (Ng), and type IVb pilins of enteropathogenic *E. coli* (Ec) and *V. cholerae* (Vc). Similar residues are shaded. Arrow indicates the site of endoproteolytic cleavage of the prepilin sequence. (B) Alignment of pilins from *P. aeruginosa* (Pa), *P. fluorescens* (Pf), *P. syringae* (Psy), *P. putida* (Pp) and *P. stutzeri* (Pst). Arrow indicates the site of endoproteolytic cleavage of the prepilin sequence. Similar residues are shaded. The C-terminal DSL is indicated. Alignments were performed using the CLUSTALW alignment tool in MacVector® Accelrys (Software Inc.).

The full-length crystal structures of *P. aeruginosa* strain PAK and *N. gonorroheae* strain MS11 pilin as well as structures of N-terminal (Δ1–28) truncations of the pilins from *P. aeruginosa* strains PAK and K122-4 and *V. cholerae* have been solved.[5,51,78,93,160] These structures indicate that overall, both type IVa and IVb pilins share a similar ladle-like architecture with an extended N-terminal α helical spine and a C-terminal globular head domain which exposes the hypervariable regions of the pilin subunits to the solvent. The first 28 residues of the N-terminal α helix are highly hydrophobic and are thought to be involved in packing of the pilin subunits into the fiber.[160] The remainder of the α helix pack onto the globular head domain which confers many of the biological functions of the tfp including adhesion to epithelial cells, phage attachment, and antigenic variability.[50]

A recent survey of pilins from more than 290 environmental, rectal, clinical, and cystic fibrosis (CF) isolates of *P. aeruginosa* has found that there are 5 distinct phylogenetic groups of *P. aeruginosa* pilin and that 4 of these are associated with a distinct accessory gene or set of genes that are located immediately downstream of *pilA*.[127] About 30% of pilins identified in this survey are not post-translationally modified and include the PilA proteins of strains PAK, PAO1, and K122-4.[127] The most common class of pilin alleles are associated with the previously identified accessory gene *pilO*,[37] the product of which is responsible for glycosylation of the C-terminal serine residue of the pilin of *P. aeruginosa* strain 1244.[37,38,46] To avoid confusion with the tfp biogenesis gene required for tfp assembly and which is also termed *pilO* (see Section 2.4), and to bring the nomenclature in line with that used for the other pilin accessory genes, Kus *et al.* have proposed that the 1244 *pilA*-associated accessory gene *pilO* should be renamed *tfpO*.[127] It is presumed that the accessory genes associated with other *pilA* alleles are also responsible for post-translational modification of the pilin subunit, though the nature of these modifications have yet to be determined.[127]

2.2. Structure of the Extracellular Tfp Filament

In 1981, Folkhard *et al.*[77] used X-ray diffraction to examine the structure of purified tfp of *P. aeruginosa* strains PAO1 and PAK.[77] They determined that the X-ray diffraction patterns were best satisfied if the tfp were comprised of pilin subunits arranged in a helical symmetry of 4 subunits per turn with each turn having a pitch of 4.1 nm. At the time of these analyses, the pilin subunits of PAK and PAO1 were estimated to be 18–19 kDa based on mobility in SDS polyacrylamide gels and amino acid composition[81,159] and Folkhard *et al.*[77] determined that at high-subunit-packing density, 4 subunits per turn could be accommodated in the PAK and PAO1 tfp filaments.[77] When more accurate estimations of pilin subunit size (~15 kDa) were obtained from amino acid sequence analysis of PAK pilin[193], the helical symmetry of the PAK and PAO1

tfp fibers was reinterpreted to be 5 subunits per turn.[224] Recently, however, computational modeling of the PAK pilus filament combining both the X-ray diffraction data of Folkhard *et al*.[77] and the X-ray crystal structure of full-length PAK pilin[51] predicts that there is insufficient space in the tfp filament to accommodate 5 subunits per turn without unacceptable steric clashes.[50] This computational modeling approach does, however, predict that the *P. aeruginosa* PAK pilin subunits can be satisfactorily arranged in a right-handed one-start helix with 4 subunits per turn and a pitch of 4.1 nm.[50] Also consistent with the structural and diffraction data is a model for the PAK tfp fiber in which pilin is arranged in a left-handed 3 start helix with 4 subunits per turn with a pitch of 12.2 nm per single-turn of a helix and strands separated by 4.1 nm.[50] In either case it appears that at least in the tfp filaments of *P. aeruginosa* PAK, pilin subunits are arranged in a helical fiber at 4 subunits per turn. Thus it appears that the original model by Folkhard *et al*.[77] may well have been accurate.

There also remains some controversy regarding the polarity of the modeled tfp filament relative to the bacterial cell as it is unclear which end of the assembled filament models represents the distal tip and which is the base. The original *N. gonorrhoeae* tfp filament model upon which the subsequent *P. aeruginosa* filament models have been based, proposed that the long hydrophobic N-terminal domains might be inserted into the bacterial membrane.[160] Recently, however, other investigators have proposed an alternate polarity to the filament model arguing greater accessibility to the DSL domain at the filament tip.[93] To add to the confusion, the recently revized 4 subunit per turn *P. aeruginosa* filament model predicts greater solvent accessibility of the DSL domain in the original filament orientation.[50] It is unclear at this time which polarity accurately reflects the biological arrangement of the tfp filament relative to the bacterial cell.

2.3. Retraction of Type IV Pili

Type IV pili are associated with an astounding number of biological functions including bacteriophage sensitivity, natural transformation, translocation across solid surfaces *via* twitching motility, biofilm development, and attachment to host cell surfaces. Many if not all of these functions are dependent upon the ability of tfp to retract.

The retractile nature of type IV pili was first proposed by David Bradley during his studies on the infection mechanisms of bacteriophage that specifically bind to the tfp of *P. aeruginosa*. In these studies, Bradley observed that tfp-specific bacteriophage were located at the bacterial cell surface at the junction of the pilus and the bacterial cell pole of wildtype *P. aeruginosa* cells.[21–29] He also noted that for many phage types, adsorption resulted in an increase in both the average number of pili/pole and the average length of the

tfp.[22,23,25,26,30] Similar results were obtained with labeling experiments using pili-specific antisera.[22] Bradley interpreted these observations as evidence that tfp are retractile and that tfp retraction was inhibited by either phage or antibody adsorption.

Bradley also investigated the effects of phage or antibody adsorption to the tfp of a class of *P. aeruginosa* mutants that were resistant to tfp-specific phage and that appeared hyperpiliated by electron microscopy. With this class of mutants, tfp-specific bacteriophages were found to be distributed randomly along the pili, and phage or antibody adsorption did not result in significant alterations in piliation levels. Bradley concluded from these observations that the tfp present on these mutants were non-retractile.[22,23,28,30] During their studies with hyperpiliated bacteriophage ϕ6-resistant derivatives of *P. syringae*, Romantschuk and Bamford also concluded that the pili to which this bacteriophage binds are retractile.[183] These ϕ6-specific pili were later confirmed to be tfp by molecular analyses.[180]

Twitching motility is a mode of surface translocation demonstrated by many bacteria that possess type IV pili.[98,99] In 1980, David Bradley proposed that the motive force for twitching motility was attributable to the retraction of tfp.[29] This hypothesis was based upon his observations that *P. aeruginosa* mutants which either lacked tfp or possessed non-retractile tfp did not demonstrate twitching motility and that prevention of tfp retraction by adsorption of pili-specific phage or antisera was inhibitory to twitching motility.[29] Bradley's hypotheses that tfp are retractile and that tfp retraction is the motive force powering twitching motility were considered somewhat controversial at the time.[99] However, both tfp retraction and its role in twitching motility have been confirmed in recent years. Molecular analysis of the hyperpiliated, non-twitching, phage resistant *P. aeruginosa* mutants used by Bradley has demonstrated that these mutants contained small deletions in the gene *pilT*[227,232] and it is now evident that PilT is a powerful molecular motor which provides the driving force for tfp retraction.[145]

In a beautiful study by Skerker and Berg, direct observation of extension and retraction of tfp *in vivo* was achieved for the first time.[206] In this study, *P. aeruginosa* cells were labeled with an amino specific fluorescent dye and pilus function directly visualized *via* internal reflection microscopy. Using this technique tfp filaments were generally seen to emerge from only one pole of the cell and to extend and retract independently of one another. This study also confirmed that cellular translocation during twitching motility is achieved through retraction and not extension of the type IV pilus as tfp that were attached to the quartz surface *via* their distal end were observed to retract and pull the *P. aeruginosa* cells forward toward the tethered end whereas pili extension was never observed to cause cellular translocation.[206]

Tfp-dependent social gliding motility in *M. xanthus* is equivalent to twitching motility seen in other type IV piliated bacteria.[143,199] During

examination of this motility on methylcellulose covered surfaces, Sun *et al.*[215] noted that occasionally wildtype *M. xanthus* cells could be seen standing perpendicular to the surface tethered by the polar tfp. These cells were able to move vertically up and down until they were retracted to the surface.[215] In contrast, vertically tethered cells of *M. xanthus pilT* mutants which possess non-retractile tfp, were never observed to move[215] confirming that tfp retraction is responsible for cellular motility and that PilT is essential for this process.

Independent confirmation that tfp retract and that tfp retraction provides the force that powers twitching motility has been obtained in series of elegant studies in which laser tweezers were used to measure the forces generated by retracting tfp of *N. gonorrhoeae*.[140,145] *N. gonorrhoeae pilT* mutants were found to be incapable of generating retractile forces indicating that PilT is the molecular motor that powers tfp retraction. In wildtype *N. gonorrhoeae*, single pilus filaments were able to retract with forces measured at greater than 100 pN making PilT the strongest molecular motor described to date.[140]

The mechanism of tfp retraction is thought to occur *via* disassembly of the tfp filament into pilin subunit components. The rate of filament disassembly has been estimated to occur at about 1000 subunits/sec.[145] It has been recently estimated that in the absence of pilin recycling, tfp biogenesis would require most of the protein synthesis capacity of the bacterial cell,[144] thus it seems highly likely that pilin subunits are recycled into the tfp filament during successive rounds of tfp filament extension and retraction. Supportive evidence for this model comes from the studies of Skerker and Berg who observed that there was no apparent loss of fluorescence intensity of the fluorescently labeled *P. aeruginosa* tfp during cycles of extension and retraction.[206] This observation suggests that the labeled pilin subunits are being recycled into newly assembled tfp. A recent study using electron microscopy and de-biotinylation protection assays in *N. gonorrhoeae* has demonstrated that tfp filament retraction involves translocation of the extracellular pilin subunits to pools in the cytoplasmic membrane.[147] Thus it appears that during tfp retraction, pilin subunits are rapidly disassembled and reside in cytoplasmic membrane pools to be recycled into new filaments, thereby providing enormous energy conservation for the cell.

2.4. Components of the *P. aeruginosa* Tfp Structure and Biogenesis Machinery

In addition to the major pilin subunit PilA, over 40 other proteins have been identified that are involved in the biogenesis, function, and regulation of tfp and twitching motility in *P. aeruginosa* (Table 1; Figures 2 and 3). In this section, proteins that are involved in the structure, assembly or mechanical function of *P. aeruginosa* tfp will be described. Section 2.5 will discuss genes involved in the regulation of tfp biogenesis and function.

Table 1. *Pseudomonas* proteins involved in the biogenesis, function, and regulation of tfp

P. aeruginosa protein	Proposed function	[a] *P. syringae* pv. *tomato* str. DC3000	[a] *P. fluorescens* PfO-1	[a] *P. putida* KT2440
Biogenesis and function				
PilA	Major pilin subunit	39/50	36/49	39/52
PilB	ATPase – assembly motor	79/80	78/89	Absent
PilC	Pilus assembly	73/87	74/86	43/69
PilD (XcpA)	Prepilin peptidase	79/87	68/82	64/78
FimT	Pseudopilin	28/44	27/42	24/44
FimU	Pseudopilin	31/47	29/47	Absent
PilV	Pseudopilin	35/71	36/70	Absent
PilW	Pseudopilin	33/55	40/59	Absent
PilX	Pseudopilin	38/66	32/66	33/57
PilY1	Retraction regulator	24/39	25/39	Absent
PilY2	Unknown	Absent	Absent	Absent
PilE	Pseudopilin	62/81	35/51	32/48
PilM	Pilus assembly	85/93	82/90	31/53 (N-terminus only)
PilN	Pilus assembly	73/86	68/83	27/51
PilO	Pilus assembly	74/85	74/82	24/40 (fusion with PilP)
PilP	Lipoprotein, PilQ assembly pilotin	72/87	61/78	37/57 (fusion with PilO)
PilQ	Secretin	72/83	72/82	59/75 (C-terminus only)
PilT	ATPase – retraction motor	87/96	87/94	50/74
PilU	ATPase – PilB/PilT modulator?	Absent	Absent	Absent
PilF	Assembly	62/81	64/80	55/71
PilZ	Assembly	88/94	88/95	Absent
FimV	Peptidoglycan remodeling?	45/58	47/60	43/54
TonB3	PMF energy transducer, component of TonB import pump	69/83	72/84	68/84
PA2982	ExbD homolog, component of TonB import pump	60/77	62/79	62/74
Regulation				
PilS	Sensor kinase; controls *pilA* transcription	62/77	60/76	Absent
PilR	Response regulator; controls *pilA* transcription	77/86	79/87	Absent

(*Continued*)

Table 1. (*Continued*)

P. aeruginosa protein	Proposed function	[a] *P. syringae* pv. *tomato* str. DC3000	[a] *P. fluorescens* PfO-1	[a] *P. putida* KT2440
RpoN	Sigma factor; controls *pilA* transcription	84/93	85/93	86 /93
FimS (AlgZ)	Atypical sensor kinase; controls *fimT-pilE* gene cluster transcription	71/81	Absent	Absent
AlgR	Response regulator; controls *fimT-pilE* gene cluster transcription	82/93	81/91	78/88
PilG	Chemosensory, CheY homolog	87/90	89/92	79/91
PilH	Chemosensory, CheY homolog	91/96	90/95	77/89
PilI	Chemosensory, CheW homolog	58/74	63/73	058/72
PilJ	Chemosensory, ligand binding MCP	77/86	78/88	76/88
PilK	Chemosensory, CheR homolog	Absent	Absent	Absent
ChpA	Chemosensory, CheA/CheY hybrid	51/65	49/62	48/64
ChpB	Chemosensory, CheB homolog	Absent	Absent	Absent
ChpC	Chemosensory, CheW homolog	59/70	57/74	55/66
FimL	Controls *vfr* transcription	Absent	Absent	Absent
Vfr	cAMP Binding transcriptional regulator; multiple targets	78/87	82/89	81/85
FimX	Multidomain signal transduction protein	Absent	Absent	Absent
Crc	Catabolite repression control; influences *pilA, pilB* transcription	86/93	88/94	86/93
RpoS	Stationary phase sigma factor	88/93	88/92	87/93
Ppk	Polyphosphate kinase	81/90	81/91	82/91

[a] percent identity/percent similarity to *P. aeruginosa* PAO1 protein. Determined by BLAST analyses of genome sequences at NCBI (http://www.ncbi.nlm.nih.gov/sutils/genom_table.cgi).

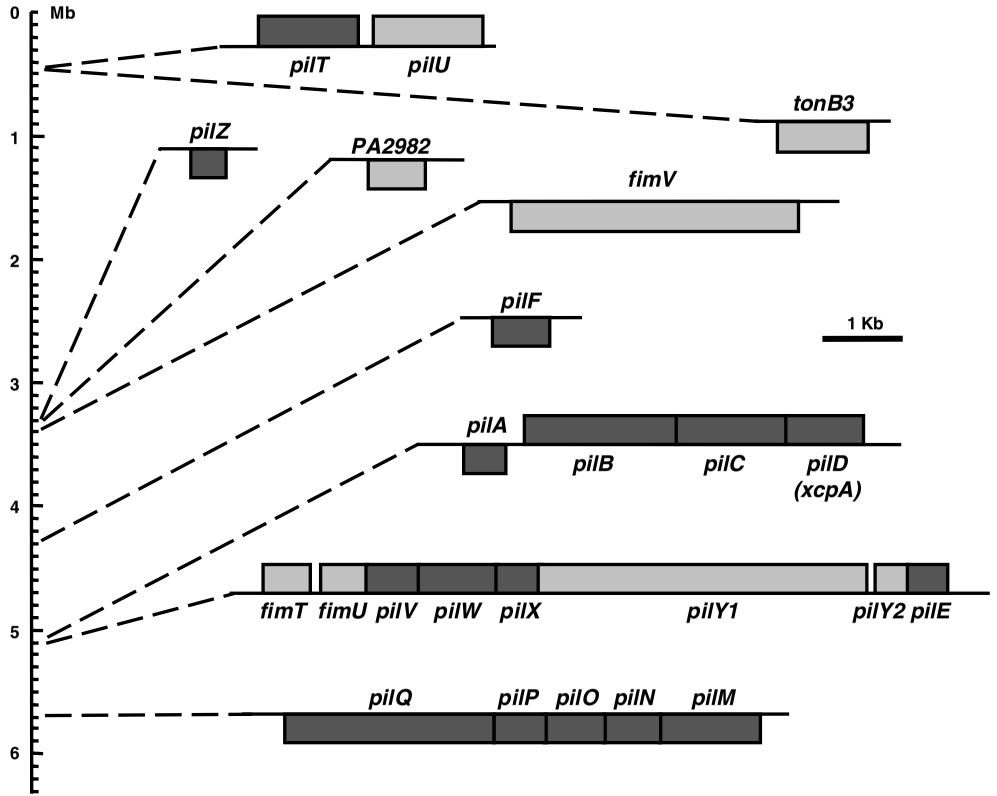


Figure 2. Genes involved in the structure, assembly and mechanical function of type IV pili in *P. aeruginosa* are dispersed around the 6.3 Mb chromosome of PAO1.[212] Genes are depicted by shaded boxes. Genes that confer resistance to tfp-specific bacteriophage when disrupted are depicted in dark shading. The direction of transcription is indicated by position of the box relative to the line (above is plus strand and below is minus strand).

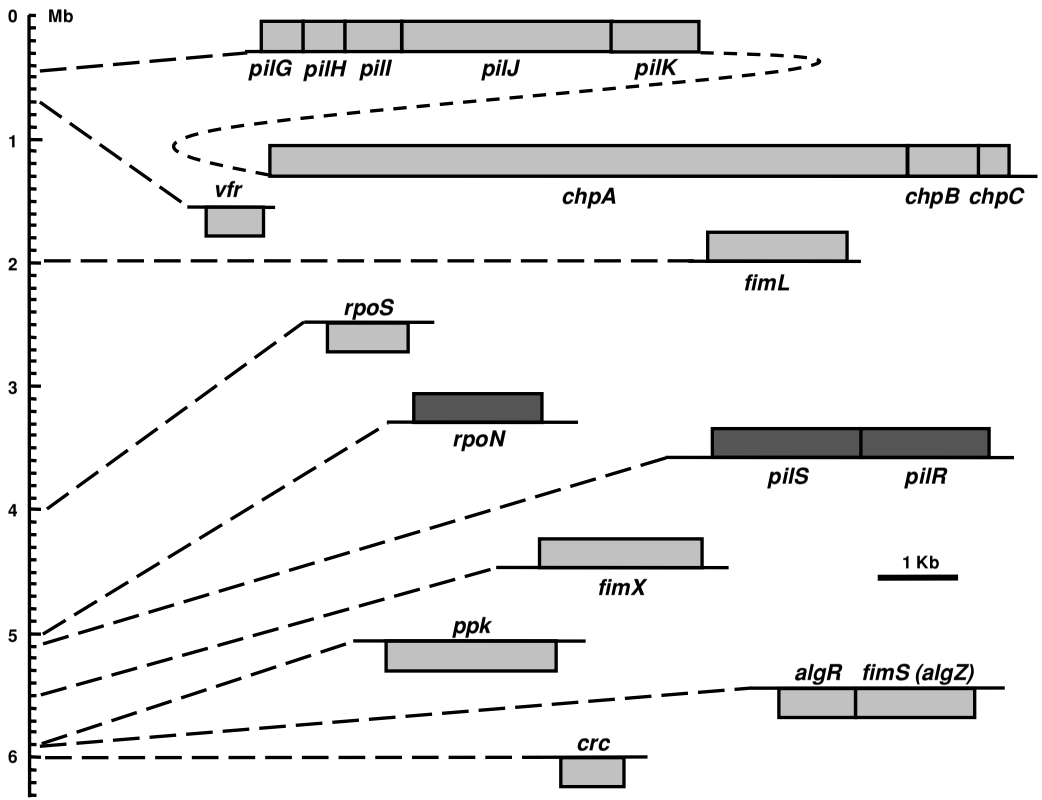


Figure 3. Genes involved in regulation of type IV pili biogenesis and function in *P. aeruginosa* are dispersed around the 6.3 Mb chromosome of PAO1.[212] Genes are depicted by shaded boxes. Genes that confer resistance to tfp-specific bacteriophage when disrupted are depicted in dark shading. The direction of transcription is indicated by position of the box relative to the line (above is plus strand and below is minus strand).

The majority of the structural and biogenesis tfp genes are found in four main gene clusters on the *P. aeruginosa* PAO1 genome (Figure 2). Many of the tfp structural and biogenesis proteins are multiply homologous to components of bacterial type II secretion systems. In recent years, it has become increasingly apparent that tfp, type II secretion systems and archeal flagella are related machineries which are thought to be comprised of "pseudopilus-like" structures in the cell wall.[162,195,220]

Situated immediately upstream of *pilA* and in the opposite transcriptional orientation are *pilB*, *pilC*, and *pilD* each of which are absolutely required for tfp assembly.[151] Mutants of *pilB*, *pilC*, and *pilD* do not produce surface assembled tfp, are non-twitching and resistant to tfp-specific bacteriophage. PilC is a polytopic inner membrane protein that is essential for tfp assembly though the exact role for this protein is unknown. PilD is a bifunctional enzyme that is required for both cleavage of the prepilin leader sequence and N-methylation of the N-terminal phenylalanine of PilA.[213] PilD is also referred to as XcpA and is responsible for processing of the prepilin-like components of the type II secretion system of *P. aeruginosa*.[7] PilB is predicted to be an NTPase that provides energy for pilus assembly and is homologous to the molecular motor PilT that powers tfp retraction as well as to PilU which also appears to play a role in tfp retraction.

Both *pilT* and *pilU* are situated together in a region of the genome separate from other tfp biogenesis genes (Figure 2). As noted above, loss of *pilT* renders the tfp of the mutant non-retractile and as a consequence these strains are hyperpiliated, non-twitching, and resistant to tfp-specific bacteriophage.[227,232] In contrast, *pilU* mutants are hyperpiliated and non-twitching yet remain sensitive to these bacteriophage.[232] This latter phenotype suggests that despite the hyperpiliated phenotype, the tfp of *pilU* mutants are sufficiently retractile to enable phage infection. The role of PilU in tfp biogenesis and function is unclear but this protein may serve to modulate the antagonistic activities of either PilB and/or PilT. PilT has recently been shown to be the molecular motor that powers tfp retraction and twitching motility in *N. gonorrhoeae* and is capable of producing forces exceeding 100 pN[140,145] (see Section 2.3).

PilT proteins from *Aquifex aeolicus* and *Synechocystis* species PCC6803 and the PilB protein from the plasmid R64 thin pili tfp systems have been experimentally demonstrated to have ATPase activity.[101,155,190] Indeed PilB, PilT, and PilU belong to a large family of bacterial ATPases involved in tfp assembly and retraction, type II secretion, type IV secretion and DNA uptake.[34,42,76,162,164,191] Members of this family are cytoplasmic proteins that are associated with the inner face of the inner membrane *via* protein–protein interactions with integral membrane components of the multimeric protein assembly.[53,76,191] PilT and related proteins from type II and type IV secretion systems have been shown to oligomerise into large hexameric ring structures that are thought to change conformation upon ATP binding.[79,177,196,242] It is not yet clear how these molecular

machines harness the energy from ATP binding and hydrolysis to power the various biological functions associated with these systems.

The third cluster of tfp structural and biogenesis genes encodes the pilin-like proteins FimT, FimU, PilV, PilW, PilX, and PilE (Figure 2). Each of these proteins possesses a prepilin-like N-terminal leader peptide which is presumably cleaved by PilD, though this has not been formally demonstrated. They also contain the signature hydrophobic N-terminal α-helical domain that is thought to be involved in protein–protein interactions and in directing the proteins to the inner membrane (see Section 2.1). It is possible that these pilin-like proteins assemble into a pseudopilus structure that spans the periplasm forming the base of the tfp structure. Alternatively they may be involved in initiating, modulating, or capping PilA assembly into the tfp filament. PilV, PilW, PilX, and PilE are each essential for tfp assembly, sensitivity to tfp-specific bacteriophage and twitching motility suggesting a crucial role in tfp assembly.[2–4,188] Interestingly, *fimU* mutants are non-twitching, do not produce detectable surface assembled tfp but retain sensitivity to the tfp-specific bacteriophage PO4, though the plaques produced are smaller and more turbid than wildtype,[4] suggesting the presence of some tfp structures at the cell-surface sufficient for phage infection. In contrast, FimT does not appear to be required for tfp assembly, PO4 phage sensitivity or twitching motility though it can functionally substitute for FimU to restore twitching motility to *fimU* mutants.[4] The exact roles for each of these pilin-like proteins in tfp assembly and function, and their intermolecular interactions have yet to be determined.

Also located within this gene cluster are the genes *pilY1* and *pilY2* (Figure 2). Mutants of *pilY1* are non-twitching, show no surface assembled tfp but have wildtype sensitivity to tfp-specific bacteriophage[2] which suggests that these mutants produce some surface tfp sufficient for phage binding and infection. PilY2 is a small protein that may play a subtle role in tfp biogenesis and is found exclusively in preparations of sheared surface pili suggesting association with the tfp filament.[2] PilY1 is also found associated with sheared tfp but unlike PilY2, is mainly found in the membrane fraction. BLAST search analyses show that the C-terminal region of PilY1 is homologous to the C-terminal regions of the PilC proteins of *N. gonorrheae* and *N. meningitidis*. The Neisserial PilC proteins have been recently found to regulate retraction of the tfp filament possibly by acting as a clip on the pilus filament.[147] Given the homology between the C-terminal domains of *P. aeruginosa* PilY1 and the Neisserial PilC proteins, it would be very interesting to determine if PilY1 functions in a similar manner to control tfp retraction.

The fourth cluster of *P. aeruginosa* tfp assembly/structural genes encodes the proteins PilM, PilN, PilO, PilP, and PilQ (Figure 2). Mutations in these genes block surface assembly of tfp and confer a non-twitching, phage resistant phenotype.[141,142] PilQ is a member of a family of bacterial integral outer membrane proteins termed secretins. Secretin family members are components of

complex multicomponent protein machineries including the type II and type III secretion pathways, the machinery responsible for filamentous phage biogenesis, and tfp.[13] Secretins form large, stable multimeric ring-shaped gated pores in the outer membrane through which macromolecules are transported. Transmission electron microscopy analyses of *P. aeruginosa* and *N. meningitidis* PilQ complexes have shown that PilQ monomers assemble into doughnut shaped complexes comprised of 12 subunits with an internal diameter of about 50–60 nm which closely resembles the diameter of the tfp filament.[14,43–45] Recent structural analysis of the *N. meningitidis* PilQ complex indicates that the PilQ homododecamer is probably comprised of a tetramer of trimers.[45] Elegant genetic evidence obtained in *N. gonorrhoeae* has confirmed that PilQ proteins form the transport channel in the outer membrane through which the tfp fiber is extruded.[235] *N. gonorrhoeae* mutants lacking both PilQ and the tfp retraction motor PilT were observed to produce tfp structures that were unable to extrude through the outer membrane due to the absence of the PilQ secretin and thus were trapped in the periplasm forming long, pili-like membrane-bound extensions.[235] Stable oligomerisation of *N. gonorrhoeae* PilQ has been shown to require the lipoprotein PilP.[70] Presumably the *P. aeruginosa* ortholog (PilP) also serves a similar function.

The functions of PilM, PilN, and PilO in tfp biogenesis are currently unknown. PilN and PilO both have long N-terminal stretches of hydrophobic residues that may serve to anchor the proteins to the cytoplasmic membrane. PilM shows homology to the actin-like prokaryotic proteins MreB (rod-shape determining) and FtsA (cell-division) and by primary sequence is predicted to belong to the actin-like ATPase superfamily of proteins.[73,136,142] Given these homologies, it is possible that PilM may undergo ATP-dependent oligomerisation. The observation that deletion mutants of PilM are dominant negative when introduced into wildtype *P. aeruginosa* does suggest that PilM may be part of a multimeric complex.[142]

Type IV pili assembly also requires the genes *pilF* and *pilZ* which are each located at distinct loci on the *P. aeruginosa* genome (Figure 2). Mutants of these genes do not assemble tfp, are non-twitching and resistant to PO4 phage indicating that these proteins have essential roles in tfp assembly.[1,223] The function of PilZ is unknown though interestingly, *pilZ* may be transcriptionally coupled to *holB* which encodes the δ subunit of DNA polymerase III.[1] PilF possesses a signal sequence and 5 tandemly arranged 34 amino acid tetratricopeptide repeat (TPR) domains that mediate protein–protein interactions and the assembly of multiprotein complexes.[15,54] The interacting target protein(s) of PilF and the role of PilF in *P. aeruginosa* tfp biogenesis are unknown. Interestingly the outer membrane located Tgl protein of *M. xanthus* is similar in size to PilF and possesses a signal sequence as well as 6 TPR domains. Mutants of *M. xanthus* that are *tgl*- and *pilA*+, or *tgl*+ and *pilA*- do not produce surface tfp and are non-motile. However when mixed

together the *tgl+/pilA-* cells can stimulate the *tgl-/pilA+* cells to assemble tfp on their surface and restore motility.[178,179,203,221,222] It remains to be determined if *P. aeruginosa* PilF plays a similar role in contact stimulation of tfp assembly.

Tfp biogenesis and twitching motility in *P. aeruginosa* also requires the activities of the gene products of *fimV, tonB3*, and PA2982.[105,198] Mutants of *fimV* and *tonB3* have been shown to produce little to no surface assembled fimbriae and are sensitive to tfp-specific bacteriophage.[105,198] These mutants are capable of some twitching motility though this is very aberrant compared to the highly co-ordinated process observed with wildtype *P. aeruginosa* (see Section 5.2). Interestingly, *P. aeruginosa fimV*, *tonB3*, and PA2982 mutants exhibit reduced virulence in the fruit fly *Drosophila melanogaster* whereas mutants that are completely impaired in tfp biogenesis retain wildtype virulence in the fruit fly model.[55] These observations suggest that tfp and twitching motility are not required for fruit fly killing and that FimV, TonB3, and PA2982 have additional roles to those involved in tfp biogenesis.

FimV is a highly acidic protein that contains a putative peptidoglycan-binding domain and is probably located in the cytoplasmic membrane.[198] Over-expression of FimV causes extreme elongation of the cells and it has been proposed that FimV might be involved in remodeling of the peptidoglycan layer to enable assembly of the tfp structure and associated machinery.[198]

The TonB systems of Gram-negative bacteria function to transduce the protonmotive energy of the cytoplasmic membrane for the active uptake of receptor-bound ligands at the outer-membrane.[166] These systems also appear to provide energy to efflux pumps for the export of antibiotics and toxic solvents from the cell.[169] The TonB systems are a complex of three membrane-spanning proteins, TonB, ExbB, and ExbD.[166] The *P. aeruginosa* genome encodes 3 *tonB* genes, only one of which (*tonB3*) is required for normal twitching motility and assembly of surface tfp.[105] Interestingly PA2982 is an ExbD homolog and like TonB3 is required for both twitching motility and *Drosophila* killing.[55] It is not yet known if this ExbD homolog complexes with TonB3 nor what role the TonB3 system of *P. aeruginosa* plays in tfp biogenesis and twitching motility. Since the Gram-negative TonB systems are involved in transducing proton-motive energy to the outer-membrane, it is possible that the *P. aeruginosa* TonB3/PA2982 system is involved in providing energy to an outer-membrane protein or outer-membrane-associated protein that is involved in tfp biogenesis, function, or response to environmental cues. Interestingly the outer-membrane tfp secretin PilQ possesses a short highly conserved domain (STN) which has been identified by the Pfam protein families database to be present in many members the bacterial secretin family as well as in many TonB-dependent receptor proteins (http://www.sanger. ac.uk/Software/Pfam).[9] Thus it is possible that the TonB3/PA2982 system may interact with PilQ to provide energy for PilQ function in the outer-membrane.

2.5. Regulation of Tfp Biogenesis and Function

Whilst much is known about the molecular genetics of tfp biogenesis in *P. aeruginosa*, very little is known about the environmental cues and signal transduction events that govern tfp biogenesis and function. In recent years, it has become evident that the control of tfp biogenesis and twitching motility in *P. aeruginosa* is a complex process involving the integration of a large number of regulatory proteins (Table 1; Figure 3). This section will address the various regulatory systems and proteins that have thus far been shown to influence tfp gene expression and to control tfp function in *P. aeruginosa*.

The two component sensor-regulator pair PilS and PilR along with the alternate sigma factor RpoN (σ^{54}) are essential for transcription of the pilin subunit gene *pilA*.[18,102,109,110] Based on current knowledge of PilS structure and function PilS can be divided into three regions, an N-terminal transmembrane region (residues 1–174), a cytosolic linker region (175–296) and a cytosolic C-terminal transmitter domain (297–530).[74] The N-terminal domain is highly hydrophobic, forms six transmembrane helices and is required for membrane insertion and probably detection of the activating stimulus which has not yet been identified. The linker domain is required for localization of PilS to both poles of the cell, placing PilS in the same subcellular region as the tfp structures and pilin subunit pools. The significance of this co-localization is unclear and may be co-incidental. Interestingly PilS is not polarly localized in *E. coli* suggesting interaction of the PilS linker domain with a specific polarly located molecule in *P. aeruginosa*.[17] Tfp and the polar flagellum have been ruled out as the polar anchor.[17] The C-terminal domain of PilS contains all of the conserved residues that are characteristic of sensor kinases including the invariant histidine residue (H319), which is probably the site of autophosphorylation. PilS is believed to transfer this high-energy phosphate to the conserved aspartate residue in the N-terminal response regulator domain of PilR.[17,19,74]

The response regulator PilR is a typical RpoN-dependent transcriptional activator and as such possesses a N-terminal response regulator domain, a central ATPase domain and a C-terminal DNA-binding domain which mediates binding of activated PilR to 4 *cis*-acting sequences upstream of the *pilA* promoter. PilR interacts with the σ^{54}-RNA polymerase holoenzyme to initiate *pilA* transcription.[102,110,113] PilS also appears to act as a phosphatase of PilR and thus may serve dual functions in the control of *pilA* transcription.[19]

PilS and PilR and the alternate sigma factor σ^{54} are essential for *pilA* transcription in *P. aeruginosa* and consequently mutants of *pilS*, *pilR*, or *rpoN* do not produce surface assembled tfp, are resistant to tfp-specific bacteriophage and are non-twitching.[18,102,110] Interestingly, expression of *pilA* from a heterologous promoter in a *P. aeruginosa rpoN* mutant is able to restore tfp production and phage sensitivity to this mutant.[123] This observation suggests that there are no other RpoN-dependent genes that are required for tfp biogenesis in *P.*

aeruginosa. This observation also suggests that there are no other targets of PilS and PilR that are essential for tfp biogenesis and function.

The atypical sensor-regulator pair comprised of FimS and AlgR is also required for tfp biogenesis and twitching motility in *P. aeruginosa*.[228] FimS is also sometimes referred to as AlgZ,[244] but since there exists another unrelated *P. aeruginosa* protein which is also termed AlgZ,[10] for the sake of simplicity the product of the gene located upstream of *algR* will be referred to as FimS here. Mutants of *fimS* or *algR* show very-defective twitching motility and are sensitive to tfp-specific bacteriophage. *fimS* and *algR* mutants produce normal levels of cell-associated pilin though only a small amount of assembled tfp can be detected at the cell surface suggesting a defect at the level of tfp assembly.[228,230]

Recent whole genome transcriptional analyses have shown that AlgR is a global positive and negative regulator of numerous *P. aeruginosa* genes, many of which are involved in *P. aeruginosa* pathogenesis.[36,134] Microarray analyses indicated that the genes *fimU*, *pilV*, *pilW*, *pilX*, *pilY1*, and *pilY2* were potentially activated by AlgR. Expression of the whole gene cluster encompassing *fimT*, *fimU*, *pilV*, *pilW*, *pilX*, *pilY1*, *pilY2*, and *pilE* from an exogenous promoter was shown to complement twitching motility to the *algR* mutant, confirming that these genes are the tfp-related transcriptional targets of AlgR. This study also concluded that the *pilM*, *pilN*, *pilO*, *pilP*, *pilQ* operon could be indirectly activated by AlgR.[134] The authors of the latter study also reported a personal communication from M. Wolfgang stating that the genes *fimU*, *pilV*, *pilW*, *pilX*, *pilY1*, *pilY2*, and *pilE* have also been identified through microarray analyses as amongst the most FimS-dependent genes.[134] Thus it appears that FimS and AlgR control tfp biogenesis and function through control of expression of the genes *fimU*, *pilV*, *pilW*, *pilX*, *pilY1*, *pilY2*, and *pilE*. Interestingly mutants of these target genes produce no detectable surface pili and are resistant to tfp-specific bacteriophage (see Section 2.4) whereas *fimS* and *algR* mutants produce small amounts of surface assembled pili and are phage sensitive.[228,230] These observations suggest that there is a low level of target tfp gene expression in *fimS* and *algR* mutants that allows assembly of tfp structures sufficient for phage infection.

AlgR possesses a conserved response regulator domain at its N-terminus and a novel LytTR-type DNA-binding domain located at the C-terminus.[150] Site-directed mutagenesis of the highly conserved phosphate-accepting aspartate (D54) in the N-terminal response regulator domain of AlgR has shown that this residue is essential for tfp biogenesis and twitching motility in *P. aeruginosa*.[230] AlgR has also been shown to be phosphorylated *in vitro* at this residue *via* the *E. coli* histidine kinase CheA.[230] Taken together these observations suggest that phosphorylation of AlgR at D54 is required for expression of the *fimT–pilE* locus.

Transcriptional activators such as AlgR are generally phosphorylated at the conserved aspartate in their response regulator module by their cognate histidine kinase.[210] The putative cognate partner of AlgR, FimS, possesses

many of the hallmarks of a 2-component histidine kinase including the presence of a conserved histidine that would be the predicted site of autophosphorylation. However, FimS lacks several of the conserved residues that are essential for nucleotide binding and thus would appear to be incapable of autophosphorylation.[228] Interestingly loss of *fimS* results in similar phenotypes with respect to tfp biogenesis and function as mutants expressing the D54N allele of AlgR,[228,230] which suggests that FimS may have some role in phosphorylation of AlgR, possibly by acting as an intermediate in a phosphorelay cascade originating with another histidine kinase. The manner in which AlgR is phosphorylated *in vivo* and interacts with FimS to control tfp gene expression, and the environmental signals to which these proteins respond remain to be determined.

Type IV pili biogenesis and twitching motility in *P. aeruginosa* is controlled by a complex signal transduction pathway which has many modules in common with chemosensory systems that control flagella rotation in bacteria. This chemotaxis-like system is referred to as the Chp system and is comprised of the *P. aeruginosa* proteins PilG, PilH, PilI, PilJ, PilK, ChpA, ChpB, and ChpC.[57–59,231] The most extensively studied bacterial chemosensory systems are those that control swimming chemotaxis in enteric species.[225,234] The enteric chemotaxis systems are comprised of membrane bound chemoreceptors referred to as methyl-accepting chemotaxis proteins (MCPs). MCPs interact directly with both the adaptor protein CheW and the multidomain histidine protein kinase CheA. The MCPs function to modulate the kinase activity of CheA in response to chemical stimuli. CheA possesses in its C-terminus conserved motifs required for nucleotide binding. Autophosphorylation occurs at a conserved histidine in the N-terminus of the protein in a conserved HPt domain. The high-energy phosphoryl group is then transferred from the histidine of CheA to a conserved aspartate in the response regulator CheY, which then interacts directly with the flagellar motor to switch the direction of flagellar rotation. The rate of autocatalytic dephosphorylation of CheY-P is enhanced through interaction with CheZ. Sensory adaptation which allows temporal control of swimming motility occurs through methylation of specific glutamate residues on the MCP to reset it into a non-signaling state. The methylation status of the MCP is adjusted *via* competing activities of the methyltransferase CheR and the methylesterase CheB. CheB also possesses a response regulator module and is competitively phosphorylated by CheA.[225,234]

The *P. aeruginosa* Chp chemosensory system controls twitching motility and tfp biogenesis at two levels: through control of pili assembly and/or retraction as well as expression of the pilin subunit gene *pilA*.[231] The Chp system consists of a putative methyl-accepting chemoreceptor (PilJ), a methyltransferase CheR-like protein (PilK), 2 CheW homologs (PilI and ChpC) and a methylesterase CheB homolog (ChpB). In addition to the response regulator domain of ChpB, the Chp system has a further 3 CheY response regulators modules (PilG, PilH and a domain located at the C-terminus of the CheA-like

histidine kinase of this pathway, ChpA). ChpA is quite possibly the most complex signaling protein yet described in nature with nine putative sites of phosphorylation: 6 histidine-containing phosphotransfer (HPt) domains, 2 novel serine- and threonine-containing phosphotransfer domains (SPt, TPt) and the CheY-like receiver domain at its C-terminus. It is likely that ChpA is responsible for phosphotransfer reactions with PilG, PilH, and ChpB as well as an intramolecular phosphotransfer to its CheY domain and may well feed into other as yet unidentified response regulator modules. By analogy with the enteric swimming chemotaxis system in which CheY interacts with the flagella motor to regulate swimming motility, either or both of the CheY homologs PilG and PilH may interact with the putative tfp motor to control tfp extension/retraction. The Chp system, like many other bacterial chemotaxis systems, lacks a homolog of the enteric CheZ which serves to accelerate dephosphorylation of phosphorylated CheY. Other bacterial chemotaxis systems such as that of *Sinorhizobium meliloti* utilize a second CheY protein to act as a phosphate sink to add a similar layer of regulation to the system.[208] It is possible that either of PilG, PilH and /or the C-terminal CheY domain of ChpA serve a similar function in this system. Interestingly the *P. aeruginosa* genome encodes at least 26 putative MCPs[52,212] and the Chp system possesses 2 CheW homologs (PilI and ChpC). This raises the possibility that PilI and ChpC may serve to complex a number of MCPs to the histidine kinase ChpA for the input of multiple environmental signals which are integrated by the Chp system to control (at least) tfp biogenesis and twitching motility.

Another *P. aeruginosa* transcriptional regulator that influences tfp biogenesis and function in *P. aeruginosa* is the cyclic AMP receptor protein Vfr.[11] Mutants of *vfr* remain sensitive to tfp-specific bacteriophage, demonstrate reduced levels of surface-assembled tfp and show severely compromised twitching motility.[11] Recent whole genome transcriptional profiling has implicated Vfr in the regulation of over 200 genes of *P. aeruginosa*.[236] Included in this list of Vfr responsive genes are the majority of the genes required for the biogenesis, function and regulation of tfp (*pilDEGHIJKNOPQSUVWXY2Z*, *fimUVSX*, *chpABC*) which places Vfr as a global regulator of tfp biogenesis and function in *P. aeruginosa.*

It appears that *P. aeruginosa* possesses another layer of gene regulation situated above Vfr. Mutants of the recently identified gene *fimL* are phenotypically very similar to isogenic *vfr* mutants with respect to twitching motility, phage sensitivity, tfp assembly, and presentation of an autolytic phenotype.[229] Expression of cloned *vfr* is able to complement these phenotypes to *fimL* mutants.[229] Interestingly, *P. aeruginosa fimL* mutants show reduced *vfr* transcription and expression when assayed from plate culture or in the presence of epithelial cells *in vitro*, but not from agitated broth culture[229] which suggests that FimL's influence over Vfr expression may be specific to conditions in which functional tfp are required. *P. aeruginosa fimL* mutants also frequently revert to

wildtype phenotypes suggesting that an extragenic suppressor mutation is able to overcome the loss of *fimL*. These revertants show elevated levels of *vfr* transcription and expression which suggests that the site of spontaneous mutation is in a gene which lies upstream of *vfr*.[229] The mechanism *via* which FimL influences *vfr* transcription has not yet been elucidated. FimL is homologous to the N-terminal 570 amino acids of ChpA, except that the putative histidine and threonine phosphotransfer sites have been replaced with glutamines and thus would be incapable of phosphate acceptance. The significance of this homology with the N-proximal domain of ChpA is not known but may suggest some interaction with the Chp chemosensory system.

Recently, there has been some controversy surrounding the involvement of the *P. aeruginosa* quorum sensing systems in twitching motility. It was initially reported that twitching motility in *P. aeruginosa* is dependent on the *las* and *rhl* quorum sensing systems[83] whereas other investigators found that *lasI*, *lasR*, *rhlI*, and *rhlR* null mutants all exhibit normal twitching motility.[12] The discrepancy between these studies is probably accounted for by the observation that *P. aeruginosa* quorum-sensing mutants are susceptible to the acquisition of second-site mutations in genes which encode regulators that are required for twitching motility such *algR* in the case of *rhlI* mutants and *vfr* in the case of *lasI* mutants.[12] The non-twitching *rhlI* mutants were found to have acquired a GGGGT→GGGGG transversion in *algR* causing a V241G substitution in the DNA-binding domain of this regulator.[12] The *lasI* mutants that demonstrated reduced twitching zones were found to have acquired a 15 nucleotide deletion in the second half of a imperfect tandem repeat situated within the cAMP-binding pocket of Vfr.[12] Interestingly, expression of this allele (VfrΔEQERS) is able to complement elastase production but not twitching motility to *vfr* null mutants, suggesting that Vfr may be able to differentially regulate discrete regulons.[11] Thus it seems that in the absence of the *las* quorum sensing system, *P. aeruginosa* strains acquire a mutation in Vfr that is specific for a subset of Vfr target genes that includes those required for twitching motility. These observations indicate that mutations in one regulatory system (such as *lasI*, *rhlI*, or *fimL*), may create distortions that select during subsequent culturing for compensatory mutations in other regulatory genes within the network. This highlights the need for vigilance to the possibility of secondary mutation when working with these regulatory systems and emphasizes the importance of complementation studies when attributing phenotype to genetic mutation.

Twitching motility in *P. aeruginosa* is also influenced by the multidomain protein FimX that possesses several modules that are often found in bacterial signal transduction proteins.[106] Mutants of *fimX* exhibit low levels of surface-assembled tfp and reduced twitching motility. FimX is polarly localized and this appears to be dependent upon Vfr but not other tfp-related genes which suggests that Vfr or a target of Vfr is responsible for targeting FimX to the cell pole.[106] FimX is necessary for the stimulation of twitching motility in

response to tryptone and mucin (see Section 5.3) though it is not known how FimX controls twitching motility in response to these environmental cues.[106] The N-terminus of FimX shows some similarity to response regulator modules though it lacks several conserved features including the critical phosphate accepting aspartate residue normally found in these domains. FimX also possesses a PAS–PAC domain which is present in a range of proteins involved in light, oxygen and redox sensing[165,248] as well as putative GGDEF and EAL domains which are thought to be involved in cyclic di-GMP (c-di-GMP) production and degradation respectively.[56,111,202] In recent years a number of studies have shown that bacterial proteins containing GGDEF and/or EAL domains are involved in the modification of cell surface adhesiveness and motility and that c-di-GMP is likely to be a global second messenger involved in the control of these processes.[56,111,202] Interestingly, Simm *et al.* observed that overexpression of a GGDEF domain protein in *Salmonella enterica* serovar Typhimurium increased c-di-GMP levels in plate-grown bacteria and significantly repressed twitching motility when the cloned gene was introduced into *P. aeruginosa*.[202] Conversely, overexpression of an EAL domain protein was found to reduce c-di-GMP levels in *S. enterica* serovar Typhimurium and enhanced twitching motility in *P. aeruginosa*.[202] Whilst c-di-GMP synthesis and degradation activities have not yet been demonstrated for FimX, it seems reasonable to expect that FimX modulates tfp biogenesis and twitching motility in *P. aeruginosa via* c-di-GMP metabolism.

The biogenesis and function of *P. aeruginosa* tfp also appears to be linked to the metabolic status of the cell *via* a number of global regulatory systems. For instance, both the alternate sigma factor RpoS and polyphosphate kinase (Ppk) are required for survival in the stationary phase of growth and in adapting to nutritional stringencies and environmental stresses[124,219] and mutants of either *rpoS* or *ppk* show reduced twitching motility.[176,214] As noted above, the alternate sigma factor RpoN is required for expression of the pilin subunit gene *pilA* and RpoN is also required for the expression of genes required for nitrogen assimilation in *P. aeruginosa*.[218] The *P. aeruginosa* catabolite repression control protein Crc which is involved in the regulation of carbon metabolism has also been found to influence expression of *pilA* and *pilB*, and *crc* mutants show reduced twitching motility and numbers of surface assembled tfp.[153] Thus it appears that control of tfp biogenesis and function in *P. aeruginosa* is a complex process with many layers of regulation, the details of which are only beginning to be deciphered.

2.6. Tfp in Other *Pseudomonas* Species

Molecular studies of the tfp of *P. stutzeri* and *P. syringae* have demonstrated that many of the structural and biogenesis components of the tfp of these

species are functionally interchangeable with the corresponding components of *P. aeruginosa* tfp.[63,84,87,180,181] Indeed BLAST analyses of the publicly available completed genome sequences of *P syringae* pv. *tomato*[32] and the unfinished genome sequences of *P. fluorescens* PfO-1 (Genbank Accession NC007492) show that both *P. syringae* and *P. fluorescens* have tfp biogenesis and regulatory systems that are very closely related to that of *P. aeruginosa* (Table 1). It seems reasonable to expect, therefore, that much of what is known about the biogenesis and function of tfp of *P. aeruginosa* will be applicable to other type IV piliated *Pseudomonas* species.

Interestingly, similar *in silico* analyses of the *P. putida* KT2440 genome[149] on the other hand suggests that this species may either have a more distantly related tfp system or is no longer capable of producing functional tfp as many of the structural and biogenesis components found in the other *Pseudomonas* species are missing or truncated in the *P. putida* KT2440 genome sequences (Table 1). This is consistent with observations made by deGroot *et al*.[63] who observed that the *pilB* gene was missing from the *pilA–pilD* gene cluster of *P. putida* WCS358. These investigators were also unable to observe surface assembled tfp on *P. putida* WCS358.[63] However, they did find that heterologous expression of the *P. putida pilA* gene in *P. aeruginosa* resulted in the production of surface assembled tfp containing *P. putida* PilA subunits[63] which indicates conservation of the tfp machinery of *P. putida* and *P. aeruginosa.* Given the high degree of conservation between the tfp machineries of other *Pseudomonas* species it seems most likely that the *P. putida* strains KT2440 and WCS358 have acquired mutations in key tfp genes and as a consequence are no longer able to produce tfp.

3. TYPE IV PILI AND NATURAL TRANSFORMATION IN *PSEUDOMONAS* SPECIES

Natural transformation involves the acquisition of naked DNA from the environment and the subsequent heritable incorporation of its genetic information. Many bacterial species from diverse taxonomic groups including archaebacteria have the capacity to transport DNA from the extracellular milieu into the cytoplasm.[135] DNA uptake is a complex process which (in most cases) involves the assembly of a competence pseudopilus that is related to tfp and type II secretion systems.[39] In naturally transformable species that also possess tfp, there is an apparent correlation between the presence of tfp and DNA uptake, though it is not clear if in these bacteria the tfp *per se* have a direct role in the transport of DNA across the bacterial cell wall or if a subset of tfp components contribute to the assembly of a competence pseudopilus that is specifically responsible for DNA uptake.[39]

The molecular genetics of natural transformation in *Pseudomonas* species strains has as yet only been studied in *P. stutzeri*.[84–87] Consistent with observations made with other naturally transformable tfp producing bacteria such as *N. gonorrhoeae*[39] and *Synechocystis* species PCC6803,[155,243] natural transformation in *P. stutzeri* is dependent upon a number of tfp components including PilC and PilT, and the tfp subunit PilA1.[84,87] *P. stutzeri* also possesses other tfp-related components that appear to have specific roles in natural transformation. For instance, co-transcribed with *pilA1* is the gene *pilAII* which encodes a closely related PilA1 homolog. PilAII is not required for tfp biogenesis or twitching motility but appears to be a negative regulator of natural transformation.[86] In a similar arrangement to that found in *P. aeruginosa*, the *P. stutzeri* gene *pilU* is located immediately downstream of *pilT*. However, unlike *P. aeruginosa pilU* mutants which are hyperpiliated and non-twitching[232], *P. stutzeri pilU* mutants are unaffected in piliation and twitching motility but have transformation efficiency reduced to about 10% of wildtype levels.[87] Interestingly, heterologous expression of *P. aeruginosa pilU* genes restores natural transformation to *P. stutzeri pilU* mutants[87] which is consistent with the high degree of similarity observed between the PilU proteins of these 2 species.

In addition to tfp-related genes, natural transformation in *P. stutzeri* also requires the non-tfp-related competence genes *comA* and *exbB* which are thought to be involved in late stages of translocation of DNA across the inner membrane.[85] BLAST search analysis of the completed genome sequences of *P. aeruginosa*, *P. putida*, *P syringae*, and *P. fluorescens* indicates that these species each possesses homologs of *P. stutzeri* ComA (69–76% similarity) and ExbB (89–92% similarity) in a similar genetic arrangement which suggests that these species may also have the genetic capability for natural transformation.

In the early 1980s, a survey of natural transformation in *Pseudomonas* species demonstrated this capacity in some strains of *P. stutzeri*, *P. pseudoalcaligenes*, *P. alcaligenes*, and *P. mendocina* but not in *P. aeruginosa*, *P. fluorescens*, *P. syringae*, and *P. putida*.[35] However, it is feasible that *P. aeruginosa* and other *Pseudomonas* species are also capable of natural transformation under specific competence conditions. Indeed, natural transformation has been recently demonstrated in *P. fluorescens* when cultured in soil microcosms but not in a range of *in vitro* conditions tested.[64] These observations suggest that specific conditions found in soil culture are required to induce competence for natural transformation in *P. fluorescens*. *P. aeruginosa* has also been found to be capable of uptake and maintenance of plasmid DNA *via* natural transformation in a tfp dependent manner, albeit at very-low frequency under the *in vitro* conditions used (C.B. Whitchurch, unpublished observations). Thus the capacity for natural transformation may in fact be widespread in *Pseudomonas* species, but the specific conditions for inducing the competence state have yet to be identified.

4. ROLES OF TYPE IV PILI IN COLONIZATION AND INFECTION BY *PSEUDOMONAS* SPECIES

Many members of the *Pseudomonas* genus are important phytopathogens (*P. syringae*), phytostimulators (*P. fluorescens*), or opportunistic pathogens of animals, plants, insects, and fungi (*P. aeruginosa*). Type IV pili-related functions that contribute to the colonization and infection process include mediating initial adhesion to host tissues, colony expansion *via* twitching motility, and the development of intractable biofilms that are recalcitrant to host immune attack and antibiotic therapy.

Type IV pili have been shown to mediate attachment of various *P. syringae* pathovars to the leaves of both susceptible and non-susceptible plants. *P. syringae* tfp mutants show significantly reduced adhesion to plant tissues and the ability to generate symptoms in host plants when spray inoculated also correlates with the presence of tfp.[181,184–185] Taken together these results indicate that tfp-mediated adhesion of *P. syringae* phytopathogens to plant tissues is an important factor in the disease process. Tfp may also play a role in colonization of tomato roots by the plant growth promoting *P. fluorescens* WCS365.[138] The identity of specific plant receptors, if any, for tfp-mediated adhesion to plant tissues has not been established.

P. aeruginosa is an important opportunistic pathogen of humans but is also a pathogen of plants, animals, and yeast. Various plant (*Arabidopsis*, lettuce, basil, and alfalfa), animal (mouse, rat, hamster, nematode, fruit fly, and moth) and yeast (*Candida albicans*) models of infection have been developed to identify and characterize virulence genes of *P. aeruginosa*. To date, the involvement of *P. aeruginosa* tfp in infection of plants or nematodes has not been investigated. As mentioned above, it has been shown that *P. aeruginosa* tfp are not required for virulence of *Drosophila melanogaster* (fruit fly).[55] It has also been demonstrated that *rpoN* mutants of *P. aeruginosa* (which do not produce tfp), are not impaired in virulence of *Galleria mellonella* (greater wax moth).[96] Thus it seems that tfp are not essential for virulence in these insect pathogenicity models. *P. aeruginosa* tfp are, however, involved in initial attachment to filamentous *C. albicans*[103] and are essential for virulence in various mouse models of infection.[40,47,75,88,217,249]

4.1. *P. aeruginosa* Tfp and Adhesion to Mammalian Tissues

It is generally accepted that a crucial component of successful colonization and infection of host tissues by bacteria is irreversible adherence to host cell surfaces.[89,182] Adhesion to human mucosal tissues is an essential step in *P. aeruginosa* pathogenesis.[114–116,239] Over the past 20 years a wealth of studies utilizing tfp mutants, competitive inhibition with purified tfp, and/or polyclonal

and monoclonal antisera have confirmed that tfp contribute significantly to the adherence of *P. aeruginosa* to mammalian cells. *P. aeruginosa* tfp have been shown to mediate adhesion to primary human buccal and tracheal cells,[68,69,240] injured dog and mouse tracheal epithelia,[174,250] injured mouse and rabbit corneal epithelia,[95,249] primary bovine tracheal epithelial monolayers,[189] and primary mouse epidermal cells.[194] *P. aeruginosa* tfp have also been established as the major adhesins that are responsible for about 90% of the attachment of the pathogen to various cultured epithelial cell lines.[40,47,75,90]

Injury or disease of epithelial tissues appears to be a prerequisite for *P. aeruginosa* adherence *in vivo*[62,172,174,175,250] presumably due to the exposure and/or upregulation of tfp receptors on the epithelial cell surface. Much attention has focused on the asialylated glycosphingolids (asialo-GSL) asiolo-GM_1 and asiolo-GM_2 as the putative epithelial tfp-receptors.[6,31,48,61,88,89,125,126,132,133,201,245] The minimum carbohydrate moiety present on asialo-GSLs that is thought to be recognized by *P. aeruginosa* tfp is GalNAcβ(1-4)Gal.[126,201] However, the role of asialo-GSLs as the major epithelial cell receptor for *P. aeruginosa* remains controversial.[60,90,107,197,204,247] To date, a candidate epithelial cell surface component that is specifically recognized by *P. aeruginosa* tfp and which contributes to the majority of tfp-mediated adhesion of *P. aeruginosa* cells has not as yet been isolated. It is also possible that *P. aeruginosa* tfp binding to epithelial cells may be either multivalent or non-specific. Various other molecules which have been identified as potential epithelial cell receptors for *P. aeruginosa* tfp include sialylated glycosphingolipids, sialic acid, *N*-acetylmannoseamine, glycoproteins, and other non-glycosylated polypeptides.[6,67,89,94,108,173]

How do *P. aeruginosa* tfp mediate attachment to epithelial cell receptors? Unlike some other fimbrial types *P. aeruginosa* tfp do not contain a specific tip associated adhesin that mediates attachment to epithelial cells. Instead, it appears that the C-terminal DSL of the pilin protein (PilA) contains the receptor-binding domain. This has been established through a series of studies which demonstrated (a) that monoclonal antibodies specific for the DSL, or proteolytic and synthetic peptides corresponding to the DSL were able to inhibit *P. aeruginosa* attachment to epithelial cells and asialo-GM_1; (b) that biotinylated DSL peptides bound to asialo-GM_1; and (c) that mutagenesis of the DSL domain of the tfp pilin subunit did not affect assembly of the filament but dramatically inhibited binding to epithelial cells to the levels of isogenic mutants that lacked tfp.[68,69,75,108,130–133,201,237,238,245,246] Immuno-electron microscopy has demonstrated that the PilA DSL is exposed only at the tip of the tfp-filament.[132] The DSL domain of *P. aeruginosa* pilin has substantial sequence variability amongst different strains yet they are able to compete for and bind to the same receptors on epithelial cells and cross-reactive antibodies specific for the DSL have been identified.[68,174,239] The solved crystal structures of the pilin subunits of *P. aeruginosa* strains PAK and K122-4 predict that the tfp pilin receptor-binding

site is formed by structurally constrained main-chain atoms which accounts for the apparent conservation in receptor specificity between divergent DSL sequences.[5,93]

Taken together the above observations indicate that up to 90% of the binding of *P. aeruginosa* to damaged epithelia is mediated *via* type IV pili through the structurally conserved DSL domain exposed at the tip of the filament and that this process is essential to the establishment of infection. However, *P. aeruginosa* binding to epithelial surfaces may not be relevant in the airways of CF patients. Instead, *P. aeruginosa* appears to preferentially bind to mucus present in the CF airways rather than epithelial cell surfaces and indeed mucin binding is believed to be important in the initial colonization of the airways of CF patients by *P. aeruginosa*.[170,241] Furthermore, established *P. aeruginosa* macrocolonies are found within the airway luminal mucus of CF patients at some distance from the airway epithelial surface.[241] These observations suggest that adhesion of *P. aeruginosa* to mucin rather than epithelia is important in the colonization of CF airways. The tendency for *P. aeruginosa* to bind to mucin may assist with the clearance of this opportunistic pathogen from healthy lungs *via* the mucus elevator, but this potential may also be a contributing factor to infection of CF airways due to the presence of thickened dehydrated mucus that is not cleared effectively and which allows *P. aeruginosa* to successfully colonize the mucus layer of CF airways.[241] Whilst it seems that *P. aeruginosa* does not access the CF epithelia, and tfp are not essential for mucin binding,[171] *P. aeruginosa* tfp may still play an important role in the colonization of CF airways as twitching motility is extremely enhanced in the presence of airway secretions including mucin (see Section 5.3).

4.2. Other Roles of *P. aeruginosa* Tfp in Mammalian Infections

Type IV pili are also important in host immune system defences against *P. aeruginosa*. *P. aeruginosa* tfp have been shown to induce T-cell responses in mice,[207] humoral responses toward tfp are protective[33,187,200] and tfp-mediated attachment to human and mouse macrophages and polymorphonuclear leukocytes is thought to contribute to non-opsonic phagocytosis of *P. aeruginosa* through interaction with macrophages, though tfp *per se* are not sufficient to trigger phagocytosis.[119,139,209]

Tfp have also been implicated in the efficient secretion of virulence factors *via* the related type II secretion pathway. This effect is believed to be mediated *via* direct interaction of the pilin subunit with components of the secretion apparatus.[137] *P. aeruginosa* tfp are also required for type III secretion system-dependent cytotoxicity and apoptosis, presumably due to a requirement for tfp to

mediate attachment of the bacterium to host epithelial cells before translocation of effector proteins can occur.[47,48,112,117,216]

As we have seen, the tfp of *P. aeruginosa* participate in numerous aspects of *P. aeruginosa* pathogenesis. Interestingly, *P. aeruginosa pilT* or *pilU* mutants, which are hyperpiliated and at least in the case of *pilT* mutants produce non-retractile tfp, show significantly reduced adhesion to and cytotoxicity of epithelial cell monolayers *in vitro*, are deficient in the ability to promote corneal disease, and are not fully virulent in a mouse model of acute pneumonia.[47,95,249] These observations suggest that many of the virulence-associated roles of tfp require fully functional tfp. It is possible that tfp retraction might be required to enable intimate adherence to the epithelial cells, possibly by engaging secondary receptors at the bacterial cell surface. These observations also suggest that twitching motility is required for successful colonization and spread of infection.

5. TWITCHING MOTILITY

Twitching motility is a form of flagella independent surface motility and as such is distinct from swimming or swarming motilities, which are both powered by flagella. As noted above, the locomotive force facilitating cellular translocation *via* twitching motility is a function of tfp retraction. Twitching motility was first described by Lautrop in 1961 when describing this motility in *Acinetobacter calcoaceticus*.[128] Lautrop coined the term "twitching" in 1965 to describe this motility.[129] The term is derived from the observation that individual cells appear to move in a jerky or "twitchy" fashion when viewed in suspension. Twitching motility has since been shown to occur in a wide range of Gram-negative bacteria including the *Pseudomonas* species *P. aeruginosa* and *P. stutzeri*[98,99,143] as well as in the Gram-positive bacterium *S. sanguis*.[100] Interestingly, twitching motility has not as yet been reported in either *P. syringae* or *P. fluorescens* despite the fact that both appear genetically capable of this motility (see Section 2.6; Table 1).

5.1. Twitching Motility in *P. aeruginosa*

When cultured in humid conditions on air-dried agar, *P. aeruginosa* colonies grow as flat, rough spreading colonies with a characteristic "ground glass" edge. In contrast, mutants which are no longer capable of twitching motility produce smooth, domed colonies.[199] This obvious distinction in colony morphology has been exploited in various genetic screens to identify twitching motility mutants for molecular genetic analysis of this phenomenon and tfp biogenesis in *P. aeruginosa*.

Over recent years it has become evident that *P. aeruginosa* exhibits extremely active twitching motility at the interstitial surface between agar and the plastic petri dish.[199] When a small inoculum of wildtype *P. aeruginosa* is "stabbed" through to the plastic at the bottom of a petri dish, a large halo of colony expansion is produced at the agar/plastic interface after overnight incubation, whereas non-twitching mutants produce no such zone. To aid visualization the zone of colony expansion is often post-stained with Coomassie Brilliant Blue or a vital dye such as tetrazolium red is incorporated into the growth medium.[199] This "stab assay" is now frequently used as a *de facto* measurement of twitching motility activity in the presence of different environmental stimuli and in *P. aeruginosa* mutants that display only subtle defects in twitching motility.

The development of rapid twitching motility at the interstitial surface between the agar and the plastic may be attributed, at least in part, to the smoothness of the surface of the agar set against the plastic. Indeed, Semmler *et al*.[199] found that colony expansion due to twitching motility can also occur very rapidly on air-exposed agar if the agar is inverted such that the surface that was set against the plastic is now exposed to the air. When this surface is inoculated and incubated overnight, the resultant zone of colony expansion approaches the size of that obtained *via* the stab assay technique.[199] This implies that the irregularities on the surface of agar that is exposed to air whilst solidifying are inhibitory to twitching motility.

There does, however, appear to be some contribution of the abiotic surface to twitching motility at the agar/plastic interface. Twitching motility is greatly enhanced on the air-exposed surface of nutrient media set with either agar or gellan gum when the inoculum is covered with glass coverslips, a phenomenon that has been exploited for microscopic examination of twitching motility in recent years.[199] We now know that translocation of individual *P. aeruginosa* cells during twitching motility on quartz is mediated by retraction of tfp after adhesion of the distal tfp tip to the quartz surface.[206] Furthermore, colony expansion at the agar–plastic interface is also greatly enhanced if plastic petri dishes that have been specially treated for use in mammalian tissue culture are used[229] presumably because tfp interactions with the treated plastic surface are enhanced.

5.2. Microscopic Examination of Colony Expansion Due to Twitching Motility

Early microscopy studies of twitching motility examined cellular movements at the outer edges of colonies grown on air-dried agar surfaces in humid conditions. Under these conditions twitching motility was observed to be a relatively slow, disorganized mode of translocation in which cells moved predominantly singly in an intermittent and jerky fashion.[97] Microscopic examination

of twitching-mediated colony expansion has been greatly facilitated in recent years by exploiting the observation that rapid colony expansion occurs at the interstitial surface between a glass coverslip and a thin layer of nutrient media solidified with gellan gum which provides greater optical clarity than agar.[199] The most striking feature of the twitching-dependent colony expansion that occurs in this system is the intricate lattice network of cells that forms behind the outgoing leading edge rafts.[199] It is clear under these conditions that twitching motility manifests as a rapid, highly organized multicellular mechanism of colony expansion presumably to enable rapid colonization of large areas.

Time-lapse video microscopy has been used to examine cellular movements during colony expansion at the gellan gum/glass interface.[199] Using this technique, Semmler *et al*.[199] observed that at the outermost edge of the expanding colony, large aggregates or rafts of cells assemble and move radially outward from the colony into unoccupied territory. The cells within these leading edge rafts are densely packed and aligned along the longitudinal axis of the cell. Occasionally individual cells are seen to reverse direction and head back toward the colony remaining confined to the area previously traversed by the outgoing raft. It is possible that the pioneering rafts lay down "trails" of an unknown substance that serves to lubricate or in some other way facilitate individual cellular movements in the ensuing network. Cells within these trails move rapidly either as individuals or as small groups, often reversing direction and sliding past each other but always following the long axis of the cell. Occasionally aggregates of cells move off in new directions to make connections with other trails of cells, a process that ultimately leads to the formation of a fine network of cells behind the leading edge rafts. *P. aeruginosa* cells seem to prefer to be aligned along the long axis of the cell and are often seen to rapidly snap into alignment after contacting a new cell or group of cells.[199]

It is now clear that individual *P. aeruginosa* cells are capable of translocation across glass surface *via* twitching motility.[205,206] However, during colony expansion on nutrient media it appears that virgin areas are only ever traversed by aggregates of cells, never by individual cells.[199] The significance of this observation is unclear but does suggest that the combined effort of a number of cells is required to traverse new areas. The mechanism that enables large aggregates of cells to move *en masse* is unknown but presumably involves tfp retraction. Furthermore, Semmler *et al*. also reported that whilst cells within the network behind the leading edge rafts are capable of moving individually, this only occurs when they are within the vicinity of other cells and that when cells become isolated they are unable to move until other cells approach to within a couple of μm[199] – a distance consistent with the length of a tfp filament. It was recently shown that individual *N. gonorrhoeae* cells are pulled into a microcolony of cells *via* tfp retraction.[145] It is conceivable therefore, that retraction of the tfp after adhesion to a neighboring cell (rather than the substrate) is the motive force behind cellular translocation of individual *P. aeruginosa* cells in the

network trails. This "chain gang" arrangement of cells might also account for the apparent confinement of cells to "trails" in the expanding colony network.

In the past few years, significant advances have been made in our understanding of how tfp facilitate translocation of individual cells across abiotic surfaces. However, we do not yet fully understand how this phenomenon is translated into the variety of multicellular behaviors that occur during twitching motility dependent colony expansion in *P. aeruginosa*. It is evident however, that twitching motility, as it occurs in the context of colony expansion, is a highly complex social phenomenon that is dependent on tfp and requires cell–cell contact.

5.3. Environmental Signals That Influence Twitching Motility

In recent years, studies that have identified environmental cues that influence twitching motility have generally utilized the subsurface stab assay as a quasi-quantitative measure of the rate of expeditionary twitching motility in *P. aeruginosa*. It should be appreciated however, that this assay is in reality a measure of the rate of colony expansion, which is a complex process with many contributing factors, including the rate of cellular movements *via* twitching motility, but is probably also influenced by the rate of cell division. It is also not clear if alterations in the rate of expeditionary twitching motility as assessed by colony expansion at the agar/petri dish interface is a function of signal transduction-mediated events such as chemotaxis (motility up or down a chemical gradient) toward unoccupied regions of the agar, or chemokinesis (increased rate of motility stimulated by a chemical), or if the effect is more passive and a consequence of physical changes to the substrate, or increased growth rate.

It has long been established that conditions of high humidity are conducive to rapid colony expansion on the agar surface.[99] The concentration of agar used in the subsurface stab assay also has as a significant influence on the resultant size of the colony expansion zone with optimal motility occurring at an agar concentration of 1% (w/vol) (C.B. Whitchurch, A.B.T Semmler, and J.S. Mattick, unpublished observations). Presumably at this concentration the relative humidity at the interstitial surface is optimal for twitching motility. We have also found that the brand of agar used can have a dramatic influence on the size of the colony expansion zone (C.B. Whitchurch and J.S Mattick, unpublished observations).

We have recently identified a number of environmental cues that stimulate or inhibit colony expansion *via* twitching motility in *P. aeruginosa* (C.B. Whitchurch, A.B.T Semmler, and J.S Mattick, unpublished observations).[106] In this survey, different chemicals were added to a common base media [0.05% (w/vol) yeast extract, 100 mM potassium phosphate buffer pH 7.0, 1% w/vol (agar)] and colony expansion at the interstitial surface between the agar and the standard plastic petri dish was assessed by the subsurface stab assay. Compounds

that were found to repress twitching motility were generally those that increase the osmolarity of the medium such as NaCl (300 mM), KCl (300 mM), sucrose (5% w/vol), glucose (10% w/vol), glycerol (5% vol/vol), and polyvynilpyrrolidone (2% w/vol). On the other hand, we found that increasing the tryptone and yeast extract components of Luria Bertani medium either together or individually to 3–5 times standard concentrations was stimulatory to colony expansion (C.B. Whitchurch, A.B.T Semmler, and J.S Mattick, unpublished observations). A similar result was obtained when tryptone (3–5% w/vol) was added to the base media. Interestingly, addition of equivalent amounts of Casamino acids to the base media had the opposite effect on colony expansion and was inhibitory (C. B. Whitchurch and J. S. Mattick, unpublished observations). Both tryptone and Casamino acids are derived from casein, the former being a tryptic digest of the protein and thus comprised of peptides, and the latter an acid digest and is comprised of amino acids. Thus it seems that peptides and not amino acids are stimulatory to twitching motility in *P. aeruginosa*. Interestingly, certain oligopeptides have been shown to be chemoattractive for swimming motility in *P. aeruginosa*[120] and thus it seems possible that peptides may also stimulate twitching motility *via* related chemosensory phenomena such as chemotaxis and/or chemokinesis.

We also found that bovine serum albumin (BSA) is extremely stimulatory to twitching motility at very low concentrations in this system.[106] It is not known if the effect of BSA on twitching motility is *via* transduction of a specific signal or through passive facilitation of twitching motility. Tfp have been shown to bind to several large polypeptides[89,158] thus twitching motility in the presence of BSA may be enhanced through increased binding of tfp to the protein in the agar substrate. It is also possible that secreted proteases of *P. aeruginosa* may play a role in digesting these proteins into stimulatory peptides that are then sensed by specific signal transducing proteins to stimulate twitching motility.

The glycoprotein mucin was also found to be extremely stimulatory to twitching motility at low concentration.[106] Mucins are high-molecular-weight glycoproteins with hundreds of branched oligosaccharide side chains linked to a protein core.[186] *P. aeruginosa* has been shown to display swimming chemotaxis toward sources of mucin and mucin-associated sugars and amino acids[148] thus it is possible that a similar chemosensory mechanism may be involved in controlling twitching motility in response to mucin. However, extensive dialysis of mucin prior to the addition to the agar base media does not reduce the stimulatory effect of mucin[106] which suggests that the glycoprotein *per se* is enhancing colony expansion either by lubricating cellular movements and/or through a specific signaling cascade for mucin or mucin derivatives that may be elicited through the action of secreted enzymes. The observation that the putative *P. aeruginosa* signal transduction protein FimX is required for stimulation of twitching motility by mucin and tryptone but not BSA[106] suggests that the response to mucin occurs *via* a transduced mucin-specific signal.

Another common protein component of airway secretions, lactoferrin has also been found to be stimulatory to twitching motility in *P. aeruginosa*.[205] This influence on twitching motility appears to be mediated by the ability of lactoferrin to sequester iron.[205] Addition of the iron chelator deforaxamine to agar plates was also found to stimulate twitching motility as assessed by the subsurface stab assay, and this effect could be reversed with the addition of free-iron.[205] Interestingly, mucin is an iron-binding glycoprotein,[49,167] thus it is possible that at least some of the stimulation of twitching motility observed in the presence of mucin is due to iron sequestration. Given the role of FimX in stimulating twitching motility in response to mucin,[106] it would be interesting to determine if FimX is involved in controlling twitching motility in response to iron availability.

It has recently been reported that twitching motility in *P. aeruginosa* is responsive to phosphatidylethanolamine (PE) isolated from late log phase *P. aeruginosa* culture.[118] These investigators demonstrated that *P. aeruginosa* colonies expand *via* directed twitching motility up gradients on non-nutrient agar of *P. aeruginosa*-derived PE or certain synthesized PE species. They also demonstrated that addition of certain synthetic PE species uniformly to the agar plate enhanced overall colony expansion.[118] Since these responses showed a degree of specificity for the fatty acid moiety of the PE these investigators favored the notion that at least part of the response to PE by *P. aeruginosa* involves chemokinesis elicited *via* signal transduction rather than merely passive facilitation of cellular movements by the surfactant properties of PE.

Recently Barker *et al.* have shown that the extracellular phospholipase C PlcB is required for directed twitching motility up gradients of certain synthetic species of PE and phosphatidylcholine (PC).[8] This observation indicates that either or both of the PlcB hydrolytic products of PE or PC are required to elicit the directed twitching motility response. It is also possible that further degradation of these products might be required to produce the actual chemical signal that triggers the observed response to PE or PC. The fact that *P. aeruginosa* displays directed twitching motility in response to *P. aeruginosa*-derived phospholipids suggests the twitching motility and colony expansion might be controlled *via* chemotaxis and/or chemokinesis in response to these endogenous signals.

Whatever the mechanisms *via* which phospholipids, lactoferrin, mucin, and albumin stimulate twitching motility, it is possible that these responses may be important in *P. aeruginosa* pathogenesis. Human lung surfactant is comprised of 80–90% lipid by weight, more than 80% of which is phospholipid and of this 60–70% is PC and 5–10% PE.[66] The PE and PC content appears to be further elevated in surfactant isolated from the lungs of young adult CF patients.[146] Albumin, lacotoferrin, and mucin are also prominent components of respiratory secretions.[16] The fact that several components of airway secretions enhance expeditionary twitching motility by *P. aeruginosa* implies that twitching motility

may also be an important component of colonization of airway mucosa. On one hand, stimulation of twitching motility by these compounds might facilitate extensive colonization of airways and other mucosal tissues by this pathogen. On the other hand, however, as postulated by Singh *et al*.[205] stimulation of twitching motility may prevent biofilm development (see Section 5.3) thus making *P. aeruginosa* more susceptible to host defence mechanisms. Therefore, it seems that up-regulation of twitching motility in response to human airway secretions may be a double-edged sword for this opportunistic pathogen and suggests that successful colonization of mucosal tissues is a complex process.

6. TFP AND BIOFILM DEVELOPMENT

Bacterial biofilms are matrix-encased communities of sessile bacteria that are attached to a surface.[211] It is now clear that the development of *P. aeruginosa* biofilms is a complex social phenomenon that is dependent upon tfp and twitching motility. The first evidence for a role of tfp in biofilm development was obtained through mutant screens to identify *P. aeruginosa* mutants which were defective in the early stages of the development of biofilms grown in static media.[154] Phase contrast time-lapse microscopy showed that, in this system, wildtype *P. aeruginosa* first formed a monolayer of cells on the abiotic surface after which microcolonies appeared which were dispersed throughout the monolayer. Flagella were identified as being responsible for mediating the initial attachment of cells to the abiotic surface and tfp were found to be required for cells to move across the abiotic surface and to aggregate into microcolonies.[154]

Microscopic examination of the development of *P. aeruginosa* biofilms grown in laminar flow cells have shown that the structure of the mature biofilm is largely dependent upon the carbon source. For instance, when defined media containing citrate as the carbon source is used, *P. aeruginosa* forms flat dynamic biofilms whereas when glucose is the available carbon source the resultant biofilm is comprised of mushroom-shaped multicellular structures separated by water-filled channels.[121] Confocal laser scanning microscopy was used to characterize the development steps involved in the formation of these different biofilms and it was found that twitching motility was the most important factor prescribing the structure of the biofilm.[121] In citrate-grown biofilms, *P. aeruginosa* cells continue to exhibit extensive twitching motility across the abiotic surface beyond the initial phases of biofilm development and that this continued motility results in the development of flat undifferentiated biofilms that lacked microcolonies. In contrast, the larger heterogenous biofilms that are formed when glucose is provided as the carbon source are due to the activities of two subpopulations of cells that form the "stalks" or the "caps" of the mushroom-like structures. The "stalks" are formed by the proliferation of non-twitching cells whereas the mushroom "caps" are comprised of a subpopulation of cells

that use twitching motility to clamber up the sides of the stalk to aggregate on the top, presumably to gain access to the flow of nutrient media.[121,122]

Twitching motility also appears to play a role in the maintenance of biofilm morphology in mature biofilms. Chiang and Burrows[41] reported recently that the hyperpiliated *pilT* mutants of *P. aeruginosa* strain PAK formed denser cell mats in the early stages of biofilm development and larger, more tightly packed mushroom-like structures in the mature biofilm than was observed with isogenic wildtype *P. aeruginosa* in flow cell biofilms using minimal media in which glucose was the provided carbon source.[41] Under the experimental conditions used by Chiang and Burrows, an isogenic PAK *pilA* mutant was non-adherent to the glass surface and did not form biofilms.[41] These observations indicated that under these experimental conditions, tfp were the major adhesin mediating attachment to the glass surface, an obvious prerequisite for biofilm formation. In contrast, Klausen *et al*.[121] observed that in their experimental conditions and with their strain background (PAO1), even in glucose minimal media their isogenic *pilA* mutant was capable of some adhesion to the glass surface and formed large, irregularly shaped mushrooms.[121] Presumably, in both cases, the three-dimensional structures formed by the non-twitching *pilT* and *pilA* mutants are due to clonal growth.

Both colony expansion and biofilm formation by *P. aeruginosa* are multicellular social phenomena that involve tfp and twitching motility. An interesting correlation between the two has recently emerged. Singh *et al*.[205] have found that iron sequestering compounds such as lactoferrin and deferoxamine inhibit the formation of *P. aeruginosa* biofilms in flow cells. Interestingly, it was also found that iron chelation causes enhanced colony expansion at the interstitial surface between agar and plastic in the twitching stab assay described above.[205] Thus from these observations it seems that conditions that promote rapid colony expansion are inhibitory to biofilm development suggesting an inverse relationship between the two colonial behaviors. Time-lapse microscopy of cellular movements during the early stages of biofilm development in the presence of lactoferrin showed that individual cells did not settle down to form the microcolonies that are required to initiate biofilm development but rather continued to move across the flow cell surface by twitching motility,[205] similar to what was observed during biofilm development using citrate as the carbon source.[121] Presumably the developmental switch that produces a subpopulation of non-motile cells upon which the mushroom structures develop is not initiated in low iron conditions or in the presence of citrate. Thus it appears that both rapid colony expansion on nutrient media and inhibition of biofilm development are due to stimulation of twitching motility. It will be interesting to determine if other environmental signals that have been found to promote rapid colony expansion on solid nutrient media (see Section 5.2) also prevent biofilm development on abiotic surfaces in flow cells, and conversely if colony expansion is enhanced in the presence of citrate.

It is now evident that *P. aeruginosa* tfp and twitching motility are required for normal biofilm development. *P. aeruginosa* biofilms are complex communities of cells that are held together by an extracellular matrix comprised of exopolysaccharides, DNA and proteins.[211,229] Given the adhesive and aggregative properties of tfp it seems plausible that tfp may also provide a structural role in the biofilm either by binding bacterial cells to the components of the matrix or by intertwining and meshing cells together. Interestingly, small colony variants (SCVs) of *P. aeruginosa* isolated from biofilm grown cultures either *in vitro* or from the lungs of chronically infected CF patients are hyperpiliated, autoaggregative, highly adherent and demonstrate increased biofilm formation capabilities.[65,71,92] It has been suggested that the SCVs represent a specific biofilm phenotype that is either switched on *via* phase variation or is the outcome of adaptive mutation.[65,71,91] Small colony variants are often observed *in vitro* to revert to wildtype phenotypes either through phase variation switching or through further adaptive mutation.[65,71,91] The rate of SCV reversion is controlled by a putative 2-component response regulator PvrR which is thought to act upstream of the phenotypic switch.[71] The genetic mechanism *via* which the phenotypic switching is occurring has yet to be determined.

7. FUTURE DIRECTIONS

The study of type IV pili began over 50 years ago and continues to be an area of active investigation. It is now evident that type IV pili are highly complex multimeric structures that facilitate a variety of biological functions in bacteria including phage sensitivity, attachment to and colonization of biotic and abiotic surfaces, natural transformation, twitching motility, colony expansion, and biofilm development. Whilst over 40 genes have now been identified which are involved in the biogenesis and regulation of tfp in *P. aeruginosa*, we still have much to learn about how the structural proteins are assembled and interact with one another to form dynamic, powerful motors that mediate assembly and retraction of tfp to promote twitching motility. We also have much to learn about how twitching motility is co-ordinated into complex multicellular behaviors such as colony expansion and biofilm development and how these processes are regulated in response to environmental stimuli. Thus it appears that this fascinating area of biology will continue to be of interest to scientists for some time to come.

ACKNOWLEDGMENTS

During preparation of this manuscript CBW was supported by an NHMRC R. Douglas Wright Career Development Award.

REFERENCES

1. Alm, A.A., Bodero, A.J., Free, P.D., and Mattick, J.S., 1996, Identification of a novel gene, *pilZ*, essential for type 4 fimbrial biogenesis in *Pseudomonas aeruginosa. J. Bacteriol.*, 178:46–53.
2. Alm, R.A., Hallinan, J.P., Watson, A.A., and Mattick, J.S., 1996, Fimbrial biogenesis genes of *Pseudomonas aeruginosa -pilW* and *pilX* increase the similarity of type 4 fimbriae to the GSP protein-secretion systems and *pilY1* encodes a gonoccocal PilC homologue. *Mol. Microbiol.*, 22:161–173.
3. Alm, R.A., and Mattick, J.S., 1995, Identification of a gene, *pilV*, required for type 4 fimbrial biogenesis in *Pseudomonas aeruginosa* whose product possesses a prepilin-like leader sequence. *Mol. Microbiol.*, 16:485–496.
4. Alm, R.A., and Mattick, J.S., 1996, Identification of two genes with prepilin-like leader sequences required for type 4 fimbrial biogenesis in *Pseudomonas aeruginosa*. *J. Bacteriol.*, 178:3809–3817.
5. Audette, G.F., Irvin, R.T., and Hazes, B., 2004, Crystallographic analysis of the *Pseudomonas aeruginosa* strain K122-4 monomeric pilin reveals a conserved receptor-binding architecture. *Biochemistry*, 43:11427–11435.
6. Baker, N.R., Minor, V., Deal, C., Shahrabadi, M.S., Simpson, D.A., and Woods, D.E., 1991, *Pseudomonas aeruginosa* exoenzyme S is an adhesin. *Infect. Immun.*, 59:2859–2863.
7. Bally, M., Filloux, A., Akrim, M., Ball, G., Lazdunski, A., and Tommassen, J., 1992, Protein secretion in *Pseudomonas aeruginosa*: characterization of seven *xcp* genes and processing of secretory apparatus components by prepilin peptidase. *Mol. Microbiol.*, 6:1121–1131.
8. Barker, A.P., Vasil, A.I., Filloux, A., Ball, G., Wilderman, P.J., and Vasil, M.L., 2004, A novel extracellular phospholipase C of *Pseudomonas aeruginosa* is required for phospholipid chemotaxis. *Mol. Microbiol.*, 53:1089–1098.
9. Bateman, A., Birney, E., Cerruti, L., Durbin, R., Etwiller, L., Eddy, S.R., Griffiths-Jones, S., Howe, K.L., Marshall, M., and Sonnhammer, E.L., 2002, The Pfam protein families database. *Nucleic Acids Res.*, 30:276–280.
10. Baynham, P.J., and Wozniak, D.J., 1996, Identification and characterization of AlgZ, an AlgT-dependent DNA-binding protein required for *Pseudomonas aeruginosa algD* transcription. *Mol. Microbiol.*, 22:97–108.
11. Beatson, S.A., Whitchurch, C.B., Sargent, J.L., Levesque, R.C., and Mattick, J.S., 2002, Differential regulation of twitching motility and elastase production by Vfr in *Pseudomonas aeruginosa*. *J. Bacteriol.*, 184:3605–3613.
12. Beatson, S.A., Whitchurch, C.B., Semmler, A.B., and Mattick, J.S., 2002, Quorum sensing is not required for twitching motility in *Pseudomonas aeruginosa*. *J. Bacteriol.*, 184:3598–3604.
13. Bitter, W., 2003, Secretins of *Pseudomonas aeruginosa*: large holes in the outer membrane. *Arch. Microbiol.*, 179:307–314.
14. Bitter, W., Koster, M., Latijnhouwers, M., de Cock, H., and Tommassen, J., 1998, Formation of oligomeric rings by XcpQ and PilQ, which are involved in protein transport across the outer membrane of *Pseudomonas aeruginosa*. *Mol. Microbiol.*, 27:209–219.
15. Blatch, G.L., and Lassle, M., 1999, The tetratricopeptide repeat: a structural motif mediating protein–protein interactions. *Bioessays*, 21:932–939.
16. Boat, T.F., and Matthews, L.W., 1973, Chemical composition of human tracheo-bronchial secretions, pp. 243–274. In M.J. Dulfano (ed.), *Sputum. Fundamentals and Clinical Pathology*. Charles C. Thomas, Springfield, IL, USA.
17. Boyd, J.M., 2000, Localization of the histidine kinase PilS to the poles of *Pseudomonas aeruginosa* and identification of a localization domain. *Mol. Microbiol.*, 36:153–162.

18. Boyd, J.M., Koga, T., and Lory, S., 1994, Identification and characterization of PilS, an essential regulator of pilin expression in *Pseudomonas aeruginosa*. *Mol. Gen. Genet.*, 243:565–574.
19. Boyd, J.M., and Lory, S., 1996, Dual function of PilS during transcriptional activation of the *Pseudomonas aeruginosa* pilin subunit gene. *J. Bacteriol.*, 178:831–839.
20. Bradley, D.E., 1965, The morphology and physiology of bacteriophages as revealed by the electron microscope. *J. R. Microsc. Soc.*, 84:257–316.
21. Bradley, D.E., 1966, The structure and infective process of a *Pseudomonas aeruginosa* bacteriophage containing ribonucleic acid. *J. Gen. Microbiol.*, 45:83–96.
22. Bradley, D.E., 1972, Shortening of *Pseudomonas aeruginosa* pili after RNA-phage adsorption. *J. Gen. Microbiol.*, 72:303–319.
23. Bradley, D.E., 1972, Evidence for the retraction of *Pseudomonas aeruginosa* RNA phage pili. *Biochem. Biophys. Res. Commun.*, 47:142–149.
24. Bradley, D.E., 1972, Stimulation of pilus formation in *Pseudomonas aeruginosa* by RNA bacteriophage adsorption. *Biochem. Biophys. Res. Commun.*, 47:1080–1087.
25. Bradley, D.E., 1972, A study of pili on *Pseudomonas aeruginosa*. *Genet. Res. Cambridge*, 19:39–51.
26. Bradley, D.E., 1973, A pilus dependent *Pseudomonas aeruginosa* bacteriophage with a long non-retractile tail. *Virology*, 51:489–492.
27. Bradley, D.E., 1973, The adsorption of the *Pseudomonas aeruginosa* filamentous bacteriophage Pf to its host. *Can. J. Microbiol.*, 19:623–631.
28. Bradley, D.E., 1974, The adsorption of *Pseudomonas aeruginosa* pilus-dependent bacteriophages to a host mutant with nonretractile pili. *Virology*, 58:149–163.
29. Bradley, D.E., 1980, A function of *Pseudomonas aeruginosa* PAO polar pili: twitching motility. *Can. J. Microbiol.*, 26:146–154.
30. Bradley, D.E., and Pitt, T.L., 1974, Pilus-dependence of four *Pseudomonas aeruginosa* bacteriophages with non-contractile tails. *J. Gen. Virol.*, 23:1–15.
31. Bryan, R., Kube, D., Perez, A., Davis, P., and Prince, A., 1998, Overproduction of the CFTR R domain leads to increased levels of asialoGM1 and increased *Pseudomonas aeruginosa* binding by epithelial cells. *Am. J. Respir. Cell. Mol. Biol.*, 19:269–277.
32. Buell, C.R., Joardar, V., Lindeberg, M., Selengut, J., Paulsen, I.T., Gwinn, M.L., Dodson, R.J., Deboy, R.T., Durkin, A.S., Kolonay, J.F., Madupu, R., Daugherty, S., Brinkac, L., Beanan, M.J., Haft, D.H., Nelson, W.C., Davidsen, T., Zafar, N., Zhou, L., Liu, J., Yuan, Q., Khouri, H., Fedorova, N., Tran, B., Russell, D., Berry, K., Utterback, T., Van Aken, S.E., Feldblyum, T.V., D'Ascenzo, M., Deng, W.L., Ramos, A.R., Alfano, J.R., Cartinhour, S., Chatterjee, A.K., Delaney, T.P., Lazarowitz, S.G., Martin, G.B., Schneider, D.J., Tang, X., Bender, C.L., White, O., Fraser, C.M., and Collmer, A., 2003, The complete genome sequence of the *Arabidopsis* and tomato pathogen *Pseudomonas syringae* pv. *tomato* DC3000. *Proc. Natl. Acad. Sci. USA*, 100:10181–10186.
33. Cachia, P.J., Glasier, L.M., Hodgins, R.R., Wong, W.Y., Irvin, R.T., and Hodges, R.S., 1998, The use of synthetic peptides in the design of a consensus sequence vaccine for *Pseudomonas aeruginosa*. *J. Pept. Res.*, 52:289–299.
34. Cao, T.B., and Saier, M.H., Jr., 2001, Conjugal type IV macromolecular transfer systems of Gram-negative bacteria: organismal distribution, structural constraints and evolutionary conclusions. *Microbiology*, 147:3201–3214.
35. Carlson, C.A., Pierson, S.L., Rosen, J.J., and Ingraham, J.L., 1983, *Pseudomonas stutzeri* and related species undergo natural transformation. *J. Bacteriol.*, 153:93–99.
36. Carterson, A.J., Morici, L.A., Jackson, D.W., Frisk, A., Lizewski, S.E., Jupiter, R., Simpson, K., Kunz, D.A., Davis, S.H., Schurr, J.R., Hassett, D.J., and Schurr, M.J., 2004, The transcriptional regulator AlgR controls cyanide production in *Pseudomonas aeruginosa*. *J. Bacteriol.*, 186:6837–6844.

37. Castric, P., 1995, *pilO*, a gene required for glycosylation of *Pseudomonas aeruginosa* 1244 pilin. *Microbiology*, 141:1247–1254.
38. Castric, P., Cassels, F.J., and Carlson, R.W., 2001, Structural characterization of the *Pseudomonas aeruginosa* 1244 pilin glycan. *J. Biol. Chem.*, 276:26479–26485.
39. Chen, I., and Dubnau, D., 2004, DNA uptake during bacterial transformation. *Nat. Rev. Microbiol.*, 2:241–249.
40. Chi, E., Mehl, T., Nunn, D., and Lory, S., 1991, Interaction of *Pseudomonas aeruginosa* with A549 pneumocyte cells. *Infect. Immun.*, 59:822–828.
41. Chiang, P., and Burrows, L.L., 2003, Biofilm formation by hyperpiliated mutants of *Pseudomonas aeruginosa*. *J. Bacteriol.*, 185:2374–2378.
42. Christie, P.J., 2004, Type IV secretion: the *Agrobacterium* VirB/D4 and related conjugation systems. *Biochim. Biophys. Acta*, 1694:219–234.
43. Collins, R.F., Davidsen, L., Derrick, J.P., Ford, R.C., and Tonjum, T., 2001, Analysis of the PilQ secretin from *Neisseria meningitidis* by transmission electron microscopy reveals a dodecameric quaternary structure. *J. Bacteriol.*, 183:3825–3832.
44. Collins, R.F., Ford, R.C., Kitmitto, A., Olsen, R.O., Tonjum, T., and Derrick, J.P., 2003, Three-dimensional structure of the *Neisseria meningitidis* secretin PilQ determined from negative-stain transmission electron microscopy. *J. Bacteriol.*, 185:2611–2617.
45. Collins, R.F., Frye, S.A., Kitmitto, A., Ford, R.C., Tonjum, T., and Derrick, J.P., 2004, Structure of the *Neisseria meningitidis* outer membrane PilQ secretin complex at 12 Å resolution. *J. Biol. Chem.*, 279:39750–39756.
46. Comer, J.E., Marshall, M.A., Blanch, V.J., Deal, C.D., and Castric, P., 2002, Identification of the *Pseudomonas aeruginosa* 1244 pilin glycosylation site. *Infect. Immun.*, 70:2837–2845.
47. Comolli, J.C., Hauser, A.R., Waite, L., Whitchurch, C.B., Mattick, J.S., and Engel, J.N., 1999, *Pseudomonas aeruginosa* gene products PilT and PilU are required for cytotoxicity in vitro and virulence in a mouse model of acute pneumonia. *Infect. Immun.*, 67:3625–3630.
48. Comolli, J.C., Waite, L.L., Mostov, K.E., and Engel, J.N., 1999, Pili binding to asialo-GM1 on epithelial cells can mediate cytotoxicity or bacterial internalization by *Pseudomonas aeruginosa*. *Infect. Immun.*, 67:3207–3214.
49. Conrad, M.E., and Umbreit, J.N., 2002, Pathways of iron absorption. *Blood Cells Mol. Dis.*, 29:336–355.
50. Craig, L., Pique, M.E., and Tainer, J.A., 2004, Type IV pilus structure and bacterial pathogenicity. *Nat. Rev. Microbiol.*, 2:363–378.
51. Craig, L., Taylor, R.K., Pique, M.E., Adair, B.D., Arvai, A.S., Singh, M., Lloyd, S.J., Shin, D.S., Getzoff, E.D., Yeager, M., Forest, K.T., and Tainer, J.A., 2003, Type IV pilin structure and assembly: x-ray and EM analyses of *Vibrio cholerae* toxin-coregulated pilus and *Pseudomonas aeruginosa* PAK pilin. *Mol. Cell*, 11:1139–1150.
52. Croft, L., Beatson, S.A., Whitchurch, C.B., Huang, B., Blakeley, R.L., and Mattick, J.S., 2000, An interactive web-based *Pseudomonas aeruginosa* genome database: discovery of new genes, pathways and structures. *Microbiology*, 146:2351–2364.
53. Crowther, L.J., Anantha, R.P., and Donnenberg, M.S., 2004, The inner membrane subassembly of the enteropathogenic *Escherichia coli* bundle-forming pilus machine. *Mol. Microbiol.*, 52:67–79.
54. D'Andrea, L.D., and Regan, L., 2003, TPR proteins: the versatile helix. *Trends Biochem. Sci.*, 28:655–662.
55. D'Argenio, D.A., Gallagher, L.A., Berg, C.A., and Manoil, C., 2001, *Drosophila* as a model host for *Pseudomonas aeruginosa* infection. *J. Bacteriol.*, 183:1466–1471.
56. D'Argenio, D.A., and Miller, S.I., 2004, Cyclic di-GMP as a bacterial second messenger. *Microbiology*, 150:2497–2502.

57. Darzins, A., 1993, The *pilG* gene product, required for *Pseudomonas aeruginosa* pilus production and twitching motility, is homologous to the enteric, single-domain response regulator CheY. *J. Bacteriol.*, 175:5934–5944.
58. Darzins, A., 1994, Characterization of a *Pseudomonas aeruginosa* gene cluster involved in pilus biosythesis and twitching motility: sequence similarity to the chemotaxis proteins of enterics and the gliding bacterium *Myxococcus xanthus*. *Mol. Microbiol.*, 11:137–153.
59. Darzins, A., 1995, The *Pseudomonas aeruginosa pilK* gene encodes a chemotactic methyltransferase (CheR) homologue that is translationally regulated. *Mol. Microbiol.*, 15:703–717.
60. Davies, J., Dewar, A., Bush, A., Pitt, T., Gruenert, D., Geddes, D.M., and Alton, E.W., 1999, Reduction in the adherence of *Pseudomonas aeruginosa* to native cystic fibrosis epithelium with anti-asialoGM1 antibody and neuraminidase inhibition. *Eur. Respir. J.*, 13:565–570.
61. de Bentzmann, S., Roger, P., Dupuit, F., Bajolet-Laudinat, O., Fuchey, C., Plotkowski, M.C., and Puchelle, E., 1996, Asialo GM1 is a receptor for *Pseudomonas aeruginosa* adherence to regenerating respiratory epithelial cells. *Infect. Immun.*, 64:1582–1588.
62. de Bentzmann, S., Roger, P., and Puchelle, E., 1996, *Pseudomonas aeruginosa* adherence to remodelling respiratory epithelium. *Eur. Respir. J.*, 9:2145–2150.
63. de Groot, A., Heijnen, I., de Cock, H., Filloux, A., and Tommassen, J., 1994, Characterization of type IV pilus genes in plant growth-promoting *Pseudomonas putida* WCS358. *J. Bacteriol.*, 176:642–650.
64. Demaneche, S., Kay, E., Gourbiere, F., and Simonet, P., 2001, Natural transformation of *Pseudomonas fluorescens* and *Agrobacterium tumefaciens* in soil. *Appl. Environ. Microbiol.*, 67:2617–2621.
65. Deziel, E., Comeau, Y., and Villemur, R., 2001, Initiation of biofilm formation by *Pseudomonas aeruginosa* 57RP correlates with emergence of hyperpiliated and highly adherent phenotypic variants deficient in swimming, swarming, and twitching motilities. *J. Bacteriol.*, 183:1195–1204.
66. Dobbs, L.G., 1989, Pulmonary surfactant. *Annu. Rev. Med.*, 40:431–446.
67. Doig, P., Paranchych, W., Sastry, P.A., and Irvin, R.T., 1989, Human buccal epithelial cell receptors of *Pseudomonas aeruginosa*: identification of glycoproteins with pilus binding activity. *Can. J. Microbiol.*, 35:1141–1145.
68. Doig, P., Sastry, P.A., Hodges, R.S., Lee, K.K., Paranchych, W., and Irvin, R.T., 1990, Inhibition of pilus-mediated adhesion of *Pseudomonas aeruginosa* to human buccal epithelial cells by monoclonal antibodies directed against pili. *Infect. Immun.*, 58:124–130.
69. Doig, P., Todd, T., Sastry, P.A., Lee, K.K., Hodges, R.S., Paranchych, W., and Irvin, R.T., 1988, Role of pili in adhesion of *Pseudomonas aeruginosa* to human respiratory epithelial cells. *Infect. Immun.*, 56:1641–1646.
70. Drake, S.L., Sandstedt, S.A., and Koomey, M., 1997, PilP, a pilus biogenesis lipoprotein in *Neisseria gonorrhoeae*, affects expression of PilQ as a high-molecular-mass multimer. *Mol. Microbiol.*, 23:657–668.
71. Drenkard, E., and Ausubel, F.M., 2002, *Pseudomonas* biofilm formation and antibiotic resistance are linked to phenotypic variation. *Nature*, 416:740–743.
72. Duguid, J.P., and Anderson, E.S., 1967, Terminology of bacterial fimbriae, or pili, and their types. *Nature*, 215:89–90.
73. Errington, J., Daniel, R.A., and Scheffers, D.J., 2003, Cytokinesis in bacteria. *Microbiol. Mol. Biol. Rev.*, 67:52–65.
74. Ethier, J., and Boyd, J.M., 2000, Topological analysis and role of the transmembrane domain in polar targeting of PilS, a *Pseudomonas aeruginosa* sensor kinase. *Mol. Microbiol.*, 38:891–903.
75. Farinha, M.A., Conway, B.D., Glasier, L.M., Ellert, N.W., Irvin, R.T., Sherburne, R., and Paranchych, W., 1994, Alteration of the pilin adhesin of *Pseudomonas aeruginosa* PAO results

in normal pilus biogenesis but a loss of adherence to human pneumocyte cells and decreased virulence in mice. *Infect. Immun.*, 62:4118–4123.
76. Filloux, A., 2004, The underlying mechanisms of type II protein secretion. *Biochim. Biophys. Acta.* 1694:163–179.
77. Folkhard, W., Marvin, D.A., Watts, T.H., and Paranchych, W., 1981, Structure of polar pili from *Pseudomonas aeruginosa* strains K and O. *J. Mol. Biol.*, 149:79–93.
78. Forest, K.T., Dunham, S.A., Koomey, M., and Tainer, J.A., 1999, Crystallographic structure reveals phosphorylated pilin from *Neisseria*: phosphoserine sites modify type IV pilus surface chemistry and fibre morphology. *Mol. Microbiol.*, 31:743–752.
79. Forest, K.T., Satyshur, K.A., Worzalla, G.A., Hansen, J.K., and Herdendorf, T.J., 2004, The pilus-retraction protein PilT: ultrastructure of the biological assembly. *Acta Crystallogr. D. Biol. Crystallogr.*, 60:978–982.
80. Frost, L.S., Carpenter, M., and Paranchych, W., 1978, N-methylphenylalanine at the N-terminus of pilin isolated from *Pseudomonas aeruginosa* K. *Nature*, 271:87–89.
81. Frost, L.S., and Paranchych, W., 1977, Composition and molecular weight of pili purified from *Pseudomonas aeruginosa* K. *J. Bacteriol.*, 131:259–269.
82. Fuerst, J.A., and Hayward, A.C., 1969, Surface appendages similar to fimbriae (pili) on *Pseudomonas* species. *J. Gen. Microbiol.*, 58:227–237.
83. Glessner, A., Smith, R.S., Iglewski, B.H., and Robinson, J.B., 1999, Roles of *Pseudomonas aeruginosa las* and *rhl* quorum-sensing systems in control of twitching motility. *J. Bacteriol.*, 181:1623–1629.
84. Graupner, S., Frey, V., Hashemi, R., Lorenz, M.G., Brandes, G., and Wackernagel, W., 2000, Type IV pilus genes *pilA* and *pilC* of *Pseudomonas stutzeri* are required for natural genetic transformation, and *pilA* can be replaced by corresponding genes from nontransformable species. *J. Bacteriol.*, 182:2184–2190.
85. Graupner, S., and Wackernagel, W., 2001, Identification and characterization of novel competence genes *comA* and *exbB* involved in natural genetic transformation of *Pseudomonas stutzeri*. *Res. Microbiol.*, 152:451–460.
86. Graupner, S., and Wackernagel, W., 2001, *Pseudomonas stutzeri* has two closely related *pilA* genes (Type IV pilus structural protein) with opposite influences on natural genetic transformation. *J. Bacteriol.*, 183:2359–2366.
87. Graupner, S., Weger, N., Sohni, M., and Wackernagel, W., 2001, Requirement of novel competence genes *pilT* and *pilU* of *Pseudomonas stutzeri* for natural transformation and suppression of *pilT* deficiency by a hexahistidine tag on the type IV pilus protein PilAI. *J. Bacteriol.*, 183:4694–4701.
88. Gupta, S.K., Berk, R.S., Masinick, S., and Hazlett, L.D., 1994, Pili and lipopolysaccharide of *Pseudomonas aeruginosa* bind to the glycolipid asialo GM1. *Infect. Immun.*, 62:4572–4579.
89. Hahn, H.P., 1997, The type-4 pilus is the major virulence-associated adhesin of *Pseudomonas aeruginosa* – a review. *Gene*, 192:99–108.
90. Hambrook, J., Titball, R., and Lindsay, C., 2004, The interaction of *Pseudomonas aeruginosa* PAK with human and animal respiratory tract cell lines. *FEMS Microbiol. Lett.*, 238:49–55.
91. Haussler, S., 2004, Biofilm formation by the small colony variant phenotype of *Pseudomonas aeruginosa*. *Environ. Microbiol.*, 6:546–551.
92. Haussler, S., Ziegler, I., Lottel, A., von Gotz, F., Rohde, M., Wehmhohner, D., Saravanamuthu, S., Tummler, B., and Steinmetz, I., 2003, Highly adherent small-colony variants of *Pseudomonas aeruginosa* in cystic fibrosis lung infection. *J. Med. Microbiol.*, 52:295–301.
93. Hazes, B., Sastry, P.A., Hayakawa, K., Read, R.J., and Irvin, R.T., 2000, Crystal structure of *Pseudomonas aeruginosa* PAK pilin suggests a main-chain-dominated mode of receptor binding. *J. Mol. Biol.*, 299:1005–1017.

94. Hazlett, L.D., Moon, M.M., Singh, A., Berk, R.S., and Rudner, X.L., 1991, Analysis of adhesion, piliation, protease production and ocular infectivity of several *P. aeruginosa* strains. *Curr. Eye Res.*, 10:351–362.
95. Hazlett, L.D., Moon, M.M., Strejc, M., and Berk, R.S., 1987, Evidence for *N*-acetylmannosamine as an ocular receptor for *P. aeruginosa* adherence to scarified cornea. *Invest. Ophthalmol. Vis. Sci.*, 28:1978–1985.
96. Hendrickson, E.L., Plotnikova, J., Mahajan-Miklos, S., Rahme, L.G., and Ausubel, F.M., 2001, Differential roles of the *Pseudomonas aeruginosa* PA14 *rpoN* gene in pathogenicity in plants, nematodes, insects, and mice. *J. Bacteriol.*, 183:7126–7134.
97. Henrichsen, J., 1972, Bacterial surface translocation: a survey and a classification. *Bacteriol. Rev.*, 36:478–503.
98. Henrichsen, J., 1975, The occurrence of twitching motility among gram-negative bacteria. *Acta Pathol. Microbiol. Scand. [B]*, 83:171–178.
99. Henrichsen, J., 1983, Twitching motility. *Annu. Rev. Microbiol.*, 37:81–93.
100. Henriksen, S.D., and Henrichsen, J., 1975, Twitching motility and possession of polar fimbriae in spreading *Streptococcus sanguis* isolates from the human throat. *Acta. Pathol. Microbiol. Scand. [B]*, 83:133–140.
101. Herdendorf, T.J., McCaslin, D.R., and Forest, K.T., 2002, *Aquifex aeolicus* PilT, homologue of a surface motility protein, is a thermostable oligomeric NTPase. *J. Bacteriol.*, 184:6465–6471.
102. Hobbs, M., Collie, E.S., Free, P.D., Livingston, S.P., and Mattick, J.S., 1993, PilS and PilR, a two-component transcriptional regulatory system controlling expression of type 4 fimbriae in *Pseudomonas aeruginosa*. *Mol. Microbiol.*, 7:669–682.
103. Hogan, D.A., and Kolter, R., 2002, *Pseudomonas–Candida* interactions: an ecological role for virulence factors. *Science*, 296:2229–2232.
104. Houwink, A., and van Iterson, W., 1950, Electron microscopical observations on bacterial cytology. II. A study on flagellation. *Biochim. Biophys. Acta*, 5:10–44.
105. Huang, B., Ru, K., Yuan, Z., Whitchurch, C.B., and Mattick, J.S., 2004, *tonB3* is required for normal twitching motility and extracellular assembly of type IV pili. *J. Bacteriol.*, 186:4387–4389.
106. Huang, B., Whitchurch, C.B., and Mattick, J.S., 2003, FimX, a multidomain protein connecting environmental signals to twitching motility in *Pseudomonas aeruginosa*. *J. Bacteriol.*, 185:7068–7076.
107. Imundo, L., Barasch, J., Prince, A., and Al-Awqati, Q., 1995, Cystic fibrosis epithelial cells have a receptor for pathogenic bacteria on their apical surface. *Proc. Natl. Acad. Sci. USA*, 92:3019–3023.
108. Irvin, R.T., Doig, P., Lee, K.K., Sastry, P.A., Paranchych, W., Todd, T., and Hodges, R.S., 1989, Characterization of the *Pseudomonas aeruginosa* pilus adhesin: confirmation that the pilin structural protein subunit contains a human epithelial cell-binding domain. *Infect. Immun.*, 57:3720–3726.
109. Ishimoto, K.S., and Lory, S., 1989, Formation of pilin in *Pseudomonas aeruginosa* requires the alternative sigma factor (RpoN) of RNA polymerase. *Proc. Natl. Acad. Sci. USA*, 86:1954–1957.
110. Ishimoto, K.S., and Lory, S., 1992, Identification of *pilR*, which encodes a transcriptional activator of the *Pseudomonas aeruginosa* pilin gene. *J. Bacteriol.*, 174:3514–3521.
111. Jenal, U., 2004, Cyclic di-guanosine-monophosphate comes of age: a novel secondary messenger involved in modulating cell surface structures in bacteria? *Curr. Opin. Microbiol.*, 7:185–191.
112. Jendrossek, V., Fillon, S., Belka, C., Muller, I., Puttkammer, B., and Lang, F., 2003, Apoptotic response of Chang cells to infection with *Pseudomonas aeruginosa* strains PAK and PAO-I: molecular ordering of the apoptosis signaling cascade and role of type IV pili. *Infect. Immun.*, 71:2665–2673.

113. Jin, S., Ishimoto, K.S., and Lory, S., 1994, PilR, a transcriptional regulator of piliation in *Pseudomonas aeruginosa*, binds to a cis-acting sequence upstream of the pilin gene promoter. *Mol. Microbiol.*, 14:1049–1057.
114. Johanson, W.G., Jr., Higuchi, J.H., Chaudhuri, T.R., and Woods, D.E., 1980, Bacterial adherence to epithelial cells in bacillary colonization of the respiratory tract. *Am. Rev. Respir. Dis.*, 121:55–63.
115. Johanson, W.G., Jr., Pierce, A.K., Sanford, J.P., and Thomas, G.D., 1972, Nosocomial respiratory infections with gram-negative bacilli. The significance of colonization of the respiratory tract. *Ann. Intern. Med.*, 77:701–706.
116. Johanson, W.G., Jr., Woods, D.E., and Chaudhuri, T., 1979, Association of respiratory tract colonization with adherence of gram-negative bacilli to epithelial cells. *J. Infect. Dis.*, 139:667–673.
117. Kang, P.J., Hauser, A.R., Apodaca, G., Fleiszig, S.M., Wiener-Kronish, J., Mostov, K., and Engel, J.N., 1997, Identification of *Pseudomonas aeruginosa* genes required for epithelial cell injury. *Mol. Microbiol.*, 24:1249–1262.
118. Kearns, D.B., Robinson, J., and Shimkets, L.J., 2001, *Pseudomonas aeruginosa* exhibits directed twitching motility up phosphatidylethanolamine gradients. *J. Bacteriol.*, 183:763–767.
119. Kelly, N.M., Kluftinger, J.L., Pasloske, B.L., Paranchych, W., and Hancock, R.E., 1989, *Pseudomonas aeruginosa* pili as ligands for nonopsonic phagocytosis by fibronectin-stimulated macrophages. *Infect. Immun.*, 57:3841–3845.
120. Kelly-Wintenberg, K., and Montie, T.C., 1994, Chemotaxis to oligopeptides by *Pseudomonas aeruginosa*. *Appl. Environ. Microbiol.*, 60:363–367.
121. Klausen, M., Aaes-Jorgensen, A., Molin, S., and Tolker-Nielsen, T., 2003, Involvement of bacterial migration in the development of complex multicellular structures in *Pseudomonas aeruginosa* biofilms. *Mol. Microbiol.*, 50:61–68.
122. Klausen, M., Heydorn, A., Ragas, P., Lambertsen, L., Aaes-Jorgensen, A., Molin, S., and Tolker-Nielsen, T., 2003, Biofilm formation by *Pseudomonas aeruginosa* wild type, flagella and type IV pili mutants. *Mol. Microbiol.*, 48:1511–1524.
123. Koga, T., Ishimoto, K., and Lory, S., 1993, Genetic and functional characterization of the gene cluster specifying expression of *Pseudomonas aeruginosa* pili. *Infect. Immun.*, 61:1371–1377.
124. Kornberg, A., Rao, N.N., and Ault-Riche, D., 1999, Inorganic polyphosphate: a molecule of many functions. *Annu. Rev. Biochem.*, 68:89–125.
125. Krivan, H.C., Ginsburg, V., and Roberts, D.D., 1988, *Pseudomonas aeruginosa* and *Pseudomonas cepacia* isolated from cystic fibrosis patients bind specifically to gangliotetraosylceramide (asialo GM1) and gangliotriaosylceramide (asialo GM2). *Arch. Biochem. Biophys.*, 260:493–496.
126. Krivan, H.C., Roberts, D.D., and Ginsburg, V., 1988, Many pulmonary pathogenic bacteria bind specifically to the carbohydrate sequence GalNAc beta 1–4Gal found in some glycolipids. *Proc. Natl. Acad. Sci. USA*, 85:6157–6161.
127. Kus, J.V., Tullis, E., Cvitkovitch, D.G., and Burrows, L.L., 2004, Significant differences in type IV pilin allele distribution among *Pseudomonas aeruginosa* isolates from cystic fibrosis (CF) versus non-CF patients. *Microbiology*, 150:1315–1326.
128. Lautrop, H., 1961, *Bacterium anitratum* transferred to the genus *Cytophaga*. *Int. Bull. Bacteriol. Nomencl.*, 11:107–108.
129. Lautrop, H., 1965, Gliding motility in bacteria as a taxonomic criterion. *Publ. Fac. Sci. Univ. J. E. Purkyne, ser K*, 35:322–327.
130. Lee, K.K., Doig, P., Irvin, R.T., Paranchych, W., and Hodges, R.S., 1989, Mapping the surface regions of *Pseudomonas aeruginosa* PAK pilin: the importance of the C-terminal region for adherence to human buccal epithelial cells. *Mol. Microbiol.*, 3:1493–1499.

131. Lee, K.K., Sastry, P.A., Paranchych, W., and Hodges, R.S., 1989, Immunological studies of the disulfide bridge region of *Pseudomonas aeruginosa* PAK and PAO pilins, using anti-PAK pilus and antipeptide antibodies. *Infect. Immun.*, 57:520–526.
132. Lee, K.K., Sheth, H.B., Wong, W.Y., Sherburne, R., Paranchych, W., Hodges, R.S., Lingwood, C.A., Krivan, H., and Irvin, R.T., 1994, The binding of *Pseudomonas aeruginosa* pili to glycosphingolipids is a tip-associated event involving the C-terminal region of the structural pilin subunit. *Mol. Microbiol.*, 11:705–713.
133. Lee, K.K., Yu, L., Macdonald, D.L., Paranchych, W., Hodges, R.S., and Irvin, R.T., 1996, Anti-adhesin antibodies that recognize a receptor-binding motif (adhesintope) inhibit pilus/fimbrial-mediated adherence of *Pseudomonas aeruginosa* and *Candida albicans* to asialo-GM1 receptors and human buccal epithelial cell surface receptors. *Can. J. Microbiol.*, 42:479–486.
134. Lizewski, S.E., Schurr, J.R., Jackson, D.W., Frisk, A., Carterson, A.J., and Schurr, M.J., 2004, Identification of AlgR-regulated genes in *Pseudomonas aeruginosa* by use of microarray analysis. *J. Bacteriol.*, 186:5672–5684.
135. Lorenz, M.G., and Wackernagel, W., 1994, Bacterial gene transfer by natural genetic transformation in the environment. *Microbiol. Rev.*, 58:563–602.
136. Lowe, J., van den Ent, F., and Amos, L.A., 2004, Molecules of the bacterial cytoskeleton. *Annu. Rev. Biophys. Biomol. Struct.*, 33:177–198.
137. Lu, H.M., Motley, S.T., and Lory, S., 1997, Interactions of the components of the general secretion pathway: role of *Pseudomonas aeruginosa* type IV pilin subunits in complex formation and extracellular protein secretion. *Mol. Microbiol.*, 25:247–259.
138. Lugtenberg, B.J., Chin, A.W. T.F., and Bloemberg, G.V., 2002, Microbe-plant interactions: principles and mechanisms. *Antonie Van Leeuwenhoek*, 81:373–383.
139. Mahenthiralingam, E., and Speert, D.P., 1995, Nonopsonic phagocytosis of *Pseudomonas aeruginosa* by macrophages and polymorphonuclear leukocytes requires the presence of the bacterial flagellum. *Infect. Immun.*, 63:4519–4523.
140. Maier, B., Potter, L., So, M., Long, C.D., Seifert, H.S., and Sheetz, M.P., 2002, Single pilus motor forces exceed 100 pN. *Proc. Natl. Acad. Sci. USA*, 99:16012–16017.
141. Martin, P.R., Hobbs, M., Free, P.D., Jeske, Y., and Mattick, J.S., 1993, Characterization of *pilQ*, a new gene required for the biogenesis of type 4 fimbriae in *Pseudomonas aeruginosa*. *Mol. Microbiol.*, 9:857–868.
142. Martin, P.R., Watson, A.A., McCaul, T.F., and Mattick, J.S., 1995, Characterization of a five-gene cluster required for the biogenesis of type 4 fimbriae in *Pseudomonas aeruginosa*. *Mol. Microbiol.*, 16:497–508.
143. Mattick, J.S., 2002, Type IV pili and twitching motility. *Annu. Rev. Microbiol.*, 56:289–314.
144. Merz, A.J., and Forest, K.T., 2002, Bacterial surface motility: slime trails, grappling hooks and nozzles. *Curr. Biol.*, 12:R297–R303.
145. Merz, A.J., So, M., and Sheetz, M.P., 2000, Pilus retraction powers bacterial twitching motility. *Nature*, 407:98–102.
146. Meyer, K.C., Sharma, A., Brown, R., Weatherly, M., Moya, F.R., Lewandoski, J., and Zimmerman, J.J., 2000, Function and composition of pulmonary surfactant and surfactant-derived fatty acid profiles are altered in young adults with cystic fibrosis. *Chest*, 118:164–174.
147. Morand, P.C., Bille, E., Morelle, S., Eugene, E., Beretti, J.L., Wolfgang, M., Meyer, T.F., Koomey, M., and Nassif, X., 2004, Type IV pilus retraction in pathogenic *Neisseria* is regulated by the PilC proteins. *EMBO J.*, 23:2009–2017.
148. Nelson, J.W., Tredgett, M.W., Sheehan, J.K., Thornton, D.J., Notman, D., and Govan, J.R., 1990, Mucinophilic and chemotactic properties of *Pseudomonas aeruginosa* in relation to pulmonary colonization in cystic fibrosis. *Infect. Immun.*, 58:1489–1495.
149. Nelson, K.E., Weinel, C., Paulsen, I.T., Dodson, R.J., Hilbert, H., Martins dos Santos, V.A., Fouts, D.E., Gill, S.R., Pop, M., Holmes, M., Brinkac, L., Beanan, M., DeBoy, R.T., Daugherty,

S., Kolonay, J., Madupu, R., Nelson, W., White, O., Peterson, J., Khouri, H., Hance, I., Chris Lee, P., Holtzapple, E., Scanlan, D., Tran, K., Moazzez, A., Utterback, T., Rizzo, M., Lee, K., Kosack, D., Moestl, D., Wedler, H., Lauber, J., Stjepandic, D., Hoheisel, J., Straetz, M., Heim, S., Kiewitz, C., Eisen, J.A., Timmis, K.N., Dusterhoft, A., Tummler, B., and Fraser, C.M., 2002, Complete genome sequence and comparative analysis of the metabolically versatile *Pseudomonas putida* KT2440. *Environ. Microbiol.*, 4:799–808.

150. Nikolskaya, A.N., and Galperin, M.Y., 2002, A novel type of conserved DNA-binding domain in the transcriptional regulators of the AlgR/AgrA/LytR family. *Nucleic Acids Res.*, 30:2453–2459.
151. Nunn, D., Bergman, S., and Lory, S., 1990, Products of three accessory genes, *pilB*, *pilC*, and *pilD*, are required for biogenesis of *Pseudomonas aeruginosa* pili. *J. Bacteriol.*, 172:2911–2919.
152. Nunn, D.N., and Lory, S., 1991, Product of the *Pseudomonas aeruginosa* gene *pilD* is a prepilin leader peptidase. *Proc. Natl. Acad. Sci. USA*, 88:3281–3285.
153. O'Toole, G.A., Gibbs, K.A., Hager, P.W., Phibbs, P.V., Jr., and Kolter, R., 2000, The global carbon metabolism regulator Crc is a component of a signal transduction pathway required for biofilm development by *Pseudomonas aeruginosa*. *J. Bacteriol.*, 182:425–431.
154. O'Toole, G.A., and Kolter, R., 1998, Flagellar and twitching motility are necessary for *Pseudomonas aeruginosa* biofilm development. *Mol. Microbiol.*, 30:295–304.
155. Okamoto, S., and Ohmori, M., 2002, The cyanobacterial PilT protein responsible for cell motility and transformation hydrolyzes ATP. *Plant Cell Physiol.*, 43:1127–1136.
156. Ottow, J.C., 1975, Ecology, physiology, and genetics of fimbriae and pili. *Annu. Rev. Microbiol.*, 29:79–108.
157. Paranchych, W., Frost, L.S., and Carpenter, M., 1978, N-Terminal amino acid sequence of pilin isolated from *Pseudomonas aeruginosa*. *J. Bacteriol.*, 134:1179–1180.
158. Paranchych, W., Sastry, P.A., Drake, D., Pearlstone, J.R., and Smillie, L.B., 1985, *Pseudomonas* pili. Studies on antigenic determinants and mammalian cell receptors. *Antibiot. Chemother.*, 36:49–57.
159. Paranchych, W., Sastry, P.A., Frost, L.S., Carpenter, M., Armstrong, G.D., and Watts, T.H., 1979, Biochemical studies on pili isolated from *Pseudomonas aeruginosa* strain PAO. *Can. J. Microbiol.*, 25:1175–1181.
160. Parge, H.E., Forest, K.T., Hickey, M.J., Christensen, D.A., Getzoff, E.D., and Tainer, J.A., 1995, Structure of the fibre-forming protein pilin at 2.6 A resolution. *Nature*, 378:32–38.
161. Pasloske, B.L., Finlay, B.B., and Paranchych, W., 1985, Cloning and sequencing of the *Pseudomonas aeruginosa* PAK pilin gene. *FEBS Lett.*, 183:408–412.
162. Peabody, C.R., Chung, Y.J., Yen, M.R., Vidal-Ingigliardi, D., Pugsley, A.P., and Saier, M.H., Jr., 2003, Type II protein secretion and its relationship to bacterial type IV pili and archaeal flagella. *Microbiology*, 149:3051–3072.
163. Pegden, R.S., Larson, M.A., Grant, R.J., and Morrison, M., 1998, Adherence of the gram-positive bacterium *Ruminococcus albus* to cellulose and identification of a novel form of cellulose-binding protein which belongs to the Pil family of proteins. *J. Bacteriol.*, 180:5921–5927.
164. Planet, P.J., Kachlany, S.C., DeSalle, R., and Figurski, D.H., 2001, Phylogeny of genes for secretion NTPases: identification of the widespread *tadA* subfamily and development of a diagnostic key for gene classification. *Proc. Natl. Acad. Sci. USA*, 98:2503–2508.
165. Ponting, C.P., and Aravind, L., 1997, PAS: a multifunctional domain family comes to light. *Curr. Biol.*, 7:R674–R677.
166. Postle, K., and Kadner, R.J., 2003, Touch and go: tying TonB to transport. *Mol. Microbiol.*, 49:869–882.
167. Powell, J.J., Jugdaohsingh, R., and Thompson, R.P., 1999, The regulation of mineral absorption in the gastrointestinal tract. *Proc. Nutr. Soc.*, 58:147–153.

168. Rakotoarivonina, H., Jubelin, G., Hebraud, M., Gaillard-Martinie, B., Forano, E., and Mosoni, P., 2002, Adhesion to cellulose of the Gram-positive bacterium *Ruminococcus albus* involves type IV pili. *Microbiology*, 148:1871–1880.
169. Ramos, J.L., Duque, E., Gallegos, M.T., Godoy, P., Ramos-González, M.I., Rojas, A., Terán, W., and Segura, A., 2002, Mechanisms of solvent tolerance in gram-negative bacteria. *Annu. Rev. Microbiol.*, 56:743–768.
170. Ramphal, R., and Arora, S.K., 2001, Recognition of mucin components by *Pseudomonas aeruginosa*. *Glycoconj. J.*, 18:709–713.
171. Ramphal, R., Koo, L., Ishimoto, K.S., Totten, P.A., Lara, J.C., and Lory, S., 1991, Adhesion of *Pseudomonas aeruginosa* pilin-deficient mutants to mucin. *Infect. Immun.*, 59:1307–1311.
172. Ramphal, R., McNiece, M.T., and Polack, F.M., 1981, Adherence of *Pseudomonas aeruginosa* to the injured cornea: a step in the pathogenesis of corneal infections. *Ann. Ophthalmol.*, 13:421–425.
173. Ramphal, R., and Pyle, M., 1983, Evidence for mucins and sialic acid as receptors for *Pseudomonas aeruginosa* in the lower respiratory tract. *Infect. Immun.*, 41:339–344.
174. Ramphal, R., Sadoff, J.C., Pyle, M., and Silipigni, J.D., 1984, Role of pili in the adherence of *Pseudomonas aeruginosa* to injured tracheal epithelium. *Infect. Immun.*, 44:38–40.
175. Ramphal, R., Small, P.M., Shands, J.W., Jr., Fischlschweiger, W., and Small, P.A., Jr., 1980, Adherence of Pseudomonas aeruginosa to tracheal cells injured by influenza infection or by endotracheal intubation. *Infect. Immun.*, 27:614–619.
176. Rashid, M.H., and Kornberg, A., 2000, Inorganic polyphosphate is needed for swimming, swarming, and twitching motilities of *Pseudomonas aeruginosa*. *Proc. Natl. Acad. Sci. USA*, 97:4885–4890.
177. Robien, M.A., Krumm, B.E., Sandkvist, M., and Hol, W.G., 2003, Crystal structure of the extracellular protein secretion NTPase EpsE of *Vibrio cholerae*. *J. Mol. Biol.*, 333:657–674.
178. Rodriguez-Soto, J.P., and Kaiser, D., 1997, Identification and localization of the Tgl protein, which is required for *Myxococcus xanthus* social motility. *J. Bacteriol.*, 179:4372–4381.
179. Rodriguez-Soto, J.P., and Kaiser, D., 1997, The *tgl* gene: social motility and stimulation in *Myxococcus xanthus*. *J. Bacteriol.*, 179:4361–4371.
180. Roine, E., Nunn, D.N., Paulin, L., and Romantschuk, M., 1996, Characterization of genes required for pilus expression in *Pseudomonas syringae* pathovar *phaseolicola*. *J. Bacteriol.*, 178:410–417.
181. Roine, E., Raineri, D.M., Romantschuk, M., Wilson, M., and Nunn, D.N., 1998, Characterization of type IV pilus genes in *Pseudomonas syringae* pv. *tomato* DC3000. *Mol. Plant Microbe Interact.*, 11:1048–1056.
182. Romantschuk, M., 1992, Attachment of plant pathogenic bacteria to plant surfaces. *Annu. Rev. Phytopathol.*, 30:225–243.
183. Romantschuk, M., and Bamford, D.H., 1985, Function of pili in bacteriophage phi 6 penetration. *J. Gen. Virol.*, 66:2461–2469.
184. Romantschuk, M., and Bamford, D.H., 1986, The causal agent of halo blight in bean, *Pseudomonas syringae* pv. *phaseolicola*, attaches to stomata via its pili. *Microb. Pathog.*, 1:139–148.
185. Romantschuk, M., Nurmiaho-Lassila, E.-L., Roine, E., and Suoniemi, A., 1993, Pilus-mediated adsorption of *Pseudomonas syringae* to host and nonhost plant leaves. *J. Gen. Microbiol.*, 139:2252–2260.
186. Rose, M.C., 1992, Mucins: structure, function, and role in pulmonary diseases. *Am. J. Physiol.*, 263:L413–L429.
187. Rudner, X.L., Berk, R.S., and Hazlett, L.D., 1993, Immunization with homologous *Pseudomonas aeruginosa* pili protects against ocular disease. *Reg. Immunol.*, 5:245–252.

188. Russell, M.A., and Darzins, A., 1994, The *pilE* gene product of *Pseudomonas aeruginosa*, required for pilus biogenesis, shares amino acid sequence identity with the N-termini of type 4 prepilin proteins. *Mol. Microbiol.*, 13:973–985.
189. Saiman, L., Ishimoto, K., Lory, S., and Prince, A., 1990, The effect of piliation and exoproduct expression on the adherence of *Pseudomonas aeruginosa* to respiratory epithelial monolayers. *J. Infect. Dis.*, 161:541–548.
190. Sakai, D., Horiuchi, T., and Komano, T., 2001, Atpase activity and multimer formation of PilQ protein are required for thin pilus biogenesis in plasmid R64. *J. Biol. Chem.*, 276:17968–17975.
191. Sandkvist, M., 2001, Biology of type II secretion. *Mol. Microbiol.*, 40:271–283.
192. Sastry, P.A., Finlay, B.B., Pasloske, B.L., Paranchych, W., Pearlstone, J.R., and Smillie, L.B., 1985, Comparative studies of the amino acid and nucleotide sequences of pilin derived from *Pseudomonas aeruginosa* PAK and PAO. *J. Bacteriol.*, 164:571–577.
193. Sastry, P.A., Pearlstone, J.R., Smillie, L.B., and Paranchych, W., 1983, Amino acid sequence of pilin isolated from pseudomonas aeruginosa PAK. *FEBS Lett.*, 151:253–256.
194. Sato, H., and Okinaga, K., 1987, Role of pili in the adherence of *Pseudomonas aeruginosa* to mouse epidermal cells. *Infect. Immun.*, 55:1774–1778.
195. Sauvonnet, N., Vignon, G., Pugsley, A.P., and Gounon, P., 2000, Pilus formation and protein secretion by the same machinery in *Escherichia coli*. *EMBO J.*, 19:2221–2228.
196. Savvides, S.N., Yeo, H.J., Beck, M.R., Blaesing, F., Lurz, R., Lanka, E., Buhrdorf, R., Fischer, W., Haas, R., and Waksman, G., 2003, VirB11 ATPases are dynamic hexameric assemblies: new insights into bacterial type IV secretion. *EMBO J.*, 22:1969–1980.
197. Schroeder, T.H., Zaidi, T., and Pier, G.B., 2001, Lack of adherence of clinical isolates of *Pseudomonas aeruginosa* to asialo-GM(1) on epithelial cells. *Infect. Immun.*, 69:719–729.
198. Semmler, A.B.T., Whitchurch, C.B., Leech, A.J., and Mattick, J.S., 2000, Identification of a novel gene, *fimV,* involved in twitching motility in *Pseudomonas aeruginosa*. *Microbiology*, 146:1331–1332.
199. Semmler, A.B.T., Whitchurch, C.B., and Mattick, J.S., 1999, A re-examination of twitching motility in *Pseudomonas aeruginosa*. *Microbiology*, 145:2863–2873.
200. Sheth, H.B., Glasier, L.M., Ellert, N.W., Cachia, P., Kohn, W., Lee, K.K., Paranchych, W., Hodges, R.S., and Irvin, R.T., 1995, Development of an anti-adhesive vaccine for *Pseudomonas aeruginosa* targeting the C-terminal region of the pilin structural protein. *Biomed. Pept. Proteins Nucleic Acids*, 1:141–148.
201. Sheth, H.B., Lee, K.K., Wong, W.Y., Srivastava, G., Hindsgaul, O., Hodges, R.S., Paranchych, W., and Irvin, R.T., 1994, The pili of *Pseudomonas aeruginosa* strains PAK and PAO bind specifically to the carbohydrate sequence beta GalNAc(1–4)beta Gal found in glycosphingolipids asialo-GM1 and asialo-GM2. *Mol. Microbiol.*, 11:715–723.
202. Simm, R., Morr, M., Kader, A., Nimtz, M., and Romling, U., 2004, GGDEF and EAL domains inversely regulate cyclic di-GMP levels and transition from sessility to motility. *Mol. Microbiol.*, 53:1123–1134.
203. Simunovic, V., Gherardini, F.C., and Shimkets, L.J., 2003, Membrane localization of motility, signaling, and polyketide synthetase proteins in *Myxococcus xanthus*. *J. Bacteriol.*, 185:5066–5075.
204. Singh, A., Hazlett, L., and Berk, R.S., 1991, Characterization of pseudomonal adherence to unwounded cornea. *Invest. Ophthalmol. Vis. Sci.*, 32:2096–2104.
205. Singh, P.K., Parsek, M.R., Greenberg, E.P., and Welsh, M.J., 2002, A component of innate immunity prevents bacterial biofilm development. *Nature*, 417:552–555.
206. Skerker, J.M., and Berg, H.C., 2001, Direct observation of extension and retraction of type IV pili. *Proc. Natl. Acad. Sci. USA*, 98:6901–6904.
207. Smart, W., Sastry, P.A., Paranchych, W., and Singh, B., 1988, Mapping of the T-cell recognition sites of *Pseudomonas aeruginosa* PAK polar pili. *Infect. Immun.*, 56:18–23.

208. Sourjik, V., and Schmitt, R., 1998, Phosphotransfer between CheA, CheY1, and CheY2 in the chemotaxis signal transduction chain of *Rhizobium meliloti*. *Biochemistry*, 37:2327–2335.
209. Speert, D.P., Loh, B.A., Cabral, D.A., and Salit, I.E., 1986, Nonopsonic phagocytosis of non-mucoid *Pseudomonas aeruginosa* by human neutrophils and monocyte-derived macrophages is correlated with bacterial piliation and hydrophobicity. *Infect. Immun*., 53:207–212.
210. Stock, A.M., Robinson, V.L., and Goudreau, P.N., 2000, Two-component signal transduction. *Annu. Rev. Biochem.*, 69:183–215.
211. Stoodley, P., Sauer, K., Davies, D.G., and Costerton, J.W., 2002, Biofilms as complex differentiated communities. *Annu. Rev. Microbiol.*, 56:187–209.
212. Stover, C.K., Pham, X.Q., Erwin, A.L., Mizoguchi, S.D., Warrener, P., Hickey, M.J., Brinkman, F.S., Hufnagle, W.O., Kowalik, D.J., Lagrou, M., Garber, R.L., Goltry, L., Tolentino, E., Westbrock-Wadman, S., Yuan, Y., Brody, L.L., Coulter, S.N., Folger, K.R., Kas, A., Larbig, K., Lim, R., Smith, K., Spencer, D., Wong, G.K., Wu, Z., Paulsen, I.T., Reizer, J., Saier, M.H., Hancock, R.E., Lory, S., and Olson, M.V., 2000, Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen. *Nature*, 406:959–964.
213. Strom, M.S., Nunn, D.N., and Lory, S., 1993, A single bifunctional enzyme, PilD, catalyzes cleavage and N-methylation of proteins belonging to the type IV pilin family. *Proc. Natl. Acad. Sci. USA*, 90:2404–2408.
214. Suh, S.J., Silo-Suh, L., Woods, D.E., Hassett, D.J., West, S.E., and Ohman, D.E., 1999, Effect of *rpoS* mutation on the stress response and expression of virulence factors in *Pseudomonas aeruginosa*. *J. Bacteriol.*, 181:3890–3897.
215. Sun, H., Zusman, D.R., and Shi, W., 2000, Type IV pilus of *Myxococcus xanthus* is a motility apparatus controlled by the *frz* chemosensory system. *Curr. Biol.*, 10:1143–1146.
216. Sundin, C., Wolfgang, M.C., Lory, S., Forsberg, A., and Frithz-Lindsten, E., 2002, Type IV pili are not specifically required for contact dependent translocation of exoenzymes by *Pseudomonas aeruginosa*. *Microb. Pathog.*, 33:265–277.
217. Tang, H., Kays, M., and Prince, A., 1995, Role of *Pseudomonas aeruginosa* pili in acute pulmonary infection. *Infect. Immun.*, 63:1278–1285.
218. Totten, P.A., Lara, J.C., and Lory, S., 1990, The *rpoN* gene product of *Pseudomonas aeruginosa* is required for expression of diverse genes, including the flagellin gene. *J. Bacteriol.*, 172:389–396.
219. Venturi, V., 2003, Control of *rpoS* transcription in *Escherichia coli* and *Pseudomonas*: why so different? *Mol. Microbiol.*, 49:1–9.
220. Vignon, G., Kohler, R., Larquet, E., Giroux, S., Prevost, M.C., Roux, P., and Pugsley, A.P., 2003, Type IV-like pili formed by the type II secreton: specificity, composition, bundling, polar localization, and surface presentation of peptides. *J. Bacteriol.*, 185:3416–3428.
221. Wall, D., and Kaiser, D., 1998, Alignment enhances the cell-to-cell transfer of pilus phenotype. *Proc. Natl. Acad. Sci. USA*, 95:3054–3058.
222. Wall, D., Wu, S.S., and Kaiser, D., 1998, Contact stimulation of Tgl and type IV pili in *Myxococcus xanthus*. *J. Bacteriol.*, 180:759–761.
223. Watson, A.A., Alm, R.A., and Mattick, J.S., 1996, Identification of a gene, *pilF*, required for type 4 fimbrial biogenesis and twitching motility in *Pseudomonas aeruginosa*. *Gene*, 180:49–56.
224. Watts, T.H., Kay, C.M., and Paranchych, W., 1983, Spectral properties of three quaternary arrangements of *Pseudomonas* pilin. *Biochemistry*, 22:3640–3646.
225. Webre, D.J., Wolanin, P.M., and Stock, J.B., 2003, Bacterial chemotaxis. *Curr. Biol.*, 13:R47–R49.
226. Weiss, R.L., 1971, The structure and occurrence of pili (fimbriae) on *Pseudomonas aeruginosa*. *J. Gen. Microbiol.*, 67:135–143.

227. Whitchurch, C.B., Hobbs, M., Livingston, S.P., Krishnapillai, V., and Mattick, J.S., 1991, Characterisation of a *Pseudomonas aeruginosa* twitching motility gene and evidence for a specialised protein export system widespread in eubacteria. *Gene*, 101:33–44.
228. Whitchurch, C.B., Alm, R.A., and Mattick, J.S., 1996, The alginate regulator AlgR and an associated sensor FimS are required for twitching motility in *Pseudomonas aeruginosa*. *Proc. Natl. Acad. Sci. USA*, 93:9839–9843.
229. Whitchurch, C.B., Beatson, S.A., Comolli, J.C., Jakobsen, T., Sargent, J.L., Bertrand, J.J., West, J., Klausen, M., Waite, L.L., Kang, P.J., Tolker-Nielsen, T., Mattick, J.S., and Engel, J.N., 2005, *Pseudomonas aeruginosa fimL* regulates multiple virulence functions by intersecting with Vfr-modulated pathways. *Mol. Microbiol.*, 55:1357–1378.
230. Whitchurch, C.B., Erova, T.E., Emery, J.A., Sargent, J.L., Harris, J.M., Semmler, A.B., Young, M.D., Mattick, J.S., and Wozniak, D.J., 2002, Phosphorylation of the *Pseudomonas aeruginosa* response regulator AlgR is essential for type IV fimbria-mediated twitching motility. *J. Bacteriol.*, 184:4544–4554.
231. Whitchurch, C.B., Leech, A.J., Young, M.D., Kennedy, D.H., Sargent, J.L., Bertrand, J.J., Semmler, A.B.T., Mellick, A.A., Martin, P.R., Alm, R.A., Hobbs, M., Beatson, S.A., Huang, B., Nguyen, L., Comolli, J.C., Engel, J.N., Darzins, A., and Mattick, J.S., 2004, Characterization of a complex chemosensory signal transduction system which controls twitching motility in *Pseudomonas aeruginosa*. *Mol. Microbiol.*, 52:873–893.
232. Whitchurch, C.B., and Mattick, J.S., 1994, Characterisation of a gene, *pilU*, required for twitching motility but not phage sensitivity in *Pseudomonas aeruginosa*. *Mol. Microbiol.*, 13:1079–1091.
233. Whitchurch, C.B., Tolker-Nielsen, T., Ragas, P.C., and Mattick, J.S., 2002, Extracellular DNA required for bacterial biofilm formation. *Science*, 295:1487.
234. Wolanin, P.M., and Stock, J.B., 2004, Bacterial chemosensing: cooperative molecular logic. *Curr. Biol.*, 14:R486–R487.
235. Wolfgang, M., van Putten, J.P., Hayes, S.F., Dorward, D., and Koomey, M., 2000, Components and dynamics of fiber formation define a ubiquitous biogenesis pathway for bacterial pili. *EMBO J.*, 19:6408–6418.
236. Wolfgang, M.C., Lee, V.T., Gilmore, M.E., and Lory, S., 2003, Coordinate regulation of bacterial virulence genes by a novel adenylate cyclase-dependent signaling pathway. *Dev. Cell*, 4:253–263.
237. Wong, W.Y., Campbell, A.P., McInnes, C., Sykes, B.D., Paranchych, W., Irvin, R.T., and Hodges, R.S., 1995, Structure-function analysis of the adherence-binding domain on the pilin of *Pseudomonas aeruginosa* strains PAK and KB7. *Biochemistry*, 34:12963–12972.
238. Wong, W.Y., Irvin, R.T., Paranchych, W., and Hodges, R.S., 1992, Antigen-antibody interactions: elucidation of the epitope and strain-specificity of a monoclonal antibody directed against the pilin protein adherence binding domain of *Pseudomonas aeruginosa* strain K. *Protein Sci.*, 1:1308–1318.
239. Woods, D.E., Bass, J.A., Johanson, W.G., Jr., and Straus, D.C., 1980, Role of adherence in the pathogenesis of *Pseudomonas aeruginosa* lung infection in cystic fibrosis patients. *Infect. Immun.*, 30:694–699.
240. Woods, D.E., Straus, D.C., Johanson, W.G., Jr., Berry, V.K., and Bass, J.A., 1980, Role of pili in adherence of *Pseudomonas aeruginosa* to mammalian buccal epithelial cells. *Infect. Immun.*, 29:1146–1151.
241. Worlitzsch, D., Tarran, R., Ulrich, M., Schwab, U., Cekici, A., Meyer, K.C., Birrer, P., Bellon, G., Berger, J., Weiss, T., Botzenhart, K., Yankaskas, J.R., Randell, S., Boucher, R.C., and Doring, G., 2002, Effects of reduced mucus oxygen concentration in airway *Pseudomonas* infections of cystic fibrosis patients. *J. Clin. Invest.*, 109:317–325.

242. Yeo, H.J., Savvides, S.N., Herr, A.B., Lanka, E., and Waksman, G., 2000, Crystal structure of the hexameric traffic ATPase of the *Helicobacter pylori* type IV secretion system. *Mol. Cell*, 6:1461–1472.
243. Yoshihara, S., Geng, X., Okamoto, S., Yura, K., Murata, T., Go, M., Ohmori, M., and Ikeuchi, M., 2001, Mutational analysis of genes involved in pilus structure, motility and transformation competency in the unicellular motile cyanobacterium *Synechocystis* sp. PCC 6803. *Plant Cell Physiol.*, 42:63–73.
244. Yu, H., Mudd, M., Boucher, J.C., Schurr, M.J., and Deretic, V., 1997, Identification of the *algZ* gene upstream of the response regulator *algR* and its participation in control of alginate production in *Pseudomonas aeruginosa*. *J. Bacteriol.*, 179:187–193.
245. Yu, L., Lee, K.K., Hodges, R.S., Paranchych, W., and Irvin, R.T., 1994, Adherence of *Pseudomonas aeruginosa* and *Candida albicans* to glycosphingolipid (Asialo-GM1) receptors is achieved by a conserved receptor-binding domain present on their adhesins. *Infect. Immun.*, 62:5213–5219.
246. Yu, L., Lee, K.K., Paranchych, W., Hodges, R.S., and Irvin, R.T., 1996, Use of synthetic peptides to confirm that the *Pseudomonas aeruginosa* PAK pilus adhesin and the *Candida albicans* fimbrial adhesin possess a homologous receptor-binding domain. *Mol. Microbiol.*, 19:1107–1116.
247. Zhao, Z., and Panjwani, N., 1995, *Pseudomonas aeruginosa* infection of the cornea and asialo GM1. *Infect. Immun.*, 63:353–355.
248. Zhulin, I.B., Taylor, B.L., and Dixon, R., 1997, PAS domain S-boxes in Archaea, Bacteria and sensors for oxygen and redox. *Trends Biochem. Sci.*, 22:331–333.
249. Zolfaghar, I., Evans, D.J., and Fleiszig, S.M., 2003, Twitching motility contributes to the role of pili in corneal infection caused by *Pseudomonas aeruginosa*. *Infect. Immun.*, 71:5389–5393.
250. Zoutman, D.E., Hulbert, W.C., Pasloske, B.L., Joffe, A.M., Volpel, K., Trebilcock, M.K., and Paranchych, W., 1991, The role of polar pili in the adherence of *Pseudomonas aeruginosa* to injured canine tracheal cells: a semiquantitative morphologic study. *Scan. Microsc.*, 5:109–126.

7

EVOLUTION OF CATABOLIC PATHWAYS IN *PSEUDOMONAS* THROUGH GENE TRANSFER

Jan Roelof van der Meer

Department of Fundamental Microbiology
Bâtiment de Biologie
University of Lausanne
1015 Lausanne, Switzerland

1. INTRODUCTION

Pseudomonads are recognized for their extraordinary capacity to catabolize a large number of different carbon compounds.[39,40] This property is evident from a relatively large number of uptake systems, *a more than median genome size*, and a large proportion of genetic material devoted to carbon compound catabolism.[120] The catabolic capacity of pseudomonads is shared with a number of other bacteria, mainly from the groups of *Gamma*-, *Beta*-, and *Alphaproteobacteria*, such as *Acinetobacter*, *Ralstonia* spp., *Burkholderia* spp., and sphingomonads, and probably many more to become analyzed in the future (e.g. *Alcanivorax*, *Rhodococcus*). Notwithstanding the catabolic capacities of individual *Pseudomonas* isolates, there has been a certain overinterpretation of the catabolic and ecological importance of the genus *Pseudomonas* as a whole, perhaps because, first of all, many original "*Pseudomonas*" isolates were later renamed to other genera. Secondly, because many pseudomonads were isolated by enrichment techniques which selected for the fastest growing microorganisms but not necessarily for the most environmentally abundant. Thirdly, because the true catabolic activity of microbes in the natural environment is very difficult to measure and can only now be beginning to be studied

Pseudomonas, Volume 4, edited by Juan-Luis Ramos and Roger C. Levesque

with the use of isotopic labeling techniques or direct *in situ* mRNA analyses. Finally, because a substantial part of the catabolic properties is part of what might be called the "mobile gene pool," to which different genera have "access," but which gives a certain volatile status, being subject to both functional gain and loss. On the other hand, whatever the status on the "true" role of pseudomonads in carbon compound catabolism in the environment may become, they have served as paradigms in many aspects.

The specific purpose of this overview is not to list and categorize all possible carbon compound catabolic pathways in pseudomonads and related "catabolic specialists," but to describe the mechanisms by which these groups of organisms (with a majority of examples from true *Pseudomonas* spp.) may have evolved a multitude of pathways and continue to expand catabolic capacities to new substrates, many of which are of synthetic nature. Many different evolutive genetic mechanisms have been clarified on a molecular level, presenting a detailed picture on the possibilities for adaptation by catabolic expansion. Pseudomonads have also here served as unique model systems for our understanding of catabolic pathway evolution. However, also in our thinking on catabolic pathway evolution we may have been biased by a choice of target organisms and catabolic pathways, and it should not be forgotten that the same sort of adaptive mechanisms are a very general feature for all bacteria.

2. CATABOLIC PATHWAYS

It is thought that the catabolic expertise of pseudomonads and other organisms is a reflection of their (natural) ecological niches as degraders of plant material,[39,63,91,93] plants being able to synthesize the most differentiated carbon structures among living beings.[187] It is not surprising then to find many potential catabolic pathways for phenolic, carboxylic, and methoxylated aromatic compounds in pseudomonads (Table 1), which must have evolved during millions of years of bacteria being exposed to (decaying) plant material. A relatively good idea of these "core" catabolic processes in *Pseudomonas* (in contrast to the "mobile and expandable parts") can nowadays be obtained from full genome sequences such as those recently finished of *P. putida* KT2440,[91] *P. syringae* or *P. aeruginosa*.[199] It is perhaps good to spend a few words on the terminology of catabolic pathways. In a more general sense catabolic pathways describe all those metabolic processes which lead to a breakdown of carbon compounds to a level at which they can be used as intermediates for building new cellular macromolecules or are used in redox processes for energy generation. For pseudomonads, catabolic pathways usually "end" at the point of pyruvate, acetyl-CoA, or as direct intermediates of the citric acid or tricarboxylic acid – TCA cycle (e.g. succinyl-CoA or fumarate). Pyruvate and acetate are transformed to

Table 1. Non-exhaustive overview of peripheral catabolic pathways in pseudomonads.

Main substrate(s)	Key intermediates or intermediate pathway	Strain(s)	Genetic nomenclature and gene location	Reference(s)
Styrene	Phenylacetate, phenylacetyl-CoA	*Pseudomonas* sp. Y2, *P. putida* CA-3, *P. fluorescens* ST, *Pseudomonas* sp. VLB120	*sty, paa1, paa2*, chromosome	13,133
Phenylacetate	Phenylacetyl-CoA	*Pseudomonas* sp. Y2, *P. putida* KT2440	*paa2, pad, pha*, chromosome	2,91
L-Phenylalanine, L-tyrosine, 3-OH-phenylacetate	Homogentisate	*P. putida* U, *P. putida* KT2440	*phh, tyr, hpd, hmgABC, fah*, chromosome	5,91
Quinoline, quinate	Protocatechuate, 3-oxoadipate	*P. putida* sp. strain 86, *P. putida* KT2440	*oxo, qorMSL*, chromosome	31,91
Ferulate, caffeate, vanillate, vanillin, eugenol, coniferyl alcohol, *p*-coumate, *p*-hydroxy-benzoate	Protocatechuate, 3-oxoadipate	*P. putida* KT2440, *P. putida* PRS2000, *Pseudomonas* sp. strain HR199	*cal, fcs, ech, vdh, van, pob, pca*, chromosome	79,91,136
Phenylhexanoate, phenylhep-tanoate phenyloc-tanoate	Phenylacetyl-CoA	*P. putida* KT2440	*fad, pha*, chromosome	91
Benzylamine, benzoate	Catechol, 3-oxoadipate	*P. putida* KT2440	*ben, catBCA*, chromosome	91
Phenylbenzoate	Benzoate plus phenol, catechol, 3-oxoadipate	*Pseudomonas* sp. strain TR3		159
Phenoxyacetic acid, phenol	Catechol, 3-oxoadipate	*Pseudomonas* sp. strain B13	*tfdA-S* from pJP4 introduced on plasmid pKJS32	46
Toluene, benzene, ethylbenzene	(Methyl-)catechol, *meta* cleavage	*P. putida* F1	*tod*, chromosome	248
Toluene, *o*-xylene	Methylcatechol, *meta* cleavage	*P. stutzeri* OX1	Chromosome	12

(*Continued*)

Table 1. (*Continued*)

Main substrate(s)	Key intermediates or intermediate pathway	Strain(s)	Genetic nomenclature and gene location	Reference(s)
Toluene, *m*- and *p*-xylene, benzylalcohols, benzylaldehyde, pseudocumene	Benzoate, catechol, *meta* cleavage	*P. putida* mt-2	*xyl*, pWW0	60,236
Toluene	*p*-Cresol, protocatechuate, 3-oxoadipate	*P. mendocina* KR1	*tmo*, chromosome	244
Benzene	Catechol, 3-oxoadipate	*P. putida* ML2	*bed*, plasmid pHMT2	203
Toluene, ethylbenzene, isopropylbenzene	(Alkyl)catechol, *meta* cleavage	*P. putida* o1G3	*ebd*, chromosome	33
Isopropylbenzene	(Alkyl)catechol, *meta* cleavage	*P. putida* RE204	*ipb*, Plasmid pRE4	55
Naphthalene	Salicylate, catechol, *meta* cleavage	*P. putida* NCIB 9816-4, *P. putida* PpG7, *Pseudomonas* sp. strain ND6	*nah,* pDTG1, NAH7, pND6-1	47,111,245
1-Methyl, 2-methylnaphthalene, benzylalcohol (not toluene or xylene)	Benzoate, catechol, *ortho* cleavage, naphthoic acids (and no further)	*P. putida* CSV86	Plasmid likely	14
Naphthalene	Salicylate, catechol, *meta* cleavage	*P. stutzeri* AN10	*nah*, chromosome	23
Benzene, toluene, *m*-, *p*-xylene, ethylbenzene, biphenyl	*Meta* cleavage	*P. putida* BP18		11
p-Cymene, cumene, *n*-propylbenzene, *n*-butylbenzene, biphenyl	*p*-Cumate, *meta* cleavage	*P. putida* F1 mutants	Mutation in *cymR, cmtE,* chromosome	37
Naphthalene, phenanthrene, anthracene	Salicylate	*P. aeruginosa* 57	*pah/nah/dox*-like, No plasmid	42
Carbazole	Anthranilate, catechol, *meta* cleavage	*P. stutzeri* OM1, *P. resinovorans* CA10	*ant*, *car,* pCAR1 plasmid	113,135

Table 1. (*Continued*)

Main substrate(s)	Key intermediates or intermediate pathway	Strain(s)	Genetic nomenclature and gene location	Reference(s)
2-Hydroxy-biphenyl, 2,2′-dihydroxy-biphenyl	Benzoate (salicylate), catechol, *meta* cleavage	*P. azelaica* HBP1	*hbp*, no plasmid	101
Biphenyl, 3-, 4-methylbiphenyl	(Methyl)benzoate	*Pseudomonas* strain IC	*bph*, chromosome	32
Biphenyl	Benzoate	*P. pseudoalcaligenes* KF707	*bph*	201
Biphenyl, salicylate	Benzoate, catechol, *meta* cleavage	*P. putida* KF715	*bph*, chromosome, Bph-Sal element	80,126
Dibenzothiophene, naphthalene	Salicylate	*Pseudomonas* sp. strain C18	*dox,* plasmid-localized	48
Dibenzofuran, dibenzo-*p*-dioxin	Salicylate, catechol	*P. veronii* PH-03	Unknown	84
Naphthalene, phenanthrene, anthracene	Salicylate	*Pseudomonas* sp. strains 5R, DFC49, DFC50	pKA plasmids	170
Phenol (transformation only of 2,4- and 2,5-dimethylphenol)	Catechol, *meta* cleavage	*P. stutzeri* OX1	*tou*, chromosome	4
3-Alkyl, 4-alkylphenol	Catechol, *meta* cleavage	*Pseudomonas* sp. strain KL28	*lap*, no plasmid	90
Phenol, 4-methylphenol	Catechol, *meta* cleavage	*Pseudomonas* sp. CF600	*dmp*, pVI150 plasmid	127,151
3,4-, 2,4-dimethylphenol	Catechol, *meta* cleavage	*P. mendocina* PC1		231
3,4-, 2,4-dimethylphenol	Protocatechuate, *ortho*	*P. fluorescens* PC18, PC24		231
Nitrobenzene	2-Aminophenol, *meta* cleavage	*P. putida* HS12	*nbz,* pNB1, pNB2	138
o-Nitrophenol	Catechol, nitrite	*P. putida* B2		246
4-Nitrotoluene, 4-nitrobenzoate	Benzoate, catechol, *meta* cleavage	*Pseudomonas* sp. strain TW3	*ntn, pnb*, chromosome	88
Nitrobenzene	2-Aminophenol, *meta* cleavage	*P. pseudoalcaligenes* JS45	Chromosome	44,188
Aromatic amines, *m*-toluate, aniline	Catechols, *meta* cleavage	*P. putida* UCC22	pTDN1 plasmid	168

(*Continued*)

Table 1. (*Continued*)

Main substrate(s)	Key intermediates or intermediate pathway	Strain(s)	Genetic nomenclature and gene location	Reference(s)
4-Chlorobenzoate	Protocatechuate, 3-oxoadipate	*Pseudomonas* sp. strain CBS3	Chromosome	172
2-Chlorobenzoate, 2,4-dichloro-benzoate	Chlorocatechol, 3-oxoadipate	*P. aeruginosa* 142	*ohb, clc*, chromosome	41,212
Toluene, chlorobenzene	3-methyl, 3-chlorocatechol, *meta* cleavage	*P. putida* GJ31	*cbz*, plasmid	116,117
Toluene, chlorobenzene	Methyl-, chlorocatechol	*Pseudomonas* sp. strain JS6		143
Monochloro, 1,2-, 1,4-dichloro, 1,2,4-trichloro-benzene	Chlorocatechol, 3-oxoadipate	*Pseudomonas* sp. strain P51	*tcb*, plasmid pP51	227
4-Chlorophenol, 3-chlorobenzoate	Chlorocatechol, 3-oxoadipate	*Pseudomonas* sp. strain B13	*clc*, mobile genomic island	52,154
1,2,3,4-Tetrachloro-benzene	Tetrachloro-catechol, chloro-oxoadipate	*P. chlororaphis* RW71	*tet*, IS1071	149,150
p-Cymene, *p*-cumate	*p*-Cumate	*P. putida* F1	*cym, cmt*, chromosome	54
Abietane diterpenoids		*P. abietaniphila* BKME-9	*dit*	118
Dodecylbenzene sulfonate, citronellol		*P. aeruginosa* W51D		29
C5–C12 alkanes, toluene, naphthalene		*Pseudomonas* sp.	Psychrophilic, OCT-like and NAH-like plasmid coexisting	234
C12–C16 alkanes		*P. fluorescens* CHA0		186
C6–C12 *n*-alkanes		*P. putida* Gpo1	*alk*, OCT	57,220
1,3-Dichloro-propene, 1,2-dibromoethane		*P. pavonaceae* 170, *Pseudomonas* sp. GJ1	*deh*	147
Atrazine	Sole N-source, to cyanuric acid	*Pseudomonas* sp. strain ADP	*atz*, pADP1	119

acetyl-CoA to feed the TCA cycle and lead to production of reduced coenzymes ($NADH_2$, $NADPH_2$, $FADH_2$), anabolic building blocks, and ATP.

Often, however, the terminology "catabolic pathways" is used for a more limited capacity, such as producing catechol from benzoate, in which case also the term "peripheral metabolism" is used.[91,120] Catechol is considered in this case as a key intermediate, feeding into an "intermediate" pathway, because it occurs as the "end point" of a number of different peripheral pathways (e.g. benzoate, benzene, toluene, salicylate, or naphthalene). Several such key intermediates have been defined, which on their turn can be seen as "nodes" of pathway convergence, "funneling" into the TCA cycle (scientific literature offers various metaphores for this concept). Pseudomonads like *P. putida* KT2440 carry four of these intermediate pathways: protocatechuate and catechol *ortho* cleavage (both converge via 3-oxoadipate to succinyl- and acetyl-CoA), the homogentisate (to fumarate and acetoacetate), and the phenylacetyl-CoA pathway.[91,120] This last pathway does not seem to be present in *P. aeruginosa* PAO1 and *P. syringae*. Similar intermediate pathways are present among other *Proteobacteria* (e.g. *Ralstonia*, *P. aeruginosa*, *Acinetobacter*), but also in other taxonomic groups (e.g. *Rhodococcus*).[91,120] Despite analogy of function, the gene orders for the intermediate pathways in different microorganisms are not very conserved and resemble mosaics, suggesting that there is no evolutionary constraint on gene order.[91] Various other intermediate pathways exist as well, such as catechol via *meta* cleavage, chlorocatechol to 3-oxoadipate, hydroquinone or gentisate, which do not form part of the "core" in e.g. *P. putida* KT2440 but are present in other pseudomonads or are part of the "mobile" genome (Table 1). Therefore, we must conclude that although the concept of "funneling" is useful to comprehend catabolic processes, there is not one specific intermediate pathway which can be considered to be an absolute requirement for the core cellular catabolism. We will see later how intermediate pathways are or can become part of the mobile gene pool.

In terms of metabolic strategies for microorganisms, the analogy of a "tree" is often used, with the stem being TCA cycle or the like, large limbs being the intermediate pathways, and several smaller twigs being peripheral pathways feeding into larger limbs. However, like any analogy this has a certain danger of not exposing the overall variability and redundancy of the system. Consider, for example, enzymes with the capacity to convert more than one substrate and feeding into different "pathways," or incomplete pathways, consider enzyme paralogs in the same cell (e.g. two catechol 1,2-dioxygenases in *P. putida* KT2440,[120] or the many other uncharacterized potential catabolic functions (e.g. 15 predicted monooxygenases in *P. putida* KT2440 with unknown specificity).

Rather than "trees" or "funnels," one might consider operonic organizations or regulons as being the structural "entities." This has some advantage in understanding hierarchical regulatory and control processes in the cell, but

again lacks the flexibility and redundancy of genetic systems for adaptation and evolutionary development. For example, although some organisms have formed "superoperonic" organization of catabolic gene functions (e.g. *Acinetobacter* sp. ADP1 *qui*, *cat*, *ben*, *pca* operons[185]), others do very well without (e.g. *Sphingomonas paucimobilis lin* genes[50]), which suggests that there is no obligatory evolutionary tendency to form large operons in order to become an "effective" catabolic microorganism. In most genomes, redundant and uncharacterized gene functions exist, which may have a silent role under most growth conditions but become more pronounced under specific niche adaptations.

3. MICROORGANISMS AND CATABOLIC ADAPTATION

Although we can speculate about the natural niche of microbes and the reasons for the evolutionary processes which have led to present day's genomes, it has become apparent that the natural ecological niche of many microbes probably no longer is the same in the age of large-scale agricultural, industrial, livestock, and waste management activities. Pollutants (e.g. industrial solvents or agricultural pesticides and herbicides) are unlike plant material; their carbon structures or substituents may be unlike anything found under natural conditions. Hence, we might expect that human activities have selected for bacteria with new characteristics, including new catabolic properties.

A first glance at Table 1 shows that in essence this expectation has been correct (as we will see below); many pseudomonads catabolize compounds which are hard to be considered "natural" (Table 1). Although it will always remain disputable whether a compound is a naturally occurring or fully synthetic one,[93,174,209,222] there are many different aspects to the selective pressure achieved by environmental pollutants, waste, and industrial activities. Consider, for example, modern wastewater treatment installations, which are large-scale reactors cultivating huge quantities of bacterial biomass on a variety of carbon substrates present in industrial or household wastewater. These bacteria certainly were not exposed to such enormous substrate quantities before or to the general process regime of the wastewater treatment plant. Due to the huge biomass densities and presence of taxonomically very different microbial groups lateral genetic interactions may have become favored which otherwise occurred perhaps only infrequently. Hence, many new niches and processes must have favored adaptation and differentiation of microorganisms, an aspect in wastewater treatment plants which still remains relatively unexplored. In other areas, such as agricultural soils or industrial dumpsites, the release of

toxic compounds (e.g. heavy metals, antimicrobial compounds, toxic solvents, or agricultural pesticides) must have eliminated sensitive microbial populations and favored resistances, thus leading to community shifts. Finally, on the level of carbon substrates the selective force of waste products and pollutants is a very strong one, since any cell metabolizing a specific carbon substrate will immediately be able to grow and divide (if not otherwise limited in nutrients) and increase its population size proportional to those who cannot gain carbon or energy from the compound. Hence, it becomes easy to envision that when chemicals are released into the environment, which naturally only occurred in trace quantities or in restricted locations (e.g. oil), they will rapidly select for faster growing microbial populations capable of using the compounds for carbon and energy metabolism. Likewise, it can be foreseen that release of carbon compounds for which no known metabolic pathways are existing, may lead to selection of any mutants with the right catabolic properties on the basis of enhanced growth rate. Bacteria with specific catabolic properties isolated from such selective environments have undergone a natural selection, and have been exposed to the full genetic potential of all the organisms present in that environment. They can be examined in order to learn how adaptation processes and catabolic pathway evolution have "worked."[93,196,208,221]

4. GENETIC ADAPTATION MECHANISMS

In order to understand genetic adaptation mechanisms we can combine information gathered from numerous different experiments. In the first place, there is what I would call "comparative historic" evidence. By comparing genetic structures (individual genes, operons, clusters, or genomes) from bacterial isolates it is easy to establish sequence relationships, find analogies in genetic organization and with the help of statistical tools, trace nucleotide abberations in genome sequences, which may point to recombined or laterally acquired DNAs. This work is like doing molecular archeology and it can help to find traces of mechanisms that must have been taking place for millions of years. More recent evolutionary events can be traced when gene sequences from bacteria thought to have been selected for a completely new catabolic property are examined and compared. In some of those cases (see below) we can get a clear idea of the genealogy, the different steps in the formation of the present day's gene and pathway structure, and even to find DNA-donating partners in the same microbial community.

Secondly, there is laboratory evidence for evolutive mechanisms on single pure cultures or simple combinations of two or more bacterial strains. Pure strains can be grown under different conditions, and due to the relatively short generation time of bacteria, mutants will develop which can be examined for genetic changes. Mutants can also be generated in specific recombinatory

pathways, which can be subsequently studied to determine the possible role of such recombinations for genetic adaptation. Insertions elements, integrons, or transposons can be examined for their activity and the mechanisms of regulation. Two or more bacterial strains can be grown together under specific conditions to observe whether offspring develops which has exchanged parts of their genetic material, or to determine when plasmids conjugate and by which mechanisms.

What can we conclude from the different lines of evidence? In a nutshell, we must assume that all microorganisms have in some way or another an evolutionary toolbox,[3] which is responsible for generating a certain genetic flexibility inside a single cell, creating mutations that can subsequently be passed on to daughter cells. In addition, all microorganisms in a microbial community have to some extent access to the genetic information of each other through the action of lateral gene transfer mechanisms that seem to operate in a selfish independent way.[24,129]

How does the content of the toolbox look like? The genomes of probably all bacteria can be subject to small-scale changes (single to a few nucleotides), which are the consequences of DNA replication errors,[53] or are caused by mismatch repair or nucleotide excision systems,[56] perhaps even by DNA topoisomerase[242] or DNA gyrase activity.[157] Despite being small, the effects of such changes can be very significant and comprise things like complete loss of gene function (e.g. frameshift mutations), alteration of enzyme or protein specificity, or altered gene expression level.[222] Many of these genetic processes and their effects have been carefully documented for several bacterial species, including pseudomonads. They are not specific for catabolic adaptation, although some nice examples exist in which catabolic enzymes acquired a change in substrate specificity by small "spontaneous" mutations and subsequently allowed the host bacterium to catabolize a new carbon compound.[37,65,152] It further has become clear that even point mutation generating processes are not completely random, but may exhibit preferences for certain DNA sequences, or DNA structures, and the mutation rates themselves may be prone to strong variations.[242] A well documented example of this are the hypermutator strains, in which the proofreading capacity of DNA polymerase or the mismatch repair system itself are mutated, leading to strongly elevated mutation rates.[58,108] Interestingly, such strains are not *per se* less viable than strains with a normal proofreading activity.[108] In addition, it has been established that starving cells in stationary phase may exhibit elevated mutation rates.[171]

Other genetic mechanisms operate at a larger scale, resulting in DNA duplications, deletions, inversions, recombinations, insertions, and the like. These effects are generally thought to be caused by RecA-dependent recombinatory mechanisms[184] or site-specific recombinases, but can also occur through illegitimate recombination (i.e. between regions without noticeable identity), which might involve the activity of DNA topoisomerases or gyrase.[157,121,200]

From fascinating research on the highly recombinatory strains *P. stutzeri* and *Acinetobacter* sp. strain ADP1 it has become clear that DNA recombination can even take place with only four identical base pairs between two targets (albeit at low frequencies) and can also give rise to large intrachromosomal duplications.[65,121,157] In addition to general recombination activities, many different site-specific recombinatory systems may operate in the bacterial cell, such as encoded by insertion (IS) elements and transposons,[114,115] integrons or integrases. The cut-and-paste mechanisms of insertion elements may lead to deleting, recombining, and inserting intervening or flanking DNA. Although this again is not something specific for catabolic pathways, a number of good examples exist in which insertion elements are flanking catabolic genes,[202,222,243] transposing parts of or complete catabolic pathway gene clusters (e.g. the *xyl* genes on the TOL plasmid pWW0) or have been assumed to be responsible for the acquisition of parts of the catabolic genes (see below and Figure 1). Quite astonishingly, a number of extremely promiscuous IS elements exist with absolute sequence conservation (Table 2). Most notably at this moment are IS*6100* and IS*1071*, two IS elements associated to a number of catabolic gene clusters, which are very widespread among completely different bacterial taxa, while remaining 100% sequence identity. This in itself already demonstrates that IS elements are being transferred between different bacterial genera and are extremely resistant against mutational drift. Furthermore, IS elements may carry promoter sequences,[96] provide regulatory signals, and thus determine activation or deactivation of silent genes.[22,47,64,96,110] A very peculiar example in this case is found in *P. stutzeri* OX1.[22] This bacterium can use toluene and *o*-xylene as sole carbon and energy source, but in addition has a silent catabolic pathway for *m*- and *p*-xylene, which is blocked by the insertion of the IS element IS*Ps1*. Spontaneous *m*- and *p*-xylene-degrading mutants develop which at the same time have lost the ability to utilize *o*-xylene. In these mutants the IS element has excised from its position upstream of the *xylC* gene and integrated into the genes for toluene/*o*-xylene monooxygenase (ToMO).[12,22]

Increasingly, also integrases from prophages and from genomic islands are implicated in the gain and loss of function through excision or integration of large genomic regions. Most bacterial genomes contain a dozen or so of predicted integrase genes[24,28,49] and very often detectable traces of large-scale inversions or insertions of "foreign" material in their genomes (see below).[68] Integrases from integrons can catalyze an interesting inverse process of capturing specific circular DNA molecules at specific target sites located downstream of the integrase,[71] forming under certain cases cassette-like structures known as superintegrons.[167]

Apart from mechanisms operating within one cell, spectacular actions are performed by lateral or horizontal gene transfer processes, in which more than one partner contribute to the outcome. Several microorganisms take up DNA molecules spontaneously from the extracellular environment during parts

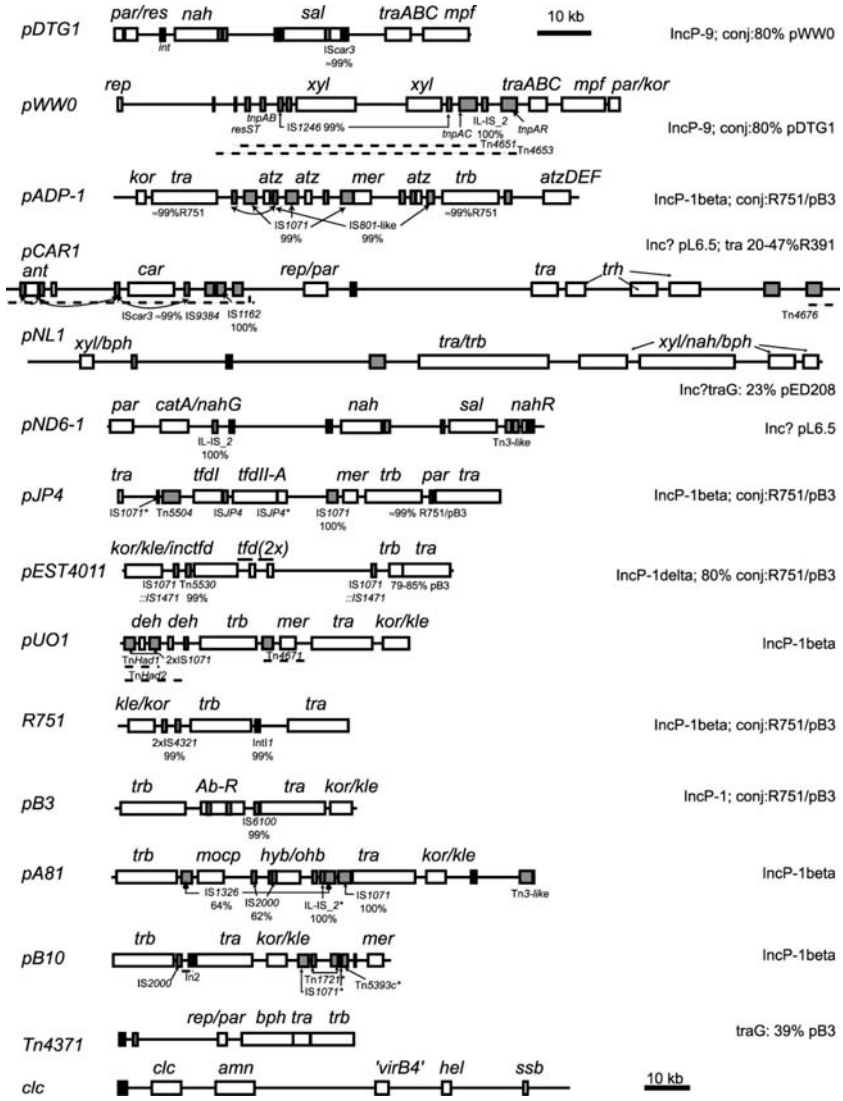


Figure 1. Schematic structures of completely sequenced "catabolic" plasmids and other mobile genetic elements. Plasmids are drawn in linear fashion as a thin black line. Regions of interest for catabolic gene functions, plasmid maintenance, conjugative transfer, and for mobile elements are schematically indicated as boxes, drawn approximately to scale. Insertion sequences are shaded gray, genes for integrases are black. The incompatibility group and homologies of conjugative systems (as % amino acid identity) are indicated on the right, if known. Transposable structures are depicted below as dotted lines. References and gene assignments: pDTG1 (Genbank accession number AF491307), *P. putida* NCIB 9816-4,[47] *par*/*res*, plasmid partitioning and resolution; *nah*, naphthalene degradation upper pathway; *sal*, salicylate degradation; *tra*, transfer and nicking;

of their normal growth cycle[137] and can efficiently incorporate (parts of) this DNA in their own genome. Among pseudomonads, *P. stutzeri* is the only known member to have this potential.[112] On the other hand, "inducible" transformation by electrical current from lightning has been implicated in DNA uptake by non-spontaneously transformable bacteria.[112] Phage transduction is a further mechanism of potential DNA transfer to new recipients, and most genomes contain prophages or parts of prophages which potentially can be transduced, cotransduced, or at some point in their history were *transducable*. There is a current revival of phage genomics and its contribution to shaping genome structure.[30] However, the role of phages in catabolic evolution *per se* is not well described.

Particularly pronounced in catabolic pathway evolution are conjugative systems, which lead to the duplication and mobilization of genes from one cell to another. In the first place, conjugation may be mediated by plasmids. Indeed, a large number of conjugative plasmids carrying catabolic genes have been described[173,237] (Table 1 and Figure 1). Conjugative plasmids are ideal vehicles for acquiring and collecting DNA from a microbial community, and they do not only carry catabolic genes, but equally frequently antibiotic resistance genes, heavy metal resistances and often additional unknown functions.[95] Plasmids probably play such an important evolutive role, first of all because they are extrachromosomal self-replicating DNA molecules and therefore possess a high level of flexibility compared to chromosomes by allowing extensive DNA rearrangements without disturbing cell integrity. Secondly, plasmids are intrinsically prone to DNA recombinations due to frequent carriage of IS elements or transposons (Figure 1). Thirdly and most importantly, large plasmids are self-transmissible and can often mobilize co-resident plasmid replicons

←

mpf, mating pair formation. Plasmid pWW0 (AJ344068) from *P. putida* mt-2,[67] *xyl*, genes for xylene degradation. Plasmid pADP-1 (U66917) from *Pseudomonas* sp. strain ADP1,[119] *atz*, atrazine catabolism; *mer*, mercury resistance; *trb*, type IV secretion system. *P. resinovorans* CA10[113] plasmid pCAR1 (AB088420), *ant*, genes for anthranilate degradation; *car*, genes for carbazole to anthranilate conversion; *trh*, conjugative transfer system. *Sphingomonas aromaticivorans*[165] plasmid pNL1 (NC_002033), *bph*, genes involved in biphenyl degradation. *Pseudomonas* sp. strain ND6[111] plasmid pND6-1 (AY208917), *cat*, genes homologous to those encoding catechol degradation via *ortho* cleavage. *Cupriavidus necator (Ralstonia eutropha*) JMP134[211] plasmid pJP4 (AY365053), *tfd*, genes involved in degradation of 2,4-dichlorophenoxyacetic acid. *A. xylosoxidans*[230] plasmid pEST4011 (AY540995). *Delftia acidovorans* strain B[189] plasmid pUO1 (AB063332), *deh*, genes for dehaloacetate dehalogenase. IncP-1beta plasmids R751,[205] pB3[82] and pB10[175] (AJ564903), Ab-R, genes for antibiotic resistances. *Achromobacter xylosoxidans* (Jencova et al., unpublished results) plasmid pA81 (AJ515144), *mocp*, chlorocatechol degradation; *hyb*, salicylate hydroxylase gene; *ohb*, ortho-halobenzoate 1,2-dioxygenase. *Ralstonia oxalatica* A5[210] biphenyl transposon Tn*4371* (AJ536756). *Pseudomonas* sp. strain B13 (Gaillard et al., unpublished results) *clc* element (AJ617740), *clc*, genes for chlorocatechol degradation, *amn*, genes for 2-aminophenol degradation, "*virB4*," gene with homology to *virB4*, *hel*, putative DNA helicase, *ssb*, putative single-stranded DNA binding protein.

Table 2. Occurrence of IS*1071* (A) and IS*6100* (B) in different bacteria.

(A) Organism[a]	Plasmid or gene cluster	Number of copies
Delftia acidovorans	pUO1	2 copies
Pseudomonas sp.	pADP1	3 copies
Arthrobacter aurescens		2 copies
Cupriavidus necator (*Ralstonia eutropha*) JMP134	pJP4	1 copy
Comamonas testosteroni	pTSA	2 copies
C. testosteroni BR60	pBR60	2 copies
Delftia acidovorans	*pdeA* gene cluster	1 copy
Delftia acidovorans	*tfd* genes	1 copy (partial)
Ralstonia metallidurens CH34	pMOL28	1 copy (partial)
Achromobacter xylooxidans	pEST4011	2 copies
Uncultured organism	pB10	1 copy, 1 partial
Pseudomonas putida	pTDN1	1 copy, 1 partial
Burkholderia cepacia	*mdc* cluster	1 copy
Pseudomonas pavonaceae	*deh* genes	1 copy
(B) Distribution of IS*6100*[b]		
Mycobacterium fortuitum	Tn*610*	
Pseudomonas aeruginosa	Plasmid R1033	
Xanthomonas campestris		
Acinetobacter baumanni	Integron	
Klebsiella pneumoniae	Integron	
Sphingomonas paucimobilis		≈10 copies
Corynebacterium diphteria		2 copies
Shigella flexneri		
Agrobacterium tumefaciens		
Serratia marcescens	R478 plasmid	
Sphingomonas herbicidovorans		
Arthrobacter sp.	pOAD2 plasmid,	5 copies
Klebsiella oxytoca	pACM1 plasmid	
Corynebacterium glutamicum	pTET3 plasmid,	3 copies
Salmonella enteritica		2 copies
Pseudaminobacter		
Ochrobactrum sp. strain mp-3		
Escherichia coli	pHSH	
Arcanobacter pyogenes	pAP2	
Aeromonas salmonicida	pRAS1	

[a] As per Genbank screening d.d. January 2005.
[b] As per Genbank screening d.d. July 2004.

and even plasmid replicons in a contacted cell (*retrotransfer*).[206] By doing so they transfer, pickup, and distribute any newly formed catabolic operons (or any other captured DNA) throughout microbial communities. The promiscuous and selfish nature of plasmids is demonstrated by their ability to influence their host to cope and survive in hostile environments[95] and to maintain their own life style. In this respect plasmid-located operons for catabolism of alternative carbon sources (e.g. xenobiotic compounds) give a similar advantage to their bacterial host. The "selfishness" of plasmids is clear from such functions as copy number control, a system for multimer resolution, partitioning, post-segregational killing, conjugative transfer, and mobilization to guarantee their own maintenance and distribution. The genes for those functions are organized in the form of inter-regulated and often juxtaposed operons collectively termed the plasmid backbone.[204] There is now growing evidence that specific chromosomal regions may also be mobilizable or self-transferable, or iterate between a plasmid replicative and a chromosomally integrated state (a class of elements known as integrative and conjugative elements, ICE), which means that some of the exclusivity and flexibility which previously had been reserved for conjugative plasmids is also available to chromosomal regions.[28,49,226]

For a number of catabolic pathways, the complete sequence of the "vehicle" carrying the pathway genes has been determined (Figure 1). Among those we now find conjugative and non-conjugative plasmids and two ICE, but this list will surely expand in the next few years. Some details of these mobile catabolic entities will be presented further below, but an overall picture emerges of more or less independent vehicles which can acquire accessory genes in regions of the "backbone" which are flexible enough to withstand insertions. Quite a number of them belong to the IncP1-beta subgroup (e.g. pJP4, pUO1, R751, pA81, pB10, pADP1), which has allowed detailed hypotheses on how the current plasmid structures evolved by insertions into an ancestor core IncP1-beta plasmid (e.g. R751 or pB10). We can also conclude that there is no big difference for vehicles carrying catabolic genes, antibiotic resistance genes, or virulence functions. This became again obvious from the IncP1-beta group,[175] which has members of both kinds (catabolic and resistance), and from ICE such as the *clc* element, which have large regions of high homology to pathogenicity islands (see further below).[49,226]

5. EXPANSION OF CATABOLIC PATHWAYS BY LATERAL GENE TRANSFER

Many ideas on catabolic gene evolution started with the dissection of (now) typical degradation pathways such as those for octane, xylene, toluene, naphthalene, or biphenyl. It became obvious very early on that the properties to

degrade octane, xylene, toluene, or naphthalene were often located on plasmid DNA, which could be transferred from the organisms degrading the compound (donor cell) to other recipient bacteria.[9] Numerous other plasmids have been described in the literature carrying genes for catabolic pathways, both fragments or individual complementary genes to a pathway, or clusters for complete "peripheral plus intermediate" pathways (e.g. toluene or naphthalene). Many of these plasmids were shown to be self-transmissible in the laboratory, or could be mobilized by other conjugative systems.

Transfer to a new recipient can result in an active degradation pathway in the recipient cell, but this must not be necessarily so. In first instance, selection for transconjugants in laboratory conjugation experiments had always been for the property to use the carbon compound for which the catabolic enzymes were encoded on the plasmid. But in other approaches in which other selectable (i.e. antibiotic resistances) or non-selectable markers (i.e. GFP) were introduced on the catabolic plasmids,[38] it became apparent that transfer is also taking place to hosts without pathway selection.[43]

There is now ample evidence to prove that plasmid transfer is not a laboratory artifact, but is actually taking place in natural environments. On several occasions, plasmids migrated from deliberately introduced donating strains to recipient cells in the same microbial community. This had been clearly shown in experiments in which researchers attempted to accelerate the degradation of 2,4-dichlorophenoxyacetic acid (2,4-D) in soils by inoculating the bacterium *Ralstonia eutropha* JMP134 (pJP4), and the pJP4 plasmid turned up among the native microbial cells.[45,125] In other experimental systems, it could be demonstrated that the same plasmid with the genes for naphthalene degradation (pDTG1) showed up in several different host bacteria when the environment was contaminated with polycyclic aromatic hydrocarbons.[81,241] A very peculiar situation occurred in which a plasmid from a deliberately released phenol-degrading *P. putida* could be reisolated after a number of years from bacteria in the same environment, although the original donor strain could not be retrieved.[142]

6. TOLUENE DEGRADATION AND THE TOL PLASMIDS

Although we cannot argue that toluene is a novel synthetic compound, exposure to toluene (and benzene, xylenes, and ethylbenzene; collectively called the BTEX compounds) is one of the most frequent human-caused pollution disasters.[197,198] Therefore, we have to assume that BTEX spills nevertheless disturbed many microbial communities and perhaps selected for bacteria with more efficient BTEX metabolic properties. Toluene degradation is very common among pseudomonads and at least five different pathways are known for aerobic degradation. The classical TOL plasmid (pWW0) harbors a fully functional

set of genes for the degradation of *m*-, *p*-xylene, and toluene via (*meta*-) toluate and (methyl-) catechol to pyruvate and acetaldehyde, including a sophisticated regulatory system for pathway expression.[9,237] In our definition from above, this means that pWW0 provides not just "peripheral" catabolism, but a complete catabolic pathway, since the end products directly enter at the level of pyruvate and acetyl-CoA, thus feeding into the TCA cycle. The intermediate catabolic pathway is known as the *meta* cleavage pathway,[76] because of the action of a catechol 2,3-dioxygenase that cleaves (methyl-) catechol at the *meta* position with respect to the two hydroxyl groups. At closer look, even the *meta* cleavage pathway for catechol degradation itself is a redundant intermediate pathway, since two branches (e.g. (i) 2-hydroxymuconic-semialdehyde dehydrogenase, 4-oxalocrotonate isomerase, and 4-oxalocrotonate decarboxylase and (ii) 2-hydroxymuconic-semialdehyde hydrolase) produce the same downstream intermediate (2-oxopentenoate).[76] The genetic organization for and regulation of the toluene and xylene catabolic pathway of TOL plasmid pWW0[236] has been studied in great detail. The genes for toluene and xylene degradation (the *xyl* genes) on pWW0 are clustered in two operons, named the *upper* and *meta* pathway operon according to their encoded functions[73–76,238] (Figure 2). Both operons are separated by some 14-kbp DNA.[60] Expression of the structural *xyl* genes on pWW0 is controlled by two positively acting regulatory proteins encoded by the genes *xylR* and *xylS*,[59,87] located adjacent to the 3′ end of the *meta* pathway operon and transcribed divergently.[192] All the other known TOL plasmids also carry the two operon types mentioned above with remarkably conserved gene order and with *xylS* and *xylR* regulatory genes, although the number of operons and regulatory genes may vary (see below).

The pWW0 plasmid with the *xyl* gene clusters contains various (potential) transposable and insertion elements. First of all, the *xyl* operons are flanked by two identical insertion sequences in direct orientation and 39 kb apart, named IS*1246*.[158] Although their activity has not been demonstrated experimentally, the two copies of IS*1246* have been associated with duplications[181] and deletions of the intervening 39-kb DNA segment.[15,89,181,239] The 39-kb region is enclosed within a 56-kb large transposon Tn*4651*. Tn*4651* itself is part of an even larger (70-kb) sized transposon Tn*4653*. Both Tn*4651* and Tn*4653* are functional class II transposons of the Tn*3* family.[214–216] On the related TOL plasmid pWW53 a 39-kb transposon (Tn*4656*) is found.[213] Its transposase TnpA and TnpR resolvase are 95% and 87%, respectively, identical and functionally interchangeable with those of Tn*4653*.[213]

Both Tn*4651* and Tn*4653* were shown to transpose into various plasmid replicons[34,89,107] and to the bacterial chromosome.[89,181] *P. putida* strain MW1000 even carries a Tn*4651*-like TOL transposon in its chromosome.[182] Transposition functions are assumed to be present on pWW15 and the related TOL plasmids pWW14 and pWW74 as well.[35] Two other IS elements are present on pWW0, one of which is 100% identical to the IL-IS_2 element

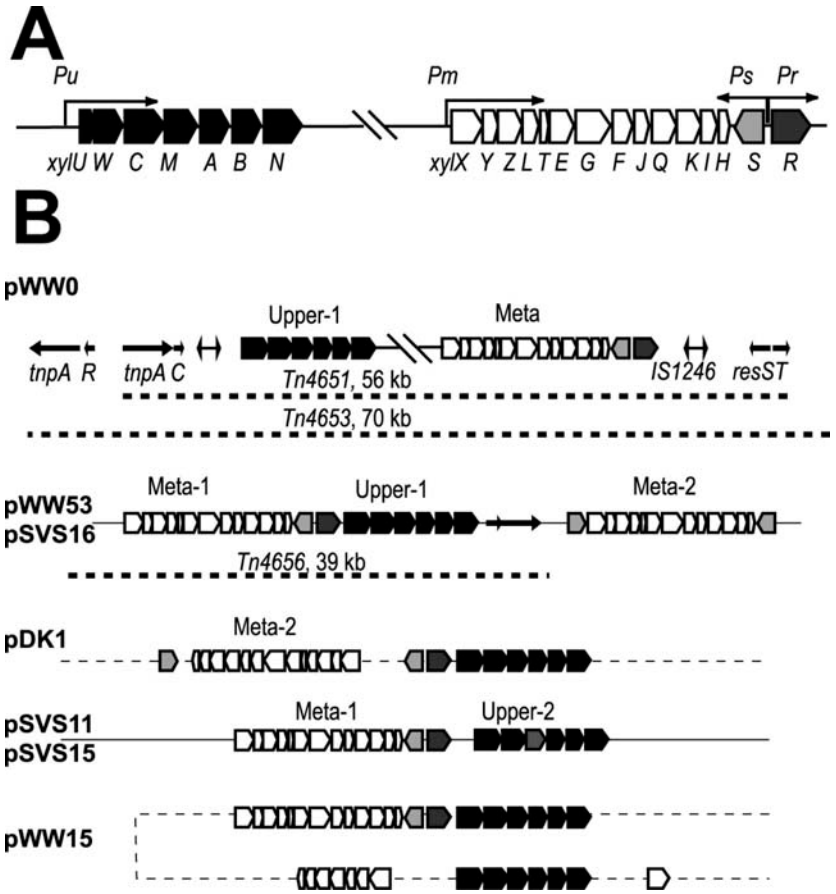


Figure 2. Mosaical variation of TOL plasmids. (a) Organization of the *xyl* genes on pWW0, with indication of transcription from mapped promoters. (b) Comparison between *xyl* clusters of pWW0 and other TOL plasmids pWW53, pDK1,[7–8,134,180] pWW15,[98,130,235] and pSVSs.[177] For further explanation, see main text. Figure courtesy of Vladimir Sentchilo.

of the deep-sea Gammaproteobacterium *Idiomarina loihiensis*. The sequence of pWW0 and its comparison to others suggest that the current structure can be derived from a single original insertion of the Tn*4563* transposon into an IncP-9 ancestor plasmid.[67] All the further insertions took subsequently place in Tn*4563*.

The self-transfer process of pWW0 is mediated by a conjugative system encoded by the *tra* and *mpf* genes. The incompatibility determinants of pWW0 belong to the IncP-9 group. Best matches of the counterparts of *tra* and *mpf* of pWW0 with around 70–80% predicted amino acid sequence identity are

currently two other conjugative plasmids in *Pseudomonas*: pBI709 (which encodes no further catabolic properties) and pDTG1, a 83.0-kb plasmid carrying gene clusters for naphthalene degradation (Figure 1).[67]

Survey of many more environmental *Pseudomonas* isolates with TOL-like plasmids have shown that they are widely distributed, and almost exclusively related to a phenotype of toluene, *m*-xylene and *meta*-toluate degradation via *meta* cleavage. Interestingly, these surveys revealed that in essence the "natural" variation of *xyl* genes is limited to (currently) three distinguishable groups (Figure 2). Plasmid pWW0 carries the *xyl* genes in single copy only (*upper* and *meta* pathway operon, one copy each of *xylS* and *xylR*). Many TOL plasmids, however, carry duplications of the *xyl* genes. For example, two highly homologous yet distinguishable copies of the *meta* pathway operon (*meta I* and *II*) are present on pWW53,[134] on the related TOL plasmids pWW5, pWW74, and pWW88,[35] and on various pSVS plasmids from *Pseudomonas* isolates from Belarus[177] (Figure 2). Nucleotide alignments showed that the *meta*, *meta I*, and *meta II* operons constitute three distinguishable groups.[177,183] Plasmids pWW15, pWW14, pWW20, and pWW74[35,98,240] carry duplications of both *xyl* operons, with one of the *meta* pathway operons being incomplete.[130,235] Two non-identical copies of the regulatory *xylS* gene are found on pDK1[6,180] and three on pWW53.[6] Three groups of *upper* pathway clusters were also postulated on the basis of a comparative survey of two dozen *Pseudomonas m*-xylene degraders from Belarus and the classical pWW0, pWW53, and pWW15 plasmids.[177] Variations among the pSVS-type TOL plasmids included single or double copies of the operons for *meta* cleavage of catechol, and small insertions and nucleotide polymorphisms in individual *xyl* genes. Whether the *xyl* gene cluster duplications occurred via recombination with other replicons or resulted from one ancient duplication event and kept afterward remains an open question. Duplications are potentially unstable and prone to deletions of intervening DNA via homologous recombination, which has often been observed in toluene-degrading organisms maintained in the lab during cultivation on benzoate.[15,102,144,180] However, double copies of *xyl* operons persist in many natural isolates pointing at some selective advantage.

An important finding was that the plasmid backbones (the "vector") into which the *xyl* genes had been inserted, were very variable, suggesting that the *xyl* operons themselves move as stable entities which can only be profitably transferred as a whole. Whereas pWW0 belongs to the IncP-9 plasmid incompatibility group,[10] most other studied TOL plasmids were shown to belong to different and even unknown incompatibility groups.[177] Only pRA1000 from *Alcaligenes eutrophus* strain 345, which exhibited a virtual identical restriction enzyme profiles as pWW0,[86] and pSVS15 from *P. graminis*,[177] were shown to belong to the IncP-9 group. Other TOL plasmids also differed considerably in size and restriction enzyme profiles from each other and from pWW0. Furthermore, not all TOL plasmids are transmissible, indicating that *xyl* genes

are carried on different plasmid backbones. Whether distribution of the whole *xyl* pathway to different plasmids is associated with the transposons Tn*4651*, Tn*4653*, and Tn*4656* cannot be firmly concluded due to lack of detailed knowledge about the plasmid backbones and the DNA regions flanking the *xyl* genes on the different TOL plasmids.

Our knowledge about the evolution of the *xyl* genes itself is mainly based on pWW0. Several studies concluded that (i) the *meta* pathway operon originated as a fusion product of two independently evolved gene blocks, *xylXYZL* and *xylTEGFJQKIH*[77]; (ii) the upper and *meta* pathway operons evolved separately[72]; and (iii) the *meta* pathway operon, in the strict sense (i.e. *xylT* through *xylH*), probably resulted from a recombination event itself.[20,77] This assumption was based on the finding that the *xylK* gene of TOL plasmid pWW102 was more homologous to *nahO* of the naphthalene catabolism plasmid pWW60-22[145] than to corresponding sequences of pWW0 and pDK1[1] Also the *xylE* gene of pWW0 was proposed to have originated via recombination between *xylE* of pDK1 and *nahH* of naphthalene plasmid NAH7.[20]

It should be noted that the possibility to catabolize toluene is not solely given by the *xyl*-encoded pathways. At least four other aerobic pathways for toluene degradation have been characterized (Table 1), none of which, however, are located on a mobile genetic element. Nevertheless, although no evidence exist for spontaneous transfer of the *tod* genes, parts of it apparently have become captured in the formation of chlorobenzene degradation pathways (see below). Also the gene clusters for the other toluene pathways show evidence for mosaic-like combinations of individual genes or gene blocks, a point that has been observed previously and discussed extensively.[221] However, based on DNA sequence relationships most of these presumed combinations must have taken place a long time ago and direct evidence for the combinatory mechanisms can no longer be retrieved.

7. POLYAROMATIC PATHWAYS

The *xyl*-encoded pathways for toluene and xylene degradation are closely related to the *nah*-encoded ones for naphthalene degradation, but strangely, the two are almost never found together in one *Pseudomonas*, although there seems to be no apparent enzymatic reason for this to be so (e.g. enzyme incompatibilities). The few existing examples in which toluene and naphthalene degradation are both present in the same cell do seem to involve a plasmid with *nah* genes, but no TOL-like plasmid (Table 1). By the pioneering work of Harayama *et al.* it was demonstrated that the *nah* and *xyl* pathways are probably ancient hybrids.[78] In fact, naphthalene degradation also proceeds via a *meta* cleavage intermediate pathway which is very similar to that operative in *m*-xylene degradation. A comparison between the two was one of the first occasions to

postulate "cassette"-like recombination mechanisms, by which bacteria could evolve new catabolic pathways not by slow single base pair polymorphisms but rather by recombining larger gene blocks in new profitable combinations, like natural recombinatorial chemistry. The blocks or cassettes which could be recognized by comparing *xyl* and *nah* genes were (i) a *meta* cleavage pathway (common to both), (ii) the *upper* pathway for xylene oxidation to *meta*-toluate (*xylUWCMABN*), (iii) a *meta*-toluate oxidation cluster (*xylXYZL*), both specific for the xylene pathway, (iv) a naphthalene *upper* pathway cluster and (v) salicylate hydroxylase (*nahG*), and finally regulatory genes (*xylR*, *xylS*, *nahR*).[76,77,78,233] As said above, the exact mechanisms of formation for these (evolutionary old) hybrids can no longer be derived with certainty.

The fact that *nah* gene clusters occur on the same replicon types as *xyl* genes may be one reason as to why they have become mutually "exclusive." However, *nah* genes are also found in chromosomal locations, such as in *P. stutzeri* AN10 or in *P. aeruginosa* 57 (Table 1). We have already mentioned the relatedness of the *tra* and *mpf* systems of pWW0 and pDTG1; the only other completely sequenced NAH plasmid, the 101-kb plasmid pND6-1, has a *par* system which is around 98% identical to that of the pL6.5 plasmid of *P. fluorescens*.[111] Available sequence information points to the existence of a few distinguishable operons for naphthalene degradation, but an overall very high degree of sequence conservation. On the classical NAH7 plasmid two clusters of *nah* genes occur, both of which are transcribed in the same direction with less than 5 kb of intervening DNA.[245] Only the *nahR* gene, located upstream of the second cluster is oriented divergently. Both *nah* clusters on the recently completely sequenced plasmids pDTG1 from *P. putida* strain NCIB 9816-4[47] and pND6-1 from *Pseudomonas* sp. strain ND6[111] are oriented face to face, although the gene order within the clusters is the same as for the NAH7 plasmid. In addition, the *nahR* gene on pND6-1 is much further away from the *nahG* gene as on NAH7 and pDTG1, which was attributed to the insertion of mobile elements. Both *nah* clusters on pND6-1 and pDTG1 have a larger intervening DNA sequence, with a few insertion elements (Figure 1). Interestingly, the *nah*-encoded *meta* cleavage pathway on pDTG1 is non-functional due to the insertion of an IS element between *nahG* and *nahT*, which probably prevents expression of the downstream located *nah* genes.[47] The strain can still grow on naphthalene by the use of a chromosomally encoded *ortho* cleavage pathway for catechol degradation.

The 83-kb pDTG1 plasmid has a typical IncP-9 backbone of about 33 kb, based on partitioning, replication, and transfer functions. The gene organizations of those are very similar to those of pWW0.[47] In addition, also other regions, but with unknown function, are conserved between pDTG1 and pWW0. Dennis and Zylstra proposed that an ancestor of pDTG1 (perhaps of the same sort as the pWW0 ancestor) underwent several DNA integration steps before the final formation of the present-day's naphthalene pathway genes. An initial

insertion of 6-kb of DNA from the plasmid pCAR1 took place into the ancestor *ruvB* gene.[47] Subsequently, a single insertion of *nah* pathway genes might have taken place in the *intA* gene of the pCAR1 fragment, and finally, several other ISs (such as IS*Pre*1) entered into this region. pND6-1, on the contrary, does not carry the IS*Pre*1 sequences, but it does show an additional duplication of the *nahG* and *nahF* genes.[111] This it has in common with the chromosomally located *nah* region of *P. stutzeri* AN10.[23]

Biphenyl degradation is also very widespread among bacteria both from gram-positive as well as gram-negative groups.[195,221] Typically, biphenyl degradation (and in some cases simultaneously monochlorobiphenyl) is mediated by the *bph* gene products, leading to hydroxylation and cleavage of one ring, and formation of benzoate and 2-hydroxypenta-2,4-dienoic acid. All known *bph* genes share common ancestry, although the individual genes in the *bph* operons have been subject to reshuffling, deletion or new insertions.[195,221] In contrast to the genes for naphthalene and toluene metabolism, those for biphenyl degradation are mostly not present on plasmids.[195] Although there is no particular reason why *bph* genes would not be on plasmids, only few reports clearly demonstrate the involvement of plasmids in biphenyl degradation.[85] On the contrary, despite being chromosomally located, genes for biphenyl degradation are reported to be transferable.[126,193] One of the best characterized elements is Tn*4371*, the biphenyl catabolic transposon from a strain now classified as *Ralstonia oxalatica* A5.[210] Although Tn*4371* itself was not capable of independent transfer, it carries genes similar to *tra* and *trb* of IncP1-beta plasmids (Figure 1), of which apparently *traI* and *traK* are missing.[210] The transposon can be mobilized on other plasmids, such as pSS50 or RP4.[195] No evidence was found for the Tn*4371* transposon in biphenyl-degrading pseudomonads,[195] despite homology of the *bph* genes. However, also the *bph* genes in *P. putida* KF715 seem to occur on a mobile region of the chromosome.[126] This 90-kb DNA region was named the "bph-sal" element and contained in addition to the *bph* genes three genes for salicylate degradation. The element could be transferred from strain KF715 to other *P. putida* and from those to again others, suggesting independent self-transfer by the element.

At least one other degradation pathway exists, which nicely illustrates the concepts seen above: reshuffling of a few operons with genes for catabolic pathways, either on conjugative plasmids or with help of transposable elements. This example consists of the degradation of carbazole, a heteroaromatic dibenzofuran-like molecule, which is degraded to anthranilate and 2-hydroxypenta-2,4-dienoic acid via angular dioxygenation.[113] Anthranilate is converted to catechol and 2-hydroxypenta-2,4-dienoic acid to acetyl-CoA.[217] The genes for these enzymatic activities are located on plasmids in several *Pseudomonas* strains. For example, in one of these, *P. resinovorans* strain CA 10, a 199-kb plasmid pCAR1 was found, which was recently completely sequenced.[113] The genes for carbazole degradation are organized in two clusters,

ant and *car*, which are both contained within the 73-kb transposon Tn*4676* (Figure 1). Several other IS elements flank the *ant* and *car* gene clusters, suggesting that they were transposed onto an ancestor pCAR1 plasmid. Like for the *xyl* and *nah* clusters, also the *ant* and *car* genes have been found on other plasmids related to pCAR1, and in the chromosome.[113]

In short, these data on toluene, naphthalene, carbazole, and biphenyl degradation demonstrate that more or less coherent gene clusters have evolved, which have common ancestry (like between *nah* and *bph*, or *nah* and *xyl*), but now remain intact and are distributed via transposons and plasmids onto different replicons and in different bacteria. By detailed analysis it was possible to derive cluster analogies and postulate some of the past events that led to formation of the current structure of pathway genes. Small polymorphisms still found between genes for related degradation pathways argue for a continuing "evolution." However, it seems that no further pathway expansion has been taking place in these classical pathways, which is in contrast to what is observed for catabolic pathways of more synthetic compounds.

8. CATABOLIC EXPANSION BY LATERAL GENE TRANSFER

The best examples in my eyes for forward catabolic expansion come from those cases in which selection substrates were used which had no natural counterparts. Although it has been argued that we can never know which compounds truly are natural or synthetic,[174] both the chemical nature and the large quantity of some applied synthetic compounds have almost certainly resulted in forward catabolic expansion of single bacterial strains by genetic mechanisms. The examples treated here consist of degradation of chlorobenzenes, atrazine, 2,4-D, and nitrobenzenes. In these cases the evidence has become very convincing that new catabolic pathways have evolved since the introduction of the synthetic compounds by acquisition of genetic material from different sources, which was successfully combined within a single organism.[93,196,208,221] From analysis of different chlorobenzene degraders, we can even assume that the same events occur reproducibly in different organisms.

9. HERBICIDE DEGRADATION

Microbial degradation of herbicides has served as outstanding example for the adaptive capabilities of microbial communities with respect to catabolic pathways. Most notably, bacterial strains have been isolated that had developed new pathways for degradation of 2,4-D, atrazine and to a lesser extent,

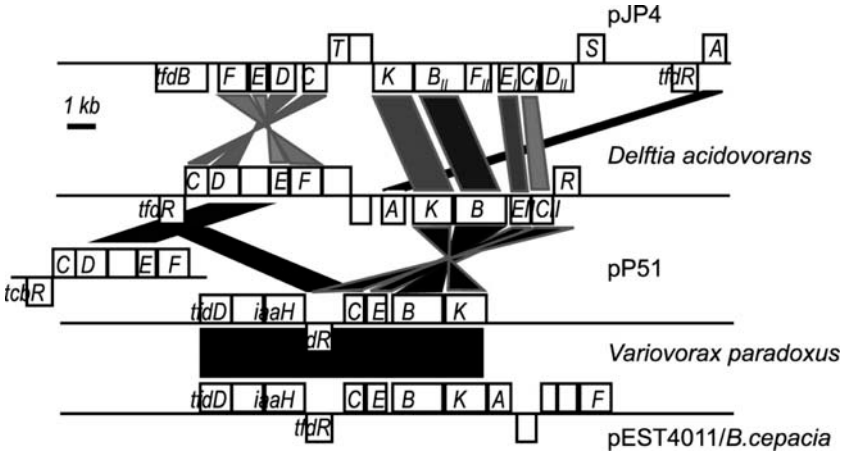


Figure 3. Mosaical patterns in gene clusters for 2,4-dichlorophenoxyacetic acid and chlorocatechol degradation. Boxes indicate open reading frames, with their gene name assignment inside, and orientation (upper, lower strand 5′–3′). Shadings compare highest percentages of identity among functionally related genes: black = ≈100%, gray (*Delftia* vs. pJP4) ≈ 82% amino acid identity, light gray (*Delftia* vs. pJP4) ≈ 55%. For further explanations, see main text. References, see Figure 1, or *Delftia acidovorans*,[83] pP51,[224] *Variovorax paradoxus*,[218] *Burkholderia cepacia*.[148]

chloroaniline or other phenoxyalkanoic acids. To illustrate the potential of herbicides to act as selective carbon sources, it is worth to mention the worldwide production. Around 50,000 tons 2,4-D have been produced yearly,[164] of which it was estimated that 100% entered the open environment. Depending on the application, target fields are treated with between 0.3 and 4.5 kg of active compound per hectare. Apparently, this has been sufficient to select for a wide range of different bacteria capable of degrading 2,4-D, most of which occur in the *Proteobacteria*.[61,62,94,97,219] The most thoroughly studied bacterium degrading 2,4-D is certainly *Ralstonia eutropha* (now *Cupriavidus necator*) JMP134 (pJP4),[51] which culminated recently in the complete genome of the plasmid pJP4 (Figure 1)[211] and a draft genome for the complete organism. For unknown reasons, true *Pseudomonas* spp. form a minority among 2,4-D degrading microorganisms. Rather, *Beta-* and *Alphaproteobacteria* dominate among 2,4-D degraders, but still the story is illustrative from the perspective of catabolic gene evolution. 2,4-D degradation capacities can be broadly divided in two groups: those of *Beta-* and *Gammaproteobacteria*, and those of *Alphaproteobacteria*, such as *Sphingomonas* or *Bradyrhizobium*[61,219]. The beta and gamma groups have genes and gene organizations for 2,4-D degradation quite similar to those found on the pJP4 plasmid of *R. eutropha* JMP134, and are very often located on plasmids as well[61,207,219] (Figure 3). Genes for 2,4-D degradation from the alpha group are quite different from those of pJP4. 2,4-D degrading organisms

are most abundant in 2,4-D exposed areas, but occur even in sites with no prior history of 2,4-D application.[94,97]

Ralstonia eutropha JMP134 (pJP4) was first described by Pemberton and colleagues.[51] It was initially identified as *Alcaligenes eutrophus* and has now seen its name changed to *Cupriavidus necator*. The ability of *R. eutropha* to degrade 2,4-D lies encoded on the pJP4 plasmid, which further confers resistance to mercuric chloride, phenylmercury acetate, and merbromin. The 87.7-kb plasmid pJP4 belongs to the IncP1-beta incompatibility group.[183] From the complete sequence it has become apparent that the plasmid backbone of pJP4 is highly related to that of other IncP1-beta plasmids (Figure 1), which has been the basis to postulate that an ancestor plasmid (perhaps like R751, pB10, or pB3) captured different catabolic and antibiotic resistance modules. The pJP4 plasmid is self-transmissible to various bacterial strains including *E. coli*, *Agrobacterium tumefaciens*, *Rhizobium* sp., *P. putida*, *P. fluorescens*, and *Acinetobacter calcoaceticus*.[51] The *tfd* genes are clustered on a 22-kb DNA fragment (Figure 3) and essentially comprise three gene cassettes from different origins[103]: *tfdA*, encoding the first enzyme needed to attack 2,4-D (α-ketoglutarate-dependent dioxygenase), *tfdTCDEFB*, and $tfdRD_{II}C_{II}E_{II}F_{II}B_{II}K$. The latter two clusters principally encode a pathway to convert chlorinated phenols via chlorocatechols and *ortho* cleavage to 3-oxoadipate.[104,105,139,140,146] The *ortho* cleavage pathway for chlorocatechols is different from the chromosomally encoded *ortho* cleavage pathway for catechol (*cat*), but does not offer the organism a complete intermediate catabolic pathway, since it ends at 3-oxoadipate which cannot directly feed into the TCA cycle. Therefore, chromosomally located enzymes which convert 3-oxoadipate to succinyl- and acetyl-CoA are necessary to confer growth on 2,4-D.

It has been argued that pathways for chlorocatechol degradation such as those encoded by *tfdCDEF* are considerably older than the 50 years or so that 2,4-D or other chlorinated substrates have been applied.[174] Its original function probably was the same as it is today, i.e. the catabolism of chlorocatechols, which are commonly found in nature as are other chloroaromatics, such as chlorophenols.[124] Interestingly, also 2,4-dichlorophenol exists as a natural compound, produced as a metabolite by a *Penicillium* sp.[164] This may explain the existence, in a pre-2,4-D era, of a gene like *tfdB* and of the chlorocatechol pathway. However, the combination of a gene cluster for chlorophenol–chlorocatechol degradation with the gene for α-ketoglutarate-dependent 2,4-D dioxygenase (*tfdA*) must have been a recent event, given the truly remarkable mosaical variety of different genes for 2,4-D utilization[61,219] (Figure 3).

Already the pJP4 plasmid itself harbors two essentially complete and similar sets of genes for chlorophenol degradation with nevertheless different origins. Gene orders for the chlorophenol cluster are not conserved. Other more recently sequenced plasmids and gene clusters for 2,4-D and chlorophenol degradation confirm this. On the pEST4011 plasmid from *Achromobacter*

denitrificans,[230] the gene order of the *tfd* genes is: *tfdD* (ORF), *iaaH*, *tfdRCEBKAF* (Figure 3). The *tfd* genes of pEST4011 are up to 99% identical to those of plasmid pIJB1 of *Burkholderia cepacia*[148] and of *Variovorax paradoxus*[218] with a conservation of gene order, except for a small duplication in the region of *tfdD* and *iaaH*. However, they are only between 74% and 90% identical in predicted amino acid sequence to the gene products of the *tfdII* cluster of pJP4. Evolutionary events have taken another course in plasmid pMAB1 of *B. cepacia*.[21] This plasmid displays an identical set of *tfdCDEFB* genes as pJP4, however, the second *tfd* cluster of pJP4 is lacking, possibly by a recombination between *tfdT* and *tfdR*. Finally, in *Delftia acidovorans* yet another combination of *tfd* clusters can be found,[83] partially related to the *tfd* of pJP4, to that of pEST4011, and to the *tcb* genes for chlorocatechol degradation of *Pseudomonas* sp. strain P51.[224]

The activity of insertion elements in the formation of the *tfd* pathways has been postulated. A potentially transposable element comprising the IS element IS*JP4* and a partial copy including one of its recognition ends is encompassing the complete $tfdRD_{II}C_{II}E_{II}F_{II}B_{II}K$ cluster.[110] The same bordering elements with an antibiotic resistance gene in between were shown to be functionally transposing in *R. eutropha* JMP289, a plasmid-free derivative of JMP134. Transposition of IS*JP4* to the chromosome of JMP134 could also be detected. In addition, pJP4 carries a copy of the widely distributed IS element IS*1071* (Table 2), which was shown to be responsible for large-scale plasmid rearrangements of pJP4 in growing cultures of *R. eutropha* and might have been inserting the first *tfd* gene module in a pJP4 ancestor plasmid.[211] Finally, a Tn*3*-related transposon with 86% amino acid sequence identity to a *tnpA* of the NAH plasmid pND6-1 is encoded on pJP4.[211] The *tfd* genes on plasmids pEST4011 and pIJB1 are part of a 48-kb transposable element (Tn*5530*) flanked by two copies of a hybrid insertion element IS*1071*::IS*1471*.[148,230] Furthermore, an insertion element related to IS*BPH* is present downstream of the *tfdF* gene on pEST4011. The *tfd* genes of *D. acidovorans* are not plasmid, but chromosomally located. They are flanked by two copies of IS*1071* and IS*1380*, which form part of a larger duplicated region involving *tfdR* and *tfdC* as well.[83]

The components of the predicted conjugative system of pJP4 (*trb* and *tra*) are up to 99% identical in amino acid sequence with those of the IncP1-beta plasmids R751 of *Enterobacter aerogenes* and pB3, an environmentally captured plasmid.[211] Although pIJB1 belongs to the IncP1-beta group, the backbone of pEST4011 is different from pJP4 as well and Tra and Trb systems point to about 80% sequence conservation with those of pB3 pseudomonads.[230]

Another fascinating example for catabolic gene evolution is that of the pathway for atrazine degradation. We have to assume that atrazine, like 2,4-D, is a fully synthetic compound, which was specifically developed as herbicide and of which large quantities were released into the environment. A wide variety of bacteria capable of degrading atrazine have now been reported, which were isolated from agricultural soils exposed to this herbicide. They include members

of the genera *Pseudomonas*, *Rhizobium*, *Acinetobacter*, *Arthrobacter*, *Nocardiodes*, and *Agrobacterium*.[119,166,169,209] For *Pseudomonas* sp. strain ADP1, the genes for atrazine degradation are localized on the 100-kb pADP1 plasmid, but form four distinct clusters: *atzA*, *atzB*, *atzC*, and *atzDEF* (Figure 1).[119] The *atzA*, *atzB*, and *atzC* genes have also been detected in other organisms with up to 100% sequence conservation, are almost exclusively present on plasmids (of different nature) and are frequently unstable and lost by recombination via insertion elements.[209] In some strains, the function encoded by *atzA* is performed by the *trzN* gene product, which in itself is very widespread and conserved as well.[169,209] Recent sequencing of a plasmid fragment from *Arthrobacter aurescens* confirmed the identity of *atzB*, *atzC*, and *trzN* in this strain and demonstrated that large fragments apparently have been exchanged and newly recombined between pADP1 plasmids and other *Arthrobacter* plasmids.[169] Further indication for the recent (independent) integration of the *atzA*, *atzB*, and *atzC* genes on pADP1 and similar vehicles is the complete absence of regulation for gene expression.[119] The wide distribution of *atz* and *trz* genes on different vehicles and in different bacterial genera shows that probably transposition and conjugative events together were responsible. Furthermore, it seems to suggest that evolution of the atrazine degradation pathway occurred independently and reproducible, and cannot be attributed to a singular arisal and further distribution.

10. CHLORO- AND NITROAROMATIC DEGRADATION

Degradation of chlorobenzenes has been taken as a good test case for demonstrating the capacity of microbial communities to expand their catabolic properties. Chlorobenzenes are only known as synthetic compounds, although other natural chlorinated aromatic compounds do exist (e.g. chlorophenols, see above). Microorganisms do not generally appreciate chlorobenzenes, probably because of their immediate toxicity to the membrane or the formation of toxic intermediates. It had been recognized very early on[160,161,163] that toluene-degrading bacteria such as *P. putida* F1 were capable of at least partially transforming chlorobenzenes to chlorocatechols by the activity of their toluene dioxygenase systems. However, toluene-degrading organisms tend to induce a catechol 2,3-dioxygenase activity when growing on toluene and this enzyme mostly cannot deal effectively with chlorocatechols as substrates.[99] Recently, some exceptions to this were described.[116,117] It was therefore attempted to combine by "natural genetic engineering" (conjugative transfer) a toluene degradation pathway and one for chlorocatechol degradation within one organism. Indeed, this could be accomplished, for example, by matings between *P. putida* F1 and *Pseudomonas* sp. strain B13, a 3-chlorobenzoate-degrading organism, in which case the genes for 3-chlorocatechol degradation are transferred from B13 to F1.[154,156,161,162] Interestingly, however, microorganisms could also be isolated

from chlorobenzene contaminated environments or from artificially contaminated natural environments in the laboratory, which were capable of growing on chlorobenzenes as sole carbon and energy source.[18,19,70,123,191,225,228] Although initially classified as *Pseudomonas*, most of these strains were afterward shown to belong to the Betaproteobacteria, but this does not change the interest in the accomplishment. Genetic analysis showed that almost exclusively and reproducibly a natural recombination between a toluene degradation pathway (of the *tod*-type) and a chlorocatechol degradation pathway had occurred. In three cases the genes for the toluene dioxygenase and *cis*-dihydrodiol dehydrogenase had been transferred to the new recipient, in one case without a complete *todF* gene (normally upstream of the toluene dioxygenase genes), but in all cases without a functional *todE* gene (normally downstream of the toluene dioxygenase genes). In all three cases, insertion elements were nearby the newly inserted genes, forming still active transposable structures (e.g. Tn*5280*, the *tcb* chlorobenzene transposon of *Pseudomonas* sp. strain P51[229]) or non-active structures (the *mcb* genes of *Ralstonia* sp. strain JS705[123]). In *Pseudomonas* sp. strain P51 the *tcb* genes had recombined on a transferable plasmid (pP51) already encoding a chlorocatechol degradation pathway (*tcbRCDEF* genes).[227] The *tcbRCDEF* genes on their turn also show up in the pathway for 2,4-D degradation in *D. acidovorans* (Figure 3), on the pA81 plasmid from *A. xylosooxidans* (Figure 1) (Jencova, V., unpublished results) and in a pathway for 4-chlorobenzoate degradation in *R. eutropha* NH9.[132] In *Ralstonia* sp. strain JS705 the *mcb* genes had recombined within the *clc* genes for chlorocatechol degradation on a self-transmissable genomic island.[123] On the basis of PCR amplification and sequencing it became apparent that the same IS elements seemed to have captured a *tod*-like region twice, but in a slightly different form and in a slightly different organism[19] (Figure 4). Most interestingly, for *Ralstonia* sp. strain JS705 it was possible to trace back the ancestor donor of the *mcb* genes, which was isolated from groundwater in the same area. This "donor" strain was a toluene-degrading organism with a 100% identical nucleotide sequence of the *mcb* genes.[123] The border of the excised and transferred fragment could be exactly determined. As more and more sequences will become available, it will probably be clear for other cases of catabolic expansion as to where the donating genes may have come from, but until then the case of *Ralstonia* sp. strain JS705 stands as one of the few where ongoing pathway evolution by "natural recombination" has been observed.

It is now known that a second possibility to form a monochlorobenzene degradation pathway is existing in bacteria in the environment. In these organisms, like *P. putida* GJ31 or *P. fluorescens* SK-1, the catechol 2,3-dioxygenase (*cbzE*), in contrast to for example *todE*, is capable of converting 3-chlorocatechol and eliminating the chlorine atom during formation of muconic semialdehyde, which can then subsequently be degraded in a classical *meta* cleavage pathway.[66,117] However, it is not known whether specific adaptive

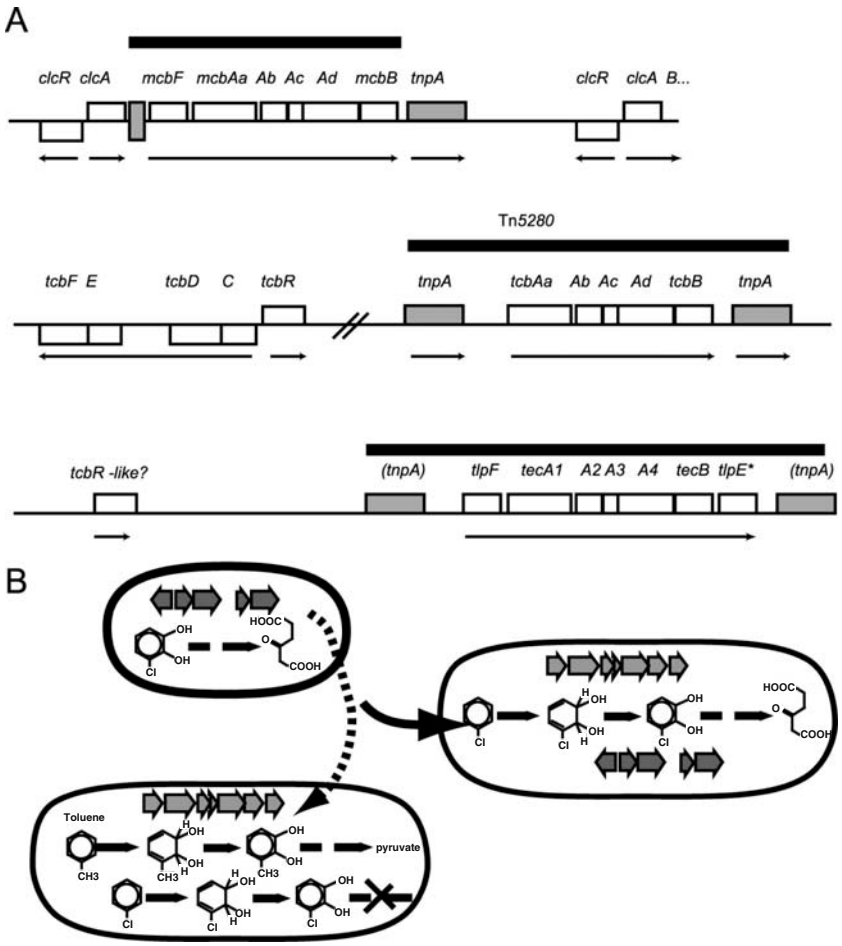


Figure 4. Reproducible gene capturing of a cluster for toluene dioxygenase in the formation of the degradation pathway for chlorobenzene. (a) *Clc-mcb* gene region of *Ralstonia* sp. strain JS705,[123,228] *tcb* gene regions of plasmid pP51 in *Pseudomonas* sp. strain P51,[226,233] and *tec* regions in *Burkholderia* sp. strain PS12.[19] (b) Conceptual idea of pathway formation by gene capture and natural recombination. For further explanation, see main text.

mechanisms were necessary to "evolve" this pathway or if it had not been detected because of slower growth of the organisms compared to those degrading chlorobenzene via a modified *ortho* cleavage pathway.

Similar evidence for pathway expansion has been retrieved from detailed studies on degradation of nitroaromatic compounds, in particular for nitrobenzene and 2,4-dinitrotoluene[190] (Table 1). Nitrobenzene degradation involves the linkage of a conserved aminophenol degradation pathway with first a reduction

of the nitro group to hydroxylaminobenzene and secondly a rearrangement to 2-aminophenol. 2-Aminophenol degradation proceeds analogously to catechol degradation via *meta* cleavage, except that the individual enzymes accommodate amino substituents. The pathways were characterized from *P. pseudoalcaligenes* JS45[93] and *P. putida* HS12.[138] Until now it is rather mysterious as from where the genes for nitrobenzene nitroreductase originated, as they only show limited homology to other flavoprotein reductases and even to 4-nitrobenzoate nitroreductases.[93] The mutase enzyme needed for the second step, on the contrary, is widely present in pathways such as for nitrophenol and 2-nitrobenzoate. The two hydroxylaminobenzene mutase genes from strains JS45 and HS12 are highly homologous, but strain JS45 carries a second paralogous enzyme with only 44% amino acid identity.[44] The mutase gene in strain HS12 is carried on a separate plasmid than the rest of the genes for 2-aminophenol degradation (pNB1, pNB2, Table 1).

The pathway for 2,4-dinitrotoluene seems composed from three modules.[92] First a multicomponent dioxygenase closely related to naphthalene dioxygenase, but with inclusion of two genes for a terminal oxygenase for salicylate hydroxylase, which, however are not needed for the dinitrotoluene pathway. Secondly, a flavoprotein monooxygenase encoded by the gene *dntB*, which is responsible for attacking the 4-methyl-5-nitrocatechol that results from the first module's activity. There is no closely related ortholog known for this enzyme at present. The final module consists of a pathway for trihydroxytoluene degradation which begins with a ring-fission oxygenase dissimilar to other known oxygenases and continues with a pathway analogous to amino acids degradation. In one organism degrading 2,4-dinitrotoluene, *Burkholderia cepacia* R34 the genes for these modules were clustered in a 20-kb region on a plasmid. In another *Burkholderia* sp. strain DNT the genes were present on a plasmid as well, but did not cluster in one region.[93]

11. THE CLC ELEMENT

We have seen the large involvement of conjugative plasmids as vehicles for acquisition and distribution of genes for catabolic pathways. One last paragraph will be devoted to describing the curious *clc* element, a conjugative element that integrates into and excises from the chromosome. This final example will show that the possibilities for acquisition of mobilization of genes is not limited to plasmids, but can in essence comprise all chromosomal genes as well. Furthermore, it serves to demonstrate that formation of catabolic pathway clusters employs the same mechanisms as generation and distribution of pathogenic functions.

In the bacterium *Pseudomonas* sp. strain B13, the first bacterium described to degrade 3-chlorobenzoate,[52] self-transfer of the capacity for

chlorocatechol degradation was well known but finally could not be attributed to plasmid mobilization.[232] Instead, this transferable element turned out to be an excisable genomic region and is now known as the *clc* element because of its ability to provide chlorocatechol degradation to the host.[154] In addition, the *clc* element carries a complete cluster of genes for 2-aminophenol degradation (Figure 1). The *clc* element is known as a genomic island (GEI), among which pathogenicity islands are found as well, although most genomic islands are (no longer) mobile. Fully functional GEI are also part of what was previously designated as ICE; integrative and conjugative elements.[26] The *clc* element has a size of 105 kb and we know of one natural variant which is very closely related. This *clc*-like element was present in *Ralstonia* sp. strain JS705 (see above), an organism degrading chlorobenzenes and was isolated from contaminated groundwater in the United States.[123,228] The *clc* element is not only self-transferable in laboratory "matings," but also transfers in more complex communities and technical systems from strain B13 to other Beta- and Gammaproteobacteria, such as *Ralstonia*, *Pseudomonas putida*, *Burkholderia* sp. and *P. aeruginosa*.[156,162,194,247] This demonstrated that the *clc* element is not a laboratory curiosity, but a type of mobile DNA element present in natural microbial communities which previously had not been recognized.

With the current explosion of bacterial genome data and detailed bioinformatic approaches for comparative genome analyses, it has become apparent that GEI are very widespread and must have been important for shaping the bacteria genome structure.[24,129] The evidence for the existence of genomic islands had been obtained long before the large-scale genome sequencing projects.[69] Mostly known as pathogenicity islands (PAI), they were known to be involved in the instability or spontaneous loss of virulence functions in pathogenic bacteria. Now, by combining information from various approaches, such as genometry and experimental genetics, genomic islands appear as a widely diverse class of DNA elements in different stages of functionality with relationships to bacteriophages, conjugative transposons, and plasmids.[26,49] Genomic islands are therefore, like plasmids, evolutionary and adaptively very interesting elements, because (i) they are plentiful in most bacterial chromosomes and seem to be responsible for shaping a large part of the bacterial genome,[28] (ii) they occur in different "stages" of functionality or evolutionary development, (iii) they seem to be hot spots for acquisition of auxiliary functions and contribute to horizontal gene transfer, and (iv) they seem to have a life style of their own, but may influence central processes of the core bacterial genome.[26,49,128,226]

A GEI has features different from conjugative plasmids. It carries one or more of the following features: a more or less specific chromosomal insertion site, a mechanism for inserting and excising from that site, functions for self-transfer (either via phage particles, conjugation, or unknown), regulatory

functions which control integration, excision or transfer, and a variable region. The core regions on genomic islands (insertion sites, integration and excision functions, and transfer and mobilization) are not free from deletions or acquisitions either (as expected).[100] In many cases, genomic data only show traces for a genomic island: a gene for an integrase nearby a gene for tRNA, but no further evidence for transfer and mobilization functions, or perhaps not even for the target site duplication.[141] In some cases, even the integrase gene is unfunctional or absent and the region of the genomic island can only be detected by *aberrant* nucleotide sequence composition compared to the rest of the chromosome of the host cell. Transfer and mobilization functions may exist,[26] although several genomic islands (as explained below) seem to have transfer systems unsimilar to those of conjugative transposons or conjugative plasmids. Variable regions have attracted much interest, because their characterization showed that genomic islands can acquire different functions whilst keeping a certain "core" necessary for integration, excision, transfer, and regulation.[16,106,131] Variable regions can contain pathogenicity determinants such as virulence factors, adhesins, small molecule metabolism, multidrug resistance genes, hence the term pathogenicity islands.[49] However, the pathogenicity character may be less pronounced and instead the variable region may encode functions such as degradation of aromatic compounds or even symbiosis factors.

As their definition says, genomic islands (GEI) reside on the chromosome. However, their position is not random and they are often located nearby tRNA genes.[26] From the few examples of which mechanistic information exist, it is known that GEI, like bacteriophages and conjugative transposons have at some point integrated at those positions, using site-specific recombination mechanisms.[25,36,155] In all hosts examined, the *clc* element resides site-specifically in the chromosome at the most 3′ 18-bp sequence of a tRNAGly gene.[154,155,194] Depending on the host, the *clc* element can be present in one, two, or more copies. Also in strain B13 two copies are present. At the other end of the element, a copy of this 18-bp sequence is found. A critical feature of the *clc* element's life style is formed by the integrase, the gene of which (*intB13*) is found close to the tRNAGly insertion site and oriented away from it. The integration process into the chromosome is mediated by the IntB13 integrase enzyme.[155] One of the important features of GEI is that they do not remain stably integrated in the chromosome, but can excise by recombination between both "ends."[27,109,155,176] Small amounts of the excised *clc* element can be detected in cultures of strain B13 itself by Southern hybridization and PCR.[155] Sequencing of this excised form showed that it consists of a closed DNA molecule in which both 18-bp repeats have recombined again to one.[155] The current hypothesis therefore is that the integrated *clc* element can excise from its chromosomal location and form a closed "circular" intermediate. The closed intermediate can either reintegrate or be transferred to a new recipient cell, in which it can again integrate. The existence of a closed circular excised GEI form was also

demonstrated for the SXT element, for HPI and several others by cloning the junction regions by inverse PCR and subsequent DNA sequencing.[27,109,155,176] In some cases, like for HPI and the *clc* element, recombination of the right and left ends positions a strong promoter at the left end in front of the integrase gene.[153,179] The idea is that this configuration results in temporary overexpression of the integrase and efficient reintegration into the chromosomal target. On the other hand, expression of the integrase in the integrated form is mostly silent.[179]

Self-transfer of the *clc* element is not a well-described and generalized process, in contrast to conjugative transposons or plasmid conjugation. It must contain all the factors essential for self-transfer, since it can transfer from the strain B13 to an intermediate recipient and from there to another species.[178] However, details on the transfer process are not known and no clear homologies exist to classical type IV conjugative systems. Other GEI and ICE have partly conjugative functions recognizable from existing type IV conjugative systems. For example, the 100-kb SXT element of *V. cholerae* is self-transmissable between gram-negative species by a conjugative process. The sequence of SXT shows a number of regions with strong similarities to the *tra* genes of plasmid R27 and the *Agrobacterium* Ti plasmid[16,17] and, to a lesser extent, to those of the pCAR1 plasmid (Figure 1). An example in which transfer functions are recognizable but partially dysfunctional is the Tn*4371* in *Ralstonia oxalatica* A5 which carries the genes for biphenyl degradation (Figure 1).[122,210] The transfer functions on Tn*4371* are similar to those of IncP and Ti plasmids, however, the element is unable to transfer by itself, because of missing TraI and TraK functions.

Interestingly, the *clc* element is a very close relative to a series of GEI in *P. aeruginosa*, *Xylella fastidiosa*, *Burkholderia xenovorans*, and *Ralstonia metallidurens*, which are not known for self-transfer but contribute to pathogenicity and catabolic functions in Beta- and Gammaproteobacteria.[100,106,128,226] Most notably, two 100-kb GEI in *P. aeruginosa* isolates, designated PAGI-2 and PAGI-3, have large regions in common to the *clc* element. Some elements nicely seem to point to a continuum between episomal and integrated forms.[100] One of these, pKLC102 of *P. aeruginosa*, shows a mosaic of fragments with different origins. Parts of PAGI-2 and PAGI-3 are recognizable, but also a clear *pil* system and an *oriV* related to *Pseudomonas syringae* and *Azotobacter vinelandii*. A XerC-type integrase seems responsible for integration of pKLC102 into tRNA-Lys. Strangely enough, in *P. aeruginosa* isolates from cystic fibrosis patients, a 23-kb integron sequence has integrated near the *pil* region on pKLC102. This integron carries resistance genes for gentamycin and tobramycin, and further bears *tra*-like and *par* genes. A further chromosomal recombination can occur between IS*6100* of the integron and a IS*6100* copy elsewhere on the chromosome, which removes the integrase from pKLC102 and renders the element incapable of excision.[100]

12. CONCLUDING REMARKS

Investigating the molecular archeology of catabolic pathways in pseudomonads and related bacteria and dissecting the different adaptation bacterial mechanisms has given an extremely detailed picture on the ways how environmentally successful mutants developed new genetic structures for pollutant metabolism. Principally, the picture is reinforced of extremely effective mechanisms to capture gene fragments from different sources, recombine them *de novo* and distribute on whatever mobile conjugative element is at hand.

ACKNOWLEDGMENTS

I am much indebted to Vladimir Sentchilo and Johan Leveau for their help in preparing part of this material.

REFERENCES

1. Aemprapa, S., and Williams, P.A., 1998, Implications of the *xylQ* gene of TOL plasmid pWW102 for the evolution of aromatic catabolic pathways. *Microbiology*, 144:1387–1396.
2. Alonso, S., Bartolome-Martín, D., del Alamo, M., Diáz, E., García, J.L., and Pérera, J., 2003, Genetic characterization of the styrene lower catabolic pathway of *Pseudomonas* sp. strain Y2. *Gene*, 319:71–83.
3. Arber, W., 2000, Genetic variation: molecular mechanisms and impact on microbial evolution. *FEMS Microbiol. Rev.*, 24:1–7.
4. Arenghi, F.L., Pinti, M., Galli, E., and Barbieri, P., 1999, Identification of the *Pseudomonas stutzeri* OX1 toluene-*o*-xylene monooxygenase regulatory gene (*touR*) and of its cognate promoter. *Appl. Environ. Microbiol.*, 65:4057–4063.
5. Arias-Barrau, E., Olivera, E.R., Luengo, J.M., Fernández, C., Galan, B., García, J.L., Diáz, E., and Minambres, B., 2004, The homogentisate pathway: a central catabolic pathway involved in the degradation of L-phenylalanine, L-tyrosine, and 3-hydroxyphenylacetate in *Pseudomonas putida*. *J. Bacteriol.*, 186:5062–5077.
6. Assinder, S.J., De Marco, P., Osborne, D.J., Poh, C.L., Shaw, L.E., Winson, M.K., and Williams, P.A., 1993, A comparison of the multiple alleles of *xylS* carried by TOL plasmids pWW53 and pDK1 and its implications for their evolutionary relationship. *J. Gen. Microbiol.*, 139:557–568.
7. Assinder, S.J., de Marco, P., Sayers, J.R., Shaw, L.E., Winson, M.K., and Williams, P.A., 1992, Identical resolvases are encoded by *Pseudomonas* TOL plasmids pWW53 and pDK1. *Nucleic Acids Res.*, 20:5476.
8. Assinder, S.J., and Williams, P.A., 1988, Comparison of the *meta* pathway operons on NAH plasmid pWW60-22 and TOL plasmid pWW53-4 and its evolutionary significance. *J. Gen. Microbiol.*, 134:2769–2778.
9. Assinder, S.J., and Williams, P.A., 1990, The TOL plasmids: determinants of the catabolism of toluene and the xylenes. *Adv. Microb. Physiol.*, 31:1–69.

10. Austen, R.A., and Dunn, N.W., 1977, A comparative study of the NAH and TOL catabolic plasmids in *Pseudomonas putida*. *Aust. J. Biol. Sci.*, 30:357–366.
11. Baldwin, B.R., Mesarch, M.B., and Nies, L., 2000, Broad substrate specificity of naphthalene- and biphenyl-utilizing bacteria. *Appl. Microbiol. Biotechnol.*, 53:748–753.
12. Barbieri, P., Arenghi, F.L., Bertoni, G., Bolognese, F., and Galli, E., 2001, Evolution of catabolic pathways and metabolic versatility in *Pseudomonas stutzeri* OX1. *Antonie van Leeuwenhoek Int. J. Microbiol.*, 79:135–140.
13. Bartolome-Martín, D., Martińez-García, E., Mascaraque, V., Rubio, J., Perera, J., and Alonso, S., 2004, Characterization of a second functional gene cluster for the catabolism of phenylacetic acid in *Pseudomonas* sp. strain Y2. *Gene*, 341:167–179.
14. Basu, A., Dixit, S.S., and Phale, P.S., 2003, Metabolism of benzyl alcohol via catechol *ortho*-pathway in methylnaphthalene-degrading *Pseudomonas putida* CSV86. *Appl. Microbiol. Biotechnol.*, 62:579–585.
15. Bayley, S.A., Duggleby, C.J., Worsey, M.J., Williams, P.A., Hardy, K.G., and Broda, P., 1977, Two modes of loss of the Tol function from *Pseudomonas putida* mt-2. *Mol. Gen. Genet.*, 154:203–204.
16. Beaber, J.W., Burrus, V., Hochhut, B., and Waldor, M.K., 2002, Comparison of SXT and R391, two conjugative integrating elements: definition of a genetic backbone for the mobilization of resistance determinants. *Cell Mol. Life Sci.*, 59:2065–2070.
17. Beaber, J.W., Hochhut, B., and Waldor, M.K., 2002, Genomic and functional analyses of SXT, an integrating antibiotic resistance gene transfer element derived from *Vibrio cholerae*. *J. Bacteriol.*, 184:4259–4269.
18. Beil, S., Happe, B., Timmis, K.N., and Pieper, D.H., 1997, Genetic and biochemical characterization of the broad spectrum chlorobenzene dioxygenase from *Burkholderia* sp. strain PS12: dechlorination of 1,2,4,5-tetrachlorobenzene. *Eur. J. Biochem.*, 247:190–199.
19. Beil, S., Timmis, K.N., and Pieper, D.H., 1999, Genetic and biochemical analyses of the *tec* operon suggest a route for evolution of chlorobenzene degradation genes. *J. Bacteriol.*, 181:341–346.
20. Benjamin, R.C., Voss, J.A., and Kunz, D.A., 1991, Nucleotide sequence of *xylE* from the TOL pDK1 plasmid and structural comparison with isofunctional catechol-2,3-dioxygenase genes from TOL, pWW0 and NAH7. *J. Bacteriol.*, 173:2724–2728.
21. Bhat, M.A., Tsuda, M., Horiike, K., Nozaki, M., Vaidyanathan, C.S., and Nakazama, T., 1994, Identification and characterization of a new plasmid carrying genes for degradation of 2,4-dichlorophenoxyacetate from *Pseudomonas cepacia* CSV90. *Appl. Environ. Microbiol.*, 60:307–312.
22. Bolognese, F., Di Lecce, C., Galli, E., and Barbieri, P., 1999, Activation and inactivation of *Pseudomonas stutzeri* methylbenzene catabolism pathways mediated by a transposable element. *Appl. Environ. Microbiol.*, 65:1876–1882.
23. Bosch, R., García-Valdes, E., and Moore, E.R., 2000, Complete nucleotide sequence and evolutionary significance of a chromosomally encoded naphthalene-degradation lower pathway from *Pseudomonas stutzeri* AN10. *Gene*, 245:65–74.
24. Boucher, Y., Douady, C.J., Papke, R.T., Walsh, D.A., Boudreau, M.E.R., Nesbø, C.L., Case, R.J., and Doolittle, W.F., 2003, Lateral gene transfer and the origins of prokaryotic groups. *Annu. Rev. Genet.*, 37:283–328.
25. Buchrieser, C., Brosch, R., Bach, S., Guiyoule, A., and Carniel, E., 1998, The high-pathogenicity island of *Yersinia pseudotuberculosis* can be inserted into any of the three chromosomal asn tRNA genes. *Mol. Microbiol.*, 30:965–978.
26. Burrus, V., Pavlovic, G., Decaris, B., and Guédon, G., 2002, Conjugative transposons: the tip of the iceberg. *Mol. Microbiol.*, 46:601–610.
27. Burrus, V., and Waldor, M.K., 2003, Control of SXT integration and excision. *J. Bacteriol.*, 185:5045–5054.

28. Burrus, V., and Waldor, M.K., 2004, Shaping bacterial genomes with integrative and conjugative elements. *Res. Microbiol.*, 155:376–386.
29. Campos-García, J., Esteve, A., Vaźquez-Duhalt, R., Ramos, J.L., and Soberoń-Chavez, G., 1999, The branched-chain dodecylbenzene sulfonate degradation pathway of *Pseudomonas aeruginosa* W51D involves a novel route for degradation of the surfactant lateral alkyl chain. *Appl. Environ. Microbiol.*, 65:3730–3734.
30. Canchaya, C., Fournous, G., Chibani-Chenoufi, S., Dillmann, M.L., and Brüssow, H., 2003, Phage as agents of lateral gene transfer. *Curr. Opin. Microbiol.*, 6:417–424.
31. Carl, B., Arnold, A., Hauer, B., and Fetzner, S., 2004, Sequence and transcriptional analysis of a gene cluster of *Pseudomonas putida* 86 involved in quinoline degradation. *Gene*, 331:177–188.
32. Carrington, B., Lowe, A., Shaw, L.E., and Williams, P.A., 1994, The lower pathway operon for benzoate catabolism in biphenyl-utilizing *Pseudomonas* sp. strain IC and the nucleotide sequence of the *bphE* gene for catechol 2,3-dioxygenase. *Microbiology*, 140:499–508.
33. Chablain, P.A., Zgoda, A.L., Sarde, C.O., and Truffaut, N., 2001, Genetic and molecular organization of the alkylbenzene catabolism operon in the psychrotrophic strain *Pseudomonas putida* 01G3. *Appl. Environ. Microbiol.*, 67:453–458.
34. Chakrabarty, A.M., Friello, D.A., and Bopp, L.H., 1978, Transposition of plasmid DNA segments specifying hydrocarbon degradation and their expression in various microorganisms. *Proc. Natl. Acad. Sci. U.S.A.*, 75:3109–3112.
35. Chatfield, L.K., and Williams, P.A., 1986, Naturally occurring TOL plasmids in *Pseudomonas* strains carry either two homologous or two nonhomologous catechol 2,3-oxygenase genes. *J. Bacteriol.*, 168:878–885.
36. Cheng, Q., Paszkiet, B.J., Shoemaker, N.B., Gardner, J.F., and Salyers, A.A., 2000, Integration and excision of a *Bacteroides* conjugative transposon, CTn*DOT*. *J. Bacteriol.*, 182:4035–4043.
37. Choi, E.N., Cho, M.C., Kim, Y., Kim, C.K., and Lee, K., 2003, Expansion of growth substrate range in *Pseudomonas putida* F1 by mutations in both *cymR* and *todS*, which recruit a ring-fission hydrolase CmtE and induce the *tod* catabolic operon, respectively. *Microbiology*, 149:795–805.
38. Christensen, B.B., Sternberg, C., Andersen, J.B., Eberl, L., Moller, S., Givskov, M., and Molin, S., 1998, Establishment of new genetic traits in a microbial biofilm community. *Appl. Environ. Microbiol.*, 64:2247–2255.
39. Clarke, P.H., 1982, The metabolic versatility of pseudomonads. *Antonie van Leeuwenhoek Int. J. Microbiol.*, 48:105–130.
40. Clarke, P.H., 1984, The evolution of degradative pathways, pp. 11–27. In D.T. Gibson (ed.), *Microbial Degradation of Organic Compounds*. Marcel Dekker Inc., New York.
41. Corbella, M.E., Garrido-Pertierra, A., and Puyet, A., 2001, Induction of the halobenzoate catabolic pathway and cometabolism of *ortho*-chlorobenzoates in *Pseudomonas aeruginosa* 142 grown on glucose-supplemented media. *Biodegradation*, 12:149–157.
42. Dagher, F., Deziel, E., Lirette, P., Paquette, G., Bisaillon, J.G., and Villemur, R., 1997, Comparative study of five polycyclic aromatic hydrocarbon degrading bacterial strains isolated from contaminated soils. *Can. J. Microbiol.*, 43:368–377.
43. Dahlberg, C., Bergström, M., and Hermansson, M., 1998, In situ detection of high levels of horizontal plasmid transfer in marine bacterial communities. *Appl. Environ. Microbiol*., 64:2670–2675.
44. Davis, J.K., Paoli, G.C., He, Z., Nadeau, L.J., Somerville, C.C., and Spain, J.C., 2000, Sequence analysis and initial characterization of two isozymes of hydroxylaminobenzene mutase from *Pseudomonas pseudoalcaligenes* JS45. *Appl. Environ. Microbiol.*, 66:2965–2971.
45. Dejonghe, W., Goris, J., El Fantroussi, S., Höfte, M., De Vos, P., Verstraete, W., and Top, E.M., 2000, Effect of dissemination of 2,4-dichlorophenoxyacetic acid (2,4-D) degradation

plasmids on 2,4-D degradation and on bacterial community structure in two different soil horizons. *Appl. Environ. Microbiol.*, 66:3297–3304.
46. de Lipthay, J.R., Barkay, T., Vekova, J., and Sörensen, S.J., 1999, Utilization of phenoxyacetic acid, by strains using either the *ortho* or *meta* cleavage of catechol during phenol degradation, after conjugal transfer of *tfdA*, the gene encoding a 2,4-dichlorophenoxyacetic acid/2-oxoglutarate dioxygenase. *Appl. Microbiol. Biotechnol.*, 51:207–214.
47. Dennis, J.J., and Zylstra, G.J., 2004, Complete sequence and genetic organization of pDTG1, the 83 kilobase naphthalene degradation plasmid from *Pseudomonas putida* strain NICB 9816-4. *J. Mol. Biol.*, 341:753–768.
48. Denome, S.A., Stanley, D.C., Olson, E.S., and Young, K.D., 1993, Metabolism of dibenzothiophene and naphthalene in *Pseudomonas* strains: complete DNA sequence of an upper naphthalene catabolic pathway. *J. Bacteriol.*, 175:6890–6901.
49. Dobrindt, U., Hochhut, B., Hentschel, U., and Hacker, J., 2004, Genomic islands in pathogenic and environmental microorganisms. *Nat. Rev. Microbiol.*, 2:414–424.
50. Dogra, C., Raina, V., Pal, R., Suar, M., Lal, S., Gartemann, K.-H., Holliger, C., van der Meer, J.R., and Lal, R., 2004, Organization of *lin* genes and IS6100 among different strains of hexachlorocyclohexane-degrading *Sphingomonas paucimobilis*-evidence for horizontal gene transfer. *J. Bacteriol.*, 186:2225–2235.
51. Don, R.H., and Pemberton, J.M., 1981, Properties of six pesticide degradation plasmids isolated from *Alcaligenes paradoxus* and *Alcaligenes eutrophus*. *J. Bacteriol.*, 145:681–686.
52. Dorn, E., Hellwig, M., Reineke, W., and Knackmuss, H.-J., 1974, Isolation and characterization of a 3-chlorobenzoate degrading pseudomonad. *Arch. Microbiol.*, 99:61–70.
53. Drake, J.W., 1991, Spontaneous mutation. *Annu. Rev. Genet.*, 25:124–146.
54. Eaton, R.W., 1997, *p*-Cymene catabolic pathway in *Pseudomonas putida* F1: cloning and characterization of DNA encoding conversion of *p*-cymene to *p*-cumate. *J. Bacteriol.*, 179:3171–3180.
55. Eaton, R.W., and Timmis, K.N., 1986, Characterization of a plasmid-specified pathway for catabolism of isopropylbenzene in *Pseudomonas putida* RE204. *J. Bacteriol.*, 168:123–131.
56. Echols, H., and Goodman, M.F., 1991, Fidelity mechanisms in DNA replication. *Annu. Rev. Biochem.*, 60:477–511.
57. Eggink, G., Lageveen, R.G., Altenburg, B., and Witholt, B., 1987, Controlled and functional expression of the *Pseudomonas oleovorans* alkane utilizing system in *Pseudomonas putida* and*Escherichia coli*. *J. Biol. Chem.*, 262:17712–17718.
58. Fowler, R.G., and Schaaper, R.M., 1997, The role of the *mutT* gene of *Escherichia coli* in maintaining replication fidelity. *FEMS Microbiol. Rev.*, 21:43–54.
59. Franklin, F.C., Lehrbach, P.R., Lurz, R., Rueckert, B., Bagdasarian, M., and Timmis, K.N., 1983, Localization and functional analysis of transposon mutations in regulatory genes of the TOL catabolic pathway. *J. Bacteriol.*, 154:676–685.
60. Franklin, F.C.H., Bagdasarian, M., Bagdasarian, M.M., and Timmis, K.N., 1981, Molecular and functional analysis of the TOL plasmid pWW0 from *Pseudomonas putida* and cloning of genes for the entire regulated aromatic ring-cleavage pathway. *Proc. Natl. Acad. Sci. U.S.A.*, 78:7458–7462.
61. Fulthorpe, R.R., McGowan, C., Maltseva, O.V., Holben, W.E., and Tiedje, J.M., 1995, 2,4-Dichlorophenoxyacetic acid-degrading bacteria contain mosaics of catabolic genes. *Appl. Environ. Microbiol.*, 61:3274–3281.
62. Fulthorpe, R.R., Rhodes, A.N., and Tiedje, J.M., 1996, Pristine soils mineralize 3-chlorobenzoate and 2,4-dichlorophenoxyacetate via different microbial populations. *Appl. Environ. Microbiol.*, 62:1159–1166.
63. Genin, S., and Boucher, C., 2004, Lessons learned from the genome analysis of *Ralstonia solanacearum*. *Annu. Rev. Phytopathol.*, 42:107–134.

64. Gerischer, U., D'Argenio, D.A., and Ornston, L.N., 1996, IS*1236*, a newly discovered member of the IS*3* family, exhibits varied patterns of insertion into the *Acinetobacter calcoaceticus* chromosome. *Microbiology*, 142:1825–1831.
65. Gerischer, U., and Ornston, L.N., 1995, Spontaneous mutations in *pcaH* and -*G*, structural genes for protocatechuate 3,4-dioxygenase in *Acinetobacter calcoaceticus*. *J. Bacteriol.*, 177:1336–1347.
66. Göbel, M., Kranz, O.H., Kaschabek, S.R., Schmidt, E., Pieper, D.H., and Reineke, W., 2004, Microorganisms degrading chlorobenzene via a *meta*-cleavage pathway harbor highly similar chlorocatechol 2,3-dioxygenase-encoding gene clusters. *Arch. Microbiol.*, 182:147–156.
67. Greated, A., Lambertsen, L., Williams, P.A., and Thomas, C.M., 2002, Complete sequence of the IncP-9 TOL plasmid pWW0 from *Pseudomonas putida*. *Environ. Microbiol.*, 4:856–871.
68. Greub, G., Collyn, F., Guy, L., and Roten, C.A., 2004, A genomic island present along the bacterial chromosome of the *Parachlamydiaceae* UWE25, an obligate amoebal endosymbiont, encodes a potentially functional F-like conjugative DNA transfer system. *BMC Microbiol.*, 4:48.
69. Hacker, J., Blum-Oehler, G., Muhldorfer, I., and Tschape, H., 1997, Pathogenicity islands of virulent bacteria: structure, function and impact on microbial evolution. *Mol. Microbiol.*, 23:1089–1097.
70. Haigler, B.E., Nishino, S.F., and Spain, J.C., 1988, Degradation of 1,2-dichlorobenzene by a *Pseudomonas* sp. *Appl. Environ. Microbiol.*, 54:294–301.
71. Hall, R.M., and Stokes, H.W., 1993, Integrons: novel DNA elements which capture genes by site-specific recombination. *Genetica*, 90:115–132.
72. Harayama, S., 1994, Codon usage patterns suggest independent evolution of two catabolic operons on toluene-degradative plasmid TOL pWW0 of*Pseudomonas putida*. *J. Mol. Evol.*, 38:328–335.
73. Harayama, S., Lehrbach, P.R., and Timmis, K.N., 1984, Transposon mutagenesis analysis of *meta*-cleavage pathway operon genes of the TOL plasmid of *Pseudomonas putida* mt-2. *J. Bacteriol.*, 160:251–255.
74. Harayama, S., Leppik, R.A., Rekik, M., Mermod, N., Lehrbach, P.R., Reineke, W., and Timmis, K.N., 1986, Gene order of the TOL catabolic plasmid upper pathway operon and oxidation of both toluene and benzyl alcohol by the *xylA* product. *J. Bacteriol.*, 167:455–461.
75. Harayama, S., Mermod, N., Rekik, M., Lehrbach, P.R., and Timmis, K.N., 1987, Roles of the divergent branches of the *meta*-cleavage pathway in the degradation of benzoate and substituted benzoates. *J. Bacteriol.*, 169:558–564.
76. Harayama, S., and Rekik, M., 1990, The *meta* cleavage operon of TOL degradative plasmid pWW0 comprises 13 genes. *Mol. Gen. Genet.*, 221:113–120.
77. Harayama, S., and Rekik, M., 1993, Comparison of the nucleotide sequences of the *meta*-cleavage pathway genes of TOL plasmid pWW0 from *Pseudomonas putida* with other *meta*-cleavage genes suggests that both single and multiple nucleotide substitutions contribute to enzyme evolution. *Mol. Gen. Genet.*, 239:81–89.
78. Harayama, S., Rekik, M., Wasserfallen, A., and Bairoch, A., 1987, Evolutionary relationships between catabolic pathways for aromatics: conservation of gene order and nucleotide sequences of catechol oxidation genes of pWW0 and NAH7 plasmids. *Mol. Gen. Genet.*, 210:241–247.
79. Harwood, C.S., and Parales, R.E., 1996, The β-ketoadipate pathway and the biology of self-identity. *Annu. Rev. Microbiol.*, 50:553–590.
80. Hayase, N., Taira, K., and Furukawa, K., 1990, *Pseudomonas putida* KF715 *bphABCD* operon encoding biphenyl and polychlorinated biphenyl degradation: cloning, analysis, and expression in soil bacteria. *J. Bacteriol.*, 172:1160–1164.

81. Herrick, J.B., Stuart-Keil, K.G., Ghiorse, W.C., and Madsen, E.L., 1997, Natural horizontal transfer of a naphthalene dioxygenase gene between bacteria native to a coal tar-contaminated field site. *Appl. Environ. Microbiol.*, 63:2330–2337.
82. Heuer, H., Szczepanowski, R., Schneiker, S., Pühler, A., Top, E.M., and Schlüter, A., 2004, The complete sequences of plasmids pB2 and pB3 provide evidence for a recent ancestor of the IncP-1β group without any accessory genes. *Microbiology*, 150:3591–3599.
83. Hoffmann, D., Kleinsteuber, S., Müller, R.H., and Babel, W., 2003, A transposon encoding the complete 2,4-dichlorophenoxyacetic acid degradation pathway in the alkalitolerant strain *Delftia acidovorans* P4a. *Microbiology*, 149:2545–2556.
84. Hong, H.B., Nam, I.H., Murugesan, K., Kim, Y.M., and Chang, Y.S., 2004, Biodegradation of dibenzo-*p*-dioxin, dibenzofuran, and chlorodibenzo-*p*-dioxins by *Pseudomonas veronii* PH-03. *Biodegradation*, 15:303–313.
85. Hooper, S.W., Dockendorff, T.C., and Sayler, G.S., 1989, Characteristics and restriction analysis of the 4-chlorobiphenyl catabolic plasmid, pSS50. *Appl. Environ. Microbiol.*, 55:1286–1288.
86. Hughes, E.J., Bayly, R.C., and Skurray, R.A., 1984, Characterization of a TOL-like plasmid from *Alcaligenes eutrophus* that controls expression of a chromosomally encoded *p*-cresol pathway. *J. Bacteriol.*, 158:73–78.
87. Inouye, S., Nakazawa, A., and Nakazawa, T., 1983, Molecular cloning of regulatory gene *xylR* and operator-promoter regions of the *xylABC* and *xylDEGF* operons of the TOL plasmid. *J. Bacteriol.*, 155:1192–1199.
88. James, K.D., Hughes, M.A., and Williams, P.A., 2000, Cloning and expression of *ntnD*, encoding a novel $NAD(P)^+$-independent 4-nitrobenzyl alcohol dehydrogenase from *Pseudomonas* sp. strain TW3. *J. Bacteriol.*, 182:3136–3141.
89. Jeenes, D.J., and Williams, P.A., 1982, Excision and integration of degradative pathway genes from TOL plasmid pWW0. *J. Bacteriol.*, 150:188–194.
90. Jeong, J.J., Kim, J.H., Kim, C.K., Hwang, I., and Lee, K., 2003, 3- and 4-Alkylphenol degradation pathway in *Pseudomonas* sp. strain KL28: genetic organization of the *lap* gene cluster and substrate specificities of phenol hydroxylase and catechol 2,3-dioxygenase. *Microbiology*, 149:3265–3277.
91. Jiménez, J.I., Miñambres, B., García, J.L., and Díaz, E., 2002, Genomic analysis of the aromatic catabolic pathways from *Pseudomonas putida* KT2440. *Environ. Microbiol.*, 4:824–841.
92. Johnson, G.R., Jain, R.K., and Spain, J.C., 2002, Origins of the 2,4-dinitrotoluene pathway. *J. Bacteriol.*, 184:4219–4232.
93. Johnson, G.R., and Spain, J.C., 2003, Evolution of catabolic pathways for synthetic compounds: bacterial pathways for degradation of 2,4-dinitrotoluene and nitrobenzene. *Appl. Microbiol. Biotechnol.*, 62:110–123.
94. Ka, J.O., Holben, W.E., and Tiedje, J.M., 1994, Genetic and phenotypic diversity of 2,4-dichlorophenoxyacetic acid (2,4-D)-degrading bacteria isolated from 2,4-D-treated field soils. *Appl. Environ. Microbiol.*, 60:1106–1115.
95. Kado, C.I., 1998, Origin and evolution of plasmids. *Antonie van Leeuwenhoek Int. J. Microbiol.*, 73:117–126.
96. Kallastu, A., Horak, R., and Kivisaar, M., 1998, Identification and characterization of IS1411, a new insertion sequence which causes transcriptional activation of the phenol degradation genes in *Pseudomonas putida*. *J. Bacteriol.*, 180:5306–5312.
97. Kamagata, Y., Fulthorpe, R.R., Tamura, K., Takami, H., Forney, L.J., and Tiedje, J.M., 1997, Pristine environments harbor a new group of oligotrophic 2,4-dichlorophenoxyacetic acid-degrading bacteria. *Appl. Environ. Microbiol.*, 63:2266–2272.
98. Keil, H., Lebens, M.R., and Williams, P.A., 1985, TOL plasmid pWW15 contains two nonhomologous, independently regulated catechol 2,3-oxygenase genes. *J. Bacteriol.*, 163:248–255.

99. Klecka, G.M., and Gibson, D.T., 1981, Inhibition of catechol 2,3-dioxygenase from *Pseudomonas putida* by 3-chlorocatechol. *Appl. Environ. Microbiol.*, 41:1159–1165.
100. Klockgether, J., Reva, O., Larbig, K., and Tummler, B., 2004, Sequence analysis of the mobile genome island pKLC102 of *Pseudomonas aeruginosa* C. *J. Bacteriol.*, 186:518–534.
101. Kohler, H.-P.E., Schmid, A., and van der Maarel, M., 1993, Metabolism of 2,2′-dihydroxybiphenyl by *Pseudomonas* sp. strain HBP1: production and consumption of 2,2′,3-trihydroxybiphenyl. *J. Bacteriol.*, 175:1621–1628.
102. Kunz, D.A., and Chapman, P.J., 1981, Isolation and characterization of spontaneously occurring TOL plasmid mutants of *Pseudomonas putida* HS1. *J. Bacteriol.*, 146:952–964.
103. Laemmli, C.M., Leveau, J.H.J., Zehnder, A.J.B., and van der Meer J.R., 2000, Characterization of a second *tfd* gene cluster for chlorophenol and chlorocatechol metabolism on plasmid pJP4 in *Ralstonia eutropha* JMP134(pJP4). *J. Bacteriol.*, 182:4165–4172.
104. Laemmli, C.M., Schönenberger, R., Suter, M.J.-F., Zehnder, A.J.B., and van der Meer J.R., 2002, TfdDII, one of the two chloromuconate cycloisomerases of *Ralstonia eutropha* JMP134 (pJP4), cannot efficiently convert 2-chloro-*cis*,*cis*-muconate to *trans*-dienelactone to allow growth on 3-chlorobenzoate. *Arch. Microbiol.*, 178:13–25.
105. Laemmli, C.M., Werlen, C., and van der Meer, J.R., 2004, Mutational analysis of the two *tfd* clusters for chloroaromatic degradation in *Ralstonia eutropha* JMP134. *Arch. Microbiol.*, 181:112–121.
106. Larbig, K., Christmann, A., Johann, A., Klockgether, J., Hartsch, T., Merkl, R., Wiehlmann, L., Fritz, H.-J., and Tümmler, B., 2002, Gene islands integrated into $tRNA^{Gly}$ genes confer genome diversity on a *Pseudomonas aeruginosa* clone. *J. Bacteriol.*, 184:6665–6680.
107. Lehrbach, P.R., Ward, J., Meulien, P., and Broda, P., 1982, Physical mapping of TOL plasmids pWWO and pND2 and various R plasmid-TOL derivatives from *Pseudomonas* spp. *J. Bacteriol.*, 152:1280–1283.
108. Lenski, R.E., Mongold, J.A., Sniegowski, P.D., Travisano, M., Vasi, F., Gerrish, P.J., and Schmidt, T.M., 1998, Evolution of competitive fitness in experimental populations of *E. coli*: what makes one genotype a better competitor than another? *Antonie van Leeuwenhoek Int. J. Microbiol.*, 73:35–47.
109. Lesic, B., Bach, S., Ghigo, J.M., Dobrindt, U., Hacker, J., and Carniel, E., 2004, Excision of the high-pathogenicity island of *Yersinia pseudotuberculosis* requires the combined actions of its cognate integrase and Hef, a new recombination directionality factor. *Mol. Microbiol.*, 52:1337–1348.
110. Leveau, J.H.J., and van der Meer, J.R., 1997, Genetic characterization of insertion sequence IS*JP4* on plasmid pJP4 from *Ralstonia eutropha* JMP134. *Gene*, 202:103–114.
111. Li, W., Shi, J., Wang, X., Han, Y., Tong, W., Ma, L., Liu, B., and Cai, B., 2004, Complete nucleotide sequence and organization of the naphthalene catabolic plasmid pND6-1 from *Pseudomonas* sp. strain ND6. *Gene*, 336:231–240.
112. Lorenz, M.G., and Wackernagel, W., 1994, Bacterial gene transfer by natural genetic transformation in the environment. *Microbiol. Rev.*, 58:563–602.
113. Maeda, K., Nojiri, H., Shintani, M., Yoshida, T., Habe, H., and Omori, T., 2003, Complete nucleotide sequence of carbazole/dioxin-degrading plasmid pCAR1 in *Pseudomonas resinovorans* strain CA10 indicates its mosaicity and the presence of large catabolic transposon Tn*4676*. *J. Mol. Biol.*, 326:21–33.
114. Mahillon, J., and Chandler, M., 1998, Insertion sequences. *Microbiol. Mol. Biol. Rev.*, 62:725–774.
115. Mahillon, J., Leonard, C., and Chandler, M., 1999, IS elements as constituents of bacterial genomes. *Res. Microbiol.*, 150:675–687.
116. Mars, A.E., Kasberg, T., Kaschabek, S.R., van Agteren, M.H., Janssen, D.B., and Reineke, W., 1997, Microbial degradation of chloroaromatics: use of the *meta*-cleavage pathway for mineralization of chlorobenzene. *J. Bacteriol.*, 179:4530–4537.

117. Mars, A.E., Kingma, J., Kaschabek, S.R., Reineke, W., and Janssen, D.B., 1999, Conversion of 3-chlorocatechol by various catechol 2,3-dioxygenases and sequence analysis of the chlorocatechol dioxygenase region of *Pseudomonas putida* GJ31. *J. Bacteriol.*, 181:1309–1318.
118. Martin, V.J., and Mohn, W.W., 2000, Genetic investigation of the catabolic pathway for degradation of abietane diterpenoids by *Pseudomonas abietaniphila* BKME-9. *J. Bacteriol.*, 182:3784–3793.
119. Martinez, B., Tomkins, J., Wackett, L.P., Wing, R., and Sadowsky, M.J., 2001, Complete nucleotide sequence and organization of the atrazine catabolic plasmid pADP-1 from *Pseudomonas* sp. strain ADP. *J. Bacteriol.*, 183:5684–5697.
120. Martins dos Santos, V.A.P., Heim, S., Moore, E.R.B., Strätz, M., and Timmis, K.N., 2004, Insights into the genomic basis of niche specificity of *Pseudomonas putida* KT2440. *Environ. Microbiol.*, 6:1264–1286.
121. Meier, P., and Wackernagel, W., 2003, Mechanisms of homology-facilitated illegitimate recombination for foreign DNA acquisition in transformable *Pseudomonas stutzeri*. *Mol. Microbiol.*, 48:1107–1118.
122. Merlin, C., Springael, D., and Toussaint, A., 1999, Tn4371: a modular structure encoding a phage-like integrase, a *Pseudomonas*-like catabolic pathway, and RP4/Ti-like transfer functions. *Plasmid*, 41:40–54.
123. Müller, T.A., Werlen, C., Spain, J., and van der Meer, J.R., 2003, Evolution of a chlorobenzene degradative pathway among bacteria in a contaminated groundwater mediated by a genomic island in *Ralstonia*. *Environ. Microbiol.*, 5:163–173.
124. Neilson, A.H., 1990, The biodegradation of halogenated organic compounds. *J. Appl. Bacteriol.*, 69:445–470.
125. Newby, D.T., Gentry, T.J., and Pepper, I.L., 2000, Comparison of 2,4-dichlorophenoxyacetic acid degradation and plasmid transfer in soil resulting from bioaugmentation with two different pJP4 donors. *Appl. Environ. Microbiol.*, 66:3399–3407.
126. Nishi, A., Tominaga, K., and Furukawa, K., 2000, A 90-kilobase conjugative chromosomal element coding for biphenyl and salicylate catabolism in *Pseudomonas putida* KF715. *J. Bacteriol.*, 182:1949–1955.
127. Nordlund, I., Powlowski, J., and Shingler, V., 1990, Complete nucleotide sequence and polypeptide analysis of multicomponent phenol hydroxylase from *Pseudomonas* sp. strain CF600. *J. Bacteriol.*, 172:6826–6833.
128. Nunes, L.R., Rosato, Y.B., Muto, N.H., Yanai, G.M., da Silva, V.S., Leite, D.B., Gonçalves, E.R., de Souza, A.A., Coletta-Filho, H.D., Machado, M.A., Lopes, S.A., and de Oliviera, R.C., 2003, Microarray analyses of *Xylella fastidiosa* provide evidence of coordinated transcription control of laterally transferred elements. *Genome Res.*, 13:570–578.
129. Ochman, H., Lawrence, J.G., and Groisman, E.A., 2000, Lateral gene transfer and the nature of bacterial innovation. *Nature*, 405:299–304.
130. O'Donnell, K.J., and Williams, P.A., 1991, Duplication of both *xyl* catabolic operons on TOL plasmid pWW15. *J. Gen. Microbiol.*, 137:2831–2838.
131. Oelschlaeger, T.A., Dobrindt, U., Janke, B., Middendorf, B., Karch, H., and Hacker, J., 2003, Analysis of pathogenicity islands of STEC. *Methods Mol. Med.*, 73:99–112.
132. Ogawa, N., and Miyashita, K., 1999, The chlorocatechol-catabolic transposon Tn5707 of *Alcaligenes eutrophus* NH9, carrying a gene cluster highly homologous to that in the 1,2,4-trichlorobenzene-degrading bacterium *Pseudomonas* sp. strain P51, confers the ability to grow on 3-chlorobenzoate. *Appl. Environ. Microbiol.*, 65:724–731.
133. O'Leary, N.D., O'Connor, K.E., and Dobson, A.D., 2002, Biochemistry, genetics and physiology of microbial styrene degradation. *FEMS Microbiol. Rev.*, 26:403–417.
134. Osborne, D.J., Pickup, R.W., and Williams, P.A., 1988, The presence of two complete homologous *meta* pathway operons on TOL plasmid pWW53. *J. Gen. Microbiol.*, 134:2965–2975.

135. Ouchiyama, N., Miyachi, S., and Omori, T., 1998, Cloning and nucleotide sequence of carbazole catabolic genes from *Pseudomonas stutzeri* strain OM1, isolated from activated sludge. *J. Gen. Appl. Microbiol.*, 44:57–63.
136. Overhage, J., Kresse, A.U., Priefert, H., Sommer, H., Krammer, G., Rabenhorst, J., and Steinbuchel, A., 1999, Molecular characterization of the genes *pcaG* and *pcaH*, encoding protocatechuate 3,4-dioxygenase, which are essential for vanillin catabolism in *Pseudomonas* sp. strain HR199. *Appl. Environ. Microbiol.*, 65:951–960.
137. Paget, E., and Simonet, P., 1994, On the track of natural transformation in soil. *FEMS Microbiol. Ecol.*, 15:109–118.
138. Park, H.S., and Kim, H.S., 2000, Identification and characterization of the nitrobenzene catabolic plasmids pNB1 and pNB2 in *Pseudomonas putida* HS12. *J. Bacteriol.*, 182:573–580.
139. Pérez-Pantoja, D., Guzman, L., Manzano, M., Pieper, D.H., and Gonzalez, B., 2000, Role of $tfdC_I D_I E_I F_I$ and $tfdD_{II} C_{II} E_{II} F_{II}$ gene modules in catabolism of 3-chlorobenzoate by *Ralstonia eutropha* JMP134(pJP4). *Appl. Environ. Microbiol.*, 66:1602–1608.
140. Pérez-Pantoja, D., Ledger, T., Pieper, D.H., and Gonzalez, B., 2003, Efficient turnover of chlorocatechols is essential for growth of *Ralstonia eutropha* JMP134 (pJP4) in 3-chlorobenzoic acid. *J. Bacteriol.*, 185:1534–1542.
141. Perna, N.T., Plunkett, G., III, Burland, V., Mau, B., Glasner, J.D., Rose, D.J., Mayhew, G.F., Evans, P.S., Gregor, J., Kirkpatrick, H.A., Posfai, G., Hackett, J., Klink, S., Boutin, A., Shao, Y., Miller, L., Grotbeck, E.J., Davis, N.W., Lim, A., Dimalanta, E.T., Potamousis, K.D., Apodaca, J., Anantharaman, T.S., Lin, J., Yen, G., Schwartz, D.C., Welch, R.A., and Blattner, F.R., 2001, Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. *Nature*, 409:529–533.
142. Peters, M., Heinaru, E., Talpsep, E., Wand, H., Stottmeister, U., Heinaru, A., and Nurk, A., 1997, Acquisition of a deliberately introduced phenol degradation operon, *pheBA*, by different indigenous *Pseudomonas* species. *Appl. Environ. Microbiol.*, 63:4899–4906.
143. Pettigrew, C.A., Haigler, B.E., and Spain, J.C., 1991, Simultaneous biodegradation of chlorobenzene and toluene by a *Pseudomonas* strain. *Appl. Environ. Microbiol.*, 57:157–162.
144. Pickup, R.W., and Williams, P.A., 1982, Spontaneous deletions in the TOL plasmid pWW20 which give rise to the B3 regulatory mutants of *Pseudomonas putida* MT20. *J. Gen. Microbiol.*, 128:1385–1390.
145. Platt, A., Shingler, V., Taylor, S.C., and Williams, P.A., 1995, The 4-hydroxy-2-oxovalerate aldolase and acetaldehyde dehydrogenase (acylating) encoded by the *nahM* and *nahO* genes of the naphthalene catabolic plasmid pWW60-22 provide further evidence of conservation of *meta*-cleavage pathway gene sequences. *Microbiology*, 141:2223–2233.
146. Plumeier, I., Pérez-Pantoja, D., Heim, S., Gonzalez, B., and Pieper, D.H., 2002, Importance of different *tfd* genes for degradation of chloroaromatics by *Ralstonia eutropha* JMP134 (pJP4) on 3-chlorobenzoate. *J. Bacteriol.*, 184:4054–4064.
147. Poelarends, G.J., Kulakov, L.A., Larkin, M.J., van Hylckama Vlieg, J.E., and Janssen, D.B., 2000, Roles of horizontal gene transfer and gene integration in evolution of 1,3-dichloropropene- and 1,2-dibromoethane-degradative pathways. *J. Bacteriol.*, 182:2191–2199.
148. Poh, R.P.-C., Smith, A.R.W., and Bruce, I.J., 2002, Complete characterization of Tn*5530* from *Burkholderia cepacia* strain 2a(pIJB1) and studies of 2,4-dichlorophenoxyacetate uptake by the organism. *Plasmid*, 48:1–12.
149. Potrawfke, T., Armengaud, J., and Wittich, R.M., 2001, Chlorocatechols substituted at positions 4 and 5 are substrates of the broad-spectrum chlorocatechol 1,2-dioxygenase of *Pseudomonas chlororaphis* RW71. *J. Bacteriol.*, 183:997–1011.
150. Potrawfke, T., Timmis, K.N., and Wittich, R.M., 1998, Degradation of 1,2,3,4-tetrachlorobenzene by *Pseudomonas chlororaphis* RW71. *Appl. Environ. Microbiol.*, 64: 3798–3806.

151. Powlowski, J., and Shingler, V., 1994, Genetics and biochemistry of phenol degradation by *Pseudomonas* sp. CF600. *Biodegradation*, 5:219–236.
152. Pries, F., Vandenwijngaard, A.J., Bos, R., Pentenga, M., and Janssen, D.B., 1994, The role of spontaneous cap domain mutations in haloalkane dehalogenase specificity and evolution. *J. Biol. Chem.*, 269:17490–17494.
153. Rakin, A., Noelting, C., Schropp, P., and Heesemann, J., 2001, Integrative module of the high-pathogenicity island of *Yersinia*. *Mol. Microbiol.*, 39:407–415.
154. Ravatn, R., Studer, S., Springael, D., Zehnder, A., J.B., and van der Meer J.R., 1998, Chromosomal integration, tandem amplification, and deamplification in *Pseudomonas putida* F1 of a 105-kilobase genetic element containing the chlorocatechol degradative genes from *Pseudomonas* sp. strain B13. *J. Bacteriol.*, 180:4360–4369.
155. Ravatn, R., Studer, S., Zehnder, A.J.B., and van der Meer, J.R., 1998, Int-B13, an unusual site-specific recombinase of the bacteriophage P4 integrase family, is responsible for chromosomal insertion of the 105-kilobase *clc* element of *Pseudomonas* sp. strain B13. *J. Bacteriol.*, 180:5505–5514.
156. Ravatn, R., Zehnder, A.J.B., and van der Meer, J.R., 1998, Low-frequency horizontal transfer of an element containing the chlorocatechol degradation genes from *Pseudomonas* sp. strain B13 to *Pseudomonas putida* F1 and to indigenous bacteria in laboratory-scale activated-sludge microcosms. *Appl. Environ. Microbiol.*, 64:2126–2132.
157. Reams, A.B., and Neidle, E.L., 2004, Gene amplification involves site-specific short homology-independent illegitimate recombination in *Acinetobacter* sp. strain ADP1. *J. Mol. Biol.*, 338:643–656.
158. Reddy, B.R., Shaw, L.E., Sayers, J.R., and Williams, P.A., 1994, Two identical copies of IS*1246*, a 1275 base pair sequence related to other bacterial insertion sequences, enclose the *xyl* genes on TOL plasmid pWW0. *Microbiology*, 140:2305–2307.
159. Reich, T., Schmidt, S., and Fortnagel, P., 1999, Metabolism of phenylbenzoate by *Pseudomonas* sp. strain TR3. *FEMS Microbiol. Lett.*, 176:477–482.
160. Reineke, W., 1984, Microbial degradation of halogenated aromatic compounds, pp. 319–360. In D.T. Gibson (ed.), Microbial degradation of organic compounds. Marcel Dekker Inc., New York.
161. Reineke, W., Jeenes, D.J., Williams, P.A., and Knackmuss, H.-J., 1982, TOL plasmid pWW0 in constructed halobenzoate-degrading *Pseudomonas* strains: prevention of *meta* pathway. *J. Bacteriol.*, 150:195–201.
162. Reineke, W., and Knackmuss, H.-J., 1979, Construction of haloaromatics utilizing bacteria. *Nature (Lond.)*, 277:385–386.
163. Reineke, W., and Knackmuss, H.-J., 1984, Microbial metabolism of haloaromatics: isolation and properties of a chlorobenzene-degrading bacterium. *Appl. Environ. Microbiol.*, 47:395–402.
164. Rippen, G., 1995, *Handbuch der Umweltchemikalien: Stoffdaten – Prüfverfahren – Vorschriften*. Ecomed, Landsber/Lech.
165. Romine, M.F., Stillwell, L.C., Wong, K.K., Thurston, S.J., Sisk, E.C., Sensen, C., Gaasterland, T., Fredrickson, J.K., and Saffer, J.D., 1999, Complete sequence of a 184-kilobase catabolic plasmid from *Sphingomonas aromaticivorans* F199. *J. Bacteriol*., 181:1585–1602.
166. Rousseaux, S., Hartmann, A., and Soulas, G., 2001, Isolation and characterisation of new Gram-negative and Gram-positive atrazine degrading bacteria from different French soils. *FEMS Microbiol. Ecol.*, 36:211–222.
167. Rowe-Magnus, D.A., Guerout, A.M., Ploncard, P., Dychinco, B., Davies, J., and Mazel, D., 2001, The evolutionary history of chromosomal super-integrons provides an ancestry for multiresistant integrons. *Proc. Natl. Acad. Sci. USA*, 98:652–657.

168. Saint, C.P., McClure, C., and Venables, W.A., 1990, Physical map of the aromatic amine and *m*-toluate catabolic plasmid pTDN1 in *Pseudomonas putida*: location of a unique *meta*-cleavage pathway. *J. Gen. Microbiol.*, 136:615–625.
169. Sajjaphan, K., Shapir, N., Wackett, L.P., Palmer, M., Blackmon, B., Tomkins, J., and Sadowsky, M.J., 2004, *Arthrobacter aerescens* TC1 atrazine catabolism genes *trzN*, *atzB*, and *atzC* are linked on a 160-kilobase region and are functional in *Escherichia coli*. *Appl. Environ. Microbiol.*, 70:4402–4407.
170. Sanseverino, J., Applegate, B.M., King, J.M., and Sayler, G.S., 1993, Plasmid-mediated mineralization of naphthalene, phenanthrene, and anthracene. *Appl. Environ. Microbiol.*, 59:1931–1917.
171. Saumaa, S., Tover, A., Kasak, L., and Kivisaar, M., 2002, Different spectra of stationary-phase mutations in early-arising versus late-arising mutants of *Pseudomonas putida*: involvement of the DNA repair enzyme MutY and the stationary-phase sigma factor RpoS. *J. Bacteriol.*, 184:6957–6965.
172. Savard, P., Péloquin, L., and Sylvestre, M., 1986, Cloning of *Pseudomonas* sp. strain CBS3 genes specifying dehalogenation of 4-chlorobenzoate. *J. Bacteriol.*, 168:81–85.
173. Sayler, G.S., Hooper, S.W., Layton, A.C., and King, J.M.H., 1990, Catabolic plasmids of environmental and ecological significance. *Microbiol. Ecol.*, 19:1–20.
174. Schlömann, M., 1994, Evolution of chlorocatechol catabolic pathways. Conclusions to be drawn from comparisons of lactone hydrolases. *Biodegradation*, 5:301–321.
175. Schlüter, A., Heuer, H., Szczepanowski, R., Forney, L.J., Thomas, C.M., Pühler, A., and Top, E.M., 2003, The 64 508 bp IncP-1b antibiotic multiresistance plasmid pB10 isolated from a waste-water treatment plant provides evidence for recombination between members of different branches of the IncP-1β group. *Microbiology*, 149:3139–3153.
176. Schubert, S., Dufke, S., Sorsa, J., and Heesemann, J., 2004, A novel integrative and conjugative element (ICE) of *Escherichia coli*: the putative progenitor of the *Yersinia* high-pathogenicity island. *Mol. Microbiol.*, 51:837–848.
177. Sentchilo, V.S., Perebituk, A.N., Zehnder, A.J., and van der Meer, J.R., 2000, Molecular diversity of plasmids bearing genes that encode toluene and xylene metabolism in *Pseudomonas* strains isolated from different contaminated sites in Belarus. *Appl. Environ. Microbiol.*, 66:2842–2852.
178. Sentchilo, V.S., Ravatn, R., Werlen, C., Zehnder, A.J.B., and van der Meer, J.R., 2003, Unusual expression of the integrase gene in the *clc* genomic island of *Pseudomonas* sp. strain B13. *J. Bacteriol.*, 185:4530–4538.
179. Sentchilo, V.S., Zehnder, A.J.B., and van der Meer, J.R., 2003, Characterization of two alternative promoters and a transcription regulator for integrase expression in the *clc* catabolic genomic island of *Pseudomonas* sp. strain B13. *Mol. Microbiol.*, 49:93–104.
180. Shaw, L.E., and Williams, P.A., 1988, Physical and functional mapping of two cointegrate plasmids derived from RP4 and TOL plasmid pDK1. *J. Gen. Microbiol.*, 134:2463–2474.
181. Sinclair, M.I., and Holloway, B.W., 1991, Chromosomal insertion of TOL transposons in *Pseudomonas aeruginosa* PAO. *J. Gen. Microbiol.*, 137:1111–1120.
182. Sinclair, M.I., Maxwell, P.C., Lyon, B.R., and Holloway, B.W., 1986, Chromosomal location of TOL plasmid DNA in *Pseudomonas putida*. *J. Bacteriol.*, 168:1302–1308.
183. Smith, C.A., and Thomas, C.M., 1987, Comparison of the organization of the genomes of phenotypically diverse plasmids of incompatibility group P: members of the IncP β subgroup are closely related. *Mol. Gen. Genet.*, 206:419–427.
184. Smith, G.R., 1988, Homologous recombination in procaryotes. *Microbiol. Rev.*, 52:1–28.
185. Smith, M.A., Weaver, V.B., Young, D.M., and Ornston, L.N., 2003, Genes for chlorogenate and hydroxycinnamate catabolism (*hca*) are linked to functionally related genes in the *dca-pca-qui-pob-hca* chromosomal cluster of *Acinetobacter* sp. strain ADP1. *Appl. Environ. Microbiol.*, 69:524–532.

186. Smits, T.H., Balada, S.B., Witholt, B., and van Beilen, J.B., 2002, Functional analysis of alkane hydroxylases from gram-negative and gram-positive bacteria. *J. Bacteriol.*, 184:1733–1742.
187. Somerville, C.C., Bauer, S., Brininstool, G., Facette, M., Hamann, T., Milne, J., Osborne, E., Paredez, A., Persson, S., Raab, T., Vorwerk, S., and Youngs, H., 2004, Towards a systems approach to understanding plant cell walls. *Science*, 306:2206–2211.
188. Somerville, C.C., Nishino, S.F., and Spain, J.C., 1995, Purification and characterization of nitrobenzene nitroreductase from *Pseudomonas pseudoalcaligenes* JS45.*J. Bacteriol.*, 177:3837–3842.
189. Sota, M., Kawasaki, H., and Tsuda, M., 2003, Structure of haloacetate-catabolic IncP-1 β plasmid pUO1 and genetic mobility of its residing haloacetate-catabolic transposon. *J. Bacteriol.*, 185:6741–6745.
190. Spain, J.C., 1995, Biodegradation of nitroaromatic compounds. *Annu. Rev. Microbiol.*, 49:523–555.
191. Spain, J.C., and Nishino, S.F., 1987, Degradation of 1,4-dichlorobenzene by a *Pseudomonas* sp. *Appl. Environ. Microbiol.*, 53:1010–1019.
192. Spooner, R.A., Lindsay, K., and Franklin, F.C., 1986, Genetic, functional and sequence analysis of the *xylR* and *xylS* regulatory genes of the TOL plasmid pWW0. *J. Gen. Microbiol.*, 132:1347–1358.
193. Springael, D., Kreps, S., and Mergeay, M., 1993, Identification of a catabolic transposon, Tn*4371*, carrying biphenyl and 4-chlorobiphenyl degradation genes in *Alcaligenes eutrophus* A5. *J. Bacteriol.*, 175:1674–1681.
194. Springael, D., Peys, K., Ryngaert, A., Van Roy, S., Hooyberghs, L., Ravatn, R., Heyndrickx, M., van der Meer, J.R., Vandecasteele, C., Mergeay, M., and Diels, L., 2002, Community shifts in a seeded 3-chlorobenzoate degrading membrane biofilm reactor: indications for involvement of in situ horizontal transfer of the *clc*-element from inoculum to contaminant bacteria. *Environ. Microbiol.*, 4:70–80.
195. Springael, D., Ryngaert, A., Merlin, C., Toussaint, A., and Mergeay, M., 2001, Occurrence of Tn*4371*-related mobile elements and sequences in (chloro)biphenyl-degrading bacteria. *Appl. Environ. Microbiol.*, 67:42–50.
196. Springael, D., and Top, E.M., 2004, Horizontal gene transfer and microbial adaptation to xenobiotics: new types of mobile genetic elements and lessons from ecological studies. *Trends Microbiol.*, 12:53–56.
197. Stapleton, R.D., Bright, N.G., and Sayler, G.S., 2000, Catabolic and genetic diversity of degradative bacteria from fuel-hydrocarbon contaminated aquifers. *Microbiol. Ecol.*, 39:211–221.
198. Stapleton, R.D., and Sayler, G.S., 1998, Assessment of the microbiological potential for the natural attenuation of petroleum hydrocarbons in a shallow aquifer system. *Microbiol. Ecol.*, 36:349–361.
199. Stover, C.K., Pham, X.Q., Erwin, A.L., Mizoguchi, S.D., Warrener, P., Hickey, M.J., Brinkman, F.S., Hufnagle, W.O., Kowalik, D.J., Lagrou, M., Garber, R.L., Goltry, L., Tolentino, E., Westbrock-Wadman, S., Yuan, Y., Brody, L.L., Coulter, S.N., Folger, K.R., Kas, A., Larbig, K., Lim, R., Smith, K., Spencer, D., Wong, G.K., Wu, Z., Paulsen, I.T., Reizer, J., Saier, M.H., Hancock, R.E., Lory, S., and Olson, M.V., 2000, Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen. *Nature*, 406:959–964.
200. Summers, D.K., 1994, The origins and consequences of genetic instability in prokaryotes. *Dev. Biol. Stand.*, 83:7–11.
201. Taira, K., Hirose, J., Hayashida, S., and Furukawa, K., 1992, Analysis of *bph* operon from the polychlorinated biphenyl-degrading strain of *Pseudomonas pseudoalcaligenes* KF707. *J. Biol. Chem.*, 267:4844–4853.
202. Tan, H.-M., 1999, Bacterial catabolic transposons. *Appl. Microbiol. Biotechnol.*, 51:1–12.

203. Tan, H.-M., and Mason, J.R., 1990, Cloning and expression of the plasmid-encoded benzene dioxygenase genes from *Pseudomonas putida* ML2. *FEMS Microbiol. Lett.*, 72:259–264.
204. Thomas, C.M. (ed.), 2000, *The Horizontal Gene Pool*. Harwood academic publishers, Singapore.
205. Thorsted, P.B., Macartney, D.P., Akhtar, P., Haines, A.S., Ali, N., Davidson, P., Stafford, T., Pocklington, M.J., Pansegrau, W., Wilkins, B.M., Lanka, E., and Thomas, C.M., 1998, Complete sequence of the IncPbeta plasmid R751: implications for evolution and organisation of the IncP backbone. *J. Mol. Biol.*, 282:969–990.
206. Top, E., Vanrolleghem, P., Mergeay, M., and Verstraete, W., 1992, Determination of the mechanism of retrotransfer by mechanistic mathematical modeling. *J. Bacteriol.*, 174:5953–5960.
207. Top, E.M., Holben, W.E., and Forney, L.J., 1995, Characterization of diverse 2,4-dichlorophenoxyacetic acid-degradative plasmids isolated from soil by complementation. *Appl. Environ. Microbiol.*, 61:1691–1698.
208. Top, E.M., and Springael, D., 2003, The role of mobile genetic elements in bacterial adaptation to xenobiotic organic compounds. *Curr. Opin. Biotechnol.*, 14:262–269.
209. Topp, E., Zhu, H., Nour, S.M., Houot, S., Lewis, M., and Cuppels, D., 2000, Characterization of an atrazine-degrading *Pseudoaminobacter* sp. isolated from Canadian and French agricultural soils. *Appl. Environ. Microbiol.*, 66:2773–2782.
210. Toussaint, A., Merlin, C., Monchy, S., Benotmane, M.A., Leplae, R., Mergeay, M., and Springael, D., 2003, The biphenyl- and 4-chlorobiphenyl-catabolic transposon Tn*4371*, a member of a new family of genomic islands related to IncP and Ti plasmids. *Appl. Environ. Microbiol.*, 69:4837–4845.
211. Trefault, N., de la Iglesia, R., Molina, A.M., Manzano, M., Ledger, T., Pérez-Pantoja, D., Sánchez, M.A., Stuardo, M., and González, B., 2004, Genetic organization of the catabolic plasmid pJP4 from *Ralstonia eutropha* JMP134 (pJP4) reveals mechanisms of adaptation to chloroaromatic pollutants and evolution of specialized chloroaromatic degradation pathways. *Environ. Microbiol.* 6:655–668
212. Tsoi, T.V., Plotnikova, E.G., Cole, J.R., Guerin, W.F., Bagdasarian, M., and Tiedje, J.M., 1999, Cloning, expression, and nucleotide sequence of the *Pseudomonas aeruginosa* 142 *ohb* genes coding for oxygenolytic *ortho* dehalogenation of halobenzoates. *Appl. Environ. Microbiol.*, 65:2151–2162.
213. Tsuda, M., and Genka, H., 2001, Identification and characterization of Tn*4656*, a novel class II transposon carrying a set of toluene-degrading genes from TOL plasmid pWW53. *J. Bacteriol.*, 183:6215–6224.
214. Tsuda, M., and Iino, T., 1987, Genetic analysis of a transposon carrying toluene degrading genes on a TOL plasmid pWW0. *Mol. Gen. Genet.*, 210:270–276.
215. Tsuda, M., and Iino, T., 1988, Identification and characterization of Tn*4653*, a transposon covering the toluene transposon Tn*4651* on TOL plasmid pWW0. *Mol. Gen. Genet.*, 213:72–77.
216. Tsuda, M., Minegishi, K., and Iino, T., 1989, Toluene transposons Tn*4651* and Tn*4653* are class II transposons. *J. Bacteriol.*, 171:1386–1393.
217. Urata, M., Miyakoshi, M., Kai, S., Maeda, K., Habe, H., Omori, T., Yamane, H., and Nojiri, H., 2004, Transcriptional regulation of the *ant* operon, encoding two-component anthranilate 1,2-dioxygenase, on the carbazole-degradative plasmid pCAR1 of *Pseudomonas resinovorans* strain CA10. *J. Bacteriol.*, 186:6815–6823.
218. Vallaeys, T., Albino, L., Soulas, G., Wright, A., and Weightman, A.J., 1998, Isolation and characterization of a stable 2,4-dichlorophenoxyacetic acid degrading bacterium, *Variovorax paradoxus*, using chemostat culture. *Biotechnol. Lett.*, 20:1073–1076.
219. Vallaeys, T., Fulthorpe, R.R., Wright, A.M., and Soulas, G., 1996, The metabolic pathway of 2,4-dichlorophenoxyacetic acid degradation involves different families of *tfdA*

and *tfdB* genes according to PCR-RFLP analysis. *FEMS Microbiol. Ecol.*, 20:163–172.

220. van Beilen, J.B., Eggink, G., Enequist, H., Bos, R., and Witholt, B., 1992, DNA sequence determination and functional characterization of the OCT-plasmid-encoded *alkJKL* genes of *Pseudomonas oleovorans*. *Mol. Microbiol.*, 6:3121–3136.
221. van der Meer, J.R., 1997, Evolution of novel metabolic pathways for the degradation of aromatic compounds. *Antonie van Leeuwenhoek Int. J. Microbiol.*, 71:159–178.
222. van der Meer, J.R., 2002, Evolution of metabolic pathways for degradation of environmental pollutants., pp. 1194–1207. In G. Bitton (ed.), *Encyclopedia of Environmental Microbiology*. John Wiley & Sons, New York.
223. van der Meer, J.R., de Vos, W.M., Harayama, S., and Zehnder, A.J.B., 1992, Molecular mechanisms of genetic adaptation to xenobiotic compounds. *Microbiol. Rev.*, 56:677–694.
224. van der Meer, J.R., Eggen, R.I.L., Zehnder, A.J.B., and de Vos, W.M. 1991, Sequence analysis of the *Pseudomonas* sp. strain P51 *tcb* gene cluster, which encodes metabolism of chlorinated catechols: evidence for specialization of catechol 1,2-dioxygenases for chlorinated substrates. *J. Bacteriol.*, 173:2425–2434.
225. van der Meer, J.R., Roelofsen, W., Schraa, G., and Zehnder, A.J.B., 1987, Degradation of low concentrations of dichlorobenzenes and 1,2,4-trichlorobenzene by *Pseudomonas* sp. strain P51 in nonsterile soil columns. *FEMS Microbiol. Ecol.*, 45:333–341.
226. van der Meer, J.R., and Sentchilo, V.S., 2003, Genomic islands and evolution of catabolic pathways in bacteria. *Curr. Opin. Biotechnol.*, 14:248–254.
227. van der Meer, J.R., van Neerven, A.R.W., de Vries, E.J., de Vos, W.M., and Zehnder, A.J.B., 1991, Cloning and characterization of plasmid-encoded genes for the degradation of 1,2-dichloro-, 1,4-dichloro-, and 1,2,4-trichlorobenzene of *Pseudomonas* sp. strain P51. *J. Bacteriol.*, 173:6–15.
228. van der Meer, J.R., Werlen, C., Nishino, S., and Spain, J.C., 1998, Evolution of a pathway for chlorobenzene metabolism leads to natural attenuation in a contaminated groundwater. *Appl. Environ. Microbiol.*, 64:4185–4193.
229. van der Meer, J.R., Zehnder, A.J.B., and de Vos, W.M., 1991, Identification of a novel composite transposable element, Tn*5280*, carrying chlorobenzene dioxygenase genes of *Pseudomonas* sp. strain P51. *J. Bacteriol.*, 173:7077–7083.
230. Vedler, E., Vahter, M., and Heinaru, A., 2004, The completely sequenced plasmid pEST4011 contains a novel IncP1 backbone and a catabolic transposon harboring *tfd* genes for 2,4-dichlorophenoxyacetic acid degradation. *J. Bacteriol.*, 186:7161–7174.
231. Viggor, S., Heinaru, E., Loponen, J., Merimaa, M., Tenno, T., and Heinaru, A., 2002, Biodegradation of dimethylphenols by bacteria with different ring-cleavage pathways of phenolic compounds. *Environ. Sci. Pollut. Res. Int.*, 1:19–26.
232. Weisshaar, M.-P., Franklin, F.C.H., and Reineke, W., 1987, Molecular cloning and expression of the 3-chlorobenzoate-degrading genes from *Pseudomonas* sp. strain B13. *J. Bacteriol.*, 169:394–402.
233. Werlen, C., Kohler, H.-P.E., and van der Meer, J.R., 1996, The broad substrate chlorobenzene dioxygenase and *cis*-chlorobenzene dihydrodiol dehydrogenase of *Pseudomonas* sp. strain P51 are linked evolutionarily to the enzymes for benzene and toluene degradation. *J. Biol. Chem.*, 271:4009–4016.
234. Whyte, L.G., Bourbonnière, L., and Greer, C.W., 1997, Biodegradation of petroleum hydrocarbons by psychrotrophic *Pseudomonas* strains possessing both alkane (*alk*) and naphthalene (*nah*) catabolic pathways. *Appl. Environ. Microbiol.*, 63:3719–3723.
235. Williams, P.A., Assinder, S.J., De Marco, P., O'Donnell, K.J., Poh, C.L., Shaw, L.E., and Winson, M.K., 1992, Catabolic gene duplications in TOL plasmids, pp. 341–353. In E. Galli, S. Silver, and B. Witholt (eds.), *Pseudomonas: Molecular Biology and Biotechnology*. American Society for Microbiology, Washington, DC.

236. Williams, P.A., and Murray, K., 1974, Metabolism of benzoate and the methylbenzoates by *Pseudomonas putida* (*arvilla*) mt-2: evidence for the existence of a TOL plasmid. *J. Bacteriol.*, 120:416–423.
237. Williams, P.A., and Sayers, J.R., 1994, The evolution of pathways for aromatic hydrocarbon oxidation in *Pseudomonas*. *Biodegradation*, 5:195–218.
238. Williams, P.A., Shaw, L.M., Pitt, C.W., and Vrecl, M., 1997, *XylUW*, two genes at the start of the upper pathway operon of TOL plasmid pWW0, appear to play no essential part in determining its catabolic phenotype. *Microbiology*, 143:101–107.
239. Williams, P.A., Taylor, S.D., and Gibb, L.E., 1988, Loss of the toluene–xylene catabolic genes of TOL plasmid pWW0 during growth of *Pseudomonas putida* on benzoate is due to a selective growth advantage of 'cured' segregants. *J. Gen. Microbiol.*, 134:2039–2048.
240. Williams, P.A., and Worsey, M.J., 1976, Ubiquity of plasmids in coding for toluene and xylene metabolism in soil bacteria: evidence for the existence of new TOL plasmids. *J. Bacteriol.*, 125:818–828.
241. Wilson, M.S., Herrick, J.B., Jeon, C.O., Hinman, D.E., and Madsen, E.L., 2003, Horizontal transfer of *phnAc* dioxygenase genes within one of two phenotypically and genotypically distinctive naphthalene-degrading guilds from adjacent soil environments. *Appl. Environ. Microbiol.*, 69:2172–2181.
242. Wright, B.E., 2000, A biochemical mechanism for nonrandom mutations and evolution. *J. Bacteriol.*, 182:2993–3001.
243. Wyndham, R.C., Cashore, A.E., Nakatsu, C.H., and Peel, M.C., 1994, Catabolic transposons. *Biodegradation*, 5:323–342.
244. Yen, K.-M., Karl, M.R., Blatt, L.M., Simon, M.J., Winter, R.B., Fausset, P.R., Lu, H.S., Harcourt, A.A., and Chen, K.K., 1991, Cloning and characterization of a *Pseudomonas mendocina* KR1 gene cluster encoding toluene-4-monooxygenase. *J. Bacteriol.*, 173:5315–5327.
245. Yen, K.-M., and Serdar, C.M., 1988, Genetics of naphthalene catabolism in *Pseudomonads*. *CRC Crit. Rev. Microbiol.*, 15:247–268.
246. Zeyer, J., Kocher, H.P., and Timmis, K.N., 1986, Influence of *para*-substituents on the oxidative metabolism of *o*-nitrophenols by *Pseudomonas putida* B2. *Appl. Environ. Microbiol.*, 52:334–339.
247. Zhou, J.Z., and Tiedje, J.M., 1995, Gene transfer from a bacterium injected into an aquifer to indigenous bacteria. *Mol. Ecol.*, 4:613–618.
248. Zylstra, G.J., and Gibson, D.T., 1989, Toluene degradation by *Pseudomonas putida* F1. *J. Biol. Chem.*, 264:14940–14946.

8

CONTROLLING REGIOSPECIFIC OXIDATION OF AROMATICS AND THE DEGRADATION OF CHLORINATED ALIPHATICS VIA ACTIVE SITE ENGINEERING OF TOLUENE MONOOXYGENASES

Ayelet Fishman, Ying Tao, Gönül Vardar, Lingyun Rui, and Thomas K. Wood

Departments of Chemical Engineering and Biology
University of Texas AdM
College Station, TX 77843-3122, USA

1. SUMMARY

Toluene monooxygenases from pseudomonads are powerful enzymes whose activities have not been fully appreciated. Recently the literature was corrected to show there was no true, predominantly *meta*-hydroxylating enzyme, and it was discovered that these enzymes perform two and three successive hydroxylations of benzene. The physiological relevance of these successive hydroxylations by toluene *p*-monooxygenase (TpMO) was discerned for the toluene degradation pathway of *Ralstonia pickettii* PKO1. More importantly, these successive hydroxylations create the possibility of producing substituted aromatics for industrially important syntheses. To take advantage of this potential and to probe the hydroxylation reaction itself, DNA shuffling has been used to determine the residues that control the regiospecific catalytic activity. This regiospecific activity was fine tuned using saturation mutagenesis to create

Pseudomonas, Volume 4, edited by Juan-Luis Ramos and Roger C. Levesque

monooxygenases that, along with the wild-type enzymes, are able to form various doubly hydroxylated, substituted aromatics including 3-nitrocatechol, 4-nitrocatechol, nitrohydroquinone, methylhydroquinone, 4-methylresorcinol, methoxyresorcinol, 3-methoxycatechol, 2-naphthol, and 1-hydroxyfluorene. The regiospecific control of hydroxylation has reached the point where toluene may be hydroxylated at the *ortho-, meta-,* and *para*-positions by altering just two amino acids near the diiron active site, and these residues may be altered to create the first predominantly *meta*-hydroxylating enzyme. Also, regiospecific oxidation of naphthalene may be controlled to form both 1- and 2-naphthol (first report of microbial formation of this compound), and the regiospecific oxidation of indole may be controlled so that a single enzyme [toluene *o*-monooxygenase (TOM) of *Burkholderia cepacia* G4] may be used to form a variety of indigoids including predominantly indigo, isoindigo, indirubin, and isatin. To control the successive hydroxylations, two-phase reactors have also been designed to produce phenol from benzene and 2-naphthol from naphthalene using toluene 4-monooxygenase (T4MO) from *Pseudomonas mendocina* KR1. To show the utility of these enzymes for pathway engineering, they have been combined with both glutathione *S*-transferases as well as epoxide hydrolases to create engineered bacteria that have accelerated degradation rates for chlorinated ethenes as well as to take advantage of the fact that only one oxygenase has been shown to oxidize tetrachloroethylene (toluene *o*-xylene monooxygenase of *Pseudomonas stutzeri* OX1). Similar monooxygenation reactions for substituted aromatics have been discovered using dinitrotoluene dioxygenases.

2. INTRODUCTION

Enzymes are attractive catalysts for chemical synthesis due to their exceptional enantio and regioselectivities; hence, enzymes can be used in both simple and complex syntheses without the need for the blocking and deblocking steps often found in their organic counterparts[105,106]. The number of biocatalytic processes that are being performed on a large scale is rapidly increasing, and it is projected that this growth will continue[48,114] to become 30% of the chemical business by 2050[122]. Although the majority of the currently used industrial enzymes are hydrolases[48], there is growing interest in the application of oxygenases[122]. Oxidative biotransformations use oxygen as an inexpensive, environmentally friendly oxidant in contrast to toxic chemical oxidants, and they exceed their chemical equivalent regiospecificity and enantioselectivity[10,58]. Among bacterial oxygenases, toluene monooxygenases have much potential for bioremediation and chemical synthesis as we have used mutagenesis to make, among other things, variants that remediate one of the world's worst pollutants (chlorinated ethenes)[15,97,99], that form a rainbow of colors from indole

with a single enzyme, that make the anticancer compound indirubin[100], and that make useful compounds like 1-naphthol (15,000 ton/year market[15]), nitrohydroquinone (precursor for therapeutics for Alzheimer's and Parkinson's disease)[127], and 4-nitrocatechol (inhibitor of nitric oxide synthase)[27]. For these industrial applications, control of regiospecific oxidation of aromatics is of vital importance. Further, regiospecific oxidation of toluene, substituted benzenes, and naphthalene by these monooxygenases presents an important challenge both in terms of the three distinct regiospecific oxidations possible for the initial hydroxylation of toluene as well as the usefulness of both the mono- and di-hydroxylated products.

TpMO, formerly known as T3MO[30] of *Ralstonia pickettii* PKO1 is a soluble, non-heme, diiron monooxygenase[52] and is composed of six proteins encoded by *tbuA1UBVA2C*[12]. TbuA1, TbuA2, and TbuU are the α, β, and γ subunits, respectively, of the hydroxylase component (209 kDa with $\alpha_2\beta_2\gamma_2$ quaternary structure) which controls regiospecificity[73,90]. TbuB is a Rieske-type [2Fe-2S] ferredoxin (12.3 kDa), and TbuV is an effector protein (11.7 kDa). TbuC is a NADH oxidoreductase (36.1 kDa) which is responsible for the transfer of electrons from reduced nucleotides via the redox center to the terminal oxygenase[12]. *R. pickettii* PKO1 cells expressing the TpMO pathway were shown to degrade trichloroethylene (TCE) when induced by toluene or TCE[53,86].

T4MO[115] of *Pseudomonas mendocina* KR1, TOM[77] of *Burkholderia cepacia* G4, and toluene/*o*-xylene monooxygenase (ToMO)[3] of *Pseudomonas stutzeri* OX1 are similar enzymes[52,79] except TOM contains a relatively unknown subunit (from *tomA0*) rather than a Rieske-type [2Fe-2S] ferredoxin. The TpMO *tbu* locus is most similar to T4MO *tmo* genes in the DNA sequence (>60%)[54], but both T4MO and TpMO are only distantly related to TOM and are substantially different enzymes (as evidenced by their different hydroxylation of toluene) since the hydroxylase alpha fragments of TpMO and T4MO share only 48% DNA identity and only 23% protein identity with TOM.

To probe the catalytic potential of these four enzymes it is necessary to clone them into a background devoid of competing monooxygenase and dioxygenase reactions. Whole *Escherichia coli* TG1 cells are an excellent host for studying these enzymes due to the lack of background oxygenases (when grown in LB medium), to the relatively strong expression of the multiple components of these enzymes (three-component hydroxylase, reductase, mediating protein, and ferredoxin), and to the ample production of the necessary cofactor NADH (which likely limits industrial production to whole cells). We have constructed plasmids like that shown in Figure 1 to express TOM, ToMO, TpMO, and T4MO in *E. coli*. Using these constructs, we have found the catalytic potential of monooxygenases is best harnessed using the random mutagenesis tool of DNA shuffling followed by saturation mutagenesis rather than site-directed mutagenesis.

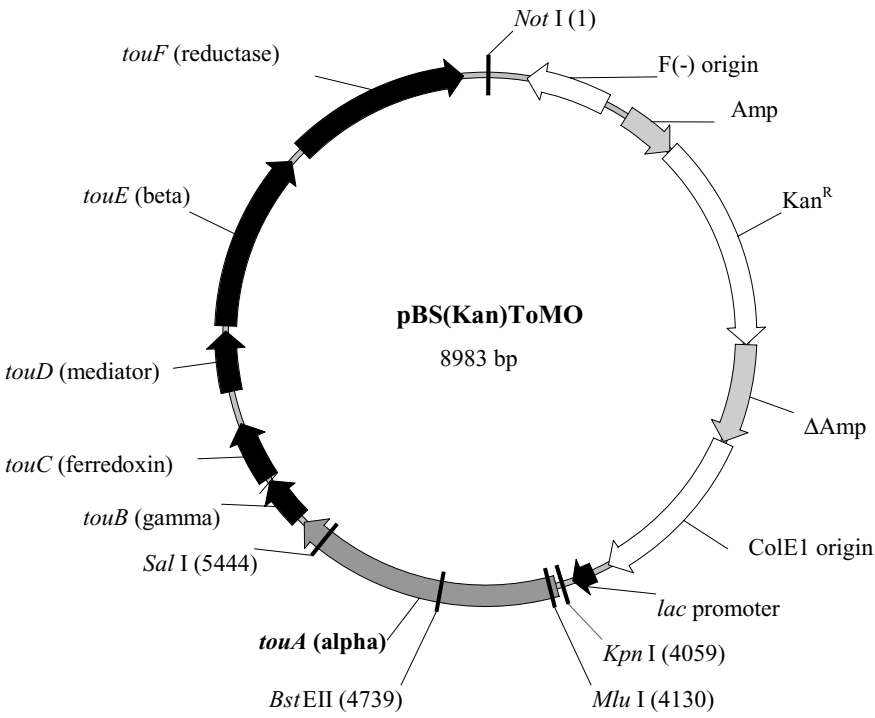


Figure 1. Vector pBS(Kan)ToMO for constitutive expression of wild-type ToMO and its mutants. Kan^R is the kanamycin resistance gene. The six genes for ToMO are *touABE* (encoding the three-component hydroxylase, $A_2B_2E_2$), *touC* (encoding ferredoxin), *touD* (encoding the mediating protein) and *touF* (encoding the NADH-ferredoxin oxidoreductase). Similar constructs pBS(Kan)TOM, pBS(Kan)T3MO, and pBS(Kan)T4MO were made.

Using this approach, we have discovered that the ToMO-equivalent alpha subunit residues I100 (gate residue)[15], A101[127], A107[100], A110[127], M180[124,128], and E214 (gate residue)[124,127] influence catalysis in this family of monooxygenases. Furthermore, purification of these complex proteins is not necessary to gauge changes in regiospecific activity. The novel enzymes produced are all active; often their activity on the wild-type substrate toluene exceeds that of the wild-type enzyme[27,98,118]. We have also found relationships between oxidation rate and regiospecificity[27,98,118], as well as between electrophilic resonance/inductive effects and regiospecificity[28]. With these enzymes and their variants, we can now synthesize nitrohydroquinone, 4-methylresorcinol, 1-hydroxyfluorene, 3-hydroxyfluorene, 4-hydroxyfluorene, 2-naphthol, 2,6-dihydroxynaphthalene, and 3,6-dihydroxyfluorene whereas there were no previous reports of the synthesis of these compounds with specific enzymes[117,127,128].

3. SUCCESSIVE HYDROXYLATIONS OF TOLUENE MONOXYGENASES

It was reported previously that no catechol derivatives were detected using T4MO with toluene and benzene as substrates[91], and that there was no further hydroxylation of *p*-cresol by the T4MO G103L mutant created by site-directed mutagenesis[73]. Also, all previous publications indicate that T4MO only hydroxylates unactivated benzene nuclei but not phenolic compounds[52,73,91]. However, it was discovered recently in our laboratory that T4MO produces catechol at physiological rates using both benzene and phenol as substrates[119] as well as produces trihydroxybenzene (THB), so toluene monooxygenases (TOM, ToMO, TpMO, and T4MO) are capable of not one but three successive hydroxylations of benzene[119].

Catechols are important intermediates for synthesis of pharmaceuticals, agrochemicals, flavors, polymerization inhibitors, and antioxidants[20,21]. Currently, catechol is produced primarily by the oxidation of phenol, *m*-diisopropylbenzene, or by coal-tar distillation[41]. However, the industrial routes to catechols are environmentally unsafe, e.g. the use of elevated metal, temperature, pressure, and solvent conditions[1,41]. These chemical routes are often lengthy, energy-intensive, multi-step reactions that require expensive starting materials and are plagued with isomerization and rearrangement problems [1,93]; hence, microbial production of catechols is attractive. Previously, catechol has been produced by transforming D-glucose with a genetically modified *E. coli* AB2834/pKD136/pKD9.069A expressing 3-dehydroshikimic acid dehydratase and protocatechuic acid decarboxylase[20,21] and by benzene oxidation with *Pseudomonas putida* 6–12 expressing toluene/benzene dioxygenase while lacking catechol 1, 2-oxygenase and catechol 2, 3-oxygenase[93].

It was discovered that T4MO of *P. mendocina* KR1, TpMO of *R. pickettii* PKO1, ToMO of *P. stutzeri* OX1, and TOM of *B. cepacia* G4 convert benzene to phenol, catechol, and 1,2,3-THB by successive hydroxylations[118,128] (Figure 2). At a concentration of 165 μM, under control of a constitutive *lac* promoter, *E. coli* TG1/pBS(Kan)T4MO expressing T4MO formed phenol from benzene at 19 ± 1.6 nmol/min/mg protein, catechol from phenol at 13.6 ± 0.3 nmol/min/mg protein, and 1,2,3-THB from catechol at 2.5 ± 0.5 nmol/min/mg protein (Table 1). The catechol and 1,2,3-THB products were identified by both high performance liquid chromatography (HPLC) and mass spectrometry. Using analogous plasmid constructs, *E. coli* TG1/pBS(Kan)T3MO expressing TpMO (plasmid constructed before realization that monooxygenase was a *para*-hydroxylating enzyme) formed phenol, catechol, 1,2,3-THB at a rate of 3 ± 1, 3.1 ± 0.3, and 0.26 ± 0.09 nmol/min/mg protein, respectively, and *E. coli* TG1/pBS(Kan)TOM expressing TOM formed 1,2,3-THB at a rate of 1.7 ± 0.3 nmol/min/mg protein (phenol and catechol formation rates were

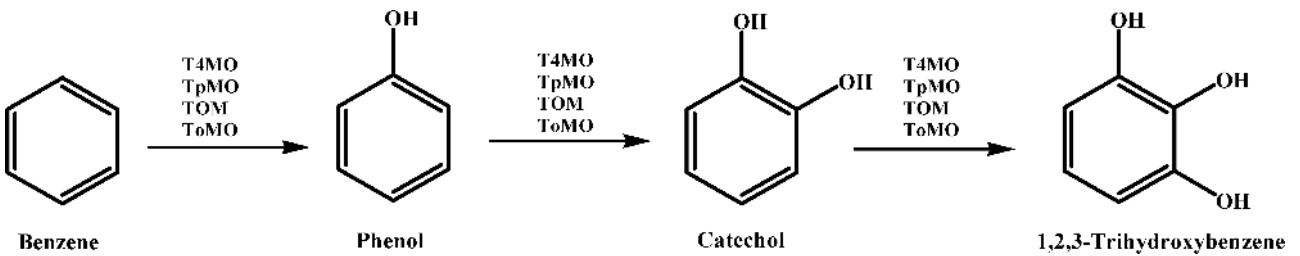


Figure 2. Successive hydroxylations of benzene by TG1(T4MO), TG1(TpMO), TG1(TOM), and TG1(ToMO).

Table 1. Synthesis[a] of phenol from benzene, catechol from phenol, and 1,2,3-THB from catechol by *E. coli* TG1 cells expressing wild-type T4MO, TpMO, and TOM. Initial concentration of substrates was 165 μM[b].

	Phenol formation from benzene		Catechol formation from phenol		1,2,3-THB formation from catechol		
Enzyme	Initial formation rate, nmol/min/mg protein	Maximum production, μM	Initial formation rate, nmol/min/mg protein	Maximum production, μM	Initial formation rate, nmol/min/mg protein	Maximum production, μM	Toluene oxidation rate, nmol/min/mg protein
T4MO[c]	19 ± 1.6	144 ± 47	13.6 ± 0.3	103 ± 10	2.5 ± 0.5	132 ± 22	10 ± 0.8
TpMO[c]	3 ± 1	122 ± 43	3.1 ± 0.3	119 ± 13	0.26 ± 0.09	73 ± 4	4 ± 0.6
TOM[d]	0.89 ± 0.07	27 ± 12	1.8 ± 0.5	140 ± 7	1.7 ± 0.3	103 ± 22	2.4 ± 0.3

[a] Based on HPLC analysis, the mean ± standard deviation of at least two independent results are shown.
[b] Initial benzene liquid concentrations of 165 μM based on a Henry's law constant of 0.22[19] (400 μM added if all the benzene in the liquid phase), and the initial toluene concentration was 165 μM based on a Henry's law constant of 0.27[19] (455 μM added if all the toluene in the liquid phase).
[c] Protein concentration 0.24 mg protein/(mL × OD).
[d] Protein concentration 0.22 mg protein/(mL × OD).

0.89 ± 0.07 and 1.5 ± 0.3 nmol/min/mg protein, respectively). Hence, the rates of synthesis of catechol by both TpMO and T4MO and the 1,2,3-THB formation rate by TOM were found to be comparable to the rates of oxidation of the natural substrate toluene for these enzymes (10.0 ± 0.8, 4.0 ± 0.6, and 2.4 ± 0.3 nmol/min/mg protein for T4MO, TpMO, and TOM, respectively, at 165 μM toluene).

Previously, it was reported that *P. mendocina* KR1 utilizes toluene as a sole carbon and energy source converting it to *p*-cresol via T4MO[131]. This single hydroxylation is followed by oxidation of the methyl group by *p*-cresol methylhydroxylase and *p*-hydroxybenzaldehyde dehydrogenase, resulting in *p*-hydroxybenzoate, which is oxidized to protocatechuate[131]. Protocatechuate is metabolized through an *ortho*-cleavage pathway[131]. *R. pickettii* PKO1 metabolizes benzene and toluene to phenol and *m*-cresol, respectively, via TpMO[13]. Phenol and *m*-cresol are then further oxidized by phenol hydroxylase to catechol and 3-methylcatechol, respectively, which are then cleaved by a *meta*-fission dioxygenase[13]. Therefore, it was surprising to find here that T4MO and TpMO further oxidize phenol and catechol as they do not appear to be physiologically relevant reactions.

At least four monooxygenases capable of successive hydroxylations of aromatics have been found. TOM[77] and toluene/benzene 2-monooxygenase of *Burkholderia* sp. strain JS150[45] transform toluene to *o*-cresol and then *o*-cresol to 3-methylcatechol. *p*-Nitrophenol monooxygenase from *Burkholderia sphaericus* JS905 converts *p*-nitrophenol to 4-nitrocatechol, and then removes the nitro group and forms the 1,2,4-THB[46]. ToMO also catalyzes toluene or *o*-xylene oxidation into methylcatechols by two subsequent monooxygenations[3] but was reported to only transform benzene to phenol[3]. Before this report, based on their ability to perform successive hydroxylations, TOM and toluene/benzene 2-monooxygenase were thought to be most similar to phenol hydroxylases rather than to other toluene monooxygenases[52,73]; however, the accumulation of catechol from benzene by TG1(T4MO) and TG1(TpMO) encoding T4MO and TpMO, respectively, indicates that these two monooxygenases possess a sequential hydroxylation pathway similar to that of TOM of *B. cepacia* G4[77], ToMO of *P. stutzeri* OX1[3], and toluene/benzene 2-monooxygenase of *Burkholderia* sp. strain JS150[45].

4. PHYSIOLOGICAL RELEVANCE OF SUCCESSIVE HYDROXYLATIONS OF TpMO

Xylene monooxygenase of *Pseudomonas putida* mt-2 hydroxylates toluene at the methyl side chain resulting in benzyl alcohol[7], and TOM of *B. cepacia* G4 hydroxylates the benzene ring at the *ortho* position to form

o-cresol which is further oxidized to 3-methylcatechol[77,109]. T4MO of *P. mendocina* KR1 is specific for *para* hydroxylations producing primarily *p*-cresol[134], and TpMO of *R. pickettii* PKO1 was reported to hydroxylate toluene at the *meta* position[83] resulting in *m*-cresol. However, using gas chromatography (GC) and ^{1}H-nuclear magnetic resonance spectroscopy, we discovered that TpMO hydroxylates monosubstituted benzenes predominantly at the *para* position[30]. *E. coli* TG1/pBS(Kan)T3MO cells expressing TpMO oxidized toluene at a maximal rate of 11.5 $\pm$ 0.33 nmol/min/mg protein with an apparent K_m value of 250 μM, and produced 90% *p*-cresol and 10% *m*-cresol. This product mixture was successively transformed to 4-methylcatechol by TpMO. T4MO, in comparison, produces 97% *p*-cresol and 3% *m*-cresol. *P. aeruginosa* POA1 harboring pRO1966 (the original TpMO-bearing plasmid) also exhibited the same product distribution as TG1/pBS(Kan)T3MO. TG1/pBS(Kan)T3MO produced 66% *p*-nitrophenol and 34% *m*-nitrophenol from nitrobenzene, 100% *p*-methoxyphenol from methoxybenzene, as well as 62% 1-naphthol and 38% 2-naphthol from naphthalene; similar results were found with TG1/pBS(Kan)T4MO. Sequencing of the *tbu* locus from pBS(Kan)T3MO and pRO1966 revealed complete identity between the two thus eliminating any possible cloning errors. Hence, there is no true *meta*-hydroxylating enzyme.

We propose a modification (Figure 3) of the degradation pathway originally described by Olsen *et al.*[83] that now relies primarily on TpMO for conversion of toluene directly to 4-methylcatechol in two successive hydroxylations. Toluene is converted primarily to *p*-cresol instead of *m*-cresol and then both *m*-cresol and *p*-cresol are oxidized to 4-methylcatechol since both *m*-cresol and *p*-cresol were found to be good substrates for *E. coli* expressing TpMO (V_{max}/K_m=0.046, 0.036, and 0.055 mM/min/mg protein for the oxidation of toluene, *m*-cresol, and *p*-cresol, respectively)[29].

In light of the broader activity of TpMO, phenol hydroxylase (encoded by *tbuD*), a flavin monooxygenase[50], appears to facilitate conversion of any *m*-cresol or *p*-cresol formed from toluene oxidation by TpMO to 4-methylcatechol; hence, the cell has a redundant method for making this important intermediate 4-methylcatechol (note that phenol hydroxylase cannot initiate the degradation of toluene) as this reaction was shown to be performed effectively by TpMO with comparable V_{max}/K_m values as toluene oxidation. The existence of multiple monooxygenases in one strain may provide the cell with a competitive advantage and with the ability to utilize a large range of chemically different substrates effectively. Note the phenol hydroxylase is also redundant with respect to hydroxylation of phenol as this reaction was also shown to be performed efficiently by TpMO[119]. Similarly, a recent study by Cafaro *et al*. shows that both ToMO and phenol hydroxylase of *P. stutzeri* OX1 oxidize benzene to phenol and phenol to catechol, albeit at different efficiencies; the phenol hydroxylase

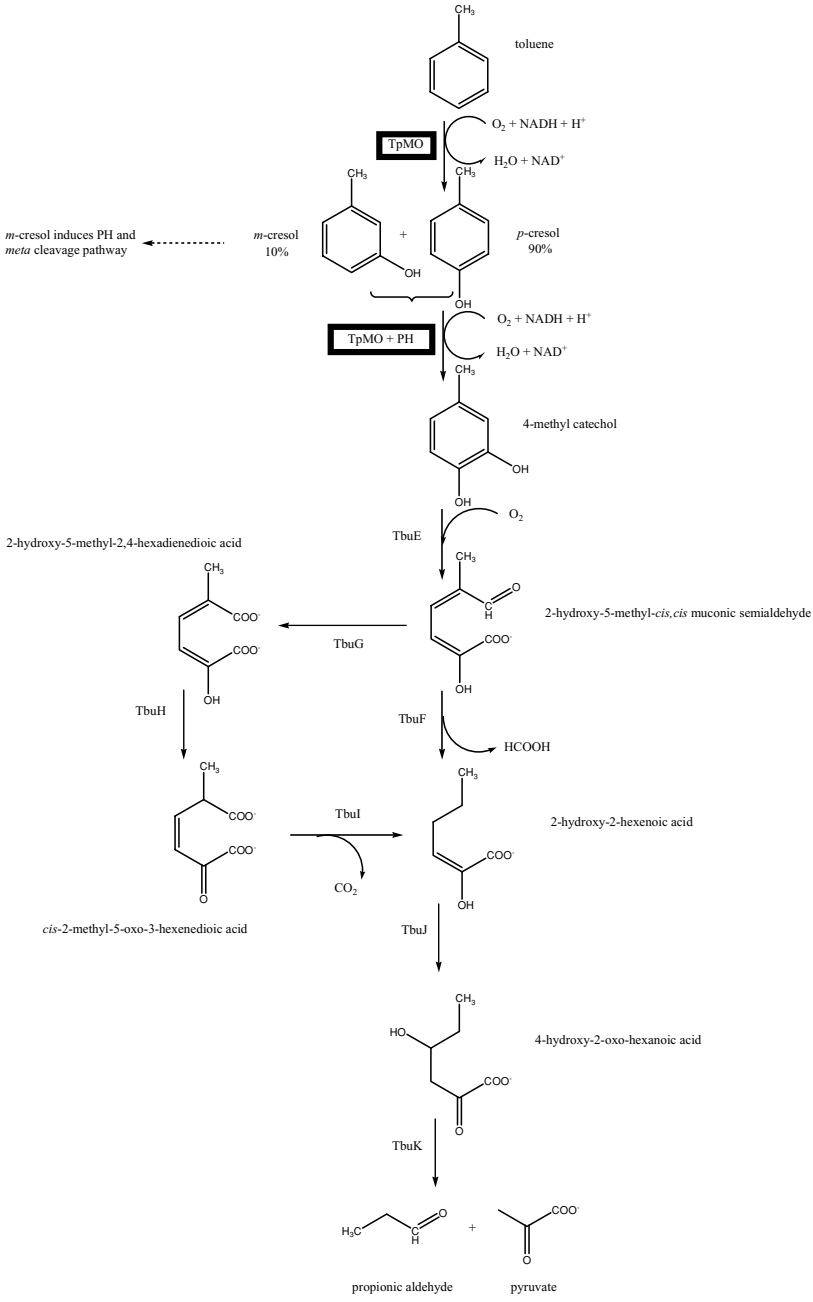


Figure 3. Proposed modification of the pathway for degradation of toluene by *R. pickettii* PKO1 based on the ability of TpMO to perform successive hydroxylations. Abbreviations are: TpMO, toluene *para*-monooxygenase; PH, phenol hydroxylase; TbuE, catechol-2,3-dioxygenase; TbuF, 2-hydroxymuconate semialdehyde hydrolase; TbuG, 2-hydroxymuconate semialdehyde dehydrogenase; TbuH, 4-oxalocrotonate isomerase; TbuI, 4-oxalocrotonate decarboxylase; TbuJ, 2-hydroxypent-2,4-dienoate hydratase; and TbuK, 4-hydroxy-2-oxovalerate aldolase.

appears to mainly drain the mono-hydroxylated compounds and prevent them for accumulating within the cell[14].

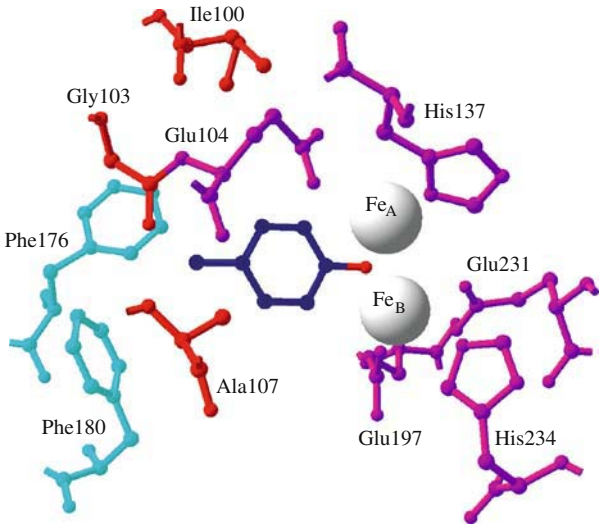
Further, we suggest that the physiological relevance of the 10% *m*-cresol formed from toluene oxidation by TpMO is for induction of the *meta*-cleavage operon *tbuWEFGKIHJ* to enable full metabolism of toluene since *p*-cresol (and *o*-cresol) do not induce the *meta*-cleavage pathway. Additionally, the 10% *m*-cresol serves to induce the redundant phenol hydroxylase. Therefore, the double hydroxylation of toluene by TpMO and its somewhat relaxed regiospecificity are of physiological relevance to *R. pickettii* PKO1. Purely *para*-hydroxylating enzymes are possible as we have constructed one with reasonable activity from T4MO (TmoA variant G103S/A107T)[118]; hence, the cell must choose to make some *m*-cresol as a result of a regulation artifact possibly related to recent acquisition of the lower *meta*-cleavage pathway.

5. DISCOVERY OF RESIDUES THAT CONTROL CATALYTIC ACTIVITY IN TOLUENE MONOOXYGENASES

The soluble methane monooxygenase (sMMO) active site residues have been identified by X-ray crystallography[24,94,95], and by comparison to sMMO, the active site residues for T4MO, TpMO, and toluene 2-monooxygenase from *Pseudomonas* sp. strain JS150 were predicted by Pikus *et al.*[91]. Based on this crystal structure, Fox and co-workers previously found mutations in T4MO of *P. mendocina* KR1 that influence its regiospecificity[73,90,91]. For the oxidation of *m*-xylene by T4MO mutant Q141C of *tmoA*, 3-methylbenzyl alcohol formation increased six-fold from 2.2% to 11.7%, and for *p*-xylene oxidation, the product distribution completely switched to 2,5-dimethylphenol (78%) from 4-methylbenzyl alcohol (22%)[91]. From toluene, the alpha subunit variant TmoA T201F yielded nearly equal amounts of *o*- and *p*-cresol (49.1 and 46.6% respectively) and a substantial amount of benzyl alcohol (11.5%) compared with the wild-type enzyme which makes 97% *p*-cresol and 3% *m*-cresol[90]. Another beneficial T4MO TmoA mutant, F205I, produced 81% *p*-cresol, 14.5% *m*-cresol and insignificant amounts of *o*-cresol and benzyl alcohol[91]. The most notable change was found with mutant TmoA G103L that influenced the selectivity for *ortho*-hydroxylation of toluene yielding 55.4% *o*-cresol[73].

Using DNA shuffling, toluene monooxygenase alpha subunit positions I100, A107, M180, and E214 of the ToMO hydroxylase have been discovered to control activity[15,124,127,128]. Leucine 110 of MmoX in sMMO of *Methylococcus capsulatus* (Bath), the analogous position to ToMO TouA I100 (residue for the similar protein TpMO is shown in Figure 4a), was shown to divide the active site pocket into two cavities and was hypothesized to function in

A) wild-type



B) A107G

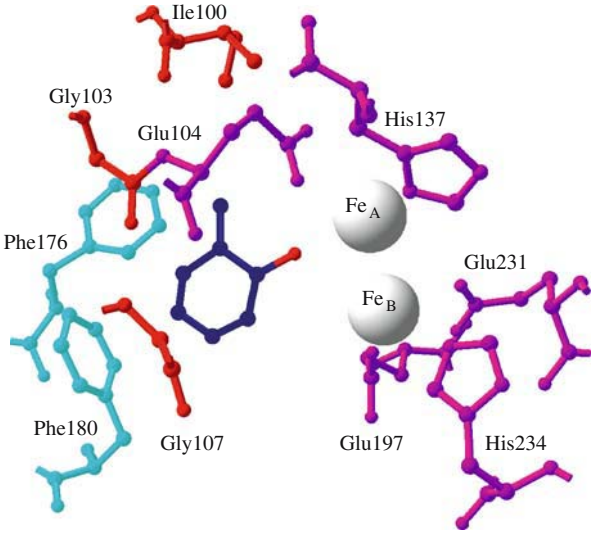
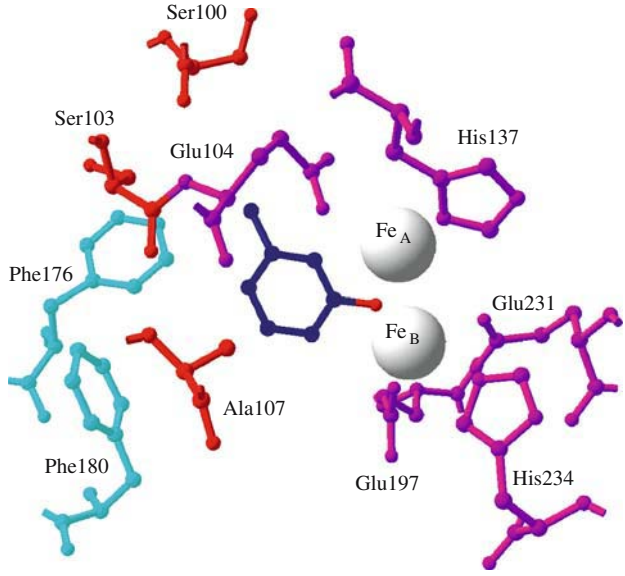


Figure 4. Docking of cresols in the active site of the TpMO TbuA1 α-subunit to indicate how *o*-, *m*-, and *p*-cresol are formed. Mutated residues in red at positions I100, G103, and A107. Residues in purple (E104, H137, E197, E231, and H234) (E134 not shown for clarity) are the coordinating residues anchoring the diiron-binding sites (silver spheres). Residues in light blue are

C) I100S/G103S



D) A107T

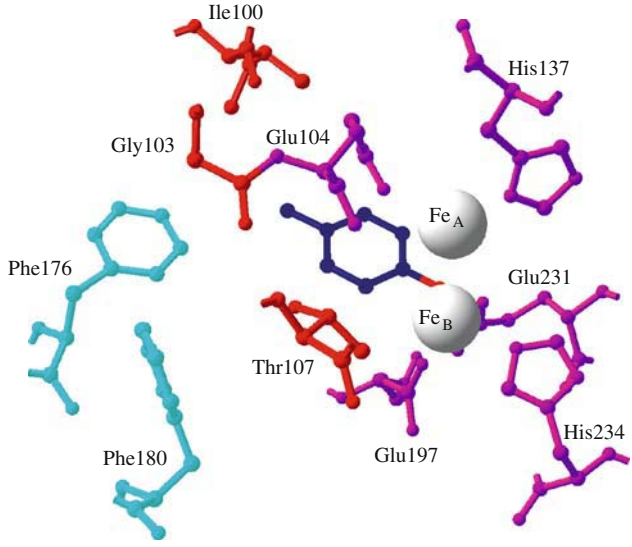


Figure 4. (Continued) part of the hydrophobic shell (F176 and F180). Cresol (dark blue with oxygen in red) is docked inside the active site taking into consideration both energy minimization and steric hindrance. Panel (a) Wild-type TpMO forming *p*-cresol, (b) Variant A107G forming *o*-cresol, (c) Variant I100S/G103S forming *m*-cresol, and (d) Variant A107T forming *p*-cresol.

controlling the access of substrates to the diiron center[94]. Support of this role was provided by Canada *et al.*[15] in the first DNA shuffling of a non-heme monooxygenase in which the function of the analogous position in TmoA3 of TOM of *B. cepacia* G4, V106, was discerned; the V106A variant was able to hydroxylate bulky three-ring polyaromatics such as phenanthrene at higher rates, indicating that a decrease in the size of the side chain allows larger substrates to enter the active site. This variant and two other mutants, V106E and V106F, had decreased regiospecificity for toluene oxidation, altering the strict *ortho*-hydroxylation capability of wild-type TOM to create relaxed catalysts that produce all three cresols[98]. With T4MO of *P. mendocina* KR1, TmoA mutants I100S and I100A exhibited higher oxidation rates (evidenced by higher apparent V_{max}/K_m values) for toluene, nitrobenzene, and nitrophenol oxidation in comparison with the wild-type enzyme[27]. The regiospecificity of toluene and nitrobenzene oxidation also changed resulting in higher percentages of the *meta* isomers. A different TmoA mutant, I100C, had altered enantioselectivity for butadiene epoxidation, forming 60% (*R*)-butadiene epoxide and 40% (*S*)-butadiene epoxide compared with 33% (*R*)- and 67% (*S*)-epoxide produced by wild-type T4MO albeit at the expense of a lower toluene oxidation rate[113]. Position I100 in TouA of ToMO of *P. stutzeri* OX1, the most non-specific of the toluene-oxidizing enzymes, was shown to influence the regiospecific oxidation of cresols and phenol to double hydroxylated products[128]. For example, variant I100Q, which was found through saturation mutagenesis, produced 80% hydroquinone and 20% catechol from phenol compared with 100% catechol formed by wild-type ToMO. In addition, this mutant produced substantial amounts of nitrohydroquinone from *m*-nitrophenol whereas wild-type ToMO produced only 4-nitrocathechol[127].

Alpha subunit position A107 (Figure 4a) is conserved in all monooxygenases studied suggesting it offers some evolutionary advantage[52]. DNA shuffling of TOM of *B. cepacia* G4 showed that this position (TomA3 A113) is involved in color formation of indigoid compounds by influencing whether the two positions of the pyrrole ring or benzene ring are hydroxylated[100]. More specifically, variant A113G (analogous to TpMO TbuA1 A107G which was discovered to be an *ortho*-hydroxylating enzyme) was found to hydroxylate the indole benzene ring rather than the pyrrole ring oxidized by wild-type TOM. Other variants at this position (A113F, A113S, and A113I) were found to produce high amounts of indirubin from indole oxidation (C-2 and C-3 pyrrole hydroxylation) while wild-type TOM produced primarily isoindigo (C-2 pyrrole hydroxylation). A T4MO site-directed mutant TmoA A107S was shown to improve the enantioselectivity of butadiene epoxidation[113] (84% *S*-butadiene epoxide formed vs. 67% for the wild-type T4MO) further demonstrating that a hydrophilic residue is tolerable at this conserved position.

DNA shuffling of ToMO from *P. stutzeri* OX1 for altered methyl- and nitro-aromatic activities (rates or regiospecific changes) led to the discoveries

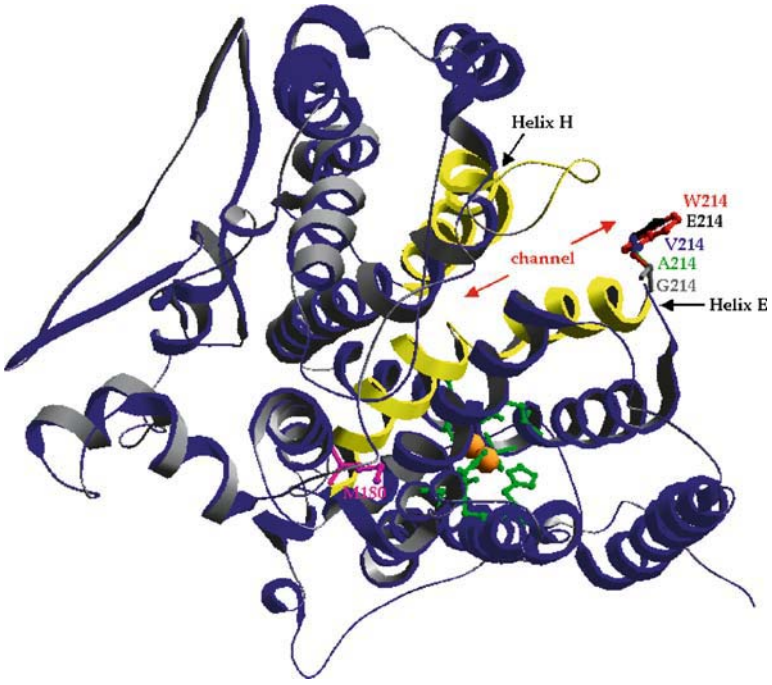


Figure 5. Catalytic residues M180 and E214 of the hydroxylase alpha subunit (TouA) of ToMO. The C-terminal loop of helix E (yellow) and the N-terminal loop of helix H (yellow) form a channel opening at the surface of TouA. Side chains shown in green are the metal-binding residues forming the diiron center (TouA-E104, E134, H137, E197, E231, H234). TouA M180 is pink and TouA E214 is black. The wild-type ToMO hydroxylase (Protein Data Bank accession code 1t0q[103]) was visualized using Swiss-Pdb Viewer program (DeepView)[35,87,107]. TouA E214 forms the entrance of the channel and is ~23 Å away from the active site, whereas TouA M180 is much closer to the active site (~8 Å). Mutations at E214 that show this residue controls the channel are indicated (E214A, E214G, E214V, and E214W). Also variants E214F, E214Q, and E214P were created but are not shown for clarity.

that residues A101, A110, M180 (Figure 5) and E214 (Figure 5) influence catalysis[124,126–128]. Regiospecific hydroxylation of *o*-cresol, *m*-cresol, *p*-cresol, phenol, catechol, and resorcinol were altered by most of the TouA M180 variants as confirmed by HPLC[124]. For example, from *o*-cresol, TouA variant M180H formed 3-methylcatechol (50%), methylhydroquinone (43%), and 4-methylresorcinol (7%), whereas wild-type ToMO formed only 3-methylcatechol (100%)[124]. Also, a shift in regiospecific toluene hydroxylation was observed for variants M180S, M180Q, M180Y, and M180N with *o*-cresol (52%, 59%, 52%, and 19%), *m*-cresol (19%, 15%, 20%, and 19%), and *p*-cresol (29%, 26%, 28%, and 62%) formed, respectively (wild-type ToMO forms 32% *o*-cresol, 21% *m*-cresol, and 47% *p*-cresol).

ToMO TouA E214 is located in the TouA E-helix (closer to the Fe_A site of the diiron center than Fe_B), and is ~23 Å away from the active site, whereas TouA M180 is closer to Fe_B diiron site and is much closer to the active site (~8 Å) (Figure 5). Hence, all of known beneficial residues that influence regiospecificity are nearby the active site, whereas position TouA E214, which influences the rate of oxidation, is not. It appears that TouA E214 is the last residue of helix E and forms an opening of a substrate channel at the northern end of the molecule[124]. Introduction of glycine at this position enhances the rate of oxidation of nitro aromatics including nitrobenzene, *o*-nitrophenol, *m*-nitrophenol, and *p*-nitrophenol. However, the substitutions phenylalanine and glutamine (which are about the same size as glutamic acid) do not enhance the oxidation rate. In contrast, *p*-nitrophenol oxidation by ToMO is enhanced by substituting alanine (four-fold) or valine (1.3-fold) at E214 which are less than that of glycine (15-fold) which is expected if size of the residue at this position is critical. Furthermore, the introduction of the large residue tryptophan reduces the rate six-fold. Therefore, it appears that having a smaller amino acid at TouA position 214 may facilitate substrate entrance/product efflux by increasing the size of the channel opening (Figure 5). This is the second rate enhancing (but not regiospecific mutation) found from DNA shuffling as the TomA3 V106A mutation in TOM (analogous residue I100 in ToMO) allowed greater access to the catalytic center with large substrates such as phenanthrene[15]. Hence, residues M180 and E214 are newly discovered alpha subunit residues of ToMO that influence regiospecific hydroxylation and reaction rates, respectively.

Table 2 summarizes the effect of mutations at these key toluene monooxygenase subunit positions. Note positions 100, 103, and 107 of the alpha subunit of the hydroxylase protein are located in the active site pocket of the B-helix close to the diiron center (Figure 4). As the residues are four amino acids apart from one another they are spatially located one above the other on the α helix[6]. All three amino acids are part of the hydrophobic region surrounding the active site similar to corresponding residues in other diiron monooxygenases[52,90].

Hydrogen bond formation in the active site is important for catalysis with toluene monooxygenases. For example, structure homology modeling suggests that hydrogen bonding interactions of the hydroxyl groups of T4MO and TpMO altered alpha subunit residues S103, S107, and T107 influence the regiospecificity of the oxygenase reaction with methoxyaromatics and toluene[28,118]. The importance of hydrogen bonding for protein function has been shown in many cases. For example, in phenol 2-monooxygenase from *Trichosporon cutaneum*, Y289 was reported to play an important role in leading to *ortho* attack of the substrate by forming a hydrogen bond with phenol substrate[71,133], and the thermophilic xylose isomerase from *Clostridium thermosulfurogenes* increased k_{cat} for glucose by 38% with an additional hydrogen bond to the C_6–OH group of the substrate upon the mutation V186T[70]. In addition, disruption of hydrogen bonds may cause the important amino acids in the catalytic site to lose their

Table 2. Summary of important residues in the alpha subunits of toluene monooxygenases.

Corresponding ToMO residue	Enzyme	Effect on catalysis
I100	ToMO I100Q	Gate residue; Changes the regiospecific oxidation of methyl and nitro aromatics and enhances TCE degradation[126–128]
	T4MO I100L	Enhances 3-methoxycatechol formation from guaiacol[118]
	T4MO I100A	Enhances 4-nitrocatechol formation from nitrobenzene[27]
	T4MO I100C	Changes the regiospecific oxidation of butadiene[113]
	TOM V106A	Enhances TCE and naphthalene oxidation[15]
	TOM V106F	Enhances chloroform oxidation and *p*-cresol formation from toluene[98]
I100/E103	TpMO I100S/G103S	Converts *para* specific TpMO into a *meta* enzyme[28]
A101T	ToMO A101T/M114T	Changes the regiospecific oxidation of toluene[127]
E103	T4MO G103L	Enhances *o*-cresol formation from toluene[73]
E103/A107	T4MO G103A/A107S	Changes the regiospecific oxidation of methoxy aromatics[118]
	T4MO G103S/A107T	Changes the regiospecific oxidation of methyl and methoxy aromatics[118]
	T4MO G103L/A107G	Converts *para* specific T4MO into an *ortho* enzyme[72]
A107/E214	ToMO A107T/E214A	Converts nonspecific ToMO into a *para* enzyme[127]
A107	TpMO A107G	Converts *para* specific TpMO into an *ortho* enzyme[28]
	TpMO A107T	Converts TpMO into a better *para* enzyme[28]
	TOM A113V	Alters indole oxidation to produce primarily indigo (wild-type produces isoindigo)[100]
	TOM A113I	Alters indole oxidation to produce primarily indirubin[100]
	TOM A113H	Alters indole oxidation to produce primarily isatin[100]
A110	ToMO A110T	Changes the regiospecific oxidation of toluene[127]
Q141	T4MO Q141C	Enhances 3-methylbenzyl alcohol formation from *m*-xylene[91]
M180	ToMO M180T	Enhances and changes the regiospecific oxidation of methyl and nitro aromatics[124,127,128]
T201	T4MO T201G	Enhances 2,5-dimethylphenol formation from *p*-xylene[90]
	T4MO T201F	Enhances benzyl alcohol formation from toluene[90]
F205	T4MO F205I	Enhances *m*-cresol formation from toluene[91]
	ToMO F205G	Changes the regiospecific oxidation of methyl and nitro aromatics[127,128]
E214	ToMO E214G	Gate residue; Enhances nitro aromatic oxidation and *cis*-DCE degradation[124,126,127]

catalytic activity[2,60]. For the toluene monooxygenases, the formation of additional hydrogen bonds with the backbone carbonyl, Fe-coordinating carbonyl, and with the substrates may cause the substrate to be oriented in a different position, and this may possibly explain the altered regiospecificity of oxidation for *o*-cresol and *o*-methoxyphenol by these mutants[118].

Through mutagenesis, we have also found that whole cells expressing the variant enzymes (e.g. V106A, V106E, and V106F alpha subunit variants of wild-type TOM[98]) often have higher activity toward toluene, the physiological substrate. However, there is usually a tradeoff in that the wild-type enzyme usually has higher regiospecificity for toluene oxidation; hence, there appears to be an evolutionary balance between activity and regiospecificity. One may increase the rate of oxidation of the physiological substrate toluene at the expense of regiospecificity[27,98,118,127] or increase the regiospecificity at the expense of rate[28,118].

6. SATURATION MUTAGENESIS

Saturation mutagenesis is extremely powerful in creating new catalysts as it can be used to introduce all possible mutations at key sites or adjacent sites to explore a larger fraction of the protein sequence space that can be achieved with site-directed mutagenesis[102]. It can provide much more comprehensive information than can be achieved by single amino acid substitutions as well as overcome the drawbacks of random mutagenesis in that a single mutation randomly placed in codons generates on average only 5.6 out of 19 possible substitutions[11].

To use saturation mutagenesis effectively, it is necessary to determine the number colonies that must be screened. To determine the number of independent clones from saturation mutagenesis that need to be screened to ensure each possible codon has been tested, a multi nomial distribution equation was used[98]. For saturation mutagenesis at one position, it is assumed that each of the 64 possible outcomes has the same probability based on the random synthesis of the primers and the fact that electroporation and plating should have no bias. A program in C language was developed to solve the following equation[96] which was used to describe the number of colonies N that should be tested to make sure the probability that each of the 64 outcomes has been sampled at least once is around 1.0[98]:

$$P\{X_1 = n_1, X_2 = n_2, \ldots, X_r = n_r\} = \frac{N!}{n_1!n_2!\cdots n_r!} p_1^{n_1} p_2^{n_2} \cdots p_r^{n_r}$$

where $p_1, p_2, \ldots, p_r$ are the respective probabilities of any one of r possible outcomes (e.g. $r = 64$ possible codons) resulting from independent and identical experiments, $A_1, A_2, \ldots, A_r$ (e.g. A_1 could be the codon "CAC");

$\sum p_i = 1$; $P\{X_1 = n_1, X_1 = n_1, \ldots, X_r = n_r\}$ is the probability that A_1 happens n_1 times, A_2 happens n_2 times, ..., A_r happens n_r times in N experiments (e.g. $N = 300$ colonies); $\sum n_i = N$; and X_i is the number of n experiments that result in outcome number i (e.g. $X_5 = 10$ means codon 5 was seen 10 times in the population). In saturation mutagenesis of one site, there are 64 possible outcomes and the program calculates the number of colonies so that each $X_i \geq 1$ which means every possible codon has been sampled at least once. With two sites subjected to simultaneous saturation mutagenesis, two independent multinomial distributions apply. The probability that all possible mutants are sampled is found from multiplying the two independent probabilities, and the determination of the number of colonies needed to be sampled is based on that $P_1\{X_1 = n_1, X_1 = n_1, \ldots, X_r = n_r\} \times P_2\{X_1 = n_1, X_1 = n_1, \ldots, X_r = n_r\} = 1.0$.

Our program indicates that 922 independent clones are needed to ensure the probability that all 64 possible outcomes from the single site random mutagenesis has been sampled is 1.0^{98}. If the probability is decreased to 0.99, only 292 colonies need to be screened[98]. If two residues are subject to simultaneous saturation mutagenesis, 342 independent clones need to be sampled to ensure the 0.99 probability that all the possible outcomes have been checked[98].

7. CONTROL OF THE REGIOSPECIFIC HYDROXYLATION OF TOLUENE

A primary goal of protein engineering is to control catalytic activity. As there is no known enzyme which hydroxylates primarily at the *meta* position for toluene[30], we were interested in performing mutagenesis on TpMO to enable it to make substantial amounts of *meta*-hydroxylated compounds from substituted benzenes. Furthermore, we also desired to fine-tune the regiospecificity of TpMO via saturation or site-directed mutagenesis, to render it an *ortho*-hydroxylating enzyme similar to TOM or a more complete *para*-hydroxylating enzyme than even wild-type T4MO. This was the first report of the mutagenesis of TpMO of *R. pickettii* PKO1[28].

We have shown[28] that through mutagenesis of three active site residues, the catalytic activity of a multi-component monooxygenase is altered so that it hydroxylates all three positions of toluene (Figure 6) as well as both positions of naphthalene (Table 3). Hence, for the first time, an enzyme has been engineered so that the complete regiospecific oxidation of a substrate can be controlled. Through the A107G mutation in the alpha subunit of TpMO, a variant was formed that hydroxylates toluene primarily at the *ortho* position while converting naphthalene to 1-naphthol. Conversely, the A107T variant produces >98% *p*-cresol and *p*-nitrophenol from toluene and nitrobenzene,

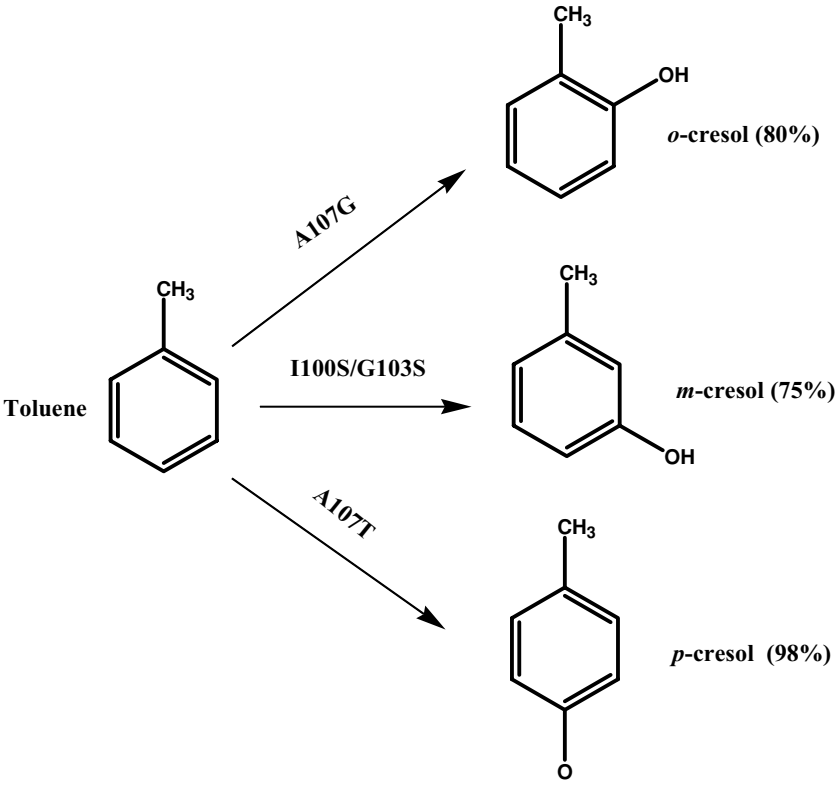


Figure 6. Control of the regiospecificity of toluene oxidation by the TpMO TbuA1 mutants A107G (*o*-cresol), I100S/G103S (*m*-cresol), and A107T (*p*-cresol). Note that wild-type TpMO produces 88% *p*-cresol, 10% *m*-cresol, and 2% *o*-cresol.

respectively, as well as produces 2-naphthol from naphthalene. The mutation I100S/G103S produced a TpMO variant that forms 75% *m*-cresol from toluene and 100% *m*-nitrophenol from nitrobenzene; thus; a true *meta*-hydroxylating toluene monooxygenase was created. This is the first report of the transformation of a single enzyme into a regiospecific catalyst that hydroxylates the aromatic ring at all possible positions producing *ortho*-, *meta*-, and *para*-cresols from toluene and producing 1- and 2-naphthol from naphthalene (there are no previous reports for making 2-naphthol with microorganisms, and the Western world demand for 2-naphthol is three-fold greater than that of 1-naphthol [41]). Evidence of these regiospecific shifts were also shown with nitrobenzene and methoxybenzene (Table 3).

Recently, the crystal structure of ToMO of *P. stutzeri* OX1 was published illustrating the importance of various active site residues on the catalytic

Table 3. Regiospecific oxidation of toluene (with rate), nitrobenzene (NB), methoxybenzene (MB), and naphthalene by TG1 cells expressing wild-type (WT) TOM, WT T4MO, WT TpMO, and TpMO TbuA1 mutants[a].

	Toluene rate[b]	Toluene oxidation[b,c]			NB oxidation[d]			MB oxidation[e]			Naphthalene oxidation[f]	
Enzyme	nmol/min/mg protein	*o*-cresol	*m*-cresol	*p*-cresol	*o*-NP	*m*-NP	*p*-NP	*o*-MP	*m*-MP	*p*-MP	1-naphthol	2-naphthol
WT TOM	1.30 ± 0.06	100	0	0	0	0	0	100	0	0	100	0
WT T4MO	4.40 ± 0.3	1	2	97	0	10	90	0	0	100	52	48
WT TpMO	2.20 ± 0.2	2	10	88	0	34	66	0	0	100	63	37
TbuA1 I100S	3.70 ± 0.7	0	33	67	0	65	35	0	0	100	17	83
TbuA1 G103S	0.74 ± 0.06	8	29	63	0	96	4	28	4	68	80	20
TbuA1 I100S/G103S	1.43 ± 0.08	0	75	25	0	100	0	0	18	82	43	57
TbuA1 A107G	0.84 ± 0.06	80	6	14	0	85	15	88	0	12	97	3
TbuA1 A107T	0.62 ± 0.1	0	2	98	0	0	100	0	0	100	27	73

[a] Each experiment with an enzyme and substrate was conducted twice with independent cultures and the results represent an average of these data (determined with two to three time points).

[b] Initial rate of toluene oxidation from a linear plot of substrate degradation with time (liquid phase concentration 90 μM based on Henry's law constant of 0.27[19]; 250 μM added if all the toluene in the liquid phase).

[c] Benzyl alcohol was formed in negligible amounts by WT T4MO and TpMO with <1.5%.

[d] Based on HPLC analysis over a 45 min time period (initial NB concentration was 200 μM).

[e] Based on HPLC analysis over a 60 min time period (initial MB concentration was 500 μM).

[f] Based on HPLC analysis over a 2 h time period (initial naphthalene concentration was 5 mM). Naphthalene solubility in water is 0.23 mM [89].

activity of this enzyme[103]. Using this structure, we docked the cresol products from toluene into the TpMO active site to reveal the mechanism of the altered regiospecific reactions (Figure 4). Docking calculations showed significant binding preferences for the favored isoform of each mutant, which agreed well with the experimental data concerning toluene oxidation regiospecificity. For variant A107G, the major change in the active site is the reduction of size of the side chain of residue 107 which allows the benzene ring of toluene to shift toward G107 and replace the methyl group of A107, thereby promoting *o*-cresol formation by oxidation of the C-2 carbon (Figure 4b). The same considerations apply for 1-naphthol formation from naphthalene.

The I100S/G103S double mutant that produces 75% *m*-cresol from toluene and 100% *m*-nitrophenol from nitrobenzene may be explained by the increased size of the TpMO channel leading to the Fe_A site (side of the active site pocket); this larger space accommodates the methyl group of toluene such that the C-3 carbon is directed toward the diiron center (Figure 4c). Additionally, the presence of serine at position 103 bears a resemblance to glutamic acid at this position in TouA of ToMO. ToMO produces 21% *m*-cresol from toluene, as well as 47% *p*-cresol and 32% *o*-cresol[128] and this lack of specificity is attributed partially to E103[103]. Similar considerations may apply to the *meta*-hydroxylation capability of TbuA1 I100S/G103S.

The A107T mutation appears to restrict the substrates into more rigid positions relative to the diiron center by shrinking the active site (Figure 4d) and may be responsible for the most specific product profile obtained (98% *p*-cresol/2% *m*-cresol and 100% *p*-nitrophenol). The hydroxyl group of T107 may form a hydrogen bond with the backbone carbonyl of G103 supporting the orientation of this residue in a way that directs the C-4 carbon in toluene towards the diiron center (forming only *p*-cresol) or the C-2 carbon in naphthalene, resulting in 2-naphthol formation.

The fact that the same enzyme can react differently with several substrates of similar structure (Table 3) underlines the importance of electrophilic resonance and inductive effects. For example, none of the enzymes formed *o*-nitrophenol from nitrobenzene including the newly constructed *ortho*-hydroxylating TpMO variant TbuA1 A107G and the *ortho*-hydroxylating wild-type TOM. Nitrobenzene is 10 million times less reactive than benzene[68] and this may explain the lack of reactivity of TOM towards this substrate. In addition, the nitro group has a strong electron withdrawing ability through inductive and resonance effects, thereby directing the electrophilic substitution to the *meta* position[68]. The product distribution of nitrobenzene oxidation by wild-type T4MO, wild-type TpMO, and its TbuA1 variants, in which a higher percentage of *m*-nitrophenol is formed from nitrobenzene relative to *m*-cresol formed from toluene, implies again that both active site orientation of the substrate and electronic effects influence the regiospecificity of the reaction. In contrast to

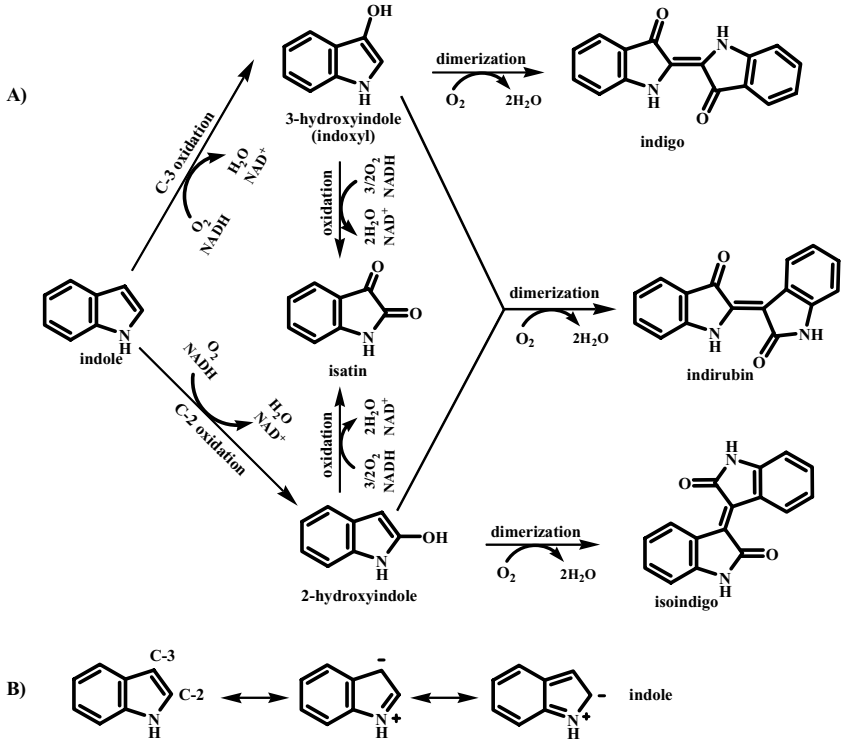


Figure 7. Proposed pathways for converting indole to indigoid compounds via TOM of *B. cepacia* G4 (a) and resonance structure of indole (b).

nitrobenzene, the methyl substituent of toluene and the methoxy substituent of methoxybenzene are *ortho*- and *para*-directing activators, with the methoxy group having more pronounced activation due to the resonance effect [68]. Consequently, *para*-hydroxylation is more complete in methoxybenzene oxidation compared with toluene oxidation (e.g. wild-type T4MO, TpMO, TbuA1 I100S, and TbuA1 A107T in Table 3).

8. CONTROL OF THE REGIOSPECIFIC HYDROXYLATION OF INDOLE

Indigo (Figure 7a) is one of the oldest dyes[26] and is still used worldwide for textiles with 22,000 tons produced annually worth $200 million[64,132]. Production of indigo is primarily by chemical syntheses such as the Adolf

von Baeyer 1890 chemical synthesis[33] which resulted in the fifth Noble Prize in chemistry. More recently, bacterial systems for commercial indigo production have been developed[74], which were inspired by the discovery that recombinant *E. coli* expressing naphthalene dioxygenase from *P. putida* PpG7 in rich medium resulted in the formation of indigo[26]. Indigo is formed as the result of the cloned enzyme oxygenating C-3 of the indole pyrrole ring (Figure 7a), and indole is produced from tryptophan via tryptophanase in *E. coli* [26]. Though limited to the production of a single hue, various monooxygenases and dioxygenases that allow growth on aromatic hydrocarbons have been identified that are capable of indole oxidation to form indigo[4,26,33,80,81], and these biological processes are inherently safer than the chemical processes since they do not produce such toxins as aromatic amines (bladder carcinogens) and cyanide[32,132].

Indirubin (Figure 7a), a pink pigment, has important therapeutic applications[40]. It is the active ingredient of a traditional Chinese medicine used to treat diseases such as chronic myelocytic leukaemia (CML) and was found to be a potent inhibitor of cyclin-dependent kinases and therefore belongs to a group of promising anticancer compounds[9,40].

B. cepacia G4 was isolated as the first pure strain that degrades TCE[76], and TOM has been shown to oxidize mixtures of chlorinated compounds, including TCE[111]. TOM originally was not considered as an indigo-forming enzyme[75,110], but our laboratory found it was responsible for color development and indole hydroxylation[61]. During growth in complex medium, recombinant *E. coli* expressing TOM forms brown colonies on rich agar plates and brown color in broth, whereas typical indole-oxygenating enzymes in whole cells form blue colonies on agar plates and blue, water-insoluble pigments in liquid medium. One TOM variant with a single amino acid change V106A of the hydroxylase alpha-subunit (TomA3) was created by us[15] and was identified as a potential indigo-forming enzyme based on the green color of its colonies on agar plates and in culture. In this variant, a single mutation was responsible for the cell color change, presumably due to the alteration in the hydroxylation of indole.

Consequently, further DNA shuffling was used to isolate a random TOM mutant that turned blue due to mutation TomA3 A113V[100]. To better understand the TOM reaction mechanism, the specificity of indole hydroxylation was studied using a spectrum of colored TOM mutants expressed in *E. coli* TG1 and formed as a result of saturation mutagenesis at TomA3 positions A113 and V106. Colonies expressing these altered enzymes range in color from blue through green and purple to orange, and enzyme products were identified using thin-layer chromatography, HPLC, and LC-mass spectroscopy. Derived from the single TOM template, enzymes were identified that produce primarily isoindigo (wild-type TOM), indigo (A113V), indirubin (A113I), and isatin (A113H and V106A/A113G) (Figure 8). The discovery that wild-type TOM

A)

B)

Figure 8. Colored chloroform LB culture extracts (no indole) of wild-type TOM and alpha subunit variants with mutations at position TomA3 A113 along with standards indigo, indirubin, isatin, and isoindigo (a). Cell color extracts of TomA3 A113G mutants (b).

forms isoindigo via C-2 hydroxylation of the indole pyrrole ring makes this the first oxygenase shown to form this compound. Variant TOM A113G is unable to form indigo, indirubin, or isoindigo (so it does not hydroxylate the indole pyrrole ring), but produces 4-hydroxyindole and unknown yellow compounds from C-4 hydroxylation of the indole benzene ring. Mutations at V106 in addition to A113G restored C-3 indole oxidation so along with C-2 indole oxidation, isatin, indigo, and indirubin were formed. Other TomA3 V106/A113 mutants with hydrophobic, polar, or charged amino acids in place of the Val and/or Ala residues hydroxylated indole at the C-3 and C-2 positions forming isatin, indigo, and indirubin in a variety of distributions. Hence, a single enzyme may be genetically modified to produce a wide range of colors from indole; a single or double amino acid change can create catalytically distinct enzyme variants that hydroxylate indole in different regiospecific positions on the pyrrole and the benzene ring.

9. CONTROL OF THE REGIOSPECIFIC HYDROXYLATION OF METHYL AND METHOXY AROMATICS BY T4MO

Another goal of our work was to seek further improvements in T4MO activity for the regiospecific hydroxylation of methoxyaromatics (Table 4)[118]; there were no previous reports for making functionalized di-hydroxylated aromatics with T4MO as these were previously undiscovered reactions[119]. Saturation mutagenesis was used to change the regiospecificity and to expand the product spectrum for three novel, industrially significant products (3-methoxycatechol, methoxyhydroquinone, and methylhydroquinone) and to convert T4MO into an *ortho*-hydroxylating enzyme as well as a better *para*-hydroxylating enzyme. Six dihydroxylated products (82% 3-methoxycatechol, 87% 4-methoxyresorcinol, 80% methoxyhydroquinone, 98% 3-methylcatechol, 100% 4-methylcatechol, and 92% methylhydroquinone) may be synthesized with high purity by a single T4MO enzyme and its variants via a one-step reaction. On the basis of the observed catalytic activity and product distribution for toluene, *o*-cresol, and *o*-methoxyphenol, we conclude that G103 and A107 are part of the substrate-binding pocket and that T4MO TmoA positions G103S and A107S have a significant role in orienting the substrate for attack at regioselective positions of these phenol substrates.

Substituted catechols, especially 3-substituted catechols, are useful precursors for making pharmaceuticals[38]; one of these, 3-methoxycatechol, is an important intermediate for the antivascular agents combretastatin A-1 and combretastatin B-1[8]. Methoxyhydroquinone is used in the synthesis of triptycene quinones that have been shown to have antileukemia cell activity[42] and methylhydroquinone has been recently reported to be used in the synthesis of (±)-helibisabonol A and puraquinonic acid which are precursors to agrochemical herbicides and antileukemia drugs, respectively[39,62]. Manufacture of these substituted dihydroxylated compounds by chemical routes is difficult due to the employment of aggressive reagents, expensive and complicated starting materials, multiple reaction steps, and low yields[38]. Therefore, alternatives to chemical synthesis of those important industrial intermediates have been investigated.

First, T4MO was discovered to oxidize *o*-cresol to 3-methylcatechol (91%) and methylhydroquinone (9%), to oxidize *m*-cresol and *p*-cresol to 4-methylcatechol (100%), as well as to oxidize *o*-methoxyphenol to 4-methoxyresorcinol (87%), 3-methoxycatechol (11%), and methoxyhydroquinone (2%). Apparent V_{max} values of 6.6 ± 0.9 to 10.7 ± 0.1 nmol/min/mg protein were obtained for *o*-, *m*-, and *p*-cresol oxidation by wild-type T4MO that are comparable to the toluene oxidation rates (15.1 ± 0.8 nmol/min/mg protein). After discovering these new reactions, saturation mutagenesis was performed near the diiron catalytic center at positions I100, G103, and A107 of

Table 4. 3-Methoxycatechol (3MxC), methoxyhydroquinone (MxH), and 4-methoxyresorcinol (4MxR) synthesis from *o*-methoxyphenol by *E. coli* TG1 expressing wild-type T4MO and saturation mutagenesis TmoA variants[a].

	Regiospecificity of *o*-methoxyphenol oxidation			3MxC formation rate (nmol/min/mg protein)			MxH formation rate[d] (nmol/min/mg protein)	*o*-methoxyphenol oxidation (nmol/min/mg protein)
Enzyme	4MxR, %	3MxC, %	MxH, %	Via HPLC	Relative Activity[b]	Via colorimetric assay		
Wild-type	87	11	2	0.21 ± 0.01	1	0.2	0.07 ± 0.03	4.1 ± 1.1
I100L	73	20	7	0.8 ± 0.2	4	0.9–1.3	~0	2.4 ± 0.8
G103A	41	52	7	1.2 ± 0.2	6	0.7–2.7	~0	2.1 ± 0.5
G103A/A107S	13	82	5	1.5 ± 0.2	7	1.9	~0	1.9
G103S	19	<1	80	~0	0	0	0.36 ± 0.01	0.4
G103S/A107T	35	30	35	ND[c]	0	0	0.08	ND[c]

[a] Activity determined at the saturation substrate concentration of 1 mM.
[b] Relative values based on HPLC analysis whereas the colorimetric assay corroborated this data.
[c] ND, not determined.
[d] The *o*-methoxyphenol substrate was oxidized to as many as 3 different products but only the MxH formation rate is shown.

the alpha subunit of the hydroxylase (TmoA) based on directed evolution of the related TOM of *B. cepacia* G4[15,100] (positions I100 & A107) and a previously reported T4MO G103L regiospecific mutation[73].

Using *o*-cresol and *o*-methoxyphenol as model substrates, regiospecific mutants of T4MO were created; for example, TmoA variant G103A/A107S produced 3-methylcatechol (98%) from *o*-cresol two-fold faster and produced 3-methoxycatechol (82%) from 1 mM *o*-methoxyphenol seven times faster than wild-type T4MO (1.5 $\pm$ 0.2 vs. 0.21 $\pm$ 0.01 nmol/min/mg protein). Variant I100L produced 3-methoxycatechol from *o*-methoxyphenol four times faster than wild-type T4MO, and G103S/A107T produced methylhydroquinone (92%) from *o*-cresol four-fold faster than wild-type T4MO and 10 times more as in terms of percentage product. Variant G103S produced methoxyhydroquinone from *o*-methoxyphenol 40-fold higher than the wild-type enzyme (80% vs. 2%) and produced methylhydroquinone (80%) from *o*-cresol. Hence the regiospecific oxidation of *o*-methoxyphenol and *o*-cresol was changed for significant synthesis of 3-methoxycatechol, methoxyhydroquinone, 3-methylcatechol, and methylhydroquinone (Table 4).

The enzyme variants also demonstrated altered mono-hydroxylation regiospecificity for toluene; for example, G103S/A107G formed 82% *o*-cresol, so saturation mutagenesis converted T4MO into an *ortho*-hydroxylating enzyme. Furthermore, G103S/A107T formed 100% *p*-cresol from toluene; hence, a better *p*-hydroxylating enzyme than wild-type T4MO was formed. Structure homology modeling suggests that hydrogen bonding interactions of the hydroxyl groups of altered residues S103, S107, and T107 influence the regiospecificity of the oxygenase reaction.

These findings not only correct the misunderstanding that wild-type T4MO can only hydroxylate benzenes to mono-hydroxylated products and cannot subsequently hydroxylate phenols[52,73,90,131], but also show that wild-type T4MO is a good biocatalyst for making some industrially significant, substituted, catecholic compounds.

10. CONTROL OF THE REGIOSPECIFIC HYDROXYLATION OF METHYL AROMATICS BY ToMO

Di- and tri-hydroxy aromatics (Figures 9–11) are important industrial chemicals with many applications as evidenced by worldwide production of catechol, resorcinol, and hydroquinone at 110,000 tons/year[84]. Catechol is used as an intermediate in the food, pharmaceutical, and agrochemical industries[84], and hydroquinone is used in photography, in cosmetics, and in both medical and industrial X-ray films[34,84]. Resorcinol and its derivatives are used to inhibit rust in

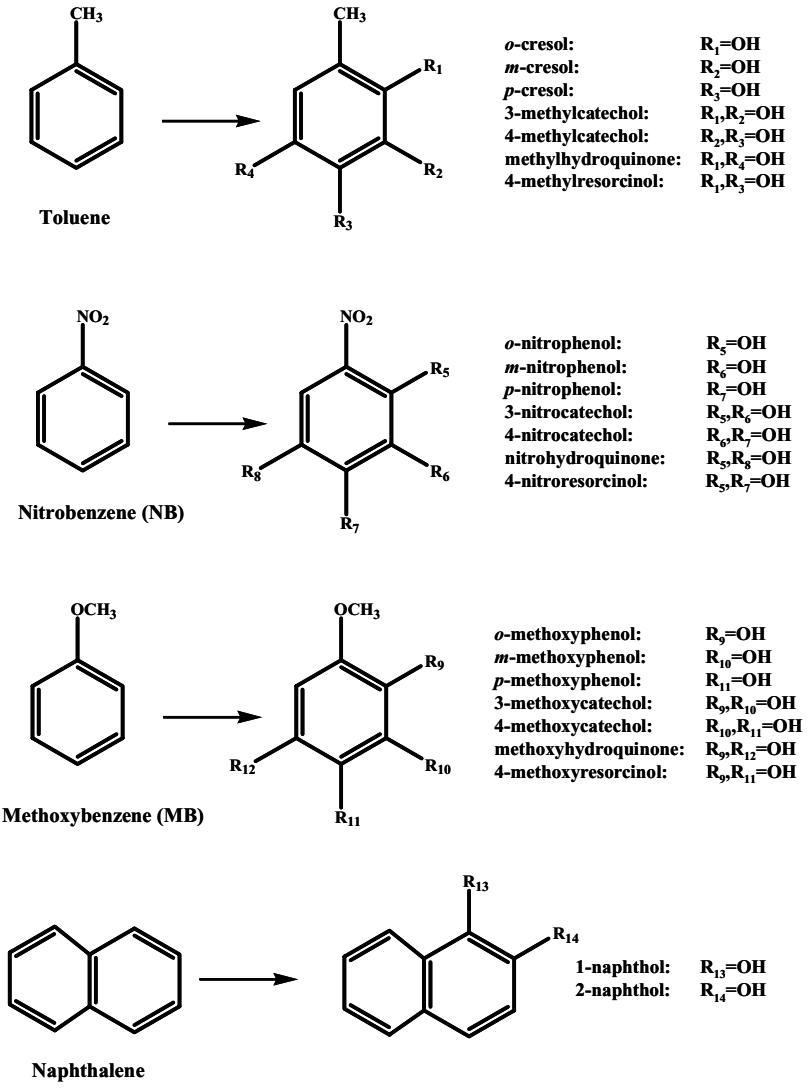


Figure 9. Structures of substituted aromatics discussed in this chapter.

paints, to regulate plant growth, and to act as capacitor electrolytes[22,49,135]. Production of 4-methylresorcinol is uncommon and prices can exceed $200,000/kg (Apin Chemicals, http://www.apinchemicals.com). 1,2,3-THB (pyrogallol), the first synthetic dye for hair, is primarily used as a modifier in oxidation dyes[120], as a pharmaceutical intermediate, and has been used as a topical antipsoriatic[84].

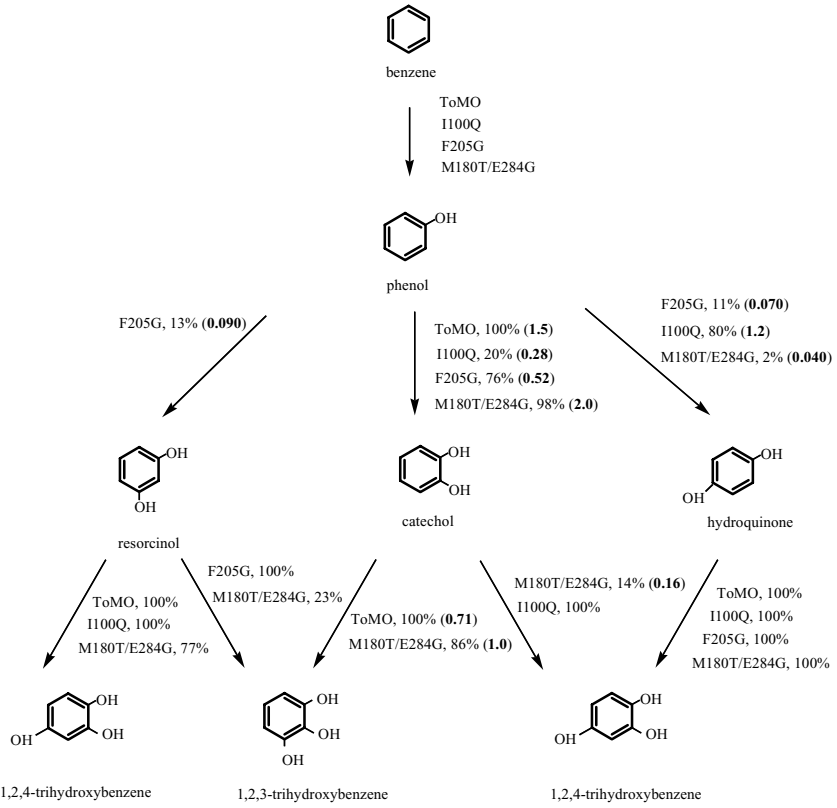


Figure 10. Pathways for the oxidation of benzene (0.8 mM) to phenol, phenol (0.8 mM) to dihydroxy-benzenes, and dihydroxy-benzenes (0.8 mM) to THBs by *E. coli* TG1/pBS(Kan)ToMO expressing wild-type ToMO and TouA variants I100Q, F205G, and M180T/E284G. Molar product percentages are shown followed by bold numbers in parenthesis, ( ), which indicate the product formation rates in nmol/min/mg protein.

Hydroxyhydroquinone (1,2,4-THB) has been used in dyes and as a corrosion inhibitor[41]. Since some of these compounds cannot be easily synthesized chemically, direct microbial synthesis of such compounds from inexpensive substrates might provide a more cost effective and more environmentally benign approach.

ToMO from *P. stutzeri* OX1 oxidizes toluene to 3- and 4-methylcatechol (Figure 11) as well as oxidizes benzene to form phenol (Figure 10); ToMO was found by us to also form catechol and 1,2,3-THB from phenol[128]. To synthesize novel dihydroxy and trihydroxy derivatives of benzene and toluene, DNA shuffling of the alpha hydroxylase fragment of ToMO (TouA) and saturation mutagenesis of the TouA active site residues I100, Q141, T201, and F205 were used to

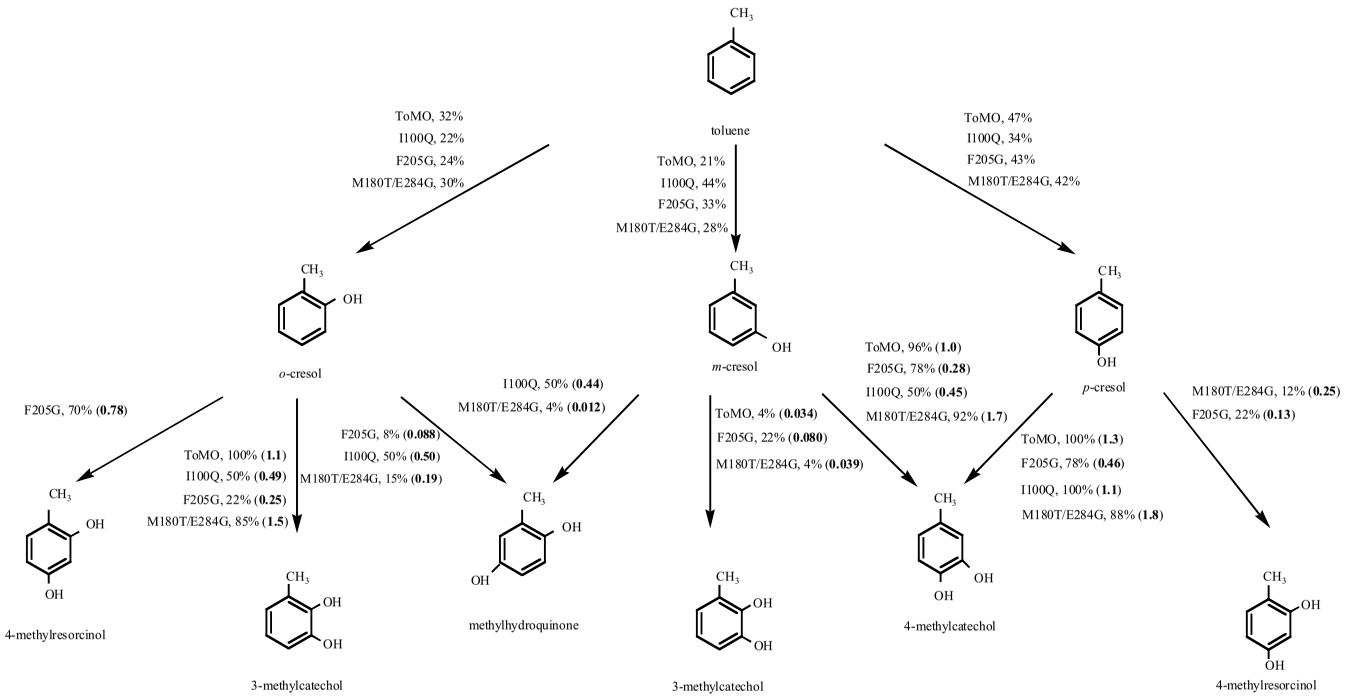


Figure 11. Pathways for the oxidation of toluene (0.8 mM) to *o*-cresol, *m*-cresol, and *p*-cresol, and oxidation of *o*-cresol (0.8 mM), *m*-cresol (0.8 mM), and *p*-cresol (0.8 mM) to methylcatechols, methyl-resorcinols, and methylhydroquinone by *E. coli* TG1/pBS(Kan)ToMO expressing wild-type ToMO and TouA variants I100Q, F205G, and M180T/E284G. Bold numbers in parenthesis, ( ), indicate the product formation rates in nmol/min/mg protein. Molar product percentages are shown before the rate values.

generate random mutants. The mutants were initially identified by screening via a rapid agar plate assay and then were further examined by HPLC and GC. Overall, five new reactions were found for wild-type ToMO (catechol formation from phenol, 1,2,3-THB from catechol, 1,2,4-THB from resorcinol, 1,2,4-THB from hydroquinone, and 3-methylcatechol from *m*-cresol), and protein engineering was used to make mutants that synthesize eight novel, industrially significant products (resorcinol, hydroquinone, 4-methylresorcinol, methylhydroquinone, 1,2,3-THB, 1,2,4-THB, 3-methylcatechol, and 4-methylcatechol)[128].

Several regiospecific mutants with high rates of activity were identified[128]; for example, *E. coli* TG1/pBS(Kan)ToMO expressing TouA saturation mutagenesis variant F205G formed 4-methylresorcinol (0.78 nmol/min/mg protein), 3-methylcatechol (0.25 nmol/min/mg protein), and methylhydroquinone (0.088 nmol/min/mg protein) from *o*-cresol whereas wild-type ToMO formed only 3-methylcatechol (1.1 nmol/min/mg protein) (Figure 11). From *o*-cresol, saturation mutagenesis mutant I100Q and DNA shuffling mutant M180T/E284G formed methylhydroquinone (0.50 and 0.19 nmol/min/mg protein, respectively) and 3-methylcatechol (0.49 and 1.5 nmol/min/mg protein, respectively). F205G formed catechol (0.52 nmol/min/mg protein), resorcinol (0.090 nmol/min/mg protein), and hydroquinone (0.070 nmol/min/mg protein) from phenol whereas wild-type ToMO formed only catechol (1.5 nmol/min/mg protein). Both I100Q and M180T/E284G formed hydroquinone (1.2 and 0.040 nmol/min/mg protein, respectively) and catechol (0.28 and 2.0 nmol/min/mg protein, respectively) from phenol. Dihydroxybenzenes were further oxidized to THBs with different regiospecificities; for example, I100Q formed 1,2,4-THB from catechol whereas wild-type ToMO formed 1,2,3-THB (pyrogallol). Regiospecific oxidation of the natural substrate toluene was also checked, for example, I100Q forms 22%, 44%, and 34% of *o*-, *m*-, and *p*-cresol, respectively, whereas wild-type ToMO forms 32%, 21%, and 47% of *o*-, *m*-, and *p*-cresol, respectively.

11. CONTROL OF THE REGIOSPECIFIC HYDROXYLATION OF NITRO AROMATICS

Nitrocatechols have been found to be useful intermediates for the synthesis of pharmaceuticals such as Flexinoxan, an antihypertensive drug[37,104]. Recently, nitrocatechol compounds were discovered as potent inhibitors of catechol-*o*-methyltransferase and are under clinical evaluation for the treatment of Parkinson's disease and other nervous system disorders[55,56]. In another study, 4-nitrocatechol and 3-nitrocatechol were found to be competitive inhibitors of nitric oxide synthase with potential antinociceptive (pain relieving) activity[85]; however, there is only one worldwide source of 3-nitrocatechol and its

price is $20,000/g (Vitas-M Laboratory, Ltd.). Furthermore, nitrohydroquinone has been used to synthesize dephostatin, an inhibitor of the protein tyrosine phosphatase[36] which is a candidate therapeutic agent for diabetes mellitus and neural diseases such as Alzheimer's disease and Parkinson's disease[121].

As chemical synthesis of these compounds is problematic in terms of yield and selectivity[85], the utilization of oxygenases is advantageous. The high redox potential of oxygenases enables them to perform reactions with chemically stable substrates as well as provide a high degree of regio- and enantio-selectivity[10,58]. Transforming selectively an inexpensive and abundant chemical as nitrobenzene into a valuable feedstock for drug production (e.g. 4-nitrocatechol) is therefore of great significance. Here we summarize our work on the mutagenesis of both T4MO and ToMO for nitroaromatics.

After discovering that T4MO of *P. mendocina* KR1 oxidizes nitrobenzene to 4-nitrocatechol, albeit at a very low rate, this reaction was improved using directed evolution and saturation mutagenesis. Screening 550 colonies from a random mutagenesis library generated by error-prone PCR of *tmoAB* using *E. coli* TG1/pBS(Kan)T4MO on agar plates containing nitrobenzene led to the discovery of nitrocatechol-producing mutants. One mutant, NB1, contained six amino acid substitutions (TmoA Y22N, I84Y, S95T, I100S, S400C; TmoB D79N). It was believed that position I100 of the alpha subunit of the hydroxylase (TmoA) is the most significant for the change in substrate reactivity due to previous results in our lab with a similar enzyme, TOM of *B. cepacia* G4[15].

Saturation mutagenesis at this position resulted in the generation of two more nitrocatechol mutants, I100A and I100S; the rate of 4-nitrocatechol formation by I100A was more than 16 times higher than that of wild-type T4MO at 200 μM nitrobenzene (0.13 ± 0.01 vs. 0.008 ± 0.001 nmol/min/mg protein). HPLC and mass spectrometry analysis revealed that variants NB1, I100A, and I100S produce 4-nitrocatechol via *m*-nitrophenol, while the wild-type produces primarily *p*-nitrophenol and negligible amounts of nitrocatechol. Relative to wild-type T4MO, whole cells expressing variant I100A convert nitrobenzene into *m*-nitrophenol with a V_{max} of 0.61 ± 0.037 vs. 0.16 ± 0.071 nmol/min/mg protein and convert *m*-nitrophenol into nitrocatechol with a V_{max} of 3.93 ± 0.26 vs. 0.58 ± 0.033 nmol/min/mg protein.

Hence the regiospecificity of nitrobenzene oxidation was changed by the random mutagenesis, and this led to a significant increase in 4-nitrocatechol production. The regiospecificity of toluene oxidation was also altered, and all of the mutants produced 20% *m*-cresol and 80% *p*-cresol, whereas the wild-type produces 96% *p*-cresol. Interestingly, the rate of toluene oxidation (the natural substrate of the enzyme) by I100A was also higher by 65% (7.2 ± 1.2 vs. 4.4 ± 0.3 nmol/min/mg protein). Homology-based modeling of TmoA suggests reducing the size of the side chain of I100 leads to an increase in the width of the active site channel which facilitates access of substrates and promotes more flexible orientations.

ToMO from *P. stutzeri* OX1 was found to oxidize nitrobenzene to form *m*-nitrophenol (72%) and *p*-nitrophenol (28%)[127]. It was also discovered that wild-type ToMO forms 4-nitrocatechol from *m*-nitrophenol and *p*-nitrophenol and 3-nitrocatechol (18%) and nitrohydroquinone (82%) from *o*-nitrophenol. To increase the oxidation rate and alter the oxidation regiospecificity of nitro aromatics as well as to study the role of the active site residues, DNA shuffling and saturation mutagenesis were performed on the alpha hydroxylase fragment of ToMO (TouA). Several mutants with higher rates of activities and with different regiospecificities were identified; for example, *Escherichia coli* TG1 cells expressing either TouA mutant M180T/E284G or E214G/D312N/M399V produce 4-NC 4.5-fold and 20-fold faster than wild-type ToMO from *p*-nitrophenol, respectively. TouA mutant A107T/E214A had the regiospecificity of nitrobenzene changed significantly from 28% to 79% *p*-nitrophenol. From 200 μM nitrobenzene, TouA variants A101T/M114T, A110T/E392D, M180T/E284G, and E214G/D312N/M399V produce 4-NC whereas wild-type ToMO does not. From *m*-nitrophenol, TouA mutant I100Q produces 4-nitrocatechol (37%) and nitrohydroquinone (63%), whereas wild-type ToMO produces only 4- nitrocatechol (100%). Variant A107T/E214A acts like a *para* enzyme and forms *p*-cresol as the major product (93%) from toluene with enhanced activity (2.3 fold), whereas wild-type ToMO forms 32%, 21%, and 47% of *o*-, *m*-, and *p*-cresol, respectively. Hence, the non-specific ToMO was converted into a regiospecific enzyme, which rivals T4MO of *Pseudomonas mendocina* KR1 and TOMO of *Burkholderia cepacia* G4 in its specificity.

12. REGIOSPECIFIC HYDROXYLATION OF NAPHTHALENE AND FLUORENE

Naphthalene and fluorene (Figure 12) serve as model bi- and tri-cyclic aromatic hydrocarbons for understanding the properties of a large class of PAHs[17], and there is no report of producing 2-naphthol with enzymes. Hence, we explored the regiospecific oxidation of naphthalene and fluorene with *E. coli* strains expressing wild-type T4MO, TpMO, TOM, ToMO, and T4MO variants (Table 5 and Figure 12)[117].

T4MO oxidized toluene, naphthalene, and fluorene faster than the other wild-type enzymes (2- to 22-fold) and produced a mixture of 1-naphthol (52%) and 2-naphthol (48%) from naphthalene which was successively transformed to a mixture of 2,3-, 2,7-, 1,7-, and 2,6-dihydroxynaphthalenes (7%, 10%, 20%, and 63%, respectively). TOM and ToMO made 1,7-dihydroxynaphthalene from 1-naphthol, and ToMO made a mixture of 2,3-, 2,6-, 2,7-, and 1,7-dihydroxynaphthalene (26%, 22%, 1%, and 44%, respectively) from

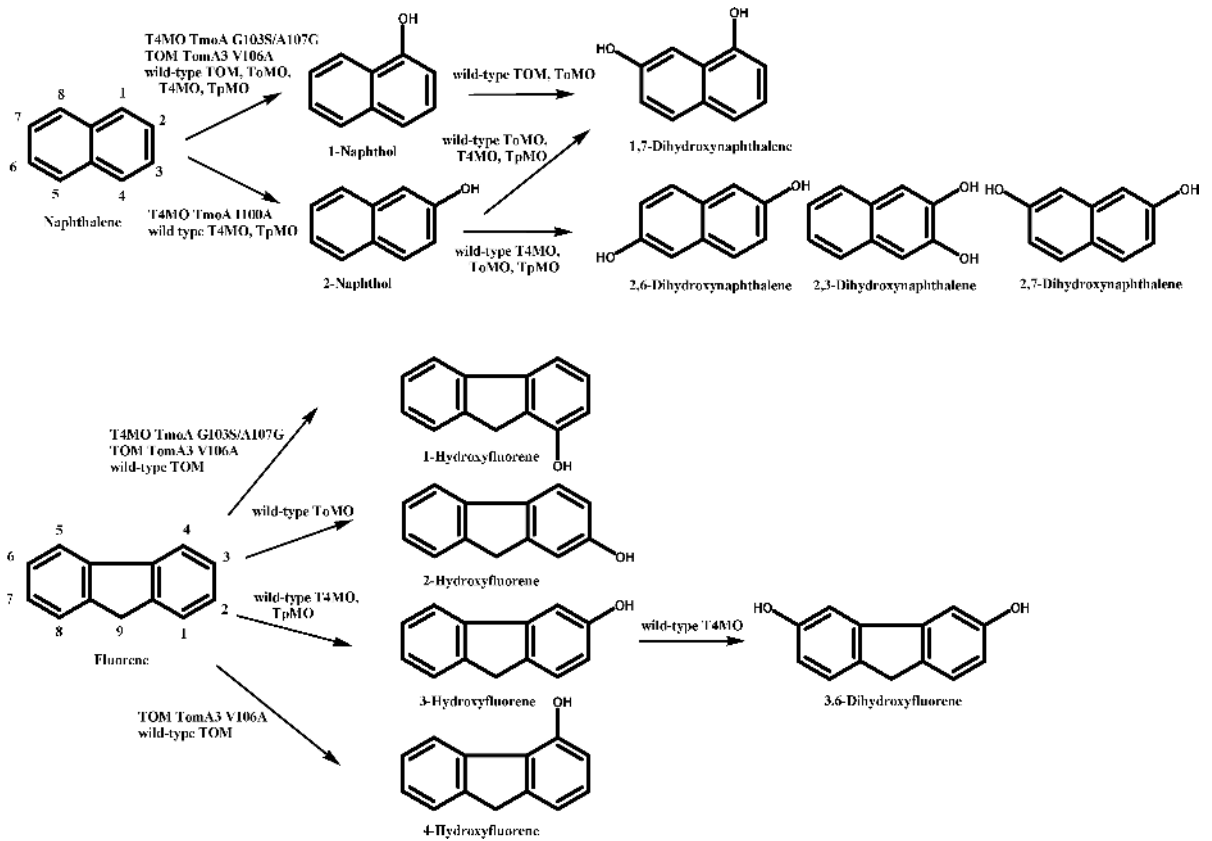


Figure 12. Pathway of naphthalene and fluorene oxidation by wild-type toluene monooxygenases, TOM variant TomA3 V106A, and T4MO TmoA variants.

Table 5. Aromatic oxidation rate and regiospecificity by TG1 cells expressing wild-type T4MO, ToMO, TOM, TOM TomA3 V106A, and T4MO TmoA variants.

	Toluene oxidation[a]				Naphthalene oxidation[b]			Fluorene oxidation[c]				
		Regiospecificity (%)				Regiospecificity (%)			Regiospecificity (%)			
Enzyme	Rate, nmol/min/mg protein	*o*-cresol	*m*-cresol	*p*-cresol	Rate, nmol/min/mg protein	1-naphthol	2-naphthol	Rate, nmol/min/mg protein	1-HF	2-HF	3-HF	4-HF
Wild-type TOM	2.5 ± 0.1	>98	0	0	0.9 ± 0.2	99	<1	0.10 ± 0.01	12	26	21	41
TOM TmoA3 V106A	4.8 ± 1.1	50	33	17	3.4 ± 0.5	99	<1	0.17 ± 0.02	33	7	8	51
Wild-type ToMO	6.1 ± 0.1	32	21	47	4.2 ± 2.3	90	10	0.03 ± 0.01	4	29[d]	67[d]	0
Wild-type TpMO	4.1 ± 0.2	0	10	90	0.9 ± 0.2	63	37	0.21 ± 0.01	0	14[d]	86[d]	0
Wild-type T4MO	12.1 ± 0.8	<1	3	96	7.7 ± 1.5	52	48	0.68 ± 0.04	0	14[d]	86[d]	0
TmoA I100A	19.1 ± 0.8	0	20	80	7.0 ± 1.5	8	92	0.57 ± 0.03	0	16[d]	84[d]	0
TmoA I100S	22.7 ± 1.6	0	20	80	5.5 ± 1.6	5	95	0.49 ± 0.13	0	17[d]	83[d]	0
TmoA I100G	15.1 ± 2.3	0	13	87	7.6 ± 0.5	12	88	0.57 ± 0.11	0	17[d]	83[d]	0
TmoA I100L	17.7 ± 0.2	7	3	90	3.2 ± 1.5	83	17	0.38 ± 0.18	0	13[d]	87[d]	0
TmoA G103S/A107G	1.5 ± 0.3	82	7	11	0.5 ± 0.1	99	<1	0.15 ± 0.04	28	18	53	0
TmoA G103A	20.1 ± 0.4	12	13	75	5.6 ± 1.5	81	19	0.38 ± 0.04	0	22[d]	78[d]	0
TmoA G103S	18.1 ± 1.7	9	15	76	5.9 ± 0.1	83	17	0.23 ± 0.02	0	14[d]	86[d]	0

[a] Toluene oxidation rate and regiospecificity determined via GC with 109 μM toluenecalculated based on Henry's law.

[b] Total naphthol formation rate and regiospecificity determined via HPLC with 5 mM (naphthalene solubility is 0.23 mM in water).

[c] Fluorene oxidation rate determined via HPLC and regiospecificity determined via GC except for TOM and TOM TomA3 V106A where regiospecificity was determined via GC-MS. Activity determined at the saturation concentration of 0.2 mM so the initial rate values represent V_{max}. Hf, hydroxy fluovene.

[d] Regiospecificity confirmed by HPLC with a supelcosil-ABZ alky-amide column.

2-naphthol. TOM had no activity on 2-naphthol, and T4MO had no activity on 1-naphthol. To take advantage of the high activity of wild-type T4MO but to increase its regiospecificity on naphthalene, 7 engineered enzymes containing mutations in T4MO alpha hydroxylase TmoA were examined; the selectivity for 2-naphthol by T4MO I100A, I100S, and I100G was enhanced to 88%–95%, and the selectivity for 1-naphthol was enhanced to 87% and 99% by T4MO I100L and G103S/A107G, respectively, while high oxidation rates were maintained except for G103S/A107G (Table 5). Therefore, the regiospecificity for naphthalene oxidation was altered to practically pure 1-naphthol or 2-naphthol. All four wild-type monooxygenases were able to oxidize fluorene to different mono-hydroxylated products; T4MO oxidized fluorene successively to 3-hydroxyfluorene and 3,6-dihydroxyfluorene, which were confirmed by GC-mass spectrometry and ^{1}H nuclear magnetic resonance analysis. TOM and its variant TomA3 V106A oxidize fluorene to a mixture of 1-, 2-, 3-, and 4-hydroxyfluorene. This is the first report of using enzymes to synthesize 1-, 3-, and 4-hydroxyfluorene, and 3,6-dihydroxyfluorene from fluorene as well as 2-naphthol and 2,6-dihydroxynaphthalene from naphthalene.

13. TWO-PHASE REACTORS FOR SYNTHESIZING PHENOL AND 2-NAPHTHOL

Phenol and naphthol compounds are important precursors for the manufacture of many dyes, drugs, perfumes, insecticides, and surfactants (e.g. pigment intermediate Tobias acid from 2-naphthol)[41]. World production for phenol is huge (6,600,000 ton/year) and that of 1-naphthol and 2-naphthol is around 15,000 and 50,000 ton/year, respectively[41,78]. Industrially, 90% of the world phenol production is manufactured by a three-step, cumene process[59] that relies on the co-production of acetone and has low yields[78]. Hence, attempts to direct conversion of benzene to phenol by one-step reactions have been investigated[51,59,63,69,88]. However, the remaining significant challenge in the chemical synthesis of phenol is the low conversion of benzene per pass of these processes ($<$10%).

Current commercial methods of manufacturing 2-naphthol include oxidation and aromatization of tetralin, caustic fusion of sodium 1-naphthalenesulfonate with sodium hydroxide in a five-step reaction, and hydroxylation of 2-isopropylnaphthalene[41]. The selectivity for 2-naphthol based on naphthol isomer distribution using those chemical methods is 90% or better. However, chemical synthesis is hampered by toxic reagents such as naphthalene-1-sulfonic acid, high concentration of acids and bases such as hydrogen fluoride, complicated, multiple-step reactions, extreme temperature conditions, and co-production of acetone[41]. For example, in order to obtain high selectivity (98%),

a fluorosulfuric acid–sulfuryl chloride fluoride solution at -60˚C to −78˚C was used for 2-naphthol production from naphthalene [82]. To our knowledge, there is no domestic producer of 2-naphthol in the United States since 1982 when the hydroperoxidation process ceased production[41] due to environmental problems, and China became the major supplier. It is predicted that production of 2-naphthol will be reduced if China intensifies its environmental protection policies, and the impact of a supply shortage may spread to the global market[18].

Since recombinant *E. coli* TG1 expressing wild-type T4MO performs successive hydroxylations of benzene to phenol, catechol, and 1,2,3-THB[119] (Figure 2), and *E. coli* TG1 expressing an evolved T4MO with alpha subunit mutation I100A may be used to hydroxylate naphthalene to 92% 2-naphthol (first production in a microbial system) and 8% 1-naphthol which are subsequently oxidized to dihydroxynaphthalenes[117] (Figure 12), we explored using a two-phase system for maximizing production of these important intermediates[116]. In the single-phase reaction, phenol production from benzene by T4MO was greatly reduced by the further conversion of phenol to catechol, and the production of 2-naphthol was limited by the toxicity of naphthalene and 2-naphthol to the biocatalyst and the low solubility of naphthalene in water. In contrast, the two-phase system allowed for the transformation of benzene to phenol and naphthalene to 2-naphthol using whole cells expressing wild-type T4MO and T4MO TmoA I100A. The solubility of naphthalene was enhanced and the toxicity of the naphthols was prevented by the use of a water (80 vol%)-dioctyl phthalate (20 vol%) system which yielded 21-fold more 2-naphthol using TG1/pBS(Kan)T4MOI100A (10.6 ± 1.3 mM in the two-phase system vs. 0.5 ± 0.05 mM in the single phase). More than 99% 2-naphthol was extracted to the dioctyl phthalate phase, further conversion of 2-naphthol was prevented, 92% 2-naphthol was formed, and 12% naphthalene was converted. In addition, using 50% vol dioctyl phthalate and an initial concentration of 39 mM, a 51 ± 9% conversion of benzene was obtained by TG1/pBS(Kan)T4MO, and phenol was produced at a purity of 97% (in the organic phase) with an average volumetric productivity of 0.4 ± 0.04 g phenol/$L_{aqueous}$/h. Relative to the one-phase system, there was a 12-fold reduction in byproduct catechol formation (96% phenol vs. 8% phenol after 3 h). This is the first report of using toluene monooxygenases in two-liquid phase systems for synthesizing industrially significant, phenolic compounds.

14. DEGRADATION OF PCE, TCE, AND CHLOROFORM

Highly chlorinated compounds are recalcitrant and threaten the environment. For example, tetrachloroethylene [perchloroethylene (PCE)] is one of the five most-frequently detected volatile organic compounds found in municipal

groundwater supplies[130] since it has been used extensively as an industrial degreasing solvent and fumigant[16]. PCE is also one of 14 volatile organic compounds on the United States EPA's Priority Pollutant List[16]. Because this solvent is toxic and is a suspected human carcinogen[57], PCE is regulated under the Safe Drinking Water Act to a maximum contaminant level of 5 parts per billion.

Bacterial degradation of PCE had never been achieved in the presence of oxygen[25,65,66] until work with TOMO[101]. However, degradation through oxygen attack of fully halogenated chlorotrifluoroethylene has been achieved by purified sMMO of *Methylosinus trichosporium* OB3b; the lack of PCE degradation by this enzyme was suggested to be due to steric hindrance[31]. Since fluorine is more electronegative and more tightly bound than chlorine, chlorotrifluoroethylene should be more difficult to oxidize than PCE. Therefore it is reasonable that a monooxygenase can degrade PCE.

It is very well established that PCE is degraded anaerobically via reductive dehalogenation to the less-chlorinated ethenes TCE, *trans*-1,2-dichloroethylene (*trans*-DCE), *cis*-1,2-dichloroethylene (*cis*-DCE), 1,1-dichloroethylene (1,1-DCE), vinyl chloride (VC), and ethene as well as to ethane[108]. The dechlorination of PCE is often incomplete when it does occur, with VC and *cis*-DCE formed primarily[66]; however, dehalorespiration of PCE to ethene is possible[65]. VC generated is a known human carcinogen[66] and both VC and *cis*-DCE are US EPA priority pollutants[5].

Surprisingly, the wastewater bacterium *P. stutzeri* OX1 was found to degrade aerobically 0.56 μmol of 2.0 μmol PCE in 21 h ($V_{max} \cong 2.5$ nmol/min/mg protein and $K_M \cong 34$ μM)[101]. These results were corroborated by the generation of 0.48 μmol of the degradation product, chloride ions. This degradation was confirmed to be a result of expression of ToMO since cloning and expressing this enzyme in *E. coli* led to the aerobic degradation of 0.19 μmol of 2.0 μmol PCE and the generation of stoichiometric amounts of chloride. PCE also induces formation of ToMO which leads to its degradation in *P. stutzeri* OX1[101]. Since this strain is also chemotactic toward PCE[125], that means a single bacterium will move toward this pollutant, the pollutant will induce the genes for the degradation enzyme, and will subsequently be degraded.

TCE is the most-frequently detected groundwater contaminant; hence, directed evolution was used to increase the activity of TOM of *B. cepacia* G4 for chlorinated ethenes. When expressed in *E. coli,* the alpha subunit variant V106A (TOM-Green) degrades more rapidly TCE (2.5 ± 0.3 vs. 1.39 ± 0.05 nmoles/min/mg protein), 1,1-DCE, and *trans*-DCE. These results show clearly that random, *in vitro* protein engineering can be used to improve a large multisubunit protein for multiple functions including environmental restoration.

Mutagenesis of the alpha subunit of ToMO also led to enhanced TCE degradation[126]. ToMO TouA saturation mutagenesis variant I100Q was

identified that degrades TCE 2.8 times better than the wild-type enzyme. Another variant, E214G/D312N/M399V, degrades *cis*-DCE 2.5 faster, and variant M180T/E284G also has enhanced *cis*-DCE activity.

Like TCE, chloroform is a widely used chlorinated solvent and a common groundwater contaminant due to improper disposal and accumulation from anaerobic dehalogenation of carbon tetrachloride at contaminated sites[23]. Chloroform is also a suspected carcinogen[67]. Whole cells expressing TOM are also able to degrade mixtures of chlorinated aliphatics including TCE and chloroform[111]. To enhance chloroform degradation, saturation mutagenesis was performed on TOM at alpha subunit position V106 since this position had been identified by DNA shuffling to be important for the degradation of chlorinated ethenes[15]. Variant A106F was identified that has 2.8 times better chloroform degradation activity based on GC (V_{max} of 2.61 vs. 0.95 nmol/min/mg protein and unchanged K_m) and chloride release (0.034 ± 0.002 vs. 0.012 ± 0.001 nmol/min/mg protein)[98]. Variant A106F also has similar protein expression as A106 and 62% better toluene oxidation activity than wild-type TOM (2.11 ± 0.3 vs. 1.30 ± 0.06 nmol/min/mg protein)[98].

15. METABOLIC ENGINEERING FOR DEGRADING CHLORINATED ETHENES

Based on the Gibbs free energy calculation of aerobic degradation of chlorinated ethenes to water, carbon dioxide and HCl, growth on nearly all chlorinated aliphatics is thermodynamically possible (e.g. the Gibbs free energy change for the aerobic mineralization of *cis*-DCE is -1143 kJ/mol)[123]. If the reactive intermediates may be effectively detoxified, the main biochemical factor that hampers chlorinated ethenes from supporting cell growth is the lack of appropriate enzymes to harvest their energy. Hence it is feasible to construct bacteria which grow on these chlorinated ethenes and create a niche for their aerobic degradation. Currently, the aerobic degradation of chlorinated ethenes is fortuitous and provides no benefit to the cell; in fact, it leads to cell death and thus is selected against (Figure 13).

The toxic epoxides generated during the aerobic biodegradation of chlorinated ethenes limit their transformation. Hydrolysis of the toxic epoxide by epoxide hydrolases represents one of the major biological detoxification strategy; however, chlorinated epoxyethanes are not accepted by known bacterial epoxide hydrolase. The epoxide hydrolase from *Agrobacterium radiobacter* AD1 (EchA), which enables growth on epichlorohydrin, was tuned by us to accept *cis*-1,2-dichloroepoxyethane as a substrate by accumulating beneficial mutations from three rounds of saturation mutagenesis at three selected active site residues: F108, I219, and C248[97]. The EchA F108L/I219L/C248I variant

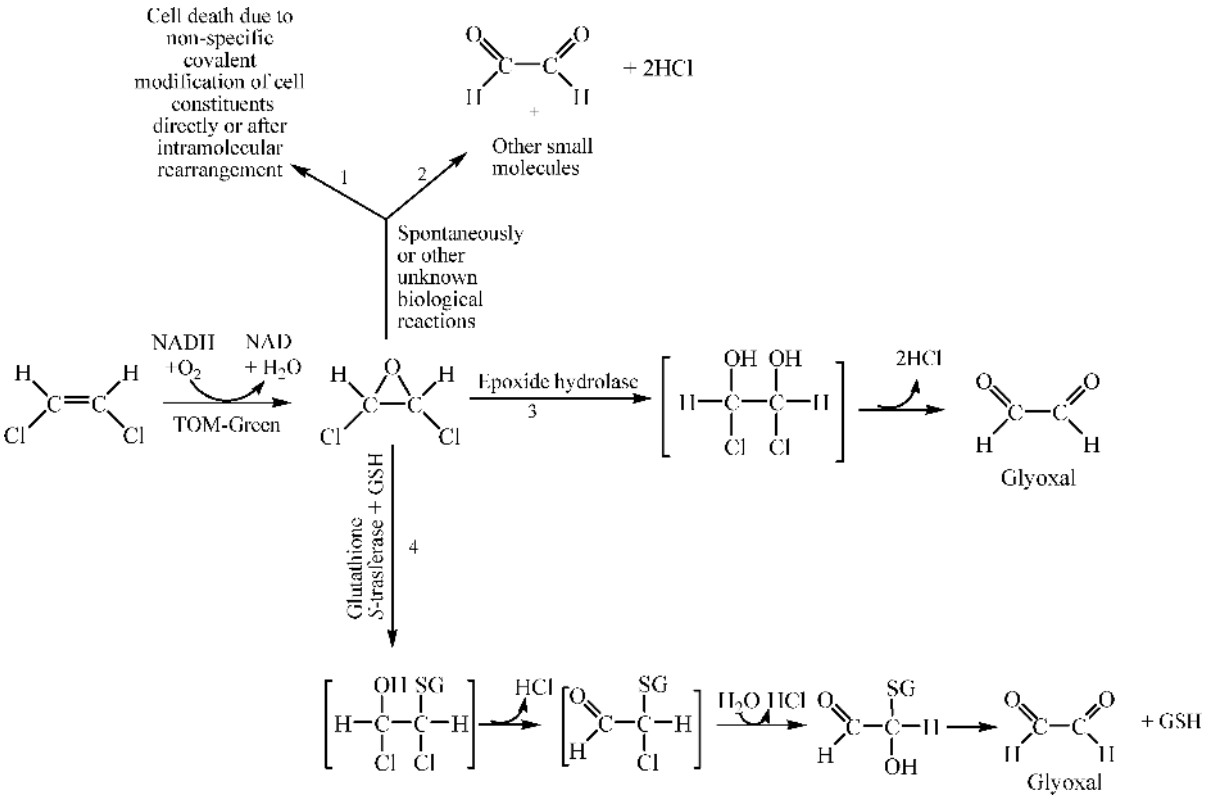


Figure 13. Metabolic engineering to enhance *cis*-DCE mineralization by cloning an evolved epoxide hydrolase (EchA) or glutathione *S*-transferase (IsoILR1) along with an evolved TOM (TOM-Green) (adapted from van Hylckama Vlieg and Janssen[123]). Steps 1 and 2 are the two possible spontaneous transformation pathways for *cis*-DCE epoxide, while step 3 and step 4 represent two major detoxification strategies in with *cis*-DCE epoxide may be biologically converted by either an epoxide hydrolase or glutathione *S*-transferase (IsoILR1).

co-expressed with DNA shuffled TOM (TOM-Green, alpha subunit V106A mutation), which initiates attack on the chlorinated ethene, allowed for the degradation of *cis*-dichloroethylene (*cis*-DCE) at low concentrations (6.8 μM) (wild-type EchA has no activity at this concentration) and enhanced degradation 10 fold at high concentrations (540 μM). EchA variants with single mutations (F108L, I219F, or C248I) enhanced *cis*-DCE mineralization 2.5 fold (540 μM), and EchA variants with double mutations, I219L/C248I and F108L/C248I, increased *cis*-DCE mineralization four-fold and seven-fold, respectively (540 μM). For complete degradation of *cis*-DCE to chloride ions, the apparent V_{max}/K_m for the recombinant *Escherichia coli* strain expressing the EchA F108L/I219L/C248I variant was increased over five-fold as a result of the evolution of EchA. The EchA F108L/I219L/C248I variant also had enhanced activity for 1,2-epoxyhexane (two-fold) and the natural substrate epichlorohydrin (six-fold)[97]. A similar approach using glutathione *S*-transferases instead of the engineered epoxide hydrolase has also been implemented by us (Figure 13)[99].

16. ENGINEERING DIOXYGENASES FOR SINGLE HYDROXYLATIONS OF AROMATICS

Like toluene monooxygenases, dioxygenases may be evolved for monohydroxylation of aromatics. Saturation mutagenesis of the 2,4-dinitrotoluene dioxygenase of *Burkholderia cepacia* R34 (DDO) at position valine 350 of the DntAc α-subunit generated mutant V350F with significantly increased activity toward *o*-nitrophenol (47 times), *m*-nitrophenol (34 times), and *o*-methoxyphenol (174 times) as well as an expanded substrate range that now includes *m*-methoxyphenol, *o*-cresol, and *m*-cresol (wild-type DDO had no detectable activity for these substrates)[47]. V350F produced both nitrohydroquinone at a rate of 0.75 ± 0.15 nmol/min/mg protein and 3-nitrocatechol at a rate of 0.069 ± 0.001 nmol/min/mg protein from *o*-nitrophenol, 4-nitrocatechol from *m*-nitrophenol at 0.29 ± 0.02 nmol/min/mg protein, methoxyhydroquinone from *o*-methoxyphenol at 2.5 ± 0.6 nmol/min/mg protein, methoxyhydroquinone from *m*-methoxyphenol 0.55 ± 0.02 nmol/min/mg protein, both methylhydroquinone at 1.52 ± 0.02 nmol/min/mg protein and 2-hydroxybenzyl alcohol at 0.74 ± 0.053 nmol/min/mg protein from *o*-cresol, and methylhydroquinone at 0.43 ± 0.12 nmol/min/mg protein from *m*-cresol. The DDO variant V350F also exhibited 10-fold enhanced activity towards naphthalene (8.2 ± 2.6 nmol/min/mg protein), forming (1*R*, 2*S*)-*cis*-1,2-dihydro-1,2-dihydroxynaphthalene. Hence, mutagenesis of the wild-type DDO through active site engineering generated variants with relatively high rates toward a previously uncharacterized class of substituted phenols for the nitroarene

dioxygenases; seven previously uncharacterized substrates were evaluated for the wild-type DDO, and four novel monooxygenase-like products were found for the DDO variant V350F (methoxyhydroquinone, methylhydroquinone, 2-hydroxybenzyl alcohol, and 3-nitrocatechol)[47].

Single hydroxylations have been reported previously for dioxygenases; for example, formation of phenol from methoxybenzene using naphthalene dioxygenase (NDO) of *Pseudomonas* sp. strain NCIB 9816-4 has been previously reported by Resnick *et al.*[92], and NDO is also capable of other monooxygenase-like reactions such as the oxidation of indene to (+)-(1*S*)-indenol[92]. However, NDO has been reported to be unable to catalyze the oxidation of the aromatic nucleus[92]. Toluene dioxygenase (TOD) from *P. putida* F1 has also been reported to produce 4-methylcatechol from *p*-cresol and 3-nitrocatechol from *m*-nitrophenol[112]; however, *o*- and *m*-cresol or *o*- and *p*-nitrophenol have not been reported to be substrates for TOD[112]. Because of these limitations in TOD and NDO, DDO was studied here.

17. CONCLUSIONS AND FUTURE DIRECTIONS

Protein engineering of toluene monooxygenases is beginning to reach the point that the regiospecificity of hydroxylation may be controlled so that a myriad of industrially relevant chemicals may be synthesized, many of which had not been synthesized previously with bacteria (e.g. 2-naphthol), and bacteria may be engineered to help remediate contaminated sites. Furthermore, the structure/function lessons learned from mutagenesis of these non-heme enzymes should be directly applicable to intractable non-heme enzymes such as sMMO of *Methylosinus trichosporium* OB3b. Active sMMO has only been expressed in *Agrobacterium tumefaciens, Rhizobium meliloti*, and pseudomonads by our group but never in *E. coli*[43,44]; hence, there are no reports of beneficial mutations of the hydroxylase for this enzyme, the best enzyme for oxidizing methane to methanol. By understanding toluene monooxygenases, we may begin to understand how the C–H bond of methane is activated by sMMO. The random mutagenesis approaches described here may also be used now to study the other, non-hydroxylase components of these multi-component enzymes to gain an even deeper understanding of the catalytic mechanism.

ACKNOWLEDGMENTS

This research was supported by the National Science Foundation (BES-9911469, BES-0331416, and BES-0124401), the US Environmental Protection Agency, and the US Army Research Office (DAAD19-00-1-0568).

REFERENCES

1. Azerad, R., 2001, Chemical biotechnology better enzyme for green chemistry. *Curr. Opin. Biotechnol.*, 12:533–534.
2. Bainbridge, G., Anralojc, P.J., Madgwick, P.J., Pitts, J.E., and Parry, M.A.J., 1998, Effect of mutation of lysine-128 of the large subunit of ribulose bisphosphate carboxylase/oxygenase from *Anacystis nidulans*. *Biochem. J.*, 336:387–393.
3. Bertoni, G., Bolognese, F., Galli, E., and Barbieri. P., 1996, Cloning of the genes for and characterization of the early stages of toluene and *o*-xylene catabolism in *Pseudomonas stutzeri* OX1. *Appl. Environ. Microbiol*., 62:3704–3711.
4. Bhushan, B., Samanta, S.K., and Jain. R.K., 2000, Indigo production by naphthalene-degrading bacteria. *Lett. Appl. Microbiol.*, 31:5–9.
5. Bradley, P.M., and Chapelle. F.H., 1998, Effect of contaminant concentration on aerobic microbial mineralization of DCE and VC in stream-bed sediments. *Environ. Sci. Technol*., 32:553–557.
6. Branden, C., and Tooze. J., 1999, *Introduction to Protein Structure*, 2nd ed. Garland Publishing Inc., New York, USA.
7. Buhler, B., Witholt, B., Hauer, B., and Schmid. A., 2002, Characterization and application of xylene monooxygenase for multistep biocatalysis. *Appl. Environ. Microbiol.*, 68:560–568.
8. Bui, V.P., Hansen, T.V., Stenstrom, Y., and Hudlicky. T., 2000, Direct biocatalytic synthesis of functionalized catechols: a green alternative to traditional methods with high effective mass yield. *Green Chem.*, 2:263–265.
9. Buolamwini, J.K., 2000, Cell cycle molecular targets in novel anticancer drug discovery. *Curr. Pharm. Des.*, 6:379–392.
10. Burton, S.G., 2003, Oxidizing enzymes as biocatalysts. *Trends Biotechnol.*, 21:543–549.
11. Bylina, E.J., Coleman, W.J., Dilworth, M.R., Robles, S.J., Tanner, M.A., Yang, M.M., and Youvan, D.C., 2000, Solid-phase enzyme screening. *ASM News*, 66:211–217.
12. Byrne, AM., Kukor, J.J., and Olsen. R.H., 1995, Sequence analysis of the gene cluster encoding toluene-3-monooxygenase from *Pseudomonas pickettii* PKO1. *Gene*, 154:65–70.
13. Byrne, A.M., and Olsen. R.H., 1996, Cascade regulation of the toluene-3-monooxygenase operon (*tbuA1UBVA2C*) of *Burkholderia pickettii* PKO1: role of the *tbuA1* promoter (*PtbuA1*) in the expression of its cognate activator, TbuT. *J. Bacteriol.*, 178:6327–6337.
14. Cafaro, V., Izzo, V., Scognamiglio, R., Notomista, E., Capasso, P., Casbarra, A., Pucci, P., and Di Donato A., 2004, Phenol hydroxylase and toluene/*o*-xylene monooxygenase from *Pseudomonas stutzeri* OX1: interplay between two enzymes. *Appl. Environ. Microbiol*., 70:2211–2219.
15. Canada, K.A., Iwashita, S., Shim, H., and Wood. T.K., 2002, Directed evolution of toluene *ortho*-monooxygenase for enhanced 1-naphthol synthesis and chlorinated ethene degradation. *J. Bacteriol*., 184:344–349.
16. Carter, S.R., and Jewell. W.J., 1993, Biotransformation of tetrachloroethylene by anaerobic attached-films at low temperatures. *Water Res*., 27:607–615.
17. Casellas, M., Grifoll, M., Bayona, J.M., and Solanas. A.M., 1997, New metabolites in the degradation of fluorene by *Arthrobacter* sp. strain F101. *Appl. Environ. Microbiol.* 63:819–826.
18. Cheng, L. Feb 16 1999, Forthcoming production reduction of 2-naphthol for environmental protection. *China Chem. Rep.* Accessed from High Beam Research, Inc., July 15 2004 <www.highbeam.com>
19. Dolfing, J., van den Wijngaard, A.J., and Janssen. D.B., 1993, Microbiological aspects of the removal of chlorinated hydrocarbons from air. *Biodegradation*, 4:261–282.

20. Draths, K.M., and Frost. J.W., 1991, Conversion of D-glucose into catechol: the not-so-common pathway of aromatic biosynthesis. *J. Am. Chem. Soc*. 113:9361–9363.
21. Draths, K.M., and Frost. J.W., 1995, Environmentally compatible synthesis of catechol from D-glucose. *J. Am. Chem. Soc*. 117:2395–2400.
22. Dressler, H. 1994, Resorcinol, Its Uses and Derivatives, 1st ed. Plenum Press, New York.
23. Egli, C., Scholtz, R., Cook, A.M., and Leisinger. T., 1987, Anaerobic dechlorination of tetrachloromethane and 1,2-dichloroethane to degradable products by pure cultures of *Desulfobacterium* sp. and *Methanobacterium* sp. *FEMS Microbiol. Lett.*, 43:257–261.
24. Elango, N., Radhakrishnan, R., Froland, W.A., Wallar, B.J., Earhart, C.A., Lipscomb, J.D., and Ohlendorf. D.H., 1997, Crystal structure of the hydroxylase component of methane monooxygenase from *Methylosinus trichosporium* OB3b. *Protein. Sci.*, 6:556–568.
25. Ensley, B.D., 1991, Biochemical diversity of trichloroethylene metabolism. *Annu. Rev. Microbiol.,* 45:283–299.
26. Ensley, B.D., Ratzkin, B.J., Osslund, T.D., Simon, M.J., Wackett, L.P., and Gibson. D.T., 1983, Expression of naphthalene oxidation genes in *Escherichia coli* results in the biosynthesis of indigo. *Science*, 222:167–169.
27. Fishman, A., Tao, Y., Bentley, W., and Wood., T.K., 2004, Protein engineering of toluene 4-monooxygenase of *Pseudomonas mendocina* KR1 for synthesizing 4-nitrocatechol from nitrobenzene. *Biotechnol. Bioeng*., 87:779–790.
28. Fishman, A., Tao, Y., Rui, L., and Wood. T.K., 2005, Controlling the regiospecific oxidation of aromatics via active site engineering of toluene *para*-monooxygenase of *Ralstonia pickettii* PKO1. *J. Biol. Chem*., 290:506–514.
29. Fishman, A., Tao, Y., and Wood. T.K., 2004, Physiological relevance of successive hydroxylations of toluene by toluene *para*-monooxygenase of *Ralstonia pickettii* PKO1. Biocatalysis Biotransformation, 22:283–289.
30. Fishman, A., Tao, Y., and Wood. T.K., 2004, Toluene 3-monooxygenase of *Ralstonia pickettii* PKO1 is a *para*-hydroxylating enzyme. *J. Bacteriol.*, 186:3117–3123.
31. Fox, B.G., Borneman, J.G., Wackett, L.P., and Lipscomb. J.D., 1990, Haloalkene oxidation by the soluble methane monooxygenase from *Methylosinus trichosporium* OB3b: mechanistic and environmental implications. *Biochem*., 29:6419–6427.
32. Frost, J.W., and Lievense. J., 1994, Prospects for biocatalytic synthesis of aromatics in the 21st century. *New J. Chem.*, 18:341–348.
33. Gillam, E.M.J., Notley, L.M., Cai, H., De Voss J.J., and Guengerich. F.P., 2000, Oxidation of indole by cytochrome P450 enzymes. *Biochem.*, 39:13817–13824.
34. Gillner, M., Moore, G.S., Cederberg, H., and Gustafsson. K., 1994, Hydroquinone, environmental health criteria 157. *International Programme on Chemical Safety*.
35. Guex, N., and Peitsch. M.C., 1997, SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modelling. *Electrophoresis*, 18:2714–2723.
36. Hamel, P., and Girard. Y., 1994, Synthesis of Dephostatin, a novel protein tyrosine phosphatase inhibitor. *Tetrahedron Lett.*, 35:8101–8102.
37. Hartog, J., and Wouters. W., 1988, Flesinoxan hydrochloride. Drugs future 1:31–33.
38. Held, M., Suske, W., Schmid, A., Engesser, K.-H., Kohler, H.-P.E., Witholt, B., and Wubbolts, M.G., 1998, Preparative scale production of 3-substituted catechols using a novel monooxygenase from *Pseudomonas azelaica* HBP1. *J. Mol. Catalysis B: Enzymatic*, 5:87–93.
39. Hisaindee, S., and Clive, D.L.J., 2001, A synthesis of puraquinonic acid. *Tetrahedron Lett.*, 42:2253–2255.
40. Hoessel, R., Leclerc, S., Endicott, J.A., Noble, M.E.M., Lawrie, A., Tunnah, P., Leost, M., Damiens, E., Marie, D., Marko, D., Niederberger, E., Tang, W., Eisenbrand, G., and Meijer. L., 1999, Indirubin, the active constituent of a Chinese antileukaemia medicine, inhibits cyclin-dependent kinases. *Nat. Cell Biol.*, 1:60–67.

41. M. Howe-Grant (ed.), 1991, Kirk-Othmer Encyclopedia of Chemical Technology, 4th ed, Vol. 13. Wiley-Interscience Publishers, New York.
42. Hua, D.H., Tamura, M., Huang, X., Stephany, H.A., Helfrich, B.A., Perchellet, E.M., Sperfslage, B.J., Perchellet, J.-P., Jiang, S., Kyle, D.E., and Chiang, P.K., 2002, Syntheses and bioactivities of substituted 9,10-dihydro-9,10-[1,2] benzenoanthracene-1,4,5,8-tetrones. Unusual reactivities with amines. *J. Org. Chem*., 67:2907–2912.
43. Jahng, D., Kim, C.S., Hanson, R.S., and Wood. T.K., 1996, Optimization of trichloroethylene degradation using soluble methane monooxygenase of *Methylosinus trichosporium* OB3b expressed in recombinant bacteria. *Biotechnol. Bioeng*., 51:349–359.
44. Jahng, D., and Wood. T.K., 1994, Trichloroethylene and chloroform degradation by a recombinant pseudomonad expressing soluble methane monooxygenase from *Methylosinus trichosporium* OB3b. *Appl. Environ. Microbiol*., 60:2473–2482.
45. Johnson, G.R., and Olsen. R.H., 1997, Multiple pathways for toluene degradation in *Burkholderia sp*. strain JS150. *Appl. Environ. Microbiol*., 63:4047–4052.
46. Kadiyala, V., and Spain, J.C., 1998, A two-component monooxygenase catalyzes both the hydroxylation of *p*-nitrophenol and the oxidative release of nitrite from 4-nitrocatechol in *Bacillus sphaericus* JS905. *Appl. Environ. Microbiol*., 64:2479–2484.
47. Keenan, B., Leungsakul, T., Smets, B., and Wood, T.K., 2004, Saturation mutagenesis of *Burkholderia cepacia* R34 2,4-DNT dioxygenase at DntAc Valine 350 for synthesizing nitrohydroquinone, methylhydroquinone, and methoxyhydroquinone. *Appl. Environ. Microbiol*., 70:3222–3231.
48. Kirk, O., Borchert, T.V., and Fuglsang, C.C., 2002, Industrial enzyme applications. *Curr. Opin. Biotechnol.*, 13:345–351.
49. Krolikowska, A., Bokszczanin, W., Kozlowski, A., and Dzikowicz, T., April 1991, Mixtures of dihydroxybenzene derivatives and alkylphenols as rust inhibitors for paints. Polland patent 153,464.
50. Kukor, J.J., and Olsen, R.H., 1992, Complete nucleotide sequence of *tbuD*, the gene encoding phenol/cresol hydroxylase from *Pseudomonas pickettii* PKO1, and functional analysis of the encoded enzyme. *J. Bacteriol.*, 174:6518–6526.
51. Laufer, W., Niederer, J.P.M., and Hoelderich, W.F., 2002, New direct hydroxylation of benzene with oxygen in the presence of hydrogen over bifunctional palladium/platinum catalysts. *Adv. Synth. Catalysis*, 344:1084–1089.
52. Leahy, J.G., Batchelor, P.J., and Morcomb, S.M., 2003, Evolution of the soluble diiron monooxygenases. *FEMS Microbiol. Rev*., 27:449–479.
53. Leahy, J.G., Byrne, A.M., and Olsen. R.H., 1996, Comparison of factors influencing trichloroethylene degradation by toluene-oxidizing bacteria. *Appl. Environ. Microbiol.*, 62:825–833.
54. Leahy, J.G., Johnson, G.R., and Olsen. R.H., 1997, Cross-regulation of toluene monooxygenases by the transcriptional activators TbmR and TbuT. *Appl. Environ. Microbiol*., 63:3736–3739.
55. Learmonth, D.A., and Freitas, A.P., 2002, Chemical synthesis and characterization of conjugates of a novel catechol-*o*-methyltransferase inhibitor. *Bioconj. Chem.*, 13:1112–1118.
56. Learmonth, D.A., Vieira-Coelho, M.A., Benes, J., Alves, P.C., Borges, N., Freitas, A.P., and Soares-da-Silva, P., 2002, Synthesis of 1-(3,4-dihydroxy-5-nitrophenyl)-2-phenyl-ethanone and derivatives as potent and long acting peripheral inhibitors of catechol-*o*-methyltransferase. *J. Med. Chem.*, 45:685–695.
57. Leung, S.W., Watts, R.J., and Miller, G.C., 1992, Degradation of perchloroethylene by Fenton's reagent: speciation and pathway. *J. Environ. Qual.*, 21:377–381.
58. Li, Z., van Beilen J.B., Duetz, W.A., Schmid, A., de Raadt, A., Grienge, H., and Witholt, B., 2002, Oxidative biotransformations using oxygenases. *Curr. Opin. Chem. Biol*., 6:136–144.

59. Liptakova, B., Bahidsky, M., and Hronec, M., 2004, Preparation of phenol from benzene by one-step reaction. Appl. Catalysis A: Gen. 263:33–38.
60. Low, D.W., and Hill, M.G., 2000, Backbone-engineered high-potential iron proteins: effects of active-site hydrogen bonding on reduction potential. *J. Am. Chem. Soc.*, 122:11039–11040.
61. Luu, P.P., Yung, C.W., Sun, A.K.-Y., and Wood, T.K., 1995, Monitoring trichloroethylene mineralization by *Pseudomonas cepacia* G4 PR1. *Appl. Microbiol. Biotechnol.*, 44:259–264.
62. Macias, F.A., Marin, D., Chinchilla, D., and Molinillo, J.M.G., 2002, First total synthesis of (+/−)-helibisabonol A. *Tetrahedron Lett*. 43:6417–6420.
63. Masumoto, Y., Hamada, R., Yokota, K., Nishiyama, S., and Tsuruya, S., 2002, Liquid-phase oxidation of benzene to phenol by vanadium catalysts in aqueous solvent with high acetic acid concentration. *J. Mol. Catalysis A: Chem.*, 184:215–222.
64. Maugard, T., Enaud, E., de La Sayette, A., Choisy, P., and Legoy, M.D., 2002, Beta-glucosidase-catalyzed hydrolysis of indican from leaves of *Polygonum tinctorium*. *Biotechnol. Prog.*, 18:1104–1108.
65. Maymo-Gatell, X., Chien, Y.-T., Gossett, J.M., and Zinder, S.H., 1997, Isolation of a bacterium that reductively dechlorinates tetrachloroethene to ethene. *Science*, 276:1568–1571.
66. McCarty, P.L., 1997, Breathing with chlorinated solvents. *Science*, 276:1521–1522.
67. McClay, K., Fox, B.G., and Steffan, R.J., 1996, Chloroform mineralization by toluene-oxidizing bacteria. *Appl. Environ. Microbiol*., 62:2716–2722.
68. McMurry, J., 2004, Organic Chemistry, 6th ed. Thomson Learning Inc., Belmont, CA, USA.
69. Meloni, D., Monaci, R., Rombi, E., Guimon, C., Martinez, H., Fechete, I., and Dumitriu. E., 2002, Synthesis and characterization of MCM-22 zeolites for the N2O oxidation of benzene to phenol. *Stud. Surface Sci. Catalysis*, 142A:167–174.
70. Meng, M., Lee, C., Bagdasarian, M., and Zeikus. J.G., 1991, Switching substrate preference of the thermophilic xylose isomerase from D-xylose to D-glucose by redesigning the substrate binding pocket. *Proc. Natl. Acad. Sci. USA,* 88:4015–4019.
71. Meyer, A., Schmid, A., Held, M., Westphal, A.H., Rothlisberger, M., Kohler, H.-P.E., van Berkel, W. J. H., and Witholt, B., 2002, Changing the substrate reactivity of 2-hydroxybiphenyl 3-monooxygenase from *Pseudomonas azelaica* HBP1 by directed evolution. *J. Biol. Chem*. 277:5575–5582.
72. Mitchell, K.H., 2002, Mutagenesis studies revealing the complementary roles of effector protein binding and the hydroxylase active site catalysis by toluene 4-monooxygenase. Ph.D. thesis. University of Wisconsin, Madison.
73. Mitchell, K.H., Studts, J.M., and Fox, B.G., 2002, Combined participation of hydroxylase active site residues and effector protein binding in a *para* to *ortho* modulation of toluene-4-monooxygenase regiospecificity. *Biochem*., 41:3176–3188.
74. Murdock, D., Ensley, B.D., Serdar, C., and Thalen, M., 1993, Construction of metabolic operons catalyzing the *de novo* biosynthesis of indigo in *Escherichia coli*. *Bio/Technol*., 11:381–386.
75. Nelson, M.J.K., Montgomery, S.O., Mahaffey, W.R., and Pritchard, P.H., 1987, Biodegradation of trichloroethylene and involvement of an aromatic biodegradative pathway. *Appl. Environ. Microbiol.*, 53:949–954.
76. Nelson, M.J.K., Montgomery, S.O., O'Neill, E.J., and Pritchard. P.H., 1986, Aerobic metabolism of trichloroethylene by a bacterial isolate. *Appl. Environ. Microbiol.*, 52:383–384.
77. Newman, L.M., and Wackett. L.P., 1995, Purification and characterization of toluene 2-monooxygenase from *Burkholderia cepacia* G4. *Biochem*., 34:14066–14076.
78. Niwa, S., Eswaramoorthy, M., Nair, J., Raj, A., Itoh, N., Shoji, H., Namba, T., and Mizukami. F., 2002, A one-step conversion of benzene to phenol with a palladium membrane. *Science*, 295:105–107.

79. Notomista, E., Lahm, A., Di Donato, A., and Tramontano, A., 2003, Evolution of bacterial and archaeal multicomponent monooxygenases. *J. Mol. Evol.*, 56:435–445.
80. O'Connor, K.E., and Hartmans, S., 1998, Indigo formation by aromatic hydrocarbon-degrading bacteria. *Biotechnol. Lett.*, 20:219–223.
81. O'Connor, K.E., Dobson, A.D., and Hartmans, S., 1997, Indigo formation by microorganisms expressing styrene monooxygenase activity. *Appl. Environ. Microbiol.*, 63:4287–4291.
82. Olah, G.A., Keumi, T., Lecoq, J.C., Fung, A.P., and Olah, J.A., 1991, Oxyfunctionalization of hydrocarbons. Acid-dependent high regioselectivity hydroxylation of naphthalene with hydrogen peroxide giving 1- or 2-naphthol. *J. Org. Chem.* 56:6148–6151.
83. Olsen, R.H., Kukor, J.J., and Kaphammer. B., 1994, A novel toluene-3-monooxygenase pathway cloned from *Pseudomonas pickettii* PKO1. *J. Bacteriol*., 176:3749–3756.
84. Othmer, K. 1991, Kirk-Othmer Encyclopedia of Chemical Technology, 4th ed, Vol. 13. Wiley-Interscience Publishers, New York.
85. Palumbo, A., Napolitano, A., and d'Ischia. M., 2002, Nitrocatechols versus nitrocatecholamines as novel competitive inhibitors of neuronal nitric oxide synthase: lack of the aminoethyl side chain determines loss of tetrahydrobiopterin-antagonizing properties. *Bioorg. Med. Chem. Lett*.,12:13–16.
86. Park, J., Kukor, J.J., and Abriola. L.M., 2002, Characterization of the adaptive response to trichloroethylene-mediated stresses in *Ralstonia pickettii* PKO1. *Appl. Environ. Microbiol.*, 68:5231–5240.
87. Peitsch, M.C., 1995, Protein modeling by e-mail. *Bio/Technol.* 13:658–660.
88. Peng, J., Shi, F., Gu, Y., and Deng. Y., 2003, Highly selective and green aqueous-ionic liquid biphasic hydroxylation of benzene to phenol with hydrogen peroxide. *Green Chem.*, 5:224–226.
89. Perry, R.H., and Chilton. C.H., 1973, Chemical Engineers' Handbook, 5th ed. McGraw-Hill Inc., New York, USA.
90. Pikus, J.D., Mitchell, K.H., Studts, J.M., McClay, K., Steffan, R.J., and Fox. B.G., 2000, Threonine 201 in the diiron enzyme toluene-4-monooxygenase is not required for catalysis. *Biochem*., 39:791–799.
91. Pikus, J.D., Studts, J.M., McClay, K., Steffan, R.J., and Fox. B.G., 1997, Changes in the regiospecificity of aromatic hydroxylation produced by active site engineering in the diiron enzyme toluene 4-monooxygenase. *Biochem.*, 36:9283–9289.
92. Resnick, S.M., Lee, K., and Gibson, D.T., 1996, Diverse reactions catalyzed by naphthalene dioxygenase from *Pseudomonas* sp. strain NCIB 9816. *J. Ind. Microbiol.*, 17:438–457.
93. Robinson, G.K., Stephens, G.M., Dalton, H., and Geary, P.J., 1992, The production of catechols from benzene and toluene by *Pseudomonas putida* in glucose fed-batch culture. *Biocatalysis*, 6:81–100.
94. Rosenzweig, A.C., Brandstetter, H., Whittington, D.A., Nordlund, P., Lippard, S.J., and Frederick. C.A., 1997, Crystal structures of the methane monooxygenase hydroxylase from *Methylococcus capsulatus* (bath): implications for substrate gating and component interactions. *Proteins Struct. Funct. Genet*., 29:141–152.
95. Rosenzweig, A.C., Nordlund, P., Takahara, P.M., Frederick, C.A., and Lippard. S.J., 1995, Geometry of the soluble methane monooxygenase catalytic diiron center in two oxidation states. *Chem. Biol.*, 2:409–418.
96. Ross, S., 1997, A First Course in Probability, p. 252. Prentice Hall. Englewood Cliffs, NJ.
97. Rui, L., Cao, L., Chen, W., Reardon, K.F., and Wood, T.K., 2004, Active site engineering of the epoxide hydrolase from *Agrobacterium radiobacter* AD1 to enhance aerobic mineralization of *cis*-1,2-dichloroethylene in cells expressing an evolved toluene *ortho*-monooxygenase. *J. Biol. Chem.*, 279:46810–46817.

98. Rui, L., Kwon, Y.-M., Fishman, A., Reardon, K.F., and Wood, T.K., 2004, Saturation mutagenesis of toluene *ortho*-monooxygenase of *Burkholderia cepacia* G4 for enhanced 1-naphthol synthesis and chloroform degradation. *Appl. Environ. Microbiol*., 70:3246–3252.
99. Rui, L., Kwon, Y.-M., Reardon, K.F., and Wood. T.K., 2004, Metabolic pathway engineering to enhance aerobic degradation of chlorinated ethenes and to reduce their toxicity by cloning a novel glutathione *S*-transferase an evolved toluene *ortho*-monooxygenase, and gamma-glutamylcysteine synthetase. *Environ. Microbiol*., 6:491–500.
100. Rui, L., Reardon, K.F., and Wood. T.K., 2005, Altering toluene *ortho*-monooxygenase of *Burkholderia cepacia* G4 for regiospecific hydroxylation of indole to form indigoid compounds. *Appl. Microbiol. Biotechnol.,* 66:422–429.
101. Ryoo, D., Shim, H., Canada, K., Barbieri, P., and Wood. T.K., 2000, Aerobic degradation of tetrachloroethylene by toluene-*o*-xylene monooxygenase of *pseudomonas stutzeri* OX1. *Nat. Biotechnol*., 18:775–778.
102. Sakamoto, T., Joern, J.M., Arisawa, A., and Arnold. F.H., 2001, Laboratory evolution of toluene dioxygenase to accept 4-picoline as a substrate. *Appl. Environ. Microbiol.*, 67:3882–3887.
103. Sazinsky, M.H., Bard, J., Di Donato, A., and Lippard, S.J., 2004, Crystal structure of the toluene/*o*-xylene monooxygease hydroxylase from *Pseudomonas stutzeri* OX1. *J. Biol. Chem.*, 279:30600–30610.
104. Scharrenburg, G.J.M., and Frankena, J., 1996, Biokatalyse helpt farmaceutische industrie bij asymmetrische synthese. *Chemisck Weekblad*, 284–286.
105. Schmid, A., Dordick, J.S., Hauer, B., Kiener, A., Wubbolts, M., and Witholt, B., 2001, Industrial biocatalysis today and tomorrow. *Nature*, 409:258–268.
106. Schoemaker, H. E., Mink, D., and Wubbolts. M.G., 2003, Dispelling the myths - biocatalysis in industrial synthesis. *Science* 299:1694–1697.
107. Schwede, T., Kopp, J., Guex, N., and Peitsch. M.C., 2003, SWISS-MODEL: An Automated Protein Homology-Modeling Server. Nucleic Acids Research 31:3381-3385.
108. Sharma, P.K., and McCarty, P.L., 1996, Isolation and characterization of a facultatively aerobic bacterium that reductively dehalogenates tetrachloroethene to *cis*-1,2-dichloroethene. *Appl. Environ. Microbiol*., 62:761–765.
109. Sheilds, M.S., Reagin, M.J., Gerger, R.R., Campbell, R., and Somerville, C., 1995, TOM, a new aromatic degradative plasmid from *Burkholderia cepacia* G4. *Appl. Environ. Microbiol.*, 61:1352–1356.
110. Shields, M.S., Montgomery, S.O., Chapman, P.J., Cuskey, S.M., and Pritchard, P.H., 1989, Novel pathway of toluene catabolism in the trichloroethylene-degrading bacterium G4. *Appl. Environ. Microbiol.*, 55:1624–1629.
111. Shim, H., and Wood, T.K., 2000, Aerobic degradation of mixtures of chlorinated aliphatics by cloned toluene-*o*-xylene monooxygenase and toluene *o*-monooxygenase in resting cells. *Biotechnol. Bioeng*., 70:693–698.
112. Spain, J.C., and Gibson, D.T., 1988, Oxidation of substituted phenols by *Pseudomonas putida* F1 and *Pseudomonas* sp. strain JS6. *Appl. Environ. Microbiol.*, 54:1399–1404.
113. Steffan, R.J., and McClay. K.R., 2000, Preparation of enantio-specific epoxides. International Patent Application patent WO 00/73425 A1.
114. Straathof, A.J.J., Panke, S., and Schmid. A., 2002, The production of fine chemicals by biotransformations. *Curr. Opin. Biotechnol*., 13:548–556.
115. Studts, J.M., Mitchell, K.H., Pikus, J.D., McClay, K., Steffan, R.J., and Fox, B.G., 2000, Optimized expression and purification of toluene 4-monooxygenase hydroxylase. *Protein Expr. Purif.*, 20:58–65.
116. Tao, Y., Bentley, W.E., and Wood. T.K., 2005, Phenol and 2-naphthol production by toluene 4-monooxygenases using two liquid phases. *Appl. Microbiol. Biotechnol*., 68:614–621.

117. Tao, Y., Bentley, W.E., and Wood, T.K., 2005, Regiospecific oxidation of naphthalene and flourene by toluene monooxygenases and engineered toluene 4-monooxygenase of *Pseudomonas mendocina* KR1. *Biotechnol. Bioeng*., 90:85–94.
118. Tao, Y., Fishman, A., Bentley, W.E., and Wood, T.K., 2004, Altering toluene 4-monooxygenase by active-site engineering for the synthesis of 3-methoxycatechol, methoxyhydroquinone, and methylhydroquinone. *J. Bacteriol*., 186:4705–4713.
119. Tao, Y., Fishman, A., Bentley, W.E., and Wood, T.K., 2004, Oxidation of benzene to phenol, catechol, and 1,2,3-trihydroxybenzene by toluene 4-monooxygenase of *Pseudomonas mendocina* KR1 and toluene 3-monooxygenase of *Ralstonia pickettii* PKO1. *Appl. Environ. Microbiol*., 70:3814–3820.
120. Tice, R., 1998, Review of toxicological literature, Pyrogallol 87-66-1. National Toxicology Program.
121. Umezawa, K., Kawakami, M., and Watanabe, T., 2003, Molecular design and biological activities of protein-tyrosine phosphatase inhibitors. *Pharm. Ther.*, 99:15–24.
122. van Beilen, J.B., Duetz, W.A., Schmid A., and Witholt, B., 2003, Practical issues in the application of oxygenases. *Trends Biotechnol.*, 21:170–177.
123. van Hylckama Vlieg, J.E.T., and Janssen, D.B., 2001, Formation and detoxification of reactive intermediates in the metabolism of chlorinated ethenes. *J. Biotechnol*., 85:81–102.
124. Vardar, G., and Wood, T.K., 2005, Alpha subunit positions methionine 180 and glutamate 214 of *Pseudomonas stutzeri* OX1 toluene-*o*-xylene monooxygenase influence catalysis. *J. Bacteriol*., 187:1511–1514.
125. Vardar, G., Barbieri, P., and Wood, T.K., 2005, Chemotaxis of *Pseudomonas stutzeri* OX1 and *Burkholderia cepacia* G4 toward chlorinated ethenes. *Appl. Microbiol. Biotechnol*., 66:696–701.
126. Vardar, G., and Wood, T.K., 2005, Protein engineering of toluene-*o*-xylene monooxygenase from *Pseudomonas stutzeri* OX1 for enhanced chlorinated ethene degradation and *o*-xylene oxidation. *Appl. Microbiol. Biotechnol*., 68:510–517.
127. Vardar, G., Ryu, K., and Wood, T.K., 2005, Protein engineering of toluene-*o*-xylene monooxygenase from *Pseudomonas stutzeri* OX1 for oxidizing nitrobenzene to 3-nitrocatechol, 4-nitrocatehcol, and nitrohydroquinone. *J. Biotechnol*., 115:145–156.
128. Vardar, G., and Wood, T.K., 2004, Protein engineering of toluene-*o*-xylene monooxygenase from *Pseudomonas stutzeri* OX1 for synthesizing 4-methylresorcinol, methylhydroquinone, and pyrogallol. *Appl. Environ. Microbiol.*, 70:3253–3262.
129. Vlieg, J.E.T.v.H., and Janssen, D.B., 2001, Formation and detoxification of reactive intermediates in the metabolism of chlorinated ethenes. *J. Biotechnol*., 85:81–102.
130. Westrick, J.J., Mello, J.W., and Thomas, R.F., 1984, The groundwater supply survey. *J. Am. Water Works Assoc.*, 76:52–59.
131. Whited, G.M., and Gibson, D.T., 1991, Separation and partial characterization of the enzymes of the toluene-4-monooxygenase catabolic pathway in *Pseudomonas mendocina* KR1. *J. Bacteriol.*, 173:3017–3020.
132. Wick, C.B., 1995, Genencor international takes a green route to blue dye. *Genet. Eng. News*, 15:22.
133. Xu, D., Ballou, D.P., and Massey, V., 2001, Studies of the mechanism of phenol hydroxylase: mutants Tyr289Phe, Asp54Asn, and Arg281Met. *Biochem.*, 40:12369–12378.
134. Yen, K.-M., Karl, M.R., Blatt, L.M., Simon, M.J., Winter, R.B., Fausset, P.R., Lu, H.S., Harcourt, A.A., and Chen, K.K., 1991, Cloning and characterization of a *Pseudomonas mendocina* KR1 gene cluster encoding toluene-4-monooxygenase. *J. Bacteriol*., 173:5315–5327.
135. Yonezawa, T., March 2003, Capacitor electrolytes containing dihydroxytoluenes for low specific resistance. Japan patent 2,003,068,585.

9

AROMATIC RING HYDROXYLATING DIOXYGENASES

Rebecca E. Parales[1] and Sol M. Resnick[2]

[1]Section of Microbiology
University of California
Davis, CA 95616, USA
and
[2]Biotechnology R&D
The Dow Chemical Company
San Diego, CA 92121, USA

Abbreviations: NDO, Naphthalene dioxygenase; TDO, toluene dioxygenase; NBDO, nitrobenzene dioxygenase; BPDO, biphenyl dioxygenase; PCB, polychlorinated biphenyl; TCE, trichloroethylene; PAH, polycyclic aromatic hydrocarbon

1. INTRODUCTION

The initial oxidation of the aromatic ring is the most difficult catalytic step in the aerobic degradation of aromatic compounds. Bacterial aromatic ring hydroxylating dioxygenases (also known as Rieske non-heme iron dioxygenases) catalyze the addition of hydroxyl groups to the highly stable aromatic ring, setting the stage for further oxidation, and eventual ring cleavage. Aromatic ring hydroxylating dioxygenases are known to catalyze the initial reaction in the bacterial biodegradation of a diverse array of aromatic and polycyclic aromatic hydrocarbons (PAHs), chlorinated aromatic, nitroaromatic, aminoaromatic, and heterocyclic aromatic compounds, and aromatic acids. Aromatic ring hydroxylating dioxygenases use molecular oxygen as a substrate, adding both atoms

Pseudomonas, Volume 4, edited by Juan-Luis Ramos and Roger C. Levesque

of O_2 to the aromatic ring of the substrate. To date, over 100 aromatic ring hydroxylating dioxygenases have been identified based on biological activity or nucleotide sequence identity. These enzymes are distributed among a variety of Gram-negative and Gram-positive bacteria and are important for the catabolism of a wide range of environmental pollutants.

Aromatic ring hydroxylating dioxygenases are multicomponent enzyme systems (E.C. 1.14.12.-) that catalyze reductive dihydroxylation of their substrates, and are distinct from aromatic ring cleavage (or ring fission) dioxygenases (E.C. 1.13.11.-), which act on the downstream catechol intermediates in many of the same catabolic pathways. Many of these enzymes are typically quite promiscuous, catalyzing the oxidation of a wide range of compounds in addition to their native substrates. At the same time, however, many of these enzyme systems are highly enantioselective, producing chiral *cis*-dihydrodiols or other chiral products in high enantiomeric purity. These properties have made aromatic ring hydroxylating dioxygenases attractive as biocatalysts.

This chapter will focus primarily on the aromatic ring hydroxylating dioxygenases present in pseudomonads and related proteobacteria, but it is important to keep in mind that closely related enzyme systems have been identified in a variety of other bacterial genera, including *Rhodococcus*, *Nocardia*, *Mycobacterium*, etc. Aspects described in this chapter include dioxygenase classification, enzymology, structure and mechanism, and applications in biotechnology. Much of our understanding of the structure and function of aromatic hydrocarbon dioxygenases comes from studies of naphthalene dioxygenase (NDO), and characteristics of this enzyme system will be described in detail and used as a basis for comparisons with other dioxygenases.

2. DISTRIBUTION OF RING HYDROXYLATING DIOXYGENASES IN CATABOLIC PATHWAYS

Bacterial pathways for the degradation of numerous aromatic hydrocarbons, PAHs, chlorinated aromatic compounds, nitroaromatic compounds, aminoaromatic compounds, aromatic acids, and heterocyclic aromatic compounds are initiated by aromatic ring hydroxylating dioxygenases (Figure 1). Many of these compounds are toxic environmental pollutants. Some, including various chlorinated and nitroarene compounds, are man-made, while many others are naturally occurring biological products or components of petroleum. Regardless of the source, bacterial pathways for the degradation of this group of chemicals are important both in the recycling of carbon on earth and in the removal of toxic pollutants at contaminated sites.

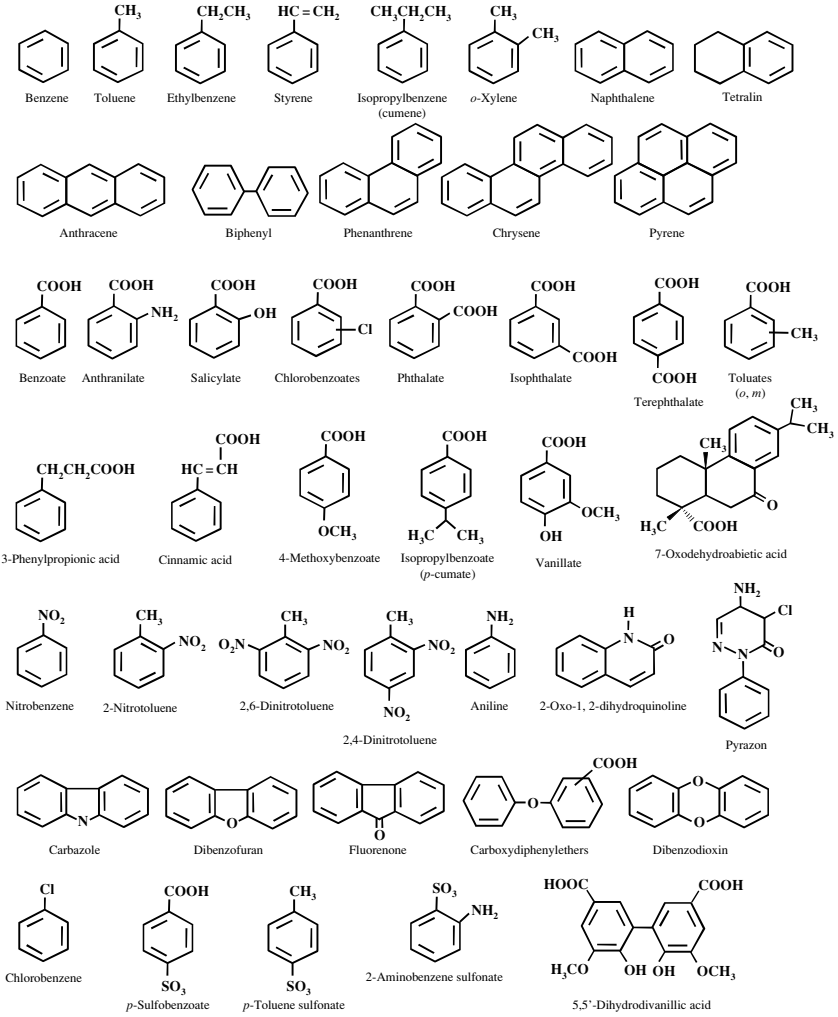
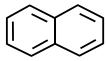
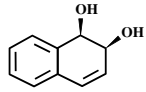
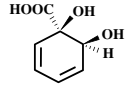
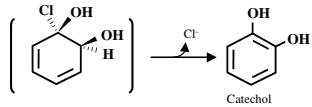
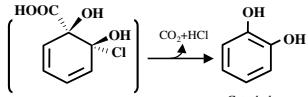
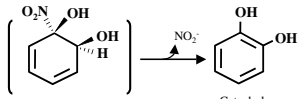
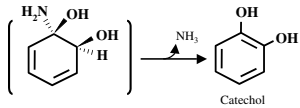
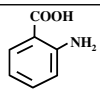
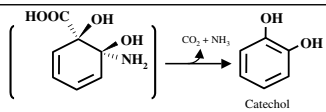
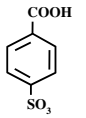
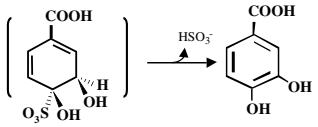
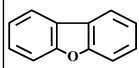
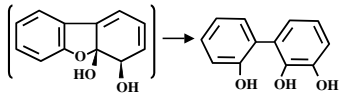


Figure 1. Aromatic compounds degraded by bacterial pathways that are initiated by aromatic ring hydroxylating dioxygenases.

2.1. Types of Reactions Catalyzed

The initial reaction catalyzed by aromatic ring hydroxylating dioxygenases on aromatic hydrocarbons and certain other substrates is a *cis*-dihydroxylation of the carbon–carbon double bond of adjacent unsubstituted carbon atoms. This reaction typically generates a chiral *cis*-dihydrodiol as seen with the reaction on naphthalene (Reaction A, Table 1). Oxidation of aromatic

Table 1. Representative types of *cis*-dihydroxylation reactions catalyzed by aromatic ring hydroxylating dioxygenases.

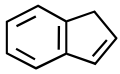
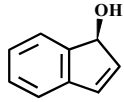
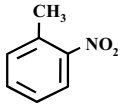
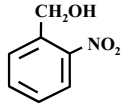
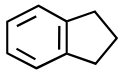
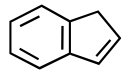
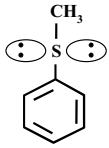
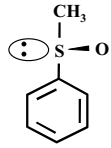
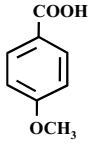
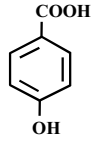
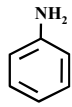
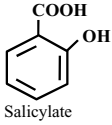
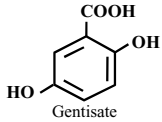
Reaction Type	Substrate	Enzyme	Product
A. *cis*-dihydroxylation	 Naphthalene	Naphthalene dioxygenase, O_2	 Naphthalene *cis* 1,2-dihydrodiol
B. *cis*-dihydroxylation	 Benzoate	Benzoate dioxygenase, O_2	 Benzoate *cis* 1,2-dihydrodiol
C. *cis*-dihydroxylation and dehalogenation	 Chlorobenzene	Chlorobenzene dioxygenase, O_2	 Catechol
D. *cis*-dihydroxylation, dehalogenation and decarboxylation	 2-Chlorobenzoate	Chlorobenzoate dioxygenase, O_2	 Catechol
E. *cis*-dihydroxylation and nitrite elimination	 Nitrobenzene	Nitrobenzene dioxygenase, O_2	 Catechol
F. *cis*-dihydroxylation and deamination	 Aniline	Aniline dioxygenase, O_2	 Catechol
G. *cis*-dihydroxylation, deamination and decarboxylation	 Anthranilate	Anthranilate dioxygenase, O_2	 Catechol
H. *cis*-dihydroxylation, and desufonation	 *p*-Sulfobenzoate	*p*-Sulfobenzoate dioxygenase, O_2	 Protocatechuate
I. Angular dihydroxylation	 Dibenzofuran	Dibenzofuran dioxygenase, O_2	 2, 2', 3-trihydroxy biphenyl

acids such as benzoate occurs at a carboxylated carbon and an adjacent unsubstituted carbon, and results in the formation of chiral *cis*-dihydroxylated cyclohexadiene carboxylic acids (Reaction B, Table 1). Dioxygenase-catalyzed dechlorination has also been demonstrated for chlorinated benzoates, benzenes, and biphenyls (Reactions C and D, Table 1). Dioxygenation at a chlorine-substituted carbon results in concomitant elimination of chloride.[88,113,193] Similar reactions have been shown with nitroaromatic, aminoaromatic, and sulfoaromatic substrates (Reactions E–H, Table 1), resulting in nitrite, ammonia, or sulfite elimination.[4,47,94,177,199,259] In these cases, the displacement reactions result in the formation of dihydroxylated (catecholic) products that are activated for further metabolism. A specialized group of ring hydroxylating dioxygenases catalyzes angular dioxygenation, as seen with dibenzofuran[48,81] (Reaction I, Table 1). Other substrates that are oxidized by angular dioxygenation include carbazole, diphenylethers, and dibenzo-*p*-dioxin.[80,239,297] Two recent reviews provide excellent summaries of angular dioxygenases and the pathways in which they operate.[200,202] Finally, certain members of this large family of multicomponent enzymes function as monooxygenases, as exemplified by methoxybenzoate monooxygenase[26] and salicylate 5-hydroxylase (Reactions N and P, Table 2).[305] Very recently, multicomponent salicylate 1-hydroxylases have been identified in three different sphingomonads.[55,68,220]

2.2. Classification of Ring Hydroxylating Dioxygenases

Historically, Rieske non-heme iron oxygenases were classified using the Batie system, which was based on the electron transfer components present in the 10 Rieske non-heme iron oxygenase systems known at that time.[20] Two-component (reductase and oxygenase; Class I) and three-component (reductase, ferredoxin, and oxygenase; Class II, III) enzyme systems were represented, and the classes were further subdivided based on the type of flavin cofactor (FAD or FMN) in the reductase, the presence or absence of an iron–sulfur center in the reductase, the number of proteins in the oxygenase, and if a ferredoxin was involved, the type of iron–sulfur center (plant or Rieske) in the ferredoxin. This classification system worked well with a small number of known enzymes, but as more enzyme systems with diverse properties were identified, not all of the new enzymes fit into this classification system. A second classification system based on amino acid sequence alignments of the available oxygenase α subunits was proposed by Werlen *et al.*, and it identified four dioxygenase families (naphthalene, toluene/benzene, biphenyl, and benzoate/toluate).[293] This classification system was based on the catalytic activity of the enzymes because the α subunit of the oxygenase plays a major role in determining substrate specificity.[17,24,186,209,212,275,307] The results of the Werlen *et al.* study demonstrated that, in general, related dioxygenases coded for enzymes with similar substrates. Nakatsu *et al.*[192] subsequently built upon this classification system,

Table 2. Other types of oxidation reactions catalyzed by aromatic ring hydroxylating dioxygenases.

Reaction Type	Substrate	Enzyme	Product
J. Benzylic hydroxylation	 Indene	Naphthalene dioxygenase, O_2	 Indenol
K. Methyl group hydroxylation	 2-Nitrotoluene	Naphthalene dioxygenase, O_2	 2-Nitrobenzyl alcohol
L. Oxygen-dependent desaturation	 Indan	Naphthalene dioxygenase, O_2	 Indene
M. Sulfoxidation	 Methyl phenyl sulfide	Naphthalene dioxygenase, O_2	 Methyl phenyl sulfoxide
N. *O*-Dealkylation	 4-Methoxybenzoate	Naphthalene dioxygenase, O_2	 4-Hydroxybenzoate
O. *N*-Dealkylation	 N-Methylaniline	Naphthalene dioxygenase, O_2	 Aniline
P. Net aromatic ring hydroxylation	 Salicylate	Salicylate 5- hydroxylase, O_2	 Gentisate

adding the α_n dioxygenases (those in the Batie system Class IA) to the Werlen system and demonstrating that these α_n oxygenases formed a separate lineage.

Two modified classification systems, also based on the phylogeny of the oxygenase α subunits have been proposed, each of which divided the known enzymes into four families or groups (Groups I–IV, or phthalate, benzoate, naphthalene, and toluene/biphenyl families respectively[106,196]). From current phylogenetic analyses, additional families or groups are apparent, as well as several enzymes that do not cluster tightly with any of the distinct families (Figure 2). In some of the families, enzymes with similar native substrates cluster together. The benzoate family (Figure 2) includes enzymes with activities toward various aromatic acids and aminoaromatic compounds. Monocyclic aromatic hydrocarbon dioxygenases, biphenyl dioxygenases (BPDOs), and chlorinated aromatic hydrocarbon dioxygenases from both Gram-negative and Gram-positive bacteria are tightly clustered in the toluene/biphenyl family. Enzymes for naphthalene and PAH degradation from Gram-negative and Gram-positive bacteria form separate clusters (naphthalene and Gram-positive PAH families, respectively), and nitroarene dioxygenases cluster tightly with NDO from Gram-negative bacteria. The phthalate family (Figure 2), which consists of all of the Rieske non-heme iron oxygenases that contain only α subunits, is a very diverse group of enzymes that catalyze the oxidation of a variety of structurally unrelated aromatic compounds. Substrates for the phthalate family include aromatic acids such as phthalate, *p*-toluene sulfonate, and phenoxybenzoate, as well as carbazole and 2-oxo-1,2-dihydroquinoline. Two phthalate dioxygenases from Gram-positive bacteria were recently identified.[76,260] These enzymes do not cluster with the phthalate family but with the Gram-positive PAH dioxygenases (Figure 2), and in addition, are comprised of both α and β subunits. A new family has recently emerged, and is designated here as the salicylate family (Figure 2). This family consists of enzymes that catalyze oxidation at either the 1- or 5-position of salicylate,[55,68,220,305] as well as other substrates. There are now a number of enzymes that are quite distantly related to those found in the core of these groups or families (Figure 2), and thus, none of the classification systems can completely describe the diversity of enzymes that have been identified to date.

2.3. Diversity of NDO and Related Nitroarene Dioxygenases

Based on sequence alignments of NDO α subunit genes, the genes from Gram-negative bacteria were found to fall into three distinct groups that have been designated *nah*-like, *dnt/ntd*-like (sometimes called *nag*-like), and *phn*-like.[165,176] An additional distantly related naphthalene/phenanthrene dioxygenase (A5 PhnA1, Figure 2) was identified from a marine *Cycloclasticus* isolate (strain A5) that grows on several PAHs.[144] Tetralin dioxygenase (TFA ThnA1,

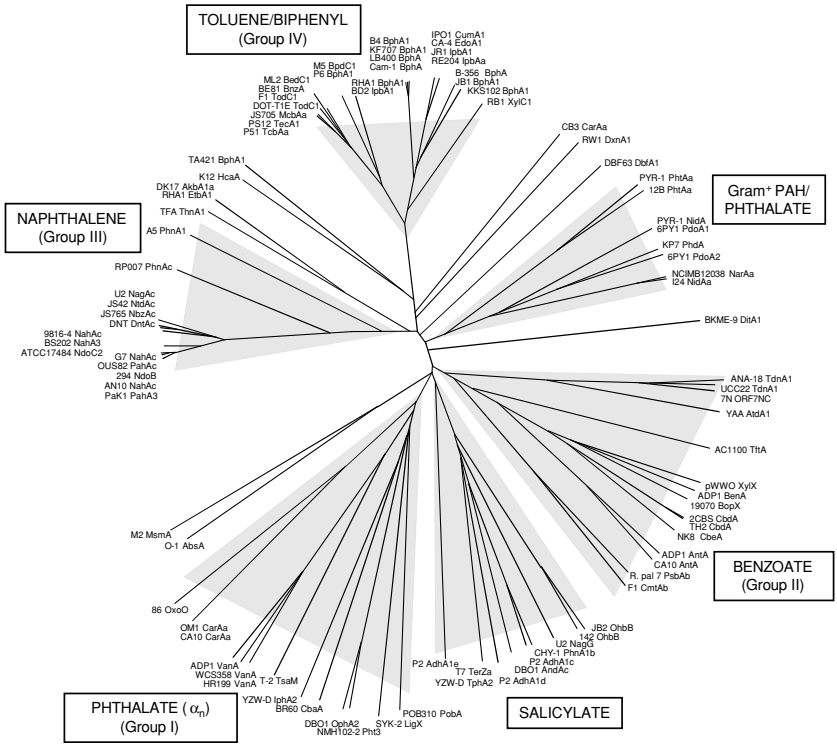


Figure 2. Phylogenetic tree of α subunits of aromatic ring hydroxylating dioxygenases. Sequences were aligned and the tree was generated using the Clustal W package. [277]. α subunit sequences are from the following organisms/dioxygenase systems. PaK1 PahA3, NDO: *P. aeruginosa* PaK1 (D84146); AN10 NahAc, NDO: *P. stutzeri* AN10 (AF039533); 294 NdoB, NDO: *P. fluorescens* 294 (AF004283); OUS82 PahAc, NDO: *P. putida* OUS82 (D16629); G7 NahAc, NDO: *P. putida* G7 (M86949); ATCC17484 NdoC2, NDO: *P. putida* ATCC17484 (AF004284); BS202 NahA3, NDO: *P. putida* BS202 (AF010471); 9816-4 NahAc, NDO: *Pseudomonas* sp. NCIB 9816-4 (U49496); U2 NagAc, NDO: *Ralstonia* sp. U2 (AF036940); JS42 NtdAc, 2-nitrotoluene dioxygenase: *Acidovorax* sp. JS42 (U49504); JS765 NbzAc, NBDO: *Comamonas* sp. JS765 (AF379638); DNT DntAc, 2,4-dinitrotoluene dioxygenase: *Burkholderia* sp. DNT (U62430); RP007 PhnAc, PAH dioxygenase: *Burkholderia* sp. RP007 (AF061751); A5 PhnA1, PAH dioxygenase: *Cycloclasticus* sp. A5 (AB102786); TFA ThnA1, tetralin dioxygenase: *Sphingopyxis macrogoltabida* TFA (AF157565); RHA1 EtbA1, ethylbenzene dioxygenase: *Rhocococcus* sp. RHA1 (AB120955); DK17 AkbA1a, alkylbenzene dioxygenase: *Rhodococcus* sp. DK17 (AY502075); K-12 HcaA, 3-phenylpropionate dioxygenase: *E. coli* K-12 (NC000913); TA421 BphA1, BPDO: *R. erythopolis* TA421 (D88020); P51 TcbAa, chlorobenzene dioxygenase: *Pseudomonas* sp. P51 (U15298); PS12 TecA1, chlorobenzene dioxygenase: *Burkholderia* sp. PS12 (U78099); JS705 McbAa, chlorobenzene dioxygenase: *Ralstonia* sp. JS705 (AJ006307); F1 TodC1, TDO: *P. putida* F1 (J04996); DOT-T1E TodC1, TDO: *P. putida* DOT-T1E (Y18245); BE81 BnzA, benzene dioxygenase *P. putida* BE81 (M17904); ML2 BedC1, benzene dioxygenase: *P. putida* ML2 (AF148496); M5 BpdC1, BPDO: *Rhodococcus* sp. M5 (U27591); P6 BphA1, BPDO: *Rhodococcus globerulus* P6 (X80041); RHA1 BphA1, BPDO: *Rhodococcus* sp. RHA1 (D32142); BD2 IpbA1, isopropylbenzene dioxygenase: *R. erythropolis* BD2 (U24277);

Figure 2 (*continued*). B4 BphA1, BPDO: *Pseudomonas* sp. B4 (AJ251217); 28, KF707 BphA1, BPDO: *P. pseudoalcaligenes* KF707 (AF049345); LB400 BphA, BPDO: *B. xenovorans* LB400 (M86348); Cam-1 BphA, BPDO: *Pseudomonas* sp. Cam-1 (AY027651); IPO1 CumA1, isopropylbenzene dioxygenase: *P. fluorescens* IPO1 (D37828); CA-4 EdoA1, ethylbenzene dioxygenase: *P. fluorescens* CA-4 (AF049851); JR1 IpbA1, isopropylbenzene dioxygenase: *Pseudomonas* sp. JR1 (U53507); RE204 IpbAa, isopropylbenzene dioxygenase: *P. putida* RE204 (AF006691); B-356 BphA, BPDO: *C. testosteroni* B-356 (U47637); JB1 BphA1, BPDO: *Burkholderia* sp. JB1 (AJ010057); KKS102 BphA1, BPDO: *Pseudomonas* sp. KKS102 (D17319); RB1 XylC1, aromatic hydrocarbon dioxygenase: *Cycloclasticus oligotrophicus* RB1 (U51165); CB3 CarAa, carbazole dioxygenase: *Sphingomonas* sp. CB3 (AF060489); RW1 DxnA1, dioxin dioxygenase: *Sphingomonas* sp. RW1 (X72850); DBF63 DbfA1, dibenzofuran dioxygenase: *Terrabacter* sp. DBF63 (AB054975); PYR-1 PhtAa, phthalate dioxygenase: *Mycobacterium vanbaalenii* PYR-1 (AY365117); 12B PhtAa, phthalate dioxygenase: *Arthrobacter keyseri* 12B (AF331043); PYR-1 NidA, naphthalene-inducible dioxygenase: *M. vanbaalenii* PYR-1 (AF249301); 6PY1 PdoA1, PAH dioxygenase: *Mycobacterium* sp. 6PY1 (AJ494745); KP7 PhdA, phenanthrene dioxygenase: *Nocardioides* sp. KP7 (AB017794); 6PY1 PdoA2, PAH dioxygenase: *Mycobacterium* sp. 6PY1 (AJ494743); NCIMB12038 NarAa, NDO: *Rhodococcus* sp. NCIMB12038 (AF082663); I24 NidAa, naphthalene-inducible dioxygenase: *Rhodococcus* sp. I24 (AF121905); BKME-9 DitA1, diterpenoid dioxygenase: *P. abietaniphila* BKME-9 (AF119621); ANA-18 TdnA1, aniline dioxygenase: *Fraturia* sp. ANA-18 (AB089795); UCC22 TdnA1, aniline dioxygenase: *P. putida* UCC22 (D85415); 7N ORF7NC, aniline dioxygenase: *Delftia acidovorans* 7N (AB177545); YAA AtdA1, aniline dioxygenase: *Acinetobacter* sp. YAA (D86080); AC1100 TftA, 2,4,5-trichlorophenoxyacetic acid oxygenase: *B. cepacia* AC1100 (U11420); pWWO XylX, toluate dioxygenase: *P. putida* pWWO (AJ344068); ADP1 BenA, benzoate dioxygenase: *Acinetobacter* sp. ADP1 (AF009224); 19070 BopX, benzoate dioxygenase: *Rhodococcus* sp. 19070 (AF279141); 2CBS CbdC, 2-halobenzoate dioxyegnase: *B. cepacia* 2CBS (X79076); TH2 CbdA, 2-halobenzoate dioxyegnase: *Burkholderia* sp. TH2 (AB035324); NK8 CbeA, chlorobenzoate dioxygenase: *Burkholderia* sp. NK8 (AB024746); ADP1 AntA, anthranilate dioxygenase: *Acinetobacter* sp. ADP1 (AF071556); AntA, CA10, anthranilate dioxygenase: *P. resinovorans* CA10 (NC_004444); *R. pal* 7 PsbAb, cumate dioxygenase: *R. palustris* 7 (AB022919); F1 CmtAb, cumate dioxygenase: *P. putida* F1 (AB042508); JB2 OhbB, *o*-halobenzoate dioxygenase: *P. aeruginosa* JB2 (AF422937); 142 OhbB, *o*-halobenzoate dioxygenase: *P. aeruginosa* 142 (AF121970); U2 NagG, salicylate 5-hydroxylase: *Ralstonia* sp. U2 (AF036940); CHY-1 PhnA1b, salicylate 1-hydroxylase: *Sphingomonas* sp. CHY-1 (AJ633552); P2 AdhA1c, salicylate 1-hydroxylase: *Sphingobium* sp. P2 (AB091693); DBO1 AndAc, anthranilate dioxygenase: *B. cepacia* DBO1 (AY223539); P2 AdhA1d, salicylate 1-hydroxylase: *Sphingobium* sp. P2 (AB091692); T7 TerZa, terephthalate dioxygenase: *Delftia* sp. T7 (AB081091); YZW-D TphA2, terephthalate dioxygenase: *C. testosteroni* YZW-D (AY923836); AdhA1e, salicylate 1-hydroxylase: *Sphingobium* sp. P2 (AB091692); POB310 PobA, phenoxybenzoate dioxygenase: *P. pseudoalcaligenes* POB310 (X78823); SYK-2 LigX, 5,5′-dehydrodivanillic acid *O*-demethylase: *S. paucimobilis* SYK-2 (AB021319); NMH102-2 Pht3, phthalate dioxygenase: *P. putida* NMH102-2 (D13229); DBO1 OphA2, phthalate dioxygenase: *B. cepacia* DBO1 (AF095748); BR60 CbaA, 3-chlorobenzoate dioxygenase: *C. testosteroni* BR60 (U18133.2); YZW-D IphA2, iosphthalate dioxygenase: *C. testosteroni* YZW-D (AY923836); T-2 TsaM, *p*-toluenesulfonate monooxygenase: C. *testosteroni* T-2 (AF311437); HR199 VanA, vanillate *O*-demethylase: *Pseudomonas* sp. HR199 (Y11521); WCS358 VanA, vanillate *O*-demethylase: *P. putida* WCS358 (Y14759); ADP1 VanA, vanillate *O*-demethylase: *Acinetobacter* sp. ADP1 (AF009672); OM1 CarAa, carbazole dioxygenase: *P. stutzeri* OM1 (AB088757); CA10 CarAa, carbazole dioxygenase: *P. resinovorans* CA10 (AB088420); 86 OxoO, 1-oxo-1,2-dihydroquinoline 8-monooxygenase: *P. putida* 86 (Y12655); O-1 AbsA, 2-aminobenzene sulfonate dioxygenase: *Alcaligenes* sp. O-1 (AF109074); M2 MsmA, methanesulfonic acid monooxygenase: *Methylsulfomonas methylovora* M2 (AF091716).

Figure 2) from *Sphingopyxis macrogoltabida* TFA is also distantly related to members of this family.[187]

The *nah*-like group contains enzymes (Figure 2) from various *Pseudomonas* species (*P. putida*, *P. fluorescens*, *P. aeruginosa*, *P. stutzeri*[34,69,274]) and includes the well-studied NDO from *Pseudomonas* sp. NCIB 9816-4 and *P. putida* G7.[256] Sequence and gene order in this group appears to be conserved. Based on further sequence analysis, the *nah*-like group has been divided into two subgroups, and a few strains were shown to contain copies of both types of NDO genes.[87]

The *dnt/ntd* group (Figure 2) is represented by dioxygenases for naphthalene and nitroarene compounds (nitrobenzene and nitrotoluenes) and these enzymes are from members of the β-proteobacteria, such as *Ralstonia*, *Burkholderia*, *Comamonas*, and *Acidovorax*, rather than the γ-proteobacteria. Recent studies have suggested that NDO in the *dnt/ntd* group may be more common in nature than previously thought.[73,295] *Comamonas* sp. strain JS765, *Acidovorax* sp. strain JS42, and *Burkholderia* sp. strain DNT contain nitroarene dioxygenases that are capable of attack at the nitro-substituted carbons of nitrobenzene, 2-nitrotoluene, or 2,4-dinitrotoluene, respectively, resulting in the removal of the nitro group as nitrite and rearomatization of the aromatic ring to form a catechol.[173,208,265] These enzyme systems are very closely related to NDO, in particular that from *Ralstonia* sp. strain U2 (*nagA*; 94% DNA sequence identity[92]), and are believed to have evolved from an ancestral NDO.[92,210]

The *phn*-like group is represented by the naphthalene/phenanthrene dioxygenase from *Burkholderia* sp. RP007 (Figure 2).[165] Additional *phn*-like dioxygenases have recently been identified from several other *Burkholderia* species.[296] The gene order of the cluster from RP007 is very different from that of the *nah*-like or *dnt/ntd* group of NDO. Genes encoding electron transport components were not found within the 11-kb sequenced region.[165] In addition, the presence of genes encoding a LysR-type regulator that is very distantly related to NahR from *P. putida* G7 (21% similarity) and a regulator of the NtrC family suggests that the genes may be controlled differently from those of the other two groups of NDO genes, possibly in a manner reminiscent of the regulation of the TOL plasmid genes.[165,226]

A series of naphthalene and/or phenanthrene and pyrene dioxygenases (Gram-positive PAH family, Figure 2) have been identified in Gram-positive organisms, including *Rhodococcus*,[5,157,164,281] *Nocardioides*,[238] and *Mycobacterium*.[149,156] The sequences of the α subunit genes from Gram-positive bacteria cluster together, and it is of note that they are quite distantly related to those from Gram-negative bacteria (Figure 2). The order of the naphthalene pathway genes differs from those of the γ- and β-proteobacteria, and the ferredoxin and reductase genes have only been identified in *Nocardioides* sp. strain KP7. In this strain, the ferredoxin and reductase genes are located approximately 3-kb downstream of the dioxygenase α and β subunit genes,[238]

but the ferredoxin and reductase genes do not appear to be co-localized with the dioxygenase genes in *Rhodococcus* or *Mycobacterium*.[149,157,281] The ferredoxin from strain KP7 has features of a [3Fe–4S] or a [4Fe–4S] type ferredoxin rather than a [2Fe–2S] type ferredoxin.[238] Another example of a [3Fe–4S] or a [4Fe–4S] type ferredoxin participating in electron transfer to a Rieske non-heme iron oxygenase is in the diterpenoid dioxygenases from *Pseudomonas abietaniphila* BKME-9 (Figure 2).[182] These two enzyme systems represent examples of oxygenases that do not fit the Batie classification system.

2.4. Diversity of Toluene/Biphenyl Family Dioxygenases

In contrast to enzymes for naphthalene and PAH degradation, the toluene/biphenyl family of dioxygenases is a cohesive group with enzymes from both Gram-positive and Gram-negative organisms represented. This family is comprised of three-component enzymes from pathways for the degradation of benzene, alkylbenzenes, chlorobenzenes, and biphenyl. Dioxygenases that initiate pathways for the degradation of benzene in *P. putida* ML2 and BE81,[131,276] and benzene, toluene, and ethylbenzene degradation in *P. putida* F1[308,309] and *P. putida* DOT-T1E,[189] are members of this family (Figure 2). However, *P. putida* ML2 uses an *ortho*-cleavage pathway,[13] *P. putida* BE-81 can use either an *ortho* or a *meta* pathway,[254] and *P. putida* F1 and DOT-T1E use a *meta* pathway for completing the conversion of these substrates to TCA cycle intermediates. Chlorobenzene dioxygenases[23,190,293] also fall into this family, and those that have been sequenced thus far cluster with benzene and toluene dioxygenases. These enzymes catalyze *cis*-dihydroxylation and dechlorination of their substrates (Reaction C, Table 1). The chlorobenzene dioxygenases from *Pseudomonas* sp. P51[293] and *Ralstonia* (formerly *Pseudomonas*) sp. PS12[23] allow these strains to grow with multiply chlorinated benzenes. Both enzymes have broad substrate specificities.[221,227] The ethylbenzene-degrading isolate *P. fluorescens* CA-4[59,60] apparently uses one pathway for the degradation of toluene, ethylbenzene, propylbenzene, and *sec*-butylbenzene, and the dioxygenase that initiates degradation of these compounds is most closely related to the isopropylbenzene dioxygenase from *Pseudomonas* sp. JR1 (see below).

Numerous biphenyl degradation pathways in both Gram-negative and Gram-positive bacteria are initiated by BPDOs. BPDOs from *Pseudomonas pseudoalcaligenes* KF707,[96] *Burkholderia xenovorans* LB400,[85] BPDO *Comamonas testosteroni* B-356,[271] and *Rhodococcus* sp. RHA1[184] have been studied in some detail. Closely related enzyme systems can be found in *Rhodococcus globerulus* P6,[12] and *P. putida* KF715,[119] and many of these enzymes have the ability to attack various polychlorinated biphenyls (PCB) congeners (see Section 4.5). One BPDO, from *Rhodococcus erythropolis* TA421,[7] does not cluster tightly with this family (Figure 2).

Isopropylbenzene (cumene) dioxygenases (Figure 2) have been identified from several bacteria. *P. putida* RE204 and *P. fluorescens* IP01 were isolated by selective enrichment for growth with isopropylbenzene as sole carbon and energy source.[6,77] *P. putida* RE204 also grew with toluene, ethylbenzene and *n*-butylbenzene, but not benzene. *P. fluorescens* IP01 also grew with toluene, ethylbenzene, *sec*-butylbenzene, and *tert*-butylbenzene, but not biphenyl or benzene.[6] *Pseudomonas* sp. JR1 was isolated from an isopropylbenzene enrichment and was able to grow with toluene. Several Gram-positive isopropylbenzene degraders have been characterized. *Rhodococcus erythropolis* BD2 was isolated from a toluene- and TCE-contaminated soil sample with isopropylbenzene as sole carbon source.[63] BD2 utilizes a linear conjugative plasmid-encoded isopropylbenzene degradation pathway identical to that described for *P. putida* RE204.[62] The isopropylbenzene dioxygenase proteins in *R. erythropolis* BD2 are most closely related to the corresponding proteins from the biphenyl pathway in *Rhodococcus* sp. RHA1, ranging from 87 to 99% amino acid sequence identity.[148]

Compared to other alkylbenzenes, there have been few reports of bacterial isolates capable of growth with *o*-xylene, and most characterized strains utilize a monooxygenase-mediated pathway.[14,29,31] Recently, however, *Rhodococcus* sp. strain DK17 was isolated for its ability to grow with *o*-xylene.[152] DK17 also grows with benzene, toluene, ethylbenzene, and isopropylbenzene, and a single mutation eliminated the ability to grow with all of these substrates as well as the ability to convert indole to indigo. These results suggested that a single oxygenase carries out the initial oxidation of all substrates, but two different pathways are used to complete the degradation of benzene and alkylbenzenes. Results of catechol dioxygenase assays demonstrated that the alkybenzenes were degraded through a *meta* pathway, but growth with benzene induced an *ortho*-cleavage dioxygenase.[152] Surprisingly, the genes encoding the oxygenase subunits, ferredoxin and reductase components (*akbA1A2A3A4)* were identical in sequence to the ethylbenzene dioxygenase components from *Rhodococcus* sp. RHA1 and the α subunits do not cluster with the toluene/biphenyl family (RHA1 EtbA1; DK17 AkbA1a, Figure 2).[151]

2.5. Diversity of Benzoate Family Dioxygenases

The benzoate family is thus far represented by two-component enzymes (reductase and dioxygenase) where the dioxygenase is composed of both α and β subunits. Substrates for these enzymes include various aromatic acids and aniline. Enzymes that have been characterized include benzoate dioxygenases from *Acinetobacter* sp. ADP1,[197] *P. putida* mt-2,[132] and *Rhodococcus* sp. strain 19070,[112] and toluate dioxygenase from *P. putida* pWWO[116] (Reaction B, Table 1). Closely related enzymes for anthranilate degradation have been identified in several organisms, including *Pseudomonas resinovorans* CA10[203] and

Acinetobacter sp. ADP1.[47,79] Interestingly, the anthranilate dioxygenase from *Burkholderia cepacia* DBO1[52] is a three-component enzyme whose α subunit sequence clusters with the salicylate family (DBO1 AndAc, Figure 2). In each case, however, these enzymes catalyze *cis*-dihydroxylation, deamination, and decarboxylation to form catechol (Reaction G, Table 1). In addition, cumate (isopropylbenzoate) dioxygenase and 2,4,5-trichlorophenoxyacetate oxygenase have been characterized from *P. putida* F1[75] and *B. cepacia* AC1100,[64] respectively. The 2-halobenzoate-1,2-dioxygenases from *Burkholderia* (formerly *Pseudomonas*) *cepacia* 2CBS and *Burkholderia* sp. TH2 and chlorobenzoate dioxygenase from *Burkholderia* sp. NK8 also fall into this family.[90,109,270] These enzymes catalyze *cis*-dihydroxylation, decarboxylation, and dehalogenation of *o*-halogenated aromatic acids (Reaction D, Table 1), although they differ in structure and sequence from enzymes catalyzing similar reactions (see Section 2.6). Finally, aniline dioxygenases from *P. putida* UCC22 (pTDN1),[94] *Acinetobacter* sp. YAA aniline dioxygenase,[93] *Delftia acidovorans* 7N,[284] and *Fraturia* sp. ANA-18[191] also group with this family. These enzymes catalyze *cis*-dihydroxylation and deamination of aniline to form catechol (Reaction F, Table 1).

2.6. Diversity of Salicylate Family Dioxygenases

Salicylate 5-hydroxylase (Figure 2) from *Ralstonia* sp. strain U2[305] catalyzes the conversion of salicylate to gentisate (Reaction P, Table 2), and is a key enzyme in the naphthalene degradation pathway in this strain (Figure 3).[92,137] Interestingly, the ferredoxin and reductase components of salicylate 5-hydroxylase are shared with NDO.[305] Five multicomponent salicylate 1-hydroxylases were recently identified in three different sphingomonads.[55,68,220] These enzymes catalyze the conversion of salicylate to catechol, but are unrelated to the well-characterized single polypeptide flavoprotein monooxygenases that are known to catalyze the same reaction in other organisms. The salicylate 1-hydroxylase from *Sphingomonas yanoikuyae* B1 was shown to share its ferredoxin and reductase components with a NDO/BPDO and toluate dioxygenase,[55] and this also appears to be the case in the other two sphingomonads.[68,220] It is not clear at this time, however, whether these multicomponent salicylate 1-hydroxylases catalyze an initial dioxygenation of the aromatic ring of salicylate with the formation of an unstable compound that rearranges, releasing water and CO_2 (analogous to Reaction G, Table 1), or if they actually function as monooxygenases that attack directly at the carboxyl-substituted carbon.

P. aeruginosa strains 142 and JB2 utilize halobenzoates using an interesting three-component dioxygenase that is related to the salicylate hydroxylases (Figure 2). The enzyme catalyzes the oxygenolytic *ortho*-dehalogenation

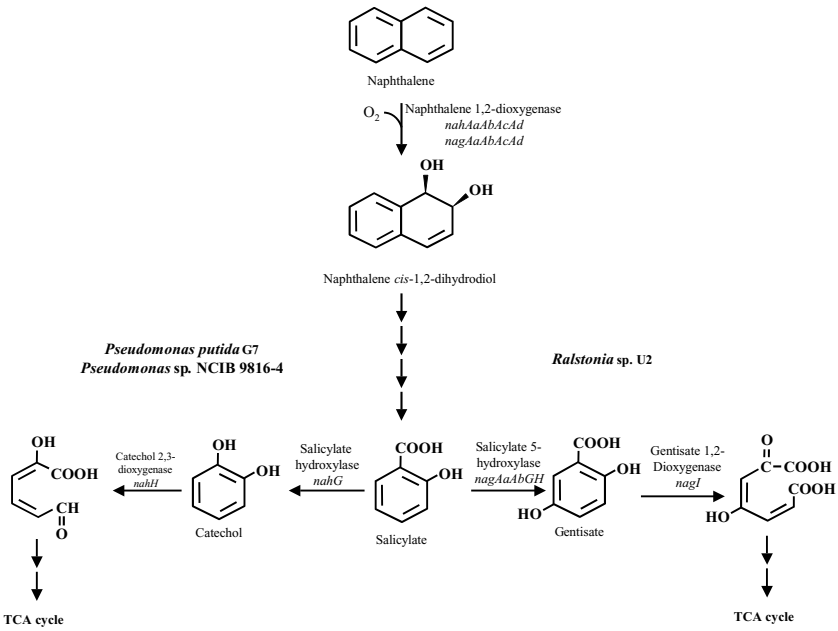


Figure 3. Pathways for the aerobic degradation of naphthalene in pseudomonads and *Ralstonia* sp. U2.

of 2-chlorobenzoate to form catechol (Reaction D, Table 1).[283] The electron transport protein-encoding genes are not co-localized with the oxygenase α and β subunit genes in either strain.[120,283] Terephthalate dioxygenase is also related to these enzymes[252,291] and is different from phthalate dioxygenases in structure and sequence (see Section 2.7). Like other members of the salicylate family, terephthalate dioxygenase has two components, reductase and oxygenase, and the oxygenase is composed of α and β subunits.[241,252,291] As mentioned in Section 2.5, the unusual anthranilate dioxygenase from *B. cepacia* DBO1 clusters with this family.[52]

2.7. Diversity of Phthalate Family Dioxygenases

Most current members of the phthalate family of Rieske non-heme iron oxygenases are two-component enzymes, each consisting of an α_n oxygenase component (lacking β subunits) and a reductase component. Phthalate dioxygenase, the namesake of this family, has been identified in a number of different Gram-positive and Gram-negative bacteria.[76,111,204,260] The well-studied phthalate dioxygenase from *Burkholderia cepacia* DBO1[19,20,61] is encoded by

the *ophA1* and *ophA2* genes, which are separated by the genes encoding the next two steps in phthalate degradation [dihydrodiol dehydrogenase (*ophB*) and decarboxylase (*ophC*)].[53] Genes encoding the oxidation of isophthalate have recently been identified from *C. testosteroni* YZW-D.[291] Like phthalate dioxygenase, isophthalate dioxygenase is a two-component enzyme with a homomultimeric oxygenase. A related enzyme is the chlorobenzoate 3,4-dioxygenase (CbaAB) from *Alcaligenes* sp. (BR60), which catalyzes the *cis*-dihydroxylation and dechlorination of 3-chlorobenzoate (Reaction C, Table 1).[193]

The phenoxybenzoate dioxygenase from *P. pseudoalcaligenes* POB310 catalyzes the angular dioxygenation (Reaction I, Table 1) of 3- and 4-carboxydiphenyl ether.[66] The unstable product of the reaction spontaneously rearranges to release phenol and protocatechuate. Like other members of this family, the enzyme has two components, PobA and PobB and the oxygenase (PobA) lacks a β subunit.

Some enzymes in this family do not appear to function as dioxygenases with their native substrates, but as monooxygenases. 2-Oxo-1,2-dihydroquinoline-8-monooxygenase from *P. putida* 86[236] is involved in the degradation of quinoline. This two-component enzyme is encoded by *oxoO* (oxygenase component) and *oxoR* (reductase component), which are separated by 15 kb of DNA.[237] The reaction results in net monooxygenation (Reaction P, Table 2) of the aromatic ring of 2-oxo-1,2-dihydroquinoline. Toluene sulfonate methyl-monooxygenase from *C. testosteroni* T-2 oxidizes the methyl group of toluene sulfonate, converting it to 4-sulfobenzyl alcohol (Reaction K, Table 2). The *tsaMB* genes encode the two components (oxygenase and reductase) of the enzyme.[139] The same strain, *C. testosteroni* T-2, also has a sulfobenzoate 3,4-dioxygenase that catalyzes the dioxygenation and desulfonation of sulfobenzoate (Reaction H, Table 1) to form protocatechuate and sulfite.[177] Vanillate demethylase, encoded by the *vanAB* genes, has been identified in *P. putida* WCS358,[286] *Acinetobacter* sp. ADP1,[245] *Pseudomonas* sp. ATCC 19151,[46] *P. fluorescens* BF13,[56]and *Pseudomonas* sp. HR199.[223] This enzyme catalyzes the oxygenative demethylation (Reaction N, Table 2) of vanillate to protocatechuate, and plays an important role in the degradation of the common plant metabolite ferulate. The product of the *ligX* gene from *Sphingomonas paucimobilis* SYK-6 encodes a 5,5′-dehydrodivanillic acid *O*-demethylase that participates in the complex pathway of lignin metabolism in this strain.[258]

Unlike other members of this family, carbazole dioxygenases from *P. resinovorans* CA10 and *P. stutzeri* OM1 are three-component enzymes, consisting of oxygenase (α_3), ferredoxin (Rieske [2Fe–2S]), and reductase (plant-type [2Fe–2S] and FAD).[207,239] Interestingly, both strains carry adjacent duplicate dioxygenase genes (CarAa). Carbazole dioxygenase catalyzes an angular dioxygenation (Reaction I, Table 1) to form 2′-aminobiphenyl-2,3-diol.

2.8. Other Ring Hydroxylating Dioxygenases

Several additional multicomponent Rieske non-heme iron oxygenases do not fall into any of the major families. The diterpenoid dioxygenase from *P. abietaniphila* BKME-9 catalyzes the dioxygenation of 7-oxo-dehydroabietic acid (Figure 1). A large gene cluster encoding several steps in the abietane diterpenoid degradation pathway has been characterized.[183,182,257] The *ditA1A2* genes encode the dioxygenase, which is not closely related to other dioxygenases (Figure 2); and the *ditA3* gene encodes an atypical [4Fe–4S] or [3Fe–4S] ferredoxin.[182] Apparently, the gene encoding the reductase has not yet been identified.

The 3-phenylpropionate dioxygenase from *Escherichia coli* K-12 catalyzes the initial dioxygenation of 3-phenylpropionate and cinnamate to the corresponding *cis*-2,3-dihydrodiols.[71] This three-component enzyme is most closely related to members of the toluene/biphenyl family of enzymes, but does not cluster with the family (Figure 2).

Several enzymes that catalyze angular dioxygenation (Reaction I, Table 1) on their respective substrates do not cluster tightly with the defined dioxygenase families (Figure 2). *Terrabacter* sp. strain DBF63 has an interesting angular dioxygenase that catalyzes the initial oxidation of dibenzofuran. The *dbfA1A2* genes encode dibenzofuran 4,4a-dioxygenase, and these proteins do not cluster with their counterparts in any of the dioxygenase families.[145] The ferredoxin component is encoded by the *dbfA3* gene, which is located 2.5 kb downstream of the *dbfA1A2* genes.[111] The ferredoxin contains a [3Fe–4S] center rather than a typical [2Fe–2S] center,[145,272] and the reductase component has not yet been identified. In contrast, a tightly clustered set of four genes encoding dibenzofuran dioxygenase was cloned from *Terrabacter* sp. YK3.[129] Several differences were noted between the *Terrabacter* sp. YK3 and *Terrabacter* sp. strain DBF63 dibenzofuran dioxygenases. The dioxygenase α and β subunits from YK3 were quite distantly related to those from DBF63, the YK3 ferredoxin carried a [2Fe–2S] cluster rather than a [3Fe–4S] center as in DBF63, and the gene encoding the YK3 FAD-containing reductase was located just downstream of the ferredoxin gene. The dibenzofuran/dibenzo-*p*-dioxin dioxygenase from *Sphingomonas wittichii* RW1[300] also catalyzes angular dioxygenation on these substrates.[48] The α and β subunits of the RW1 dioxygenase are only distantly related to those from DBF63 and YK.[39,129,145] Two isofunctional putidaredoxin-type (rather than the more common Rieske type) ferredoxins and two isofunctional flavin-containing reductases were shown to have the ability to transfer electrons to the oxygenase, and the genes encoding these components are apparently dispersed on the chromosome.[8,10,11] The carbazole dioxygenase from *Sphingomonas* sp. CB3 also does not group with any of the main families.[251] This three-component carbazole dioxygenase is very different from that from *P. resinovorans* CA10. The oxygenase has both α and β subunits unlike that

from CA10, and the α subunits from these two proteins only share 13% amino acid sequence identity. At the time it was identified, the CB3 carbazole dioxygenase was actually most closely related to isopropylbenzene dioxygenases and BPDOs from Gram-positive bacteria, but it does not cluster with the toluene/biphenyl family. More recently, it was found to be more closely related to the dioxin dioxygenase from *Sphingomonas wittichii* RW1.[200] Nojiri *et al*.[200] have argued that the enzyme from *Sphingomonas* sp. CB3 may not actually be a carbazole dioxygenase, because the only evidence for its role in carbazole degradation is the fact that the genes are induced during growth with carbazole. However, it is unlikely to be a BPDO, because biphenyl is not a growth substrate for CB3 and carbazole-grown CB3 cells do not oxidize biphenyl.[251]

2-Aminobenzene sulfonate dioxygenase is a two-component enzyme system comprised of an oxygenase with both α and β subunits, and a reductase. The oxygenase has been purified and characterized.[181] The enzyme (O-1 AbsA, Figure 2) from *Alcaligenes* sp. O-1 initiates the pathway for 2-aminobenzoate degradation by catalyzing *cis*-dihydroxylation and deamination to form 3-sulfocatechol (Reaction F, Table 1). The most closely related enzyme to AbsA, surprisingly, is methanesulfonic acid monooxygenase (M2 MsmA, Figure 2) from *Methylosulfonomonas methylovora* M2, an enzyme that oxidizes a non-aromatic substrate.[65] The enzyme has three components. The oxygenase is composed of α and β subunits and has an unusual iron–sulfur center-binding motif: while most α subunits of Rieske centers have 15–17 residues between the two Cys-His pairs, this enzyme has 26. The ferredoxin and reductase components are more like those from diiron center-containing monooxygenases such as toluene 4-monooxygenase from *P. mendocina* KR1, than from aromatic ring hydroxylating dioxygenase systems, making this quite an unusual enzyme.[65]

3. RING HYDROXYLATING DIOXYGENASE STRUCTURE AND FUNCTION

To date, protein components from several dioxygenase systems have been purified and studied in detail. In this chapter, we will focus on NDO as a model system and as a basis for comparison to other related enzyme systems.

3.1. Enzymology of NDO from *Pseudomonas* sp. NCIB 9816-4

Two pathways for the aerobic degradation of naphthalene have been described (Figure 3). Both pathways utilize naphthalene 1,2-dioxygenase (E.C. 1.14.12.12) to convert naphthalene to (+)-naphthalene *cis*-(1*R*,2*S*)-dihydrodiol (naphthalene *cis*-dihydrodiol). Naphthalene *cis*-dihydrodiol

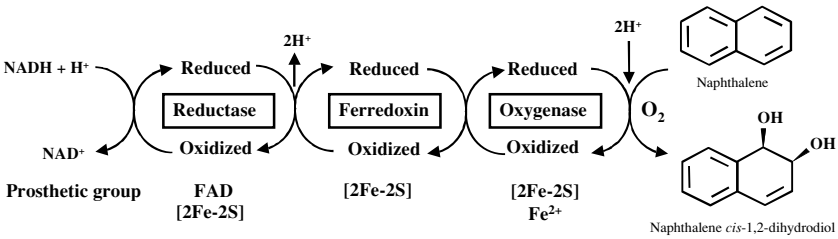


Figure 4. Reaction catalyzed by the three-component naphthalene dioxygenase (NDO) system. Electrons from NADH are transferred by the iron–sulfur flavoprotein reductase to the Rieske [2Fe–2S] ferredoxin. An electron is then transferred from the ferredoxin to one of the Rieske centers in the oxygenase. Reduced oxygenase catalyzes the addition of both atoms of O_2 to naphthalene, forming enantiomerically pure (+)-naphthalene *cis*-(1*R*, 2*S*)-dihydrodiol.

dehydrogenase (E.C. 1.3.1.29) then oxidizes naphthalene *cis*-dihydrodiol to 1,2-dihydroxynaphthalene. The oxidized ring is then cleaved and further degraded to form salicylate (Figure 3). Pseudomonads and some Gram-positive organisms convert salicylate to catechol, which is degraded by a standard *meta* cleavage pathway,[34,50,158,301] although variants that use the *ortho* pathway have been reported.[50,89] In contrast, *Ralstonia* sp. strain U2 and some strains of *Rhodococcus* convert salicylate to gentisate and utilize a gentisate dioxygenase pathway to complete naphthalene degradation (Figure 3).[3,70,305,306]

NDO is a three-component enzyme system that catalyzes the addition of both atoms of oxygen to the aromatic ring of naphthalene (Figure 4). All three protein components of NDO have been purified from *Pseudomonas* sp. NCIB 9816-4.[82,114,115] The reductase is a 35 kDa monomer that contains one molecule of FAD and a plant-type iron–sulfur center. It can accept electrons from either NADH or NADPH.[115,256] The ferredoxin is a Rieske [2Fe–2S] center-containing monomer of approximately 11.4 kDa.[114,256] The catalytic oxygenase component is an $\alpha_3\beta_3$ hexamer consisting of large (α) and small (β) subunits.[147] Each α subunit contains two redox centers, a Rieske [2Fe–2S] center and mononuclear Fe^{2+}at the active site. Individually purified α and β subunits of the oxygenase were reconstituted,[263,264] demonstrating that both subunits are essential for activity, a result consistent with those obtained with BPDO[127] and toluene dioxygenase.[135] As purified, the Rieske center of the oxygenase is oxidized, and the iron at the active site is reduced. Electrons are transferred sequentially from NAD(P)H to the reductase, to the ferredoxin, to the Rieske center of the oxygenase, and finally to the iron at the active site of the oxygenase. The reduced oxygenase catalyzes the addition of both atoms of O_2 to the aromatic ring. Two electrons are necessary to complete the reaction cycle.

Among the aromatic ring hydroxylating dioxygenases that have been identified to date, two types of oxygenase structures are known: those with both α and β subunits, such as NDO, and those consisting of only α subunits,

such as phthalate dioxygenase. Based on studies of hybrid dioxygenases, in which the individual α and β subunits from different enzymes were substituted, it appears that the α subunits of NDO and the closely related enzymes 2-nitrotoluene dioxygenase and 2,4-dinitrotoluene dioxygenase control substrate specificity.[209,212] Similar results were reported with BPDO hybrids, toluene dioxygenase-tetrachlorobenzene dioxygenase hybrids, and benzene-BPDO hybrids.[17,24,186,275,307] These results are consistent with the crystal structures of NDO and related dioxygenases (see Section 3.2), which showed that no β subunit residues are near the active site.[91,98,147] Other studies have suggested, however, that β subunits may play a role determining substrate specificity in toluene, toluate, and other BPDOs.[54,117,121,128] Therefore, it seems as though the β subunit has a structural function in most dioxygenases, but in some cases, the β subunit may be capable of modulating substrate specificity.

3.2. Dioxygenase Structure and Mechanism

Crystal structures of the oxygenase component of NDO have been determined in the presence and absence of various substrates.[51,143,147] NDO is an $\alpha_3\beta_3$ hexamer. The β subunits contain no redox centers and based on the currently available crystal structures of NDO, nitrobenzene dioxygenase (NBDO) from *Comamonas* sp. strain JS765, cumene dioxygenase from *P. fluorescens* IP01, and BPDO from *Rhodococcus* sp. strain RHA1, the β subunit residues are distant from the active sites in each enzyme.[91,98,147] The $\alpha_3\beta_3$ hexamer of NDO contains three active sites, which are located at the junctions of adjacent α subunits. Each α subunit contains a Rieske [2Fe–2S] center, which is coordinated by two histidines (His83; His104) and two cysteines (Cys81; Cys101), and mononuclear non-heme iron. This iron at the active site is in a distorted octahedral conformation coordinated by His208, His213, Asp362 and a water molecule (Figure 5).[147] These two redox center-binding motifs (C-X-H-X_{15-26}-C-X_2-H and H-X_{4-5}-H-X_n-D) are conserved in all sequenced Rieske non-heme iron oxygenases. The 2-His-1-carboxylate mode of mononuclear iron coordination has also been identified in a variety of non-heme iron-containing enzymes that catalyze a wide range of reactions, including not only Rieske non-heme iron oxygenases, but *meta*-cleavage dioxygenases, pterin-dependent hydroxylases such as tyrosine hydroxyase, and a family of β-lactam biosynthesis enzymes.[163,224] The recently published crystal structures of the *Comamonas* sp. strain JS765 NBDO, *Rhodococcus* sp. RHA1 BPDO, and *P. fluorescens* IP01 cumene dioxygenase showed very similar overall structures compared to NDO, including similar relative positions of the redox centers.[74,98] One difference was apparent between NDO and the two members of the toluene/biphenyl family of Rieske non-heme iron oxygenases. The conserved Asp residue that coordinates the active site iron in NDO (Asp 362) does so in a bidentate fashion with both coordinating carboxyl oxygen atoms approximately equidistant from the

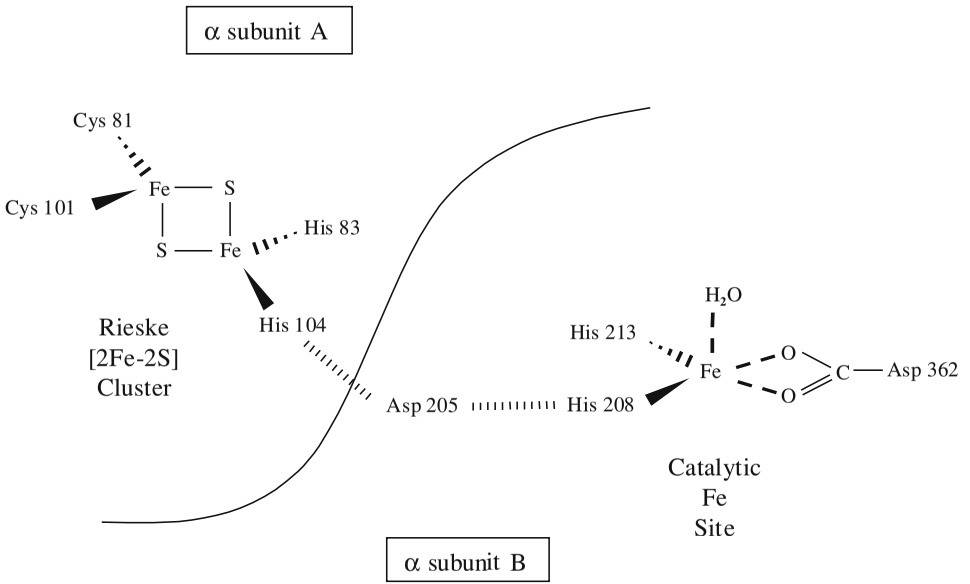


Figure 5. The junction between two α subunits in NDO based on the crystal structure of the enzyme. Shown are the Rieske [2Fe–2S] center and mononuclear iron at the active site. Amino acids Cys81, His83, Cys101, and His104 coordinate the Rieske center; His208, His213, and Asp362 coordinate mononuclear iron at the active site. See text for additional details.

iron.[147] In contrast, the corresponding aspartate residues in BPDO and cumene dioxygenase (Asp 378 and 388, respectively), are coordinated to the iron via a single carboxyl oxygen; in both cases, the other carboxyl oxygen is apparently too far away to coordinate the iron. Therefore, the coordination geometry of the mononuclear iron in both BPDO and cumene dioxygenase is distorted tetrahedral.[74,98]

Substitution of Asp362 with an alanine inactivated NDO, demonstrating that this iron-coordinating residue is essential.[214] In an earlier study, site directed mutagenesis of toluene dioxygenase (TDO) demonstrated that the conserved residues Glu214, Asp219, His222, and His228 (corresponding to Glu200, Asp205, His208, and His213 in NDO) were essential for enzyme activity, and were suggested to be mononuclear iron ligands.[136] Substitution of the corresponding histidines in benzene dioxygenase also resulted in loss of enzyme activity.[49] Of the four potential iron ligands, only the histidines were found to coordinate iron based on the NDO structure.[147] The loss of activity in the other two TDO mutants can be explained using the NDO structure. The structure revealed that Glu200 provides an important contact between adjacent α subunits by forming a salt link with Arg84.[146] Asp205 in NDO is hydrogen bonded to the Rieske center ligand His104 and mononuclear iron ligand His208. Based on this structural conformation, Asp205 appeared to participate in the electron transfer pathway between the NDO Rieske center and active site iron in adjacent α subunits in (Figure 5).[147] This hypothesis was supported by the analysis of mutant forms of NDO in which Asp205 was replaced by Ala, Glu, Gln, or Asn.[215] The modified proteins had little or no dioxygenase activity, although they appeared to have no major structural defects. Additional support for this role for Asp205 was provided when benzene, an efficient uncoupling agent for wild-type NDO,[168] was unable to stimulate oxygen uptake by the purified Gln205-containing enzyme. This result suggests that electrons cannot be passed from the Rieske center to the active site iron in the mutant form of the enzyme.[215] An alternative role for this conserved aspartate was proposed in studies of anthranilate dioxygenase, an enzyme that catalyzes the first step in anthranilate catabolism in *Acinetobacter* sp. strain ADP. In anthranilate dioxygenase, substitution of the aspartate corresponding to Asp205 in NDO with Ala, Glu, or Asn resulted in completely inactive enzymes. Analysis of wild-type and mutant forms of anthranilate dioxygenase at various pH values suggested that this aspartate may play a role in maintaining the reduction potential and the protonation state of the Rieske center of the oxygenase.[22]

As purified, NDO contains an oxidized Rieske center and ferrous iron at the active site. Spectroscopic studies with the *Burkholderia cepacia* DBO1 phthalate dioxygenase, which catalyzes the *cis*-dihydroxylation of phthalate, demonstrated that the iron is in a six-coordinate octahedral configuration in the absence of substrate, but a five-coordinate configuration results when substrate

is bound.[101,217,282] This result is in contrast to crystal structure data for NDO, which indicates that the coordination of iron does not change upon substrate binding, and in fact no significant conformational changes occur anywhere in the active site when substrate binds.[51,143] The crystal structure of the *Rhodococcus* sp. RHA1 BPDO revealed that when biphenyl was bound, significant conformational changes occurred at the active site.[98] These conformational changes in BPDO resulted in the formation of an entryway and an expansion of the substrate-binding pocket to allow biphenyl to enter and bind the active site pocket. The active site pockets of the *Rhodococcus* sp. RHA1 BPDO and the *P. fluorescens* IP01 cumene dioxygenase are strikingly similar,[74] which is consistent with their overall level of primary amino acid sequence identity (67%).

The nature of the iron-oxygen species involved in catalysis is not known. Studies of 4-methoxybenzoate monooxygenase suggested an attack by a ferric-peroxo intermediate in the monooxygenation and *cis*-dihydroxylation reactions,[27] and recent work with NDO supports this possibility.[51,143,168] The participation of a high valent iron-oxo species, as in the methane monooxygenase[175] and cytochrome $P450^{242}_{cam}$ reaction cycles, has also been proposed.[225,298,299] Single turnover studies with NDO demonstrated that 0.85 units of product were formed per electron added, and both the Rieske center and mononuclear iron were oxidized during the reaction,[299] and peroxide-mediated turnover with NDO was recently reported.[298] The observation that the reaction stopped after one turnover is consistent with the idea that the release of product requires the input of a second electron. NDO is known to catalyze other reaction types in addition to *cis*-dihydroxylation (dioxygenation), including benzylic monohydroxylation (monooxygenation[234,288]), *O*- and *N*-dealkylation,[229] sulfoxidation,[169] and desaturation[107,280] reactions (Table 2). However, the *cis*-dihydroxylation reaction is unique to bacterial aromatic ring-hydroxylating dioxygenases, and it has implications for the mechanism of the enzyme. Recent crystal structure data demonstrated that oxygen binds side-on to the iron at the active site of NDO, and the substrates naphthalene or indole bind in a long cleft in the active site approximately 4 Å from the iron atom.[51,143] Alone, dioxygen binds side-on approximately 2.2 Å from the iron, but measurably closer to the iron when substrate is present.[143] This series of NDO crystal structures with substrate, oxygen, both substrate and oxygen, or product (naphthalene *cis*-dihydrodiol) bound suggest a concerted reaction mechanism in which both atoms of molecular oxygen react with the carbon atoms of the substrate double bond in a way that would explain the *cis*-specific dioxygenation catalyzed by this class of enzymes. Based on the structures of NDO alone and in complex with substrate, oxygen, and product, the reaction cycle shown in Figure 6 was proposed.[143]

Three ferredoxin reductase structures from three different dioxygenase systems are currently available. The phthalate dioxygenase reductase

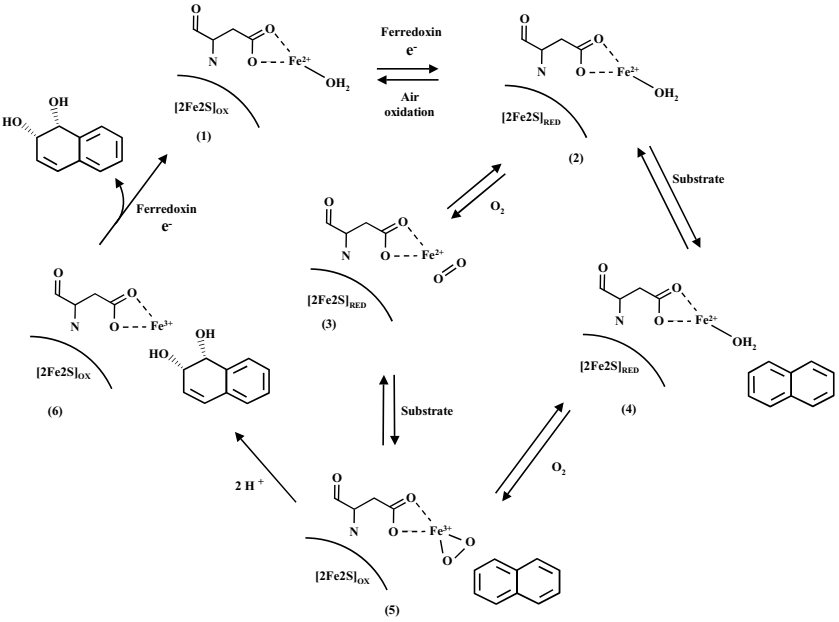


Figure 6. Proposed mechanism of NDO based on crystal structures of the enzyme alone, with bound substrate (naphthalene or indole), with bound O_2, with bound substrate and O_2, or with bound product (naphthalene *cis*-dihydrodiol). (1) As purified, the enzyme has an oxidized Rieske site. (2) One electron is transferred from the ferredoxin. Either oxygen (3) or substrate (naphthalene) (4) can bind at the active site. (5) Ternary complex with naphthalene and oxygen bound at the active site. (6) Product formation; release of product requires input of an additional electron from the ferredoxin. Reproduced wth permission from Karlsson *et al*. [143].

(PDR) from *Burkholderia* (formerly *Pseudomonas*) *cepacia* DB01[61] and the benzoate dioxygenase reductase (BenC) from *Acinetobacter* sp. strain ADP1[142]are members of the $NADP^+$-ferredoxin reductase superfamily. PDR is the most well-studied of the electron transfer partners from this family of dioxygenases.[19,61,99,100,102] Benzoate and phthalate dioxygenases are two-component systems consisting of an oxygenase and a ferredoxin reductase.[20] The ferredoxin reductases are both composed of three distinct domains, for binding a plant-type [2Fe–2S] center, NAD, and flavin (FMN in PDR; FAD in BenC), but the arrangement of the domains differs in the two proteins.[61] The BPDO reductase (BphA4) from the three-component enzyme in *Pseudomonas* sp. strain KKS102 has very low sequence identity with other aromatic ring-hydroxylating dioxygenase reductase components.[150] It contains NAD- and FAD-binding domains, but no iron–sulfur cluster; electrons are transferred from BphA4 to the [2Fe–2S] center in an intermediary ferredoxin protein. The

protein fold of BphA4 is actually similar to that of glutathione reductases[248] and putidaredoxin reductase from *P. putida* cytochrome $P450_{cam}$.[250] Although these reductases belong to two distinct families, the overall structures of the NAD- and FAD-binding domains are quite similar.[142] On the basis of modeling studies, the NDO reductase structure is predicted to be similar to that of benzoate dioxygenase reductase.[142]

The structure of the BPDO ferredoxin component (BphF) from *Burkholderia* sp. strain LB400 was determined at 1.6 Å resolution.[57] The structure of the carbazole dioxygenase ferredoxin component (CarAc) from *P. resinovorans* CA10 was solved by molecular replacement with the BphF structure.[194] Although the sequences of the two proteins are only 34% identical, their structures were very similar, and the location of the Rieske center at the surface of the proteins may have implications for docking with the reductase and oxygenase components to allow electron transfer. One notable difference was in the surface charges of the two proteins, which probably determine the ability to interact productively (and specifically) with the respective oxygenases of each enzyme system.[194] To our knowledge, no reports have yet been published demonstrating co-crystallization of non-covalent complexes of ferredoxins with their corresponding oxygenases or ferredoxin reductases for any of the three-component aromatic ring hydroxylating dioxygenase systems, and information regarding specific protein–protein interactions is limited. However, cross-linking studies with the three toluene dioxygenase components suggested the involvement of electrostatic interactions between the proteins, and argued against the formation of a ternary complex.[167] These results imply that the ferredoxin in three-component enzyme systems alternately binds to the reductase and oxygenase components using the same surface near the exposed Rieske center in order to carry out electron transfer.

3.3. Substrate Specificity of NDO and Other Dioxygenases

In general, aromatic ring hydroxylating dioxygenases are capable of initiating oxidative attack on a very wide range of substrates.[41,232] Many of these enzymes display remarkable diversity in the number of substrates oxidized and the types of reactions catalyzed. Our understanding of the mechanism of substrate oxidation has been based on studies with blocked mutant strains that accumulate the products of dioxygenase-catalyzed reactions, recombinant strains expressing dioxygenase genes, and in some cases with purified enzymes. NDO from *Pseudomonas* sp. NCIB 9816-4 is known to catalyze the oxidation of more than 75 different substrates by reactions including *cis*-dihydroxylation, monooxygenation,[234,288] desaturation,[107,280] *O*- and *N*-dealkylation,[229] and sulfoxidation[169] (Tables 2 and 3). The complete list of substrates oxidized and products formed by NDO is summarized in a useful

Table 3. Representative reactions catalyzed by Rieske non-heme iron oxygenases for specific pathway substrates.

	Type of oxidation reaction[a,b]																
	Dihydroxylations									Monohydroxylations							
Substrate	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	References
Aromatic hydrocarbons																	
Naphthalene	X																33,69,83,133,164,256,274
Benzene	X																103,130,276
Toluene	X										T						171,189,262,309
Ethylbenzene	X									T		T					59,171
Styrene	X/T																141,140,170,292
Isopropylbenzene	X																6,63,77
o-Xylene	X										T						152,171
Biphenyl	X																7,12,85,96,119,184,271
Anthracene	X/T																1,134,201
Phenanthrene	X/T																134,165,238
9,10-Dihydrophenanthrene	T									T							40,230
Chrysene	T																35,39,104
Pyrene	X																149,156,290
Fluoranthene	X																201,261
Acenaphthene										T							246
Fluorene	X/T									T							108,231,247
Substituted aromatic compounds																	
Aniline						X											93,94,284
Anthranilate							X										47,52,191
Benzoate		X															112,132,197
Salicylate																X	55,68,220,305
o-Halobenzoates			X	X													90,109,270,283
Chlorobenzoates	T	X	X														90,192,193

(*Continued*)

Table 3. (*Continued*)

Substrate	Type of oxidation reaction[a,b]																References
	Dihydroxylations									Monohydroxylations							
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	
Isopropylbenzoate	X																75
3-Phenylpropionate	X																71
Cinnamate	X																71
Chlorobenzenes	T		X														23,126,285
(poly)Chlorobiphenyls	T		T														15,21,105,206,249
Phthalate (*o*-)	X																53,76,111,195,207
Isophthalate		X															291
Terephthalate		X															241,252,291
Toluates (*o*-, *m*-)		X															116
4-Methoxybenzoate														X			25
Vanillate														X			46,223,245
p-Toluene sulfonate											X						139,178
Aminobenzene sulfonate						X											181
p-Sulfobenzoate								X									177
Nitrobenzene					X												173
2-Nitrotoluene	T				X						T						4,208,212,235
3-Nitrotoluene					T						T						173,235
4-Nitrotoluene	T				T						T						173,235
2,4-Dinitrotoluene					X						T						173,266
2,6-Dinitrotoluene					X												198
2,2′-Dinitrobiphenyl					T												244
1,2,4-Trimethylbenzene											T						246
1- and 2-substituted naphthalenes	T	X															30,44,67,154,294
Dimethylnaphthalenes											T						30,246
7-Oxodehydroabietic acid	X																182
5,5′-Dihydrodivanillic acid														X			258
Heterocyclic aromatic compounds																	
Carbazole	T								X								110,201,207,233,239,251
Dibenzofuran	T								X								48,129,145,231,244,272
Dibenzodioxin	T								X								8–11,244

Table 3. (*Continued*)

Substrate	Type of oxidation reaction[a,b]																References
	Dihydroxylations									Monohydroxylations							
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	
Dibenzothiophene	X/T												T				155,201,231
Indole	T																84
Fluorenone	T								X								246,247
Quinolines	T																38,41
2-Oxo-1,2-dihydroquinoline																X	236,237
Pyrazon	X																240
3-Methyl benzothiophene											T		T				246
2-Methylbenzo-1,3-thiole													T				2
Carbocyclic, alkyl-aryl ether, thioether or N-alkyl substrates																	
Indan										T		T					43,107,288
Indene	T									T							58,107,281
1,2-Dihydronaphthalene	T									T		T					78,280
Tetralin	X																187,255
Methyl phenyl sulfide													T				2,169
Ethyl phenyl sulfide													T				2,169
Methyl *p*-tolyl sulfide													T				169
p-Methoxyphenyl methyl sulfide													T				169
Methyl *p*-nitrophenyl sulfide													T				169
Anisole	T													T			229
Phenetole	T											T		T			229
Carboxydiphenylethers									X								66
N-Methylindole															T		232
N-Methylaniline															T		166
N,*N*-Dimethylaniline															T		166

[a] The types of reactions are shown in Tables 1 and 2: A, *cis*-dihydroxylation (C=C); B, *cis*-dihydroxylation (at and adjacent to a carboxyl bearing carbon); C, *cis*-dihydroxylation and dehalogenation; D, *cis*-dihydroxylation, decarboxylation, and dehalogenation; E, *cis*-dihydroxylation and nitrite elimination; F, *cis*-dihydroxylation and deamination; G, *cis*-dihydroxylation, deamination, and decarboxylation; H, dihydroxylation and desulfonation; I, angular dihydroxylation; J, benzylic hydroxylation; K, methyl group hydroxylation; L, oxygen-dependent desaturation; M, sulfoxidation; N, *O*-dealkylation; O, *N*-dealkylation; P, net aromatic ring monohydroxylation.

[b] "*X*" denotes reactions catalyzed for growth substrates; "*T*" indicates a transformation reaction catalyzed for a non-growth substrate. This list is not exhaustive and additional information, particularly on transformation substrates, can be found in recent reviews.[36,42,125,232]

review.[232] TDO from *P. putida* F1 is even more impressive in its substrate range, with the ability to catalyze over 100 oxidation reactions on monocyclic aromatic compounds, fused and linked aromatic compounds, heterocyclic aromatic compounds, and a variety of halogenated and non-halogenated aliphatic olefins.[42,125,161] TDO has overlapping but frequently distinct specificity from NDO, and as with NDO, multiple reaction types are catalyzed by TDO [dioxygenation, monooxygenation,[288] desaturation,[280] and sulfoxidation[2,169] (Tables 2 and 3)]. For example, TDO oxidizes indan to (1*R*)-indanol and converts indene to *cis*-(1*S*, 2*R*)-indandiol and (1*R*)-indenol[288] while NDO oxidizes indan to (1*S*)-indanol and oxidizes indene to *cis*-(1*R*, 2*S*)-indandiol and (1*S*)-indenol.[107] Similar trends in enantioselectivity are typically observed for the *cis*-dihydroxylation, benzylic monohydroxylation, and sulfoxidation reactions catalyzed by NDO.[232] The range of multi-ring substrates oxidized by the angular dioxygenase carbazole 1,9a-dioxygenase from *P. resinovorans* CA10 has been reported and these results were compared with the products formed by the *Terrabacter* sp. dibenzofuran dioxygenase.[201,273] Both enzymes are capable of standard *cis*-dihydroxylation, angular dioxygenation, and monooxygenation depending on the substrate, suggesting that the specificity is determined by the position of the substrate in the active site relative to the active site iron atom.

3.4. Critical Amino Acids at the Active Site of NDO and Other Dioxygenases

The active site iron of NDO is located in a hydrophobic pocket that accommodates the predominantly hydrophobic substrates for the enzyme. The amino acids located near the active site were identified from the crystal structure of NDO.[51,143] Site-directed mutagenesis of residues near the active site iron has identified those that are important in determining the substrate specificity and enantioselectivity of NDO. These studies demonstrated that NDO was able to tolerate a wide range of single amino acid substitutions near the active site.[214,216] Enzymes with substitutions at position 352 of the oxygenase α subunit (phenylalanine in the wild-type enzyme) had the most striking changes in substrate specificity. These enzymes had altered enantioselectivity with naphthalene, biphenyl, phenanthrene, and anthracene, and changes in the regioselectivity were seen when biphenyl and phenanthrene were provided as substrates.[214,216] Replacement of Phe352 with smaller amino acids (Gly, Ala, Val, Ile, Leu, Thr) resulted in enzymes that produced increased amounts of biphenyl *cis*-3,4-dihydrodiol relative to biphenyl *cis*-2,3-dihydrodiol, and the stereochemistry of the biphenyl *cis*-3,4-dihydrodiol was altered. The NDO-F352V and NDO-F352T enzymes formed significant amounts of (−)-biphenyl *cis*-(3*S*,4*R*)-dihydrodiol, a compound not produced by wild-type NDO. A major shift in the regioselectivity was also seen with enzymes carrying substitutions

at Phe352 with phenanthrene as a substrate. In addition, enzymes with substitutions at position 206 (NDO-A206I and NDO-A206I/L253T) formed significantly more phenanthrene *cis*-1,2-dihydrodiol than wild type.[302] Several of the enzymes formed phenanthrene *cis*-9,10-dihydrodiol, a new product not formed by the wild type.

An attempt to engineer NDO into an enzyme with the substrate specificity of 2-nitrotoluene dioxygenase by making the corresponding amino acid substitutions in the active site was only partially successful. The α subunits of the two enzymes are 84% identical in primary amino acid sequence, with only five amino acid differences at the active site. The substrate specificities of the two enzymes are significantly different.[209] Only two amino acid substitutions at the active site of NDO (F352I, A206I) were necessary to change the enantioselectivity with naphthalene to that of 2-nitrotoluene dioxygenase, which forms 70% (+)-(1*R*,2*S*)-*cis*-naphthalene dihydrodiol.[302] Wild-type NDO makes >99% (+)-(1*R*,2*S*)-*cis*-naphthalene dihydrodiol. None of the mutant forms of NDO was capable of oxidizing the aromatic ring of 2-nitrotoluene with release of nitrite. In fact, enzymes with four or more amino acid substitutions were inactive,[302] and all of the mutant enzymes had lower overall product formation rates with all tested substrates.[214,216,302]

The nitrobenzene and 2-nitrotoluene dioxygenase components have been purified[213] and the crystal structure of the oxygenase component of NBDO has been solved.[91] The overall structure of the $\alpha_3\beta_3$ hexamer of NBDO looks strikingly similar to that of NDO,[147] and an overlay of the $\alpha\beta$ heterodimers from the two enzymes highlights the similarity (Figure 7a). However, several amino acid differences at the active site can account for the observed differences in substrate specificity. In particular, an asparagine at position 258 in the α subunit of NBDO (corresponding to Val260 in NDO) was shown to form a hydrogen bond with the nitro groups of nitro-substituted aromatic compounds (Figure 7b). This hydrogen bond positions the nitroarene substrates for oxidation at the nitro-substituted carbon.[91] Replacement of this asparagine with a valine in NBDO resulted in an enzyme with a significantly reduced ability to oxidize the aromatic ring of nitro-substituted aromatics.[138] A similar result was obtained with the same substitution in 2-nitrotoluene dioxygenase.[172] It therefore appears that the orientation of the nitroarene substrate in the active site governs the regioselectivity of the enzyme.

Recent studies have used the crystal structure of the NDO active site to predict active site residues in related dioxygenases. Based on such a model, the active site of the tetrachlorobenzene dioxygenase from *Ralstonia* sp. PC12 (TecA) was modified to identify specific amino acids involved in controlling both regioselectivity and product formation rate.[222] Substitution of Phe366 in the α subunit of TecA (corresponds to Phe352 in NDO) with tyrosine or tryptophan resulted in inactive enzymes,[222] a result similar to that obtained with the corresponding NDO mutants.[216] Substitution of smaller residues at this position

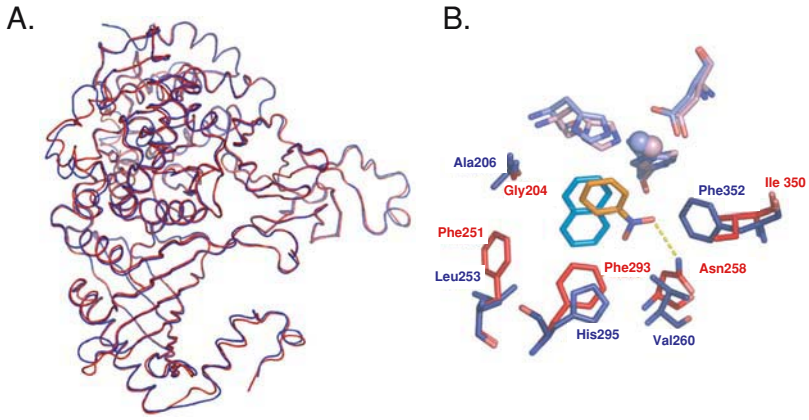


Figure 7. Comparison of the NDO and NBDO structures. (a) Overlay of αβ heterodimers (NDO in blue; NBDO in red). The α subunits are toward the top and β subunits toward the bottom of the figure. The active site iron and coordinating histidine side chains are visible in the α subunit. (b) Overlay of active site structures of NDO and NBDO (NDO in blue; NBDO in red). Histidine and aspartate residues coordinating iron at the active site are in light blue and pink for NDO and NBDO, respectively. Five residues at the active that differ between the two enzymes are shown. Bound naphthalene in NDO is aqua; bound nitrobenzene in NBDO is yellow.

in TecA resulted in a shift from dioxygenation of the aromatic ring to monooxygenation of the methyl group of mono- and dichlorotoluenes. Replacement of Leu272 with tryptophan or phenylalanine resulted in enzymes with improved rates of dichlorotoluene transformation but little or no change in regioselectivity. This residue corresponds to Val260 in NDO, and Asn258 in NBDO and 2-nitrotoluene dioxygenase. Therefore, based on substrates tested to date, the residue at this position plays a role in controlling regioselectivity in NBDO and 2-nitrotoluene dioxygenase, but not NDO or TecA.[91,210,214,222] In a similar study, the α subunit of BPDO from *P. pseudoalcaligenes* KF707 was subjected to site-directed mutagenesis based on modeling with the NDO coordinates. This study demonstrated that Phe227, Ile335, Thr376, and Phe377 were important for determining the position of oxidation with various PCB congeners.[269] Thr376 had been previously identified as playing an important role in determining substrate specificity based on sequence comparisons of BPDOs with different specificities.[153,186] In contrast, changing the corresponding residue in NDO (Thr351) did not seem to affect substrate specificity.[214] Similarly, substitutions in NDO at Phe202 (corresponds to Phe227 in $BPDO_{KF707}$) had little or no effect on substrate specificity. Phe377 corresponds to Phe352 in NDO, however, so the amino acid at this position has now been shown to be critical in TecA, NDO, and $BPDO_{KF707}$. From these studies, it appears that certain residues, such as Phe352 in NDO, may be important in determining the specificity in many

(or all) dioxygenase active sites, while others may play a specific role in just a small subset of enzymes. With more crystal structures becoming available, other key residues effecting substrate catalysis and dioxygenase selectivity may be identified.

4. BIOTECHNOLOGY APPLICATIONS

A number of biotechnology applications have been enabled by the high regioselectivity and enantioselectivity of aromatic hydrocarbon dioxygenases. Several examples include: (1) dioxygenase catalyzed synthesis of a wide range of chiral intermediates, some which have been subsequently used in the preparation of natural products, polyfuntionalized metabolites, or pharmaceutical intermediates, (2) recombinant expression of NDO in a multi-step pathway engineered for the fermentation based production of indigo from glucose, and (3) target-specific agents for biodegradation of environmental pollutants.

4.1. Production of Chiral Synthons

As noted in Section 3.3, several well-characterized aromatic hydrocarbon dioxygenases have been shown to catalyze diverse types of oxidation reactions for a wide range of substrates.[36,125,232] TDOs from *P. putida* F1 and UV4 have extremely broad specificity with respect to the monocyclic, substituted aromatic substrates, which are typically oxidized to arene-*cis*-dihydrodiols of high enantiopurity. Similarly, NDO,[232] BPDO,[36] and carbazole dioxygenase[201] catalyze *cis*-dihydroxyation, as well as other oxidations on a range of polycyclic, heterocyclic, and substituted aromatic compounds. Interested readers are referred to recent reviews detailing the chiral *cis*-dihydrodiols obtained to date, rational approaches for expanding asymmetric (synthetic) methodology, and their versatile application in synthesis of natural products.[34,36,125] The range of available metabolites resulting mainly from *cis*-dihydroxylation have been utilized in numerous syntheses for the preparation of fine chemicals, natural products, pharmaceutical intermediates, and compounds with biological activity (Table 4). It should be noted that many, but not all, metabolites described are amenable to large-scale production using whole cell biocatalysis. This whole-cell or resting-cell biotransformation approach has facilitated the production of multi-kilogram quantities of chiral metabolites and relies on the integrity of multicomponent dioxygenases activity, typically expressed in recombinant hosts with reduced cofactors supplied through the metabolism of exogenous carbon substrates (e.g. glucose). Alternatively, the use of purified oxygenases can be facilitated by the inclusion systems for enzymatic[243] or electrochemical[122,123] regeneration of reduced NAD(P)H cofactor.

Table 4. Examples of synthetic reactions and compounds accessible from *cis*-dihydrodiols.

Compounds that can be synthesized from chiral cis-dihydrodiols[a]
Azasugars, aminosugars, fluorodeoxysugars, perdeuterosugars
Conduritols, conduramines
Deoxysugar analogs
Erythroses
Inositols, inositol oligimers
Kifunensine
Lycoricidine
Narcliclasine
Nojirimycin analogs
Pancreastatin, (+) and (−)-7-deoxypancreastatin
Prostaglandin intermediates ($PGE_{2\alpha}$)
Pinitols
Pyrrolizidine alkaloids
Shikimic acid, methyl shikimate
Specionin
Zeylena
Synthetic reactions conducted utilizing chiral cis-dihydrodiols[a]
Cycloadditions of dienediols and their acetonides
Diels–Alder reactions
Oxidative cleavage-reductive cyclization
Cyclopropanation

[a] See Hudlicky *et al.* [125] for a detailed review of chiral synthons, product structures, and primary literature.

4.2. Indigo Production

The oxidation of indole to indigo was initially shown in *E. coli* strains expressing NDO from *Pseudomonas* sp. NCIB 9816-4.[84] NDO was shown to oxidize indole to an unstable *cis*-dihydroindolediol, which dehydrates to indoxyl, and undergoes spontaneous oxidation to indigo. Since the reaction is catalyzed by many related dioxygenases, this discovery has been widely utilized for detection and isolation of strains expressing mono- and dioxygenases. Commercial interest in the reaction led Genencor International to genetically engineer recombinant *E. coli* for the cost-competitive, multi-step production of indigo from glucose.[28] The fermentation process was based on a single strain expressing a modified tryptophan pathway (allowing high level indole production) and NDO from *P. putida*.[28] Several strategies were employed (e.g. gene dosage, gene amplification, gene inactivation) to improve metabolite flux, enzymatically eliminate the formation of isatin byproduct, and ultimately increase production of indigo to levels exceeding 18 g/l. Despite technical successes, the process for indigo production has not been implemented at an industrial scale.

4.3. Indinavir Production

Interest in enantiopure (−)-*cis*-(1*S*, 2*R*)-indandiol is based on its direct conversion to *cis*-(1*S*)-amino-(2*R*)-indanol, a key intermediate in the chemical synthesis of Merck's HIV-1 protease inhibitor Indinivir Sulfate (Crixivan®).[58] Biotransformations conducted with *P. putida* F39/D (a mutant strain lacking *cis*-dihydrodiol dehydrogenase activity) or a recombinant TDO expressed in *E. coli* indicated that wild-type TDO oxidized indene to (−)-*cis*-(1*S*, 2*R*)-indandiol (∼30% e.e.) and (1*R*)-indenol as main products, with traces of 1-indenone.[228,288] The (−)-*cis*-(1*S*, 2*R*)-indandiol could be obtained in >98% e.e. in late stages of indene conversion with wild-type *P. putida* F1[58] or by co-expression of *cis*-dihydrodiol dehydrogenase (*todD*) with recombinant TDO in *E. coli*.[228] The upgraded *cis*-(1*S*, 2*R*)-indandiol enantiopurity occurs at the expense of total indandiol yield as a result of kinetic resolution catalyzed by the *cis*-dihydrodiol dehydrogenase, which is selective for the unwanted (+)-*cis*-(1*R*, 2*S*)-indandiol.[228] Directed evolution of TDO was conducted to select for reduced levels of the indene by-products, 1-indenol and 1-indenone, while maintaining the highest (−)-*cis*-(1*S*, 2*R*)-indandiol enantiopurity.[303] After three rounds of mutagenesis, variants were obtained that produced significantly more *cis*-indandiol relative to the undesired by-product indenol. However, the stereoselectivity was changed to favor the production of the undesired (+)-*cis*-indandiol enantiomer.[303] These strategies, as well as the application of *Rhodococcus*-derived oxygenases,[205] were unable to alleviate formation of indene by-products and limited the maximum yields to <60% (−)-*cis*-(1*S*, 2*R*)-indandiol. The TDO-catalyzed enantioselective monohydroxylation of 2-indanol to (−)-*cis*-(1*S*, 2*R*)-indandiol in high e.e. represents an alternative route to the vicinal aminoindanol and served as the basis for a process to prepare chiral 1-hydroxy-2-substituted indan intermediates.[32] TDO expressed in *P. putida* strain UV-4[37] and strain *P. putida* F39/D oxidized 2-indenol to (−)-*cis*-(1*S*, 2*R*)-indandiol in >98% e.e. and in >85% yield, while minor products included *trans*-1,2-indandiol (<15%) and 2-hydroxy-1-indanone (<2%).[160]

4.4. Bioremediation of TCE Contamination

Trichloroethylene (TCE) is widely used as a solvent and is classified as an Environmental Protection Agency priority pollutant. It is difficult to degrade and no bacteria are known to utilize it as a carbon and energy source. TDO and other dioxygenases are capable of oxidizing TCE.[148,180,287] TDO converts TCE into non-toxic products: formic acid and glyoxylic acid,[174] but the host strain does not obtain energy from the reaction. TCE is actually capable of inducing the genes encoding TDO,[253] but *P. putida* F1 cells in resting cultures exposed to TCE rapidly lost TCE oxidation ability.[287] A recent study demonstrated that the addition of benzene or toluene restored TCE-degrading activity to *P. putida*

F1,[188] suggesting TCE degradation by this strain could be optimized using this strategy. In another study, a hybrid TDO–BPDO was found to have an enhanced ability to degrade TCE.[97] Characterization of the purified hybrid protein showed that it had higher catalytic efficiency and a lower K_m for TCE than wild-type TDO.[180] A cooxidation strategy for bioremediation of TCE in the field has been reported.[124] In this study, toluene or phenol and oxygen or hydrogen peroxide were added as co-substrates to stimulate *in situ* TCE degradation.

4.5. Improved Biodegradation of PCBs

Genetic engineering has been used to improve a variety of bacterial enzymes and pathways for more efficient degradation of environmental pollutants.[185,211,218,219,278,279] Engineered forms of various dioxygenases have been generated with a variety of methods, and have led to the development of enzymes with novel or enhanced activities, and have provided insights into the control of substrate specificity. Various BPDOs have very different abilities to oxidize PCBs.[105] Using sequence alignments of BPDOs from *B. xenovorans* LB400 and *P. pseudoalcaligenes* KF707 as a guide, Mondello and coworkers generated hybrid and site-directed mutant forms of BPDO. This work resulted in the identification of regions of the protein and specific amino acid residues in the α subunit that controlled substrate specificity.[86,186] These and other studies[153] produced engineered enzymes with the ability to oxidize a wider range of PCB congeners. In particular, this work identified the important role of the residue at position 377 in the *B. xenovorans* LB400 enzyme in the oxidation of 2,5,2′,5′-tetrachlorobiphenyl. Using DNA shuffling,[304] variant forms of BPDO with enhanced abilities to degrade single-ring aromatic hydrocarbons, PCBs, and heterocyclic aromatic compounds have been generated.[16,95,159,267,268] Recently, Barriault *et al*.[18] used random mutagenesis of the *B. xenovorans* LB400 BPDO *bphA* gene to improve enzymatic activity with PCB congeners. Their results showed that amino acid substitutions at Thr335 and Phe336 in BphA caused changes in regiospecificity toward 2,2′-dichlorobiphenyl.

4.6. Construction of New Biodegradation Pathways

Ring hydroxylating dioxygenases have been used to construct new pathways for the degradation of particularly recalcitrant compounds. An attempt was made to assemble a pathway for 2-chlorotoluene degradation by expressing the *todC1C2BA* genes encoding TDO from *P. putida* F1 (to convert 2-chlorotoluene to 2-chlorobenzylalcohol) with the upper TOL pathway genes from pWWO of *P. putida* strain mt-2 encoding benzylalcohol dehydrogenase and benzaldehyde dehydrogenase (to convert 2-chlorobenzylalcohol to 2-chlorobenzoate).[118] These genes were expressed in two different host strains carrying either the *ortho*

or modified *ortho* pathways. Unfortunately, due to the production of dead-end products and unfavorable metabolic flux, the pathway was not functional – even though each individual part of the pathway was functional.[118]

In a more successful pathway construction strategy, expression of genes encoding TDO from *P. putida* F1 and cytochrome $P450_{cam}$ monooxygenase resulted in an engineered *Pseudomonas* strain capable of metabolizing polyhalogenated compounds through sequential reductive and oxidative reactions. In this constructed pathway, cytochrome $P450_{cam}$ catalyzed the conversion of polyhalogenated ethanes, such as pentachloroethane, to TCE under low oxygen tension, and TDO oxidized TCE to glyoxalate and formate.[289]

Many bacterial strains degrade only a limited number of toxic aromatic compounds, and genetic engineering has been used to increase the substrate range of specific organisms. A constructed cassette carrying genes for the conversion of styrene to phenylacetate was introduced into phenylacetate-degrading bacteria to generate strains capable of using styrene as the sole source of carbon and energy.[179] When the plasmid-borne styrene cassette was introduced into *P. putida* F1 carrying the TOL plasmid, the new strain was capable of growth on an increased number of aromatic hydrocarbons, including benzene, toluene, ethylbenzene, *m*-xylene, *p*-xylene, and styrene. To reduce the possibility of horizontal transfer of genetically engineered DNA among bacteria, the cassette was inserted into a minitransposon carrying an engineered gene containment system.[72]

The *tod* genes encoding TDO have also been cloned into the chromosome of *Deinococcus radiodurans*, a bacterium that is highly resistant to radiation. This new strain was capable of degrading toluene and related aromatic hydrocarbons in the presence of high levels of radiation.[162] Expression of the mercury resistance gene (*merA*) together with the toluene dioxygenase genes in *D. radiodurans* resulted the integration of multiple remediation functions in a single engineered strain, which could be used for amelioration of mixed radioactive waste containing aromatic hydrocarbon pollutants and the heavy metal mercury.[45]

5. CONCLUSIONS

Aromatic ring hydroxylating dioxygenases play a key role in the biodegradation of numerous environmental pollutants, both in the natural environment (via natural attenuation) and in the engineered bioremediation systems. Recent structural and mechanistic information, together with enzyme engineering and strain construction strategies should allow the development of engineered microorganisms with new and/or optimized degradation abilities. The continued application of these approaches should also facilitate the development of Rieske non-heme iron dioxygenases with requisite selectivities for specific opportunities in target direct biocatalysis or metabolic engineering.

ACKNOWLEDGMENTS

Research in the author's laboratory is supported by the Army Research Office, the Strategic Environmental Research and Development Program, and the University of California Toxic Substances Research and Teaching Program (REP). We thank Rosmarie Friemann and Juan Parales for providing figures.

REFERENCES

1. Akhtar, M.N., Boyd, D.R., Thompson, N.J., Koreeda, M., Gibson, D.T., Mahadevan, V., and Jerina, D.M., 1975, Absolute stereochemistry of the dihydroanthracene-*cis*- and -*trans*-1,2-diols produced from anthracene by mammals and bacteria. *J. Chem. Soc. Perkin Trans.*, 1:2506–2511.
2. Allen, C.C.R., Boyd, D.R., Dalton, H., Sharma, N.D., Haughey, S.A., McMordie, R.A.S., McMurray, B.T., Sheldrake, G.N., and Sproule, K., 1995, Sulfoxides of high enantiopurity from bacterial dioxygenase-catalyzed oxidation. *J. Chem. Soc. Chem. Commun.*, 119–120.
3. Allen, C.C.R., Boyd, D.R., Larkin, M.J., Reid, K.A., Sharma, N.D., and Wilson, K., 1997, Metabolism of naphthalene, 1-naphthol, indene and indole by *Rhodococcus* sp. strain NCIMB12038. *Appl. Environ. Microbiol*., 63:151–155.
4. An, D., Gibson, D.T., and Spain, J.C., 1994, Oxidative release of nitrite from 2-nitrotoluene by a three-component enzyme system from *Pseudomonas* sp. strain JS42. *J. Bacteriol*., 176:7462–7467.
5. Andreoni, V., Bernasconi, S., Colombo, M., van Beilen, J.B., and Cavalca, L., 2000, Detection of genes for alkane and naphthalene catabolism in *Rhodococcus* sp. strain 1BN. *Environ. Microbiol*., 2:572–577.
6. Aoki, H., Kimura, T., Habe, H., Yamane, H., Kodama, T., and Omori, T., 1996, Cloning, nucleotide sequence, and characterization of the genes encoding enzymes involved in the degradation of cumene to 2-hydroxy-6-oxo-7-methylocta-2,4-dienoic acid in *Pseudomonas fluorescens* IP01. *J. Ferment. Bioeng*., 81:187–196.
7. Arai, H., Kosono, S., Taguchi, K., Maeda, M., Song, E., Fuji, F., Chung, S.-Y., and Kudo, T., 1998, Two sets of biphenyl and PCB degradation genes on a linear plasmid in *Rhodococcus erythropolis* A421. *J. Ferment. Bioeng*., 86:595–599.
8. Armengaud, J., Gaillard, J., and Timmis, K.N., 2000, A second [2Fe-2S] ferredoxin from *Sphingomonas* sp. strain RW1 can function as an electron donor for the dioxin dioxygenase. *J. Bacteriol*., 182:2238–2244.
9. Armengaud, J., Happe, B., and Timmis, K.N., 1998, Genetic analysis of dioxin dioxygenase of *Sphingomonas* sp. strain RW1: catabolic genes dispersed on the genome. *J. Bacteriol*., 180:3954–3966.
10. Armengaud, J., and Timmis, K.N., 1997, Molecular characterization of Fdx1, a putidaredoxin-type [2Fe-2S] ferredoxin able to transfer electrons to the dioxin dioxygenase of *Sphingomonas* sp. RW1. *Eur. J. Biochem*., 247:833–842.
11. Armengaud, J., and Timmis, K.N., 1998, The reductase RedA2 of the multi-component dioxin dioxygenase system of *Sphingomonas* sp. RW1 is related to class-I cytochrome P450-type reductases. *Eur. J. Biochem*., 253:437–444.
12. Asturias, J.A., Díaz, E., and Timmis, K.N., 1995, The evolutionary relationship of biphenyl dioxygenase from Gram-positive *Rhodococcus globerulus* P6 to multicomponent dioxygenases from Gram-negative bacteria. *Gene*, 156:11–18.

13. Axcell, B.C., and Geary, P.J., 1973, The metabolism of benzene by bacteria- purification and some properties of the enzyme *cis*-1,2-dihydroxycyclohexa-3,5-diene (nicotinamide adenine dinucleotide) oxidoreductase (*cis*-benzene glycol dehydrogenase). *Biochem. J*., 136:927–934.
14. Baggi, G., Barbieri, P., Galli, E., and Tollari, S., 1987, Isolation of a *Pseudomonas stutzeri* strain that degrades *o*-xylene. *Appl. Environ. Microbiol*., 64:2473–2478.
15. Barriault, D., Pelletier, C., Hurtubise, Y., and Sylvestre, M., 1997, Substrate selectivity pattern of *Comamonas testosteroni* strain B-356 towards dichlorobiphenyls. *Int. Biodeterior. Biodegradation*, 39:311–316.
16. Barriault, D., Plante, M.-M., and Sylvestre, M., 2002, Family shuffling of a targeted *bphA* region to engineer biphenyl dioxygenase. *J. Bacteriol*., 184:3794–3800.
17. Barriault, D., Simard, C., Chatel, H., and Sylvestre, M., 2001, Characterization of hybrid biphenyl dioxygenases obtained by recombining *Burkholderia* sp. strain LB400 *bphA* with the homologous gene of *Comamonas testosteroni* B-356. *Can. J. Microbiol*., 47:1025–1032.
18. Barriault, D., and Sylvestre, M., 2004, Evolution of the biphenyl dioxygenase BphA from *Burkholderia xenovorans* LB400 by random mutagenesis of multiple sites in region III. *J. Biol. Chem*., 279:47480–47488.
19. Batie, C.J., and Ballou, D.P., 1990, Phthalate dioxygenase. *Methods Enzymol*., 188:61–70.
20. Batie, C.J., Ballou, D.P., and Corell, C.C., 1991, Phthalate dioxygenase reductase and related flavin-iron-sulfur containing electron transferases, pp. 543–556. In F. Müller (ed.), *Chemistry and Biochemistry of Flavoenzymes*. CRC Press, Boca Raton, FL.
21. Bedard, D.L., and Haberl, M.L., 1990, Influence of chlorine substitution pattern on the degradation of polychlorinated biphenyls by eight bacterial strains. *Microb. Ecol*., 20:87–102.
22. Beharry, Z.M., Eby, D.M., Coulter, E.D., Viswanathan, R., Neidle, E.L., Phillips, R.S., and Kurtz, D.M.J., 2003, Histidine ligand protonation and redox potential in the Rieske dioxygenases: role of a conserved aspartate in anthranilate 1,2-dioxygenase. *Biochemistry*, 42:13625–13636.
23. Beil, S., Happe, B., Timmis, K.N., and Pieper, D.H., 1997, Genetic and biochemical characterization of the broad spectrum chlorobenzene dioxygenase from *Burkholderia* sp. strain PS12–dechlorination of 1,2,4,5-tetrachlorobenzene. *Eur. J. Biochem*., 247:190–199.
24. Beil, S., Mason, J.R., Timmis, K.N., and Pieper, D.H., 1998, Identification of chlorobenzene dioxygenase sequence elements involved in dechlorination of 1,2,4,5-tetrachlorobenzene. *J. Bacteriol*., 180:5520–5528.
25. Bernhardt, F.-H., Bill, E., Trautwein, A.X., and Twilfer, H., 1988, 4-Methoxybenzoate monooxygenase from *Pseudomonas putida*: isolation, biochemical properties, substrate specificity, and reaction mechanisms of the enzyme components. *Methods Enzymol*., 161:281–294.
26. Bernhardt, F.-H., Erdin, N., Staudinger, H., and Ullrich, V., 1973, Interactions of substrates with a purified 4-methoxybenzoate monooxygenase system (*O*-demethylating) from *Pseudomonas putida*. *Eur. J. Biochem*., 35:126–134.
27. Bernhardt, F.-H., and Kuthan, H., 1981, Dioxygen activation by putidamonooxin: the oxygen species formed and released under uncoupling conditions. *Eur. J. Biochem*., 120:547–555.
28. Berry, A., Dodge, T.C., Pepsin, M., and Weyler, W., 2002, Application of metabolic engineering to improve both the production and use of biotech indigo. *J. Ind. Microbiol. Biotechnol*., 28:127–133.
29. Bertoni, G., Bolognese, F., Galli, E., and Barbieri, P., 1996, Cloning of the genes for and characterization of the early stages of toluene and *o*-xylene catabolism in *Pseudomonas stutzeri* OX1. *Appl. Environ. Microbiol*., 62:3704–3711.
30. Bestetti, G., Bianchi, D., Bosetti, A., Di Gennaro, P., Galli, E., Leoni, B., Pelizzoni, F., and Sello, G., 1995, Bioconversion of substituted naphthalenes to the corresponding 1,2-dihydro-1,2-dihydroxy derivatives. Determination of the regio- and stereochemistry of the oxidation reactions. *Appl. Microbiol. Biotechnol*., 44:306–313.

31. Bickerdike, S.R., Holt, R.A., and Stephens, G.M., 1997, Evidence for metabolism of *o*-xylene by simultaneous ring and methyl group oxidation in a new soil isolate. *Microbiology*, 143:2321–2329.
32. Blacker, A.J., Boyd, D.R., Dalton, H., and Bowers, N., May 20, 1996, Preparation of hydroxyl compounds by conversion with dioxygenase patent WO 96/37628.
33. Bosch, R., Garcia-Valdes, E., and Moore, E.R.B., 2000, Complete nucleotide sequence and evolutionary significance of a chromosomally encoded naphthalene-degradation lower pathway from *Pseudomonas stutzeri* AN10. *Gene*, 245:65–74.
34. Boyd, D.R., and Bugg, T.D.H., 2006, Arene *cis*-dihydrodiol formation: from biology to application. *Org. Biomol. Chem.* In Press (D01: 10.1039/b513226f).
35. Boyd, D.R., Sharma, N.D., Agarwal, R., Resnick, S.M., Schocken, M.J., Gibson, D.T., Sayer, J.M., Yagi, H., and Jerina, D.M., 1997, Bacterial dioxygenase-catalyzed dihydroxylation and chemical resolution routes to enantiopure *cis*-dihydrodiols of chrysene. *J. Chem. Soc. Perkin Trans*., 1:1715–1723.
36. Boyd, D.R., Sharma, N.D., and Allen, C.C.R., 2001, Aromatic dioxygenases: molecular biocatalysis and applications. *Curr. Opin. Biotechnol*., 12:564–573.
37. Boyd, D.R., Sharma, N.D., Bowers, N.I., Goodrich, P.A., Groocock, M.R., Blaker, A.J., Clarke, D.A., Howard, T., and Dalton, H., 1996, Stereoselective dioxygenase-catalyzed benzylic hydroxylation at prochiral methylene groups in the chemoenzymatic synthesis of enantiopure vicinal aminoindanols. *Tetrahedron Asymmetry*, 7:1559–1562.
38. Boyd, D.R., Sharma, N.D., Dorrity, M.R.J., Hand, M.V., McMordie, R.A.S., Malone, J.F., Porter, H.P., Chima, J., Dalton, H., and Sheldrake, G.N., 1993, Structure and stereochemistry of *cis*-dihydrodiol and phenol metabolites of bicyclic azaarenes from *Pseudomonas putida* UV4. *J. Chem. Soc. Perkin Trans*., 1:1065–1071.
39. Boyd, D.R., Sharma, N.D., Hempenstall, F., Kennedy, M.A., Malone, J.F., Allen, C.C.R., Resnick, S.M., and Gibson, D.T., 1999, *bis-cis*-Dihydrodiols: a new class of metabolites resulting from biphenyl dioxygenase-catalyzed sequential asymmetric *cis*-dihydroxylation of polycyclic arenes and heteroarenes. *J. Org. Chem*., 64:4005–4011.
40. Boyd, D.R., Sharma, N.D., Kerley, N.A., McMordie, R.A.S., Sheldrake, G.N., Williams, P., and Dalton, H., 1996, Dioxygenase-catalyzed oxidation of dihydronaphthalenes to yield arene hydrate and *cis*-dihydronaphthalenediols. *J. Chem. Soc. Perkin Trans*., 1:67–74.
41. Boyd, D.R., Sharma, N.D., Modyanova, L.V., Carroll, J.G., Malone, J.F., Allen, C.C.R., Hamilton, J.T.G., Gibson, D.T., Parales, R.E., and Dalton, H., 2002, Dioxygenase-catalyzed *cis*-dihydroxylation of pyridine-ring systems. *Can. J. Chem*., 80:589–600.
42. Boyd, D.R., and Sheldrake, G.N., 1998, The dioxygenase-catalysed formation of vicinal *cis*-diols. *Nat. Prod. Rep*., 15:309–324.
43. Brand, J.M., Cruden, D.L., Zylstra, G.J., and Gibson, D.T., 1992, Stereospecific hydroxylation of indan by *E. coli* containing the cloned toluene dioxygenase genes from *Pseudomonas putida* F1. *Appl. Environ. Microbiol*., 58:3407–3409.
44. Brilon, C., Beckmann, W., and Knackmuss, H.-J., 1981, Catabolism of naphthalenesulfonic acids by *Pseudomonas* sp. A3 and *Pseudomonas* sp. C22. *Appl. Environ. Microbiol*., 42:44–55.
45. Brim, H., McFarlan, S.C., Fredrickson, J.K., Minton, K.W., Zhai, M., Wackett, L.P., and Daly, M.J., 2000, Engineering *Deinococcus radiodurans* for metal remediation in radioactive mixed waste environments. *Nat. Biotechnol*., 18:85–90.
46. Brunel, F., and Davison, J., 1988, Cloning and sequencing of *Pseudomonas* genes encoding vanillate demethylase. *J. Bacteriol*., 170:4924–4930.
47. Bundy, B.M., Campbell, A.L., and Neidle, E.L., 1998, Similarities between the *antABC*-encoded anthranilate dioxygenase and the *benABC*-encoded benzoate dioxygenase of *Acinetobacter* sp. strain ADP1. *J. Bacteriol*., 180:4466–4474.

48. Bünz, P.V., and Cook, A.M., 1993, Dibenzofuran 4,4a-dioxygenase from *Sphingomonas* sp. strain RW1: angular dioxygenation by a three-component enzyme system. *J. Bacteriol*., 175:6467–6475.
49. Butler, C.S., and Mason, J.R., 1997, Structure-function analysis of the bacterial aromatic ring-hydroxylating dioxygenases. *Adv. Microb. Physiol*., 38:47–84.
50. Cane, P.A., and Williams, P.A., 1986, A restriction map of naphthalene catabolic plasmid pWW60-1 and the location of some of its catabolic genes. *J. Gen. Microbiol*., 132:2919–2929.
51. Carredano, E., Karlsson, A., Kauppi, B., Choudhury, D., Parales, R.E., Parales, J.V., Lee, K., Gibson, D.T., Eklund, H., and Ramaswamy, S., 2000, Substrate binding site of naphthalene 1,2-dioxygenase: functional implications of indole binding. *J. Mol. Biol*., 296:701–712.
52. Chang, H.K., Mohseni, P., and Zylstra, G.J., 2003, Characterization and regulation of the genes for a novel anthranilate 1,2-dioxygenase from *Burkholderia cepacia* DBO1. *J. Bacteriol*., 185:5871–5881.
53. Chang, H.-K., and Zylstra, G.J., 1998, Novel organization of the genes for phthalate degradation from *Burkholderia cepacia* DBO1. *J. Bacteriol*., 180:6529–6537.
54. Chebrou, H., Hurtubise, Y., Barriault, D., and Sylvestre, M., 1999, Catalytic activity toward chlorobiphenyls of purified recombinant His-tagged oxygenase component of *Rhodococcus globerulous* strain P6 biphenyl dioxygenase and of chimeras derived from it and their expression in *Escherichia coli* and in *Pseudomonas putida*. *J. Bacteriol*., 181:4805–4811.
55. Cho, O., Choi, K.Y., Zylstra, G.J., Kim, Y.-S., Kim, S.-K., Lee, J.H., Sohn, H.-Y., Kwon, G.-S., Kim, Y.M., and Kim, E., 2005, Catabolic role of a three-component salicylate oxygenase from *Sphingomonas yanoikuyae* B1 in polycyclic aromatic hydrocarbon degradation. *Biochem. Biophys. Res. Commun*., 327:656–662.
56. Civolani, C., Barghini, P., Roncetti, A.R., Ruzzi, M., and Schiesser, A., 2000, Bioconversion of ferulic acid into vanillic acid by means of a vanillate-negative mutant of *Pseudomonas fluorescens* strain BF13. *Appl. Environ. Microbiol*., 66:2311–2317.
57. Colbert, C.L., Couture, M.M.-J., Eltis, L.D., and Bolin, J.T., 2000, A cluster exposed: structure of the Rieske ferredoxin from biphenyl dioxygenase and the redox properties of Rieske Fe-S proteins. *Structure*, 8:1267–1278.
58. Connors, N., Prevoznak, R., Chartrain, M., Reddy, J., Singhvi, R., Patel, Z., Olewinski, R., Salmon, P., Wilson, J., and Greasham, R., 1997, Conversion of indene to *cis*-(1*S*,2*R*)-indandiol by mutants of *Pseudomonas putida* F1. *J. Ind. Microbiol. Biotechnol*., 18:353–359.
59. Corkery, D.M., and Dobson, A.D., 1998, Reverse transcription-PCR analysis of the regulation of ethylbenzene dioxygenase gene expression in *Pseudomonas fluorescens* CA-4. *FEMS Microbiol. Lett*., 166:171–176.
60. Corkery, D.M., O'Connor, K.E., Buckley, C.M., and Dobson, A.D., 1994, Ethylbenzene degradation by *Pseudomonas fluorescens* strain CA-4. *FEMS Microbiol. Lett*., 124:23–27.
61. Correll, C.C., Batie, C.J., Ballou, D.P., and Ludwig, M.L., 1992, Phthalate dioxygenase reductase: a modular structure for electron transfer from pyridine nucleotides to [2Fe-2S]. *Science*, 258:1604–1610.
62. Dabrock, B., Keßeler, M., Averhoff, B., and Gottschalk, G., 1994, Identification and characterization of a transmissible linear plasmid from *Rhodococcus erythropolis* BD2 that encodes isopropylbenzene and trichloroethene catabolism. *Appl. Environ. Microbiol*., 60:853–860.
63. Dabrock, B., Riedel, J., Bertram, J., and Gottschalk, G., 1992, Isopropylbenzene (cumene)–a new substrate for the isolation of trichloroethene-degrading bacteria. *Arch. Microbiol*., 158:9–13.
64. Danganan, C.E., Ye, R.W., Daubaras, D.L., Xun, L., and Chakrabarty, A.M., 1994, Nucleotide sequence and functional analysis of the genes encoding 2,4,5-trichlorophenoxyacetic

acid oxygenase in *Pseudomonas cepacia* AC1100. *Appl. Environ. Microbiol*., 60:4100–4106.

65. de Marco, P., Moradas-Ferreira, P., Higgins, T.P., McDonald, I., Kenna, E.M., and Murrell, J.C., 1999, Molecular analysis of a novel methanesulfonic acid monooxygenase from the methylotroph *Methylosulfonomonas methylovora*. *J. Bacteriol*., 181:2244–2251.
66. Dehmel, U., Engesser, K.-H., Timmis, K.N., and Dwyer, D.F., 1995, Cloning, nucleotide sequence, and expression of the gene encoding a novel dioxygenase involved in metabolism of carboxydiphenyl ethers in *Pseudomonas pseudoalcaligenes* POB310. *Arch. Microbiol*., 163:35–41.
67. Deluca, M.E., and Hudlicky, T., 1990, Microbial oxidation of naphthalene derivatives. Absolute configuration of metabolites. *Tetrahedron Lett*., 31:13–16.
68. Demaneche, S., Meyer, C., Micoud, J., Louwagie, M., Willison, J.C., and Jouanneau, Y., 2004, Identification and functional analysis of two aromatic ring-hydroxylating dioxygenases from a *Sphingomonas* strain that degrades various polycyclic aromatic hydrocarbons. *Appl. Environ. Microbiol*., 70:6714–6725.
69. Denome, S.A., Stanley, D.C., Olsen, E.S., and Young, K.D., 1993, Metabolism of dibenzothiophene and naphthalene in *Pseudomonas* strains: complete DNA sequence of an upper naphthalene catabolic pathway. *J. Bacteriol*., 175:6890–6901.
70. Di Gennaro, P., Rescalli, E., Galli, E., Sello, G., and Bestetti, G., 2001, Characterization of *Rhodococcus opacus* R7, a strain able to degrade naphthalene and *o*-xylene isolated from a polycyclic aromatic hydrocarbon-contaminated soil. *Res. Microbiol*., 152:641–651.
71. Díaz, E., Ferrández, A., and García, J.L., 1998, Characterization of the *hca* cluster encoding the dioxygenolytic pathway for initial catabolism of 3-phenylpropionic acid in *Escherichia coli* K-12. *J. Bacteriol*., 180:2915–2923.
72. Díaz, E., Munthali, M., de Lorenzo, V., and Timmis, K.N., 1994, Universal barrier to lateral spread of specific genes among microorganisms. *Mol. Microbiol*., 13:855–861.
73. Dionisi, H.M., Chewning, C.S., Morgan, K.H., Menn, F.-M., Easter, J.P., and Sayler, G.S., 2004, Abundance of dioxygenase genes similar to *Ralstonia* sp. strain U2 *nagAc* is correlated with naphthalene concentrations in coal tar-contaminanted freshwater sediments. *Appl. Environ. Microbiol*., 70:3988–3995.
74. Dong, X., Fushinobu, S., Fukuda, E., Terada, T., Nakamura, S., Shimizu, K., Nojiri, H., Omori, T., Shoun, H., and Wakagi, T., 2005, Crystal structure of the terminal oxygenase component of cumene dioxygenase from *Pseudomonas fluorescen*s IP01. *J. Bacteriol*., 187:2483–2490.
75. Eaton, R.W., 1996, *p*-Cumate catabolic pathway in *Pseudomonas putida* F1: cloning and characterization of DNA carrying the *cmt* operon. *J. Bacteriol*., 178:1351–1362.
76. Eaton, R.W., 2001, Plasmid-encoded phthalate catabolic pathway in *Arthrobacter keyseri* 12B. *J. Bacteriol*., 183:3689–3703.
77. Eaton, R.W., and Timmis, K.N., 1986, Characterization of a plasmid-specified pathway for catabolism of isopropylbenzene in *Pseudomonas putida* RE204. *J. Bacteriol*., 168:123–131.
78. Eaton, S.E., Resnick, S.M., and Gibson, D.T., 1996, Initial reactions in the oxidation of 1,2-dihydronaphthalene metabolism by *Sphingomonas yanoikuyae* strains. *Appl. Environ. Microbiol*., 62:4388–4394.
79. Eby, D.M., Beharry, Z.M., Coulter, E.D., Kurtz, D.M.J., and Neidle, E.L., 2001, Characterization and evolution of anthranilate 1,2-dioxygenase from *Acinetobacter* sp. strain ADP1. *J. Bacteriol*., 183:109–118.
80. Engesser, K.H., Fietz, W., Fischer, P., Schulte, P., and Knackmuss, H.-J., 1990, Dioxygenolytic cleavage of aryl ether bonds: 1,2-dihydro-1,2-dihydroxy-4-carboxybenzophenone as evidence for initial 1,2-dioxygenation in 3- and 4-carboxy biphenyl ether degradation. *FEMS Microbiol. Lett*., 69:317–322.
81. Engesser, K.H., Strubel, V., Christoglou, K., Fischer, P., and Rast, H.G., 1989, Dioxygenolytic cleavage of aryl ether bonds: 1, 10-dihydro-1,10-dihydroxyfluoren-9-one, a novel arene

dihydrodiol as evidence for angular dioxygenation of dibenzofuran. *FEMS Microbiol. Lett*., 65:205–210.
82. Ensley, B.D., and Gibson, D.T., 1983, Naphthalene dioxygenase: purification and properties of a terminal oxygenase component. *J. Bacteriol*., 155:505–511.
83. Ensley, B.D., Gibson, D.T., and Laborde, A.L., 1982, Oxidation of naphthalene by a multicomponent enzyme system from *Pseudomonas* sp. strain NCIB 9816. *J. Bacteriol*., 149:948–954.
84. Ensley, B.D., Ratzkin, B.J., Osslund, T.D., Simon, M.J., Wackett, L.P., and Gibson, D.T., 1983, Expression of naphthalene oxidation genes in *Escherichia coli* results in the biosynthesis of indigo. *Science*, 222:167–169.
85. Erickson, B.D., and Mondello, F.J., 1992, Nucleotide sequencing and transcriptional mapping of the genes encoding biphenyl dioxygenase, a multicomponent polychlorinated biphenyl-degrading enzyme in *Pseudomonas* strain LB400. *J. Bacteriol*., 174:2903–2912.
86. Erickson, B.D., and Mondello, F.J., 1993, Enhanced biodegradation of polychlorinated biphenyls after site-directed mutagenesis of a biphenyl dioxygenase gene. *Appl. Environ. Microbiol*., 59:3858–3862.
87. Ferrero, M., Llobet-Brossa, E., Lalucat, J., Garciá-Valdez, E., Rosselló-Mora, R., and Bosch, R., 2002, Coexistence of two distinct copies of naphthalene degradation genes in *Pseudomonas* strains isolated from the western Mediterranean region. *Appl. Environ. Microbiol*., 68:957–962.
88. Fetzner, S., Müller, R., and Lingens, F., 1992, Purification and some properties of 2-halobenzoate 1,2-dioxygenase, a two-component enzyme system from *Pseudomonas cepacia* 2CBS. *J. Bacteriol*., 174:279–290.
89. Filonov, A.E., Karpov, A.V., Kosheleva, I.A., Puntus, I.F., Balashova, N.V., and Boronin, A.M., 2000, The efficiency of salicylate utilization by *Pseudomonas putida* strains catabolizing naphthalene via different biochemical pathways. *Process Biochem*., 35:983–987.
90. Francisco, P.J., Ogawa, N., Suzuki, K., and Miyashita, K., 2001, The chlorobenzoate dioxygenase genes of *Burkholderia* sp. strain NK8 involved in the catabolism of chlorobenzoates. *Microbiology*, 147:121–133.
91. Friemann, R., Ivkovic-Jensen, M.M., Lessner, D.J., Yu, C.-L., Gibson, D.T., Parales, R.E., Eklund, H., and Ramaswamy, S., 2005, Structural insights into the dioxygenation of nitroarene compounds: the crystal structure of the nitrobenzene dioxygenase. *J. Mol. Biol*., 348:1139–1151.
92. Fuenmayor, S.L., Wild, M., Boyles, A.L., and Williams, P.A., 1998, A gene cluster encoding steps in the conversion of naphthalene to gentisate in *Pseudomonas* sp. strain U2. *J. Bacteriol*., 180:2522–2530.
93. Fujii, T., Takeo, M., and Maeda, Y., 1997, Plasmid-encoded genes specifying aniline oxidation from *Acinetobacter* sp. strain YAA. *Microbiology*, 143:93–99.
94. Fukumori, F., and Saint, C., 1997, Nucleotide sequences and regulational analysis of genes involved in conversion of aniline to catechol in *Pseudomonas putida* UCC22(pTDN1). *J. Bacteriol*., 179:399–408.
95. Furukawa, K., 2000, Engineering dioxygenases for efficient degradation of environmental pollutants. *Curr. Opin. Biotechnol*., 11:244–249.
96. Furukawa, K., Arimura, N., and Miyazaki, T., 1987, Nucleotide sequence of the 2,3-dihydroxybiphenyl dioxygenase gene of *Pseudomonas pseudoalcaligenes*. *J. Bacteriol*., 169:427–429.
97. Furukawa, K., Hirose, J., Hayashida, S., and Nakamura, K., 1994, Efficient degradation of trichloroethylene by a hybrid aromatic ring dioxygenase. *J. Bacteriol*., 176:2121–2123.
98. Furusawa, Y., Nagarajan, V., Tanokura, M., Masai, E., Fukuda, M., and Senda, T., 2004, Crystal structure of the terminal oxygenase of biphenyl dioxygenase from *Rhodococcus* sp. strain RHA1. *J. Mol. Biol*., 342:1041–1052.

99. Gassner, G., Wang, L., Batie, C., and Ballou, D.P., 1994, Reaction of phthalate dioxygenase reductase with NADH and NAD: kinetic and spectral characterization of intermediates. *Biochemistry*, 33:12184–12193.
100. Gassner, G.T., and Ballou, D.P., 1995, Preparation and characterization of a truncated form of phthalate dioxygenase reductase that lacks an iron–sulfur domain. *Biochemistry*, 34:13460–13471.
101. Gassner, G.T., Ballou, D.P., Landrum, G.A., and Whittaker, J.W., 1993, Magnetic circular dichroism studies on the mononuclear ferrous active site of phthalate dioxygenase from *Pseudomonas cepacia* show a change of ligation state on substrate binding. *Biochemistry*, 32:4820–4825.
102. Gassner, G.T., Ludwig, M.L., Gatti, D.L., Correll, C.C., and Ballou, D.P., 1995, Structure and mechanism of the iron–sulfur flavoprotein phthalate dioxygenase reductase. *FASEB J*., 9:1411–1418.
103. Geary, P.J., Mason, J.R., and Joannou, C.L., 1990, Benzene dioxygenase from *Pseudomonas putida* ML2 (NCIB 12190). *Methods Enzymol*., 188:52–60.
104. Gibson, D.T., 1999, *Beijerinckia* sp. strain B1: a strain by any other name. *J. Ind. Microbiol. Biotechnol*., 23:284–293.
105. Gibson, D.T., Cruden, D.L., Haddock, J.D., Zylstra, G.J., and Brand, J.M., 1993, Oxidation of polychlorinated biphenyls by *Pseudomonas* sp. strain LB400 and *Pseudomonas pseudoalcaligenes* KF707. *J. Bacteriol*., 175:4561–4564.
106. Gibson, D.T., and Parales, R.E., 2000, Aromatic hydrocarbon dioxygenases in environmental biotechnology. *Curr. Opin. Biotechnol*., 11:236–243.
107. Gibson, D.T., Resnick, S.M., Lee, K., Brand, J.M., Torok, D.S., Wackett, L.P., Schocken, M.J., and Haigler, B.E., 1995, Desaturation, dioxygenation and monooxygenation reactions catalyzed by naphthalene dioxygenase from *Pseudomonas* sp. strain 9816-4. *J. Bacteriol*., 177:2615–2621.
108. Grifoll, M., Selifonov, S.A., Gatlin, C.V., and Chapman, P.J., 1995, Actions of a versatile fluorene-degrading bacterial isolate on polycyclic aromatic compounds. *Appl. Environ. Microbiol*., 61:3711–3723.
109. Haak, B., Fetzner, S., and Lingens, F., 1995, Cloning, nucleotide sequence, and expression of the plasmid-encoded genes for the two-component 2-halobenzoate 1,2-dioxygenase from *Pseudomonas cepacia* 2CBS. *J. Bacteriol*., 177:667–675.
110. Habe, H., Ashikawa, Y., Saiki, Y., Yoshida, T., Nojiri, H., and Omori, T., 2002, *Sphingomonas* sp. strain KA1, carrying a carbazole dioxygenase gene homologue, degrades chlorinated dibenzo-*p*-dioxins in soil. *FEMS Microbiol. Lett*., 211:43–49.
111. Habe, H., Miyakoshi, M., Chung, J., Kasuga, K., Yoshida, T., Nojiri, H., and Omori, T., 2003, Phthalate catabolic gene cluster is linked to the angular dioxygenase gene in *Terrabacter* sp. strain DBF63. *Appl. Microbiol. Biotechnol*., 61:44–54.
112. Haddad, S., Eby, D.M., and Neidle, E.L., 2001, Cloning and expression of the benzoate dioxygenase genes from *Rhodococcus* sp. strain 19070. *Appl. Environ. Microbiol*., 67:2507–2514.
113. Haddock, J.D., Horton, J.R., and Gibson, D.T., 1995, Dihydroxylation and dechlorination of chlorinated biphenyls by purified biphenyl 2,3-dioxygenase from *Pseudomonas* sp. strain LB400. *J. Bacteriol*., 177:20–26.
114. Haigler, B.E., and Gibson, D.T., 1990, Purification and properties of ferredoxin$_{NAP}$, a component of naphthalene dioxygenase from *Pseudomonas* sp. strain NCIB 9816. *J. Bacteriol*., 172:465–468.
115. Haigler, B.E., and Gibson, D.T., 1990, Purification and properties of NADH-ferredoxin$_{NAP}$ reductase, a component of naphthalene dioxygenase from *Pseudomonas* sp. strain NCIB 9816. *J. Bacteriol*., 172:457–464.

116. Harayama, S., Rekik, M., Bairoch, A., Neidle, E.L., and Ornston, L.N., 1991, Potential DNA slippage structure acquired during evolutionary divergence of *Acinetobacter calcoaceticus* chromosomal *benABC* and *Pseudomonas putida* TOL pWWO plasmid *xylXYZ*, genes encoding benzoate dioxygenases. *J. Bacteriol*., 173:7540–7548.
117. Harayama, S., Rekik, M., and Timmis, K.N., 1986, Genetic analysis of a relaxed substrate specificity aromatic ring dioxygenase, toluate 1,2-dioxygenase, encoded by TOL plasmid pWWO of *Pseudomonas putida*. *Mol. Gen. Genet*., 202:226–234.
118. Haro, M.A., and de Lorenzo, V., 2001, Metabolic engineering of bacteria for environmental applications: construction of *Pseudomonas* strains for biodegradation of 2-chlorotoluene. *J. Biotechnol*., 85:103–113.
119. Hayase, N., Taira, K., and Furukawa, K., 1990, *Pseudomonas putida* KF715 *bphABCD* operon encoding biphenyl and polychlorinated biphenyl degradation: Cloning, analysis, and expression in soil bacteria. *J. Bacteriol*., 172:1160–1164.
120. Hickey, W.J., and Sabat, G., 2001, Integration of matrix-assisted laser desorption ionization-time of flight mass spectrometry and molecular cloning for the identification and functional characterization of mobile *ortho*-halobenzoate oxygenase genes in *Pseudomonas aeruginosa* strain JB2. *Appl. Environ. Microbiol*., 67:5648–5655.
121. Hirose, J., Suyama, A., Hayashida, S., and Furukawa, K., 1994, Construction of hybrid biphenyl (*bph*) and toluene (*tod*) genes for functional analysis of aromatic ring dioxygenases. *Gene*, 138:27–33.
122. Hollmann, F., Lin, P.-C., Witholt, B., and Schmid, A., 2003, Stereospecific biocatalytic epoxidation: the first example of direct regeneration of a FAD-dependent monooxygenase for catalysis. *J. Am. Chem. Soc*., 125:8209–8217.
123. Hollmann, F., Schmid, A., and Steckhan, E., 2001, First synthetic application of a monooxygenase employing indirect electrochemical NADH regeneration. *Angew. Chem. Int. Ed*., 40:169–171.
124. Hopkins, G.D., and McCarty, P.L., 1995, Field evaluation of in-situ aerobic cometabolism of trichloroethene and three dichloroethene isomers using phenol and toluene as the primary substrates. *Environ. Sci. Technol*., 29:1628–1637.
125. Hudlicky, T., Gonzalez, D., and Gibson, D.T., 1999, Enzymatic dihydroxylation of aromatics in enantioselective synthesis: expanding asymmetric methodology. *Aldrichimica Acta*, 32:35–62.
126. Hudlicky, T., Stabile, M.R., Gibson, D.T., and Whited, G.M., 1999, 1-Chloro-(2*S*,3*S*)-dihydroxycyclohexa-4,6-diene. *Org. Syntheses*, 76:77–85.
127. Hurtubise, Y., Barriault, D., and Sylvestre, M., 1996, Characterization of active recombinant His-tagged oxygenase component of *Comamonas testosteroni* B-356 biphenyl dioxygenase. *J. Biol. Chem*., 271:8152–8156.
128. Hurtubise, Y., Barriault, D., and Sylvestre, M., 1998, Involvement of the terminal oxygenase β subunit in the biphenyl dioxygenase reactivity pattern toward chlorobiphenyls. *J. Bacteriol*., 180:5828–5835.
129. Iida, T., Mukouzaka, Y., Nakamura, K., and Kudo, T., 2002, Plasmid-borne genes code for an angular dioxygenase involved in dibenzofuran degradation by *Terrabacter* sp. strain YK3. *Appl. Environ. Microbiol*., 68:3716–3723.
130. Irie, S., Doi, S., Yorifuji, T., Takagi, M., and Yano, K., 1987, Nucleotide sequencing and characterization of the genes encoding benzene oxidation enzymes of *Pseudomonas putida*. *J. Bacteriol*., 169:5174–5179.
131. Irie, S., Shirai, K., Doi, S., and Yorifuji, T., 1987, Cloning of genes encoding oxidation of benzene in *Pseudomonas putida* and their expression in *Escherichia coli* and *P. putida*. *Agric. Biol. Chem*., 51:1489–1493.
132. Jeffrey, W.H., Cuskey, S.M., Chapman, P.J., Resnick, S., and Olsen, R.H., 1992, Characterization of *Pseudomonas putida* mutants unable to catabolize benzoate: cloning and

characterization of *Pseudomonas* genes involved in benzoate catabolism and isolation of a chromosomal DNA fragment able to substitute for *xylS* in activation of the TOL lower-pathway promoter. *J. Bacteriol*., 174:4986–4996.

133. Jeon, C.O., Park, W., Padmanabhan, P., DeRito, C., Snape, J.R., and Madsen, E.L., 2003, Discovery of a bacterium, with distinctive dioxygenase, that is responsible for in situ biodegradation in contaminated sediment. *Proc. Natl. Acad. Sci. U. S. A.*, 100:13591–13596.
134. Jerina, D.M., Selander, H., Yagi, H., Wells, M.C., Davey, J.F., Mahadevan, V., and Gibson, D.T., 1976, Dihydrodiols from anthracene and phenanthrene. *J. Am. Chem. Soc*., 98:5988–5996.
135. Jiang, H., Parales, R.E., and Gibson, D.T., 1999, The α subunit of toluene dioxygenase from *Pseudomonas putida* F1 can accept electrons from reduced ferredoxin$_{TOL}$ but is catalytically inactive in the absence of the β subunit. *Appl. Environ. Microbiol*., 65:315–318.
136. Jiang, H., Parales, R.E., Lynch, N.A., and Gibson, D.T., 1996, Site-directed mutagenesis of conserved amino acids in the alpha subunit of toluene dioxygenase: potential mononuclear non-heme iron coordination sites. *J. Bacteriol*., 178:3133–3139.
137. Jones, R.M., Britt-Compton, B., and Williams, P.A., 2003, The naphthalene catabolic (*nag*) genes of *Ralstonia* sp. strain U2 are an operon that is regulated by NagR, a LysR-type transcriptional regulator. *J. Bacteriol*., 185:5847–5853.
138. Ju, K.-S., and Parales, R.E., 2006, Control of substrate specificity by active site residues in nitrobenzene dioxygenase. *Appl. Environ. Microbiol*. In Press.
139. Junker, F., Kiewitz, R., and Cook, A.M., 1997, Characterization of the *p*-toluenesulfonate operon *tsaMBCD* and *tsaR* in *Comamonas testosteroni* T-2. *J. Bacteriol*., 179:919–927.
140. Junker, L., and Hay, A.G., 2002, Preliminary evidence for bacterial growth on styrene via the *tod* pathway. In *Abstracts of the 102rd General Meeting of the American Society for Microbiology 2002*. American Society for Microbiology.
141. Junker, L., and Hay, A.G., 2003, Bacterial growth on styrene using a mosaic pathway. In *Abstracts of the 103rd General Meeting of the American Society for Microbiology 2003*. American Society for Microbiology.
142. Karlsson, A., Beharry, Z.M., Eby, M., Coulter, E.D., Neidle, E.L., Kurtz, D.M.J., Eklund, H., and Ramaswamy, S., 2002, X-ray crystal structure of benzoate 1,2-dioxygenase reductase from *Acinetobacter* sp. strain ADP1. *J. Mol. Biol*., 318:261–272.
143. Karlsson, A., Parales, J.V., Parales, R.E., Gibson, D.T., Eklund, H., and Ramaswamy, S., 2003, Crystal structure of naphthalene dioxygenase: side-on binding of dioxygen to iron. *Science*, 299:1039–1043.
144. Kasai, Y., Shindo, K., Harayama, S., and Misawa, N., 2003, Molecular characterization and substrate preference of a polycyclic aromatic hydrocarbon dioxygenase from *Cycloclasticus* sp. strain A5. *Appl. Environ. Microbiol*., 69:6688–6697.
145. Kasuga, K., Habe, H., Chung, J.S., Yoshida, T., Nojiri, H., Yamane, H., and Omori, T., 2001, Isolation and characterization of the genes encoding a novel oxygenase component of angular dioxygenase from the gram-positive dibenzofuran-degrader *Terrabacter* sp. strain DBF63. *Biochem. Biophys. Res. Commun*., 283:195–204.
146. Kauppi, B., 1997. Doctoral thesis. Swedish University of Agricultural Sciences, Uppsala.
147. Kauppi, B., Lee, K., Carredano, E., Parales, R.E., Gibson, D.T., Eklund, H., and Ramaswamy, S., 1998, Structure of an aromatic ring-hydroxylating dioxygenase- naphthalene 1,2-dioxygenase. *Structure*, 6:571–586.
148. Kesseler, M., Dabbs, E.R., Averhoff, B., and Gottschalk, G., 1996, Studies on the isopropylbenzene 2,3-dioxygenase and 3-isopropylcatechol 2,3-dioxygenase genes encoded by the linear plasmid of *Rhodococcus erythropolis* BD2. *Microbiology*, 142:3241–3251.
149. Khan, A.A., Wang, R.F., Cao, W.W., Doerge, D.R., Wennerstrom, D., and Cerniglia, C.E., 2001, Molecular cloning, nucleotide sequence, and expression of genes encoding a polycyclic aromatic ring dioxygenase from *Mycobacterium* sp. strain PYR-1. *Appl. Environ. Microbiol*., 67:3577–3585.

150. Kikuchi, Y., Nagata, Y., Hinata, M., Kimbara, K., Fukuda, M., Yano, K., and Takagi, M., 1994, Identification of the *bphA4* gene encoding ferredoxin reductase involved in biphenyl and polychlorinated biphenyl degradation in *Pseudomonas* sp. strain KKS102. *J. Bacteriol*., 176:1689–1694.
151. Kim, D., Chae, J.-C., Zylstra, G.J., Kim, Y.-S., Kim, S.-K., Nam, M.H., Kim, Y.M., and Kim, E., 2004, Identification of a novel dioxygenase involved in metabolism of *o*-xylene, toluene, and ethylbenzene by *Rhodococcus* sp. strain DK17. *Appl. Environ. Microbiol*., 70:7086–7092.
152. Kim, D., Kim, Y.S., Kim, S.K., Kim, S.W., Zylstra, G.J., Kim, Y.M., and Kim, E., 2002, Monocyclic aromatic hydrocarbon degradation by *Rhodococcus* sp. strain DK17. *Appl. Environ. Microbiol*., 68:3270–3278.
153. Kimura, N., Nishi, A., Goto, M., and Furukawa, K., 1997, Functional analyses of a variety of chimeric dioxygenases constructed from two biphenyl dioxygenases that are similar structurally but different functionally. *J. Bacteriol*., 179:3936–3943.
154. Knackmuss, H.-J., Beckmann, W., and Otting, W., 1976, Microbiological synthesis of (+)-*cis*-1,2-dihydroxy-1,2-dihydronaphthalene-2-carboxylic acid. *Angew. Chem. Int. Ed. Engl*., 15:549.
155. Kodama, K., Umehara, K., Shimizu, K., Nakatani, S., Minoda, Y., and Yamada, K., 1973, Identification of microbial products from dibenzothiophene and its proposed oxidation pathway. *Agric. Biol. Chem*., 37:45–50.
156. Krivobok, S., Kuony, S., Meyer, C., Louwagie, M., Willison, J.C., and Jouanneau, Y., 2003, Identification of pyrene-induced proteins in *Mycobacterium* sp. strain 6PY1: evidence for two ring-hydroxylating dioxygenases. *J. Bacteriol*., 185:3828–3841.
157. Kulakov, L.A., Chen, S., Allen, C.C.R., and Larkin, M.J., 2005, Web-type evolution of *Rhodococcus* gene clusters associated with utilization of naphthalene. *Appl. Environ. Microbiol*., 71:1754–1764.
158. Kulakova, A.N., Reid, K.A., Larkin, M.J., Allen, C.R., and Kulakov, L.A., 1995, Plasmid controlling 1-chloroalkane degradation by *Rhodococcus rhodochrous* NCIMB13064. *Plasmid*, 33:208–217.
159. Kumamaru, T., Suenaga, H., Mitsuoka, M., Watanabe, T., and Furukawa, K., 1998, Enhanced degradation of polychlorinated biphenyls by directed evolution of biphenyl dioxygenase. *Nat. Biotechnol*., 16:663–666.
160. Lakshman, M.K., Chaturvedi, S., Zajc, B., Gibson, D.T., and Resnick, S.M., 1998, A general chemoenzymatic synthesis of enantiopure *cis* β-amino alcohols from microbially derived *cis*-glycols. *Synthesis*, 9:1352–1356.
161. Lange, C.C., and Wackett, L.P., 1997, Oxidation of aliphatic olefins by toluene dioxygenase: enzyme rates and product identification. *J. Bacteriol*., 179:3858–3865.
162. Lange, C.C., Wackett, L.P., Minton, K.W., and Daly, M.J., 1998, Engineering a recombinant *Deinococcus radiodurans* for organopollutant degradation in radioactive mixed waste environments. *Nat. Biotechnol*., 16:929–933.
163. Lange, S.J., and Que, L.J., 1998, Oxygen activating nonheme iron enzymes. *Curr. Opin. Chem. Biol*., 2:159–172.
164. Larkin, M.J., Allen, C.C.R., Kulakov, L.A., and Lipscomb, D.A., 1999, Purification and characterization of a novel naphthalene dioxygenase from *Rhodococcus* sp. strain NCIMB12038. *J. Bacteriol*., 181:6200–6204.
165. Laurie, A.D., and Lloyd-Jones, G., 1999, The *phn* genes of *Burkholderia* sp. strain RP007 constitute a divergent gene cluster for polycyclic aromatic hydrocarbon catabolism. *J. Bacteriol*., 181:531–540.
166. Lee, K., 1995, Biochemical studies on toluene and naphthalene dioxygenases. Ph.D. dissertation. The University of Iowa, Iowa City, Iowa.
167. Lee, K., 1998, Involvement of electrostatic interactions between the components of toluene dioxygenase from *Pseudomonas putida* F1. *J. Microbiol. Biotechnol*., 8:416–421.

168. Lee, K., 1999, Benzene-induced uncoupling of naphthalene dioxygenase activity and enzyme inactivation by production of hydrogen peroxide. *J. Bacteriol*., 181:2719–2725.
169. Lee, K., Brand, J.M., and Gibson, D.T., 1995, Stereospecific sulfoxidation by toluene and naphthalene dioxygenases. *Biochem. Biophys. Res. Commun*., 212:9–15.
170. Lee, K., and Gibson, D.T., 1996, Stereospecific dihydroxylation of the styrene vinyl group by purified naphthalene dioxygenase from *Pseudomonas* sp. strain 9816-4. *J. Bacteriol*., 178:3353–3356.
171. Lee, K., and Gibson, D.T., 1996, Toluene and ethylbenzene oxidation by purified naphthalene dioxygenase from *Pseudomonas* sp. strain NCIB 9816-4. *Appl. Environ. Microbiol*., 62:3101–3106.
172. Lee, K.-S., Parales, J.V., Friemann, R., and Parales, R.E., 2005, Active site residues controlling substrate specificity in 2-nitrotolnene dioxygenase from *Acidovorax* sp. strain JS42. J. Ind. Microbiol. Biotechnol. 32:465–473.
173. Lessner, D.J., Johnson, G.R., Parales, R.E., Spain, J.C., and Gibson, D.T., 2002, Molecular characterization and substrate specificity of nitrobenzene dioxygenase from *Comamonas* sp. strain JS765. *Appl. Environ. Microbiol*., 68:634–641.
174. Li, S., and Wackett, L.P., 1992, Trichloroethylene oxidation by toluene dioxygenase. *Biochem. Biophys. Res. Commun*., 185:443–451.
175. Lipscomb, J.D., 1994, Biochemistry of the soluble methane monooxygenase. *Annu. Rev. Microbiol*., 48:371–399.
176. Lloyd-Jones, G., Laurie, A.D., Hunter, D.W.F., and Fraser, R., 1999, Analysis of catabolic genes for naphthalene and phenanthrene degradation in contaminated New Zealand soils. *FEMS Microbiol. Ecol*., 29:69–79.
177. Locher, H.H., Leisinger, T., and Cook, A.M., 1991, 4-Sulphobenzoate 3,4-dioxygenase: purification and properties of a desulphonative two-component enzyme system from *Comamonas testosteroni* T-2. *Biochem. J*., 274:833–842.
178. Locher, H.H., Leisinger, T., and Cook, A.M., 1991, 4-Toluene sulfonate methyl-monooxygenase from *Comamonas testosteroni* T-2: purification and some properties of the oxygenase component. *J. Bacteriol*., 173:3741–3748.
179. Lorenzo, P., Alonso, S., Velasco, A., Diaz, E., Garcia, J.L., and Perera, J., 2003, Design of catabolic cassettes for styrene biodegradation. *Antonie Van Leeuwenhoek*., 84:17–24.
180. Maeda, T., Takahashi, Y., Suenaga, H., Suyama, A., Goto, M., and Furukawa, K., 2001, Functional analyses of Bph-Tod hybrid dioxygenase, which exhibits high degradation activity toward trichloroethylene. *J. Biol. Chem*., 276:29833–29838.
181. Mampel, J., Ruff, J., Junker, F., and Cook, A.M., 1999, The oxygenase component of the 2-aminobenzenesulfonate dioxygenase system from *Alcaligenes* sp. strain O-1. *Microbiology*, 145:3255–3264.
182. Martin, V.J.J., and Mohn, W.W., 1999, A novel aromatic-ring-hydroxylating dioxygenase from the diterpenoid-degrading bacterium *Pseudomonas abietaniphila* BKME-9. *J. Bacteriol*., 181:2675–2682.
183. Martin, V.J.J., and Mohn, W.W., 2000, Genetic investigation of the catabolic pathway for degradation of abietane diterpenoids by *Pseudomonas abietaniphila* BKME-9. *J. Bacteriol*., 182:3784–3793.
184. Masai, E., Yamada, A., Healy, J., Hatta, T., Kimbara, K., Fukuda, M., and Yano, K., 1995, Characterization of biphenyl catabolic genes of gram-positive polychlorinated biphenyl degrader *Rhodococcus* sp. strain RHA1. *Appl. Environ. Microbiol*., 61:2079–2085.
185. Minshull, J., 1995, Cleaning up our own back yard: developing new catabolic pathways to degrade pollutants. *Curr. Biol*., 2:775–780.
186. Mondello, F.J., Turcich, M.P., Lobos, J.H., and Erickson, B.D., 1997, Identification and modification of biphenyl dioxygenase sequences that determine the specificity of polychlorinated biphenyl degradation. *Appl. Environ. Microbiol*., 63:3096–3103.

187. Moreno-Ruíz, E., Hernaez, M.J., Martínez-Pérez, O., and Santero, E., 2003, Identification and functional characterization of *Sphingomonas macrogolitabida* strain TFA genes involved in the first two steps of the tetralin catabolic pathway. *J. Bacteriol*., 185:2026–2030.
188. Morono, Y., Unno, H., Tanji, Y., and Hori, K., 2004, Addition of aromatic substrates restores trichloroethylene degradation activity in *Pseudomonas putida* F1. *Appl. Environ. Microbiol*., 70:2830–2835.
189. Mosqueda, G., Ramos-González, M.I., and Ramos, J.L., 1999, Toluene metabolism by the solvent-tolerant *Pseudomonas putida* DOT-T1 strain, and its role in solvent impermeabilization. *Gene*, 232:69–76.
190. Muller, T.A., Werlen, C., Spain, J.C., and Van Der Meer, J.R., 2003, Evolution of a chlorobenzene degradative pathway among bacteria in a contaminated groundwater mediated by a genomic island in *Ralstonia*. *Environ. Microbiol*., 5:163–173.
191. Murakami, S., Hayashi, T., Maeda, T., Takenaka, S., and Aoki, K., 2003, Cloning and functional analysis of aniline dioxygenase gene cluster, from *Frateuria* species ANA-18, that metabolizes aniline via an *ortho*-cleavage pathway of catechol. *Biosci. Biotechnol. Biochem*., 67:2351–2358.
192. Nakatsu, C.H., Straus, N.A., and Wyndham, R.C., 1995, The nucleotide sequence of the Tn5271 3-chlorobenzoate 3,4-dioxygenase genes (*cbaAB*) unites the class IA oxygenases in a single lineage. *Microbiology*, 141:485–495.
193. Nakatsu, C.H., and Wyndham, C., 1993, Cloning and expression of the transposable chlorobenzoate-3,4-dioxygenase genes of *Alcaligenes* sp. strain BR60. *Appl. Environ. Microbiol*., 59:3625–3633.
194. Nam, J.W., Noguchi, H., Fujimoto, Z., Mizuno, H., Ashikawa, Y., Abo, M., Fushinobu, S., Kobashi, N., Wakagi, T., Iwata, K., Yoshida, T., Habe, H., Yamane, H., Omori, T., and Nojiri, H., 2005, Crystal structure of the ferredoxin component of carbazole 1,9a-dioxygenase of *Pseudomonas resinovorans* strain CA10, a novel Rieske non-heme iron oxygenase system. *Proteins*, 58:779–789.
195. Nam, J.-W., Nojiri, H., Noguchi, H., Uchimura, H., Yoshida, T., Habe, H., Yamane, H., and Omori, T., 2002, Purification and characterization of carbazole 1,9a-dioxygenase, a three component dioxygenase system of *Pseudomonas resinovorans* strain CA10. *Appl. Environ. Microbiol*., 68:5882–5890.
196. Nam, J.-W., Nojiri, H., Yoshida, T., Habe, H., Yamane, H., and Omori, T., 2001, New classification for oxygenase components involved in ring-hydroxylating oxygenations. *Biosci. Biotechnol. Biochem*., 65:254–263.
197. Neidle, E.L., Hartnett, C., Ornston, L.N., Bairock, A., Rekik, M., and Harayama, S., 1991, Nucleotide sequences of the *Acinetobacter calcoaceticus benABC* genes for benzoate 1,2-dioxygenase reveal evolutionary relationships among multicomponent oxygenases. *J. Bacteriol*., 173:5385–5395.
198. Nishino, S.F., Paoli, G.C., and Spain, J.C., 2000, Aerobic degradation of dinitrotoluenes and the pathway for bacterial degradation of 2,6-dinitrotoluene. *Appl. Environ. Microbiol*., 66:2139–2147.
199. Nishino, S.F., and Spain, J.C., 1995, Oxidative pathway for the biodegradation of nitrobenzene by *Comamonas* sp. strain JS765. *Appl. Environ. Microbiol*., 61:2308–2313.
200. Nojiri, H., Habe, H., and Omori, T., 2001, Bacterial degradation of aromatic compounds via angular dioxygenation. *J. Gen. Appl. Microbiol*., 47:279–305.
201. Nojiri, H., Nam, J.-W., Kosaka, M., Morii, K., Takemura, T., Furihata, K., Yamane, H., and Omori, T., 1999, Diverse oxygenations catalyzed by carbazole 1,9a-dioxygenase from *Pseudomonas* sp. strain CA10. *J. Bacteriol*., 181:3105–3113.
202. Nojiri, H., and Omori, T., 2002, Molecular bases of aerobic bacterial degradation of dioxins: involvement of angular dioxygenation. *Biosci. Biotechnol. Biochem*., 66:2001–2016.

203. Nojiri, H., Sekiguchi, H., Maeda, K., Urata, M., Nakai, S., Yoshida, T., Habe, H., and Omori, T., 2001, Genetic characterization and evolutionary implications of a *car* gene cluster in the carbazole degrader *Pseudomonas* sp. strain CA10. *J. Bacteriol*., 183:3663–3679.
204. Nomura, Y., Nakagawa, M., Ogawa, N., Harashima, S., and Oshima, Y., 1992, Genes in PHT plasmid encoding the initial degradation pathway of phthalate in *Pseudomonas putida*. *J. Ferment. Bioeng*., 6:333–344.
205. O'Brien, X.M., Parker, J.A., Lessard, P.A., and Sinskey, A.J., 2002, Engineering an indene bioconversion process for the production of *cis*-aminoindanol: a model system for the production of chiral synthons. *Appl. Microbiol. Biotechnol*., 59:389–399.
206. Ohtsubo, Y., Shimura, M., Delawary, M., Kimbara, K., Takagi, M., Kudo, T., Ohta, A., and Nagata, Y., 2003, A novel approach to the improvement of biphenyl and polychlorinated biphenyl degradation activity: promoter implantation by homologous recombination. *Appl. Environ. Microbiol*., 69:146–153.
207. Ouchiyama, N., Miyachi, S., and Omori, T., 1998, Cloning and nucleotide sequence of carbazole catabolic genes from *Pseudomonas stutzeri* strain OM1, isolated from activated sludge. *J. Gen. Appl. Microbiol*., 44:57–63.
208. Parales, J.V., Kumar, A., Parales, R.E., and Gibson, D.T., 1996, Cloning and sequencing of the genes encoding 2-nitrotoluene dioxygenase from *Pseudomonas* sp. JS42. *Gene*, 181:57–61.
209. Parales, J.V., Parales, R.E., Resnick, S.M., and Gibson, D.T., 1998, Enzyme specificity of 2-nitrotoluene 2,3-dioxygenase from *Pseudomonas* sp. strain JS42 is determined by the C-terminal region of the α subunit of the oxygenase component. *J. Bacteriol*., 180:1194–1199.
210. Parales, R.E., 2000, Molecular biology of nitroarene degradation, pp. 63–89. In J.C. Spain, J.B. Hughes, and H.-J. Knackmuss (eds.), *Biodegradation of Nitroaromatic Compounds and Explosives*. CRC Press, Boca Raton, FL.
211. Parales, R.E., and Ditty, J.L., 2005, Laboratory evolution of catabolic enzymes and pathways. *Curr. Opin. Biotechnol*., 16:315–325.
212. Parales, R.E., Emig, M.D., Lynch, N.A., and Gibson, D.T., 1998, Substrate specificities of hybrid naphthalene and 2,4-dinitrotoluene dioxygenase enzyme systems. *J. Bacteriol*., 180:2337–2344.
213. Parales, R.E., Huang, R., Yu, C.-L., Parales, J.V., Lee, F.K.N., Ivkovic-Jensen, M.M., Liu, W., Lessner, D.J., Friemann, R., Ramaswamy, S., and Gibson, D.T., 2005, Purification, characterization, and crystallization of the components of the nitrobenzene and 2-nitrotoluene dioxygenase enzyme systems. *Appl. Environ. Microbiol*., 71:3806–3814.
214. Parales, R.E., Lee, K., Resnick, S.M., Jiang, H., Lessner, D.J., and Gibson, D.T., 2000, Substrate specificity of naphthalene dioxygenase: effect of specific amino acids at the active site of the enzyme. *J. Bacteriol*., 182:1641–1649.
215. Parales, R.E., Parales, J.V., and Gibson, D.T., 1999, Aspartate 205 in the catalytic domain of naphthalene dioxygenase is essential for activity. *J. Bacteriol*., 181:1831–1837.
216. Parales, R.E., Resnick, S.M., Yu, C.L., Boyd, D.R., Sharma, N.D., and Gibson, D.T., 2000, Regioselectivity and enantioselectivity of naphthalene dioxygenase during arene *cis*-dihydroxylation: control by phenylalanine 352 in the α subunit. *J. Bacteriol*., 182:5495–5504.
217. Pavel, E.G., Martins, L.J., Ellis, W.R.J., and Solomon, E.I., 1994, Magnetic circular dichroism studies of exogenous ligand and substrate binding to the non-heme ferrous active site in phthalate dioxygenase. *Chem. Biol*., 1:173–183.
218. Pieper, D.H., and Reineke, W., 2000, Engineering bacteria for bioremediation. *Curr. Opin. Biotechnol*., 11:262–270.
219. Pieper, D.H., Timmis, K.N., and Ramos, J.L., 1996, Designing bacteria for the degradation of nitro- and chloroaromatic pollutants. *Naturwissenschaften*, 83:201–213.

220. Pinyakong, O., Habe, H., Yoshida, T., Nojiri, H., and Omori, T., 2003, Identification of three novel salicylate 1-hydroxylases involved in the phenanthrene degradation of *Sphingobium* sp. strain P2. *Biochem. Biophys. Res. Commun*., 301:350–357.
221. Pollmann, K., Beil, S., and Pieper, D.H., 2001, Transformation of chlorinated benzenes and toluenes by *Ralstonia* sp. strain PS12 *tecA* (tetrachlorobenzene dioxygenase) and *tecB* (chlorobenzene dihydrodiol dehydrogenase) gene products. *Appl. Environ. Microbiol*., 67:4057–4063.
222. Pollmann, K., Wray, V., Hecht, H.-J., and Pieper, D.H., 2003, Rational engineering of the regioselectivity of TecA tetrachlorobenzene dioxygenase for the transformation of chlorinated toluenes. *Microbiology*, 149:903–913.
223. Priefert, H., Rabenhorst, J., and Steinbüchel, A., 1997, Molecular characterization of genes of *Pseudomonas* sp. strain HR199 involved in bioconversion of vanillin to protocatechuate. *J. Bacteriol*., 179:2595–2607.
224. Que, L.J., 2000, One motif- many different reactions. *Nat. Struct. Biol*., 7:182–184.
225. Que, L.J., and Ho, R.Y.N., 1996, Dioxygen activation by enzymes with mononuclear non-heme iron active sites. *Chem. Rev*., 96:2607–2624.
226. Ramos, J.L., Marqués, S., and Timmis, K.N., 1997, Transcriptional control of the *Pseudomonas* TOL plasmid catabolic operons is achieved through an interplay of host factors and plasmid-encoded regulators. *Annu. Rev. Microbiol*., 51:341–373.
227. Raschke, H., Meier, M., Burken, J.G., Hany, R., Müller, M.D., Van Der Meer, J.R., and Kohler, H.P., 2001, Biotransformation of various substituted aromatic compounds to chiral dihydrodihydroxy derivatives. *Appl. Environ. Microbiol*., 67:3333–3339.
228. Reddy, J., Lee, C., Neeper, M., Greasham, R., and Zhang, J., 1999, Development of a bioconversion process for production of *cis*-1*S*, 2*R*-indandiol from indene by recombinant *Escherichia coli* constructs. *Appl. Microbiol. Biotechnol*., 51:614–620.
229. Resnick, S.M., and Gibson, D.T., 1993, Biotransformation of anisole and phenetole by aerobic hydrocarbon-oxidizing bacteria. *Biodegradation*, 4:195–203.
230. Resnick, S.M., and Gibson, D.T., 1996, Regio- and stereospecific oxidation of 9,10-dihydroanthracene and 9,10-dihydrophenanthrene by naphthalene dioxygenase: structure and absolute stereochemistry of metabolites. *Appl. Environ. Microbiol*., 62:3355–3359.
231. Resnick, S.M., and Gibson, D.T., 1996, Regio- and stereospecific oxidation of fluorene, dibenzofuran, and dibenzothiophene by naphthalene dioxygenase from *Pseudomonas* sp. strain NCIB 9816-4. *Appl. Environ. Microbiol*., 62:4073–4080.
232. Resnick, S.M., Lee, K., and Gibson, D.T., 1996, Diverse reactions catalyzed by naphthalene dioxygenase from *Pseudomonas* sp. strain NCIB 9816. *J. Ind. Microbiol*., 17:438–457.
233. Resnick, S.M., Torok, D.S., and Gibson, D.T., 1993, Oxidation of carbazole to 3-hydroxycarbazole by naphthalene 1,2-dioxygenase and biphenyl 2,3-dioxygenase. *FEMS Microbiol. Lett*., 113:297–302.
234. Resnick, S.M., Torok, D.S., Lee, K., Brand, J.M., and Gibson, D.T., 1994, Regiospecific and stereoselective hydroxylation of 1-indanone and 2-indanone by naphthalene dioxygenase and toluene dioxygenase. *Appl. Environ. Microbiol*., 60:3323–3328.
235. Robertson, J.B., Spain, J.C., Haddock, J.D., and Gibson, D.T., 1992, Oxidation of nitrotoluene by toluene dioxygenase: evidence for a monooxygenase reaction. *Appl. Environ. Microbiol*., 58:2643–2648.
236. Rosche, B., Tshisuaka, B., Fetzner, S., and Lingens, F., 1995, 2-Oxo-1,2-dihydroquinoline 8-monooxygenase, a two-component enzyme system from *Pseudomonas putida* 86. *J. Biol. Chem*., 270:17836–17842.
237. Rosche, B., Tshisuaka, B., Hauer, B., Lingens, F., and Fetzner, S., 1997, 2-Oxo-1,2-dihydroquinoline 8-monooxygenase: phylogenetic relationships to other multicomponent nonheme iron oxygenases. *J. Bacteriol*., 179:3549–3554.

238. Saito, A., Iwabuchi, T., and Harayama, S., 2000, A novel phenanthrene dioxygenase from *Nocardioides* sp. strain KP7: expression in *Escherichia coli*. *J. Bacteriol*., 182:2134–2141.
239. Sato, S.-I., Nam, J.-W., Kasuga, K., Nojiri, H., Yamane, H., and Omori, T., 1997, Identification and characterization of genes encoding carbazole 1,9a-dioxygenase in *Pseudomonas* sp. strain CA10. *J. Bacteriol*., 179:4850–4858.
240. Sauber, K., Fröhner, C., Rosenberg, G., Eberspächer, J., and Lingens, F., 1977, Purification and properties of pyrazon dioxygenase from pyrazon-degrading bacteria. *Eur. J. Biochem*., 74:89–97.
241. Schläfli, H.R., Weiss, M.A., Leisinger, T., and Cook, A.M., 1994, Terephthalate 1,2-dioxygenase system from *Comamonas testosteroni* T-2: purification and some properties of the oxygenase component. *J. Bacteriol*., 176:6644–6652.
242. Schlichting, I., Berendzen, J., Chu, K., Stock, A.M., Maves, S.A., Benson, D.E., Sweet, R.M., Ringe, D., Petsko, G.A., and Sligar, S.G., 2000, The catalytic pathway of cytochrome P450cam at atomic resolution. *Science*, 287:1615–1622.
243. Schmid, A., Vereyken, I., Held, M., and Witholt, B., 2001, Preparative regio- and chemoselective functionalization of hydrocarbons catalyzed by cell free preparations of 2-hydroxybiphenyl 3-monooxygenase. *J. Mol. Catal. B Enzym*., 11:455–462.
244. Seeger, M., Camara, B., and Hofer, B., 2001, Dehalogenation, denitration, dehydrogenation, and angular attack on substituted biphenyls and related compounds by biphenyl dioxygenase. *J. Bacteriol*., 183:3548–3555.
245. Segura, A., Bunz, P.V., D'Argenio, D.A., and Ornston, L.N., 1999, Genetic analyis of a chromosomal region containing *vanA* and *vanB*, genes required for conversion of either ferulate or vanillate to protocatechuate in *Acinetobacter*. *J. Bacteriol*., 181:3439–3504.
246. Selifonov, S.A., Grifoll, M., Eaton, R.W., and Chapman, P.J., 1996, Oxidation of naphthenoaromatic and methyl-substituted aromatic compounds by naphthalene 1,2-dioxygenase. *Appl. Environ. Microbiol*., 62:507–514.
247. Selifonov, S.A., Grifoll, M., Gurst, J.E., and Chapman, P.J., 1993, Isolation and characterization of (+)-1,1a-dihydroxy-1-hydrofluoren-9-one formed by angular dioxygenation in the bacterial catabolism of fluorene. *Biochem. Biophys. Res. Commun*., 193:67–76.
248. Senda, T., Yamada, T., Sakurai, N., Kubota, M., Nishizaki, T., Masai, E., Fukuda, M., and Mitsui, Y., 2000, Crystal structure of NADH-dependent ferredoxin reductase component in biphenyl dioxygenase. *J. Mol. Biol*., 304:397–410.
249. Seto, M., Kimbara, K., Shimura, M., Hatta, T., Fukuda, M., and Yano, K., 1995, A novel transformation of polychlorinated biphenyls by *Rhodococcus* sp. strain RHA1. *Appl. Environ. Microbiol*., 61:3353–3358.
250. Sevrioukova, I.F., Li, H., and Poulos, T.L., 2004, Crystal structure of putidaredoxin reductase from *Pseudomonas putida*, the final structural component of the cytochrome P450cam monooxygenase. *J. Mol. Biol*., 336:889–902.
251. Shepherd, J.M., and Lloyd-Jones, G., 1998, Novel carbazole degradation genes of *Sphingomonas* CB3: sequence analysis, transcription, and molecular ecology. *Biochem. Biophys. Res. Commun*., 247:129–135.
252. Shigematsu, T., Yumihara, K., Ueda, Y., Morimura, S., and Kida, K., 2003, Purification and gene cloning of the oxygenase component of the terephthalate 1,2-dioxygenase system from *Delftia tsuruhatensis* strain T7. *FEMS Microbiol. Lett*., 220:255–260.
253. Shingleton, J.T., Applegate, B.M., Nagel, A.C., Bienkowski, P.R., and Sayler, G.S., 1998, Induction of the *tod* operon by trichloroethylene in *Pseudomonas putida* TVA8. *Appl. Environ. Microbiol*., 64:5049–5052.
254. Shirai, K., 1986, Screening of microorganisms for catechol production from benzene. *Agric. Biol. Chem*., 50:2875–2880.
255. Sikkema, J., and De Bont, J.A.M., 1993, Metabolism of tetralin (1,2,3,4-tetrahydronaphthalene) in *Corynebacterium* sp. strain C125. *Appl. Environ. Microbiol*., 59:567–572.

256. Simon, M.J., Osslund, T.D., Saunders, R., Ensley, B.D., Suggs, S., Harcourt, A., Suen, W.-C., Cruden, D.L., Gibson, D.T., and Zylstra, G.J., 1993, Sequences of genes encoding naphthalene dioxygenase in *Pseudomonas putida* strains G7 and NCIB 9816-4. *Gene*, 127:31–37.
257. Smith, D.J., Martin, V.J.J., and Mohn, W.W., 2004, A cytochrome P450 involved in the metabolism of abietane diterpenoids by *Pseudomonas abietaniphila* BKME-9. *J. Bacteriol*., 186:3631–3639.
258. Sonoki, T., Obi, T., Kubota, S., Higashi, M., Masai, E., and Katayama, Y., 2000, Coexistence of two different O demethylation systems in lignin metabolism by *Sphingomonas paucimobilis* SYK-6: cloning and sequencing of the lignin biphenyl-specific O-demethylase (LigX) gene. *Appl. Environ. Microbiol*., 66:2125–2132.
259. 259. Spanggord, R.J., Spain, J.C., Nishino, S.F., and Mortelmans, K.E., 1991, Biodegradation of 2,4-dinitrotoluene by a *Pseudomonas* sp. *Appl. Environ. Microbiol*., 57:3200–3205.
260. Stingley, R.L., Brezna, B., Khan, A.A., and Cerniglia, C.E., 2004, Novel organization of genes in a phthalate degradation operon of *Mycobacterium vanbaalenii* PYR-1. *Microbiology*, 150:3749–3761.
261. Story, S.P., Parker, S.H., Hayasaka, S.S., Riley, M.B., and Kline, E.L., 2001, Convergent and divergent points in catabolic pathways involved in utilization of fluoranthene, naphthalene, anthracene, and phenanthrene by *Sphingomonas paucimobilis* var. EPA505. *J. Ind. Microbiol. Biotechnol*., 26:369–382.
262. Subramanian, V., Liu, T.-N., Yeh, W.-K., and Gibson, D.T., 1979, Toluene dioxygenase: purification of an iron-sulfur protein by affinity chromatography. *Biochem. Biophys. Res. Commun*., 91:1131–1139.
263. Suen, W.-C., and Gibson, D.T., 1993, Isolation and preliminary characterization of the subunits of the terminal component of naphthalene dioxygenase from *Pseudomonas putida* NCIB 9816-4. *J. Bacteriol*., 175:5877–5881.
264. Suen, W.-C., and Gibson, D.T., 1994, Recombinant *Escherichia coli* strains synthesize active forms of naphthalene dioxygenase and its individual α and β subunits. *Gene*, 143:67–71.
265. Suen, W.-C., Haigler, B.E., and Spain, J.C., 1996, 2,4-Dinitrotoluene dioxygenase from *Burkholderi*a sp. strain DNT: similarity to naphthalene dioxygenase. *J. Bacteriol*., 178:4926–4934.
266. Suen, W.-C., and Spain, J.C., 1993, Cloning and characterization of *Pseudomonas* sp. strain DNT genes for 2,4-dinitrotoluene degradation. *J. Bacteriol*., 175:1831–1837.
267. Suenaga, H., Goto, M., and Furukawa, K., 2001, Emergence of multifunctional oxygenase activities by random priming recombination. *J. Biol. Chem*., 276:22500–22506.
268. Suenaga, H., Mitsuoka, M., Ura, Y., Watanabe, T., and Furukawa, K., 2001, Directed evolution of biphenyl dioxygenase: emergence of enhanced degradation capacity for benzene, toluene, and alkylbenzenes. *J. Bacteriol*., 183:5441–5444.
269. Suenaga, H., Watanabe, T., Sato, M., Ngadiman, and Furukawa, K., 2002, Alteration of regiospecificity in biphenyl dioxygenase by active-site engineering. *J. Bacteriol*., 184:3682–3688.
270. Suzuki, K., Ogawa, N., and Miyashita, K., 2001, Expression of 2-halobenzoate dioxygenase genes (*cbdSABC*) involved in the degradation of benzoate and 2-halobenzoate in *Burkholderia* sp. TH2. *Gene*, 262:137–145.
271. Sylvestre, M., Sirois, M., Hurtubise, Y., Bergeron, J., Ahmad, D., Shareck, F., Barriault, D., Guillemette, I., and Juteau, J.M., 1996, Sequencing of *Comamonas testosteroni* strain B-356-biphenyl/chlorobiphenyl dioxygenase genes: evolutionary relationships among Gram-negative bacterial biphenyl dioxygenases. *Gene*, 174:195–202.
272. Takagi, T., Habe, H., Yoshida, T., Yamane, H., Omori, T., and Nojiri, H., 2005, Characterization of [3Fe-4S] ferredoxin DbfA3, which functions in the angular dioxygenase system of *Terrabacter* sp. strain DBF63. *Appl. Microbiol. Biotechnol*., 68:336–345.

273. Takagi, T., Nojiri, H., Yoshida, T., Habe, H., and Omori, T., 2002, Detailed comparison between the substrate specificities of two angular dioxygenases, dibenzofuran dioxygenase from *Terrabacter* sp. and carbazole dixygenase from *Pseudomonas resinovorans*. *Biotechnol. Lett*., 24:2099–2106.
274. Takizawa, N., Kaida, N., Torigoe, S., Moritani, T., Sawada, T., Satoh, S., and Kiyohara, H., 1994, Identification and characterization of genes encoding polycyclic aromatic hydrocarbon dioxygenase and polycyclic aromatic hydrocarbon dihydrodiol dehydrogenase in *Pseudomonas putida* OUS82. *J. Bacteriol*., 176:2444–2449.
275. Tan, H.-M., and Cheong, C.-M., 1994, Substitution of the ISPα subunit of biphenyl dioxygenase from *Pseudomonas* results in a modification of the enzyme activity. *Biochem. Biophys. Res. Commun*., 204:912–917.
276. Tan, H.-M., Tang, H.-Y., Joannou, C.L., Abdel-Wahab, N.H., and Mason, J.R., 1993, The *Pseudomonas putida* ML2 plasmid-encoded genes for benzene dioxygenase are unusual in codon usage and low in G+C content. *Gene*, 130:33–39.
277. Thompson, J.D., Higgins, D.G., and Gibson, T.J., 1994, CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Res*., 22:4673–4680.
278. Timmis, K.N., Rojo, F., and Ramos, J.L., 1989, Design of new pathways for the catabolism of environmental pollutants. *Adv. Appl. Biotechnol*., 4:61–82.
279. Timmis, K.N., Steffan, R.J., and Unterman, R., 1994, Designing microorganisms for the treatment of toxic wastes. *Annu. Rev. Microbiol*., 48:525–557.
280. Torok, D.S., Resnick, S.M., Brand, J.M., Cruden, D.L., and Gibson, D.T., 1995, Desaturation and oxygenation of 1,2-dihydronaphthalene by toluene and naphthalene dioxygenase. *J. Bacteriol*., 177:5799–5805.
281. Treadway, S.L., Yanagimachi, K.S., Lankenau, E., Lessard, P.A., Stephanopoulos, G., and Sinskey, A.J., 1999, Isolation and characterization of indene bioconversion genes from *Rhodococcus* strain I24. *Appl. Microbiol. Biotechnol*., 51:786–793.
282. Tsang, H.-T., Batie, C.J., Ballou, P.D., and Penner-Halm, J.E., 1989, X-ray absorption spectroscopy of the [2Fe-2S] Reiske cluster in *Pseudomonas cepacia* phthalate dioxygenase : determination of core dimentions and iron ligand. *Biochemistry*, 28:7233–7240.
283. Tsoi, T.V., Plotnikova, E.G., Cole, J.R., Guerin, W.F., Bagdasarian, M., and Tiedje, J.M., 1999, Cloning, expression, and nucleotide sequence of the *Pseudomonas aeruginosa* 142 *ohb* genes coding for oxygenolytic *ortho* dehalogenation of halobenzoates. *Appl. Environ. Microbiol*., 65:2151–2162.
284. Urata, M., Uchida, E., Nojiri, H., Omori, T., Obo, R., Miyaura, N., and Ouchiyama, N., 2004, Genes involved in aniline degradation by *Delftia acidovorans* strain 7N and its distribution in the natural environment. *Biosci. Biotechnol. Biochem*., 68:2457–2465.
285. van der Meer, J.R., Van Neerven, A.R.W., De Vries, E.J., De Vos, W.M., and Zehnder, A.J.B., 1991, Cloning and characterization of plasmid-encoded genes for the degradation of 1,2-dichloro-, 1,4-dichloro, and 1,2,4-trichlorobenzene of *Pseudomonas* sp. strain P51. *J. Bacteriol*., 173:6–15.
286. Venturi, V., Zennaro, F., Degrassi, G., Okeke, B.C., and Bruschi, C.V., 1998, Genetics of ferulic acid bioconversion to protocatechuic acid in plant-growth-promoting *Pseudomonas putida* WCS358. *Microbiology*, 144:965–973.
287. Wackett, L.P., and Gibson, D.T., 1988, Degradation of trichloroethylene by toluene dioxygenase in whole cell studies with *Pseudomonas putida* F1. *Appl. Environ. Microbiol*., 54:1703–1708.
288. Wackett, L.P., Kwart, L.D., and Gibson, D.T., 1988, Benzylic monooxygenation catalyzed by toluene dioxygenase from *Pseudomonas putida*. *Biochemistry*, 27:1360–1367.
289. Wackett, L.P., Sadowsky, M.J., Newman, L.M., Hur, H.-G., and Li, S., 1994, Metabolism of polyhalogenated compounds by a genetically engineered bacterium. *Nature*, 368:627–629.

290. Walter, U., Beyer, M., Klein, J., and Rehm, H.J., 1991, Degradation of pyrene by *Rhodococcus* sp. UW1. *Appl. Microbiol. Biotechnol*., 34:671–676.
291. Wang, Y.Z., Zhou, Y., and Zylstra, G.J., 1995, Molecular analysis of isophthalate and terephthalate degradation by *Comamonas testosteroni* YZW-D. *Environ. Health Perspect*., 103(Suppl 5):9–12.
292. Warhurst, A.M., Clarke, K.F., Hill, R.A., Holt, R.A., and Fewson, C.A., 1994, Metabolism of styrene by *Rhodococcus rhodochrous* NCIMB 13259. *Appl. Environ. Microbiol*., 60:1137–1145.
293. Werlen, C., Kohler, H.-P., and van der Meer, J.R., 1996, The broad substrate chlorobenzene dioxygenase and *cis*-chlorobenzene dihydrodiol dehydrogenase of *Pseudomonas* sp. strain P51 are linked evolutionarily to the enzymes for benzene and toluene degradation. *J. Biol. Chem*., 271:4009–4016.
294. Whited, G.M., Downie, J.C., Hudlicky, T., Fearnley, S.P., Dudding, T.C., Olivo, H.F., and Parker, D., 1994, Oxidation of 2-methoxynaphthalene by toluene, naphthalene and biphenyl dioxygenases: structure and absolute stereochemistry of metabolites. *Bioorg. Med. Chem*., 2:727–734.
295. Widada, J., Nojiri, H., Kasuga, K., Yoshida, T., Habe, H., and Omori, T., 2002, Molecular detection and diversity of polycyclic aromatic hydrocarbon-degrading bacteria isolated from geographically diverse sites. *Appl. Microbiol. Biotechnol*., 58:202–209.
296. Wilson, M.S., Herrik, J.B., Jeon, C.O., Hinman, D.E., and Madsen, E.L., 2003, Horizontal transfer of *phnAc* dioxygenase genes within one of two phenotypically and genotypically distinctive naphthalene-degrading guilds from adjacent soil environments. *Appl. Environ. Microbiol*., 69:2172–2181.
297. Wittich, R.-M., Wilkes, H., Sinnwell, V., Francke, W., and Fortnagel, P., 1992, Metabolism of dibenzo-*p*-dioxin by *Sphingomonas* sp. strain RW1. *Appl. Environ. Microbiol*., 58:1005–1010.
298. Wolfe, M.D., and Lipscomb, J.D., 2003, Hydrogen peroxide-coupled *cis*-diol formation catalyzed by naphthalene 1,2-dioxygenase. *J. Biol. Chem*., 278:829–835.
299. Wolfe, M.D., Parales, J.V., Gibson, D.T., and Lipscomb, J.D., 2001, Singe turnover chemistry and regulation of O_2 activation by the oxygenase component of naphthalene dioxygenase. *J. Biol. Chem*., 276:1945–1953.
300. Yabuuchi, E., Yamamoto, H., Terakubo, S., Okamura, N., Naka, T., Fujiwara, N., Kobayashi, K., Kosako, Y., and Hiraishi, A., 2001, Proposal of *Sphingomonas wittichii* sp. nov. for strain $RW1^T$, known as a dibenzo-*p*-dioxin metabolizer. *Int. J. Syst. Evol. Microbiol*., 51:281–292.
301. Yen, K.-M., and Gunsalus, I.C., 1982, Plasmid gene organization: naphthalene/salicylate oxidation. *Proc. Natl. Acad. Sci. U. S. A.*, 79:874–878.
302. Yu, C.-L., Parales, R.E., and Gibson, D.T., 2001, Multiple mutations at the active site of naphthalene dioxygenase affect regioselectivity and enatioselectivity. *J. Ind. Microbiol. Biotechnol*., 27:94–103.
303. Zhang, N., Stewart, B.G., Moore, J.C., Greasham, R.L., Robinson, D.K., Buckland, B.C., and Lee, C., 2000, Directed evolution of toluene dioxygenase from *Pseudomonas putida* for improved selectivity toward *cis*-indandiol during indene bioconversion. *Metab. Eng*., 2:339–348.
304. Zhao, H., and Arnold, F.H., 1997, Optimization of DNA shuffling for high fidelity recombination. *Nucleic Acids Res*., 6:1307–1308.
305. Zhou, N.-Y., Al-Dulayymi, J., Baird, M.S., and Williams, P.A., 2002, Salicylate 5-hydroxylase from *Ralstonia* sp. strain U2: a monooxygenase with close relationships to and shared electron transport proteins with naphthalene dioxygenase. *J. Bacteriol*., 184:1547–1555.
306. Zhou, N.-Y., Fuenmayor, S.L., and Williams, P.A., 2001, *nag* genes of *Ralstonia* (formerly *Pseudomonas*) sp. strain U2 encoding enzymes for gentisate catabolism. *J. Bacteriol*., 183:700–708.

307. Zielinski, M., Backhaus, S., and Hofer, B., 2002, The principle determinants for the structure of the substrate-binding pocket are located within a central core of a biphenyl dioxygenase α subunit. *Microbiology*, 148:2439–2448.
308. Zylstra, G.J., and Gibson, D.T., 1989, Toluene degradation by *Pseudomonas putida* F1: nucleotide sequence of the *todC1C2BADE* genes and their expression in *E. coli. J. Biol. Chem*., 264:14940–14946.
309. Zylstra, G.J., and Gibson, D.T., 1991, Aromatic hydrocarbon degradation: a molecular approach. *Genet. Eng*., 13:183–203.

10

THE *PSEUDOMONAS* GENOME DATABASE

Fiona S. L. Brinkman

Department of Molecular Biology and Biochemistry
Simon Fraser University
Burnaby, B.C., Canada

1. PREAMBLE – THE *PSEUDOMONAS* GENOME DATABASE PHILOSOPHY AND SCOPE

The *Pseudomonas* Genome Database (www.*pseudomonas*.com) was initially developed with the goal of providing genome sequence and annotation data for *Pseudomonas aeruginosa* strain PAO1 in

- A user-friendly web-based format
- A format that facilitates continually updated, community-based annotation

The database continues to maintain these priorities, but a new version (version 2) has since developed further to address the following needs

- Comparative analysis of multiple *Pseudomonas* genomes
- Further enabling alternate annotations to be viewed, or development of alternative views of the same annotations
- Providing new computational analyses that facilitate the identification of genome components of medical importance

In the future, this resource aims to continue to meet these goals while also further expanding its collection of information directly relevant to *P. aeruginosa*

Pseudomonas, Volume 4, edited by Juan-Luis Ramos and Roger C. Levesque

infectious disease pathogenesis, vaccine development, and treatment. This collection of information will continue to be derived from new computational and laboratory analyses being performed by groups directly involved in this database development, and other research groups worldwide. One key to this approach will be continued focus on community involvement and providing more access to multiple different views of a given research problem or research data. It is hoped that the resulting diverse data collection, customized to help facilitate the development of new approaches for *P. aeruginosa* control, will become a significant research aid that may serve as a model for how a community can come together to reach a common goal. I and the other project coordinators of course encourage suggestions for improvement to the database structure, possible analyses, and content.

A review of this database is therefore provided below, with these goals in mind. This review provides further explanation of the history, goals, and philosophy of this project, along with what types of analyses are possible, or may be possible in the future. This will hopefully help more researchers realize how they may use this database to aid their research and spur them to participate in helping to maintain this database as a critical resource for *Pseudomonas* research in general.

2. INITIAL DATABASE DEVELOPMENT AND COMMUNITY ANNOTATION APPROACH – THE *PSEUDOMONAS* COMMUNITY COMES TOGETHER

When the *Pseudomonas aeruginosa* PAO1 Genome Project was first initiated in 1997, the major participants in this project wrestled with how to best annotate gene descriptions and other information for this genome, as have many first-time genome project organizers. At the time, genome annotation approaches were very much centralized, and concepts like open annotation, annotation jamborees, and web-based annotation submission had not yet been developed. *Pseudomonas aeruginosa*, an important pathogen and major cause of morbidity and mortality for Cystic Fibrosis patients, was the largest bacterial genome sequenced to date, with an estimated 6000 genes, and the third most cited bacterium in Medline. Robert Hancock and I decided to initiate a community-aided approach to capitalize on the large number of researchers studying this bacterium. Furthermore, we decided that the project should be conducted entirely through the Internet, in an effort to keep project costs down.

Our approach therefore involved enlisting the expertise of volunteer researchers from the *Pseudomonas* research community to annotate genes or gene families with which they were familiar. One moderator coordinated this

project, termed the *Pseudomonas* Community Annotation Project or PseudoCAP, and provided supportive documentation, including computer-generated annotation, and example curated annotations of selected genes, through the project's dedicated web site. On September 15, 1998, a formal request for participants was sent by email to members of the *Pseudomonas* research community. In all, 47 members of the research community initially expressed interest. Annotations provided by the participants were subjected to peer-review by a *Pseudomonas aeruginosa* genome analysis committee (PAGAC), before overlaying these annotations on a genome viewer console containing layers of other automatically generated analyses and literature reference information. This resource, coupled with a critical, conservative annotation approach, was used to generate the final genome annotations which were made available at www.*pseudomonas*.com[31].

One of the most surprising results from our investigation of this internet-based community annotation approach was the higher than anticipated participation rate in PseudoCAP. Of the 47 participants who initially expressed interest in the project, over 90% submitted annotations and, as word spread of the project, a total of 61 participants from 11 countries ended up making 1741 submissions. This is an exceptional participation rate, considering that participants provided their expertise on a volunteer basis only. Notably, most of the later participants were researchers not involved in *Pseudomonas* research, but rather researchers studying certain classes of genes who wished to help annotate the homologous genes in this particular genome. The high participation rate observed may reflect the research community's significant interest in increasing the quality of genome annotation[3].

We also observed that, even on a volunteer basis, some participants submitted very high-quality annotations. Regardless, if such high-quality annotations had been used exclusively, the final annotations would have contained numerous significant inconsistencies. Therefore, we found that one important step in the project involved using a small set of curators to provide a filtering step for consistency, and ensure conservative, critical annotation. This meant the final annotations were based almost exclusively on functional studies of the gene in question, or homology only to functionally studied proteins. This process essentially allowed us to incorporate expert knowledge from researchers who had been studying a particular set of genes for years – while formatting this knowledge into a consistent, relatively conservative, format that could potentially aid future genome-wide analyses.

We have therefore proposed[3] that this PseudoCAP method, complemented with a more decentralized approach using a distributed annotation system[6] such as GBrowse[30] (mentioned below), is an approach that best suits many genome projects, as it permits both the development of centralized annotation for all to use as a reference (i.e. gene names) combined with a system to facilitate easy access to alternate annotations. A number of other community annotation

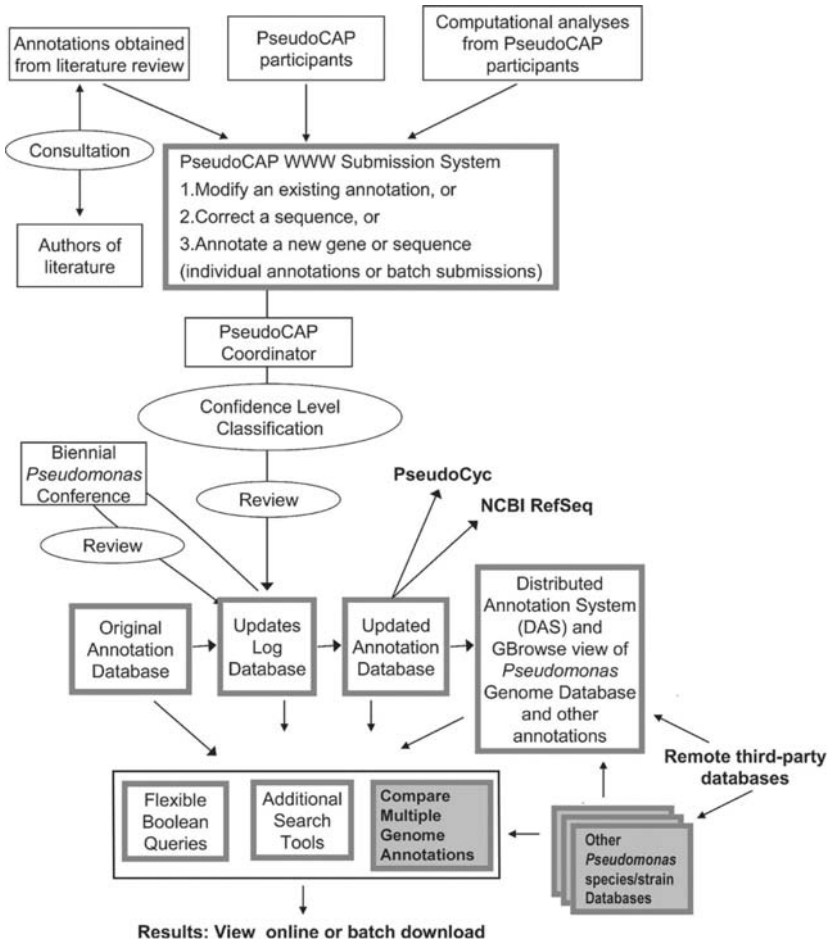


Figure 1. Schematic overview of the *Pseudomonas* Genome Database and PseudoCAP annotation update process. Squares with thick lines/borders denote database and application components. Circles reflect processes in the annotation update approach. Elements with thin lines/borders reflect human intervention. Shaded boxes denote new features added in version 2 of the database.

approaches have been developed with components of this philosophy in mind, including WormBase[29], the PeerGAD system[5], and ASAP[12]. Our currently expanded annotation approach, as outlined in Figure 1, has now been utilized by other genome projects, reflecting the interest in combining both peer-reviewed centralized, and unreviewed decentralized, data resources that utilize a combination of previously reported approaches, including the "open annotation" approach[15] and community annotation project (CAP) input[3].

3. PSEUDOCAP TODAY – FACILITATING CONTINUALLY UPDATED ANNOTATION

While PseudoCAP was invaluable during initial annotation efforts, this CAP approach continues to be used today in a more flexible, expanded fashion to enable annotations to be continually updated. Annotation updates now occur through a combination of researcher submissions, reviews of the literature, and computational analyses[35]. A log file is maintained of the updates that is searchable, and indicates what update type was made, who was involved, when the update occurred, and what rationale was used for the update (literature reference, computational analysis, or other). The ability of the database to perform concise Boolean searches and sorting of results is an important feature, since such functionality will become increasingly important as log files for updates become larger and more complex.

3.1. Annotation Updates by PseudoCAP Participants

Submission by a researcher can be made either by simply filling in a web-based form, sending an email to the project coordinator, or by creating a user ID and logging into the system (the latter is useful when making many submissions). All submissions are subject to initial review by the PseudoCAP coordinator. The coordinator examines the submission and then responds with any requests for additional information/clarification, if required. At this point a protein name confidence level is assigned or changed, using our previously developed classification system[31] (see Box 1).

In addition, a confidence level is assigned/changed for protein subcellular localization, as part of our expanded confidence system. Any suitable Gene Ontology[2] (GO) annotation is also now evaluated and associated with the annotation change.

The submission is then reviewed by at least one additional reviewer from the research community and a collection of all annotations made over a 2-year period are subject to additional review at the biennial International *Pseudomonas* conference. We feel this latter review step is important to provide the community with an efficient mechanism to review annotation updates collectively and examine and discuss any systematic annotation issues. Management of the review stages by a coordinator is also important to ensure consistency in annotation updates and to ensure that additional reviewers chosen from the research community are appropriate for a given annotation update case.

Note that both sequence information and annotations can be updated. Sequence changes, if they involve an insertion or deletion of a nucleotide, can affect the nucleotide coordinates of all annotations downstream of the change, however our unique ID system for numbering genes ensures that researchers

Box 1. Confidence level classification

One of the most frequent questions we receive is "what is the confidence level classification?". This system was designed to aid researchers who wish to identify only those proteins annotated with a certain level of confidence, such as identifying all those proteins with an experimentally determined subcellular localization, versus identifying all those who subcellular localization was computationally predicted using a particular method. For subcellular localization annotations, this classification system is relatively straightforward. For the protein description this classification is a little more subjective, however, in general a protein description is given a class 1 rating (meaning "function experimentally determined in *Pseudomonas aeruginosa*") if the protein name primarily reflects the function that has been experimentally studied. It is this class 1 rating that has apparently been the most useful, as it reflects a dataset of genes/proteins that have been experimentally studied. Of course this constantly needs to be updated, and so researchers are encouraged to send the project coordinator a quick note about any annotation that should be designated as a class 1. For more information about the classification system, see the "Frequently asked questions" (FAQ) list on the *Pseudomonas* Genome Database website.

can refer to a given gene using a coordinate (the unique ID) that will not change (see Box 2).

3.2. Annotation Updates Based on Literature Review

In addition to PseudoCAP participant submissions, the *Pseudomonas* research literature is also reviewed weekly using PubCrawler[14] and papers are noted that report new gene names, gene functions, or other information that may impact on the genome annotation. The corresponding author of the paper is contacted with a proposal for an annotation submission that is based on the paper's work. If agreed to, the submission is directly accepted (because it has been already subject to peer-review during publication) and the log file notes both that this was a submission based on literature review and provides information on the accepting author and journal citations.

3.3. Annotation Updates Based on Computational Analyses

In addition to annotation updates based on experimental data, the integration of computationally based predictions are also encouraged – as long as they

Box 2. The benefit of function-independent unique identifiers (or "why you should cite the PA number")

For the *P. aeruginosa* strain PAO1 genome annotation, records in both the original and updated annotations are uniquely identified by a locus ID consisting of "PA" (for *P. aeruginosa*) followed by a four-digit number representing the order of ORFs around the chromosome starting at the origin of replication. In consultation with the *Pseudomonas* research community, we have adapted the convention of using a decimal system to account for newly described ORFs (i.e. PA1000.1 would be between PA1000 and PA1001; PA1001.01 would be between PA1000 and PA1000.1)[35]. These identifiers can link to external databases such as TIGR's comprehensive microbial resource (CMR[24]), which represents an alternate annotation view, and the NCBI RefSeq[25] that we submit updates to. In the future we would also like to implement a similar unique ID number system for other genomic features, such as promoters and other regulatory elements.

But why have such a numbering system, rather than using a unique gene name (or other annotation name) for each annotation, which could be easier to remember?

A unique ID that is function-independent is useful because it will not need to be changed if a change in function is identified for a given gene. For example, based on an initial analysis, a (fictional) protein could be found to be involved in arginine uptake, and named ArgQ, but then later it is discovered that in fact the protein has a more general function involving uptake of basic amino acids. Meanwhile, another protein is found also to be involved more specifically in arginine uptake – prompting a proposed gene name change. Therefore, we strictly maintain the PA number as a unique identifier of a gene/protein that will never change, while allowing the use of multiple gene/protein names (a primary name and "alt", or alternative, names). This best reflects the continual evolution of protein descriptions, and in some cases gene names, as functions are better elucidated. This also handles the inevitable updates involving corrections to the sequence, which may change sequence coordinates for a given gene). Therefore, in short, a researcher should always refer to the PA number (or other annotation unique identifier) to best clarify what gene/protein is being referred to in a given analysis/experiment. We encourage the use of similar identifiers for other future *Pseudomonas* genome projects. We also discourage the development of gene IDs in related genomes on the basis of orthology (since, like function a proposed orthology may change as further information is obtained about the true relatedness of genes between given strains). One of the simplest conventions is therefore to number the genes around the genome as we, and many other genome projects, have done, and so I encourage this simple numbering system for other future genome projects.

Box 3. Literature-based annotation updates – constraints that should perhaps change?

Note that currently we only automatically review literature that involves a study of *P. aeruginosa* strain PAO1 specifically (i.e. we do not annotate the PAO1 strain genome-based on a published study in another *P. aeruginosa* strain). This means that it is important that a researcher clearly indicate in their publication what strain was used for a given experiment. However, in the interests of gathering more data, we are now considering in the future incorporating information from other strains, as long as we clearly differentiate such data from PAO1-based data. Of course, as illustrated in the first *P. aeruginosa* genome publication[31], there are multiple versions of the PAO1 strain in existence in research laboratories – some with notable inversions versus the sequenced strain. Therefore, even studies based on the PAO1 strain are not clear cut. Also, some annotations previously developed may have been based on data involving other strains that we were not aware of. Therefore, we are interested in incorporating more data from other strains, however, we feel it is important to emphasize the need to maintain strain specific information as much as possible, to maintain a database of the highest quality.

are clearly classified as such by our classification system, and the computational approach used is clearly noted. Since the accuracy of computational methods may change dramatically depending on the method used, indicating the specific method used and version number is critical. We also plan to incorporate more information about the accuracy of each method as small help files that would be associated with each computational analysis. For example, a signal peptide prediction method may have only about 80% precision, while the PSORTb method for protein subcellular localization prediction has over 95% precision. If a researcher knows of a particular computational analyses that they have found very useful, we encourage them to contact the project manager, Geoff Winsor, about it, as usually we can easily incorporate the analysis into the database (see the description of GBrowse, below).

In general, I encourage as much community participation in this project, as is possible, recognizing that the project must review the literature regularly to identify suitable annotation updates that may not be reported to us. Through the combination of community participation, literature review, and the addition of additional high-throughput analyses, such as computational analyses, this resource will hopefully continue to be current and maintain its usefulness in facilitating *Pseudomonas* research.

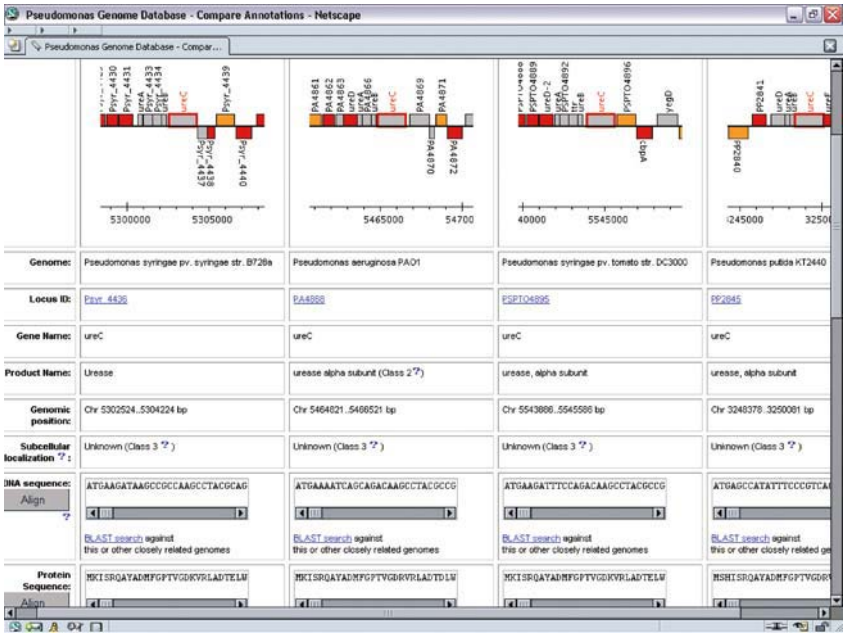


Figure 2. Screenshot of a comparison between annotations from different *Pseudomonas* species in the *Pseudomonas* Genome Database. To obtain this comparison, first a search was performed against all *Pseudomonas* species genomes available, and then the genes of interest selected and added to a "clipboard". Then a "compare" button is clicked, resulting in the illustrated display (a section of which is shown). Note how one can easily view difference in annotation (such as product name) and differences in gene order in the chosen sequence region. In addition, selecting the "Align" button for DNA or protein sequences generates a multiple sequence alignment using Clustal[33] of the selected genes/proteins being compared. The resulting alignment may be viewed directly or downloaded for further analysis.

4. DATABASE STRUCTURE OVERVIEW

An overview of the database structure for versions 1 and 2 of the database is shown schematically in Figure 2 (new features available in version 2 are shaded).

The *Pseudomonas* Genome Database currently comprises (1) the original *P. aeruginosa* PAO1 genome annotation published in the year 2000 (an important reference dataset[31]), (2) the continually updated *P. aeruginosa* PAO1 annotation and log file of updates[35] and (3) annotations of other *Pseudomonas* species/strain genomes that may be examined for comparative purposes. Visitors entering the site can examine any of these genomes and choose to browse, search or compare specific annotations as well as download tab-delimited files of information from

a search/comparison. Alternatively, one can submit an annotation update, either as a guest or as a registered PseudoCAP participant (the latter is required for more complex submissions). Eventually, we will develop this database further, to facilitate continual updating of the annotations of other *Pseudomonas* strains, in particular the *P. aeruginosa* strain data that we will be hosting as part of genome projects we are involved with. We are happy to assist researchers who either wish to host their *Pseudomonas* strain data on our system, or wish to host their strain data themselves on their own system – but using our database structure. The database structure and source code is freely available under an open source license (GNU GPL).

As mentioned in Box 2, a unique identifier ("PA number") is assigned to each gene. Records for each ORF contain information on the primary name associated with the ORF and its product as well as any alternate names that have been used. Furthermore, multiple functional classifications, genomic context, structure, predicted localization of the product as well as reactions and predicted pathways the product is involved in, PubMed references and DNA and protein sequences are stored. Fields are searchable using a Boolean search interface with the ability to flexibly sort the data and then either view the data directly or download the search results in tab-delimited format. ORFs can also be browsed by their order around the genome or by functional classification of their product. With regard to nucleotide and protein sequences, a BLAST search can be performed against genomic DNA and protein sequences using BLASTALL from NCBI (16). In addition, subsequences of the PAO1 genome (DNA or translated sequences) can be downloaded by specifying the base-pair coordinates of the DNA sequence. Finally, the amino acid sequence of proteins can be obtained by specifying the PA number of the gene encoding it.

5. COMPARATIVE GENOMICS CAPABILITY

In addition to the *P. aeruginosa* PAO1 genome annotations (original and updated), a new version (version 2) of this database now also contains annotations and sequence information for other *Pseudomonas* species/strains[4,21]. While the focus of the database is still centered around the *P. aeruginosa* PAO1 sequence and annotation, there has been a growing appreciation of the need to facilitate research discoveries through comparative genomics. Therefore, the database is moving further towards the development of more comparative views/analyses of multiple *Pseudomonas* genomes at once. Ironically the "*Pseudomonas* Genome Database" may finally now live up to its name. One can now perform a search involving multiple *Pseudomonas* genomes at once, and perform a "compare" of multiple annotations – either across multiple different genomes, or within one genome, or a combination of the two. An example of a comparison is illustrated in Figure 2. From such comparisons, one can then

perform additional comparative analyses, such as multiple sequence alignment of the selected gene/protein sequences.

This comparative feature is useful for comparing gene order between species or between genes that are a member of a particular gene family (for example, efflux proteins). It is also useful for comparing what is known about gene function for particular genes across species (or within gene families) or even for detecting anomalies or errors in gene annotation between species. Sequence alignments may also be easily obtained for the set of genes being compared, and links are available to more detailed annotations.

In the future, comparative analyses will be expanded to facilitate additional types of comparisons, such as comparative identification of insertions/genomic islands. The comparisons will also be made more flexible, for example allowing one to select which annotations they wish to view in the comparison and, for gene order graphics, invert a particular gene order in one organism, versus another organism, to facilitate making gene order comparisons between species whose genomes have large inversions in some regions.

6. GBROWSE – A DISTRIBUTED ANNOTATION SYSTEM THAT ALLOWS RESEARCHERS TO VIEW ALTERNATE ANNOTATIONS AND ADDITIONAL ANALYSES

Comparisons can also be made through GBrowse – an application that allows one to view multiple different annotations for a given sequence region/gene at once. With the increase in annotations available to microbiologists via the internet, it is becoming increasingly necessary to visualize genomic annotation information from multiple sources in a single viewer. We also feel that it is important to encourage alternate scientific views by allowing researchers to view any alternate annotations relative to our database's primary, peer-reviewed, annotation information. To facilitate this, we have incorporated a platform-independent web application called GBrowse developed by Stein *et al.*[30]. Using checkboxes, one can select diverse annotation information to view, including alternate gene annotations, motifs/structures, metabolic pathway data, gene knockout data, and ortholog data (the latter reflects comparisons with other genomes). One can perform a search based on criteria specified (for example, particular base-pair coordinates or a gene name) or just link to GBrowse from a particular gene entry in the main database. GBrowse then fetches the region of the genome specified by the search criteria and presents a view containing one or more horizontal tracks representing different sequence features for that area. The user is free to zoom in and out according to level of magnification/resolution desired. "Landmarks" (icons or glyphs) on each track usually

contain a link to detailed information contained on additional websites. One can select, and even reorder, what tracks are shown, and can also upload their own data (for example functional genomics data), for viewing alongside the other tracks. Uploaded data can be privately viewed by implementing the Distributed Annotation System[6]. Through this system, researchers can upload annotation data to a server where others can also view this data within a GBrowse track by entering the URL of the reference server. In essence, *Pseudomonas* GBrowse

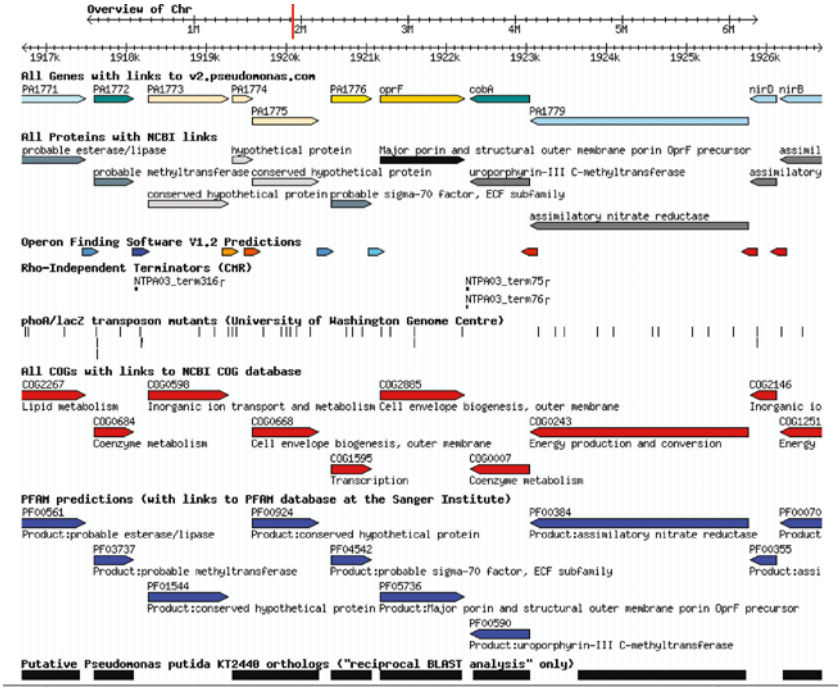


Figure 3. Screenshot showing selected tracks in a GBrowse display, illustrating the variety of different annotations that may be shown. Genomic coordinates for the selected region are shown at top, followed by a series of tracks containing "landmarks" (glyphs, or colored icons) that link to additional information. For example, the first track links to *P. aeruginosa* PAO1 gene annotations at *Pseudomonas*.com, while the second links to NCBI data and the third links to operon prediction information for these genes. Information about each track can be set to either be verbose or compact. For example, for the operon prediction track, the color of the landmarks between genes reflects the probability that the two genes are cotranscribed without a promoter between them[34]. However under "track options" this track can be changed to show also the specific *p*-value for the prediction. Conversely, the track of "all proteins with NCBI links" can be changed to the compact form, which doesn't list the long protein names, permitting the track to be displayed as one line. Track order can also be changed and one can zoom in or out of the sequence at any level desired. For a full list of all currently available tracks, see Table 1.

and DAS allow researchers to easily view alternate annotations while promoting open discussion leading to better annotations for the *Pseudomonas* research community.

An example of a GBrowse view of a genomic region is illustrated in Figure 3 and a list of all currently available tracks of annotations is described in Table 1.

7. INTEGRATION WITH OTHER DATA VIEWS, SUCH AS PATHWAY-BASED VISUALIZATIONS

In addition to the gene/genome-centric views of the main *Pseudomonas* Genome Database, a metabolic pathway-based view of the data called PseudoCyc is also maintained. It is based on a database schema developed by Karp *et al.*[17] using their Pathway Tools software. PseudoCyc, created by Romero and Karp in 2003[27] (curation taken over by PseudoCAP in 2004[35]), contains information on *Pseudomonas aeruginosa* PAO1 genes and proteins along with 821 enzymatic reactions and 141 biochemical pathways involving 738 enzymes and 651 compounds. Updates to gene and protein names that have changed since the original annotation in 2000 have also been updated in PseudoCyc, though annotation updates that were made during the initial curation of the PseudoCyc pathways[27] have been maintained. PseudoCyc is fully integrated with the *Pseudomonas* Genome Database and contains additional interactive functionalities including the ability to upload and view gene expression data in the context of the pathways. Eventually it would be suitable to expand this database to include signaling pathways, as is now occurring for other selected Pathway Tools-based databases.

8. EXAMPLES OF ANALYSIS TOOLS ASSOCIATED WITH THIS DATABASE: UTILITY FOR RESEARCH AND DRUG/VACCINE DISCOVERY

There are many tools and analyses that we are interested in incorporating into this database, including functional genomics data, other computational predictions (particularly medically relevant ones, such as improved epitope prediction), and other inter-genome comparison tools. Below is a short highlight of some tools/analyses currently available, illustrating how diverse types of analyses/tool are being made available, some of which provide insight into *Pseudomonas* biology and some of which may also be useful for development of *Pseudomonas* infection controls.

Table 1. GBrowse tracks of annotation data that is currently available.

GBrowse track	Database version that track was first available for *P. aeruginosa* PAO1	Track available for other *Pseudomonas* genomes
All genes with links to www.Pseudomonas.com	1	Yes
All Proteins with NCBI links	1	Yes
Operon Finding Software V1.2[34] Predictions (probability that adjacent genes on the same strand are cotranscribed)[b]	1	No[a]
Putative *Pseudomonas putida* KT2440 orthologs ("reciprocal BLAST analysis" only)	1	Yes[c]
Putative *Pseudomonas syringae* DC3000 orthologs ("reciprocal BLAST analysis" only)	1	Yes[c]
Putative *Pseudomonas syringae* B728a orthologs ("reciprocal BLAST analysis" only)	2	Yes[c]
Putative *Pseudomonas aeruginosa* PAO1 orthologs ("reciprocal BLAST analysis" only)	N/A	Yes
All COGs[32] with links to NCBI COG database	2	Yes
PFAM predictions (with links to PFAM database)[28]	2	Yes
Predicted TIGRFAMs (with links to TIGRFAMs)[24]	2	Yes
Predicted TIGR Roles and Sub Roles[24]	2	Yes
Subcellular localization (with confidence level)[10,11,18]	2	Yes[d]
Non-Coding RNA Genes	1	No
KEGG Metabolic Pathways[23]	1	Yes
PRODORIC Regulons[20]	1	No
Protein Data Bank (Hits to known 3D structures by PEDANT)[9]	1	No[a]
Intergenic Sequences (for ease of retrieval)	1	No
Rho-independent Terminators (from CMR)[7]	1	No[a]
Rho-independent Terminators (additional terminators identified by the Brinkman group)[35]	1	No
phoA/lacZ transposon mutants (University of Washington Genome Centre)[16]	1	No
TN5 lux transposon mutants (UBC)[19]	1	No
DNA sequence (at high resolution <= 130 bp)/% GC Content (at low resolution > 130 bp)[30]	1	Yes
3-frame translation (forward)[30]	1	Yes
3-frame translation (reverse)[30]	1	Yes

[a] These tracks will likely be implemented by the end of 2006.

[b] More accurately, this track containing operon predictions could be referred to as the probability that two adjacent genes encoded on the same strand do not contain a promoter between them. Note that it is possible to have two genes be cotranscribed but also have an additional promoter located between the two genes to further modulate expression of the downstream gene.

[c] Not applicable for the GBrowse database for the same *Pseudomonas* strain as is listed here for this track

[d] Only annotations computationally predicted by PSORTb are provided for other *Pseudomonas* genomes, while the *P. aeruginosa* PAO1 genome annotation additionally contains information about localizations experimentally confirmed in laboratory experiments[18,35].

8.1. Motif Search

We have created a user-friendly search tool that can be used to find specified DNA motifs within the strain PAO1 genomic DNA sequence[35]. This search tool accepts IUPAC-formatted variable length DNA sequence motifs as a query, and then, upon search completion, an online report or downloadable tab-delimited file is produced containing information on all regions the motif is found in. We used this tool to discover a previously undescribed rho-independent terminator subset containing a common sequence string (Sequence: AAAGC{3,4}SN{5,30}SGGGCTTT; occurrences not previously reported in the CMR terminator GBrowse track are viewable under the Brinkman terminator track under GBrowse)[35]. The *P. aeruginosa* PAO1 genome had the highest total number of occurrences of this terminator subset compared to all complete genomes available, with related *Pseudomonas* species genomes containing slightly higher than average occurrences as a proportion of their genome, perhaps reflecting an evolutionary relationship in the evolution of this terminator sequence. The discovery of this surprisingly terminator-specific and species-specific sequence motif is just one example of how integration of such elementary tools has led to new insights through *Pseudomonas* genome analysis.

8.2. Subcellular Localization Predictions

The protein subcellular localization information that we have recently made available is of particular note since subcellular localization information can be an important component of vaccine/drug target identification, and is critical for many downstream protein studies. We previously developed the most precise predictor of subcellular localization, termed PSORTb[11] which the project has used to make predictions of localization, in addition to laboratory-derived data that we have obtained from the literature[26] and experiments[18]. This PSORTb method was updated in 2005, permitting more predictions to be made while retaining the same high level of precision (greater than 95%)[10]. We have found that the level of accuracy is now exceeding that of most high-throughput laboratory studies of subcellular localization. This resource is an example of an analysis that can aid many researchers in identifying proteins of interest (such as cell-surface proteins) that are not well studied, since frequently predictions of subcellular localization can be made for proteins of unknown function. For example, using the database search, one can easily identify all proteins predicted to be secreted which are annotated as "hypothetical protein".

We hope to incorporate more such analyses into the database, focusing primarily on those that have high precision, to ensure that a quality database is maintained. In addition, this database project will move further toward using more advanced data mining techniques to perform more sophisticated analyses of the genomic data, along with incorporating more analyses specifically

Box 4. Database "Helpful Hints"

- Click on the "blue question mark" symbol near anything in the database, if you want to learn more about it.
- If you can't find something during a search, try a "wildcard" search (which allows text before and after your query text). For example, you may search for "urease" under product name for the strain PAO1 annotation, but not find anything because the name is actually "urease alpha subunit". If you search for "urease" with a wildcard specification though, you will find it.
- When searching by gene name, try to search under both gene name and alt gene name.
- Remember, you can combine searches under the advanced search option. For example, identifying all those proteins predicted to be outer membrane proteins which are not class 1 (i.e. have not been experimentally studied).
- Additional GBrowse annotations are best reached by first going to a relevant gene entry in the main database (usually through a search), and then clicking on the GBrowse link.
- In GBrowse, you can turn on and off what annotation tracks you wish to view, and sort them in a different order. If you zoom into the sequence enough, some tracks change to show additional information, such as more descriptive names, or nucleotide sequence (for the G+C track).
- In GBrowse, here are tracks of information about orthologs between different *Pseudomonas* strains, and you can click on these to view data on the orthologous gene (gene that is most likely the "same" in the other species) and therefore move between genomes easily through these orthologs.
- For some genome sequences, there may be little gene name information, and annotation quality/degree of detail will vary markedly. Therefore, just because you don't see a ureC gene in one genome for example, that doesn't mean its not there. You should always ensure that you use complementary sequence search capabilities, in addition to annotation text searches.
- If you have a very large search you want to perform (i.e. that would result in greater than 1000 "hits" in your search) you should consider either contacting us to perform a custom search for you, or download a copy of the annotations so you may perform the query on your own machine. Very large searches = very slow on a website that gets many queries in a day as this database does. Note that we do not allow excessively large searchers, regardless (an automated message will appear that will indicate that the search is too big and you should contact us for help).

relevant to uncovering new approaches for *P. aeruginosa* infection control (immune modulators, epitope identification, small molecule interactions, etc).

9. CONCLUSIONS

The initial sequencing and annotation of the *P. aeruginosa* PAO1 genome sequence was a landmark event for *Pseudomonas* research. It revealed that this microbe contained a complex metabolic capacity with surprisingly large gene families identified that explain in part the versatility of this organism. Subsequent laboratory and computational studies of the encoded genes, proteins, and regulatory regions are revealing a further level of complexity. Such analyses are also identifying notable new targets for new therapeutics, vaccines, diagnostics, as well as providing insights into *P. aeruginosa* function, antimicrobial resistance, and pathogenicity. There is much promise in *P. aeruginosa* research today[1,8,13,22,36,37].

This database of genome annotation information was initiated with the simple goals of providing a user-friendly resource that capitalized on input from the *Pseudomonas* research community. This community input has continued, with the number of PseudoCAP participants increasing from 61 in the year 2000 to 107 as of the end of 2004. These individuals, from over 11 countries worldwide, have contributed over 1000 annotation updates, not including the submissions made prior to the genome sequence publication, or particular high-throughput analyses. However, additional annotation updates made through review of recent peer-reviewed publications were also useful and are increasingly forming a larger proportion of annotation updates being made. To make such literature review as effective as possible, it is critical that researchers clearly indicate which strain is being used in a given study, so that through text searches, strain-specific studies can be identified and any associated annotation updated. Of course it only benefits the community if individuals notify the genome database coordinator or manager of any suitable changes/corrections as they notice them.

Updates such as nucleotide sequence corrections reinforce the need for having a clear unique identifier for all genes in a genome, since the nucleotide coordinates can easily change for all genes downstream of corrections that involve an insertion or deletion. We have utilized the PA number, with its additional decimal system as described above, as a primary key, and discourage the use of the gene name as a primary identifier, due to its potential to change in some cases, as knowledge of a gene's function increases. We therefore encourage researchers to always reference the PA number (or similar number in another strain) for a particular study, while often continuing to also use the gene name as a useful term because of its functional meaning.

This database is centered around the philosophy that both centralized annotation, as well as access to decentralized annotations, is critical. Continually

updated annotation approaches that do not have a reference annotation that is subject to review are susceptible to increased confusion, as researchers sort through which annotation information to trust or report in the context of their laboratory or computational analysis. Continually updated annotation approaches that only provide a single primary reference annotation run the risk of stifling alternate biological views. Of course, static annotations that are not subject to updating risk becoming obsolete and the workload involved in high-quality, re-curation of whole genomes can render occasional re-annotation of a whole genome unfeasible, in contrast to the incremental approach we perform as we keep abreast of the research literature and PseudoCAP submissions. In addition, we propose strong links between different *Pseudomonas* databases, and have started to implement this as part of the comparative genomics component of the database. Each *Pseudomonas* database has unique expertise to offer, and through the power of internet links, these expertises can be integrated at each site, potentially focusing on slightly different viewpoints suitable for the needs of particular *Pseudomonas* communities. As significantly more genomes are sequenced, comparative genomics tools will become increasingly necessary if we are to manage the continuing flood of genomic data.

PseudoCAP and the *Pseudomonas* Genome Database has hopefully facilitated the collaboration of *Pseudomonas* researchers, capitalized on their experience, and stimulated interaction and collaboration while furthering our understanding of *Pseudomonas* biology. Through the centralized peer-review process, we have been able to attain a high-quality genome annotation for a model *P. aeruginosa* strain that takes into account levels of confidence for each annotation. This has been achieved with minimal cost by taking advantage of the internet as an effective means of collecting, distributing, and analyzing information. Through its availability as an open source package, with additional utilization of other open source packages such as GBrowse that are already available, this database, and the associated continually updated annotation approach will potentially act as base for additional *Pseudomonas* genome projects in the future. The *P. aeruginosa* PAO1 genome sequence and database is clearly only the beginning of many *Pseudomonas* genome projects. Many interesting strains still need to be sequenced from this facilitating genus, and there is optimism that we will make great strides in understanding *P. aeruginosa* and other *Pseudomonas* through a combination of sequence analysis, other biological study, and integration of the resulting data.

ACKNOWLEDGMENTS

The author is indebted to Robert E. W. Hancock, Geoff Winsor, and Raymond Lo for their critical coordination, guidance, and database development efforts. All members of the *Pseudomonas* research community who provided

important input on this project, annotation information, and corrections are thankfully acknowledged – in particular those involved in galvanizing the initial CAP. Matthew Laird and other members of the Brinkman Laboratory are thanked for their critical input regarding database development, and Peter Karp and Randy Gobbel are acknowledged for their assistance with implementing the necessary changes to PseudoCyc for its integration. Primary funding for this project was provided by the Cystic Fibrosis Foundation. Ancillary funding was provided by Genome Prairie, Genome BC, Inimex Pharmaceuticals and the Canadian Cystic Fibrosis Foundation. The author is a Michael Smith Foundation for Health Research Scholar and a Canadian Institutes of Health Research New Investigator.

REFERENCES

1. Arnold, H., Bumann, D., Felies, M., Gewecke, B., Sorensen, M., Gessner, J.E., Freihorst, J., von Specht, B.U., and Baumann. U., 2004, Enhanced immunogenicity in the murine airway mucosa with an attenuated Salmonella live vaccine expressing OprF-OprI from *Pseudomonas aeruginosa*. *Infect. Immun.,* 72:6546–6553.
2. Ashburner, M., Ball, C.A., Blake, J.A., D. Botstein, Butler, H., Cherry, JM., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., Harris, M.A., Hill, D.P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J.C., Richardson, J.E., Ringwald, M., Rubin, G.M., and Sherlock. G., 2000, Gene ontology: tool for the unification of biology. *The Gene Ontology Consortium. Nat. Genet*., 25:25–29.
3. Brinkman, F.S., Hancock, R.E., and Stover, C.K., 2000, Sequencing solution: use volunteer annotators organized via Internet. *Nature*, 406:933.
4. Buell, C.R., Joardar, V., Lindeberg, M., Selengut, J., Paulsen, I.T., Gwinn, M.L., Dodson, R.J., Deboy, R.T., Durkin, A.S., Kolonay, J.F., Madupu, R., Daugherty, S., Brinkac, L., Beanan, M.J., Haft, D.H., Nelson, W.C., Davidsen, T., Zafar, N., Zhou, L., Liu, J., Yuan, Q., Khouri, H., Fedorova, N., Tran, B., Russell, D., Berry, K., Utterback, T., Van Aken, S.E., Feldblyum, T.V., D'Ascenzo, M., Deng, W.L., Ramos, A.R., Alfano, J.R., Cartinhour, S., Chatterjee, A.K., Delaney, T.P., Lazarowitz, S.G., Martin, G.B., Schneider, D.J., Tang, X., Bender, C.L., White, O., Fraser, C.M., and Collmer. A., 2003, The complete genome sequence of the Arabidopsis and tomato pathogen *Pseudomonas* syringae pv. tomato DC3000. *Proc. Natl. Acad. Sci. USA,* 100:10181–10186.
5. D'Ascenzo, M.D., Collmer, A., and Martin, G.B., 2004, PeerGAD: a peer-review-based and community-centric web application for viewing and annotating prokaryotic genome sequences. *Nucleic Acids Res.,* 32:3124–3135.
6. Dowell, R.D., Jokerst, R.M., Day, A., Eddy, S.R., and Stein. L., 2001, The distributed annotation system. *BMC Bioinformatics,* 2:7.
7. Ermolaeva, M.D., Khalak, H.G., White, O., Smith, H.O., and Salzberg. S.L., 2000, Prediction of transcription terminators in bacterial genomes. *J. Mol. Biol.,* 301:27-33.
8. Erwin, A.L., and VanDevanter. D.R., 2002, The *Pseudomonas aeruginosa* genome: how do we use it to develop strategies for the treatment of patients with cystic fibrosis and *Pseudomonas* infections? *Curr. Opin. Pulm. Med.,* 8:547–551.
9. Frishman, D., Mokrejs, M., Kosykh, D., Kastenmuller, G., Kolesov, G., Zubrzycki, I., Gruber, C., Geier, B., Kaps, A., Albermann, K., Volz, A., Wagner, C., Fellenberg, M., Heumann, K., and Mewes, H.W., 2003, The PEDANT genome database. *Nucleic Acids Res*., 31:207–211.

10. Gardy, J.L., Laird, M.R., Chen, F., Rey, S., Walsh, C.J., Ester, M., and Brinkman. F.S., 2005, PSORTb v.2.0: expanded prediction of bacterial protein subcellular localization and insights gained from comparative proteome analysis. *Bioinformatics,* 21:617–623.
11. Gardy, J.L., Spencer, C., Wang, K., Ester, M., Tusnady, G.E., Simon, I., Hua, S., deFays, K., Lambert, C., Nakai, K., and Brinkman, F.S., 2003, PSORT-B: Improving protein subcellular localization prediction for Gram-negative bacteria. *Nucleic Acids Res*., 31:3613–3617.
12. Glasner, J.D., Liss, P., Plunkett 3rd G, Darling, A., Prasad, T., Rusch, M., Byrnes, A., Gilson, M., Biehl, B., Blattner, F.R., and Perna. N.T., 2003, ASAP, a systematic annotation package for community analysis of genomes. *Nucleic Acids Res*., 31:147–151.
13. Goodman, A.L. and Lory. S., 2004, Analysis of regulatory networks in *Pseudomonas aeruginosa* by genomewide transcriptional profiling. *Curr. Opin. Microbiol.*, 7:39-44.
14. Hokamp, K. and Wolfe. K., 1999, What's new in the library? What's new in GenBank? let PubCrawler tell you. *Trends Genet*., 15:471–472.
15. Hubbard, T. and Birney. E., 2000, Open annotation offers a democratic solution to genome sequencing. *Nature,* 403:825.
16. Jacobs, M.A., Alwood, A., Thaipisuttikul, I., Spencer, D., Haugen, E., Ernst, S., Will, O., Kaul, R., Raymond, C., Levy, R. Chun-Rong, L., Guenthner, D., Bovee, D., Olson, M.V., and Manoil. C., 2003, Comprehensive transposon mutant library of *Pseudomonas aeruginosa*. *Proc. Natl. Acad. Sci. USA,* 100:14339–14344.
17. Karp, P.D., Paley, S., and Romero. P., 2002, The pathway tools software. *Bioinformatics,* 18(Suppl 1):S225–32.
18. Lewenza, S., Gardy, J.L., Brinkman, F.S., and Hancock, R.E., 2005, Genome-wide identification of *Pseudomonas aeruginosa* exported proteins using a consensus computational strategy combined with a laboratory-based PhoA fusion screen. *Genome Res*., 15:321–329.
19. Lewenza, S., Falsafi, R.K., Winsor, G., Gooderham, W.J., McPhee, J.B., Brinkman, F.S., and Hancock. R.E., 2005, Construction of a mini-Tn5-luxCDABE mutant library in *Pseudomonas aeruginosa* PAO1: a tool for identifying differentially regulated genes. *Genome Res*., 15:583–589.
20. Munch, R., Hiller, K., Barg, H., Heldt, D., Linz, S., Wingender, E., and Jahn. D., 2003, PRODORIC: prokaryotic database of gene regulation. *Nucleic Acids Res.*, 31:266–269.
21. Nelson, K.E., Weinel, C., Paulsen, I.T., Dodson, R.J., Hilbert, H., Martins dos Santos, V.A., Fouts, D.E., Gill, S.R., Pop, M., Holmes, M., Brinkac, L., Beanan, M., DeBoy, R.T., Daugherty, S., Kolonay, J., Madupu, R., Nelson, W., White, O., Peterson, J., Khouri, H., Hance, I., Chris Lee P., Holtzapple, E., Scanlan, D., Tran, K., Moazzez, A., Utterback, T., Rizzo, M., Lee, K., Kosack, D., Moestl, D., Wedler, H., Lauber, J., Stjepandic, D., Hoheisel, J., Straetz, M., Heim, S., Kiewitz, C., Eisen, J.A., Timmis, K.N., Dusterhoft, A., Tummler, B., and Fraser., C.M., 2002, Complete genome sequence and comparative analysis of the metabolically versatile *Pseudomonas* putida KT2440. *Environ. Microbiol*., 4:799–808.
22. Nouwens, A.S., Walsh, B.J., and Cordwell., S.J., 2003, Application of proteomics to *Pseudomonas aeruginosa*. *Adv. Biochem. Eng. Biotechnol.,* 83:117–140.
23. Ogata, H., Goto, S., Sato, K., Fujibuchi, W., Bono, H., and Kanehisa. M., 1999, KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.,* 27:29-34.
24. Peterson, J.D., Umayam, L.A., Dickinson, T., Hickey, E.K., and White. O., 2001, The Comprehensive Microbial Resource. *Nucleic Acids Res.,* 29:123–125.
25. Pruitt, K.D. and Maglott. D.R., 2001, RefSeq and LocusLink: NCBI gene-centered resources. *Nucleic Acids Res*., 29:137–140.
26. Rey, S., Acab, M., Gardy, J.L., Laird, M.R., deFays, K., Lambert, C., and Brinkman, F.S., 2005, PSORTdb: a protein subcellular localization database for bacteria. *Nucleic Acids Res*. 33:D164–8.
27. Romero, P. and Karp. P., 2003, PseudoCyc, a pathway-genome database for *Pseudomonas aeruginosa*. *J. Mol. Microbiol. Biotechnol.,* 5:230–239.

28. Sonnhammer, E.L., Eddy, S.R., and Durbin. R., 1997, Pfam: a comprehensive database of protein domain families based on seed alignments. *Proteins,* 28:405–420.
29. Stein, L., Sternberg, P., Durbin, R., Thierry-Mieg, J., and Spieth. J., 2001, WormBase: network access to the genome and biology of Caenorhabditis elegans. *Nucleic Acids Res*., 29:82–86.
30. Stein, L.D., Mungall, C., Shu, S., Caudy, M., Mangone, M., Day, A., Nickerson, E., Stajich, J.E., Harris, T.W., Arva, A., and Lewis. S., 2002, The generic genome browser: a building block for a model organism system database. *Genome Res*., 12:1599–1610.
31. Stover, C.K., Pham, X.Q., Erwin, A.L., Mizoguchi, S.D., Warrener, P., Hickey, M.J., Brinkman, F.S., Hufnagle, W.O., Kowalik, D.J., Lagrou, M., Garber, R.L., Goltry, L., Tolentino, E., Westbrock-Wadman, S., Yuan, Y., Brody, L.L., Coulter, S.N., Folger, K.R., Kas, A., Larbig, K., Lim, R., Smith, K., Spencer, D., Wong, G.K., Wu, Z., Paulsen, I.T., Reizer, J., Saier, M.H., Hancock, R.E., Lory, S., and Olson. M.V., 2000, Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen. *Nature*, 406:959–964.
32. Tatusov, R.L., Galperin, M.Y., Natale, D.A., and Koonin. E.V., 2000, The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res*., 28:33–36.
33. Thompson, JD., Higgins, D.G., and Gibson. T.J., 1994, CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.,* 22:4673–4680.
34. Westover, B.P., Buhler, J.D., Sonnenburg, J.L., and Gordon. J.I., 2005, Operon prediction without a training set. *Bioinformatics,* 21:880–888.
35. Winsor, G.L., Lo, R., Sui, S.J., Ung, K.S., Huang, S., Cheng, D., Ching, W.K., Hancock, R.E., and Brinkman. F.S., 2005, *Pseudomonas aeruginosa* Genome Database and PseudoCAP: facilitating community-based, continually updated, genome annotation. *Nucleic Acids Res.,* 33:D338–43.
36. Woods, D.E., 2004. Comparative genomic analysis of *Pseudomonas aeruginosa* virulence. *Trends Microbiol.*, 12:437–439.
37. Worgall, S., Krause, A., Rivara, M., Hee, K.K., Vintayen, E.V., Hackett, N.R., Roelvink, P.W., Bruder, J.T., Wickham, T.J., Kovesdi, I., and Crystal. R.G., 2005, Protection against *P. aeruginosa* with an adenovirus vector containing an OprF epitope in the capsid. *J. Clin. Invest*., 115:1281–1289.

INDEX