INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

UMI[®]

**IDENTIFICATION AND CHARACTERIZATION OF CbaR,
A REPRESSOR OF *cbaABC*, THE 3-CHLOROBENZOIC ACID CATABOLIC
GENES OF *Comamonas testosteroni* BR60 (pBRC60)**

By

Miguel A. Providenti

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

Ottawa-Carleton Institute of Biology
Carleton University
Ottawa, Ontario

© 2000 Miguel Providenti

National Library
of Canada

Acquisitions and
Bibliographic Services

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale
du Canada

Acquisitions et
services bibliographiques

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file *Votre référence*

Our file *Notre référence*

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-48351-7

The undersigned hereby recommend to
the Faculty of Graduate Studies and Research
acceptance of the thesis

**Identification and Characterization of CbaR, A Repressor of
cbaABC, the 3-Chlorobenzoic Acid Catabolic Genes of
Camamonas testosteroni BR60 (pBRC60)**

Submitted by

Miguel Providenti, B.Sc., M.Sc.

in partial fulfilment of the requirements
for the degree of Doctor of Philosophy

Chair, Department of Biology

Thesis Supervisor

External Examiner

Carleton University

ABSTRACT

Regulation of *cbaABC*-mediated catabolism of 3-chlorobenzoate (3Cba) was characterized. These genes code for enzymes that convert 3Cba to protocatechuate (Pca) and 5-chloro-Pca and are found on transposon Tn5271 (located on conjugative plasmid pBRC60) from *Comamonas testosteroni* BR60. O₂ uptake experiments and expression studies with various *cbaA-lacZ* constructs showed that 3Cba catabolism is inducible by 3Cba, the substrate of the pathway, as well as benzoate, a compound that is not metabolized by this pathway and is therefore a gratuitous inducer. Determination of the *cbaA* transcriptional start site by primer extension and studies with the aforementioned expression constructs suggest that *cbaA* has a σ^{70} -type promoter. The gene expression studies also show that *cbaA* transcription is controlled by at least two regulators: an unidentified, chromosomally-encoded protein that induces *cbaABC* in response to benzoate or 3Cba; and CbaR, a Tn5271-encoded repressor linked to *cbaABC* that modulates gratuitous induction of *cbaA* by benzoate. According to gel shift and DNaseI protection assays, CbaR binds specifically to two regions of the *cbaA* promoter. The first region (Site I) is between the transcriptional start site (+1) and the *cbaA* start codon, while the second (Site II) overlaps +1. CbaR forms two complexes with *cbaA* promoter DNA, referred to as C1 and C2. C1 reflects CbaR bound to Site I while C2 reflects CbaR bound to Site I and II. CbaR shows different apparent dissociation constants ($K_{d_{app}}$) towards these sites, binding to Site I relatively strongly ($K_{d_{app}} \sim 10^{-8}$ M) and to Site II less strongly ($K_{d_{app}} \sim 10^{-6}$ M). CbaR binding can be reversed by 3Cba but not benzoate. A wide variety of other substituted benzoates were also tested for their effect on CbaR binding and several other binding inhibitors were identified (e.g. benzoate with a nitro or methyl group in position 2 or 3; or dichlorinated benzoates with at least one *meta* chlorine [i.e. in position 3 or 5]). Interestingly, compounds that improved binding were also identified (e.g. benzoate with a hydroxy or carboxy in position 3). Based on these experiments, it would appear that Tn5271 evolved such that expression of the *cba* pathway relies on a chromosomally-encoded regulator that induces activity in response to specific aromatic compounds, some of which are not substrates for *cbaABC*. However, to minimize gratuitous induction of the genes, Tn5271 retains the gene for CbaR.

ACKNOWLEDGEMENTS

Without a doubt, I would never have accomplished a fraction of what I am presenting in this thesis if not for the contributions of the following people. First and foremost, I would like to thank my supervisor, Dr. R. Campbell Wyndham (whose mentorship, support, and patience throughout the last five years will always inspire me in how I approach future work) and Dr. Suzanne Paterson (for her invaluable assistance with a variety of crucial experiments). I would also like to thank Drs. Myron Smith and Doug Johnson, official members of my advisory committee, and Dr. John Vierula, whom I consider to be an unofficial advisor, for their excellent guidance and kind donations of time.

I would also like to thank the following people from the Biology Department for their frequent assistance and most importantly their friendship, which made my time here extremely pleasant. These include colleagues (Christine Morgan, Michelle Peel, Sarah Rosolen, Dorothy McColl, Jacqui Whiteway, Christina Micali, Reza Nokbeh, Craig Carrol, Rashdan Rashid, Nadine Auger), staff (Joan Mallet, Cathy Merrit, Ed Bruggink, Art Goodenough, Glen Kit, Mike Weber, Adam Baker, Angela Kiss, Dave Omond), and faculty (Drs. Iain Lambert, George Carmody, Byron Johnson, and Jim Cheetham).

Lastly, I thank Allison, whose motivating influence ensured that I complete this thesis on time, and Isabella Aida, who waited until after the hardest part was over before popping out into the world and enriching my life in ways I never imagined.

DEDICATION

To my father

For his unconditional love and unstinting support.

He has always been there for me and I will never forget this.

TABLE OF CONTENTS

ACCEPTANCE SHEET.....	ii
ABSTRACT.....	iii
ACKNOWLEDGEMENTS.....	iv
DEDICATION.....	v
TABLE OF CONTENTS.....	vi
LIST OF FIGURES.....	x
LIST OF TABLES.....	xiii
LIST OF APPENDICES.....	xiv
LIST OF ABBREVIATIONS.....	xv
INTRODUCTION.....	1
CHAPTER 1: LITERATURE REVIEW.....	2
REGULATION OF GENE EXPRESSION IN BACTERIA.....	2
Regulation at initiation of transcription.....	2
σ -factors and promoter recognition.....	3
σ^{70} promoters.....	4
σ^{54} promoters.....	7
Global regulatory networks.....	8
Regulation after transcription initiation.....	9
Termination of transcription.....	9
Antisense RNA.....	10
ORGANIZATION AND REGULATION OF CATABOLIC GENES FOR AROMATIC COMPOUNDS.....	10
Overview of aerobic catabolic pathways for aromatic compounds.....	10
Catabolic operons regulated by LysR-type transcriptional regulators.....	14
Catabolic operons regulated by two-component systems.....	21
Catabolic operons dependent on σ^{54} and controlled by NtrC-like regulators.....	25
Catabolic operons regulated by GylR/IciR-like transcriptional regulators.....	28
Catabolic operons regulated by AraC-like transcriptional regulators.....	33
Catabolic operons regulated by repressors.....	33
Catabolic pathways dependent on two or more regulatory factors.....	36
Concluding statements.....	41

RESEARCH OBJECTIVES AND RATIONALE FOR PROJECT	43
Degradation of chlorinated aromatic compounds and the <i>cba</i> pathway.....	43
Rationale for the project.....	47
Research objectives.....	47
CHAPTER 2: CONTROL OF <i>cbaABC</i> -MEDIATED CATABOLISM OF 3- CHLOROBENZOATE	50
INTRODUCTION.....	50
MATERIALS AND METHODS	53
Bacterial strains, plasmids, and <i>E. coli</i> growth conditions	53
Classification of strain BR60 (pBRC60)	53
DNA manipulation, sequencing, and sequence analysis	53
Determination of <i>cbaA</i> transcriptional start site	54
Regulation of 3Cba degradative activity	54
Determining whether <i>cbaABC</i> -proficient cells convert benzoate to a vicinal diol intermediate	56
Construction of gene expression constructs	56
<i>cbaA-lacZ</i> constructs	56
<i>cbaR-lacZ</i> constructs	60
Regulation of <i>cbaA</i> and <i>cbaR</i> expression.....	61
Insertion of expression constructs.....	61
Confirmation that constructs inserted as a single copy via transposition	62
Measurement of gene expression levels	62
RESULTS	67
Classification of strain BR60 (pBRC60)	67
Identification of <i>cbaR</i> and determination of <i>cbaA</i> transcriptional start site	67
Induction of 3Cba degradative activity	72
Determining whether <i>cbaABC</i> -proficient cells convert benzoate to a vicinal diol intermediate	74
Regulation of <i>cbaA</i> and <i>cbaR</i> expression.....	74
Confirmation that single copies of expression constructs transposed into BR6020 chromosome	74
Expression of <i>cbaA</i>	77
Expression of <i>cbaR</i>	84

DISCUSSION AND SUGGESTIONS FOR ADDITIONAL RESEARCH	85
Classification of strain BR60 (pBRC60)	85
The <i>cbaA</i> promoter	85
Substrates that induce 3Cba degradative activity and <i>cbaA</i> expression	88
Growth on 3Cba or benzoate induces degradative activity for other aromatic compounds	89
Regulators controlling <i>cbaA</i> expression	89
General statements	89
CbaR: role in controlling <i>cbaA</i> , characterization of <i>cbaR</i> expression, and comparison to other MarR-like proteins	90
CbaS: speculation on identity and role in controlling aromatic catabolic pathways	96
CHAPTER 3: FUNCTIONAL CHARACTERIZATION OF CbaR, A REPRESSOR OF THE <i>cba</i> PATHWAY	98
INTRODUCTION	98
MATERIALS AND METHODS	99
Bacterial strains and plasmids	99
Identification and alignment of CbaR homologues	99
DNA manipulation, PCR amplification conditions, and sequencing	99
Cloning of <i>cbaR</i> , over-expression, and purification of CbaR	100
Gel shift assays	102
Binding of CbaR to DNA templates containing P _{<i>cbaR</i>} or P _{<i>cbaA</i>}	102
Interactions of CbaR with P _{<i>cbaA</i>} DNA	102
DNaseI protection assays	106
RESULTS	109
CbaR is a MarR-like protein	109
Expression and purification of CbaR	114
Binding of CbaR to DNA templates containing P _{<i>cbaR</i>} or P _{<i>cbaA</i>}	114
Stoichiometry of CbaR binding to P _{<i>cbaA</i>} DNA and lowest concentration at which 3Cba disrupts CbaR binding	120
Effectors of CbaR binding to P _{<i>cbaA</i>} DNA	123
Binding sites of CbaR	134

DISCUSSION AND SUGGESTIONS FOR ADDITIONAL RESEARCH	139
Comparison of CbaR primary sequence to other MarR-family members and structure information	139
Expression of recombinant CbaR.....	143
Nature and specificity of CbaR binding to <i>cbaR-cbaA</i> intergenic region.....	144
The effect of aromatic compounds on CbaR binding to P _{cbaA}	146
General statements	146
Effectors that inhibit binding.....	148
Effectors that improve binding.....	150
Role of CbaR in controlling <i>cbaA</i> expression.....	151
CHAPTER 4: ATTEMPT TO CLONE GENE FOR CbaS, THE PROPOSED ACTIVATOR OF OF THE <i>cba</i> PATHWAY	153
INTRODUCTION.....	153
MATERIALS AND METHODS	156
Bacterial strains and plasmids.....	156
DNA sequence manipulation	156
General information.....	156
Generating transcriptional interference vector for CbaS	156
Generating a plasmid library of <i>C. testosteroni</i> BR6020 chromosomal DNA	159
Screening the plasmid library for clones containing CbaS.....	160
First round of screening.....	160
Second round of screening and Southern blot analysis	162
RESULTS	164
DISCUSSION AND SUGGESTION FOR FURTHER RESEARCH.....	167
CHAPTER 5: CONCLUSIONS	169
Proposed scenario for control of <i>cba</i> pathway.....	169
Evolution of the (mobilizable) <i>cba</i> pathway.....	172
LITERATURE CITED.....	176
APPENDICES	192

LIST OF FIGURES

Figure 1.1 Consensus sequence and organization of σ^{70} and σ^{54} promoters	6
Figure 1.2 Biodegradation pathways of various aromatic compounds funnel structurally diverse aromatic compounds towards catechol, protocatechuate, or gentisate.....	13
Figure 1.3 Catabolic genes for non-chlorinated aromatic compounds that are regulated by LysR-type transcriptional regulators	16
Figure 1.4 Catabolic genes for chlorinated aromatic compounds that are regulated by LysR-type transcriptional regulators	20
Figure 1.5 Catabolic genes for aromatic compounds that are regulated by two-component systems	24
Figure 1.6 Catabolic genes for degradation of aromatic compounds that possess σ^{54} -dependent promoters and are activated by NtrC-like proteins	27
Figure 1.7 Catabolic genes for aromatic compounds regulated by GylR/IciR-like regulators.....	30
Figure 1.8 Catabolic genes that are negatively regulated.....	35
Figure 1.9 Catabolic genes for aromatic compounds that are regulated by at least two regulatory elements	39
Figure 1.10 Degradation pathways for chlorinated aromatics channel these compounds towards dihydroxymetabolites.	45
Figure 1.11 The <i>cba</i> pathway of <i>C. testosteroni</i> BR60 (pBRC60)	49
Figure 2.1 Summary of degradation pathways for chlorobenzoates.....	52
Figure 2.2 Restriction map of 3-chlorobenzoate catabolic transposon Tn5271 and relative organization of genes and open reading frames	58
Figure 2.3 Nucleotide sequence of the upper strand of Tn5271 from the <i>Nsi</i> I site in IS1071L to the <i>Sca</i> I site in <i>cbaA</i>	69

Figure 2.4 Transcriptional initiation site of <i>cbaA</i> , as determined by primer extension analysis	71
Figure 2.5 Confirmation that single copies of expression constructs transposed into <i>C. testosteroni</i> BR6020.....	76
Figure 2.6 <i>cbaA</i> expression levels in <i>C. testosteroni</i> BR6020 containing a single chromosomal insert of the indicated <i>lacZ</i> construct	80
Figure 2.7. <i>cbaA</i> expression levels in <i>C. testosteroni</i> BR6020 with <i>cbaRA</i> '-' <i>lacZ</i> or <i>cbaA</i> '-' <i>lacZ</i> 1 following growth in medium containing succinate, succinate and benzoate, or succinate and 3-chlorobenzoate.....	83
Figure 2.8 Degradation pathways in <i>C. testosteroni</i> BR60 funnel various substituted benzoates towards protocatechuate, which is then metabolized by an extradiol ring fission pathway.....	87
Figure 2.9 Organization of various genes regulated by MarR-like transcriptional regulators, relative to the gene for the MarR-like proteins	95
Figure 3.1 Schematic representation of DNA templates used in gel shift assays	104
Figure 3.2 Alignment of CbaR to nineteen other MarR-like proteins.....	113
Figure 3.3 Expression and purification of histidine-tagged CbaR, as monitored by electrophoresis in a denaturing sodium dodecylsulfate gel.....	117
Figure 3.4 Gel shift assay to map area bound by CbaR	119
Figure 3.5 Gel shift assays to study stoichiometry of CbaR-P _{cbaA} DNA complex formation and minimum concentration at which 3-chlorobenzoate inhibits CbaR binding ..	122
Figure 3.6 Gel shift assays to study the effect on CbaR binding to P _{cbaA} DNA by mono-, di-, and some tri-chlorobenzoates; benzoate with a halogen in position 3; and several chlorobenzene compounds with various groups <i>meta</i> to chlorine.....	126
Figure 3.7 Gel shift assays to study the effect on CbaR binding to P _{cbaA} DNA by benzoate substituted with a carboxy group; sulfo or nitro group; and amino or methyl group	128

Figure 3.8 Gel shift assays to study the effect on CbaR binding to P _{cbaA} DNA by benzoate with no substituent or substituted with hydroxy groups; and methoxy and hydroxy groups.....	130
Figure 3.9 Gel shift assays to better characterize the effect of 3-carboxybenzoate on CbaR-P _{cbaA} DNA complex formation	132
Figure 3.10 DNaseI protection assay to determine specific region of P _{cbaA} DNA bound by CbaR.....	136
Figure 3.11 DNA sequence of the 186 bp template used in the DNaseI protection assay indicating the two regions protected by CbaR.....	138
Figure 3.12 Comparison of the CbaR operator to those of MarR and PecS	142
Figure 4.1 General strategy for selection of clones expressing proteins that bind to specific DNA sequences.....	155
Figure 4.2 Selection of putative CbaS DNA binding site and cloning into transcriptional interference vector	158
Figure 4.3 Restriction analysis of clones from a plasmid library of <i>C. testosteroni</i> BR6020 chromosomal DNA recovered in the first round of screening with transcriptional interference vector pNN388OP.1	166
Figure 5.1 Regulation of the <i>cba</i> pathway	171
Figure 5.2 Mobilization of <i>cba</i> pathway	175

LIST OF TABLES

Table 2.1. Bacterial strains and plasmids used for experiments described in Chapter 2.	64
Table 3. Oxygen consumption rates by <i>C. testosteroni</i> BR60 (pBRC60) towards various carbon sources.	73
Table 3.1 <i>E. coli</i> strains and plasmids used for experiments described in Chapter 3.....	108
Table 3.2. Proteins showing homology to CbaR or MarR.....	110
Table 3.3 Summary of the effects of various aromatic compounds on CbaR binding to P _{cbaA} DNA.....	133
Table 4.1 <i>E. coli</i> strains and plasmids used for experiments described in Chapter 4.....	163

LIST OF APPENDICES

Appendix A. Sequence of DNA coding for 16S rRNA from strain BR60 (pBRC60) and comparison to other sequences in the database	192
Appendix B. Schematic diagram of <i>cbaA-lacZ</i> constructs cloned into suicide delivery vector pUTCm.	193
Appendix C. Schematic diagram of <i>cbaR-lacZ</i> constructs and control <i>lacZ</i> construct cloned into suicide-delivery vector pUTCm or pUTCmMCS.3.	194
Appendix D. DNA sequence of junctions between ' <i>lacZ</i> and <i>cbaA</i> ' and <i>cbaR</i> '; and the junction between the DNA coding for the six-histidine tag and the N-terminus of CbaR.....	195

LIST OF ABBREVIATIONS

Amp, ampicillin

bp, basepair

C1, complex 1 formed between CbaR and P_{cbaA} DNA

C2, complex 2 formed between CbaR and P_{cbaA} DNA

Cba, chlorobenzoate

cba, chlorobenzoate catabolic pathway

cbaABC, genes coding for 3-chlorobenzoate dioxygenase (CbaAB) and the dehydrogenase for the vicinal diol intermediate (CbaC)

cbaR, gene for CbaR

CbaR, repressor of *cbaABC*

CbaS, putative activator of *cbaABC*

cm, centimetre

Cm, chloramphenicol

DNA, deoxyribonucleic acid

dNTP, deoxynucleotide triphosphate

ddUTP, dideoxyuridine triphosphate

EDTA, (ethylene dinitrilo)tetraacetic acid

g, gram

g, gravity

hrs, hours

IPTG, isopropyl- β -D-thiogalactopyranoside

K_{d_{app}}, apparent dissociation constant

Km, kanamycin

kb, kilobases

kbp, kilobasepairs

kDa, kiloDaltons

L, litre

LacZ, β -galactosidase

lacZ, β -galactosidase gene
 LB, Luria-Bertani medium
 LTTR, LysR type transcriptional regulator
 M, molar
 MarR, repressor of multiple antibiotic resistance genes
 mg, milligram
 mL, millilitre
 mM, millimolar
 MMA, minimal medium A
 mRNA, messenger RNA
 MCS, multiple cloning site
 MW, molecular weight
 μ g, microgram
 μ L, microlitre
 μ M, micromolar
 ng, nanogram
 nt, nucleotide
 NTA, nitrilotriacetate
 OD₆₀₀, optical density measured with 600 nm wavelength light
 ORF, open reading frame
 Pca, protocatechuate (3,4-dihydroxybenzoate)
 P_{cbmA}, the *cbmA* promoter
 PCR, polymerase chain reaction
pmd, protocatechuate *meta* dioxygenase genes/pathway
 pmol, picomol
 ppm, parts per million
 PSI, pounds per square inch
 r, resistant/resistance
 RNA, ribonucleic acid

RBS, ribosome binding site

rpm, revolutions per minute

SDS PAGE, sodium dodecylsulfate polyacrylamide gel electrophoresis

TBE, Tris base/boric acid/EDTA

TCA, tricarboxylic acid

Tris, Tris[hydroxymethyl]aminoethane

Trp, tryptophan

U, unit

UV, ultraviolet

v, volume

V, volts

w, weight

X-gal, 5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside

INTRODUCTION

This doctoral thesis describes a project investigating the regulation of *cbaABC*, genes from the bacterium *Comamonas testosteroni* BR60 (pBRC60) coding for the enzymes that convert 3-chlorobenzoate (3Cba) to protocatechuate and 5-chloroprotocatechuate. The thesis is divided into five chapters. Chapter 1 is a review of relevant literature. It initially provides a general overview of how gene expression is regulated in bacteria, follows with a description of the organization and regulation of catabolic pathways for aromatic compounds, and ends by describing the objectives of the project. Chapters 2 to 4 describe the research that was conducted for this thesis and each chapter contains its own introduction, materials and methods, results, discussion, and suggestions for additional research. Chapter 2 describes studies investigating regulation of 3Cba degradation, Chapter 3 describes the functional characterization of a repressor involved in regulating *cbaABC*, and Chapter 4 describes attempts to clone the gene for another protein that is believed to be involved in controlling *cbaABC*. Chapter 5 summarizes results of my research and provides a scenario for how *cbaABC* expression is controlled and some speculation on the evolution of the *cba* pathway.

CHAPTER 1: LITERATURE REVIEW

REGULATION OF GENE EXPRESSION IN BACTERIA

Bacteria in natural environments encounter different nutrients. The metabolic pathways that are active will be determined by the available nutrient sources. To optimize growth, expression of catabolic genes is tightly controlled. Because translation of mRNA occurs soon after gene transcription begins and the half-life of mRNA is very short, gene expression in bacteria is almost always controlled at the transcriptional level (Lodish and Darnell, 1995). Most often, control is exerted at the point of initiation, but in some cases, it is exerted after transcription has begun by prematurely terminating transcription (Rojo, 1999).

Regulation at initiation of transcription

DNA transcription is initiated by RNA polymerase holoenzyme, a five-subunit protein consisting of the core enzyme ($\alpha_2\beta\beta'$) and σ -factor. Although core enzyme can bind loosely to any DNA sequence, σ -factor encourages strong binding of holoenzyme to the promoters of genes. After binding, holoenzyme separates the DNA strands approximately 10 nucleotides downstream of the promoter, transcription begins, and σ -factor is released. Core enzyme is therefore responsible for chain elongation while σ -factor is responsible for recognition of the promoter (Lewin, 1987).

Various mechanisms have evolved in bacteria to control transcription initiation but genes are, in essence, either negatively regulated (repressible) or positively regulated (activatable). Negatively regulated genes are expressed unless something prevents transcription. Transcription initiation is inhibited by a repressor protein which binds to the operator region of the gene and either impedes access of RNA polymerase holoenzyme to the promoter or prevents it from initiating transcription. In practically all cases, the operator is at or adjacent to the promoter (see below). Transcription in some negatively regulated genes occurs when an effector molecule (inducer) decreases the affinity of the repressor for the operator. In others, transcription is prevented when an effector molecule (co-repressor) increases the affinity of a repressor for the operator (Lewin, 1987).

Positively regulated genes, in contrast, are not expressed unless a regulatory protein activates transcription. The activator protein (apoinducer) binds to a specific "activation" site on the DNA and promotes transcription initiation by RNA polymerase holoenzyme. Activation sites are located at various distances relative to promoters (discussed below). Expression in some positively regulated genes occurs when apoinducer is converted to an active form by an effector molecule. In others, expression is halted when apoinducer is converted to an inactive form by an effector molecule (Lewin, 1987).

An obvious question is whether the expression of genes coding for regulatory proteins is also subject to strict control. Expression of genes for regulatory proteins is controlled by two mechanisms. Some regulatory proteins are expressed constitutively but at a low level. This appears to be the case for many repressors. Repressors are inactivated when a surplus of effector molecules overwhelm the repressor protein, allowing derepression of negatively regulated genes. Other regulatory proteins directly control their own expression in a process known as autogenous regulation (Maloy and Stewart, 1993). The gene for the regulatory protein can be in the operon it regulates or in a separate operon. In some cases when it is in a separate operon, the gene for the regulatory protein is transcribed in the opposite direction from the operon it regulates and the regulatory protein simultaneously exerts control over its own expression and the target operon (Beck and Warren, 1988).

σ -factors and promoter recognition

Promoters can be divided into two classes: simple and complex (Collado-Vides *et al.*, 1991). A simple promoter is isolated from others and controlled by a single regulator. A complex promoter is either controlled by more than one regulator or is "multiple" (i.e. the regulatory region is part of another promoter and the two are controlled by a common regulator or the two transcribe the same gene).

Crucial to promoter recognition and transcription initiation is σ -factor. Even though a large variety of σ -factors have been discovered in bacteria, differing in molecular weight and the kinds of genes with which they are associated, two families of σ -factors are

recognized; the σ^{70} and σ^{54} family. They are distinguished by their phylogenetic unrelatedness, different mechanisms of action, and recognition of distinct promoter sequences (Lonetto *et al.*, 1992). Many different operons containing promoters from both families have been characterized extensively and the following sections describe major properties of each.

σ^{70} promoters

σ^{70} is the most abundant bacterial σ -factor and recognizes both simple and complex promoters. The consensus sequence of σ^{70} promoters includes a pair of hexamers found at approximately -10 and -35 nucleotides from the transcription initiation site, designated as +1 (Figure 1.1A). Most operons characterized to date have σ^{70} promoters and several features common to negatively and positively regulated operons have been uncovered.

In almost all negatively regulated operons containing a σ^{70} promoter, the operator region is proximal to the transcription initiation site, proximal defined as nucleotides spanning approximately -65 to +20 (Figure 1.1A). This highlights the importance of close contact between the repressor and RNA polymerase holoenzyme, which occupies DNA from nucleotides -45 to +20 (Collado-Vides *et al.*, 1991). Within a negatively regulated regulon (i.e. a set of distinct operons controlled by a common regulatory element), the degree of repression can differ. This is achieved either by changes in the position of the operator relative to the promoter (which will presumably affect how well the repressor interferes with RNA polymerase binding) or by changes to the sequence of the operator itself (which affects the affinity of the repressor for the operator [Collado-Vides *et al.*, 1991]).

Approximately half of characterized negatively regulated operons contain a duplicated operator, and two-thirds of these duplicated operators are proximal. Why some operators are duplicated is not certain. In operons containing simple promoters (see above), duplication may strengthen repression by enhancing the interference with RNA polymerase. This could be because repression by a single operator is not very strong. It

Figure 1.1 Consensus sequence and organization of A) σ^{70} and B) σ^{54} promoters.

The transcription initiation site is indicated with a +1. In σ^{70} promoters, proximal is defined as the area spanning +20 to -65. This represents the binding site for the cAMP-binding protein in the lac operon (Collado-Vides *et al.*, 1991). In σ^{54} promoters, the activator binding site (ABS) is an inverted repeat approximately 100 to 200 bp upstream of +1 (Morett and Segovia, 1993). Consensus sequences from Gross *et al.*, (1992). C) Activators of σ^{54} promoters are modular proteins and each domain performs a distinct function, as indicated. D) General organization of a σ^{54} -dependent gene and the gene for the activator, which may be transcribed divergently from its target genes. The activator attaches to the upstream inverted repeat. It contacts σ^{54} -RNA polymerase holoenzyme (intervening DNA loops out) and in response to some signal, ATP hydrolysis occurs (not indicated) and transcription is initiated.

A. σ^{70} promoter

B. σ^{54} promoter

may also allow a more flexible response to signals, so that the operon is not simply on or off. In operons containing complex promoters, which may be subject to multiple signals, stronger repression may not be necessary and duplications may increase flexibility. Each operator might be acting as its own regulatory unit, allowing a finer control of gene expression (Collado-Vides *et al.*, 1991).

In positively regulated operons, the activator binding site is, with only one exception, proximal to the promoter. Furthermore, the majority of these sites overlap the nucleotides close to the -35 hexamer recognized by σ^{70} , while none span the -10 hexamer. This contrasts to repressor binding sites, which are located at various positions relative to the transcription initiation site (see above; [Collado-Vides *et al.*, 1991]). The -35 hexamer of positively regulated promoters often differs from the consensus sequence. Activators appear to assist in binding of the RNA polymerase holoenzyme and may therefore be acting as an alternative to the -35 hexamer, possibly explaining why the binding position of the activator is proximal to the promoter. No activators span the -10 hexamer for the simple reason that strand separation occurs there and the mechanics of transcription initiation require that this area be free (Collado-Vides *et al.*, 1991). It is noteworthy that some activators can also act as repressors. This is accomplished by their binding at alternative sites on the promoter which prevent gene transcription.

σ^{54} promoters

σ^{54} is an alternative bacterial σ -factor which does not appear to be phylogenetically related to σ^{70} . It is involved in a wide range of physiological processes (e.g. NH_4^+ assimilation, N_2 fixation, H_2 oxidation, dicarboxylic acid transport, catabolism of aromatic compounds) but is not crucial to cell survival (Morett and Segovia, 1993). The consensus sequence for σ^{54} includes pentamer and hexamer sequences located at -12 and -24 nucleotides, respectively (Figure 1.1B). A σ^{54} promoter is characterized at the very least by a GC pair at approximately -13 followed by a GG pair exactly 10 bp further upstream (Kustu *et al.*, 1989).

All operons containing σ^{54} promoters are positively regulated. Activators of σ^{54} promoters are modular proteins containing three distinct, structurally independent domains. The N-terminal domain receives the physiological signal responsible for converting the protein to an active form, the central domain is responsible for ATP hydrolysis (see below) and transcription initiation, and the C-terminal domain is responsible for binding to the activation site (Figure 1.1C). The N-terminal domain is highly variable between different activators of σ^{54} -dependent genes, while the central domain and parts of the C-terminal domain are conserved. This indicates that these proteins, or at least the latter two domains, are descended from a common ancestor (Morett and Segovia, 1993).

Binding sites for the activator proteins are inverted repeats 100 to 200 nucleotides upstream of the transcription initiation site (Figure 1.1D; [Morett and Segovia, 1993]). This is in contrast to positively regulated σ^{70} promoters, where binding sites for the activator are always proximal (see above). Like any σ -factor, σ^{54} enhances binding of RNA polymerase holoenzyme to a promoter. However, σ^{54} -holoenzyme is transcriptionally inactive because it is unable to separate the DNA strands. Transcription initiation occurs when some physiological signal converts the activator protein to an active form. The activator protein, which is still bound to the upstream site, contacts the σ^{54} -holoenzyme and catalyses strand separation at the σ^{54} promoter. Contact between the activator and σ^{54} -holoenzyme is accomplished by looping of the intervening DNA (Figure 1.1D; [Kustu *et al.*, 1991; Kustu *et al.*, 1989]). Strand separation is an energy-dependent reaction and ATP is hydrolysed by the activator (Morett and Segovia, 1993).

Global regulatory networks

In response to certain environmental or nutritional changes, whole sets of bacterial genes which share a common regulatory element are either expressed or have the potential to be expressed. For example, σ^S is an alternative σ -factor belonging to the σ^{70} family. It is involved in the transcription of genes which prepare bacteria for survival in harsher environments, especially those with low concentrations of nutrients. During exponential growth, σ^S is present at relatively low amounts. As cells enter stationary phase,

transcription of σ^S increases 20-fold and expression of σ^S -dependent operons is made possible (Loewen and Hengge-Aronis, 1994).

Another regulator common to many diverse genes is the cAMP binding protein (CBP). In response to various stimuli, most notably carbon starvation, intracellular levels of cAMP rise, cAMP binds to CBP, and the cAMP-CBP complex attaches to DNA. cAMP-CBP can then either positively or negatively regulate transcription of various operons. Most are operons involved in catabolism of carbon sources other than glucose, but several are unrelated to metabolic functions (e.g. flagellum synthesis, pilus synthesis; [Botsford and Harman, 1992]). Transcriptional activation is either direct (cAMP-CBP alone) or indirect (cAMP-CBP in conjunction with another positive regulator). Repression by cAMP-CBP is also either direct (e.g. it binds within the promoter and interferes with transcription) or indirect (e.g. it improves binding of a repressor; [Kolb *et al.*, 1993]). Related to CPB but lacking its ability to bind cAMP is FNR (fumarate and nitrate reduction), another global regulator which can activate many different genes involved in the use of alternative electron acceptors when anaerobic conditions prevail (Kolb *et al.*, 1993).

Regulation after transcription initiation

Termination of transcription

Expression of some genes is regulated by transcriptional attenuation, which prematurely terminates transcription of an operon. Transcriptional attenuation occurs mainly in biosynthetic operons (e.g. those involved in formation of pyrimidine bases and some amino acids) but it is also found in several catabolic operons (e.g. β -glucosidase). Transcriptional attenuation is accomplished through a signal region (attenuation site) in the first 100 to 200 nucleotides of the transcript. The attenuation site is an area of the mRNA which can form hairpin structures and terminate transcription. Transcription can only proceed if formation of these secondary structures is prevented. In biosynthetic operons, this is accomplished by the ribosome during translation of the transcript. The attenuation site codes for a small peptide which precedes the first structural gene. If the amount of the amino acid or the pyrimidine base for which the biosynthetic operon codes

is low, the ribosome stalls over the attenuation site, thus preventing formation of the hairpin and allowing mRNA elongation to continue. If there is a surplus of the amino acid or the pyrimidine base, the translational machinery follows closely behind RNA polymerase, the terminator hairpin forms, and mRNA elongation is aborted. In some catabolic operons, formation of hairpins is prevented by a regulatory protein which is activated in response to a signal (e.g. dephosphorylation) and binds to the attenuation site (Landick and Turnbough, 1992).

Antisense RNA

Antisense RNA is used to control expression of some genes (e.g. transfer genes on some plasmids, transposase of *IS10*). Antisense RNA accomplishes this either by complexing with a complete transcript, thus preventing translation, or by complexing with the growing mRNA and creating a duplex region just behind RNA polymerase which mimics termination loops, thus stopping transcription (Simons and Kleckner, 1988).

ORGANIZATION AND REGULATION OF CATABOLIC GENES FOR AROMATIC COMPOUNDS

Overview of aerobic catabolic pathways for aromatic compounds

Microorganisms are crucial to the recycling of organic matter in the environment. Aromatic compounds are ubiquitous in nature and bacteria have evolved the ability to metabolize a wide variety of these substrates (Hagedorn *et al.*, 1988; Kirk and Farrell, 1987). Extensive research has been conducted to isolate these bacteria and characterize their aerobic and anaerobic catabolic pathways towards aromatic substrates. In the past several decades, this topic has received an added impetus because of the realization that many bacteria have the intrinsic ability to biodegrade many synthetic aromatic pollutants, leading to their use in remediation of contaminated sites (bioremediation). However, for bioremediation to be successfully applied to new contaminants and contaminant mixtures, a complete understanding of the inherent metabolic diversity of bacteria toward these substrates is necessary.

Aerobic biodegradation pathways for many structurally diverse aromatic compounds is initiated by oxygenases and CoA ligases, a diverse class of enzymes that channel these compounds towards a few common metabolites before final ring fission of the benzene ring (Figure 1.2A). These metabolites contain two hydroxyl groups which are either adjacent (catechol, protocatechuate) or opposite each other (gentisate) on the aromatic ring. Ring fission can occur between the two hydroxyl groups (intradiol or *ortho*-fission) or outside the two hydroxyl groups (extradiol or *meta*-fission). The ring-fission products are ultimately converted to tri-carboxylic acid (TCA) cycle intermediates which enter central metabolic pathways (Figure 1.2B; [Harayama and Timmis, 1989; Harayama and Timmis, 1992; Harwood and Parales, 1996]).

Catabolic pathways for aromatic compounds are often divided into “upper” and “lower” pathways. Upper pathways funnel aromatic compounds towards monoaromatic dihydroxylated intermediates while lower pathways are involved in ring fission of these intermediates and their subsequent conversion to TCA cycle intermediates. Genes for both types of pathways have been characterized for a wide variety of bacteria and what follows is a summary of the various ways these genes are organized and how their expression is regulated. For the purposes of this review, these genes were grouped according to the nature of the regulatory protein.

Figure 1.2 A) Biodegradation pathways of various aromatic compounds funnel structurally diverse aromatic compounds towards catechol, protocatechuate, or gentisate. B) The dihydroxylated aromatic compounds then undergo intradiol or extradiol ring fission, resulting in the production of metabolites that enter the tricarboxylic acid cycle. Structural analogues (e.g. homoprotocatechuate, homogentisate, and chlorinated variants of catechol and protocatechuate) undergo similar reactions and are not included. Ring fission of gentisate is mechanistically and enzymatically akin to extradiol ring fission (Lipscomb and Orville, 1992). Adapted from various references (Harayama and Timmis, 1989; Harayama and Timmis, 1992).

A

B

Catabolic operons regulated by LysR-type transcriptional regulators

Several catabolic operons involved in the metabolism of a variety of aromatic compounds are positively controlled by regulators belonging to a large family of transcriptional activators known as LysR-type transcriptional regulators (LTTRs). LTTRs regulate genes with diverse functions (e.g. anabolic, catabolic, virulence, antibiotic resistance), all of which are transcribed from σ^{70} -dependent promoters. LTTRs are modular proteins and possess at least three domains. The N-terminus is involved in binding to the DNA and shows sequence conservation between different members of the protein family. The central region is highly variable and is involved in co-inducer recognition. The C-terminus is conserved and is also believed to be involved in DNA binding and/or co-inducer recognition (Figure 1.3A; [Schell, 1993]).

The genes for LTTRs are transcribed divergently from the operons they control and LTTRs negatively regulate their own synthesis. An LTTR interacts with two regions of the promoter. One is called the repressor binding site (RBS). It is approximately 65 nucleotides upstream of the transcriptional start site of the target gene but overlaps or is downstream of the transcriptional start site of the divergently transcribed gene for the LTTR, thus mediating repression of the LTTR gene. The RBS contains a sequence of interrupted dyad symmetry and the LTTR binds to it regardless of whether co-inducer is present. The second site is called the activator binding site (ABS) and is adjacent to the -35 hexamer of the target gene's promoter. The LTTR interacts with it when co-inducer is present and thus stimulates transcription of the gene (Figure 1.3B; [Schell, 1993]).

Chromosomally-encoded pathways for aromatic compound metabolism controlled by LTTRs include the *Acinetobacter* sp. ADP1 *ben-cat* genes for metabolism of benzoate and catechol by intradiol ring fission, regulated by BenM and CatM. BenM induces *benABCD*, which codes for the dioxygenase and dehydrogenase that convert benzoate to catechol. CatM induces two separate operons: *catA*, involved in conversion of catechol to *cis, cis*-muconate; and *catBCIJFD*, involved in further catabolism of *cis, cis*-muconate to TCA cycle intermediates. BenM can also induce *catA*, and to a lesser degree *catBCIJFD*,

Figure 1.3 Catabolic genes for non-chlorinated aromatic compounds that are regulated by LysR-type transcriptional regulators (LTTRs). A) LTTRs are modular proteins and each domain performs a distinct function, as indicated. B) Relative organization of the catabolic gene regulated by the LTTR and the gene for the LTTR, which is transcribed divergently from at least one of its target genes. The LTTR interacts with two sites in the intergenic region. It binds to the repressor site regardless of whether its co-inducer is present and represses transcription of the LTTR gene. If co-inducer is present, it binds to the activator site and activates transcription of the target gene. C) to G) Relative organization of catabolic genes for diverse aromatic compounds and the gene coding for the LTTR. Indicated are the bacteria from which the genes were isolated and whether the genes are on a plasmid. Gene regulation by LTTRs is represented by dashed lines and co-inducers (if known) are shown below the dashed line. Gene lengths are not to scale. Adapted from various references listed in text.

but CatM cannot induce *benABCD*. The co-inducers for BenM are benzoate or *cis, cis*-muconate for induction of either the *ben* or *cat* genes. If present together, synergistic induction of the *ben* genes occurs. CatM only responds to *cis, cis*-muconate (Figure 1.3C; (Collier *et al.*, 1998; Romero-Arroyo *et al.*, 1995)).

Similarly, in several *Pseudomonas putida* strains, the intradiol ring fission pathway of catechol (*catBCA*) is activated by CatR in response to *cis, cis*-muconate (Figure 1.3D; [Houghton *et al.*, 1995; Parsek *et al.*, 1992; Rothmel *et al.*, 1990; Rothmel *et al.*, 1991b]). Further degradation of the metabolite produced by CatBCA is mediated by PcaDIJ, and the genes for these proteins are activated by a GylR/IciR-type regulator (discussed below; Figure 1.7). In addition to binding to the promoter of *catBCA*, CatR binds at an internal position within *catB* (~180 nt downstream of +1 and ~100 nt downstream from the ATG), which lowers expression (Chugani *et al.*, 1998). Furthermore, CatR in conjunction with *cis, cis*-muconate is also necessary for activation of *pheBA*, a plasmid-encoded operon originally isolated in *Pseudomonas* sp. EST1001 and involved in phenol hydroxylation and intradiol ring fission of the catechol that is generated (Figure 1.3D; [Kasak *et al.*, 1993; Parsek *et al.*, 1995]).

Another chromosomally-encoded operon regulated by an LTTR includes the initial genes of the intradiol ring fission pathway for protocatechuate of *Agrobacterium tumefaciens* A348, which is activated by PcaQ in response to the co-inducers β -carboxy-*cis, cis*-muconate and γ -carboxymucunolactone (Figure 1.3E; [Parke, 1993; Parke, 1995; Parke, 1996]). This contrasts with *Acinetobacter* and *Pseudomonas*, in which the homologous pathways are regulated by an IciR/GylR-type regulator (described below). The *hca* operon of *Escherichia coli* K-12, which codes for the dioxygenase and dehydrogenase that convert 3-phenylpropionate or cinnamate to a 2,3-dihydroxyphenyl-compound, also requires an LTTR (designated HcaR), but little else is known regarding control of its expression (Figure 1.3F; [Díaz *et al.*, 1998]). Further metabolism of the latter two compounds requires the unlinked *mhp* extradiol ring fission pathway, which may be controlled by an IciR/GylR-like regulator (described below).

LTTRs are also involved in controlling the expression of various plasmid-encoded catabolic operons for aromatic compounds. However, unlike CatR regulation of *pheBA* described above, the genes for these LTTRs are physically linked to the genes they control. In *P. putida* PpG7, NahR with the co-inducer salicylate activate the two operons for degradation of naphthalene via an extradiol ring fission pathway (Figure 1.3G; [Schell, 1993]). In *C. testosteroni* T-2, TsaR apparently regulates expression of the *tsa* genes coding for the enzymes that oxidize the methyl side group *para* to carboxy- or sulfate-substituted benzene, but characterization of TsaR is in the initial stages and features such as the co-inducer are not yet known (Figure 1.3H; [Junker *et al.*, 1997; Oppenberg *et al.*, 1995]). Similarly, in *P. putida* UCC22, the *tdn* genes for conversion of aniline to catechol are associated with *tdnR* which codes for an LTTR. TdnR is necessary for expression of the *tdn* genes and growth on aniline, but characterization is in the initial stages and little else is known (Figure 1.3I; [Fukumori and Saint, 1997]).

LTTRs also regulate several genes involved in the modified intradiol-ring fission pathways for chlorinated aromatics. All are plasmid-encoded and the LTTR gene is physically linked to the catabolic genes. In *P. putida* RB1, ClcR activates *clcABD*, which is involved in the ring fission of 3-chlorocatechol, produced during the catabolism of 3-chlorobenzoic acid by the non-specific activity of chromosomally-encoded benzoate dioxygenase. The highest expression occurs with the co-inducer 4-chloromuconate, a product of the pathway, but some transcription does take place without this compound (Figure 1.4A; [Coco *et al.*, 1994; Coco *et al.*, 1993; McFall *et al.*, 1997a; McFall *et al.*, 1997b]). Similarly, in *Pseudomonas* sp. P51, TcbR activates *tcbCDEF*, which is also involved in the modified intradiol ring fission pathway for the chlorinated catechols produced by TcbAB-mediated catabolism of 1,2-dichloro-, 1,4-dichloro-, and 1,2,4-trichlorobenzene, but the co-inducer has not been identified (Figure 1.4B; [Leveau *et al.*, 1994; van der Meer *et al.*, 1991a; van der Meer *et al.*, 1991b]).

In *Ralstonia eutropha* (formerly *Alcaligenes eutrophus*) JMP134, the *tfd* pathway for complete degradation of 2,4-dichlorophenoxyacetate (2,4-D) and 3-chlorobenzoate is apparently controlled by two LTTRs, TfdS and TfdR (Figure 1.4C). These proteins,

Figure 1.4 Catabolic genes for chlorinated aromatic compounds that are regulated by LysR-type transcriptional regulators (LTTRs), all of which are located on plasmids. A) to C) Relative organization of the catabolic genes and the physically-linked gene for the LTTR. Indicated are the bacteria from which the genes were isolated. Gene regulation by LTTRs is represented by dashed lines and co-inducers (if known) are shown below the dashed line. Gene lengths are not to scale. Adapted from various references listed in text.

A. *P. putida* RB1 plasmid-encoded *clc* genes

B. *Pseudomonas* sp. P51 plasmid-encoded *tcb* genes

C. *Ralstonia eutropha* JMP134 plasmid-encoded *tfd* genes

which are identical (Matrubutham and Harker, 1994), have apparently contradictory roles, depending on the study. TfdS can either activate or repress *tfdB*, the gene involved in conversion of 2,4-dichlorophenol to 2,4-dichlorocatechol, depending on whether effector is present (Kaphammer and Olsen, 1990). TfdS appears to activate *tfdA*, the gene involved in conversion of 2,4-dichlorophenoxyacetate to 2,4-dichlorophenol (You and Ghosal, 1995). TfdR was initially reported to repress *tfdCDEF*, the genes coding for enzymes involved in intradiol ring fission of 2,4-dichlorocatechol and metabolism of the product (Kaphammer *et al.*, 1990), but another report shows that it in fact activates this operon (Leveau and van der Meer, 1996). The effector of the pathway is probably dichloromuconate, the product of chlorocatechol ring fission (Filer and Harker, 1997). More work is necessary to fully understand how expression of the *tfd* pathway is controlled.

Catabolic operons regulated by two-component systems

One method used by bacteria to regulate expression of several different types of genes (e.g. N₂-fixation, PO₄-uptake, transport of carboxylic acids) is through a family of proteins consisting of two separate components. One component, the sensor, is a histidine protein kinase (HPK) which phosphorylates itself at a histidine group in response to a signal. The phosphate group from the phosphorylated HPK is then transferred to an aspartate group of the second component, the response regulator (Figure 1.5A). Phosphorylation of the response regulator improves its ability to bind to DNA and increase transcription of the target gene. Dephosphorylation is an autocatalytic event. The sensor is usually a membrane-bound protein while the response regulator is cytoplasmic (Stock *et al.*, 1989).

Both the sensor and response regulator are modular proteins. The sensor protein possesses two domains. Located at the N-terminus is the "input" domain, which detects the signal, and at the C-terminus the "transmitter" domain, which contains the HPK activity. With membrane-bound sensors, the input domain spans the lipid bilayer while the transmitter domain extends into the interior of the cell. The response regulator at the very least has a "receiver" domain at the N-terminus, which contains the aspartate group, and

the "output" domain at the C-terminus, which transmits the signal (Figure 1.5A). The output module is often a DNA-binding domain (Parkinson and Kofoed, 1992). Some σ^{54} -dependent genes are controlled by two-component systems, and the response regulators which activate σ^{54} promoters have an extra module responsible for ATP hydrolysis and transcription initiation (described above; Figure 1.1C).

Two σ^{70} -dependent, chromosomally-encoded catabolic operons for aerobic degradation of aromatic compounds are regulated by two-component systems; the *tod* genes from *P. putida* F1 for toluene degradation and the *bpd* genes from *Rhodococcus* sp. M5 for biphenyl degradation, both of which are metabolised by extradiol ring fission pathways (Furukawa, 1994; Zylstra *et al.*, 1988). The *tod* genes are toluene-inducible and require *todS* and *todT*, which code for a sensor and a response regulator, respectively (Lau *et al.*, 1997; Wang *et al.*, 1995). The *todST* operon is located downstream of and is transcribed in the same direction as the *tod* structural genes. TodT binds to the region between the *tod* structural genes and *todR*, an open reading frame read divergently from the *tod* structural genes (Lau *et al.*, 1997). *todR* appears to code for a shortened LTTR lacking the co-inducer recognition domain and the conserved C-terminal domain. It plays little or no role in regulation of the *tod* genes and likely represents the remnants of an ancestral regulatory element used before the TodT/TodS system evolved to control the *tod* genes (Figure 1.5B; [Wang *et al.*, 1995]). Expression of the *bpd* genes, or at least *bpdC1C2BADE*, requires *bpdS* and *bpdT*, which appear to code for a sensor and a response regulator, respectively, based on sequence similarity with TodS/TodT (Figure 1.5C). Biphenyl induces expression of both sets of genes, but little else is known regarding this system as its description is in the initial stages (Labbé *et al.*, 1997). Similarly, the *tut* pathway of strain T1 for anaerobic metabolism of toluene require genes that code for proteins similar to TodT/TodS, but little else is known (Coschigano *et al.*, 1998; Coschigano and Young, 1997).

Figure 1.5 Catabolic genes for aromatic compounds that are regulated by two-component systems. A) Two component systems include a sensor and response regulator. In response to a signal, the sensor auto-phosphorylates at a histidine (H) group in the transmitter domain. The phosphate group is then transferred to an aspartate group (D) in the receiver domain of the response regulator. In response regulators which affect DNA transcription, the output domain is a DNA-binding element. Response regulators which activate σ^{54} promoters possess an extra module responsible for ATP hydrolysis and transcription initiation. Not all σ^{54} promoters are controlled by two-component systems. Figure adapted from Parkinson and Kofoed (1992). B) Organization of *tod* genes for degradation of toluene, which are co-induced by toluene. *todR* codes for what appears to be a truncated LysR-type transcriptional regulator which has a negligible role in regulation of the *tod* genes. C) *bpd* genes for degradation of biphenyl, which are co-induced by biphenyl. Studies of the interactions between the sensors (TodS, BpdS) and response regulators (TodT, BpdT) are in the initial stages, but some aspects of promoter-binding by the regulators are known, and they are indicated. Gene lengths are not to scale. Adapted from various references listed in text.

A.

B. *P. putida* F1 *tod* genes

C. *Rhodococcus* sp. M5 *bpd* genes

Catabolic operons dependent on σ^{54} and controlled by NtrC-like regulators

σ^{54} -dependent promoters have been discovered in several catabolic operons for aromatic compounds. As described above, all σ^{54} -dependent genes are positively regulated, and it appears that for the aromatic catabolic pathways described below, transcription of the genes for the cognate activators (most of which are σ^{70} -dependent) is constitutive but at a low level. Thus, a basal amount of activator is present which is bound to the upstream activator binding sites. Induction of the target gene, though, only occurs when co-inducer is present. The activators of σ^{54} -dependent genes described in this section all belong to the NtrC family of transcriptional regulators, named after proteins involved in regulating nitrogen metabolism in various bacteria (Morett and Segovia, 1993). It is noteworthy that some members of the NtrC family are also the response regulators of two-component systems (Kustu *et al.*, 1989; Parkinson and Kofoed, 1992), but the activators described here do not possess a separate sensor which phosphorylates them in response to a signal. In effect, the sensor is the N-terminal domain, which responds to co-inducers of gene expression and transmits the signal to the central domain, which then initiates transcription.

In *Pseudomonas* sp. P35X and *Pseudomonas* sp. CF600, expression of the chromosomally-encoded *phh* operon and plasmid-encoded *dmp* operon, respectively, both for conversion of phenol/methylphenols to catechol/methylcatechols and subsequent ring fission by the extradiol pathway, require σ^{54} (Figure 1.6C and 1.6D). The genes for the positive regulators of these operons (PhhR and DmpR, respectively) are divergently transcribed from the target operons and the co-inducers are the substrates of the pathways (Ng *et al.*, 1995; Shingler *et al.*, 1993). Similarly, the chromosomally-encoded *mop* operon of *Ac. calcoaceticus* NCIB8250 for conversion of phenol to catechol and subsequent metabolism by the intradiol ring fission pathway also requires σ^{54} . Expression is controlled by MopR, also encoded by a divergently transcribed gene, in response to the substrate (Figure 1.6E; [Ehrt *et al.*, 1994; Ehrt *et al.*, 1995; Schirmer *et al.*, 1997]).

Figure 1.6 Catabolic genes for degradation of aromatic compounds that possess σ^{54} -dependent promoters and are activated by NtrC-like proteins. Summary of A) modular nature of activators for σ^{54} -dependent genes and B) the general organization of a σ^{54} -dependent gene and the gene for the activator. More detail is provided in Figure 1.1D. C) to I) Organization of various σ^{54} -dependent genes involved in catabolism of aromatic compounds relative to the gene for the NtrC-like protein. Bacteria from which the genes were isolated are indicated. Dashed lines indicate genes regulated by the NtrC-like protein and co-inducers are below the dashed lines. Gene lengths are not to scale. Adapted from various references listed in the text.

Some other σ^{54} -dependent catabolic operons differ in interesting ways. For example, the regulatory and structural genes of the *tbm* pathway of *Burkholderia* sp. JS150 for metabolism of toluene by a 2- or 4-monooxygenase are arranged as described above (Figure 1.6F), but the regulator (TbmR) responds to either the parent substrates or the downstream metabolites (Johnson and Olsen, 1997). In contrast, the *tbu* and *tou* pathways of *Burkholderia pickettii* PKO1 and *P. stutzeri* OX1, respectively, for metabolism of toluene or xylene by monooxygenases are upstream and transcribed in the same direction as the regulatory genes (Figure 1.6G and 1.6H). TbuT, the activator of the *tbu* pathway responds to the parent substrates or the downstream products, but TouR, the activator of the *tou* pathway only responds to the downstream products (Arengi *et al.*, 1999; Byrne and Olsen, 1996). With both sets of catabolic genes, a basal level of expression appears to always take place, permitting production of a small amount of the downstream metabolites. When the parent substrate is present, a cascade is set up as both sets of compounds increase expression of both the catabolic and regulatory gene (Arengi *et al.*, 1999; Byrne and Olsen, 1996). Finally, the *aph* genes of *C. testosteroni* TA441 for phenol degradation via catechol and an extradiol ring fission pathway are induced by an NtrC-like protein from a divergently transcribed gene (Figure 1.6I). However, the *aph* genes were only uncovered after a lengthy adaptation phase (2-3 weeks) during which an unidentified protein, possibly a repressor, was apparently disabled, allowing expression of *aph* to occur and therefore growth on phenol (Arai *et al.*, 1998).

Catabolic operons regulated by GylR/IciR-like transcriptional regulators

Regulation of the catabolic genes for a variety of aromatic catabolic operons are regulated by GylR/IciR-type regulators. GylR is an activator of glycerol catabolic genes in *Staphylococcus coelicolor* and IciR is a repressor of catabolic genes for acetate by the glyoxylate pathway in *E. coli* and *Salmonella typhimurium* (DiMarco *et al.*, 1993). In *Acinetobacter* sp. ADP1 and *P. putida* PRS2000, proteins from this family control metabolism of quinate, shikimate, and/or 4-hydroxybenzoate via protocatechuate and the intradiol ring fission pathway, all of which are σ^{70} -dependent (Figure 1.7).

Figure 1.7 Catabolic genes for aromatic compounds regulated by GylR/IciR-like regulators. Indicated are the relative organization, position, and direction of transcription for catabolic genes and the gene for the regulators. The best characterized examples are the evolutionarily related genes of A) *Acinetobacter* sp. ADP1 and B) *P. putida* PRS2000 for metabolism of several substrates via protocatechuate and an intradiol ring fission pathway. For homologous enzymes in the pathways of the two bacteria, the reactions catalyzed are identical but some interesting contrasts include different arrangements of homologous genes and how regulators exert control. In *Acinetobacter* sp. ADP1, the *pca* genes form a contiguous operon with *quiBCX* while *quiA* is believed to form a separate operon. The *qui* genes code for dehydrogenases that funnel quinate and shikimate towards protocatechuate. In *P. putida* PRS2000, the various genes for protocatechuate degradation are distributed into several distinct operons. Dashed lines indicate genes activated by the indicated regulator and co-inducers are shown below the dashed lines. Gene lengths are not to scale. C) The β -ketoadipate pathway for benzoate degradation via catechol and 4-hydroxybenzoate degradation via protocatechuate. Indicated are the genes identified in *Acinetobacter* and *P. putida* for catabolism of the metabolites, the regulatory proteins (if known), and the co-inducers. CatM/CatR and BenM are LysR-type regulators and were discussed previously (Figure 1.3). Note that in *Pseudomonas*, the shared metabolites of the two branches are metabolized by one set of enzymes (*pcaDIJF*) while in *Acinetobacter* sp. ADP1 there are redundant enzymes coded for by *pcaDIJF* and *catDIJF* which degrade the shared metabolites. The pathway that is active will depend on the original source of the metabolites. Adapted from (Harwood and Parales, 1996) and various references listed in the text. Recall that the homologous protocatechuate catabolic genes in *Ag. tumefaciens* A348 are regulated by an LTTR (see Figure 1.3).

A. *Acinetobacter* sp. ADP1 *pca-qui-pob* genes

B. *P. putida* PRS2000 *pca* genes

C. β -ketoadipate pathway of *Acinetobacter* and *P. putida*

In *Acinetobacter* sp. ADP1, the genes for protocatechuate catabolism form part of a cluster of operons involved in the dissimilation of quinate and shikimate (non-aromatic, cyclic organic acids of plant origin) and of 4-hydroxybenzoate (Figure 1.7A). The structural genes for protocatechuate fission and conversion to TCA cycle intermediates, *pcaIJFBDKCGH*, form a contiguous operon with *quiBCX*, the structural genes for conversion of dehydroquinate to dehydroshikimate (*quiB*) and dehydroshikimate to protocatechuate (*quiC*; [Elsemore and Ornston, 1994]). These genes are activated by PcaU in response to protocatechuate (Gerischer and Ornston, 1995). *quiA* codes for the enzyme that dehydrogenates quinate or shikimate and is also expressed in response to protocatechuate (Stanier and Ornston, 1973). *quiA* can be transcribed independently from the *pca-qui* operon, leading to the suggestion that it is a separate operon (Elsemore and Ornston, 1995), but more research is necessary to clearly elucidate what controls expression of this gene. Another gene in the *pca-qui* catabolic operon cluster of *Acinetobacter* sp. ADP1 is *pobA*, the structural gene for conversion of 4-hydroxybenzoate to protocatechuate, which is positively regulated by PobR in response to the parent substrate (DiMarco *et al.*, 1993).

Both PcaU and PobR bind to the intergenic region between the catabolic gene and the divergently transcribed gene for themselves. They bind to the intergenic DNA regardless of whether co-inducer is present and negatively regulate their own synthesis. Binding of the regulator therefore does not initiate expression of the catabolic gene unless co-inducer is present, a mechanism of action that is similar to that of an LTTR (see above). When their respective co-inducer is present, though, the two activators differ in their response: PcaU induces expression of both *pcaU* and the *pca* genes while PobR only induces expression of *pobA* and *pobR* expression is unaffected (DiMarco and Ornston, 1994; Gerischer *et al.*, 1998).

In *P. putida* PRS2000, the genes involved in catabolism of protocatechuate are divided into five unlinked operons (Figure 1.7B). *pcaHG*, the structural genes for intradiol fission of protocatechuate, is induced by protocatechuate, but the regulatory protein has not been characterized. *pcaBDC*, *pcaIJ*, and *pcaF*, the structural genes for further

catabolism of the protocatechuate fission product, are positively regulated by PcaR and the co-inducer β -ketoadipate. However, the exact organization and direction of transcription of the regulatory and structural genes are not known (Hughes *et al.*, 1988; Nichols and Harwood, 1995; Parales and Harwood, 1993). Like PobR and PcaU (see above), PcaR binds to the promoter region of one of its target operons (*pcaIJ*) regardless of whether co-inducer is present, but gene expression takes place when co-inducer appears (Romero-Steiner *et al.*, 1994). The exact region of DNA bound by PcaR has not been defined, but as little as 30 bp preceding the transcription initiation site are required for full activation of *pcaIJ* (Parales and Harwood, 1993). A 15 bp sequence within that region has also been found in the promoter areas of *pcaF*, *pcaK* (see below), and *pcaR*, leading to the suggestion that this sequence plays a role in binding of PcaR to promoter DNA of the genes it regulates (Romero-Steiner *et al.*, 1994), but more work is necessary for that to be shown conclusively.

PcaR and β -ketoadipate also activate *pcaK*, a gene coding for a protein involved in uptake of 4-hydroxybenzoate (Harwood *et al.*, 1994; Nichols and Harwood, 1995). However, if benzoate is present in the medium along with 4-hydroxybenzoate, benzoate appears to repress *pcaK* expression, as well as lowering levels of PobA (4-hydroxybenzoate hydroxylase) and PcaHG (protocatechuate 3,4-dioxygenase), without affecting *pcaF* expression (Nichols and Harwood, 1995). In *P. putida* (and *Acinetobacter* sp. ADP1), benzoate is initially catabolized via a pathway parallel to but distinct from the 4-hydroxybenzoate/protocatechuate pathway. The two degradation pathways converge at the point where β -ketoadipate enol-lactone is converted to β -ketoadipate (Figure 1.7C). PcaF, PcaD, and PcaIJ are used when catabolizing benzoate, yet how *pcaF* (and presumably *pcaBDC* and *pcaIJ*) is activated while *pcaK* is repressed, even though all require the same positive regulator, is unknown. CatR, the activator of *catBC* (Figure 1.3D and 1.6), is not involved (Nichols and Harwood, 1995). Conceivably, a benzoate- or *cis*, *cis*-muconate-requiring repressor may be involved. Benzoate is probably used in preference to 4-hydroxybenzoate because catabolism of benzoate does not require any net

input of reducing equivalents, while catabolism of 4-hydroxybenzoate does (Figure 1.6C; [Nichols and Harwood, 1995]).

Several other catabolic pathways for aromatic compounds have been cloned and sequenced (*mhp* and *hpp* genes for 3-hydroxyphenylpropionate hydroxylation and subsequent metabolism by an extradiol ring fission pathway in *E. coli* K-12 and *Rhodococcus globerulus* PWD1, respectively; the *pca* homologues of *Rhodococcus opacus* 1CP for protocatechuate metabolism; the *ohb* genes for halobenzoate degradation by *P. aeruginosa* 142). Closely linked ORFs that may code for IciR/GylR-type regulators have been identified, but their role, if any, in controlling expression has not yet been elucidated (Barnes *et al.*, 1997; Eulberg *et al.*, 1998b; Ferrández *et al.*, 1997; Tsoi *et al.*, 1999).

Catabolic operons regulated by AraC-like transcriptional regulators

AraC, the activator of the arabinose catabolic operon of *E. coli*, is the prototypic member of what appears to be another family of regulators involved in controlling aromatic catabolic operons (Gallegos *et al.*, 1997). To date, the only catabolic operon that is confirmed to be regulated by an AraC-like transcriptional regulator is the lower pathway of the xylene catabolic pathway (*xyl*), which is activated by XylS (discussed below; see Figure 1.9). Several other catabolic pathways, though, have been sequenced and closely associated ORFs that may code for AraC-like regulators have been identified. One is the plasmid-encoded *cbd* genes of *B. cepacia* 2CBS for the enzymes that convert 2-halobenzoates to catechol (Haak *et al.*, 1995). Another is the chromosomally-encoded *hpa* genes of *E. coli* W for 4-hydroxyphenylacetate metabolism (Prieto *et al.*, 1996), but these are discussed below as an example of a pathway that may be controlled by two regulators.

Catabolic operons regulated by repressors

Aromatic catabolic pathways appear to be subject mostly to positive control, but some pathways have been uncovered that appear to be under some form of negative control. Several catabolic operons are regulated by proteins that are similar to MarR, a repressor controlling the multiple antibiotic resistance response of *E. coli* (Seoane and Levy, 1995). All pathways controlled by MarR-like regulators appear to be σ^{70} -dependent.

Figure 1.8 Catabolic genes that are negatively regulated. These include genes for the degradation of A) homoprotocatechuate in *E. coli* C; B) isopropylbenzene (*cym*) and isopropylbenzoate (*cmt*) in *P. putida* F1; and C) phenylacetate in *E. coli* W. Dashed lines indicate genes repressed by the indicated regulator and de-repressors are shown below the dashed lines. The orientation of *cymR* relative to the genes it regulates was not reported. Gene lengths are not to scale. Adapted from various references listed in the text.

A. *E. coli* C *hpc* genes

B. *P. putida* F1 *cym-cmt* genes

C. *E. coli* W *paa* genes

For example, the *hpc* genes of *E. coli* C for homoprotocatechuate metabolism by an extradiol ring fission pathway are negatively regulated by HpcR. *hpcR* is upstream of and transcribed divergently from the *hpc* structural genes, which are induced by homoprotocatechuate (Figure 1.8A; [Jenkins and Cooper, 1988; Roper *et al.*, 1993]). An almost identical operon exists in *E. coli* W, but it is part of a larger set of catabolic genes for 4-hydroxyphenylacetate metabolism, and because these may be controlled by two regulators they are discussed below. Similarly, another MarR-like protein has been implicated in anaerobic benzoate degradation by the *bad* pathway of *Rhodopseudomonas palustris*, but this is also discussed below as an example of a pathway controlled by two regulators.

Repressors unrelated to MarR and involved in regulating aromatic catabolic pathways include CymR of *P. putida* F1, which controls the chromosomally-encoded *cym* and *cmt* genes for conversion of 4-isopropyltoluene (cymene) to 4-isopropylbenzoate (cumate) and subsequent metabolism by an extradiol ring fission pathway (Figure 1.8B). CymR responds to 4-isopropylbenzoate (but not 4-isopropyltoluene) and shows similarity to MtrR, a regulator of antibiotic efflux genes in *Neisseria gonorrhoeae* (Eaton, 1996; Eaton, 1997). In *E. coli* W, PaaX controls expression of the chromosomally-encoded *paa* genes for metabolism of phenylacetate by a CoA-ligase pathway (Figure 1.8C). PaaX responds to phenylacetyl-CoA (but not phenylacetate) and shows similarity to GntR, a repressor of gluconate metabolism in *Bacillus subtilis* (Ferrández *et al.*, 1998; Fujita and Miwa, 1989).

Catabolic pathways dependent on two or more regulatory factors

Several catabolic pathways for aromatic compounds are known or believed to be controlled by more than one regulator. One of the best characterized set of the operons which can be classified in this group is the plasmid-encoded *xyl* pathway of *P. putida* mt-2 for degradation of toluene and *m*- and *p*-xylene. A plethora of factors is involved in controlling *xyl**CMABN*, the σ^{54} -dependent upper operon which codes for the enzymes that convert toluene/xylene to benzoate/methylbenzoate, and *xyl**XYZLTEGFJQKIH*, the σ^{70} -/ σ^S -dependent lower operon which codes for the enzymes that convert

benzoate/methylbenzoate to catechol/methylcatechol and subsequently cleaves them via an extradiol ring fission pathway (Figure 1.9A). The upper operon is positively controlled by XylR (an NtrC-like activator; discussed above) while the lower operon is positively controlled by XylS (an AraC-like activator; discussed above). XylR is also the activator of *xylS*, a σ^{54} -dependent gene. The upper operon is induced in response to toluene or xylene, effectors of XylR. The lower operon can be induced in response to benzoate or methylbenzoate, effectors of XylS, or by hyperproduction of XylS, which occurs when effectors of XylR are present and *xylS* transcription is induced (Figure 1.9A; [Marques and Ramos, 1993]).

The lower operon can also experience XylS-independent induction in response to benzoate. Expression is about 10% of its normal level, it occurs during post-exponential and stationary phase instead of the exponential phase (see below), and is apparently activated by a chromosomally-encoded regulator, possibly one from the chromosomal benzoate degradation pathway. σ^{54} is involved, but it is not certain whether it is used instead of σ^{70} to transcribe the lower operon or if it is necessary for transcription of the alternative activator (Cuskey and Sprengle, 1988; Jeffrey *et al.*, 1992; Kessler *et al.*, 1994). As well, both the upper and lower operons are subject to catabolite repression and this appears to be mediated by lowered expression of XylR, which thus affects expression of the σ^{54} -dependent genes that require this activator for induction (Ramos *et al.*, 1997).

Expression of the upper and lower *xyl* operons demonstrates growth-phase dependence. Expression of the σ^{70} -dependent lower operon begins during exponential phase, while that of the σ^{54} -dependent upper operon begins during late-exponential and stationary phase (de Lorenzo *et al.*, 1993). As well, an alternate σ -factor may be used to transcribe the lower operon, depending on the growth-phase. During exponential phase, σ^{70} is used, but as cells enter late-exponential and early-stationary phases, σ^S is used (Marques *et al.*, 1995).

Figure 1.9 Catabolic genes for aromatic compounds that are regulated by at least two regulatory elements. Dashed lines indicate genes controlled by the indicated regulator. Effectors, if known, are shown below the dashed lines. Gene lengths are not to scale. A) The *xyl* genes of *P. putida* mt-2 for degradation of toluene, *o*-xylene, *p*-xylene are controlled by XylR, an NtrC-like regulator, and XylS, an AraC-like regulator. Indicated also is the σ factor upon which the gene is dependent. B) The *badCDEFG* genes of *Rhodopseudomonas palustris* CGA009 for several steps of the anaerobic metabolic pathway of benzoate are induced by BadR, a MarR-like regulator, and AadR, an FnrR-like regulator coded for by a gene unlinked to the *bad* genes. C) The *hpa* genes of *E. coli* W for metabolism of 4-hydroxyphenylacetate have closely associated genes that may code for a MarR-like protein (*hpaR*) and an AraC-like protein (*hpaA*), but no role for either has been demonstrated yet. D) The *phn* genes of *Burkholderia* sp. RP007 for initial steps in the metabolism of phenanthrene or naphthalene have closely associated genes that may code for an NtrC-like protein (*phnS*) and a LysR type transcriptional regulator (*phnR*), but no role for either has been demonstrated yet. Adapted from various references listed in the text.

A. *P. putida* mt-2 plasmid-encoded *xyl* genes

B. *Rhodopseudomonas palustris* CGA009 *bad* genes

C. *E. coli* W *hpa* genes

D. *Burkholderia* sp. RP007 *phn* genes

As with the activator of any σ^{54} -dependent promoter, XylR binds to a site relatively far upstream of the transcription initiation site of the upper operon and *xylS*. Transcription of the target genes occurs when the DNA between bound XylR and σ^{54} -RNA polymerase holoenzyme loops out and contact between the activator and holoenzyme is made (Figure 1.1D). Two sequence-specific DNA-binding proteins that often cause bends in DNA are involved in controlling expression of these σ^{54} -dependent genes. For the upper operon, integration host factor (IHF) assists in DNA looping and is critical for expression (Calb *et al.*, 1996; Marques and Ramos, 1993). For *xylS*, IHF can bind to the intervening DNA but this inhibits expression of the gene (Holtel *et al.*, 1995). Looping of the intervening DNA in the *xylS* promoter is assisted by HU, a histone-like protein (Perez-Martin and de Lorenzo, 1995).

Another aromatic catabolic pathway controlled by two regulators is the chromosomally-encoded σ^{70} -dependent *bad* genes of *Rhodopseudomonas palustris* for anaerobic metabolism of benzoate by a CoA pathway. The two activators involved are BadR, a MarR-like protein (discussed above, Figure 1.8A) coded for by a linked gene, and AadR, a member of the FNR family of activators (see above, "Global regulatory networks") coded for by an unlinked gene (Figure 1.9B). BadR is probably inducing expression in response to benzoate while AadR is likely responding to the redox state of cells. In the absence of either regulator, expression of the genes and growth on benzoate is impaired, but in the absence of both, cells are unable to grow even though a low level of *badE* expression occurs (Dispensa *et al.*, 1992; Egland and Harwood, 1999; Egland *et al.*, 1997). It is noteworthy that BadR is an activator while most MarR-like regulators are repressors.

Several other chromosomally-encoded aromatic catabolic pathways have been isolated that possess closely linked open reading frames that may code for proteins showing identity to known regulators. For example, the *hpa* genes of *E. coli* W for 4-hydroxyphenylacetate metabolism have *hpaR* (a MarR-like regulator) and *hpaS* (an AraC-

like regulator; [Prieto *et al.*, 1996]). As well, the *phn* genes of *Burkholderia* sp. RP007 for naphthalene or phenanthrene metabolism have *phnS* (an LTTR) and *phnR* (an NtrC-like regulator; Laurie and Lloyd-Jones, 1999]). The arrangement of these genes relative to linked catabolic genes is shown in Figure 1.9C and 1.9D, but more research is necessary to determine whether these open reading frames code for active proteins that are indeed involved in regulating expression of their respective genes.

Concluding statements

The organization of aromatic catabolic genes and modes of control vary greatly. A wide assortment of transcriptional regulators belonging to separate families of proteins are known to be involved, and often the genes for these proteins are closely linked to the catabolic operons they regulate. For the most part, pathways appear to be controlled by positive regulators (activators), but the involvement of negative regulators (repressors) is known for some. Furthermore, regulators can respond to the substrates of the pathway or downstream metabolites. For the latter to occur, a certain basal level of expression of the catabolic genes is necessary, which allows production of the co-inducing compound when parent substrate is present (e.g. CatR-mediated induction of various pathways and TouR-mediated induction of the *tou* pathway).

Metabolism of many aromatic compounds often involves both “upper” and “lower” pathways. Some upper and lower pathways are not linked and induction of the two involves the coordination of various distinct elements. It is noteworthy that unlinked pathways can reflect those that are far apart on the chromosome (e.g. *hca-mhp* pathway of *E. coli* in which *hca* mediates dihydroxylation of 3-phenylpropionate or cinnamate and *mhp* mediates metabolism of the dihydroxymetabolite) or those in which one pathway is on the chromosome while the other is found on a plasmid (e.g. *ben-clc* pathway in which chromosomal *ben* genes mediate conversion of 3-chlorobenzoate to 3-chlorocatechol and *clc* on a plasmid mediates chlorocatechol metabolism).

Although it is beyond the scope of this review to go into any depth on the topic, it is noteworthy that many aromatic degradative operons have clear evolutionary links. This is reflected not only by sequence similarities for proteins with comparable function, but in

the organization of catabolic genes, their mode of regulation, and the structure and sequence of the promoters. Examples include the modified intradiol ring fission genes for metabolism of chlorocatechols (*clcAB*, *tcbCD*, *tfdCD*; Figure 1.4) and the functionally similar genes for intradiol metabolism of catechol (*catAB*, Figure 1.3; [Eulberg *et al.*, 1998a]); and corresponding genes in the various σ^{54} -dependent operons for phenol/methylphenol metabolism by the *phh*, *dmp*, *mop*, and *aph* pathways (Figure 1.6; [Arai *et al.*, 1998]). Between some related operons, though, upper and lower pathways can differ, such as in the aforementioned phenol/methylphenol pathways where the upper pathways are homologous but the lower pathways of *phh*, *dmp*, and *aph* metabolize the catecholic intermediates by an extradiol ring fission pathway while the *mop* operon uses an intradiol ring fission pathway (Figure 1.6A-D; reviewed by Arai *et al.* [1998]). Between homologous pathways, other variations that can arise include differences in the relative orientation of corresponding catabolic or regulatory genes or the type of controlling element, such as in the *pca* genes of *Acinetobacter* ADP1, *P. putida*, and *Agrobacterium tumefaciens* (Figures 1.7A, 1.7B, and 1.3F, respectively; [Eulberg *et al.*, 1998b]).

Because of the evolutionary links between some pathways, regulators are sometimes able to cross-activate non-target promoters. This was alluded to above (e.g. BenM-mediated induction of *cat* genes or XylS-independent induction of the lower *xyl* operon), but very well studied examples include the NtrC-like activators PhhR, DmpR, and XylR, the activators of the *phh*, *dmp*, and two *xyl* operons, respectively (Figure 1.6C, 1.6D, and 1.9A). DmpR and XylR are similar enough that each can activate the other's promoter, provided the co-inducers of the respective activator are present (Fernandez *et al.*, 1994). Furthermore, the effector range of DmpR and XylR can be changed simply by switching their N-terminal domains (Shingler and Moore, 1994). Interestingly, *Pseudomonas* sp. CF600 is unable to grow on dimethylphenol as a carbon source not because of an inability of the *phh* catabolic enzymes to recognize this substrate but because PhhR does not activate transcription in response to it (Ng *et al.*, 1995). The ability of some regulators to induce non-target promoter often reflects similarities between non-cognate promoters as well as the ability of regulators to respond to a broad spectrum

of compounds, but whether this “cross-activation” is crucial to proper regulation of the catabolic pathway or just a “hold over” reflecting an evolutionary link depends on the particular operon.

It appears that bacterial metabolic pathways possess a large amount of innate plasticity in the degree to which they are controlled and the type of regulators that can exert an influence. This is apparently crucial in the evolution of different gene arrangements for biodegradation of aromatic compounds, many of which are novel to the environment (de Lorenzo and Perez-Martin, 1996).

RESEARCH OBJECTIVES AND RATIONALE FOR PROJECT

Degradation of chlorinated aromatic compounds and the *cba* pathway

Chlorinated aromatic compounds are integral to our modern society and have found various applications in industry, agriculture, and the home. Their extensive use and at times accidental release has resulted in widespread contamination of the environment, and because many of these chemicals are toxic, pollution by these compounds has been a topic of concern for several decades.

Until quite recently, halogenated organics were thought to be strictly man-made and considered novel additions to the environment. Newer research, though, has shown that many of these chemicals are synthesized by a variety of organisms (de Jong and Field, 1997; Gribble, 1992; van Pee, 1996). Consequently, these compounds (or structurally similar variants) have always been present, and many bacteria have the innate ability to degrade a wide variety of chloroaromatics. It would appear, therefore, that human activities have, if anything, just increased their concentration in the environment.

As with non-chlorinated aromatic substrates, bacterial metabolic pathways channel these compounds toward several common dihydroxylated intermediates before final ring fission of the benzene ring, and a summary is provided in Figure 1.10. The best characterized ones proceed via intradiol ring fission for the diol intermediates using the β -ketoadipate pathway (*cat* or *pca* genes) or a modified form of the β -ketoadipate pathway that has evolved to handle chlorocatecholic intermediates (*clc* or *clc*-like genes; [Chaudhry and Chapalamadugu, 1991; Reineke, 1998; Reineke and Knackmuss, 1988]). Many of

Figure 1.10 Degradation pathways for chlorinated aromatics channel these compounds towards dihydroxymetabolites A) The most frequently encountered biodegradation pathways funnel chlorinated aromatics towards vicinal diol intermediates which are then metabolized via intradiol ring fission. Common intermediates are catechol (Cat) or chlorinated variants of catechol, but a CoA ligase pathway that converts 4-chlorobenzoate via 4-hydroxybenzoate to protocatechuate (Pca) has been reported. B) Less often reported are lower pathways that metabolize the dihydroxy intermediates by alternate ring fission pathways. The best characterized one converts 3-chlorobenzoate (3Cba) to Pca or 5-Cl-Pca, which are then degraded by extradiol ring fission enzymes. As well, a CoA ligase pathway that degrades 3Cba via 4-hydroxybenzoate to gentisate has been described. Adapted from various references listed in the text. Abbreviations: PCBs, polychlorinated biphenyls; 2,4-D, 2,4-dichlorophenoxyacetic acid; 2,4,5-T, 2,4,5-trichlorophenoxyacetic acid; Cl-Ba, chlorinated benzoic acid; Cl-Ph, chlorinated phenol; Cl-Bz, chlorinated benzene.

A) intradiol ring fission pathways

B) extradiol ring fission and gentisate

these pathways have been characterized at a molecular level and were discussed above (e.g. Figure 1.4). Less frequently reported, and consequently less well characterized, are pathways that metabolize the dihydroxy intermediates by alternate lower ring fission mechanisms. For example, *Alcaligenes* sp. L6 degrades 3-chlorobenzoate (3Cba) by a CoA ligase pathway to 3-hydroxybenzoate and then gentisate (Krooneman *et al.*, 1998; Krooneman *et al.*, 1996), but no information on the genes involved has been published.

A better characterized example of the latter was described in *C. testosteroni* BR60 for growth on 3Cba, 4Cba, or 3,4-diCba, which are converted to protocatechuate (Pca) or 5-Cl-Pca by 3-Cba-3,4-(4,5)-dioxygenase, the product of *cbaAB*, and chlorobenzoate-dihydrodiol dehydrogenase, the product of *cbaC* (Figure 1.11B; [Nakatsu *et al.*, 1997; Nakatsu and Wyndham, 1993]). These genes are on Tn5271, a ~17 kbp transposable element found in pBRC60, an 88 kbp broad-host range conjugative plasmid of the IncP β group (Figure 1.11A; [Nakatsu *et al.*, 1991; Wyndham *et al.*, 1988]). Further metabolism of the products, which is necessary for growth, requires the chromosomally-encoded Pca extradiol ring fission genes (*pmd* for protocatechuate *meta* dioxygenase), making *pmd* the lower pathway of *cba*-mediated metabolism of 3Cba by this organism (Providenti *et al.*, 1999b).

The *cba* genes were originally detected because of their inherent instability: unless constant selection for the trait is maintained, homologous recombination between the IS1071 copies (Figure 1.11A) results in loss of the intervening DNA so that *C. testosteroni* BR60 loses the ability to grow on 3Cba (Wyndham *et al.*, 1988). The lower *pmd* pathway, though, is not linked to Tn5271 and is a stable trait. Some initial studies on induction of the *cba* pathway have shown that 3Cba oxidation levels by *C. testosteroni* BR60 are highest if cells are grown on 3Cba or 3,4-diCba. Interestingly, growth on benzoate, a compound that is not thought to be metabolized by this pathway, also induces some activity (~20%; [Nakatsu and Wyndham, 1993]). Little else, though, is known regarding control of the *cba* pathway.

The *cba* pathway is located on a broad-host range plasmid and therefore has the potential for horizontal transfer to a wide variety of β - and γ -proteobacteria. However, it is

most often associated with bacteria possessing a lower pathway that metabolizes Pca by an extradiol ring fission pathway (Fulthorpe and Wyndham, 1989; Fulthorpe and Wyndham, 1991; Fulthorpe and Wyndham, 1992; Nakatsu *et al.*, 1995; Peel and Wyndham, 1999; Wyndham *et al.*, 1994). Whether this is because spread of the *cba* genes is limited to hosts that can integrate the *cba* pathway with *pmd*-like pathways or because these hosts possess specific regulatory factors necessary for proper *cba* expression is unknown.

Rationale for the project

The *cba* genes of *C. testosteroni* BR60 (pBRC60) represent another catabolic pathway in what appears to be a very diverse metabolic armamentarium available to bacteria for growth on chlorinated aromatic compounds. Certain fundamental features of the *cba* pathway, though, are poorly understood. Because a better understanding of factors controlling this pathway may improve bioremediation of sites contaminated with refractory chemicals like polychlorinated biphenyls; and because this work may provide some insight into how catabolic pathways and regulatory determinants co-evolve and what affects the spread of transmissible elements under various selective pressures, in this case novel growth substrates, I undertook a project to better characterize the *cba* genes.

Research objectives

The specific objectives of this project were to:

- (1) understand how *cbaABC*-mediated catabolism of 3Cba is controlled.
- (2) identify the regulatory elements involved and characterize at a molecular level how they mediate their control.

Figure 1.11 The *cba* pathway of *C. testosteroni* BR60 (pBRC60) A) Plasmid pBRC60 (approximately 88 kbp) harbours Tn5271, which contains *cbaABC*, the catabolic genes for degradation of 3-chlorobenzoate (3Cba). The *cba* genes are flanked by two 3.2 kbp direct repeats, designated IS1071L and IS1071R, each of which codes for a transposase (*mpA*). Site specific recombination between these elements can lead to loss of the *cba* genes. Also indicated on the plasmid are the relative positions of several other loci, such as *oriV* (origin of vegetative replication), *oriT* (origin of transfer), Tra (conjugal transfer functions), and *pri* (primase). B) *cbaABC* codes for: CbaAB, a relatively non-specific enzyme that dioxygenates 3Cba at the 3,4 or 4,5 position; and CbaC, which dehydrogenates one of the *cis* diol intermediates (indicated) to 5-Cl-protocatechuate. Dehydrogenation of the other to protocatechuate is spontaneous. Both compounds are then further catabolized by the chromosomally-encoded protocatechuate extradiol ring fission pathway, which initiates metabolism by a 4,5 dioxygenase (not indicated). Adapted from various references listed in the text.

A

B

CHAPTER 2: CONTROL OF *cbaABC*-MEDIATED CATABOLISM OF 3-CHLOROBENZOATE

INTRODUCTION

Chlorinated benzoates are a common class of pollutants encountered in the environment, resulting from direct introduction (e.g. in the form of herbicides) or incomplete bacterial metabolism of some accidentally released chemicals (e.g. polychlorinated biphenyls; [Peel and Wyndham, 1997]). An assortment of bacterial pathways for their degradation has been characterized at the genetic and biochemical level (Figure 2.1), and for the most part these pathways proceed via intradiol ring fission of the diol intermediates, as for the “*ben*”¹-*clc* pathway for 3Cba degradation via chlorocatechol, the *cbd*-“*cat*” or *ohb*-“*cat*” pathway for 2Cba degradation via catechol, and the *fc**b*-“*pob-pca*” pathway for 4Cba degradation via 4-hydroxybenzoate and protocatechuate (Peel and Wyndham, 1999). Confirmed or potential regulatory factors for parts of the first three pathways have been identified on genes that are closely linked to the catabolic genes. They include ClcR, a LysR-type transcriptional regulator for *clc*, CbdR, an AraC-like regulator for *cbd*, and OhbR, an IciR/GylR-like regulator (Coco *et al.*, 1993; Haak *et al.*, 1995; Tsoi *et al.*, 1999). An alternative pathway that degrades 3Cba via 3-hydroxybenzoate to gentisate has also been described, but there is no information as of yet on the catabolic genes or their regulation (Krooneman *et al.*, 1998; Krooneman *et al.*, 1996).

Another metabolic pathway for chlorobenzoate degradation is the *cba-pmd* pathway for 3Cba, 4Cba, or 3,4-diCba degradation via protocatechuate (Pca) or 5-Cl-Pca and extradiol ring fission of these diol intermediates (Nakatsu *et al.*, 1997; Nakatsu and Wyndham, 1993; Providenti *et al.*, 1999b). Little is known about how expression of the upper part of the pathway is controlled and I therefore characterized regulation by investigating induction of 3Cba degradative activity and transcriptional control of *cbaABC*.

¹ The involvement of a *ben*-, *cat*- or *pca*-pathway was inferred from the host background and by characterization of metabolic intermediates. The degree of homology to sequenced and characterized *ben*-, *cat*-, and *pca* genes is not known, nor have the regulatory mechanisms of these pathways been confirmed.

Figure 2.1 Summary of degradation pathways for chlorobenzoates (Cba). Some bacteria funnel these substrates towards lower pathways that metabolize the diol intermediate by intradiol ring fission (A to D). Other bacteria have been shown to funnel these substrates to alternative lower pathways (E to F). The genes involved (not to scale) and their location (plasmid or chromosome) are indicated, if known, as is the bacterium in which the pathways have been described. Unlinked genes are not given directionality and are shown as solid lines. Adapted from various references listed in the text.

A. *P. putida* RB1 *ben-clc* pathway for 3Cba

chromosomally-encoded upper pathway

plasmid-encoded lower pathway

B. *B. cepacia* 2CBS *cbd-cat* pathway for 2Cba

plasmid-encoded upper pathway

chromosomally-encoded lower pathway

C. *P. aeruginosa* 142 *ohb-cat* pathway for 2Cba

upper pathway

lower pathway

D. *Arthrobacter globiformis* KZT1 *fcg-pob-pca* pathway for 4Cba

upper pathway

lower pathway

E. *Alcaligenes* sp. L6 degradation pathway for 3Cba

upper pathway

lower pathway

F. *C. testosteroni* BR60 *cba-pmd* pathway for 3Cba

plasmid-encoded upper pathway

chromosomally-encoded lower pathway

MATERIALS AND METHODS

Bacterial strains, plasmids, and *E. coli* growth conditions

Bacterial strains and plasmids described in this chapter are listed in Table 2.1. All bacterial media were prepared by standard methods (Gerhardt *et al.*, 1994). *E. coli* strains were grown at 37°C in Luria-Bertani medium (LB; 1% w/v tryptone, 0.5% w/v yeast extract, 0.5% w/v NaCl) containing ampicillin (250 mg/L), kanamycin (40 mg/L), or chloramphenicol (50 mg/L; Sigma-Aldrich Ltd., Oakville, Ontario, Canada), as required.

Classification of strain BR60 (pBRC60)

The identity of 3Cba-degrading strain BR60 (pBRC60), originally described as an *Alcaligenes* sp. (Wyndham and Straus, 1988), was not firmly established, therefore the first ~500 nt of the 16S rRNA gene were sequenced at MIDI Labs (Newark, Delaware, USA) using primers that correspond to positions 5 and 531 of the *E. coli* 16S rRNA gene.

DNA manipulation, sequencing, and sequence analysis

All enzymes used to generate clones described in Table 2.1 were purchased from New England Biolabs (Mississauga, Ontario, Canada) and used according to manufacturer's instructions. Plasmids were isolated and DNA was manipulated (e.g. agarose electrophoresis, ethidium bromide staining, gel purification of DNA) according to standard methods (Sambrook *et al.*, 1989). *E. coli* was transformed by electroporation using a Bio-Rad Gene Pulser according to the manufacturer's recommendations. All DNA sequences were determined by chain-terminating dideoxy sequencing on an ABI Prism Automated Sequencer (Biotechnology Research Institute, University of Ottawa, Ottawa, Ontario, Canada). Oligonucleotides were synthesized at the Biotechnology Research Institute.

Sequence of Tn5271 between IS1071L and *cbaA* was determined by sequencing relevant regions from clones pBRE11, pBRCW27, pBRCW34, and p60N.1 (Table 2.1). The DNA sequence of this region and deduced proteins were analyzed for similarities to entries in the GenBank non-redundant database using the BLASTN, BLASTX and BLASTP (Altschul *et al.*, 1990; Altschul *et al.*, 1997) network service of the National Center for Biotechnology Information (NCBI) at Bethesda, Maryland, USA

(<http://www.ncbi.nlm.nih.gov>). A divergently transcribed open reading frame (ORF) upstream of *cbaA* was identified and called *cbaR*. The putative product of this ORF showed similarity to various other transcriptional regulators that are known to respond to aromatic compounds or regulate catabolic genes for aromatic compounds (Chapter 3), therefore this gene was taken into consideration when designing some of the *lacZ* expression constructs described below. Downstream DNA had already been sequenced and was not taken into consideration as it contains ORFs (Figure 2.2) that show similarity to proteins involved in 3Cba uptake (Providenti *et al.*, 1999a) and a truncated aryl-CoA ligase (Di Gioia *et al.*, 1998).

Determination of *cbaA* transcriptional start site

The *cbaA* promoter was characterized by determining the transcriptional start site of *cbaA*. This was accomplished by primer extension analysis, as described elsewhere (Rothmel *et al.*, 1991a). In brief, *C. testosteroni* BR60 (pBRC60) was grown to mid-log phase in a minimal medium containing 4mM 3Cba. mRNA was isolated and a primer complementary to *cbaA* (5'-CATGAGGCCGCCCCATCG-3') was annealed to it and extended with MMLV Reverse Transcriptase (New England Biolabs). During extension, [α -³²P]dCTP (Amersham Canada Ltd., Oakville, Ontario, Canada) was included in the reaction so as to label the product. The extension product was resolved by electrophoresis on a denaturing sequencing gel (8% w/v polyacrylamide/8 M urea) along with a sequencing reaction conducted with the same primer using pMP14.10 as template.

Regulation of 3Cba degradative activity

Regulation of 3Cba degradative activity in whole cells of *C. testosteroni* BR60 (pBRC60) was studied by measuring O₂ uptake of cultures grown on a variety of aromatic carbon sources that induce the lower protocatechuate extradiol ring fission pathway (Figure 2.8), and succinate, a non-aromatic TCA cycle intermediate, which does not. Degradative activity of 3Cba-grown cells was measured by growing BR60 (pBRC60) in 1000 mL of minimal medium A (MMA) containing 4 mM 3Cba, 0.001% w/v tryptone, and 0.0005% w/v yeast extract (Wyndham, 1986). Tryptone and yeast extract are not necessary but improve growth. Cultures were incubated at 32°C and agitated by rotary

shaking (200 rpm) for 18 to 24 hours until mid- to late-log phase ($OD_{600} = 0.130$ to 0.350). They were then placed on ice for approximately 30 minutes, harvested at 4°C by centrifugation, washed once in ice cold MMA, resuspended in ice cold MMA (final $OD_{600} = 0.5$), and kept on ice until O_2 uptake was measured. O_2 uptake was measured with an Orion oxygen electrode on 23 mL aliquots of cells which had been equilibrated to 32°C . The endogenous O_2 uptake rate (measured over a 5 to 10 minute period) was subtracted from rates measured following addition of the various carbon sources (1 mM final concentration). One to three independent trials were performed for each carbon source and O_2 uptake was measured on replicate samples for each trial.

Degradative activity of cells grown on four unchlorinated, aromatic carbon sources that also induce the extradiol ring fission pathway of protocatechuate (4 mM benzoate, 3-hydroxybenzoate, 4-hydroxybenzoate, or protocatechuate; Figure 2.8) or a non-aromatic carbon source (10 mM succinate) was measured by first pregrowing BR60 (pBRC60) cells for approximately 18 hours in 100 mL of MMA containing 4mM 3Cba, 0.001% w/v tryptone, and 0.0005% w/v yeast extract. This maximized retention of Tn5271, which can be lost by homologous recombination if selective pressure is removed (Nakatsu *et al.*, 1991; Wyndham *et al.*, 1988). Cells were then harvested by centrifugation, washed once in MMA, and added to 1000 mL of MMA containing one of the other carbon sources, 0.001% w/v tryptone, and 0.0005% w/v yeast extract (initial OD_{600} between 0.030 and 0.060). Cells were then incubated to mid- to late-log phase (final OD_{600} between 0.200 and 0.350) and O_2 uptake measured as described above for cells grown on 3Cba.

As mentioned above, spontaneous deletion of Tn5271 occurs at a high frequency. Loss of the 3Cba degradative genes for cultures grown on carbon sources other than 3Cba was assessed by diluting and plating samples on solid MMA containing 4 mM 3Cba (Tn5271-proficient cells) or 10 mM succinate (total count). Within the incubation period of these experiments, there was no measureable loss of Tn5271, indicating that all cultures used in this study were Tn5271-proficient.

Determining whether *cbaABC*-proficient cells convert benzoate to a vicinal diol intermediate

According to O₂ uptake data, growth on benzoate induces 3Cba degradative activity (see Results, Table 2.2). The possibility that benzoate was being metabolized by the *cba* pathway was studied by a colourimetric plate assay which determines whether enzymes are present that convert an aromatic substrate into catechol or protocatechuate. Production of either vicinal diol intermediate results in a dark reddish-brown halo around the bacteria (Parke, 1992). *C. testosteroni* BR6020 with pBRCN1, pBRCN6, or pBRCN10 (Table 2.1) were cultured on solid minimal medium A (Wyndham, 1986) containing 10 mM succinate, 4 mM benzoate, 12 mg/L tetracycline, and amended 1 hr before plating of cells with 70 µL of a 50 mM aqueous FeCl₃ solution and 100 µL of a 0.1 M dimethylformamide solution of *p*-toluidine. Plates were incubated at 32°C. As a control, the same strains were cultured in a medium identical to the previous one except 4 mM 3Cba was used instead of benzoate.

Construction of gene expression constructs

***cbaA-lacZ* constructs**

Expression of *cbaABC* was studied with four fusion constructs of *cbaA* and *lacZ* (Figure 2.2). The *cbaA*'-'*lacZ* construct from which all ultimately stem was made by first cloning an 835 bp *NruI*-*ScaI* fragment from pBRE11 into the *SrfI* site of pCR-Script (SK⁺) so that the *ScaI* end was proximal to α-*lacZ*, yielding pBRCW34. *lacZ* lacking its natural promoter and the first nine amino acids was cloned downstream and in frame with *cbaA*' as a *Bam*HI-*Hind*III fragment from pBW30::Tn*phoA*'-4 into pBRCW34 digested with the same enzymes, resulting in pBRCW34L. This construct includes *cbaA*' fused to '*lacZ* and the putative promoter of *cbaABC*. *cbaRA*'-'*lacZ*, the expression construct which contains *cbaR* (Figure 2.2), was made by cloning a 1199 bp *Fsp*I-*Not*I fragment from pBRE11 into pBRCW34L which had been digested with *Sac*II (blunt-ended with Klenow) and *Not*I, resulting in pMP68.3. A 4.8 kb fragment released by digesting pMP68.3 with

Figure 2.2 A) Restriction map of 3-chlorobenzoate (3Cba) catabolic transposon Tn5271 and relative organization of genes and open reading frames. Indicated are *cbaABC*, which codes for the dioxygenase and dehydrogenase that convert 3Cba to protocatechuate (Pca) or 5-Cl-Pca, and *cbaR*, which may code for a repressor. Also indicated are *cbaDEFGH*, which may code for proteins involve in uptake of 3Cba, a potential stem-loop structure that may act as a transcriptional terminator, and ORF8, which codes for a truncated aryl-CoA ligase-like protein. B) Restriction map of *EcoRI* fragment 11 (E11), which was used to make all the *cbaA* and *cbaR* expression constructs, shown below in schematic format (see text for details). *lacZ* is not to scale. Abbreviations: IRL, inverted repeat left of IS1071; IRR, inverted repeat right of IS1071; *tnpA*, transposase of IS1071.

A) Tn5271

B) Fragment E11

NsiI and *HindIII* was blunt-ended with Klenow and then subcloned into pUTCm digested with *NotI* and also blunt-ended with Klenow, resulting in pMP70.5.

The remaining three *cbaA'*-*lacZ* expression constructs did not include *cbaR* but contained decreasing amounts of DNA upstream of the *cbaA* start codon (Figure 2.2). *cbaA'*-*lacZ*1, the expression construct containing upstream DNA until the *NotI* site, was made by sub-cloning the 3.7 kb *BglII* fragment from pBRCW34L into the *BamHI* site of pUC18Not, resulting in pMP14.10. A 3.4 kb *NotI* fragment was then sub-cloned into the *NotI* site of pUTCm, resulting in pMP17.6. Constructs *cbaA'*-*lacZ*2 and *cbaA'*-*lacZ*3 were created by PCR-generated deletions of upstream DNA in ~30 bp intervals (Figure 2.3) using the 3.4 kb *NotI* fragment from pMP14.10 as a template. Both PCR-generated fragments were amplified with Vent DNA Polymerase in a 100 µL reaction volume containing 200 µM of each dNTP and 20 pmol of forward and reverse primer (see below). Fragments were amplified under the following conditions: 5 minute hot start at 95°C followed by 30 cycles of 30 seconds at 48°C for primer annealing, 30 seconds at 72°C for extension, and 60 seconds at 94°C for template melting. The “deleted” promoter of *cbaA'*-*lacZ*2 was generated by amplifying a 298 bp fragment with forward primer PCBA4478 (5'-ggctgcaGGCTCGGGTAAACACCTAGA-3') and reverse primer LACZ65 (5'-CGATTAAGTTGGGTAACG-3'; Appendix D). Lower case nucleotides in PCBA4478 are not native to the putative *cbaA* promoter but contain the restriction site for *PstI* (underlined). LACZ65 is complementary to a region approximately 65 nt downstream of the DNA coding for the N-terminus of *lacZ*. The product was digested with *PstI* and *BamHI* (present internally at the *cbaA-lacZ* junction; Appendix D) and the 240 bp fragment was cloned into pMP28.1 digested with the same enzymes, resulting in pMP111.1. This placed *cbaA'* upstream of and in frame with *lacZ*. The 3.3 kb *NotI* fragment from pMP111.1 was then subcloned into pUTCm digested with *NotI*, resulting in pMP112.3. The “deleted” promoter of *cbaA'*-*lacZ*3 was generated by amplifying a 268 bp fragment with forward primer PCBA4510 (5'-ggctgcaggAAACCTCCGTATAGTTAC-3') and reverse primer LACZ65. As above, the lower case nucleotides of PCBA4510 are not native to the construct and contain the restriction site for *PstI* (underlined). The

product was digested with *Pst*I and *Bam*HI and the 210 bp fragment was cloned into pMP28.1 digested with the same enzymes, resulting in pMP117.1. The 3.3 kb *Not*I fragment from pMP117.1 was then subcloned into pUTCm digested with *Not*I, resulting in pMP118.3.

A control expression construct was made by cloning a promoterless '*lacZ*' as a *Not*I fragment from pMP10.1 into the *Not*I site of pUTCm, resulting in pMP19.9. All pUTCm-based vectors contain the insert oriented so that *lacZ* is transcribed in the opposite direction of the *Cm^r* gene (Appendix B). The nucleotide sequence of the junction between *cbaA'* and '*lacZ*' in pBRCW34L (Appendix D) and of the PCR-generated fragments in pMP111.1 and pMP117.1 were confirmed by chain terminating dideoxy sequencing .

***cbaR-lacZ* constructs**

Expression of *cbaR* was studied with three transcriptional fusion constructs of *cbaR* and *lacZ*. Each contained decreasing amounts of DNA upstream of the *cbaR* start codon (Figure 2.2). The *cbaR'*-*lacZ* construct from which all three ultimately stem was made by cloning an 835 bp *Nru*I-*Sca*I fragment from pBRE11 into the *Srf*I site of pCR-Script (SK+) so that the *Nru*I end was proximal to α -*lacZ*. This resulted in pBRCW27. '*lacZ*' was cloned downstream and in frame of *cbaR'* by removing '*lacZ*' as a *Bam*HI-*Hind*III fragment from pBW30::Tn*phoA'*-4 and inserting it into pBRCW27 digested with the same enzymes, resulting in pBRCW27L. *cbaR'*-*lacZ*1, the expression construct containing DNA upstream of *cbaR* until the *Sca*I site in *cbaA*, was created by first removing the 2.7 kb *Sac*I fragment from pBRCW27L and cloning it into pMP28.1 digested with *Sac*I, resulting in pMP105.1. The 2.7 kbp fragment contained the complete 835 bp insert and the first 1.9 kb of *lacZ*. pMP28.1 digested with *Sac*I released the first 1.9 kbp of *lacZ* but retained the last 1.2 kb. The 2.7 kb fragment was cloned into pMP28.1 such that a complete *lacZ* was reconstructed. The 3.9 kb *Sac*II-*Xba*I fragment from pMP105.1 containing the complete 835 bp insert and *lacZ* gene was sub-cloned into pUC128 digested with *Sac*II and *Xba*I, resulting in pMP123.2. pMP123.2 was digested with *Nsi*I and partially digested with *Not*I and the 3.9 kb fragment containing the complete

835 bp insert was isolated and sub-cloned into pUTCmMCS.3 digested with *Nsi*I and *Not*I, resulting in pMP125.1. *cbaR'*-*'lacZ*2, the expression construct containing DNA upstream of *cbaR* until the *Not*I site, was created by removing the 3.7 kb *Not*I fragment from pMP105.1 and sub-cloning it into pUTCm digested with *Not*I, resulting in pMP106.1. *cbaR'*-*'lacZ*3, the expression construct containing DNA upstream of *cbaR* until the *Bgl*II site, was created by first sub-cloning the 3.3 kb *Bgl*II fragment from pBRCW27L into pUC18Not digested with *Bam*HI, resulting in pMP50.5. The 3.3 kb *Not*I fragment was then removed and sub-cloned into pUTCm digested with *Not*I, resulting in pMP54.9.

All pUTCm-based vectors contain the insert oriented so that *lacZ* is transcribed in the opposite direction of the *Cm^r* gene (Appendix C). The junction between *cbaR'* and *'lacZ* in pBRCW27L was confirmed by chain terminating dideoxy sequencing (Appendix D).

Regulation of *cbaA* and *cbaR* expression

Insertion of expression constructs

The various expression constructs cloned into the suicide delivery vector pUTCm were mobilized from *E. coli* CC118 λ *pir* and inserted into the chromosome of *C. testosteroni* BR6020 by triparental mating (Nakatsu and Wyndham, 1993). BR6020 is a pBRC60-cured variant derived from BR60 (pBRC60) and allows *cbaA* expression to be measured in a genetic Tn5271-deficient background. Transconjugants were recovered by plating cells on solid MMA (Wyndham, 1986) containing 4 mM 4-hydroxybenzoate, 50 mg/L chloramphenicol, and 100 mg/L 5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside (X-gal). Because insertion of the expression constructs into the chromosome is reported to be random, we tested whether the location of transposition affected LacZ levels. Four independent transconjugants were isolated for each construct and LacZ levels of succinate-grown cells were measured (described below). Because levels could vary by as much as two fold between independent transconjugants containing the same construct, the transconjugant that had the lowest LacZ activity was selected as the representative for that construct and used in all subsequent expression studies. A Tn5271-proficient background

was created for the strains containing an expression construct by transferring pBRC60 from *C. testosteroni* BR6024 (pBRC60) and recovering transconjugants on solid MMA containing 4 mM 3Cba, 50 mg/L chloramphenicol, and 100 mg/L X-gal (Nakatsu and Wyndham, 1993).

Confirmation that constructs inserted as a single copy via transposition

Random, single insertion via transposition of the *cbaA* and *cbaR* expression constructs into the chromosome of BR6020, as opposed to multiple insertion, site-specific insertion, or integration of the whole vector (de Lorenzo and Timmis, 1994), was confirmed by Southern blotting, as follows. Chromosomal DNA was isolated from the various BR6020 strains (Ausubel *et al.*, 1989), digested with *Bam*HI (see Results for rationale), electrophoresed in 0.9% agarose, depurinated by illuminating the gel with ultraviolet light for approximately 10 minutes, and then transferred to a positively charged nylon membrane (Roche Molecular Biochemicals, Laval, Québec, Canada) with a VacuGene XL vacuum blotter (Amersham-Pharmacia Biotech, Baie d'Urfé, Québec, Canada), according to manufacturer's instructions. Blots were probed under high stringency conditions with DNA fragments that had been randomly labelled with digoxigenin, according to manufacturer's instructions (Roche Molecular Biochemicals). A *lacZ* probe (obtained by digesting pMP28.1 with *Not*I and gel-purifying the ~3.5 kb fragment) was used to confirm that constructs had inserted only once and that insertion was random. A probe for *oriT* of pUTCm (obtained by digesting pUTCm with *Bam*HI and gel-purifying the ~1.8 kb fragment) was used to distinguish between transposition of the construct (*oriT*⁻) or integration of the whole vector (*oriT*⁺).

Measurement of gene expression levels

Gene expression levels for cells with the various *cbaA-lacZ* and *cbaR-lacZ* constructs were determined with cells grown essentially as described for O₂ uptake studies (e.g. incubated at 32°C with rotary shaking at 200 rpm; see above), except cultures were grown in 30 mL of medium. Tn5271-deficient cells were grown until mid- to late-log phase (OD₆₀₀=0.2 to 0.35) on the test carbon source. Tn5271-proficient cells were first

pre-grown on 3Cba to mid-log phase to ensure that all cells contained Tn5271, subcultured into medium containing the test carbon source, and incubated until mid- to late-log phase. A 1 mL aliquot of cells was removed, placed on ice for at least 30 minutes, and 200 μ L were assayed for β -galactosidase as described in Sambrook *et al.* (1989). At least three independent trials were performed for each carbon source tested, and for each trial, a 1 mL sample was removed from the growing cultures at two separate time points and expression was measured on three 200 μ L aliquots from each sample. Results were averaged and represent the mean for one trial. The statistical significance of the two observed effects (induction levels relative to succinate-grown cells and the reduction in *cbaA* expression levels following introduction of pBRC60 [see Results]) was determined via a paired t-Test of the averages of the data from the three trials ($p=0.05$).

Table 2.1. Bacterial strains and plasmids. See text for greater detail on cloning strategy for the various expression constructs. Schematic diagrams of the expression constructs are provided in Appendix B and C.

Strain and plasmid	Relevant phenotypic/genotypic characteristics ¹	Reference
<i>E. coli</i>		
MM294 (pRK2013)	Km ^r , RK2 tra. Used in triparental mating to mobilize pUTCm-based plasmids	(Figurski and Helinski, 1979)
XL-1Blue ² (pCR-Script SK(+))	Amp ^r , α -lacZ/MCS	Stratagene
JM101 (pUC18)	Amp ^r , α -lacZ/MCS	(Yanisch-Perron <i>et al.</i> , 1985)
CC118 (pUC18Not)	Amp ^r , α -lacZ/MCS. Identical to pUC18 except with <i>NotI</i> sites flanking either end of MCS	(de Lorenzo and Timmis, 1994)
TB1 (pNEB193)	Amp ^r , α -lacZ/MCS	New England Biolabs
DH5 α (pUC128)	Amp ^r , α -lacZ/MCS	(Keen <i>et al.</i> , 1988)
CC118:: λ pir (pUTCm)	Amp ^r , Cm ^r . Tn5-based suicide delivery vector	(de Lorenzo <i>et al.</i> , 1990)
CC118:: λ pir (pUTCmMCS.3)	Amp ^r , Cm ^r . pUTCm with oligonucleotide containing additional restriction sites cloned into <i>NotI</i> site ³	this study
BW16478 (pBW30::TnphoA'-4)	Amp ^r , Km ^r . Vector used as source of promoterless lacZ	(Wilmes-Riesenberg and Wanner, 1992)
XL-1Blue (pMP10.1)	Amp ^r . Promoterless <i>lacZ</i> missing first nine amino acids from pBW30::TnphoA'-4 cloned as a <i>Bam</i> HI- <i>Hind</i> III fragment into pUC18Not digested with <i>Bam</i> HI- <i>Hind</i> III	this study
XL-1Blue (pMP28.1)	Amp ^r . Promoterless <i>lacZ</i> missing first nine amino acids from pBW30::TnphoA'-4 cloned as a <i>Bam</i> HI- <i>Bgl</i> II fragment into pUC18Not digested with <i>Bam</i> HI. <i>LacZ</i> is inserted such that it is transcribed in the same direction as α -lacZ	this study
XL-1Blue (p60N.1)	Amp ^r . <i>Nhe</i> I fragment from pBRC60 containing Tn5271 without the flanking IS1071. Cloned into <i>Xba</i> I site of pNEB193	this study
JM101 (pBRE11)	Amp ^r . <i>Eco</i> RI fragment 11 from pBRC60 cloned into pUC18	(Nakatsu <i>et al.</i> , 1991)

¹ Abbreviations: MCS, multiple cloning site; Amp, ampicillin; Km, kanamycin; Cm, chloramphenicol; r, resistant; Trp-, tryptophan auxotroph; Tn5271, composite transposon containing *cba* genes for growth on 3Cba

² *E. coli* XL-1 Blue is Km^r

³ The sequence of the modified cloning site in pUTCmMCS.3 is 5'-GCGGCCGCAA GCTTGCATGC CTGCAGGTCG ACTCTAGAGG ATCCCCCGGG GGATTAAAT ATGCATCGGC CGC-3'. This sequence contains restriction sites for *NotI*, *Hind*III, *Sph*I, *Pst*I, *Sal*I, *Xba*I, *Bam*HI, *Sma*I, *Swa*I, *Nsi*I, and *Eag*I. Only the restriction sites for *NotI*, *Swa*I, and *Nsi*I are unique to this vector. The 5' end is proximal to the Cm^r gene.

Table 2.1 (continued)

Strain and plasmid	Relevant phenotypic/genotypic characteristics ¹	Reference
<i>cbaR</i> '- <i>lacZ</i> and <i>cbaA</i> '- <i>lacZ</i> constructs		
XL-1Blue (pBRCW34)	Amp ^r . <i>NruI</i> - <i>ScaI</i> fragment from pBRE11 cloned into pCR-Script SK(+) so that <i>ScaI</i> end (i.e. <i>cbaA</i> ') is adjacent to α - <i>lacZ</i>	this study
XL-1Blue (pBRCW34L)	Amp ^r . <i>lacZ</i> cloned downstream of <i>NruI</i> - <i>ScaI</i> fragment in pBRCW34, placing it in frame with <i>cbaA</i> '	this study
XL-1Blue (pMP68.3)	Amp ^r . Contains construct <i>cbaR</i> '- <i>lacZ</i> . <i>FspI</i> - <i>NotI</i> fragment from pBRE11 cloned into pBRCW34L so that DNA containing <i>cbaR</i> , a divergently transcribed open reading frame between <i>cbaA</i> and IS1071L, is restored upstream of <i>cbaA</i> '	this study
CC118:: λ pir (pMP70.5)	Amp ^r , Cm ^r . <i>NsiI</i> - <i>HindIII</i> fragment from pMP68.3 cloned into pUTCm	this study
XL-1Blue (pMP14.10)	Amp ^r . Contains construct <i>cbaA</i> '- <i>lacZ</i> 1. <i>BglIII</i> fragment from BRCW34L cloned into pUC18Not	this study
CC118:: λ pir (pMP17.6)	Amp ^r , Cm ^r . <i>NotI</i> fragment from pMP14.10 cloned into pUTCm	this study
CC118:: λ pir (pMP111.1)	Amp ^r . Contains construct <i>cbaA</i> '- <i>lacZ</i> 2	this study
CC118:: λ pir (pMP112.3)	Amp ^r , Cm ^r . <i>NotI</i> fragment from pMP111.1 cloned into pUTCm	this study
CC118:: λ pir (pMP117.1)	Amp ^r . Contains construct <i>cbaA</i> '- <i>lacZ</i> 3	this study
CC118:: λ pir (pMP118.1)	Amp ^r , Cm ^r . <i>NotI</i> fragment from pMP117.1 cloned into pUTCm	this study
<i>cbaR</i> '- <i>lacZ</i> constructs		
XL-1Blue (pBRCW27)	Amp ^r . <i>NruI</i> - <i>ScaI</i> fragment from pBRE11 cloned into pCR-Script SK(+) so that <i>NruI</i> end (i.e. <i>cbaR</i> ') is adjacent to α - <i>lacZ</i>	this study
XL-1Blue (pBRCW27L)	Amp ^r . <i>lacZ</i> cloned downstream of <i>ScaI</i> - <i>NruI</i> fragment in pBRCW27, placing it in frame with <i>cbaR</i> '	this study
XL-1Blue (pMP105.1)	Amp ^r . <i>SacI</i> fragment from pBRCW27L cloned into pMP28.1	this study
CC118:: λ pir (pMP123.2)	Amp ^r . <i>SacII</i> - <i>XbaI</i> fragment from pMP105.1 cloned into pUC128	this study
CC118:: λ pir (pMP125.1)	Amp ^r , Cm ^r . Contains construct <i>cbaR</i> '- <i>lacZ</i> 1. <i>NsiI</i> - <i>NotI</i> fragment from pMP123.2 cloned into pUTCmMCS.3	this study
CC118:: λ pir (pMP106.1)	Amp ^r , Cm ^r . Contains construct <i>cbaR</i> '- <i>lacZ</i> 2. <i>NotI</i> fragment from pMP105.1 cloned into pUTCm	this study
XL-1Blue (pMP50.5)	Amp ^r . <i>BglIII</i> fragment from pBRCW27L cloned into pUC18Not	this study
CC118:: λ pir (pMP54.9)	Amp ^r , Cm ^r . Contains construct <i>cbaR</i> '- <i>lacZ</i> 3. <i>NotI</i> fragment from pMP50.5 cloned into pUTCm	this study
Control <i>lacZ</i> construct		
CC118:: λ pir (pMP19.9)	Amp ^r , Cm ^r . <i>lacZ</i> from pMP10.1 cloned into pUTCm.	this study

¹Abbreviations: Amp, ampicillin; Cm, chloramphenicol; r, resistant; *cbaR*, gene for regulatory protein involved in controlling *cbaABC*.

Table 2.1 (continued)

Strain and plasmid	Relevant phenotypic/genotypic characteristics ¹	Reference
<i>C. testosteroni</i>		
BR60 (pBRC60)	Wild type. Tn5271-proficient	(Wyndham <i>et al.</i> , 1988)
BR6020	BR60 cured of pBRC60. Tn5271-deficient	(Wyndham <i>et al.</i> , 1988)
BR6024 (pBRC60)	Trp ⁺ and Cm ^r strain derived from BR6020. Tn5271-proficient. Used to conjugate pBRC60 into BR6020 strains	(Ng and Wyndham, 1993)
BR6020 (pBRCN1)	Tet ^r . Part of Tn5271 cloned as a <i>HpaI</i> fragment into vector pBW13. Contains all <i>cba</i> genes but lacks complete flanking IS1071 elements	(Nakatsu and Wyndham, 1993)
BR6020 (pBRCN5)	Tet ^r . Part of Tn5271 cloned as a <i>PstI-HindIII</i> fragment into fragment into vector pBW13. Contains <i>cbaABCDE</i> ¹	(Nakatsu and Wyndham, 1993)
BR6020 (pBRCN10)	Tet ^r . Contains only <i>cbaABC</i> . Generated by exonuclease digestion of pBRCN5	(Nakatsu and Wyndham, 1993)
BR6020::(<i>cbaRA'</i> - <i>lacZ</i>)	Cm ^r , CbaR+. Contains single chromosomal insert of <i>cbaRA'</i> - <i>lacZ</i> delivered by pMP70.5	this study
BR6020::(<i>cbaRA'</i> - <i>lacZ</i>) (pBRC60)	Cm ^r , CbaR+, Tn5271-proficient. Previous strain with pBRC60. Note that CbaR is provided both by the construct and pBRC60	this study
BR6020::(<i>cbaA'</i> - <i>lacZ</i> 1)	Cm ^r . Contains single chromosomal insert of construct <i>cbaA'</i> - <i>lacZ</i> 1 delivered by pMP17.6	this study
BR6020::(<i>cbaA'</i> - <i>lacZ</i> 1) (pBRC60)	Cm ^r , Tn5271-proficient. Previous strain with pBRC60	this study
BR6020::(<i>cbaA'</i> - <i>lacZ</i> 2)	Cm ^r . Contains single chromosomal insert of construct <i>cbaA'</i> - <i>lacZ</i> 2 delivered by pMP112.3	this study
BR6020::(<i>cbaA'</i> - <i>lacZ</i> 2) (pBRC60)	Cm ^r , Tn5271-proficient. Previous strain with pBRC60	this study
BR6020::(<i>cbaA'</i> - <i>lacZ</i> 3)	Cm ^r . Contains single chromosomal insert of construct <i>cbaA'</i> - <i>lacZ</i> 3 delivered by pMP118.1	this study
BR6020::(<i>cbaA'</i> - <i>lacZ</i> 3) (pBRC60)	Cm ^r , Tn5271-proficient. Previous strain with pBRC60	this study
BR6020::(<i>cbaR'</i> - <i>lacZ</i> 1)	Cm ^r . Contains single chromosomal insert of construct <i>cbaR'</i> - <i>lacZ</i> 1 delivered by pMP125.1	this study
BR6020::(<i>cbaR'</i> - <i>lacZ</i> 1) (pBRC60)	Cm ^r , Tn5271-proficient. Previous strain with pBRC60	this study
BR6020::(<i>cbaR'</i> - <i>lacZ</i> 2)	Cm ^r . Contains single chromosomal insert of construct <i>cbaR'</i> - <i>lacZ</i> 2 delivered by pMP106.1	this study
BR6020::(<i>cbaR'</i> - <i>lacZ</i> 2) (pBRC60)	Cm ^r , Tn5271-proficient. Previous strain with pBRC60	this study
BR6020::(<i>cbaR'</i> - <i>lacZ</i> 3)	Cm ^r . Contains single chromosomal insert of construct <i>cbaR'</i> - <i>lacZ</i> 3 delivered by pMP54.9	this study
BR6020::(<i>cbaR'</i> - <i>lacZ</i> 3) (pBRC60)	Cm ^r , Tn5271-proficient. Previous strain with pBRC60	this study
BR6020:: <i>lacZ</i>	Cm ^r . Contains single chromosomal insert of promoterless <i>lacZ</i> delivered by pMP19.9	this study
BR6020:: <i>lacZ</i> (pBRC60)	Cm ^r , Tn5271-proficient. Previous strain with pBRC60	this study

¹Abbreviations: Cm, chloramphenicol; Tet, tetracycline; r, resistant; Trp⁺, tryptophan auxotroph; Tn5271, composite transposon containing *cbaABC*, the catabolic genes necessary for growth on 3Cba; *cbaR*, an upstream divergently oriented gene coding for a repressor; and *cbaDEFGH*, downstream genes that may code for uptake proteins.

RESULTS

Classification of strain BR60 (pBRC60)

The identity of strain BR60 (pBRC60), which was not firmly established, was determined by sequencing a portion of the 16S rRNA gene, and it is identical to that of *C. testosteroni* type strain (Appendix A). Because BR60 (pBRC60) has several genetic and metabolic similarities to other *C. testosteroni* strains (see Discussion and original characterization in Wyndham and Straus [1988]), *Alcaligenes* sp. BR60 (pBRC60) is reclassified as *C. testosteroni* BR60 (pBRC60).

Identification of *cbaR* and determination of *cbaA* transcriptional start site

Genes coding for regulatory elements controlling catabolic genes are often closely associated with the genes they control. The DNA sequence from the *Nsi*I site in IS1071L to the *Sca*I site in *cbaA* was therefore determined and results are presented in Figure 2.3. BLASTX analysis revealed the presence of a divergently oriented open reading frame, designated *cbaR*, beginning 567 bp upstream of *cbaA* (Figure 2.3) that codes for a 19.4 kDa protein that shows significant amino acid similarity to various members of the MarR family of transcriptional regulators (Chapter 3). MarR is a repressor involved in controlling expression of the multiple antibiotic resistance (*mar*) genes of *E. coli* and responds to aromatic compounds like 2-hydroxybenzoate (Ariza *et al.*, 1994; Cohen *et al.*, 1993; Seoane and Levy, 1995), and because many other proteins showing similarity to it also respond to specific aromatic compounds or are involved in regulating genes for metabolism of aromatic compounds, we focussed on the effect of *cbaR* in studies investigating the expression of *cbaA*. The functional characterization of CbaR is described in Chapter 3.

The *cbaA* promoter was characterized by determining the transcriptional start site for this gene (Figure 2.4) and it is an adenine 99 bp upstream of the putative start codon (Figure 2.3). A hexamer starting at -12, TATAGT, is virtually identical to the consensus enterobacterial -10 hexamer for σ^{70} -dependent genes (TATAAT). No sequence similar to TTGACA, the consensus enterobacterial -35 hexamer, is apparent, nor are any features of

Figure 2.3 Nucleotide sequence of the upper strand of Tn5271 from the *Nsi*I site in IS1071L to the *Sca*I site in *cbaA*. Nucleotide 1 is defined as the first G in IS1071L (not shown). Indicated are several features of *cbaR* (deduced translation product and putative ribosome binding site [RBS]) and *cbaA* (putative -10 hexamer, transcriptional start site (+1), and putative RBS). Also indicated are the right inverted repeat of IS1071L (nucleotides in bold); relevant restriction sites (underlined); and the boundaries of all *cbaA-lacZ* expression constructs.

3030 *NsiI*
 |==> *cbaRA*'-'*lacZ* Right inverted repeat of IS1071L
 ATGCATCGATAACTGGCACCAATGGATGAAAACCGACCACGGTTTGACGCCACGAATCGCAGATTTGAAAAGTGAACAGGAA
 3110 *NheI*
 AGTCAATGAAATCAACGATCTACCAACACCACCTCCGCGCCAGTGTAGCTTTTCGTACCGTCACTTATTGCACTGAAAA
 3190
 CGAGGAGACCCCATGCCCTGCTCGACGCCCTTGCTGGGCTGCGTGTGCTGCTGAGCCGGTGTACCTCGATGTCCCCGAG
 3270 *NruI*
 CCCAACCGCGAGGCCATCGCGATGGCCGAGCGCCAGGGCATGGTGCAGGGCTTTCCGACCGCGCGCATGTACCGGGGCC
 3350
 GCGCGCCGCGCTGGACCTGGGGCGTTGTTATGCCCTTGACGGCCTTGGAGGTGCGCTGAGCCTGGTGGCTTGATCAGGGCA
 3430 *NruI*
 GCGCTTGGCG CTA CTC CTG AGG AGA TTC GCG ATG CCA GTG CGC CTT GAC GTA GAG CGA AGA
 * Glu Gln Pro Ser Glu Arg His Trp His Ala Lys Val Tyr Leu Ser Ser
 3491
 ATC GAT ATC CAC CCT GTT GAT GTT CTC CTT CAG GCG CTG GAG AAG GCG AAC CAG GGT CTC
 Asp Ile Asp Val Arg Asn Ile Asn Glu Lys Leu Arg Gln Leu Leu Arg Val Leu Thr Glu
 3551 *NruI*
 TCG CTC CGA GGC GGT CAG CAC AGA CAG GAG CGA ATT TTC GCG ATG GAG CGC GTA CAG AAT
 Arg Glu Ser Ala Thr Leu Val Ser Leu Leu Ser Asn Glu Arg His Leu Ala Tyr Leu Ile
 3611
 GAT CTT GTC GTG CAT TGC CTG GCC TTC TTC GGT CAG CTG CGT GAG CCT GAC CCG CTT GTC
 Ile Lys Asp His Met Ala Gln Gly Glu Glu Thr Leu Gln Thr Leu Arg Val Arg Lys Asp
 3671
 GCG GTC CGC CGT TGC CAG AAC GAT CAG GCC CTT GGC CAA CAT TTG CTT CAT CAC CTT GCT
 Arg Asp Ala Thr Ala Leu Val Ile Leu Gly Lys Ala Leu Met Gln Lys Met Val Lys Ser
 3731
 GAC GGA GCC CTT GTC CAT GCC AGT GAA CTG GCA AAT GCT CTG TGC CGT GGT TTC GCC TGC
 Val Ser Gly Lys Asp Met Gly Thr Phe Gln Cys Ile Ser Gln Ala Thr Thr Glu Gly Ala
 3791
 CAC TGC TAT GAG CGC CAG GAC CCG CCA GGT CTC GAT GCC GAT GCC GAA ATT CTT TAA ATA
 Val Ala Ile Leu Ala Leu Val Arg Trp Thr Glu Ile Gly Ile Gly Phe Asn Lys Leu Tyr
 3851 *NruI*
 GTC TTT TGA GGC ACC GCT CGC GAG CTT GTT GGC GAT GGC GGT AAA AAA TGC CGG GGT GTA
 Asp Lys Ser Ala Gly Ser Ala Leu Lys Asn Ala Ile Ala Thr Phe Phe Ala Pro Thr Tyr
 3911 RBS
 GGT ATC GAG ATT GAG GGT CAC CCG GCT TTT TCG AGG ATC TCT TGC AAG CAT TCAGTATCCAGC
 Thr Asp Leu Asn Leu Thr Val Arg Ser Lys Arg Pro Asp Arg Ala Leu Met <CbaR
 3974
 GTTGATGCGTGTAGGCCCCAGACTCGGTGGGGCGGTAAAAGCCATAAACCCCTTTCCGGATACCGCTGCAACCAGCTATC
 4054 *BglII PstI*
 GACCCCGTTGACATGGCGCTTGGCGCAGGCGCGTGTGAAAGAGATCTGCAGGGATCACAACAATGCTGCGCGGTCTCT
 4134
 CTGTCTCCTCCGGGCCATTGCGAGCAGGCTCATGATTCGATGCCGAGCTTAGCACCTTGGTTCTGAGCAATGGCCAGGGT
 4214
 CCTTTGGTTCTGATTGGCCACCAACAGCATTGCGCGTGACGCCCCATTGAGCGCCGCGCTCCTCCGGAGACCGGCAATTC
 4294
 CCTCGGCCGGTGACCGGCTGGCGCTGCCGATTGCGCGCTTCTCCTCCCGCCCGGCTTGCCGCTGCCCCGACGCGATC
 4374 *cbaA*'-'*lacZ1* |==> *NotI*
 TGGCGGCGCTCGGCATGAGCCGCACTGAGCGGCAGCGCGCGCCGCCACCGCGCTTCACGCGCCGCACAGCGGCGCT
 4454 *cbaA*'-'*lacZ2* |==> *cbaA*'-'*lacZ3* |==> -10 +1
 GCGCGCTGCATCACCTGCCGCTCGGCTCGGGTAAACACCTAGATTTTCGATCTGAAAACCTCCGTATAGTTACAAAAC
 4534
 CAACTACATGGATCGAATCATCTCAGGTTTTCCCTCATAAGGGTTTGACACAAGTTGCTAAACCAACACATCCGCAAA
 4614 RBS
 ATTTCAGGAGCAAGCT ATG GCA TTG ACC AAG GAC AAC GAA GAT TTG GTG CGA GTG GGC CGT GGA
 CbaA=> Met Ala Leu Thr Lys Asp Asn Glu Asp Leu Val Arg Val Gly Arg Gly
 4678 *ScaI*
 ACC CCG ATG GGC GGC CTC ATG CGC GAG TAC T
 Thr Pro Met Gly Gly Leu Met Arg Glu Tyr

Figure 2.4 Transcriptional initiation site (+1) of *cbaA*, as determined by primer extension analysis of mRNA isolated from *C. testosteroni* BR60 (pBRC60). The extension product (indicated) was loaded onto a sequencing gel next to the four sequencing reactions (lanes GTCA) conducted with the same primer.

A C T G

Primer extension
product

5' G T A T A G T T A C A A A A C C A A C T A C A 3'

a σ^{34} -dependent promoter, such as the invariant GG and GC at -24 and -12, respectively (Steinbüchel *et al.*, 1992).

Induction of 3Cba degradative activity

Preliminary investigations of factors controlling 3Cba catabolism by *C. testosteroni* BR60 (pBRC60) were conducted by measuring O₂ uptake rates with whole cells grown on a variety of aromatic carbon sources that also induce the protocatechuate extradiol ring fission pathway (designated *pmd*). As a control, 3Cba degradative activity was determined with cells grown on succinate, a TCA cycle intermediate which does not induce *pmd* (Table 2.2). Maximal 3Cba degradative activity was induced by growth on 3Cba. Growth on benzoate also induced 3Cba degradative activity (~25% of the maximum). Similarly, benzoate degradative activity was induced by growth on 3Cba (~22% of the maximum observed following growth on benzoate). Growth on 4-hydroxybenzoate, 3-hydroxybenzoate, and Pca (aromatic carbon sources that are channelled towards *pmd*; Figure 2.8) and succinate did not induce 3Cba or benzoate degradative activity.

Oxygen uptake rates following growth on succinate, 3Cba, and the hydroxylated benzoates showed good reproducibility between independent trials, with standard errors often less than 10% (Table 2.2). However, for unknown reasons, benzoate-grown BR60 (pBRC60) cultures exhibited unpredictable lag phases (24 to 48 hrs) between trials before initiation of logarithmic growth. This apparently led to variable oxygen uptake rates towards benzoate, although not towards the other carbon sources (Table 2.2). The highest rates were achieved in the trial where log phase was achieved within 24 hrs, approximately the same time period necessary for the other growth substrates (i.e. ~18 hrs).

Table 2.2. Oxygen consumption rates by *C. testosteroni* BR60 (pBRC60) towards various carbon sources. Rates are corrected for endogenous oxygen uptake. Averages of the means from the independent trials (with standard errors in brackets) are presented.

Growth substrate	O ₂ uptake rate (mg O ₂ L ⁻¹ min ⁻¹) when exposed to:					
	3-chlorobenzoate	benzoate	3-hydroxybenzoate	4-hydroxybenzoate	protocatechuate	succinate
3-chlorobenzoate	1.10 (0.05)	0.69 (0.02)	0.06 (0.01)	0.40 (0.03)	1.36 (0.05)	0.13 (0.01)
Benzoate	0.27 (0.02)	0.54 (0.08) to 3.08 (0.31) ¹	0.13 (0.01)	0.25 (0.05)	n.d. ²	0.12 (0.02)
3-hydroxybenzoate	<0.05	<0.05	3.07 (0.20)	<0.05 to 0.60 (0.01) ³	1.36 (0.04)	0.35 (0.02)
4-hydroxybenzoate	<0.05	<0.05	<0.05 to 0.40 (0.01) ⁴	3.51 (0.13)	1.47 (0.04)	0.22 (0.02)
Protocatechuate	<0.05	<0.05	<0.05	<0.05	1.15 (0.04)	0.22 (0.01)
Succinate	<0.05	<0.05	<0.05	<0.05	n.d.	0.74 (0.03)

¹ highly variable activity was observed between the three trials and the range is shown

² n.d., not determined. However, for benzoate-grown cultures, protocatechuate degradative activity is at levels comparable to those of 3Cba-grown cultures (Nakatsu and Wyndham, 1993).

³ in two of the three trials, activity >0.05 mg O₂/L/min was observed.

⁴ in one of the three trials, activity >0.05 mg O₂/L/min was observed.

Determining whether *cbaABC*-proficient cells convert benzoate to a vicinal diol intermediate

The possibility that the *cba* pathway could convert benzoate to vicinal diol intermediates was studied by a colourimetric plate assay. No colour indicative of catechol or protocatechuate was produced by various *cbaABC*-proficient cultures (BR6020 with pBRCN1, pBRCN5, or pBRCN10) grown on medium containing benzoate, even after 7 days incubation. Growth of these cultures on medium containing 3Cba resulted in a dark red colour within 24 hrs.

Regulation of *cbaA* and *cbaR* expression

Confirmation that single copies of expression constructs transposed into BR6020 chromosome

Random transposition of single copies of the *cbaA-lacZ* and *cbaR-lacZ* expression constructs was confirmed by Southern blots of *Bam*HI-digested chromosomal DNA from all transconjugants. Transposition as opposed to integration was shown by using a probe against the *oriT* of the suicide delivery vector. This region should be present if integration had occurred and would be released as a ~1.8 kb fragment in *Bam*HI-digested DNA (Appendix B or C). However, a probe-positive band is only present in the lane with the positive control and not in chromosomal DNA (Figure 2.5C). Random transposition was shown by probing the same blot for *lacZ*. *Bam*HI will cut once in the expression construct (at the junction between '*lacZ* and *cbaA*' or *cbaR*'; Appendix B or C) and at some undetermined spot in the chromosome. If transposition had occurred into a "hot spot", one would expect *Bam*HI-digested DNA to result in a probe-positive band of similar or identical size in all or most transconjugants, which was not the case (Figure 2.5C). Furthermore, if multiple insertions had occurred either at distinct loci or in one locus, more than one probe-positive band would be expected, but only one was detected for each transconjugant, indicative of single insertions.

Figure 2.5 Confirmation that single copies of expression constructs transposed into *C. testosteroni* BR6020. A) Rationale for use of *Bam*HI when digesting chromosomal DNA. Any sequence cloned between the I and O end of pUTCm, the suicide delivery vector, transposes randomly into the bacterial chromosome, represented by the thick dashed line (de Lorenzo and Timmis, 1994). A *Bam*HI site is present at the extreme “left end” of *lacZ* and in the chromosome at some unknown position to the “right” of the insertion point. If transposition is random, digesting chromosomal DNA from transconjugants with *Bam*HI should result in *lacZ*-positive fragments of variable length. See text for additional detail. B) Ethidium bromide-stained 0.9% agarose gel of *Bam*HI-digested chromosomal DNA from *C. testosteroni* BR6020 parent strain and the BR6020 transconjugants containing the control expression construct (*lacZ*) or one of the *cbaA-lacZ* or *cbaR-lacZ* expression constructs (indicated on gel). The rightmost lane contains *Bam*HI-digested pRR1, a vector derived from pUTCm which harbours the same origin of conjugal transfer (*oriT*), origin of vegetative replication (*oriV*), and ampicillin resistance gene (*bla*). This DNA was used as a positive control when a Southern blot of this gel was probed for *oriT*. Sizes of DNA standards (λ digested with *Hind*III) are indicated, in kilobasepairs. C) A Southern blot of this gel was probed for *oriT*, which was not detected in chromosomal DNA but detected in positive control lane (not shown). The membrane was stripped, probed for *lacZ*, and results are demonstrated. Faint probe-positive bands are indicated with an arrow.

B

C. testospermum BR6020
with indicated expression construct

Expression of *cbaA*

Control of 3Cba catabolism was further characterized by studying factors affecting expression of *cbaA*. Expression was measured with four *cbaA-lacZ* constructs (Figure 2.2). One retained *cbaR*, the only gene on Tn5271 which potentially codes for a regulatory element (see more detail in Materials and Methods and in Discussion), while the other three lacked it and retained ever decreasing amounts of DNA upstream of the *cbaA* start codon. It was initially thought that *cbaA* was subject simply to repression by CbaR, and the first two constructs would allow the effect of *cbaR* to be studied. In addition, the latter three constructs would permit mapping of the upstream DNA necessary for proper *cbaA* expression. The host strain was BR6020, a pBRC60-cured derivative of BR60 (pBRC60). This strain is Tn5271-deficient (i.e. unable to grow on 3Cba; Table 2.1) and allows one to measure expression levels in a genetic background lacking any pBRC60-encoded and, more importantly, Tn5271-encoded elements, in particular CbaR. Furthermore, it permitted the possible involvement of chromosomally-encoded proteins to be investigated. Because most catabolic genes for aromatic compounds are regulated by products from linked genes (see Chapter 1), this genetic background allowed the potential involvement of genes unlinked to *cbaA* to be uncovered. A Tn5271-proficient background was created by re-introducing pBRC60 to the transconjugants (Table 2.1), and this allowed expression levels to be measured under normal conditions. Expression levels were investigated with cells grown on 3Cba and benzoate (two substrates that according to O₂ uptake studies induce 3Cba degradative activity) and compared to levels observed with cells grown on succinate (a TCA cycle intermediate that does not induce activity; Table 2.2).

In cultures harbouring the control construct (promoterless '*lacZ*') grown on succinate or benzoate (for Tn5271-deficient cells) or grown on succinate or 3Cba (for Tn5271-proficient cells), LacZ levels were negligible (~1 Miller unit). LacZ levels in cells harbouring the various *cbaA-lacZ* constructs are summarized in Figure 2.6. Induction ratios for benzoate-grown cells relative to succinate-grown cells were as follows: In

Tn5271-deficient cells with *cbaRA'*-*lacZ*, 3.2-fold; *cbaA'*-*lacZ1*, 7.4-fold; *cbaA'*-*lacZ2*, 4.4-fold, and *cbaA'*-*lacZ3*, 2.4-fold. In Tn5271-proficient cells, induction ratios for benzoate relative to succinate were approximately the same for cells with the first three constructs (4.2- to 4.5-fold), but with construct *cbaA'*-*lacZ3* a 3.3-fold induction was observed. Induction ratios on 3Cba relative to succinate were approximately the same for all constructs (5.4- to 5.9-fold). The differences in LacZ levels observed between uninduced and induced cultures were all statistically significant.

Absolute expression levels for cells with *cbaA'*-*lacZ3* were approximately 10- to 40-fold lower than levels observed for cells with *cbaRA'*-*lacZ*, *cbaA'*-*lacZ1*, or *cbaA'*-*lacZ2* (Figure 2.6). As well, when pBRC60 was introduced to cells so as to create a Tn5271-proficient background, *cbaA* expression levels relative to Tn5271-deficient cells tended to decrease. For succinate-grown cells containing *cbaRA'*-*lacZ*, *cbaA'*-*lacZ1*, or *cbaA'*-*lacZ2*, the effect was slight to moderate (1.1- to 2.1-fold reduction), but with construct *cbaA'*-*lacZ3*, a 4.6-fold reduction was observed (Figure 2.6). Similarly, for benzoate-grown cells, the effect was slight to moderate (1.2- to 2.0-fold reduction), but with *cbaA'*-*lacZ3* a 3.4-fold decrease was observed (Figure 2.6). All reductions were statistically significant, except for the 1.1-fold decrease observed in succinate-grown cells with *cbaA'*-*lacZ1*.

Figure 2.6 *cbaA* expression levels in *C. testosteroni* BR6020 containing a single chromosomal insert of the indicated *lacZ* construct. Expression was measured with growth substrates that, according to O₂ uptake experiments, did not induce 3-chlorobenzoate degradative activity (succinate) or did (benzoate [Ba] or 3-chlorobenzoate [3Cba]). Cultures in two genetic backgrounds were tested, Tn5271-deficient (- pBRC60) and Tn5271-proficient (+ pBRC60). *C. testosteroni* BR6020 is a pBRC60-cured strain derived from *C. testosteroni* BR60 (pBRC60) and allows *cbaA* expression levels to be determined in a Tn5271-deficient background. pBRC60 carries Tn5271, which contains *cbaABC*, the catabolic genes for initial metabolism of 3Cba, and *cbaR*, which codes for a regulatory protein that controls *cbaABC* expression. Recall that Tn5271-proficient cultures were pre-grown on 3Cba before subculturing into succinate or benzoate to ensure maximum retention of Tn5271. Each bar represents the mean of an independent trial ($\pm$ standard deviation). Note the difference in scale for LacZ activity in the *cbaA*'-'*lacZ3* graph.

LacZ Activity (Miller Units)

A final set of *cbaA* expression studies was conducted to further characterize induction of *cbaA* expression by 3Cba, in particular to determine whether a chromosomally-encoded protein was potentially involved. Because Tn5271-deficient cells are unable to grow on 3Cba, cells were grown on succinate but exposed to the inducing compound 3Cba during growth. Only cells with *cbaRA*'-'*lacZ* or *cbaA*'-'*lacZ*1 were tested. As a control, the effect of growing cells on succinate in the presence of benzoate was also measured to determine how LacZ levels under these conditions compared to those observed when benzoate was the sole carbon source. In Tn5271-deficient cultures harbouring either construct, growth in the presence of 3Cba led to a moderate but significant increase (1.6- to 1.8-fold) in expression relative to cells grown on succinate alone (Figure 2.7). Growth in the presence of benzoate elicited a small but statistically significant effect on expression. For cultures with *cbaA*'-'*lacZ*1, expression increased 1.3-fold, while with *cbaRA*'-'*lacZ* growth in the presence of benzoate lowered expression 1.2-fold (Figure 2.7). To ensure that the low induction levels observed under these conditions was not because a Tn5271-deficient background was used (i.e. a Tn5271-encoded function was lacking), growth of Tn5271-proficient cultures containing either construct was measured under the same conditions, and comparable LacZ activity was observed (Figure 2.7).

Figure 2.7. *cbaA* expression levels in *C. testosteroni* BR6020 with *cbaRA'*-*lacZ* or *cbaA'*-*lacZ* following growth in medium containing succinate (Succi.), succinate and benzoate (Succi.+Ba), or succinate and 3-chlorobenzoate (Succi.+3Cba). Expression was measured in Tn5271-deficient cells (- pBRC60) and Tn5271-proficient cells (+ pBRC60). *C. testosteroni* BR6020 is a pBRC60-cured strain derived from *C. testosteroni* BR60 (pBRC60) and allows *cbaA* expression levels to be determined in a Tn5271-deficient background. pBRC60 carries Tn5271, which contains *cbaABC*, the catabolic genes for initial metabolism of 3Cba, and *cbaR*, which codes for a regulatory protein that controls *cbaABC* expression. Recall that Tn5271-proficient cultures were pre-grown on 3Cba before subculturing into the media described here to ensure maximum retention of Tn5271. Each bar represents the mean of a trial ($\pm$ standard deviation).

Expression of *cbaR*

Because all indications are that *cbaR* codes for a repressor involved in controlling *cbaA* expression (see Discussion and Chapter 3), expression of *cbaR* was investigated with various constructs (Figure 2.2) so as to roughly map its promoter and determine if expression was affected by the growth conditions and genetic backgrounds that affected *cbaA* expression. LacZ activity in cells harbouring constructs *cbaR'*-*lacZ*1 and *cbaR'*-*lacZ*2 was very low, relative to *cbaA*, but measurable (10 to 20 Miller Units). These levels were observed in Tn5271-deficient cells grown on succinate or benzoate and Tn5271-proficient cells grown on succinate or 3Cba. In cells harbouring construct *cbaR'*-*lacZ*3 in both genetic backgrounds and grown under the same conditions, LacZ activity was negligible (~1 Miller Unit).

DISCUSSION AND SUGGESTIONS FOR ADDITIONAL RESEARCH

Classification of strain BR60 (pBRC60)

Alcaligenes sp. BR60 (pBRC60) is reclassified to *C. testosteroni* BR60 (pBRC60) based on 16S rRNA sequence information and several metabolic similarities between BR60 (pBRC60) and other strains of *C. testosteroni* (formerly *Pseudomonas testosteroni*; [Tamaoka *et al.*, 1987; Wheelis *et al.*, 1967]). Metabolic similarities include inability to grow on various carbohydrates (e.g. glucose, fuctose, mannitol, gluconate) and induction of the Pca extradiol ring fission pathway (*pmd*) by growth on benzoate, 3-hydroxybenzoate, 4-hydroxybenzoate, or Pca (Figure 2.8; [Nakatsu and Wyndham, 1993; Wyndham and Straus, 1988]). It is beyond the scope of this thesis to go into much detail on the topic, but it is noteworthy that while benzoate induces the Pca extradiol ring fission pathway in *C. testosteroni* BR60 and *C. testosteroni* strain 16 (Nakatsu and Wyndham, 1993; Wheelis *et al.*, 1967), BR60 at least does not require functional *pmd* genes for growth on this substrate (Providenti *et al.*, 1999b). This may be because this substrate is in fact metabolized via gentisate (Figure 2.8), as reported for *C. testosteroni* B-356 (M. Sylvestre, personal communication) and not Pca, as reported originally for strain 16 (Wheelis *et al.*, 1967). More work is therefore necessary to better characterize the benzoate degradative pathway in *C. testosteroni*, especially because elements controlling this pathway appear to affect expression of the *cba* pathway (see below).

The *cbaA* promoter

Primer extension analysis identified the *cbaA* transcriptional start site (Figure 2.4) and suggests that *cbaABC* possesses a typical σ^{70} promoter. A hexamer was located upstream of the transcriptional start site that is virtually identical to the consensus -10 sequence (Figure 2.3), but no hexamer similar to the consensus -35 hexamer is apparent. This is a common feature of positively regulated genes (Collado-Vides *et al.*, 1991) and may provide some additional insight into the nature of elements that control expression of *cbaA* (see below).

Figure 2.8 Degradation pathways in *C. testosteroni* BR60 funnel various substituted benzoates towards protocatechuate, which is then metabolized by an extradiol ring fission pathway, designated *pmd* (protocatechuate *meta* *d*ioxygenase). Benzoate induces protocatechuate degradative activity and was originally reported to be channeled towards this intermediate via 3-hydroxybenzoate. However, mutants in which the *pmd* pathway is inactivated are still able to grow on benzoate, suggesting that this substrate may be metabolized by an alternate intermediate, perhaps gentisate. Based on various references listed in text.

The various translational fusions of LacZ to CbaA used to study expression also helped characterize the promoter of *cbaA*. LacZ levels for cells with constructs *cbaRA'*-*lacZ*, *cbaA'*-*lacZ1*, or *cbaA'*-*lacZ2*, which contain decreasing amounts of DNA upstream of *cbaA*, were within the same range, but decreased drastically in cells with *cbaA'*-*lacZ3* (Figure 2.5). This suggests that the 30 bp missing in this construct relative to *cbaA'*-*lacZ2*, the next largest construct (Figure 2.3), are necessary for normal *cbaA* expression. This dramatic reduction in expression is likely because this area spans the putative binding site for RNA polymerase holoenzyme, which normally occupies DNA from -45 to +20 (Collado-Vides *et al.*, 1991). As well, additional regulatory elements (see below) that induce expression may act in this region. This is suggested by the lower induction levels observed in benzoate-grown cells with *cbaA'*-*lacZ3*, relative to succinate-grown cells (Figure 2.6). Interestingly, 3Cba-grown cells had identical induction levels relative to the other constructs, implying that even though overall expression was reduced, ability to induce in response to 3Cba may have remained unaffected.

Substrates that induce 3Cba degradative activity and *cbaA* expression

Which growth substrates induce 3Cba catabolism was determined by studying 3Cba degradative activity at a whole-cell level and measuring LacZ levels with various *cbaA-lacZ* constructs. In a Tn5271-proficient background, O₂ uptake measurements and expression studies indicated that 3Cba catabolic activity (Table 2.2) and *cbaA* transcription (Figure 2.6) were induced by growth on 3Cba and, to a lesser degree, benzoate. A colourimetric assay indicates that benzoate cannot be metabolized by CbaABC, implying therefore that it is gratuitously inducing the *cba* pathway. Other inducers of the Pca lower pathway (designated *pmd*) did not induce 3Cba degradative activity (Table 2.2), suggesting therefore that when the *cba* pathway is induced, it is responding to the parent substrates (i.e. 3Cba or benzoate) and not downstream metabolites that are common to these degradative pathways (e.g. Pca or the ring fission products; Figure 2.8).

An exhaustive search for potential inducers of the *cba* pathway was not carried out. There is the possibility that 5-Cl-Pca, the predominant product of 3Cba metabolism

by the *cba* pathway (Nakatsu, 1992), is an inducer but this compound is not commercially available. As well, the effect on 3Cba degradative activity of a variety of other aromatic substrates that can serve as growth substrates should be studied. These include gentisate, 2-carboxybenzoate, and several methoxylated benzoates (3,4-dimethoxybenzoate; 3-methoxy, 4-hydroxybenzoate; and 3-hydroxy, 4-methoxybenzoate; Figure 2.8). Studies such as this may give some additional insight into potential interactions between the plasmid-encoded *cba* pathway and chromosomally-encoded catabolic pathways for other aromatic compounds.

Growth on 3Cba or benzoate induces degradative activity for other aromatic compounds

It is noteworthy that the induction spectra of 3Cba and benzoate were quite similar with respect to degradative activity towards the aromatic substrates 3- and 4-hydroxybenzoate (Table 2.2). Activity was low but levels were comparable. As mentioned previously (Figure 2.8), benzoate is thought to be degraded via 3-hydroxybenzoate to either Pca (Wheelis *et al.*, 1967) or gentisate (M. Sylvestre, pers. commun.), so induction of degradative activity towards 3-hydroxybenzoate by benzoate-grown cells would be expected. However, induction by 3Cba-grown cells is not expected as 3Cba is not metabolized to 3-hydroxybenzoate. Nor is induction of 4-hydroxybenzoate degradative activity expected for cells grown on benzoate or 3Cba, as 4-hydroxybenzoate is not a metabolite. It appears therefore that gratuitous induction of several aromatic catabolic pathways occurs when cells are grown on 3Cba or benzoate, and even though completely different metabolic pathways are used for these substrates, very similar induction patterns are observed following growth.

Regulators controlling *cbaA* expression

General statements

Most catabolic operons are controlled by closely linked genes coding for regulatory proteins (Chapter 1). The possible involvement in *cba* regulation of a protein encoded by a gene located on pBRC60 but outside of Tn5271 was discounted because the *cba* genes in other environmental isolates have been found integrated into completely

unrelated plasmids or even in the chromosome (Di Gioia *et al.*, 1998; Peel and Wyndham, 1999; Wyndham *et al.*, 1988). This suggests that there has been no selection for functions encoded outside Tn5271 on pBRC60. Therefore, the possibility that a regulator is present within Tn5271 was considered. Regions of Tn5271 other than those coding for *cbaABC* were searched for ORFs that show similarity to conceptual translation products in the public nucleotide sequence databases. The DNA downstream of *cbaABC* contains five ORFs (*cbaDEFGH*) that may code for proteins involved in uptake (Providenti *et al.*, 1999a) and a truncated ORF (ORF8) for part of an aryl CoA ligase (Di Gioia *et al.*, 1998). These ORFs therefore do not appear to code for regulatory elements. Also, because previous research had shown that the minimal region of Tn5271 necessary for growth on 3Cba harbours only *cbaABC* (Nakatsu and Wyndham, 1993), the downstream ORFs from *cbaD* to ORF8 were not taken into consideration when designing the *lacZ* expression constructs. In contrast, upstream DNA was found to contain a divergently transcribed ORF that codes for CbaR, a protein that is part of a family of transcriptional regulators known to respond to aromatic compounds or to regulate catabolic pathways for aromatic compounds (Chapter 3). The expression studies therefore focussed on the effect of this protein on transcription.

CbaR: role in controlling *cbaA*, characterization of *cbaR* expression, and comparison to other MarR-like proteins

Based on data presented in this chapter and in Chapter 3, CbaR appears to be a repressor of the *cba* pathway whose major role is to modulate gratuitous induction of *cbaA* by benzoate. If control of *cbaA* expression was simply a case of negative regulation by CbaR, one would expect maximal expression levels under non-inducing conditions in Tn5271-deficient lacking CbaR. Instead, expression levels for succinate-grown cells with *cbaA*'-'*lacZ*1 and *cbaA*'-'*lacZ*2 were roughly the same as those of Tn5271-deficient cells with *cbaRA*'-'*lacZ*, which provides the gene for CbaR in *cis*. When grown on benzoate, though, LacZ levels in cells with *cbaA*'-'*lacZ*1 and *cbaA*'-'*lacZ*2 were approximately 2-fold higher than in cells with *cbaRA*'-'*lacZ*, implying that benzoate-induced expression was

substantially reduced by the product of *cbaR* (Figure 2.6). Other data that support the proposed role of CbaR as a modulator of benzoate-induced expression include gel shift data which show CbaR binding to DNA containing the *cbaA* promoter despite the presence benzoate but not in the presence of 3Cba (Chapter 3). As well, when Tn5271-deficient cells (i.e. no *trans*-encoded CbaR) with *cbaA'*-*lacZ*1 were grown on succinate in the presence of benzoate, LacZ levels in cells with *cbaA'*-*lacZ*1 increased 1.3-fold relative to cultures grown on succinate alone, whereas in cultures with *cbaRA'*-*lacZ*, a 1.2-fold reduction occurred (Figure 2.7). These induction levels, though not high, were statistically significant. Furthermore, for each construct, the experiment was conducted with matched cultures (i.e. one culture was subcultured into the three different media at the same time and cultures were grown and expression measured under identical conditions with the same materials), such that the the differences observed can not be ascribed to experimental error. The substantially lower expression levels observed under these culture conditions, relative to levels observed with benzoate- or 3Cba-grown cells, is most likely due to a catabolite repression-like effect brought about by the succinate, but this question was not studied further.

In addition to its proposed role as a modulator, CbaR may have a general repressing effect on *cbaA* expression. When Tn5271 (on pBRC60) is re-introduced to the strains with the expression constructs, thus providing *trans*-encoded *cbaR*, expression relative to Tn5271-deficient cells for succinate- or benzoate-grown cells with *cbaA'*-*lacZ*1 or *cbaA'*-*lacZ*2 was reduced by as much as 2-fold (see Results). The effect on cells with *cbaRA'*-*lacZ* was not as great, especially for benzoate-grown cells, most likely because CbaR is already provided by the construct. A significantly greater reduction was observed when cells with *cbaA'*-*lacZ*3 were made Tn5271 proficient (3.4- to 4.6-fold). This is likely because with this construct the overall ability to express *cbaA* was substantially lowered (Figure 2.6) while the ability of CbaR to repress was not affected as the binding sites for CbaR remained intact (Chapter 3). There is also the possibility that the general decrease in expression observed when pBRC60 is introduced reflects titration of CbaS, the proposed chromosomally-encoded activator of *cbaA* (see below), by binding sequences on the

plasmid, which would be present in greater than one copy, such that there is less activator to induce the chromosomally-located expression constructs. Therefore, to determine if *trans*-encoded *cbaR* lowers *cbaA* expression in general and, more specifically, modulates benzoate-induced expression of *cbaA*, more definitive studies should be conducted. For example, *cbaR* could be exclusively cloned on a plasmid and LacZ studies similar to the ones described above could be conducted to confirm its proposed role.

Expression of *cbaR* appears to be constant but at a very low level. Growth conditions that induce *cbaA* did not increase expression of *cbaR* (see Results). No expression was observed with construct *cbaR'*-*lacZ*3, implying that *cbaR* transcription begins between nt 4098 (*Bgl*II site) and nt. 4446 (*Not*I site; Figure 2.3). CbaR does not bind to DNA spanning this region (Chapter 3) and *cbaR* expression is not affected by *trans*-encoded *cbaR* (i.e. Tn5271-proficient cells; see Results), implying that CbaR does not repress its own expression.

CbaR, as mentioned above, belongs to the MarR family of transcriptional regulators, a group of proteins involved in controlling diverse cellular functions, including antibiotic resistance, exclusion or efflux mechanisms, catabolism of various substrates, and plant and animal virulence and pathogenicity. A common feature, though, of MarR-like proteins is their ability to respond to aromatic compounds (discussed in greater detail in Chapter 3). The organization of genes for MarR-like proteins relative to the genes they regulate and how expression of the gene for the MarR-like protein is controlled can vary substantially, and a summary is provided in Figure 2.9. Interesting contrasts include how the genes for MarR-like proteins can be divergently transcribed from their target genes, as with *hpcR*, *mexR*, and *nhhD* (Kobayashi and Shimizu, 1998; Poole *et al.*, 1996; Roper *et al.*, 1993); part of the same operon, as with *emrR* and *marR* (Alekshun and Levy, 1997; Lomovskaya *et al.*, 1995), or part of a different operon but transcribed in the same direction, as with *badR* (Egland and Harwood, 1999). Expression of some is subject to induction by other regulators, as with *nhhD* and *marR* (Alekshun and Levy, 1997; Kobayashi and Shimizu, 1998), while auto-repression has been shown for *emrR*, *mexR*, and *marR* (Alekshun and Levy, 1997; Poole *et al.*, 1996; Lomovskaya *et al.*, 1995).

For the most part, MarR-like proteins appear to act alone on target genes but, as discussed in Chapter 3, the MarR family of proteins was recognized relatively recently and characterization of many members from this family has not been extensive, so no gross generalizations are yet possible. In some cases, though, they are known to act in concert with at least one other regulator, such as with BadR which works together with AadR to fully activate the *bad* genes of *Rhodopseudomonas palustris* (Egland and Harwood, 1999). The MarR/MarA system of *E. coli* provides another interesting example. MarR represses its own expression and that of the gene for MarA, an AraC-like activator. MarA activates a whole host of other genes involved in response to various stresses (e.g. pH, antibiotic resistance, superoxides) but can also induce expression of *marRAB* (i.e. autoactivate its own expression; Figure 2.9). MarR is thus the first level of control for a very complex regulatory cascade (Alekhun and Levy, 1997). Similarly, PecS is part of an intricate regulatory network which, in concert with a variety of other regulators (both repressors and activators), controls expression of a host of unlinked operons spread out over the *Erwinia chrysanthemi* chromosome and involved in pectinolysis and plant pathogenicity (Hugouvieux-Cotte-Pattat *et al.*, 1996). For some operons, the involvement of PecS is quite subtle and its role has been suggested to be that of an attenuator rather than a full-fledged repressor (Rouanet *et al.*, 1999).

Figure 2.9 Organization of various genes regulated by MarR-like transcriptional regulators, relative to the gene for the MarR-like proteins. Indicated are the organism in which these genes are located, the function of the genes, whether the MarR-like protein negatively regulates (-) or positively regulates (+) the target gene, and whether the MarR-like protein represses its own expression. For some of the genes indicated, additional regulators are involved either in expression of the gene for the MarR-like protein or in controlling expression of the target gene along with the MarR-like protein. Note that *cbaS* and *aadR* are unlinked to the genes they control and no directionality is shown relative to them. Additional information is provided in the text and in Chapter 3 (Table 3.2). Based on various references listed in the text.

C. testosteroni BR60 3-chlorobenzoate catabolic genes (*cba*)

E. coli C homoprotocatechuate catabolic genes (*hpc*)

Rhodopseudomonas palustris CGA009 genes for anaerobic catabolism of benzoate (*bad*)

Rhodococcus rhodochrous J1 nitrile hydratase genes (*nhh*)

E. coli multiple antibiotic resistance genes (*mar*)

E. coli antibiotic efflux pump genes (*emr*)

P. aeruginosa antibiotic efflux pump genes (*mex*)

CbaS: speculation on identity and role in controlling aromatic catabolic pathways

It was initially thought that CbaR might be the sole protein involved in controlling *cbaA*, but indications are that another protein appears to be required for activation of the *cba* pathway. As mentioned above, the *cbaABC*-encoded pathway is induced maximally by 3Cba and gratuitously by benzoate (Table 2.2 and Figure 2.6). Studies with Tn5271-deficient cells showed that growth on benzoate (Figure 2.6) or growth on succinate in the presence of 3Cba (Figure 2.7) induced *cbaA* expression, regardless of whether CbaR was present. Expression of *cbaA* in response to benzoate in Tn5271-deficient cells lacking *cis*-encoded CbaR or other *trans*-encoded proteins from Tn5271 can be at levels that are higher or comparable to levels observed with Tn5271-proficient cells grown on 3Cba or benzoate (Figure 2.6). It appears therefore that a chromosomally-encoded regulatory element, tentatively designated CbaS, responds to either 3Cba or benzoate and induces *cbaABC* expression, and that CbaS and CbaR therefore function in concert to control *cbaABC* (Figure 2.9; see also Chapter 5). A description of attempts to clone the gene for CbaS is given in Chapter 4.

Some features of the *cbaA* promoter (see above) suggest that CbaS is a positive regulator. Furthermore, most catabolic genes for aromatic compounds are regulated exclusively or in part by activators (Chapter 1), a sensible arrangement which ensures their expression only under the necessary conditions (i.e. when the substrates for the pathway are present). The lower induction levels in benzoate-grown cells with *cbaA*'-'*lacZ*3 indicate that the 30 bp upstream of nt 4511 (Figure 2.3) may be necessary for proper activation, typical of positively regulated σ^{70} -dependent promoters (Chapter 1). CbaS may somehow be involved with control of the benzoate catabolic pathway because 3Cba, which is not metabolized by this pathway (Nakatsu and Wyndham, 1993), gratuitously induces benzoate degradative activity (Table 2.2). As well, the induction spectra of 3Cba and benzoate were quite similar with respect to other aromatic substrates (Table 2.2).

Expression of various plasmid-encoded aromatic pathways is known to require or at least be affected by unlinked, chromosomally-encoded regulators. As discussed in Chapter 1, CatR, the activator of chromosomally-encoded *catBCA*, is required for full activation of *pheBA*, a plasmid-encoded operon involved in phenol hydroxylation and intradiol ring fission of the catechol that is generated (Figure 1.3D; [Kasak *et al.*, 1993; Parsek *et al.*, 1995]). In addition, some chromosomally-encoded proteins are able to cross-activate non-cognate promoters controlling plasmid-encoded pathways. Induction can be at a level sufficient to allow growth on the relevant substrate, but growth rates can be impaired. Examples include the XylS-independent induction of the lower pathway of the *xyl* genes in response to benzoate (Figure 1.9A; [Cuskey and Sprenkle, 1988; Jeffrey *et al.*, 1992; Kessler *et al.*, 1994]), CatR-mediated induction of the *clc* pathways, and for that matter, ClcR-mediated induction of the *cat* pathway (McFall *et al.*, 1997a; Parsek *et al.*, 1994).

CHAPTER 3: FUNCTIONAL CHARACTERIZATION OF CbaR, A REPRESSOR OF THE *cba* PATHWAY

INTRODUCTION

Initial analysis of DNA upstream of *cbaABC* identified *cbaR*, a divergently transcribed open reading frame (ORF) putatively coding for a 19.4 kDa protein that shows similarity to various members of the MarR family of transcriptional regulators. MarR is a repressor involved in controlling expression of the multiple antibiotic resistance (*mar*) genes of *E. coli* and responds to aromatic compounds like 2-hydroxybenzoate and 2,4-dinitrophenol (Alekshun and Levy, 1999a; Ariza *et al.*, 1994; Cohen *et al.*, 1993; Seoane and Levy, 1995). Many other proteins showing similarity to MarR also respond to specific aromatic compounds or are involved in regulating genes for metabolism of aromatic compounds (Sulavik *et al.*, 1995)]; see also Chapter 1 [Figure 1.8 and 1.9]), therefore CbaR appeared to be a good candidate for a regulatory protein controlling expression of the *cba* pathway.

Results presented in Chapter 2 indicate that *cbaR* is transcribed at a low basal level and suggest that its presence in *cis* (i.e. from construct *cbaRA'*-'*lacZ*) and possibly in *trans* (i.e. from Tn5271) modulates gratuitous induction of *cbaA* in response to compounds like benzoate, and perhaps lowers expression in general. Effects, though statistically significant, were quite subtle, and a clear involvement was not shown. It was therefore decided to clone this gene and functionally characterize the product.

MATERIALS AND METHODS

Bacterial strains and plasmids

Bacterial strains and plasmids used in this study are listed in Table 3.1. Unless otherwise stated, *E. coli* strains were grown at 37°C in Luria-Bertani medium (LB; 1% w/v tryptone, 0.5% w/v yeast extract, 0.5% w/v NaCl) containing ampicillin (250 mg/L) and/or kanamycin (40 mg/L), as required. Antibiotics were purchased from Sigma-Aldrich Ltd. (Oakville, Ontario, Canada).

Identification and alignment of CbaR homologues

The deduced protein sequence of CbaR (Chapter 2) was used to search for similar proteins in the GenBank non-redundant database using the BLASTP (Altschul *et al.*, 1990; Altschul *et al.*, 1997) network service of the National Center for Biotechnology Information (NCBI) at Bethesda, Maryland, USA (<http://www.ncbi.nlm.nih.gov>). CbaR belongs to the MarR family of transcriptional regulators, a highly divergent group of proteins (see Results), therefore a second round of BLASTP analysis was performed using CbaR homologues as query terms to identify other potential members. A third round was performed but no new sequences were identified. CbaR homologues and several other MarR family proteins identified in the second round of BLASTP analysis (see Results) were aligned using MaxHom (Sander and Schneider, 1991) and the secondary structure of the aligned sequences was predicted by PHDsec (Rost and Sander, 1993; Rost and Sander, 1994; Rost *et al.*, 1994), both on the network service of the European Molecular Biology Laboratory (EMBL) at Heidelberg, Germany (<http://www.embl-heidelberg.de/predictprotein>).

DNA manipulation, PCR amplification conditions, and sequencing

All enzymes used to generate the clones described below were purchased from New England Biolabs (Mississauga, Ontario, Canada). Plasmids were isolated and DNA was manipulated according to standard methods (Sambrook *et al.*, 1989). *E. coli* was transformed by electroporation using a Gene Pulser (Bio-Rad Laboratories, Ltd., Mississauga, Ontario, Canada) according to the manufacturer's recommendations. All PCR amplification of DNA was in a 100 µL reaction volume containing 1 U Vent DNA Polymerase, 200 µM of each dNTP, 50 pmol of each primer, and the appropriate template

DNA (see below). Amplification conditions were 5 minute hot start at 95°C followed by 30 cycles of 30 seconds at 48°C for primer annealing, 30 seconds at 72°C for extension, and 60 seconds at 94°C for template melting. DNA sequencing (see below) was performed by the chain-terminating dideoxy method on an ABI Prism Automated Sequencer (Biotechnology Research Institute, University of Ottawa, Ottawa, Ontario, Canada). All oligonucleotides were synthesized at the Biotechnology Research Institute.

Cloning of *cbaR*, over-expression, and purification of CbaR

Purified CbaR for gel shift studies and DNaseI protection assays (described below) was obtained by cloning and over-expressing *cbaR* in *E. coli*. The gene coding for CbaR was PCR-amplified using forward primer CBAR3958 (5'-cgggatccCTTGCAAGAGATCCTCGA-3') and reverse primer CBAR3440 (5'-ccgagctcCTACTCCTGAGGAGATTC-3'). Nucleotides in lower case represent DNA that is not native to *cbaR* and underlined nucleotides are restriction sites for *Bam*HI and *Sac*I, respectively. CBAR3958 is complementary to the portion of *cbaR* coding for the N-terminus but lacks the native start codon, while CBAR3440 is complementary to the C-terminal coding portion of *cbaR* up to and including the native stop codon. Template DNA was an ~850 bp gel purified *Pst*I fragment from p60N.4 that contains the complete gene for CbaR. The amplification product was cloned into pCR-Script SK(+) digested with *Sma*I (resulting in pMP132.4) and the nucleotide sequence of the insert was confirmed. *cbaR* was placed downstream and in frame of DNA coding for the N-terminal tag MRGSH₆GS by digesting pMP132.4 with *Bam*HI and *Sac*I, agarose gel purifying the ~550 bp insert, and cloning this fragment into pQE30 digested with *Bam*HI and *Sac*I, resulting in pQE30cbaR. Initially, the ligation was transformed into *E. coli* JM103 or JM109, which contain a chromosomal LacI^q gene, but none of the strains that harboured the predicted plasmid expressed any protein, as determined by sodium dodecylsulfate polyacrylamide gel electrophoresis, conducted under standard conditions (Sambrook *et al.*, 1989). The ligation was thus transformed directly into *E. coli* M15 (pREP4) and a strain was obtained that harboured the expected plasmid and was capable of expressing a protein which matched the predicted size of CbaR.

CbaR was over-expressed and affinity purified as follows: *E. coli* M15 (pQE30cbaR, pREP4) was grown at 30°C with rotary shaking (200 rpm) to mid-log phase ($OD_{600} \sim 0.6$) in 100 mL 2YT medium (1.6% w/v tryptone, 1% w/v yeast extract, 0.5% w/v NaCl) amended with ampicillin and kanamycin. Expression was induced with IPTG (final concentration 1mM) and cells were allowed to incubate for an additional 6 hours. Cells were harvested by centrifugation, resuspended in 20 mL lysis solution (300 mM NaCl, 10 mM imidazole, 50 mM NaH_2PO_4 , pH 8.0), frozen at -20°C then thawed, and then ruptured by four passages at 20000 PSI through a C5 cell homogenizer (Avestin, Ottawa, Ontario, Canada). All subsequent steps were at 4°C. Ruptured cells were centrifuged for 15 minutes at 12100×g. The supernatant was then combined with 2 mL Ni-NTA agarose (Qiagen Inc., Mississauga, Ontario, Canada) and held for 1 hour with occasional mixing. The agarose was then washed three times with 10 mL wash buffer (300 mM NaCl, 20 mM imidazole, 50 mM NaH_2PO_4 , pH 8.0) and once with 10 mL wash buffer containing 50% (v/v) glycerol. Protein was eluted in a total of 3 mL elution buffer (300 mM NaCl, 250 mM imidazole, 50 mM NaH_2PO_4 , 50% v/v glycerol, pH 8.0), dialyzed two times for ~18 hrs against 200 mL glycerol storage buffer (300 mM NaCl, 50 mM NaH_2PO_4 , 50% v/v glycerol, pH 8.0) using a Pierce Slide-A-Lyzer cassette, 10 kDa molecular weight cutoff (MJS BioLynx, Brockville, Ontario, Canada) and stored at -20°C. Protein concentration was determined using the Bradford reagent (Bio-Rad Laboratories Ltd) with bovine serum albumin as a standard. Some contaminating proteins were present in affinity-purified CbaR (see Results), and to eliminate the possibility that these were affecting migration of the DNA templates (see Results), *E. coli* M15 (pQE30, pREP4) (i.e. containing the expression construct with no insert) was IPTG-induced and treated as described before. This extract was then tested for its ability to affect migration of a gel shift template (see Results). One protein preparation was used in all the assays described below.

Gel shift assays

Binding of CbaR to DNA templates containing P_{cbaR} or P_{cbaA}

Initial studies investigating potential interactions between CbaR and DNA fragments containing *cbaR* or *cbaA* promoter sequences were investigated by gel shift assays. Three different templates were used. The first template contained the putative transcriptional start site of *cbaR* (Chapter 2) and spanned nt 4098 (*Bgl*III site) to nt 4446 (*Not*I site; Figure 3.1). The second template contained the transcriptional start site of *cbaA* and spanned nt 4446 (*Not*I site) to nt 4705 (*Sca*I site; Figure 3.1). The third template also contained the transcriptional start site of *cbaA* but lacked upstream DNA necessary for normal expression (see Chapter 2). It spanned nt 4511 to nt 4705 (*Sca*I site; Figure 3.1). The first and second template were obtained by digesting pMP14.10 with *Not*I and *Bam*HI and gel purifying the ~400 bp and ~260 bp fragments, respectively. The third template was obtained by digesting pMP117.1 with *Pst*I and *Bam*HI and gel purifying the ~190 bp fragment. The three templates were then end-labelled with digoxigenin as described below, mixed with 0.4 μ g CbaR, and handled as described below with respect to binding, electrophoresis, transfer, and chemiluminescent detection.

Interactions of CbaR with P_{cbaA} DNA

More in depth studies investigating interactions between CbaR and DNA containing P_{cbaA} were conducted with a 186 bp template (MW 114.7 kDa) generated by PCR. This template is based on the third template described above, which defined roughly a minimum area necessary for binding. It is identical with respect to the “left” end but is slightly shorter at the “right” end (Figure 3.1). The primers used were PCBA4480B (5'-CTCGGGTAAACACCTAGA-3') and CBAA4665 (5'-TCGCACCAAATCTTCGT-3'; Figure 3.1). Template DNA was the ~3.5 kb *Not*I fragment from pMP14.10. The amplification product was agarose gel purified, quantified by UV spectroscopy, end-labelled by Terminal transferase with digoxigenin-11-ddUTP (Roche Molecular Biochemicals, Laval, Québec, Canada), and purified by spin-column chromatography (Qiagen Inc.). One preparation was used for all gel shift assays.

Figure 3.1 Schematic representation of DNA templates used in gel shift assays.

The region of catabolic transposon Tn5271 between *cbaR* and *cbaA* start codons (ATG) is presented, along with some relevant restriction sites. Template 1 contains the putative transcriptional start site (+1) of *cbaR*, while Templates 2 and 3 contain the transcriptional start site of *cbaA*. Template 3 is ~60 nt shorter than Template 2 at one end. These three templates were obtained by restriction digestion of appropriate constructs and were used in initial assays to roughly map the regions of *cbaR-cbaA* intergenic DNA bound by CbaR. All subsequent assays used a PCR-generated template, whose boundaries relative to the other templates are indicated. See Materials and Methods for additional detail.

All DNA binding assays were performed in a total volume of 10 μ L and each reaction contained 2 μ L 5 $\times$ binding buffer (100 mM Hepes, pH 7.6, 5 mM EDTA, 50 mM $(\text{NH}_4)_2\text{SO}_4$, 5 mM dithiothreitol, 1% w/v Tween-20, 150 mM KCl), 0.5 μ L poly[d(I-C)] (1 μ g/ μ L), 0.5 μ L poly[L-lysine] (0.1 μ g/ μ L) (Roche Molecular Biochemicals), 4 to 40 ng template DNA, various amounts of amendments depending on the experiment (see below), 1 to 3 μ L of CbaR in glycerol storage buffer, glycerol storage buffer containing no CbaR if required so that the total amount present was 3 μ L, and sterile distilled H_2O as required to bring it up to volume. The final concentration of glycerol (15% v/v) in the binding reactions made loading buffer for subsequent electrophoresis unnecessary.

Samples were held at room temperature for 15-30 minutes, placed on ice for at least 5 minutes, then an aliquot (4 to 5 μ L) was applied to a non-denaturing polyacrylamide gel (10 $\times$ 10 $\times$ 0.08 cm) at 4°C that had been pre-electrophoresed to a constant current (1-2 hours) in a running buffer containing Tris-base/boric acid/EDTA (TBE, final concentration 22.25 mM, 22.25 mM, and 0.5 mM, respectively, pH 8.0). The first third of the gel was 3.5% w/v polyacrylamide and the last two-thirds was 12% w/v polyacrylamide and both parts of the gel contained TBE at the same concentration as the running buffer. Samples were electrophoresed at 10 V/cm for 3 to 4 hours and then transferred to a positively charged nylon membrane (Roche Molecular Biochemicals) by electroblotting with TBE for 60 to 90 minutes at 300 mA in a Panther Semi-Dry Electroblotter (Owl Separation Systems, Portsmouth, New Hampshire, USA). DNA was detected by chemiluminescence using an anti-digoxigenin system, according to manufacturer's recommendations (Roche Molecular Biochemicals).

The various tests that characterized CbaR binding to P_{cbaA} DNA, most of which were conducted at least two times to confirm results, were as follows:

(1) Initial gel shift assays showed that when CbaR is combined with P_{cbaA} DNA, two protein-DNA complexes appear (C1 and C2; see Results), suggesting the presence of at least two binding sites. To determine whether CbaR bound to sites with different affinity, the effect of various amounts of CbaR (final concentration 200 ng/ μ L to 0.08 ng/ μ L) was tested. The apparent dissociation constant (K_{dapp} ; [Praillet *et al.*, 1996]) for C1 (the CbaR

concentration which resulted in equal amount of C1 and free DNA) and C2 (the CbaR concentration which resulted in equal amounts of C1 and C2) were judged by visual inspection.

(2) Initial experiments had shown that 4 mM 3Cba is disrupts CbaR binding (see Results), and to better characterize this phenomenon, the minimum concentration at which an effect could be observed was determined. CbaR was present at 40 ng/ μ L (a concentration which results in equal amounts of both retarded complexes) while the amount of 3Cba was varied (final concentration 4 μ M to 4 mM). 3Cba was added as an aqueous solution.

(3) The effect on CbaR binding by a variety of other aromatic compounds, in particular substituted benzoates (summarized in Table 3.3), was also tested, for reasons given in the Results and Discussion section. The final concentration of all compounds tested was 4 mM (which reflects concentrations typically used when cells are grown on these substrates) and CbaR was present at a final concentration of 40 ng/ μ L (for reasons given above). All aromatic compounds were purchased from Sigma-Aldrich Ltd., except for 3-chlorobiphenyl, which was purchased from ICN Biochemicals Inc. (K&K Laboratories, Cleveland, Ohio, USA). Chemicals were of 98% purity or better. All were added as aqueous solutions, except for 3-chlorobiphenyl, 3-chlorotoluene, and 1-chloro-3-nitrobenzene, which were added as dimethylformamide solutions.

(4) Some compounds improved CbaR binding (see Results) and to better characterize this phenomenon, the minimum concentration at which an effect could be observed was determined with one model compound, 3-carboxybenzoate. CbaR was present at a constant concentration (40 ng/ μ L) while the amount of 3-carboxybenzoate was varied (final concentration 4 mM to 10 μ M). As well, to characterize the stoichiometry of this phenomenon, the effect of 3-carboxybenzoate at a constant concentration (4 mM) was tested with various amounts of CbaR (final concentration 30 to 0.075 ng/ μ L).

DNaseI protection assays

Binding of CbaR to P_{cbaA} DNA was further characterized by DNaseI protection assays, which identify specific areas bound by DNA-binding proteins. The two templates used, ³²P-PCBA and ³²P-CBAA, were identical to the 186 bp gel shift template and

generated by PCR as described above except that during amplification, one of the primers was end-labelled with ^{32}P so that only the non-coding strand (^{32}P -PCBA) or the coding strand (^{32}P -CBAA) was labelled. Primers (50 pmol) had been phosphorylated previously with $[\gamma\text{-}^{32}\text{P}]\text{-ATP}$ (Amersham Canada Ltd, Oakville, Ontario, Canada) and T4 polynucleotide kinase (New England Biolabs). PCR products were purified by spin-column chromatography (Qiagen Inc.) and ~6 ng were bound to various amounts of CbaR (with and without 4 mM 3Cba or 3-carboxybenzoate [see Results]) as described for gel shift assays, the only difference being that 5 mM MgCl_2 and 2.5 mM CaCl_2 were present and the reaction volume was scaled up to 12 μL . After binding, 2.5×10^{-3} U DNaseI (Sigma-Aldrich Ltd) was added and template digested at 22°C as follows: For ^{32}P -PCBA, 4 μL of sample was removed from the reaction mix at 30, 60 and 120 seconds and combined into one tube containing 1.5 μL stop solution (500 mM EDTA), while for ^{32}P -CBAA, 1.5 μL stop solution was added after 60 seconds digestion. Three μL loading buffer were then added and samples were placed in a boiling water bath for 3 minutes. Fragments were resolved electrophoretically by applying a 3 μL aliquot to denaturing sequencing gel (8% w/v polyacrylamide/8 M urea). A sequencing ladder conducted with the appropriate primer (using pMP14.10 as a template) was included in adjacent lanes.

Table 3.1 *E. coli* strains and plasmids used for experiments described in Chapter 3.

Strain and plasmid	Relevant characteristics of plasmid(s) ¹	Reference
XL-1Blue ² (pCR-Script SK(+))	Amp ^r , α - <i>lacZ</i> /MCS	Stratagene, LaJolla, California, USA
M15 (pQE30, pREP4)	pQE30 is Amp ^r , IPTG-inducible expression vector that adds six histidine tag to N-terminus of proteins of cloned genes pREP4 is Km ^r and codes for LacI ^q , which represses expression from pQE30 promoter	Qiagen Inc., Mississauga, Ontario, Canada
XL-1Blue (p60N.4)	Amp ^r , contains Tn5271 cloned as an <i>NheI</i> fragment (i.e. without the flanking IS1071)	Chapter 2
XL-1Blue (pMP14.10)	Amp ^r , contains <i>cbaA</i> promoter	Chapter 2
XL-1Blue (pMP117.1)	Amp ^r , contains <i>cbaA</i> promoter	Chapter 2
XL-1Blue (pMP130.4)	Amp ^r , PCR-amplified <i>cbaR</i> cloned into pCR-Script SK(+)	This study
M15 (pQE30cbaR, pREP4)	Amp ^r , Km ^r . pQE30cbaR contains <i>cbaR</i> sub-cloned into pQE30 while pREP4 codes for LacI ^q , which represses expression from pQE30 promoter	This study

¹Abbreviations: MCS, multiple cloning site; IPTG, isopropyl- β -D-thiogalactopyranoside; Amp, ampicillin; Km, kanamycin; r, resistant

²Strain XL-1 Blue is Km^r

RESULTS

CbaR is a MarR-like protein

BLASTP analysis identified CbaR as a member of the MarR family of transcriptional regulators. This group is named after MarR, a repressor involved in controlling expression of the multiple antibiotic resistance (*mar*) genes of *E. coli* (Ariza *et al.*, 1994; Cohen *et al.*, 1993; Seoane and Levy, 1995). HpcR and MarR were two of the first members described (Table 3.2) and since then more than forty new members have been shown to exist in various genera of Gram positive and Gram negative bacteria. Most, though, are ORFs of unknown function and represent data generated by several genome or plasmid sequencing projects (see Table 3.2 for examples). Because a presentation of all CbaR- or MarR-like sequences would not be informative, only a select group of nineteen sequences is shown in alignment with CbaR (Figure 3.2). One criterion for presentation was that the protein have a known or putative function, although several ORFs of unknown function were included to demonstrate the ubiquity of this family. *B. subtilis*, *St. sciuri*, and *E. coli* harbour more than one hypothetical MarR-like protein, and several homologues between *S. typhimurium* and *E. coli* were extremely similar (>85% identity). Only the ones showing greatest similarity to CbaR were selected for presentation.

MarR-family proteins harbour the consensus sequence DxRxxxxx(L/I)TxxG, where x represents any amino acid (Sulavik *et al.*, 1995). A more extensive consensus sequence based on an alignment of 38 MarR-like proteins (not shown) is presented in Figure 3.2. A predicted secondary structure based on this alignment is also presented. At least two helix-turn-helix domains are proposed; an N-terminal domain (amino acids 18 to 56) and a C-terminal domain (amino acid 111 to 151). An alternative to the former is an internal helix-turn-helix domain (amino acids 46 to 85) which includes part of the conserved central region (Figure 3.2). An extended β -sheet is predicted for sections of the highly conserved region.

Table 3.2. Proteins (hypothetical and characterized) showing homology to CbaR or MarR and included in the alignment

Sequence	Accession number	% identity/ similarity to CbaR ¹	Organism	Function	Reference
BadR_RhPa	U75363	n.a. ²	<i>Rhodopseudomonas palustris</i> CGA009	Activates genes required for anaerobic growth on benzoate. May respond to benzoyl-CoA	(Egland and Harwood, 1999; Egland <i>et al.</i> , 1997)
CinR_BuFi	U64802	n.a.	<i>Butyrivibrio fibrisolvens</i> E14	Represses genes for cinnamoyl ester hydrolase. Responds to feruloyl-containing sugars	(Dalrymple and Swadling, 1997)
EmrR_EsCl	U19993	20/53	<i>E. coli</i>	Represses multidrug resistance pump genes. Responds to 2,4-dinitrophenol	(Brooun <i>et al.</i> , 1999; Lomovskaya <i>et al.</i> , 1995)
HpcR_EsCl	S56952	19/45	<i>E. coli</i> C	Represses homoprotocatechuate catabolic genes	(Jenkins and Cooper, 1988; Roper <i>et al.</i> , 1993)
MarR_EsCl	M96235	28/50	<i>E. coli</i>	Represses multiple antibiotic resistance genes. Responds to 2-hydroxybenzoate and 2,4-dinitrophenol	(Aleksun and Levy, 1999a; Ariza <i>et al.</i> , 1994; Seoane and Levy, 1995)
MexR_PsAe	U23763	28/48	<i>Pseudomonas aeruginosa</i>	Activates multiple antibiotic resistance genes. May respond to chloramphenicol	(Poole <i>et al.</i> , 1996)
NapB_EhHi	AJ000346	30/52	<i>Enterococcus hirae</i>	hypothetical protein	Unpublished
NhhD_RhRh	D67027	n.a.	<i>Rhodococcus rhodochrous</i> J1	Activates nitrile hydratase genes	(Komeda <i>et al.</i> , 1996)
ORF1_AcCa	Y09102	n.a.	<i>Acinetobacter</i> sp. ADP1	hypothetical protein. ORF is adjacent to genes involved in dodecane catabolism	(Parche <i>et al.</i> , 1997)

¹ because these values are based on BLASTP search results, the numbers may reflect what is observed for only a portion of the protein or ORF

² n.a., not applicable. BLASTP search with CbaR as a query did not identify this protein as a CbaR homologue. However, the protein is included because it is a MarR-family member based on sequence similarity using CbaR homologues as a query in a second round of BLASTP analysis

Table 3.2 (continued)

Sequence	Accession number	% identity/ similarity to CbaR ¹	Organism	Function	Reference
ORF158_SpAr	AF079317	24/41	<i>Sphingomonas aromaticavorans</i> F199	hypothetical protein located on catabolic plasmid	(Romine <i>et al.</i> , 1999)
ORF_ThAr	AJ001830	27/49	<i>Thauera aromatica</i> K172	hypothetical protein. ORF is upstream of genes required for anaerobic catabolism of 4-hydroxybenzoate	(Breese and Fuchs, 1998)
ORF1_MyLe	Z99494	n.a.	<i>Mycobacterium leprae</i>	hypothetical protein	Unpublished
ORF145_StSc	Y13094	24/49	<i>Staphylococcus sciuri</i> K11	hypothetical protein	(Wu <i>et al.</i> , 1998)
ORF7_BcFi	U89914	n.a.	<i>Bacillus firmus</i> OF4	hypothetical protein	Unpublished
PapX_EsCl	X61239	20/42	<i>E. coli</i>	hypothetical protein	(Marklund <i>et al.</i> , 1992)
PecS_ErCh	X74409	22/48	<i>Erwinia chrysanthemi</i>	Represses a variety of pectinase and cellulase genes	(Reverchon <i>et al.</i> , 1994)
PetP_RbCa	Z12113	28/54	<i>Rhodobacter capsulatus</i>	may control expression of genes required for photosynthetic growth	(Tokito and Daldal, 1992)
RV14_MyTu	Z80108	21/39	<i>Mycobacterium tuberculosis</i> H37Rv	hypothetical protein	(Cole <i>et al.</i> , 1998)
SlyA_SITy	U03842	25/55	<i>Salmonella typhimurium</i>	Activates genes required for haemolysis and infection of and survival in mouse macrophages	(Buchmeier <i>et al.</i> , 1997; Libby <i>et al.</i> , 1994; Oscarsson <i>et al.</i> , 1996)
YkoM_BcSu	AJ002571	28/45	<i>Bacillus subtilis</i>	Hypothetical protein	Unpublished

¹ because these values are based on BLASTP search results, the numbers may reflect what is observed for only a portion of the protein or ORF

Figure 3.2 Alignment of CbaR to nineteen other MarR-like proteins. See Table 3.2 for GenBank accession numbers and information regarding function, if known, for the proteins. Because of space considerations, the N-terminus of some sequences was shortened by up to eleven amino acids. Identical amino acids present in 33-49% of the sequences are shaded in light gray; in 50-74% of the sequences, dark gray; and in $\geq 75\%$ of the sequences, black. A consensus sequence highlighting amino acids or classes of amino acids present in 38 MarR-like sequences is shown below the alignment. Identical amino acids or classes of similar amino acids are underlined if present in $\geq 75\%$ of the sequences. A class of similar amino acids (italicized) is only shown if present in $\geq 75\%$ of the sequences and the amino acid indicated is the most common one. Amino acids in upper case are present in $\geq 50\%$ of the sequences and those in lower case are present in 33-49% of the sequences. A secondary structure prediction is also presented below the alignment (H=helix, E=sheet, and L=loop). Predictions have a reliability ≥ 0.6 . Also indicated is a secondary structure prediction for this family made by Dalrymple and Swadling (1997), and the signature sequence for MarR-like proteins, according to Sulavik *et al.* (1995).

Expression and purification of CbaR

IPTG-induced cultures of *E. coli* M15 (pQE30cbaR, pREP4) produced large amounts of a protein that co-migrated with the 20.1 kDa standard (Figure 3.3). This is very close to the predicted molecular weight of CbaR with the affinity tag (20.7 kDa). Following Ni-NTA column purification and dialysis, the majority of extraneous protein was removed so that the preparation used in all assays was judged to represent CbaR at near homogeneity (Figure 3.3).

Binding of CbaR to DNA templates containing P_{cbaR} or P_{cbaA}

Initial studies determined if CbaR, putatively a repressor, bound DNA templates containing the putative transcriptional start site of *cbaR* or the known start site of *cbaA*. The former would indicate if CbaR potentially repressed its own expression while the latter would indicate if CbaR in any way interacted with the promoter it was proposed to regulate. CbaR did not affect the migration of DNA containing the putative transcriptional start site of *cbaR* but affected the migration of the two templates containing the transcriptional start site of *cbaA* (Figure 3.4). CbaR caused the appearance of two prominent slower-migrating forms of DNA, referred to as complex 1 (C1) and complex 2 (C2). The effects of CbaR were reversed by addition of 4 mM 3Cba (Figure 3.4), which was tested because it is a substrate for the *cba* catabolic pathway and would presumably reverse binding by any repressors. As a control, the effect of purified extract from *E. coli* M15 (pQE30, pREP4) (i.e. expression vector with no insert) was tested on the second and third template and no retarded forms were observed (not shown).

Figure 3.3 Expression and purification of histidine-tagged CbaR, as monitored by electrophoresis in a denaturing sodium dodecylsulfate gel (15% w/v polyacrylamide). A) Gel demonstrating that CbaR expression is controlled. A culture was allowed to reach log phase, divided in two, and one half was treated with IPTG (induced) while the other was not (uninduced). Both cultures were then incubated for ~6 hrs and proteins from whole cells were analysed. The arrow indicates CbaR, which is over-expressed in induced cultures. B) Gel demonstrating various steps in the CbaR purification procedure. Indicated are proteins from: induced cells; the pellet that results following cell rupture and centrifugation; the flow-through following binding of the post-rupture supernatant to the affinity matrix; and successive eluates from the the matrix following washes. Molecular weight of the size standards, in kDa, is indicated to the right of the gel.

Figure 3.4 Gel shift assay to map area bound by CbaR. A) Three templates were tested, representing various areas of the intergenic region between *cbaR* and *cbaA*. Template 1 contains the putative transcriptional start site of *cbaR* while Templates 2 and 3 contain the transcriptional start site of *cbaA* but differ in the amount of upstream DNA. B) Gel demonstrating that CbaR does not affect mobility of Template 1 but binds to Template 2 and 3, resulting in two protein-DNA complexes, designated C1 and C2 (indicated). The effects of CbaR could be reversed by addition of 4 mM 3-chlorobenzoate (3Cba).

Stoichiometry of CbaR binding to P_{cbaA} DNA and lowest concentration at which 3Cba disrupts CbaR binding

Preliminary work investigating interactions between CbaR and P_{cbaA} DNA showed that two protein-DNA complexes result, implying two binding sites. To determine if affinity towards these sites was equal, the stoichiometry of their formation was studied and results are shown in Figure 3.5A. Small amounts of the first complex (C1) appeared when CbaR and P_{cbaA} DNA were at approximately equimolar concentrations, and equal amounts of unbound DNA and C1 DNA resulted with a 10-fold molar excess of CbaR. Small amounts of the second complex (C2) were already apparent by this point, but equal amounts of C1 and C2 did not appear until there was an ~250-fold molar excess of CbaR. C2 did not predominate until there was at least an ~2500-fold molar excess of CbaR. The apparent dissociation constant (Praillet *et al.*, 1996) for C1 was judged to be $\sim 10^{-8}$ M, and for C2 $\sim 10^{-6}$ M. Preliminary work presented in Figure 3.4 had shown that CbaR-P_{cbaA} DNA complexes could be disrupted by 3Cba at 4 mM and a more detailed study found that 3Cba at 40 μ M still elicited a significant effect (Figure 3.5B).

Figure 3.5 Gel shift assays to study A) stoichiometry of CbaR-P_{cbaA} DNA complex formation and B) minimum concentration at which 3-chlorobenzoate (3Cba) inhibits CbaR binding. In A), the amount of CbaR per reaction was varied, while the amount of P_{cbaA} DNA was kept constant (4 ng). For reference purposes, the approximate molar ratio of CbaR to P_{cbaA} DNA is indicated. In B), the amount of 3Cba per reaction was varied, while the amount of CbaR and P_{cbaA} DNA were kept constant (0.4 µg and 4 ng, respectively, per reaction). Indicated for both gels are unbound template (Free DNA), CbaR-P_{cbaA} DNA complex 1 (C1), and CbaR-P_{cbaA} DNA complex 2 (C2).

Effectors of CbaR binding to P_{cbaA} DNA

To better characterize how compounds that are structurally similar to 3Cba affect binding of CbaR to P_{cbaA} DNA, gel shift assays with a wide variety of benzoic acids and several benzene compounds were conducted. The main reason these studies were conducted was to determine if all substituted benzoates disrupt CbaR binding and, if not, which features of a benzoate make it a CbaR effector. As well, many of the compounds tested are potential growth substrates for BR60 (pBRC60) and/or other *C. testosteroni* strains (indicated in Table 3.2), and this work would potentially provide some insight into how substrates for other degradative pathways interact with factors controlling *cbaABC* expression.

Not all compounds affected CbaR-P_{cbaA} DNA complexes, but of the ones that did three types of effects were exhibited: inhibition of binding comparable or identical to what is observed with 4 mM 3Cba (i.e. decrease or disappearance of C1 and C2 and increase in free DNA); weak inhibition of binding (i.e. slight decrease in C2 with a concomitant increase in C1 and slight increase in free DNA); and unexpectedly, improvement of binding (i.e. increase of C2 and decrease of C1 and free DNA). The gels demonstrating the effect of all the test compounds are presented in Figures 3.5 to 3.8 and results are summarized in Table 3.3.

Whether a substituted benzoic acid had any effect depended on the number, nature, and position of the substituent(s). With mono-substituted benzoates, no definite trends were evident. A Cl, NO₂, or CH₃ group in position 2 or 3 inhibited binding but in position 4 had no effect (Figures 3.6A, 3.7B, and 3.7C). An OH and COOH group in position 3 or an OCH₃ in position 4 improved binding, but in other positions either had no effect or weakly inhibited binding, at best (Figures 3.7A, 3.8A, and 3.8B). An NH₂ or SO₄ group had no effect on binding, regardless of position (for the isomers tested; Figures 3.7A and 3.7B).

With some di-substituted benzoates, a trend was evident. For example, inhibition of binding occurred with dichlorobenzoates provided one Cl was in the *meta* position (i.e. position 3 or 5; Figure 3.6A). Similarly, the two dihydroxybenzoates tested inhibited

binding and each had one OH in the *meta* position (Figure 3.8A). Results with other di-substituted benzoates varied. 3,4-dimethoxybenzoate and 3-methoxy, 4-hydroxybenzoate improved binding but 3-hydroxy, 4-methoxybenzoate had no effect (Figure 3.8B). Similarly, 3-nitro, 4-chlorobenzoate inhibited binding (Figure 3.7B) but 3-chloro, 4-hydroxybenzoate had no effect (not shown). None of the tri-substituted benzoates tested affected CbaR binding (Figures 3.6A and 3.8A).

A study to investigate how substituent size affects CbaR binding was conducted by testing the effect of three other halogens when in position 3 of benzoate. 3-fluoro and 3-bromobenzoate inhibited binding as well as 3Cba, while 3-iodobenzoate had no effect on binding (Figure 3.6B). To determine how important the COOH group of 3Cba is in determining its ability to disrupt CbaR binding, the effect of 3-chlorotoluene or 1-chloro, 3-nitrobenzene was tested (i.e. the COOH *meta* to Cl was replaced with CH₃ or NO₂, respectively), and these compounds were not able to disrupt binding (Figure 3.6C). Because biodegradation of 3-chlorobiphenyl by some bacteria results in production of 3Cba, its effect on CbaR binding was also tested and it weakly inhibited binding (Figure 3.6C).

The unexpected observation that some substituted benzoates improved binding was better characterized using as a model compound 3-carboxybenzoate. The lowest concentration at which an effect could be observed was studied, and 10 μ M 3-carboxybenzoate was able to improve binding, as shown by the small amounts of C2 that began to appear with this amount. However, a significant decrease in free DNA was not apparent until 100 μ M 3-carboxybenzoate (Figure 3.9A). Whether binding enhancement takes precedence over inhibition was studied by testing the effect of 3Cba and 3-carboxybenzoate together, and under these conditions CbaR binding was disrupted (Figure 3.9A). A study was conducted to quantify the improved level of CbaR binding to P_{cbaA} DNA, and 0.75 ng CbaR with 4 mM 3-carboxybenzoate was able to retard DNA as well as 30 ng CbaR with no 3-carboxybenzoate (Figure 3.9B).

Figure 3.6 Gel shift assays to study the effect on CbaR binding to P_{cbaA} DNA by A) mono-, di-, and some tri-chlorobenzoates; B) benzoate with a halogen in position 3; and C) several chlorobenzene compounds with various groups *meta* to chlorine. Each reaction contained 4 mM aromatic compound, 0.4 µg CbaR, and 4 ng P_{cbaA} DNA. Indicated for all gels are unbound template (Free DNA), CbaR-P_{cbaA} DNA complex 1 (C1), and CbaR-P_{cbaA} DNA complex 2 (C2).

Figure 3.7 Gel shift assays to study the effect on CbaR binding to P_{cbaA} DNA by benzoate substituted with a A) carboxy group; B) sulfo or nitro group; and C) amino or methyl group. Each reaction contained 4 mM aromatic compound, 0.4 µg CbaR, and 4 ng P_{cbaA} DNA. Indicated for all gels are unbound template (Free DNA), CbaR-P_{cbaA} DNA complex 1 (C1), and CbaR-P_{cbaA} DNA complex 2 (C2).

Figure 3.8 Gel shift assays to study the effect on CbaR binding to P_{cbaA} DNA by benzoate A) with no substituent or substituted with hydroxy groups and B) methoxy and hydroxy groups. Each reaction contained 4 mM aromatic compound, 0.4 µg CbaR, and 4 ng P_{cbaA} DNA. Indicated for all gels are unbound template (Free DNA), CbaR-P_{cbaA} DNA complex 1 (C1), and CbaR-P_{cbaA} DNA complex 2 (C2).

Figure 3.9 Gel shift assays to better characterize the effect of 3-carboxybenzoate on CbaR-P_{cbaA} DNA complex formation. A) The stoichiometry of the improved binding elicited by this compound was studied by varying the amount of CbaR per reaction, while keeping the amount of P_{cbaA} DNA (~40 ng) and 3-carboxybenzoate (4 mM) constant. B) The minimum concentration at which 3-carboxybenzoate elicits an effect was determined by varying the amount of this compound while keeping CbaR and P_{cbaA} DNA constant (0.4 µg and 4 ng, respectively, per reaction). As well, the effect of including 3-chlorobenzoate (3Cba) in a reaction mix with 3-carboxybenzoate was studied. Indicated for both gels are unbound template (Free DNA), CbaR-P_{cbaA} DNA complex 1 (C1), and CbaR-P_{cbaA} DNA complex 2 (C2).

Table 3.3 Summary of the effects of various aromatic compounds on CbaR binding to P_{cbaA} DNA. Based on gel shift assays shown in Figures 3.6 To 3.8. Compounds that are growth substrates for various *C. testosteroni* strains are indicated in bold.

	Inhibits binding	Weakly inhibits binding	Improves binding	No effect on binding
BENZOIC ACIDS				
No substituent		benzoate		
Halogen				
mono-Cl	2-Cl 3-Cl			4-Cl
di-Cl	2,3-diCl 2,5-diCl 3,4-diCl 3,5-diCl			2,4-diCl 2,6-diCl
tri-Cl				2,3,5-triCl 2,3,6-triCl 2,4,6-triCl
Fl, Br, or I	3-Fl 3-Br			3-I
Methyl	2-CH ₃ 3-CH ₃			4-CH ₃
Carboxy			3-COOH	2-COOH 4-COOH
Amino				2-NH ₂ 3-NH ₂ 4-NH ₂
Nitro ¹	3-NO ₂			4-NO ₂
Hydroxy		2-OH	3-OH	4-OH
mono-				
di and tri-	2,5-diOH ² 3,4-diOH			3,4,5-triOH
Methoxy			4-OCH ₃ 3,4-diOCH ₃	2-OCH ₃ 3-OCH ₃
mono-				
di-			3-OCH ₃ , 4-OH	4-OCH ₃ , 3-OH
Hydroxy and Methoxy				
Sulfo				3-SO ₄ 4-SO ₄
Other	3-NO ₂ , 4-Cl			3-Cl, 4-OH 3-OCH ₂ COOH
BENZENE COMPOUND		3-Cl-biphenyl		3-Cl-toluene 1-Cl, 3-NO ₂ -benzene

¹2-nitrobenzoate was not tested because the only commercially available source contained ~5% contamination of 3- and 4-nitrobenzoate

²2,5-dihydroxybenzoate, gentisate; 3,4-dihydroxybenzoate, protocatechuate

Binding sites of CbaR

The binding sites of CbaR on P_{cbaA} DNA were determined by DNaseI protection assays and results are presented in Figure 3.10. Approximately 6 ng of DNA were used per binding reaction, and recall that according to gel shift assays, 9.4 ng of CbaR (i.e. a ~9-fold molar excess of CbaR to P_{cbaA} DNA) will result mostly in C1 with very little C2, while 600 ng CbaR (i.e. a ~600-fold molar excess of CbaR to P_{cbaA} DNA) will result in approximately equal amounts of C1 and C2 (Figure 3.5A). This assay identified two protected regions: Site I, which was clearly evident with 9.4 ng CbaR; and Site II, which only became apparent with 600 ng CbaR (or if 3-carboxybenzoate, which improves binding of CbaR, was included). If 3Cba was included during DNaseI digestion, neither site was protected. The improvement in binding elicited by 3-carboxybenzoate, estimated previously by gel shift assay (Figure 3.9B), was also estimated by visual inspection of the autoradiogram. According to this assay, 9.4 ng of CbaR with 4 mM 3-carboxybenzoate protected Site I and Site II as well as 600 ng CbaR with no 3-carboxybenzoate (Figure 3.10).

The complete sequence of the 186 bp template along with certain features of P_{cbaA} and *cbaA* are shown in Figure 3.11. Site II overlaps the *cbaA* transcriptional start site and Site I is approximately 40 nt downstream. Site I and II share many identical nucleotides (indicated on Figure 3.11). As well, Site I possesses a 4 nt inverted repeat separated by 6 or 9 nt (5'-GTTG(N)_{6/9}CAAC-3'). Modified forms of the two possible inverted repeats are present on the lower strand of Site II (5'-GTTG(N)₆TAAC-3' or 5'-GTAG(N)₉TAAC-3').

Figure 3.10 DNaseI protection assay to determine specific region of P_{cbaA} DNA bound by CbaR. Template DNA was end-labelled with $^{32}\text{P}\text{O}_4$ on the non-coding strand ($^{32}\text{P}\text{-PCBA}$) or the coding strand ($^{32}\text{P}\text{-CBAA}$). Various amounts of CbaR were added, and the effect of 4 mM 3-chlorobenzoate (3Cba) or 3-carboxybenzoate (*isophthalate*) on CbaR protection of two regions (Site I and II) were tested. Samples were digested with DNaseI and loaded onto a sequencing gel next to sequencing reactions conducted with the appropriate primer.

Figure 3.11 DNA sequence of the 186 bp template used in the DNaseI protection assay. Indicated are the two regions protected by CbaR (Site I and II), shared bases between these two regions (nucleotides in bold), and two potential inverted repeats in Site I and the similar structures found in Site II (arrows). Also indicated are several features of *cbaA*, such as the transcriptional start site (+1), the upstream hexamer typical of σ^{70} -dependent promoters (-10), and the start codon (ATG).

|----- Site II -----|
 | -10 | +1
 5' CTCGGGTAACACCTAGATTTCGATCTGAAACCTCCGTATAGTTAC**AAAACCAACTACATGG**ATCGAATCATCTCAGGTTTTCCTCCCTCATATA
 3' GAGCCCAATTGTGGATCTAAAGCTAGACTTTTGGAGGCAT**ATCAATGTTTGGTTGATGT**ACCCTAGCTTAGTAGAGTCCAAAAGGGGAGTAT

← → ← →

|----- Site I -----|
 5' AGGGTTTGACAC**CAAGTTGCTAAACCAACAACAT**CCGCAAAATTTTCAGGAGCAAGCT**ATGGCAT**TGACCAAGGACAACGAAGATTGTGGTCCGAG
 3' TCCCAAAC**TGTGTTCAACGATTTTGGTTGTTG**TAGGCGGTTTAAAGTCCTCGTTCGATACCGTAACCTGGTTCCTGTTCCTTAACCAACGCTC

← → ← → ← →

cbaA==>

DISCUSSION AND SUGGESTIONS FOR ADDITIONAL RESEARCH

Comparison of CbaR primary sequence to other MarR-family members and structure information

The proposed role of CbaR (Chapter 2) is consistent with that of other MarR-like proteins, which can respond to aromatic effectors and are known to be involved in controlling expression of genes for the catabolism of certain aromatic substrates (Table 3.2). For the most part, regulators from this family are repressors (Table 3.2), but BadR, NhdR, and SlyA are activators of the operons they control (Egland and Harwood, 1999; Komeda *et al.*, 1996; Oscarsson *et al.*, 1996), while MexR appears to be both a repressor and an activator (Poole *et al.*, 1996). The MarR family of proteins has been defined relatively recently and characterization of members at both a functional and structural level is in the initial stages. For example, MarR from *E. coli* and PecS from *Erwinia chrysanthemi* are reported to be dimers (Martin and Rosner, 1995; Praillet *et al.*, 1996). This may be a common feature of MarR-like proteins, but whether CbaR acts as one too is unknown.

Membership in the MarR-family, especially for the more distantly related proteins, is based mostly on similarity around a central region of the protein (Sulavik *et al.*, 1995). Based on an alignment of 10 sequences, Sulavik *et al.* (1995) proposed the consensus sequence DxRxxxxx(L/I)TxG, where x represents any amino acid, as a signature for this family of proteins. Based on an alignment of 38 MarR-like proteins, this motif appears in fact to be part of a longer signature sequence that extends in both directions (Figure 3.2). The significance of this conserved sequence, though, remains unknown. Dalrymple and Swadling (1997) predicted a helix-turn-helix domain in the conserved central region for MarR like proteins (indicated in Figure 3.2), while the secondary structure prediction presented here suggests the presence of at least two helix-turn-helix domains; one at either the N-terminus or the central region, and another at the C-terminus (Figure 3.2).

Whether any of these proposed domains are involved in DNA binding remains to be shown. Deletion studies with MarR initially suggested that the N-terminus mediates attachment to DNA (Seoane and Levy, 1995). Nevertheless, as several studies showed,

point mutations (single and double) that change the primary sequence and affect the ability of MarR to bind target DNA and/or repress expression of the *mar* genes are not confined to any one area (Alekshun and Levy, 1999b; Ariza *et al.*, 1994; Asako *et al.*, 1997; Cohen *et al.*, 1993; Maneewannakul and Levy, 1996; Martin and Rosner, 1995; Oethinger *et al.*, 1998). This result could be because of inhibition of dimerization between the monomers, and therefore no clear role in DNA binding for any particular region of the primary sequence can be suggested (Alekshun and Levy, 1997).

An interesting aspect, though, of two potential helix-turn-helix domains proposed in Figure 3.2 is that the amino acids of the first helix in each domain are poorly conserved between different members of the family. If either of these domains is involved in DNA binding, it may explain why the operator regions recognized by CbaR (see below), MarR (Martin and Rosner, 1995), and PecS (Praillet *et al.*, 1996) do not have any significant nucleotide sequence homology (Figure 3.12). Individual members of the MarR family of repressors appear to bind to distinct DNA sequences, and this may be determined by unique amino acids in one of these domains. The more highly conserved central region (Figure 3.2) could conceivably mediate effector recognition, as proteins from this family share the common feature of responding to aromatic compounds. Obviously, crystallographic studies would be invaluable for elucidating the functional regions of MarR-like proteins.

Figure 3.12 Comparison of the CbaR operator to those of MarR and PecS.

Indicated is the sequence of the binding sites (boxed nucleotides), their spacing, and the position of these sites relative to the transcriptional start site (+1), the -10 and/or -35 hexamer, and the start codon (ATG or GTG) of the gene controlled by the repressor. Some secondary structures within these sites are also indicated (arrows), but it not known if they have any significance (for example, nucleotides directly contacted by the protein). Relevant information for MarR is in Alekshun and Levy (1997) and in Martin and Rosner (1995). Relevant information for PecS is in Praillet *et al.* (1996), who do not indicate the start codon of *celZ* nor the exact transcriptional start site of *outC* in this article. It is noteworthy that PecS binds to a variety of other promoters but the regions protected could not be determined (Praillet *et al.* 1996).

CbaR

modulates expression of genes for CbaABC (dioxxygenase and dehydrogenase involved in 3-chlorobenzoate degradation)

MarR

represses expression of gene for MarR, MarA (activator of various stress-responsive genes), and MarB (unknown function)

PecS

represses gene for CelZ (cellulase)

represses gene for OutC (secretion of pectinases and cellulases)

Expression of recombinant CbaR

The gene for CbaR was cloned into an expression vector that placed a histidine tag at the N-terminus of CbaR. The recombinant ~20 kDa protein was affinity purified by Ni-NTA chromatography to near homogeneity, as judged by SDS-PAGE (Figure 3.3). One preparation was used in all subsequent assays to avoid variations that could occur with different preparations. Contaminating proteins were present but they themselves do not appear to bind the DNA templates with P_{cbaA} (see Results) nor do they appear to interfere with CbaR binding. However, these contaminating proteins interfere with determinations of the CbaR concentration such that any quantitative conclusions drawn from experiments in which the amount of added CbaR is important (for example, estimations of $K_{d,app}$ for C1 and C2) are to be taken as slight to moderate underestimates. More accurate experiments would either require highly purified CbaR or densitometric determination of the amount of contaminating protein in the CbaR preparation. Furthermore, by all indications, the N-terminal histidine tag does not interfere with CbaR binding, but it may be worthwhile to confirm this by using CbaR lacking the tag.

It appears that pQE30cbaR must be maintained in pREP4-containing *E. coli*. When the ligation for this plasmid was transformed into *E. coli* JM103 or JM109, strains that express $LacI^q$ from the chromosome, the expected plasmid was recovered but transformants did not express CbaR in response to IPTG (see Materials and Methods). This may be because of mutations or deletions occurred in the promoter upstream of the multiple cloning site of pQE30, which has been observed for this vector when genes coding for potentially “toxic” proteins are cloned and expression is not tightly controlled (R. Nokbeh, personal communication). This would appear to imply that the presence of CbaR is somehow detrimental to *E. coli* such that the low levels of expression which presumably occur in JM103 and JM109 are not permissible and selection for promoter mutants occurs. This question was not investigated further.

Nature and specificity of CbaR binding to *cbaR-cbaA* intergenic region

Gel shift assays showed that CbaR does not bind to DNA containing the putative transcriptional start site of *cbaR* (Figure 3.4), implying no autorepression by CbaR. CbaR does bind to DNA containing the +1 site of *cbaA*, the area which would presumably possess a binding region for this protein if CbaR were involved in controlling *cbaA* expression. Furthermore, nucleotides upstream of -21 are not necessary for binding, as shown by the identical retardation pattern elicited by the two templates that only differ in this region (Figure 3.4). These results also suggest that CbaR does not exert any “long distance” upstream effect on *cbaA* expression by interacting with DNA from -21 to approximately -400. Whether CbaR exerts a “long distance” downstream effect by binding DNA beyond the twentyfifth codon for *cbaA* (the limit of the gel shift templates; Figure 3.4) is not known, or for that matter whether it interacts with any other regions of Tn5271. This may be worth studying (see more below) as at least one other regulator of aromatic catabolic genes is known to bind downstream of the target gene’s start codon and affect expression (Chapter 1, discussion on CatR-mediated control of *cbaBCA*).

The two CbaR-P_{cbaA} DNA complexes observed in initial gel shift assays (Figure 3.4) were characterized by studying the stoichiometry of the phenomenon (Figure 3.5) and correlating those data with DNaseI protection work (Figure 3.10). C1, the first complex, represents CbaR bound to a region between the transcriptional start point for *cbaA* and the *cbaA* start codon, designated Site I. C2, the second complex, represents CbaR bound to both Site I and another region overlapping the transcriptional start site, designated Site II (Figure 3.10). It is not known if one CbaR protein (whether it act as a monomer or a dimer) is interacting with both sites or if two distinct proteins are bound at each site. If the former, this would imply bending of the intervening DNA, which can be studied with gel shift assays (Crothers *et al.*, 1991). If the latter, it would be interesting to determine if cooperation between CbaR bound at the two sites occurs such that, for example, CbaR bound at Site I facilitates binding of another CbaR to Site II. Until more definitive experiments are performed (for example, showing that binding at either Site I or II occurs without the other site in *cis*), no firm conclusions are possible.

Site I appears to be a high affinity binding site ($K_{d_{app}} \sim 10^{-8}$ M) while Site II is of lower affinity ($K_{d_{app}} \sim 10^{-6}$ M; see Results). When binding inhibitors are present, binding to Site II is affected first as disappearance of C2 always precedes disappearance of C1 (for example, Figure 3.5). This is most likely because of the lower affinity of CbaR towards Site II, which would presumably render it more susceptible to compounds that disrupt the protein-DNA complex. It appears that P_{cbaA} evolved such that it contains two operators for CbaR. One, though, is of lower affinity and overlaps the +1 of *cbaA* (Figure 3.11), which may allow a more flexible response such that compounds which are potential substrates for the pathway can lower CbaR-mediated inhibition of transcription (see more below).

Site I and II are extremely similar and an inverted repeat is present on Site I and in modified forms on Site II (Figure 3.11). It is unknown whether the inverted repeat of Site I represents the nucleotides recognized by CbaR, but interestingly, many of the nucleotide differences between these sites lie within the two possible modified forms of this structure in Site II (indicated on Figure 3.11), which may explain the dissimilar binding affinities. More research is necessary to better define the exact nucleotides that CbaR directly contacts, either by hydroxyl radical footprinting (Dixon *et al.*, 1991) or site-directed mutagenesis of either Site I and II and additional gel-shift assays. This would also help determine if Tn5271 possesses other potential CbaR binding sites and which in some way affect expression of *cbaA* or other genes. Lastly, this may also help establish whether the lower affinity of Site II is fundamental to transcriptional regulation of *cbaA* (see more below).

CbaR shares various functional features with MarR and PecS (Table 3.2), the only other MarR family regulators that to date have undergone a fair degree of characterization. With respect to apparent binding constants, the $K_{d_{app}}$ determined for MarR (Martin and Rosner, 1995) and high and low affinity $K_{d_{app}}$ determined for PecS towards several different promoters (Praillet *et al.*, 1996) are comparable to the values estimated for CbaR towards Site I and Site II. As well, MarR and PecS form two to three protein-DNA complexes in a concentration-dependent fashion when combined with DNA

containing the respective cognate promoter, and the number of complexes reflect the number of binding sites (Martin *et al.*, 1996; Martin and Rosner, 1995; Praillet *et al.*, 1996; Seoane and Levy, 1995). However, the sequence, spacing, size, and position (relative to transcriptional start site and start codons) of binding sites for these regulators are quite variable (Figure 3.12), indicating that the manner in which members from this family exert control differs greatly.

SlyA and CinR are two other MarR family members (Table 3.2) that have undergone some functional characterization. SlyA may also form two protein-DNA complexes when combined with DNA containing target promoter sequences (Oscarsson *et al.*, 1996), but CinR seems to only form one (Dalrymple and Swadling, 1997). Unfortunately, further characterization of these and other MarR-like proteins has not been extensive and further comparisons are not possible.

The effect of aromatic compounds on CbaR binding to P_{cbaA}

General statements

CbaR responds to a variety of aromatic compounds (Table 3.2). Initial tests studied the effect of 3Cba, but in an effort to determine whether other aromatic compounds affect CbaR interactions with P_{cbaA} DNA, the influence on binding of a large number of benzoate compounds as well as several benzene compounds was studied. The major reason for these studies was to better understand features of the aromatic compound which make it a CbaR effector. However, another reason is that natural environments (contaminated and uncontaminated) contain a mixture of aromatic compounds and most degrading bacteria are capable of metabolizing more than one aromatic substrates. How expression of the *cba* pathway would be affected by competing growth substrates is worth determining because, as discussed in Chapter 1, the *cba* pathway is on a transmissible element and therefore has the potential for co-existing with host-encoded metabolic pathways for other aromatic substrates (see for example Table 3.3 which indicates those metabolized by *C. testosteroni* strains). Test compounds were restricted almost exclusively to benzoates. Nevertheless, catechol-like metabolites of pathways for non-carboxylated compounds may also occur in cells expressing *cbaA*,

therefore these metabolites should also be investigated for their ability to modulate CbaR binding.

Two major effects were elicited with the compounds tested: inhibition of binding, consistent with the proposed role of CbaR as a repressor of *cbaA* expression; and improvement of binding, which was unexpected. Except for 3Cba, benzoate, 3-hydroxybenzoate, 4-hydroxybenzoate, and protocatechuate, though, whose effects on 3Cba degradative activity and in the case of 3Cba and benzoate, *cba* expression, were studied in greater detail (Chapter 2), it is not known if these *in vitro* observations translate into an *in vivo* effect. Because of the complex nature of *cba* expression with the proposed involvement of CbaR as repressor and an unidentified protein as activator, an effect on CbaR binding to P_{cbaA} does not necessarily point toward an effect on *cbaA* transcription and an overall effect on 3Cba degradative activity, and detailed studies similar to those described in Chapter 2 would be necessary to study this question. It is not known, for example, if a compound that improves CbaR binding leads to “super-repression” of *cbaA*. Conceivably, this would be a beneficial effect which ensures that the *cba* pathway is not expressed when alternate growth substrates that are metabolized via other pathways are present. More studies, though, are necessary to determine whether there is a correlation.

It would appear, though, that even if a compound inhibits CbaR binding, it may not necessarily lead to de-repression of *cbaA*, as 3,4-dihydroxybenzoate (protocatechuate) disrupts CbaR binding (Figure 3.8A), but O₂ uptake data show that growth on this compound does not induce 3Cba degradative activity (Chapter 2, Table 2.2). Disruption of CbaR binding by protocatechuate, though, is conceivably a potentially beneficial effect, especially when one recalls that it is a metabolite of 3Cba, 4Cba, and 3,4-diCba degradation by the *cba* pathway (Chapter 1) and may represent a positive feedback mechanism for ensuring de-repression of *cba* under conditions in which protocatechuate is being produced by chlorobenzoate metabolism. This question warrants further studies.

Effectors that inhibit binding

3Cba inhibits CbaR binding to P_{cbaA} DNA, as shown by gel shift (Figure 3.4) and DNaseI protection assays (Figure 3.10). In these assays, inhibition of binding by 4 mM 3Cba was tested, which reflects the concentration typically used when cells are grown on this substrate (Chapter 2). It is noteworthy, though, that at 40 μ M (~6 ppm), 3Cba is still able to exert an effect (Figure 3.5B), and this may reflect an adaptation which allows BR60 (pBRC60) to respond to the small amounts of contaminants often encountered in the environment (Wyndham and Straus, 1988).

Disruption of CbaR binding appears to require that the group *meta* to Cl be a COOH (or COO⁻ at pH 7). This is suggested because neither 3-chlorotoluene nor 1-chloro-3-nitrobenzene elicited a response (Figure 3.6C). In contrast, 3-chlorobiphenyl (3-CBp) is a weak inhibitor of binding, even though the phenyl side group is much bulkier and more hydrophobic than COOH (Figure 3.6C). Whether this effect has any biological significance is unknown (see above), but it is worthy of notice because 3Cba in the environment may be a result of incomplete metabolism of 3-CBp by some bacteria (reviewed by Peel and Wyndham [1997]). Furthermore, the *cba* pathway co-exists with chlorobiphenyl degradative pathways in various environmental isolates (Peel and Wyndham, 1999). The weak effect this substrate exerts on CbaR may reflect a “priming” of the *cba* genes for chlorobenzoates that result from CBp metabolism.

A variety of chlorinated benzoates other than 3Cba were also tested and interestingly, the ability to elicit an effect was not a function of whether the compound was a substrate for the *cba* pathway. For example, 2Cba, 2,3-diCba, 2,5-diCba, and 3,4-diCba inhibited CbaR binding but only 3,4-diCba is oxidized by the *cba* pathway, while 4Cba, which is also oxidized (Nakatsu and Wyndham, 1993), did not inhibit CbaR binding (Figure 3.6A). It is noteworthy that 4Cba-utilizing BR60 (pBRC60) strains arise at a frequency typical of mutagenic events (Wyndham *et al.*, 1988), and the mechanism by which this adaptation occurs could involve changes to CbaR and/or P_{cbaA} such that *cba* can be expressed in response to 4Cba. This warrants further study.

Inhibition of binding was also elicited by a variety of other mono- and di-substituted benzoates, and for functional groups that elicited a response the importance of position is quite apparent. With mono-substituted benzoates, it appears that “binding inhibitor” groups (Cl, CH₃, NO₂) are equally effective whether in position 2 or 3 but are not able to elicit a response in position 4 (Figures 3.6A, 3.7B, 3.7C). For di-substituted benzoates, it appears that at least one of the “binding inhibitor” groups (Cl or OH) has to be in the *meta* position for the substrate to elicit a response (Figure 3.6A and 3.8A). More experiments, though, are necessary, to determine if some of the trends observed are in fact rules, by for example, studying whether all “binding inhibitor” groups that elicit a response in position 3 (F, Br, NO₂) also elicit a response in position 2; whether all groups that inhibit binding as mono-substituents (e.g. F, Br, CH₃, NO₂) do so as di-substituents; and whether all di-substituted compounds that inhibit binding require a *meta* substituent.

Another factor that determines whether binding is inhibited appears to be size, as F, Cl, or Br in position 3 of benzoate are “binding inhibitors”, while iodine is not (Figure 3.6B). The available data, though, do not allow a clear distinction to be made between steric, polar, or ionic effects in determining whether a substituent inhibits binding of CbaR to P_{cbaA} DNA. It appears that a complex interplay of factors is at work when it comes to whether a substituted benzoate inhibits binding, not only chemical nature, position, size, and number of substituents, but presumably the amino acids of CbaR which interact with the effector. Until more work is done to clearly establish which determinants are crucial, few predictions appear possible. For instance, neither oxidation state nor polarity nor hydrophobicity can be suggested as indicators of whether a monosubstituent will be a “binding inhibitor”. For example, 3-aminobenzoate elicits no response while 3-nitrobenzoate inhibits binding; and 2 and 3-methylbenzoate both inhibit binding while 2-carboxybenzoate has no effect and 3-carboxybenzoate improves binding. Nor is there any easy explanation why addition of another substituent to monosubstituted benzoates obviates inhibition of binding in some cases (for example, 3-Cl-4-hydroxybenzoate versus 3Cba; and 2,4- or 2,6-diCba versus 2Cba) but not another (e.g. 3-nitro-4-chlorobenzoate inhibits binding as well as 3-nitrobenzoate).

Effectors that improve binding

As mentioned above, some benzoates had the unexpected effect of improving CbaR binding to P_{cbaA} DNA (Table 3.3). Two separate methods were used to quantify the degree of improved binding mediated by 3-carboxybenzoate, the model compound used to study this effect. Based on gel shift assays and DNaseI footprinting work (see Results), it appears that a ~50-fold improvement occurs (i.e. the amount of binding or protection exhibited by CbaR with 3-carboxybenzoate is equivalent to that exhibited when ~50× more CbaR, in µg, is present without 3-carboxybenzoate). Because a molar excess of CbaR is present in these experiments, it could be argued that the improved binding observed reflects a shift in the equilibrium of unbound CbaR to bound CbaR, but until the exact nature of binding by CbaR to Site I and II is determined (see above), this conclusion is speculative. When 3Cba (a binding inhibitor) and 3-carboxybenzoate (a binding enhancer) are present together, the inhibitor obviates the effect of the enhancer (Figure 3.9B). This may be because both substrates are competing for the same site on CbaR and 3Cba conceivably binds more strongly or readily than 3-carboxybenzoate. However, the observed effect could also be a result of the compounds tested, so it may be necessary to study different combinations of substrates before the generality of this observation is firmly concluded.

The only common feature of substrates that enhanced binding is that none possessed a substituent in position 2 (Table 3.3). Like inhibitors of binding, no clear trend is apparent nor is any facile explanation available to describe the results. For example, why does benzoate at best weakly inhibit binding, 3-hydroxybenzoate improve binding, 3,4-dihydroxybenzoate inhibit binding, 3-hydroxy-4-methoxybenzoate have no effect, but 3-methoxy-4-hydroxybenzoate or 3,4-dimethoxybenzoate improve binding? No trend may be apparent because of the limited number of substrates tested, therefore additional experiments, in particular with di-substituted benzoates, may be in order (for example, test all possible dimethoxybenzoates and hydroxylated methoxybenzoates).

The serendipitous discovery that some substrates enhanced binding of CbaR to P_{cbaA} may provide some insight into the nature of MarR-like proteins and perhaps the

evolution of the *cba* pathway. As mentioned previously, several members from this family are activators of the genes they control (Table 3.2) and as discussed in Chapter 1, the effect exerted by a transcriptional regulator on a target gene can be a function of the position of its DNA binding site (e.g. LTTRs autorepress their own expression while activating expression of divergently transcribed proximal catabolic genes). It would be interesting to determine whether CbaR's ability to repress *cbaA* is merely a result of the location of its two binding sites (i.e. overlapping +1 as well as downstream of +1; Figure 3.11). Conceivably, CbaR or a CbaR-like protein may be a positive regulator of other metabolic pathways provided the binding sites are in the appropriate location and orientation relative to the transcriptional start site. Under these conditions, CbaR may activate transcription in response to substrates that enhance binding. It is noteworthy that all binding enhancers uncovered in the course of this research are growth substrates for BR60 (pBRC60) or other *C. testosteroni* strains, and whether this is a co-incidence or reflects CbaR's ancestry (e.g. a regulator of catabolic genes for one of these substrates) is a fascinating question.

Role of CbaR in controlling *cbaA* expression

All indications are that CbaR is a repressor of the *cba* pathway. CbaR binds specifically to P_{cbaA} and under conditions where expression of these catabolic genes is required (e.g. presence of 3Cba or 3,4-diCba), CbaR binding is disrupted. Benzoate, a compound that gratuitously induces this pathway, elicits a very weak effect on CbaR binding. Two binding sites are present on P_{cbaA} and CbaR's binding affinity to each differs. CbaR binds strongly to Site I (a region between +1 and the *cbaA* start codon) and ~100× less strongly to Site II (which overlaps +1). Binding to either site is almost completely inhibited by 4 mM 3Cba, but an effect can be elicited at Site II at concentrations as low as 40 μM. This level of binding disruption may be enough to allow a certain level of *cbaA* expression to occur, such that metabolism of substrates for this pathway may occur at the low concentrations which are normally encountered in contaminated environments. The interesting observation that protocatechuate (3,4-dihydroxybenzoate) also disrupts binding suggests that CbaR may have evolved such that it can respond to both the parent substrate

and at least one downstream metabolite of the pathway. The two together may have an additive or synergistic effect on CbaR binding, such that *cba4* expression is ensured at low concentrations. Unfortunately, the effect of 5-Cl-protocatechuate, the major 3Cba metabolite, could not be tested because it is not commercially available.

CHAPTER 4: ATTEMPT TO CLONE GENE FOR CbaS, THE PROPOSED ACTIVATOR OF OF THE *cba* PATHWAY

INTRODUCTION

Results presented in Chapter 2 suggest that a regulator encoded by the chromosome of *C. testosteroni* BR6020 can induce expression of the plasmid-encoded *cba* pathway in response to benzoate or 3Cba. This regulator was tentatively designated CbaS and so as to better characterize factors controlling *cbaA* expression, I attempted to clone the gene coding for this protein with a transcriptional interference vector (Figure 4.1). The principle of these vectors is to allow a DNA-binding protein to interfere with transcription from a strong, constitutive promoter, which is preventing expression, in the opposite direction, of an antibiotic resistance gene.

Figure 4.1 General strategy for selection of clones expressing proteins that bind to specific DNA sequences. A) Transcription of *aadA*, the gene conferring spectinomycin (Sp) and streptomycin (Str) resistance, is repressed because convergent transcription from P_{lac} is stronger than transcription driven by P_{aadA} , making cells sensitive to these antibiotics. B) Cells are transformed with a gene library which, in the work presented here, is maintained in a pUC-based plasmid with an ampicillin resistance (Amp^r) determinant. Clones expressing a protein (filled circles) capable of binding to a specific DNA sequence (Op) cloned downstream of P_{lac} can be selected for because cells are now spectinomycin and streptomycin resistant as *aadA* expression is no longer interfered with. Adapted from Elledge *et al.* (1989).

A.

Sp'/Str'

B.

Sp'/Str' and Amp^r

MATERIALS AND METHODS

Bacterial strains and plasmids

Bacterial strains and plasmids used in this study are listed in Table 4.1. Unless otherwise stated, *E. coli* strains were grown at 37°C in Luria-Bertani medium (LB; 1% w/v tryptone, 0.5% w/v yeast extract, 0.5% w/v NaCl) containing ampicillin (250 mg/L), chloramphenicol (50 mg/L), streptomycin (100 mg/L), and/or spectinomycin (100 mg/L), as required. When necessary, 100 mg/L 5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside (X-gal) was included in solid media. Antibiotics and X-gal were purchased from Sigma-Aldrich Ltd. (Oakville, Ontario, Canada).

DNA sequence manipulation

General information

All enzymes used to generate the clones described below were purchased from New England Biolabs (Mississauga, Ontario, Canada). Plasmids were isolated and DNA was manipulated according to standard methods (Sambrook *et al.*, 1989) and *E. coli* was transformed by electroporation using a Gene Pulser (Bio-Rad Laboratories, Ltd., Mississauga, Ontario, Canada), according to the manufacturer's recommendations. Oligonucleotides were synthesized and purified by high performance liquid chromatography at the Biotechnology Research Institute (University of Ottawa, Ottawa, Ontario, Canada).

Generating transcriptional interference vector for CbaS

Expression studies with various *cbaA-lacZ* constructs implied that nucleotides from -52 to -21 (relative to the *cbaA* transcription start site) are important for proper expression of *cbaA* and, more importantly, benzoate-induced transcription (Chapter 2). Nucleotides spanning -52 to +1 (Figure 4.2) were therefore used in designing oligonucleotides containing a potential binding site for CbaS. The area that was eventually selected included upstream sequence that was interrupted in *cbaA'*-*lacZ3* (Figure 4.2) and therefore was presumed to represent DNA necessary for CbaS interaction with P_{cbaA}. It also contained an intriguing inverted repeat that may play a role in regulator binding (Figure 4.2).

Figure 4.2 Selection of putative CbaS DNA binding site and cloning into transcriptional interference vector. A) Sequence around the *cbaA* promoter and area selected as a potential CbaS binding site (underlined nucleotides). Also indicated are an inverted repeat present in this region, the boundaries of the various *cbaA-lacZ* expression constructs described in Chapter 2 and several features of *cbaA* (transcriptional start site [+1], the putative -10 hexamer, the putative ribosome binding site of *cbaA* [RBS], and the N-terminus of *cbaA*) B) The DNA fragment that results when oligonucleotides CBAXOPU and CBAXOPL2 are annealed. Indicated is the inverted repeat present in this putative CbaS binding site (nucleotides in bold). This fragment was cloned into the *Bam*HI and *Pst*I site of pNN396 (shown in C), which places it 10 nucleotides downstream of +1 (shown in D), the start site of the transcript driven by P_{lac}, a constitutive strong promoter. See text for additional detail regarding construction of the transcriptional interference vectors.

A) cbaA'- 'lacZ1 | ==>
4374 TGGCGGCGTCGGCATGAGCCGCACTGAGCGGCAGCGCGCGCCGCCACCGCCGCTTCACGCGCCCGCACAGCGGCCGCT

4454 cbaA'- 'lacZ2 | ==> cbaA'- 'lacZ3 | ==> |-10 | +1
GCCCGCTGCATCACCTGCCGTCGTCGGCTCGGCTAAACACCTAGATTTTCGATCTTGAAAAACCTCCGTATAGTTACAAAAC

4534 <----->
CAACTACATGGATCGAATCATCTCAGGTTTTCCTCCTCATAAGGGTTTGACACAAAGTTGCTAAACCAACATCCGCAAA

4614 |RBS|
ATTTCAGGAGCAAGCT ATG GCA TTG ACC AAG GAC AAC GAA GAT TTG GTG CGA GTG GGC CGT GGA
CbaA⇒ Met Ala Leu Thr Lys Asp Asn Glu Asp Leu Val Arg Val Gly Arg Gly

B) <----->
CBAXOPU: 5' GATCCACACCTAGATTTTCGGATCTTGAAAAACCTCCGTCTTGCA 3'
CBAXOPL2: 3' GTGTGGATCTAAAAAGCTAGACTTTTGGAGGCAG 5'
BamHI PstI

C) |-10 | +1
TTGACAAATTAATCATCGGCTCGTATAATGGTACCCGGGGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTT
KpnI SmaI BamHI XbaI Sali PstI SphI HindIII

D) |-10 | +1 <----->
TTGACAAATTAATCATCGGCTCGTATAATGGTACCCGGGGATCCACACCTAGATTTTCGATCTGAAAAACCTCCGTCTGCA
KpnI SmaI BamHI PstI

Two oligonucleotides were synthesized: CBAXOPU (5'-gatccACACCTAGATTTTCGATCTGAAAACCTCCGTctgca-3') and CBAXOPL2 (5'-gACGGAGGTTTTCAGATCGAAAATCTAGGTGTg-3'). Lower case nucleotides represent DNA that is not native to P_{cbaA} and underlined nucleotides represent the *Bam*HI- and *Pst*I-complementary overhangs, respectively, generated when the two oligos are annealed. Equimolar amounts of the oligos were heated to 95°C for 5 minutes, allowed to cool slowly to room temperature, and then cloned into pNN396 digested with *Bam*HI and *Pst*I, resulting in pNN396OP.1. This placed the putative CbaS binding site proximal to the start site for the transcript that is driven by P_{tac} , a strong constitutive promoter (Figure 4.2). The strong promoter and the putative binding site were then removed as a *Not*I/*Hind*III fragment and sub-cloned into pNN388 digested with the same enzymes, resulting in pNN388OP.1. Transcription from P_{tac} thus silences transcription from P_{aadA} and makes cells harbouring this vector sensitive to spectinomycin and streptomycin (Elledge and Davis, 1989; Elledge *et al.*, 1989).

Generating a plasmid library of *C. testosteroni* BR6020 chromosomal DNA

A plasmid library of chromosomal DNA from *C. testosteroni* BR6020 was generated to create a pool of clones that (may) contain DNA coding for CbaS. BR6020 is the pBRC60-cured variant of BR60 (pBRC60) (Wyndham *et al.*, 1988) used for the expression studies described in Chapter 2. Chromosomal DNA was isolated from BR6020 (Ausubel *et al.*, 1989) and quantified by UV spectrophotometry.

Using procedures described in Boulnois (1987) as a guide, the plasmid library was generated as follows. In a total reaction volume of 260 μ L, 400 μ g DNA was mixed with 50 U *Sau*3AI and digested at 37°C. After 10 minutes and 20 minutes, 130 μ L aliquots were removed, mixed with 300 μ L H₂O, and extracted twice with equal volumes of phenol-chloroform. DNA was precipitated with *isopropanol* and redissolved in 50 μ L sterile H₂O. The samples were centrifuged to *pellet all* undissolved matter and DNA in the supernatants was resolved by electrophoresis in a 0.9% NuSieve GTG low melt agarose

gel (FMC BioProducts, Mandel Scientific Company Ltd, Guelph, Ontario, Canada) at 1 V/cm for ~18 hrs. A pool of DNA in the size range of ~16.5 to ~6.8 kb was extracted from the gel and 2 µg was added to a ligation mix containing 1.5 µg of pUC18Not.1 digested with *Bam*HI and dephosphorylated with calf alkaline phosphatase. The mix (30 µL total volume, 400 U T4 ligase) was incubated at 15°C for ~18 hrs, aliquots electrotransformed into *E. coli* JM109, and cells were plated on solid LB+ampicillin+X-gal at a dilution that ensured adequate spacing between colonies. Transformants were either white (~50%), dark blue (~50%), or light blue (<1%). Dark blue colonies were ignored as they were presumed to represent religated vector. The other colonies were screened as follows: the average size of inserts and the number of false positives or false negatives was determined by picking twenty five white colonies and five light-blue colonies at random, purifying plasmid DNA, and digesting with *Not*I. This enzyme was chosen because the multiple cloning site of pUC18Not.1 contains a *Not*I site at either end (Table 4.1) and if an insert were present, *Not*I digestion would completely release it. As well, because this enzyme recognizes an 8 bp sequence, it would cut the insert rarely, if at all, and few fragments would be generated. Digested DNA was then subjected to agarose gel electrophoresis with appropriate size standards.

All white colonies tested contained an insert (i.e. none were false positives), as did the light blue colonies (i.e. all were false negatives), therefore transformants chosen for the library were restricted exclusively to these types of colonies. A library of 2304 clones was created by transferring well-isolated colonies to sterile 96 well plates containing 100 µL LB+ampicillin per well, incubating plates overnight, and adding 100 µL 50% w/v glycerol per well. Twenty four plates were filled and these were stored at -70°C.

Screening the plasmid library for clones containing CbaS

First round of screening

Twenty four initial pools of library DNA, each representing 96 clones, were created by transferring bacteria from each 96 well plate to solid LB+ampicillin, incubating the petri dishes overnight, harvesting growth from one plate with 6 mL of sterile resuspension buffer (50 mM glucose, 25 mM Tris-HCl, 10 mM EDTA, pH 8.0), purifying

plasmid DNA, and redissolving it in 200 μL H_2O . Aliquots (0.5 μL) from these twenty four pools were combined such that the two subsequent pools, each representing 1152 clones, contained library plasmids from pool 1 to 12 (designated pLIB1-12) or pool 13 to 24 (designated pLIB13-24). Two *E. coli* JM109 (pNN388OP.1) cultures were then transformed with a 1.5 μL aliquot of either pool and potential clones coding for CbaS were recovered by plating cells on solid LB+chloramphenicol+streptomycin+spectinomycin+ampicillin. As well, appropriate dilutions of cells were plated on LB+chloramphenicol to estimate total number of bacteria; LB+chloramphenicol+ampicillin to estimate number of transformants; and LB+chloramphenicol+streptomycin+spectinomycin to estimate number of cells that spontaneously develop resistance to the latter two antibiotics.

Plates were incubated at 32°C and 18 transformants that arose within 24 hrs on the first medium were picked and their plasmid DNA was extracted. The ten that came from the first pool were designated pLIB1-12.1 to pLIB1-12.10 and the eight that came from the second pool were designated pLIB13-24.1 to pLIB13-24.8. Two plasmids were present in each transformant extract (pLIB clone and pNN388OP.1) and to avoid interference by the latter in restriction analyses (see below), an aliquot of each mixture was transformed into *E. coli* JM109 and cells were recovered on LB+ampicillin, which selected for the pLIB vector. To ensure that pNN388OP.1 had not been taken up, transformants were tested for sensitivity towards chloramphenicol.

Plasmid DNA was isolated, digested with *NotI* or *PstI*, and analysed by agarose gel electrophoresis with appropriate size standards. *NotI* was used to estimate the size of the inserts (see above for rationale) and *PstI* was selected because it was found to cut inserts very frequently and thus generate many restriction fragments. This “restriction fragment length polymorphism” approach was used as an initial screen to determine whether one clone had been recovered multiple times, and/or if clones possessed common sequences (i.e. plasmids had been recovered that represented overlapping regions of the BR6020 chromosome).

Second round of screening and Southern blot analysis

Initial analysis suggested that these clones represented eighteen distinct loci, therefore a second round of selection was conducted as follows: aliquots of the eighteen clones isolated in the first round of selection were combined as before to create a more restricted pool. This pool was then subjected to a second round of selection as described above to see which plasmid, if any, predominated, the assumption being that if any coded for a DNA binding protein (i.e. a “true” positive), it would be selected over the other “false” positives as the pool represented far fewer potential clones. Forty three transformants were recovered on the selection medium, plasmid DNA was extracted, and the identity of the pLIB vectors determined by comparing the *Pst*I restriction pattern to those of the original clones from the eighteen member pool. Five clones were recovered at a frequency deemed to be higher than random (i.e. $>1/18$ or $>5\%$).

The possibility that these clones contained shared regions that were not detected by restriction analysis was studied by Southern blotting, as follows: Inserts from the five clones recovered in the second round of selection were removed as *Not*I fragments and randomly labelled with digoxigenin (Roche Molecular Biochemicals, Laval, Québec, Canada). Plasmid DNA from the eighteen clones recovered in the first round was digested with *Not*I or *Pst*I, electrophoretically resolved in 0.9% agarose, depurinated by illuminating with UV light for approximately 10 minutes, and then transferred to a positively charged nylon membrane (Roche Molecular Biochemicals) with a VacuGene XL vacuum blotter (Amersham-Pharmacia Biotech, Baie d’Urfé, Québec, Canada), according to manufacturer’s instructions. Blots were probed under high stringency with each of the digoxigenin-labelled fragments and probe-positive fragments were detected colourimetrically, according to manufacturer’s instructions (Roche Molecular Biochemicals).

Table 4.1 *E. coli* strains and plasmids used for experiments described in Chapter 4.

Strain and plasmid	Relevant characteristics of plasmid(s) ¹	Reference
JM107 (pNN396)	Amp ^r . MCS. Used for subcloning of protein binding site sequences into pNN388	(Elledge and Davis, 1989; Elledge <i>et al.</i> , 1989)
JM107 (pNN388)	Cm ^r , Str ^r /Sp ^r . When subclone P _{tac} with putative DNA binding site into pNN388, this constitutive promoter silences <i>aadA</i> (gene for Str ^r and Sp ^r) because of strong transcription towards P _{aadA} , a weaker promoter	(Elledge and Davis, 1989; Elledge <i>et al.</i> , 1989)
JM109 (pNN396OP.1)	Amp ^r . pNN396 with putative binding site for CbaS cloned into <i>Bam</i> HI and <i>Pst</i> I site and thus proximal to +1 of P _{tac}	This study
JM109 (pNN388OP.1)	Cm ^r , Str ^r /Sp ^r . pNN388 with P _{tac} and putative binding site for CbaS sub-cloned as a <i>Not</i> I/ <i>Hind</i> III fragment	This study
JM109 (pUC18Not.1)	Amp ^r , α - <i>lacZ</i> / pUC18Not-based plasmid (see Chapter 2) with a modified MCS ²	This study
JM109 (pLIB)	Amp ^r . Generic designation for a plasmid from a library of <i>Sau</i> 3AI-digested <i>C. testosteronei</i> BR6020 chromosomal DNA cloned into pUC18Not.1	This study

¹ Abbreviations: MCS, multiple cloning site; Amp, ampicillin; Cm, chloramphenicol; Sp, spectinomycin; Str, streptomycin; r, resistant

² The sequence of the MCS is GCGGCCGCGC GAATTCGAGC TCCACCGCGG TGGCGGCCGA TGCATATTTA AATCCCCCGG GGGATCCTCT AGAGTCGA CCTGCAGGCA TGCAAGCTTG CGGCCGCCC AGCTTGGCAC TGGCC, which contains, in order of appearance, the restriction sites for *Not*I, *Eco*RI, *Sac*I, *Sac*II, *Eag*I, *Nsi*I, *Swa*I, *Sma*I, *Bam*HI, *Xba*I, *Sal*I, *Pst*I, *Sph*I, *Hind*III, and *Not*I. The first *Not*I site is proximal to the start codon to the of α -*lacZ*.

RESULTS

The average size of cloned fragments in the plasmid library of *C. testosteroni* BR6020 chromosomal DNA was $\sim 5.5 \pm 1.7$ kbp ($n=30$). Assuming a genome size of 4.2 Mbp (i.e. similar to *E. coli*), there is a 95% probability that a sequence will be represented in a library of ~ 2300 clones (Boulnois, 1987). This library was screened with a transcriptional interference vector which was meant to select for clones that code for a protein capable of binding to a specific section of P_{cbaA} (Figure 4.2). Eighteen clones obtained in the first round (pLIB1-12.1 to 10 and pLIB13-24.1 to 8) were studied by restriction analysis of the inserts, but none appeared to have DNA fragments in common (Figure 4.3). This suggested that each clone represented a distinct locus of the BR6020 chromosome and there was no way of determining which clone deserved more in-depth analysis.

A second round of selection was thus conducted with a pool of these eighteen clones under the assumption that if one truly coded for a DNA-binding protein, it would be selected for to the exclusion of others as it now represented $\sim 5\%$ of the library (i.e. 1 in 18 clones) instead of $\sim 0.09\%$ (i.e. 1 in 1152 clones). The five clones detected at a significant frequency ($n=43$) were pLIB1-12.10 (11/43), pLIB1-12.9 (6/43), pLIB13-24.2 (4/43), pLIB13-24.5 (4/43), and pLIB13-24.8 (4/43). The possibility that these five clones contained shared regions, either between themselves or with the other thirteen clones, and not detected by *Pst*I restriction analysis was studied by Southern blot analysis. Probes made from the former five clones only hybridized with the clone from which they were originally obtained, implying that these clones represent five distinct loci.

In the first and second round of selection, the rate at which potential CbaS-coding clones appeared (i.e. resistance to all antibiotics) was $\sim 2 \times 10^{-6}$ while the spontaneous rate at which cells became spectinomycin- and streptomycin-resistant was $\sim 4 \times 10^{-6}$.

Figure 4.3 Restriction analysis of clones from a plasmid library of *C. testosteroni* BR6020 chromosomal DNA recovered in the first round of screening with transcriptional interference vector pNN388OP.1. Clones were digested with *Pst*I, subjected to electrophoresis in a 1.5% agarose gel, and the gel was stained with ethidium bromide. Indicated on the left is the size, in basepairs, of several standards from the 100 bp ladder, and on the right the size, in kilobasepairs, of several standards from the λ -*Hind*III ladder.

DISCUSSION AND SUGGESTION FOR FURTHER RESEARCH

It does not appear that the clones obtained in this screen are expressing a DNA-binding protein. The main line of evidence in support of this conclusion is that the frequency at which transformed cells are spectinomycin and streptomycin resistant (i.e. the indicator for transcriptional interference) was practically identical to the spontaneous rate at which cells become spectinomycin and streptomycin resistant. Furthermore, neither the eighteen clones isolated in the first round of selection nor the five which predominated in the second round appear to have any common DNA, as suggested by restriction digestion analysis (Figure 4.3) and Southern hybridization (see Results). If true selection were occurring, one would expect either one clone, or at least a set of clones containing overlapping DNA, to predominate. The likelihood that eighteen different loci in BR6020 code for proteins capable of binding to the putative binding site used in the transcriptional interference vector is low. In the second round of selection, clone pLIB1-12.10 showed up in ~25% of the transformants (n=43) but it was not studied further because it was also the smallest of the eighteen (1.86 kb insert) and most likely represented the plasmid which would be taken up with the highest frequency. However, because one of these clones may in fact be expressing a DNA binding protein, gel shift assays similar to the ones described in Chapter 3 should be conducted to ensure that these are in fact false positives.

Possible reasons this method may not have worked are that the putative DNA binding sequence cloned into the transcriptional interference vector does not represent the CbaS binding site; only represents a portion of it; and/or may not be proximal enough to the $P_{lac} +1$ site so that even if CbaS binding occurred, it was not able to interfere with transcription effectively. As well, CbaS may not be expressed very in *E. coli*, or at least not at levels which allow these vectors to be effectively used. Either way, if this method is to be used, the binding site must be characterized more carefully. This could be accomplished by testing crude protein extracts from *C. testosteroni* BR6020 for their ability to bind to P_{cbaA} -containing DNA (e.g. gel shift assays) and by characterizing protein-DNA interactions (e.g. mapping the area bound via DNaseI protection assays and determining if aromatic substrates affect binding). This information could be used to

rationaly design oligonucleotides containing the DNA binding sequencem which could subsequently be cloned into the transcriptional interference vector and the plasmid library could be screened once more. Alternatively, the sequence which is bound can be used to affinity purify CbaS, either by column chromatography (Singh *et al.*, 1988) or with commercially-available magnetic bead products (e.g. Dynal Inc, Lake Success, New York, USA or Roche Molecular Biochemicals, Laval, Québec, Canada), and the N-terminus of this protein could conceivably be sequenced. Degenerate oligonucleotides based on this information could then be used to screen the plasmid library by Southern hybridization.

The likelihood that the plasmid library is somehow at fault would appear to be low. Because of the average size of inserts, chances are good (~95%) that the complete genome is represented (see Results). As well, the library described here has already been used to isolate the *pmd* genes (coding for the protocatechuate extradiol dioxygenase enzymes). Three clones were obtained containing overlapping regions of the chromosome by a colourimetric assay which requires expression of active enzymes (Providenti *et al.*, 1999b). This indicates that the library contains some functional genes that can be maintained and expressed in *E. coli*, even though they are from an unrelated bacterium (*C. testosteronei* BR6020) .

CHAPTER 5: CONCLUSIONS

Proposed scenario for control of *cba* pathway

Based on available data, I propose the following model for control of the *cba* genes in *C. testosteroni* BR60 (pBRC60), summarized in Figure 5.1. CbaR, coded for by a divergently transcribed gene upstream of *cbaABC*, binds to the *cbaA* promoter (P_{cbaA}) and represses expression (Chapter 2 and 3). CbaS, coded for by a chromosomally-encoded gene unlinked to the *cba* pathway, is a positive regulator that can induce these genes in response to either 3Cba, a substrate of the pathway, or benzoate, which is not a substrate (Chapter 2). In response to 3Cba, repression by CbaR is reversed because 3Cba inhibits binding of CbaR to P_{cbaA} (Chapter 3). However, in response to benzoate, gratuitous induction of *cbaA* by CbaS is modulated by CbaR (Chapter 2), which responds minimally to this compound and thus remains bound to the *cbaA* promoter (Chapter 3).

CbaR is expressed at a constant, constitutively low level, and does not appear to repress its own expression (Chapter 2 and 3). CbaS remains unidentified and thus uncharacterized (Chapter 4). CbaS may be a positive activator involved in the regulation of the chromosomally-encoded benzoate degradative pathway. In response to 3Cba, CbaS has the added ability to activate *cbaABC*. Alternatively, CbaS could be involved in the control of another aromatic pathway and is therefore just cross-activating the benzoate degradative pathway in response to 3Cba (Chapter 2). Either way, expression of the *cba* pathway appears to require at least one other element that is not linked to the catabolic genes.

Figure 5.1 Regulation of the *cba* pathway. Plasmid-encoded *cbaABC* is activated by chromosomally-encoded CbaS in response to 3-chlorobenzoate, the substrate for the *cba* genes. CbaS can also respond to benzoate and may be involved in regulating the chromosomally-encoded benzoate degradative genes. Gratuitous induction of *cbaABC* in response to benzoate is modulated by CbaR, a repressor encoded by a divergently transcribed gene physically linked to *cbaABC*.

Evolution of the (mobilizable) *cba* pathway

The origins of *cbaABC* are unknown. As mentioned in Chapter 1, chloroaromatics occur naturally and these catabolic genes may have evolved long before the appearance of the modern chemical industry. Presumably, these genes allow some bacteria to utilize some chlorobenzoates as growth substrates, even if these compounds are not very plentiful. The *cba* genes appear to always be present in the environment because, in response to localized increases of chloroaromatics brought about by pollution, the incidence of *cbaABC* will increase from undetectable levels to a significant proportion of catabolic genotypes (Fulthorpe and Wyndham, 1991; Fulthorpe and Wyndham, 1992; Peel and Wyndham, 1999). In various independent environmental isolates screened to date, the *cba* genes appear to always be part of a transposable element (Tn5271 or Tn5271-like elements) and are often located on conjugative plasmids. The sequence of *cbaABC* does not appear to have diverged significantly (<1%), implying that even though the genes may be “old”, their horizontal transmission is recent (Di Gioia *et al.*, 1998; Peel and Wyndham, 1999). Furthermore, several Tn5271-like elements have been characterized and these show variation at the junctions between DNA containing the catabolic genes and the flanking IS1071 elements, indicating that mobilization into a transmissible element has occurred more than once. However, all Tn5271-like elements, in addition to containing *cbaABC* (the catabolic genes), appear to retain *cbaR* (a regulatory gene) and *cbaDEFGH* (genes that may mediate uptake of the substrate; [Di Gioia *et al.*, 1998; Peel and Wyndham, 1999]), implying that positive selection has acted for all of these genes.

The evidence presented in this thesis indicates that the *cba* pathway evolved such that it requires an unlinked, host-encoded regulator to fully induce expression. If this scenario is true, an interesting question would be: Why has regulation evolved this way. As discussed in Chapter 1, most mobile catabolic pathways have genes for dedicated regulators closely associated with the catabolic genes, a sensible arrangement that would presumably allow careful control of expression and permit these self-sufficient catabolic modules to be shared between bacteria (i.e. “selfish genes” that ensure their spread). As mentioned above, recent mobilization of the *cba* genes seems to have occurred in response

to environmental pollution by chlorinated aromatics (represented in schematic format in Figure 5.2). It may be that in the original host (i.e. a bacterium which has not yet been isolated but which acts as a “reservoir” for this operon), the *cba* genes are part of a larger pathway for degradation of the biphenyl-like compounds or other complex substrates that result when lignin is broken down (Kirk and Farrell, 1987). Biodegradation of these polyaromatic substrates would result in monoaromatic metabolites (e.g. substituted and unsubstituted benzoates) that feed into downstream pathways. These pathways may not be closely linked to the upstream pathway. In this case, a “master” regulator may coordinately induce both upstream and associated downstream pathways. To ensure that the pathways for monoaromatic substrates are only induced under the proper conditions (i.e. when the intended substrate is present), catabolic genes such as *cbaABC* possess closely linked genes for regulators that attenuate expression caused by the “master” regulator. Under environmental pressures that mobilize the *cba* pathway, the gene for the unlinked activator cannot “tag along”, so the mobilized *cba* pathway is forced to rely on host-encoded regulators for activation. However, the *cba* pathway has retained *cbaR* as a “fail safe” mechanism so as to modulate gratuitous induction by aromatic compounds which are not substrates of the pathway.

Figure 5.2 Mobilization of *cba* pathway. A) The *cba* genes may be part of a larger pathway for degradation of polyaromatic substrates, encoded on the chromosome (grey line) of an unidentified bacterium. A “master” activator (CbaS) can induce unlinked genes involved in degrading some of the monoaromatic metabolites produced (indicated by dashed arrows with a + sign). The *cba* genes possess a repressor coded for by a closely linked gene (*cbaR*) that ensures these genes are only active when the proper substrate is present (indicated by a dashed arrow with a – sign). Plasmids (dark circle) containing IS1071 appear to be ubiquitous and conceivably could insert close to the *cba* genes by a replicative transposition mechanism (solid arrow). The plasmid thus integrates into the chromosome and two copies of IS1071 result. B) Under environmental pressures that would favour spread of the *cba* genes, one IS1071 may transpose intramolecularly, resulting in C) mobilized *cba* genes (Tn5271-like element). *cbaS* is too far to be mobilized with the *cba* genes, but in the proper host it can still exert control in *trans*. To ensure that gratuitous induction is minimized, the mobilized catabolic genes retain *cbaR*. The intramolecular transposition event described in B occurs at some random point downstream of the *cba* genes so that a section of the adjoining DNA is also mobilized. This area may in fact be part of the benzoate degradative operon as the truncated ORFs which have been isolated to date (ORF8 and a shorter version of ORF8) show homology to aryl-CoA ligases (Di Gioia *et al.*, 1998). More detail on how IS1071 may mobilize genes is found in Wyndham *et al.* (1994).

A

B

C

LITERATURE CITED

- Alekshun, M.N., and S.B. Levy. 1997. Regulation of chromosomally mediated multiple antibiotic resistance: the *mar* regulon. *Antimicrob. Agents Chemother.* 41:2067-2075.
- Alekshun, M.N., and S.B. Levy. 1999a. Alteration of the repressor activity of MarR, the negative regulator of the *Escherichia coli marRAB* locus, by multiple chemicals in vitro. *J. Bacteriol.* 181:4669-4672.
- Alekshun, M.N., and S.B. Levy. 1999b. Characterization of MarR superrepressor mutants. *J. Bacteriol.* 181:3303-3306.
- Altschul, S.F., W. Gish, W. Miller, E.W. Myers, and D.J. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215:403-410.
- Altschul, S.F., T.L. Madden, A.A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D.J. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25:3389-402.
- Arai, H., S. Akahira, T. Ohishi, M. Maeda, and T. Kudo. 1998. Adaptation of *Comamonas testosteroni* TA441 to utilize phenol: organization and regulation of the genes involved in phenol degradation. *Microbiology.* 144:2895-2903.
- Arengi, F.L., M. Pinti, E. Galli, and P. Barbieri. 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.
- Ariza, R.R., S.P. Cohen, N. Bachhawat, S.B. Levy, and B. Dimple. 1994. Repressor mutations in the *marRAB* operon that activate oxidative stress genes and multiple antibiotic resistance in *Escherichia coli*. *J. Bacteriol.* 176:143-148.
- Asako, H., H. Nakajima, K. Kobayashi, M. Kobayashi, and R. Aono. 1997. Organic solvent tolerance and antibiotic resistance increased by overexpression of *marA* in *Escherichia coli*. *Appl. Environ. Microbiol.* 63:1428-1433.
- Ausubel, F.M., R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl. 1989. Short protocols in molecular biology : a compendium of methods from Current Protocols in Molecular Biology. Greene Publishing Associates and Wiley-Interscience, New York.
- Barnes, M.R., W.A. Duetz, and P.A. Williams. 1997. A 3-(3-hydroxyphenyl)propionic acid catabolic pathway in *Rhodococcus globerulus* PWD1: cloning and characterization of the *hpp* operon. *J. Bacteriol.* 179:6145-6153.
- Beck, C.F., and R.A. Warren. 1988. Divergent promoters, a common form of gene organization. *Microbiol. Rev.* 52:318-326.
- Botsford, J.L., and J.G. Harman. 1992. Cyclic AMP in prokaryotes. *Microbiol. Rev.* 56:100-122.

- Boulnois, G.J. 1987. Gene cloning and analysis : a laboratory guide. Blackwell Scientific Publications, Oxford ; Boston.
- Breese, K., and G. Fuchs. 1998. 4-Hydroxybenzoyl-CoA reductase (dehydroxylating) from the denitrifying bacterium *Thauera aromatica*: prosthetic groups, electron donor, and genes of a member of the molybdenum-flavin-iron-sulfur proteins. *Eur. J. Biochem.* 251:916-923.
- Brooun, A., J.J. Tomashek, and K. Lewis. 1999. Purification and ligand binding of EmrR, a regulator of a multidrug transporter. *J. Bacteriol.* 181:5131-5133.
- Buchmeier, N., S. Bossie, C.Y. Chen, F.C. Fang, D.G. Guiney, and S.J. Libby. 1997. SlyA, a transcriptional regulator of *Salmonella typhimurium*, is required for resistance to oxidative stress and is expressed in the intracellular environment of macrophages. *Infect. Immun.* 65:3725-3730.
- Byrne, A.M., and R.H. Olsen. 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.
- Calb, R., A. Davidovitch, S. Koby, H. Giladi, D. Goldenberg, H. Margalit, A. Holtel, K. Timmis, J.M. Sanchez-Romero, V. de Lorenzo, and A.B. Oppenheim. 1996. Structure and function of the *Pseudomonas putida* integration host factor. *J. Bacteriol.* 178:6319-6326.
- Chaudhry, G.R., and S. Chapalamadugu. 1991. Biodegradation of halogenated organic compounds. *Microbiol. Rev.* 55:59-79.
- Chugani, S.A., M.R. Parsek, and A.M. Chakrabarty. 1998. Transcriptional repression mediated by LysR-type regulator CatR bound at multiple binding sites. *J. Bacteriol.* 180:2367-2372.
- Coco, W.M., M.R. Parsek, and A.M. Chakrabarty. 1994. Purification of the LysR family regulator, ClcR, and its interaction with the *Pseudomonas putida* *clcABD* chlorocatechol operon promoter. *J. Bacteriol.* 176:5530-5533.
- Coco, W.M., R.K. Rothmel, S. Henikoff, and A.M. Chakrabarty. 1993. Nucleotide sequence and initial functional characterization of the *clcR* gene encoding a LysR family activator of the *clcABD* chlorocatechol operon in *Pseudomonas putida*. *J. Bacteriol.* 175:417-427.
- Cohen, S.P., H. Hächler, and S.B. Levy. 1993. Genetic and functional analysis of the multiple antibiotic resistance (*mar*) locus in *Escherichia coli*. *J. Bacteriol.* 175:1484-1492.

- Cole, S.T., R. Brosch, J. Parkhill, T. Garnier, C. Churcher, D. Harris, S.V. Gordon, K. Eiglmeier, S. Gas, C.E. Barry, 3rd, F. Tekaiia, K. Badcock, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. Davies, K. Devlin, T. Feltwell, S. Gentles, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels, A. A. Krogh, J. McLean, S. Moule, L. Murphy, S. Oliver, J. Osborne, M.A. Quail, M.A. Rajandream, J. Rogers, S. Rutter, K. Seeger, S. Skelton, S. Squares, R. Sqaers, J.E. Sulston, K. Taylor, S. Whitehead, and B.G. Barrell. 1998. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature*. 393:537-544.
- Collado-Vides, J., B. Magasanik, and J.D. Gralla. 1991. Control site location and transcriptional regulation in *Escherichia coli*. *Microbiol. Rev.* 55:371-394.
- Collier, L.S., G.L. Gaines, III, and E.L. Neidle. 1998. Regulation of benzoate degradation in *Acinetobacter* sp. strain ADP1 by BenM, a LysR-type transcriptional activator. *J. Bacteriol.* 180:2493-2501.
- Coschigano, P.W., T.S. Wehrman, and L.Y. Young. 1998. Identification and analysis of genes involved in anaerobic toluene metabolism by strain T1: putative role of a glycine free radical. *Appl. Environ. Microbiol.* 64:1650-1656.
- Coschigano, P.W., and L.Y. Young. 1997. Identification and sequence analysis of two regulatory genes involved in anaerobic toluene metabolism by strain T1. *Appl. Environ. Microbiol.* 63:652-660.
- Crothers, D.M., M.R. Gartenberg, and T.E. Shrader. 1991. DNA bending in protein-DNA complexes. *Methods Enzymol.* 208:118-46.
- Cuskey, S.M., and A.B. Sprenkle. 1988. Benzoate-dependent induction from the OP2 operator-promoter region of the TOL plasmid pWWO in the absence of known plasmid regulatory genes. *J. Bacteriol.* 170:3742-3746.
- Dalrymple, B.P., and Y. Swadling. 1997. Expression of a *Butyrivibrio fibrisolvens* E14 gene (*cinB*) encoding an enzyme with cinnamoyl ester hydrolase activity is negatively regulated by the product of an adjacent gene (*cinR*). *Microbiology* 143:1203-1210.
- de Jong, E., and J.A. Field. 1997. Sulfur tuft and turkey tail: biosynthesis and biodegradation of organohalogens by Basidiomycetes. *Annu. Rev. Microbiol.* 51:375-414.
- de Lorenzo, V., I. Cases, M. Herrero, and K.N. Timmis. 1993. Early and late responses of TOL promoters to pathway inducers: identification of postexponential promoters in *Pseudomonas putida* with *lacZ-tet* bicistronic reporters. *J. Bacteriol.* 175:6902-6907.
- de Lorenzo, V., M. Herrero, U. Jakubzik, and K.N. Timmis. 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.

- de Lorenzo, V., and J. Perez-Martin. 1996. Regulatory noise in prokaryotic promoters: how bacteria learn to respond to novel environmental signals. *Molec. Microbiol.* 19:1177-1184.
- de Lorenzo, V., and K.N. Timmis. 1994. Analysis and construction of stable phenotypes in gram-negative bacteria with Tn5- and Tn10-derived minitransposons. *Methods Enzymol.* 235:386-405.
- Di Gioia, D., M. Peel, F. Fava, and R.C. Wyndham. 1998. Structures of homologous composite transposons carrying *cbaABC* genes from Europe and North America. *Appl. Environ. Microbiol.* 64:1940-1946.
- Díaz, E., A. Ferrández, and J.L. García. 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.
- DiMarco, A.A., B. Averhoff, and L.N. Ornston. 1993. Identification of the transcriptional activator *pobR* and characterization of its role in the expression of *pobA*, the structural gene for *p*-hydroxybenzoate hydroxylase in *Acinetobacter calcoaceticus*. *J. Bacteriol.* 175:4499-4506.
- DiMarco, A.A., and L.N. Ornston. 1994. Regulation of *p*-hydroxybenzoate hydroxylase synthesis by PobR bound to an operator in *Acinetobacter calcoaceticus*. *J. Bacteriol.* 176:4277-4284.
- Dispensa, M., C.T. Thomas, M.K. Kim, J.A. Perrotta, J. Gibson, and C.S. Harwood. 1992. Anaerobic growth of *Rhodopseudomonas palustris* on 4-hydroxybenzoate is dependent on AadR, a member of the cyclic AMP receptor protein family of transcriptional regulators. *J. Bacteriol.* 174:5803-5813.
- Dixon, W.J., J.J. Hayes, J.R. Levin, M.F. Weidner, B.A. Dombroski, and T.D. Tullius. 1991. Hydroxyl radical footprinting. *Methods Enzymol.* 208:380-413.
- 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.
- 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.
- Egland, P.G., and C.S. Harwood. 1999. BadR, a new MarR family member, regulates anaerobic benzoate degradation by *Rhodopseudomonas palustris* in concert with AadR, an Fnr family member. *J. Bacteriol.* 181:2102-2109.
- Egland, P.G., D.A. Pelletier, M. Dispensa, J. Gibson, and C.S. Harwood. 1997. A cluster of bacterial genes for anaerobic benzene ring biodegradation. *Proc. Natl. Acad. Sci. USA.* 94:6484-6489.
- Ehrt, S., L.N. Ornston, and W. Hillen. 1994. RpoN (σ^{54}) is required for conversion of phenol to catechol in *Acinetobacter calcoaceticus*. *J. Bacteriol.* 176:3493-3499.

- Ehrt, S., F. Schirmer, and W. Hillen. 1995. Genetic organization, nucleotide sequence and regulation of expression of genes encoding phenol hydroxylase and catechol 1,2-dioxygenase in *Acinetobacter calcoaceticus* NCIB8250. *Mol. Microbiol.* 18:13-20.
- Elledge, S.J., and R.W. Davis. 1989. Position and density effects on repression by stationary and mobile DNA-binding proteins. *Genes Dev.* 3:185-197.
- Elledge, S.J., P. Sugiono, L. Guarente, and R.W. Davis. 1989. Genetic selection for genes encoding sequence-specific DNA-binding proteins. *Proc. Natl. Acad. Sci. USA.* 86:3689-3693.
- Elsemore, D.A., and L.N. Ornston. 1994. The *pca-pob* supraoperonic cluster of *Acinetobacter calcoaceticus* contains *quiA*, the structural gene for quinate-shikimate dehydrogenase. *J. Bacteriol.* 176:7659-7666.
- Elsemore, D.A., and L.N. Ornston. 1995. Unusual ancestry of dehydratases associated with quinate catabolism in *Acinetobacter calcoaceticus*. *J. Bacteriol.* 177:5971-5978.
- Eulberg, D., E.M. Kourbatova, L.A. Golovleva, and M. Schlömann. 1998a. Evolutionary relationship between chlorocatechol catabolic enzymes from *Rhodococcus opacus* 1CP and their counterparts in proteobacteria: sequence divergence and functional convergence. *J. Bacteriol.* 180:1082-1094.
- Eulberg, D., S. Lakner, L.A. Golovleva, and M. Schlömann. 1998b. Characterization of a protocatechuate catabolic gene cluster from *Rhodococcus opacus* 1CP: evidence for a merged enzyme with carboxymuconolactone-decarboxylating and 3-oxoadipate enol-lactone-hydrolyzing activity. *J. Bacteriol.* 180:1072-1081.
- Fernandez, S., V. Shingler, and V. De Lorenzo. 1994. Cross-regulation by XylR and DmpR activators of *Pseudomonas putida* suggests that transcriptional control of biodegradative operons evolves independently of catabolic genes. *J. Bacteriol.* 176:5052-5058.
- Ferrández, A., J.L. García, and E. Díaz. 1997. Genetic characterization and expression in heterologous hosts of the 3- (3-hydroxyphenyl)propionate catabolic pathway of *Escherichia coli* K-12. *J. Bacteriol.* 179:2573-2581.
- Ferrández, A., B. Minambres, B. García, E.R. Olivera, J.M. Luengo, J.L. García, and E. Díaz. 1998. Catabolism of phenylacetic acid in *Escherichia coli*: characterization of a new aerobic hybrid pathway. *J. Biol. Chem.* 273:25974-25986.
- Figurski, D.H., and D.R. Helinski. 1979. Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided *in trans*. *Proc. Natl. Acad. Sci. USA.* 76:1648-1652.
- Filer, K., and A.R. Harker. 1997. Identification of the inducing agent of the 2,4-dichlorophenoxyacetic acid pathway encoded by plasmid pJP4. *Appl. Environ. Microbiol.* 63:317-320.

- Fujita, Y., and Y. Miwa. 1989. Identification of an operator sequence for the *Bacillus subtilis* *gnt* operon. *J. Biol. Chem.* 264:4201-4216.
- Fukumori, F., and C.P. Saint. 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.
- Fulthorpe, R.R., and R.C. Wyndham. 1989. Survival and activity of a 3-chlorobenzoate-catabolic genotype in a natural system. *Appl. Environ. Microbiol.* 55:1584-1590.
- Fulthorpe, R.R., and R.C. Wyndham. 1991. Transfer and expression of the catabolic plasmid pBRC60 in wild bacterial recipients in a freshwater ecosystem. *Appl. Environ. Microbiol.* 57:1546-1553.
- Fulthorpe, R.R., and R.C. Wyndham. 1992. Involvement of a chlorobenzoate-catabolic transposon, Tn5271, in community adaptation to chlorobiphenyl, chloroaniline, and 2,4-dichlorophenoxyacetic acid in a freshwater ecosystem. *Appl. Environ. Microbiol.* 58:314-325.
- Furukawa, K. 1994. Molecular genetics and evolutionary relationship of PCB-degrading bacteria. *Biodegradation.* 5:289-300.
- Gallegos, M.T., R. Schleif, A. Bairoch, K. Hofmann, and J.L. Ramos. 1997. Arac/XylS family of transcriptional regulators. *Microbiol. Molec. Biol. Rev.* 61:393-410.
- Gerhardt, P., R.G.E. Murray, W.A. Wood, and N.R. Krieg. 1994. Methods for general and molecular bacteriology. American Society for Microbiology, Washington, D.C.
- Gerischer, U., and L.N. Ornston. 1995. Spontaneous mutations in *pcaH* and *-G*, structural genes for protocatechuate 3,4-dioxygenase in *Acinetobacter calcoaceticus*. *J. Bacteriol.* 177:1336-1347.
- Gerischer, U., A. Segura, and L.N. Ornston. 1998. PcaU, a transcriptional activator of genes for protocatechuate utilization in *Acinetobacter*. *J. Bacteriol.* 180:1512-1524.
- Gribble, G.W. 1992. Naturally occurring organohalogen compounds: a survey. *J. Natural Products.* 55:1353-1395.
- Gross, C.A., M. Lonetto, and R. Losick. 1992. Bacterial sigma factors. In *Transcriptional regulation*. Vol. 1. S.L. McKnight and K.R. Yamamoto, editors. Cold Spring Harbor Laboratory Press, Plainview, N.Y. 129-176.
- Haak, B., S. Fetzner, and F. Lingens. 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.
- Hagedorn, S.R., R.S. Hanson, and D.A. Kunz. 1988. Microbial metabolism and the carbon cycle : a symposium in honor of Stanley Dagley. Harwood Academic Publishers, Chur, Switzerland ; New York.

- Harayama, S., and K.N. Timmis. 1989. Catabolism of aromatic hydrocarbons by *Pseudomonas*. In *Genetics of bacterial diversity*. D.A. Hopwood and K.F. Chater, editors. Academic Press, London ; San Diego. 151-174.
- Harayama, S., and K.N. Timmis. 1992. Aerobic biodegradation of aromatic hydrocarbons by bacteria. In *Degradation of environmental pollutants by microorganisms and their metalloenzymes*. Vol. 28. H. Sigel and A. Sigel, editors. Marcel Dekker, New York. 99-156.
- Harwood, C.S., N.N. Nichols, M.K. Kim, J.L. Ditty, and R.E. Parales. 1994. Identification of the *pcaRKF* gene cluster from *Pseudomonas putida*: involvement in chemotaxis, biodegradation, and transport of 4- hydroxybenzoate. *J. Bacteriol.* 176:6479-6488.
- Harwood, C.S., and R.E. Parales. 1996. The β -ketoadipate pathway and the biology of self-identity. *Annu. Rev. Microbiol.* 50:553-590.
- Holtel, A., D. Goldenberg, H. Giladi, A.B. Oppenheim, and K.N. Timmis. 1995. Involvement of IHF protein in expression of the Ps promoter of the *Pseudomonas putida* TOL plasmid. *J. Bacteriol.* 177:3312-3315.
- Houghton, J.E., T.M. Brown, A.J. Appel, E.J. Hughes, and L.N. Ornston. 1995. Discontinuities in the evolution of *Pseudomonas putida cat* genes. *J. Bacteriol.* 177:401-412.
- Hughes, E.J., M.K. Shapiro, J.E. Houghton, and L.N. Ornston. 1988. Cloning and expression of *pca* genes from *Pseudomonas putida* in *Escherichia coli*. *J. Gen. Microbiol.* 134:2877-2887.
- Hugouvieux-Cotte-Pattat, N., G. Condemine, W. Nasser, and S. Reverchon. 1996. Regulation of pectinolysis in *Erwinia chrysanthemi*. *Annu. Rev. Microbiol.* 50:213-257.
- Jeffrey, W.H., S.M. Cuskey, P.J. Chapman, S. Resnick, and R.H. Olsen. 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.
- Jenkins, J.R., and R.A. Cooper. 1988. Molecular cloning, expression, and analysis of the genes of the homoprotocatechuate catabolic pathway of *Escherichia coli* C. *J. Bacteriol.* 170:5317-5324.
- Johnson, G.R., and R.H. Olsen. 1997. Multiple pathways for toluene degradation in *Burkholderia* sp. strain JS150. *Appl. Environ. Microbiol.* 63:4047-4052.
- Junker, F., R. Kiewitz, and A.M. Cook. 1997. Characterization of the *p*-toluenesulfonate operon *tsaMBCD* and *tsaR* in *Comamonas testosteroni* T-2. *J. Bacteriol.* 179:919-927.

- Kaphammer, B., J.J. Kukor, and R.H. Olsen. 1990. Regulation of *tfdCDEF* by *tfdR* of the 2,4-dichlorophenoxyacetic acid degradation plasmid pJP4. *J. Bacteriol.* 172:2280-2286.
- Kaphammer, B., and R.H. Olsen. 1990. Cloning and characterization of *tfdS*, the repressor-activator gene of *tfdB*, from the 2,4-dichlorophenoxyacetic acid catabolic plasmid pJP4. *J. Bacteriol.* 172:5856-5862.
- Kasak, L., R. Horak, A. Nurk, K. Talvik, and M. Kivisaar. 1993. Regulation of the catechol 1,2-dioxygenase- and phenol monooxygenase-encoding *pheBA* operon in *Pseudomonas putida* PaW85. *J. Bacteriol.* 175:8038-8042.
- Keen, N.T., S. Tamaki, D. Kobayashi, and D. Trollinger. 1988. Improved broad-host-range plasmids for DNA cloning in gram-negative bacteria. *Gene.* 70:191-197.
- Kessler, B., S. Marques, T. Kohler, J.L. Ramos, K.N. Timmis, and V. de Lorenzo. 1994. Cross talk between catabolic pathways in *Pseudomonas putida*: XylS-dependent and -independent activation of the TOL *meta* operon requires the same *cis*-acting sequences within the *P_m* promoter. *J. Bacteriol.* 176:5578-5582.
- Kirk, T.K., and R.L. Farrell. 1987. Enzymatic "combustion": the microbial degradation of lignin. *Annu. Rev. Microbiol.* 41:465-505.
- Kobayashi, M., and S. Shimizu. 1998. Metalloenzyme nitrile hydratase: structure, regulation, and application to biotechnology. *Nat. Biotechnol.* 16:733-736.
- Kolb, A., S. Busby, H. Buc, S. Garges, and S. Adhya. 1993. Transcriptional regulation by cAMP and its receptor protein. *Annu. Rev. Biochem.* 62:749-795.
- Komeda, H., M. Kobayashi, and S. Shimizu. 1996. Characterization of the gene cluster of high-molecular-mass nitrile hydratase (H-NHase) induced by its reaction product in *Rhodococcus rhodochrous* J1. *Proc. Natl. Acad. Sci. U S A.* 93:4267-4272.
- Krooneman, J., E.R.B. Moore, J.C.L. Van Velzen, R.A. Prins, L.J. Forney, and J.C. Gottschal. 1998. Competition for oxygen and 3-chlorobenzoate between two aerobic bacteria using different degradation pathways. *FEMS Microbiol. Ecol.* 26:171-179.
- Krooneman, J., E.B. Wieringa, E.R. Moore, J. Gerritse, R.A. Prins, and J.C. Gottschal. 1996. Isolation of *Alcaligenes* sp. strain L6 at low oxygen concentrations and degradation of 3-chlorobenzoate via a pathway not involving (chloro)catechols. *Appl. Environ. Microbiol.* 62:2427-2434.
- Kustu, S., A.K. North, and D.S. Weiss. 1991. Prokaryotic transcriptional enhancers and enhancer-binding proteins. *Trends Biochem. Sci.* 16:397-402.
- Kustu, S., E. Santero, J. Keener, D. Popham, and D. Weiss. 1989. Expression of σ^{54} (*ntrA*)-dependent genes is probably united by a common mechanism. *Microbiol. Rev.* 53:367-376.

- Labbé, D., J. Garnon, and P.C. Lau. 1997. Characterization of the genes encoding a receptor-like histidine kinase and a cognate response regulator from a biphenyl/polychlorobiphenyl-degrading bacterium, *Rhodococcus* sp. strain M5. *J. Bacteriol.* 179:2772-2776.
- Landick, R., and C.L. Turnbough. 1992. Transcriptional attenuation. *In* Transcriptional regulation. Vol. 1. S.L. McKnight and K.R. Yamamoto, editors. Cold Spring Harbor Laboratory Press, Plainview, N.Y. 407-446.
- Lau, P.C., Y. Wang, A. Patel, D. Labbé, H. Bergeron, R. Brousseau, Y. Konishi, and M. Rawlings. 1997. A bacterial basic region leucine zipper histidine kinase regulating toluene degradation. *Proc. Natl. Acad. Sci. U S A.* 94:1453-1458.
- Laurie, A.D., and G. Lloyd-Jones. 1999. The *phn* genes of *Burkholderia* sp. strain RP007 constitute a divergent gene cluster for polycyclic aromatic hydrocarbon catabolism. *J. Bacteriol.* 181:531-540.
- Leveau, J.H., and J.R. van der Meer. 1996. The *tfdR* gene product can successfully take over the role of the insertion element-inactivated TfdT protein as a transcriptional activator of the *tfdCDEF* gene cluster, which encodes chlorocatechol degradation in *Ralstonia eutropha* JMP134(pJP4). *J. Bacteriol.* 178:6824-6832.
- Leveau, J.H.J., W.M. De Vos, and J.R. Van der Meer. 1994. Analysis of the binding site of the LysR-type transcriptional activator TcbR on the *tcbR* and *tcbC* divergent promoter sequences. *J. Bacteriol.* 176:1850-1856.
- Lewin, B. 1987. Genes. Wiley, New York
- Libby, S.J., W. Goebel, A. Ludwig, N. Buchmeier, F. Bowe, F.C. Fang, D.G. Guiney, J.G. Songer, and F. Heffron. 1994. A cytolysin encoded by *Salmonella* is required for survival within macrophages. *Proc. Natl. Acad. Sci. U S A.* 91:489-493.
- Lipscomb, J.D., and A.M. Orville. 1992. Mechanistic aspects of dihydroxybenzoate dioxygenases. *In* Degradation of environmental pollutants by microorganisms and their metalloenzymes. Vol. 28. H. Sigel and A. Sigel, editors. Marcel Dekker, New York. 243-298.
- Lodish, H.F., and J.E. Darnell. 1995. Molecular cell biology. Scientific American Books : Distributed by W.H. Freeman and Co., New York.
- Loewen, P.C., and R. Hengge-Aronis. 1994. The role of the sigma factor σ^S (KatF) in bacterial global regulation. *Annu. Rev. Microbiol.* 48:53-80.
- Lomovskaya, O., K. Lewis, and A. Matin. 1995. EmrR is a negative regulator of the *Escherichia coli* multidrug resistance pump EmrAB. *J. Bacteriol.* 177:2328-2334.
- Lonetto, M., M. Gribskov, and C.A. Gross. 1992. The σ^{70} family: sequence conservation and evolutionary relationships. *J. Bacteriol.* 174:3843-9.
- Maloy, S., and V. Stewart. 1993. Autogenous regulation of gene expression. *J. Bacteriol.* 175:307-316.

- Maneewannakul, K., and S.B. Levy. 1996. Identification of *mar* mutants among quinolone-resistant clinical isolates of *Escherichia coli*. *Antimicrob. Agents Chemother.* 40:1695-1698.
- Marklund, B.I., J.M. Tennent, E. García, A. Hamers, M. Baga, F. Lindberg, W. Gaastra, and S. Normark. 1992. Horizontal gene transfer of the *Escherichia coli* *pap* and *prs* pili operons as a mechanism for the development of tissue-specific adhesive properties. *Mol. Microbiol.* 16:2225-2242.
- Marques, S., M.T. Gallegos, and J.L. Ramos. 1995. Role of σ^S in transcription from the positively controlled Pm promoter of the TOL plasmid of *Pseudomonas putida*. *Mol. Microbiol.* 18:851-857.
- Marques, S., and J.L. Ramos. 1993. Transcriptional control of the *Pseudomonas putida* TOL plasmid catabolic pathways. *Mol. Microbiol.* 9:923-929.
- Martin, R.G., K.W. Jair, R.E. Wolf, Jr., and J.L. Rosner. 1996. Autoactivation of the *marRAB* multiple antibiotic resistance operon by the MarA transcriptional activator in *Escherichia coli*. *J. Bacteriol.* 178:2216-2223.
- Martin, R.G., and J.L. Rosner. 1995. Binding of purified multiple antibiotic-resistance repressor protein (MarR) to *mar* operator sequences. *Proc. Natl. Acad. Sci. U S A.* 92:5456-5460.
- Matrubutham, U., and A.R. Harker. 1994. Analysis of duplicated gene sequences associated with *tfdR* and *tfdS* in *Alcaligenes eutrophus* JMP134. *J. Bacteriol.* 176:2348-2353.
- McFall, S.M., T.J. Klem, N. Fujita, A. Ishihama, and A.M. Chakrabarty. 1997a. DNase I footprinting, DNA bending and in vitro transcription analyses of ClcR and CatR interactions with the *clcABD* promoter: evidence of a conserved transcriptional activation mechanism. *Mol Microbiol.* 24:965-76.
- McFall, S.M., M.R. Parsek, and A.M. Chakrabarty. 1997b. 2-chloromuconate and ClcR-mediated activation of the *clcABD* operon: in vitro transcriptional and DNase I footprint analyses. *J. Bacteriol.* 179:3655-3663.
- Morett, E., and L. Segovia. 1993. The σ^{54} bacterial enhancer-binding protein family: mechanism of action and phylogenetic relationship of their functional domains. *J. Bacteriol.* 175:6067-6074.
- Nakatsu, C., J. Ng, R. Singh, N. Straus, and C. Wyndham. 1991. Chlorobenzoate catabolic transposon Tn5271 is a composite class I element with flanking class II insertion sequences. *Proc. Natl. Acad. Sci. USA.* 88:8312-8316.
- Nakatsu, C.H. 1992. The biochemical pathway and catabolic genes for chlorobenzoate degradation on the transposon Tn5271 of *Alcaligenes* sp. strain BR60. In Ottawa-Carleton Institute of Biology. Carleton University, Ottawa, Ontario, Canada, Ottawa, Ontario, Canada. 218.

- Nakatsu, C.H., R.R. Fulthorpe, B.A. Holland, M.C. Peel, and R.C. Wyndham. 1995. The phylogenetic distribution of a transposable dioxygenase from the Niagara River watershed. *Molec. Ecol.* 4:593-603.
- Nakatsu, C.H., M. Providenti, and R.C. Wyndham. 1997. The *cis*-diol dehydrogenase *chaC* gene of Tn5271 is required for growth on 3-chlorobenzoate but not 3,4-dichlorobenzoate. *Gene*. 196:209-218.
- Nakatsu, C.H., and R.C. Wyndham. 1993. Cloning and expression of the transposable chlorobenzoate-3,4-dioxygenase genes of *Alcaligenes* sp. strain BR60. *Appl. Environ. Microbiol.* 59:3625-3633.
- Ng, J., and R.C. Wyndham. 1993. IS1071-mediated recombinational equilibrium in *Alcaligenes* sp. BR60 carrying the 3-chlorobenzoate catabolic transposon Tn5271. *Can. J. Microbiol.* 39:92-100.
- Ng, L.C., C.L. Poh, and V. Shingler. 1995. Aromatic effector activation of the NtrC-like transcriptional regulator PhhR limits the catabolic potential of the (methyl)phenol degradative pathway it controls. *J. Bacteriol.* 177:1485-1490.
- Nichols, N.N., and C.S. Harwood. 1995. Repression of 4-hydroxybenzoate transport and degradation by benzoate: a new layer of regulatory control in the *Pseudomonas putida* β -ketoadipate pathway. *J. Bacteriol.* 177:7033-7040.
- Oethinger, M., I. Podglajen, W.V. Kern, and S.B. Levy. 1998. Overexpression of the *marA* or *soxS* regulatory gene in clinical topoisomerase mutants of *Escherichia coli*. *Antimicrob. Agents Chemother.* 42:2089-2094.
- Oppenberg, H.R.S., G. Chen, T. Leisinger, and A.M. Cook. 1995. Regulation of the degradative pathways from 4-toluenesulphonate and 4-toluenecarboxylate to protocatechuate in *Comamonas testosteroni* T-2. *Microbiology*. 141:1891-1899.
- Oscarsson, J., Y. Mizunoe, B.E. Uhlin, and D.J. Haydon. 1996. Induction of haemolytic activity in *Escherichia coli* by the *slyA* gene product. *Mol. Microbiol.* 20:191-199.
- Parales, R.E., and C.S. Harwood. 1993. Regulation of the *pcaIJ* genes for aromatic acid degradation in *Pseudomonas putida*. *J. Bacteriol.* 175:5829-5838.
- Parche, S., W. Geissdorfer, and W. Hillen. 1997. Identification and characterization of *xcpR* encoding a subunit of the general secretory pathway necessary for dodecane degradation in *Acinetobacter calcoaceticus* ADP1. *J. Bacteriol.* 179:4631-4634.
- Parke, D. 1992. Application of *p*-toluidine in chromogenic detection of catechol and protocatechuate, diphenolic intermediates in catabolism of aromatic compounds. *Appl. Environ. Microbiol.* 58:2694-2697.
- Parke, D. 1993. Positive regulation of phenolic catabolism in *Agrobacterium tumefaciens* by the *pcaQ* gene in response to β -carboxy-*cis*, *cis*-muconate. *J. Bacteriol.* 175:3529-3535.
- Parke, D. 1995. Supraoperonic clustering of *pca* genes for catabolism of the phenolic compound protocatechuate in *Agrobacterium tumefaciens*. *J. Bacteriol.* 177:3808-3817.

- Parke, D. 1996. Characterization of *pcaQ*, a LysR-type transcriptional activator required for catabolism of phenolic compounds, from *Agrobacterium tumefaciens*. *J. Bacteriol.* 178:266-272.
- Parkinson, J.S., and E.C. Kofoid. 1992. Communication modules in bacterial signaling proteins. *Annu. Rev. Genet.* 26:71-112.
- Parsek, M.R., M. Kivisaar, and A.M. Chakrabarty. 1995. Differential DNA bending introduced by the *Pseudomonas putida* LysR-type regulator, CatR, at the plasmid-borne *pheBA* and chromosomal *catBC* promoters. *Mol. Microbiology.* 15:819-828.
- Parsek, M.R., S.M. McFall, D.L. Shinabarger, and A.M. Chakrabarty. 1994. Interaction of two LysR-type regulatory proteins CatR and ClcR with heterologous promoters: functional and evolutionary implications. *Proc. Natl. Acad. Sci. U S A.* 91:12393-12397.
- Parsek, M.R., D.L. Shinabarger, R.K. Rothmel, and A.M. Chakrabarty. 1992. Roles of CatR and *cis*, *cis*-muconate in activation of the *catBC* operon, which is involved in benzoate degradation in *Pseudomonas putida*. *J. Bacteriol.* 174:7798-7806.
- Peel, M.C., and R.C. Wyndham. 1997. The impact of industrial contamination on microbial chlorobenzoate degradation in the Niagara watershed. *Microb. Ecol.* 33:59-68.
- Peel, M.C., and R.C. Wyndham. 1999. Selection of *clc*, *cba*, and *fcB* chlorobenzoate-catabolic genotypes from ground-water and surface waters adjacent to the Hyde Park, Niagara Falls, chemical landfill. *Appl. Environ. Microbiol.* 65:1627-1635.
- Perez-Martin, J., and V. de Lorenzo. 1995. The σ^{54} -dependent promoter Ps of the TOL plasmid of *Pseudomonas putida* requires HU for transcriptional activation *in vivo* by XylR. *J. Bacteriol.* 177:3758-3763.
- Poole, K., K. Tetro, Q. Zhao, S. Neshat, D.E. Heinrichs, and N. Bianco. 1996. Expression of the multidrug resistance operon *mexA-mexB-oprM* in *Pseudomonas aeruginosa*: *mexR* encodes a regulator of operon expression. *Antimicrob. Agents Chemother.* 40:2021-2028.
- Praillet, T., W. Nasser, J. Robert-Baudouy, and S. Reverchon. 1996. Purification and functional characterization of PecS, a regulator of virulence-factor synthesis in *Erwinia chrysanthemi*. *Mol. Microbiol.* 20:391-402.
- Prieto, M.A., E. Díaz, and J.L. García. 1996. Molecular characterization of the 4-hydroxyphenylacetate catabolic pathway of *Escherichia coli* W: engineering a mobile aromatic degradative cluster. *J. Bacteriol.* 178:111-120.
- Providenti, M.A., S. MacSween, B. Harrison, C.H. Nakatsu, and R.C. Wyndham. 1999a. Identification and characterization of *cbaDEFGH*, which may code for proteins involved in the uptake of 3-chlorobenzoate by *Comamonas testosteroni* BR60. *Unpublished results*.

- Providenti, M.A., J. Mampel, S. MacSween, A.M. Cook, and R.C. Wyndham. 1999b. Protocatechuate *meta* ring fission catabolic pathway (*pmd*) of *Comamonas testosteroni* is required for growth on various substituted benzoates but not benzoate. *Unpublished results*.
- Ramos, J.L., S. Marques, and K.N. Timmis. 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.
- Reineke, W. 1998. Development of hybrid strains for the mineralization of chloroaromatics by patchwork assembly. *Annu. Rev. Microbiol.* 52:287-331.
- Reineke, W., and H.J. Knackmuss. 1988. Microbial degradation of haloaromatics. *Annu. Rev. Microbiol.* 42:263-287.
- Reverchon, S., W. Nasser, and J. Robert-Baudouy. 1994. *pecS*: a locus controlling pectinase, cellulase and blue pigment production in *Erwinia chrysanthemi*. *Mol. Microbiol.* 11:1127-1139.
- Rojo, F. 1999. Repression of transcription initiation in bacteria. *J. Bacteriol.* 181:2987-2991.
- Romero-Arroyo, C.E., M.A. Schell, G.L. Gaines, 3rd, and E.L. Neidle. 1995. *catM* encodes a LysR-type transcriptional activator regulating catechol degradation in *Acinetobacter calcoaceticus*. *J. Bacteriol.* 177:5891-5898.
- Romero-Steiner, S., R.E. Parales, C.S. Harwood, and J.E. Houghton. 1994. Characterization of the *pcaR* regulatory gene from *Pseudomonas putida*, which is required for the complete degradation of *p*-hydroxybenzoate. *J. Bacteriol.* 176:5771-5779.
- Romine, M.F., L.C. Stillwell, K.K. Wong, S.J. Thurston, E.C. Sisk, C. Sensen, T. Gaasterland, J.K. Fredrickson, and J.D. Saffer. 1999. Complete sequence of a 184-kilobase catabolic plasmid from *Sphingomonas aromaticivorans* F199. *J. Bacteriol.* 181:1585-1602.
- Roper, D.I., T. Fawcett, and R.A. Cooper. 1993. The *Escherichia coli* C homoprotocatechuate degradative operon: *hpc* gene order, direction of transcription and control of expression. *Mol. Gen. Genet.* 237:241-250.
- Rost, B., and C. Sander. 1993. Prediction of protein secondary structure at better than 70% accuracy. *J. Mol. Biol.* 232:584-599.
- Rost, B., and C. Sander. 1994. Combining evolutionary information and neural networks to predict protein secondary structure. *Proteins.* 19:55-72.
- Rost, B., C. Sander, and R. Schneider. 1994. PHD: an automatic mail server for protein secondary structure prediction. *Comput. Appl. Biosci.* 10:53-60.
- Rothmel, R.K., T.L. Aldrich, J.E. Houghton, W.M. Coco, L.N. Ornston, and A.M. Chakrabarty. 1990. Nucleotide sequencing and characterization of *Pseudomonas putida catR*: a positive regulator of the *catBC* operon is a member of the LysR family. *J. Bacteriol.* 172:922-931.

- Rothmel, R.K., A.M. Chakrabarty, A. Berry, and A. Darzins. 1991a. Genetic systems in *Pseudomonas*. *Methods Enzymol.* 204:485-514.
- Rothmel, R.K., D.L. Shinabarger, M.R. Parsek, T.L. Aldrich, and A.M. Chakrabarty. 1991b. Functional analysis of the *Pseudomonas putida* regulatory protein CatR: transcriptional studies and determination of the CatR DNA-binding site by hydroxyl-radical footprinting. *J. Bacteriol.* 173:4717-4724.
- Rouanet, C., K. Nomura, S. Tsuyumu, and W. Nasser. 1999. Regulation of *pelD* and *pelE*, encoding major alkaline pectate lyases in *Erwinia chrysanthemi*: involvement of the main transcriptional factors. *J. Bacteriol.* 181:5948-5957.
- Sambrook, J., E.F. Fritsch, and T. Maniatis. 1989. Molecular cloning : a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. 3 v. pp.
- Sander, C., and R. Schneider. 1991. Database of homology-derived protein structures and the structural meaning of sequence alignment. *Proteins.* 9:56-68.
- Schell, M.A. 1993. Molecular biology of the LysR family of transcriptional regulators. *Annu. Rev. Microbiol.* 47:597-626.
- Schirmer, F., S. Ehrt, and W. Hillen. 1997. Expression, inducer spectrum, domain structure, and function of MopR, the regulator of phenol degradation in *Acinetobacter calcoaceticus* NCIB8250. *J. Bacteriol.* 179:1329-1336.
- Seoane, A.S., and S.B. Levy. 1995. Characterization of MarR, the repressor of the multiple antibiotic resistance (*mar*) operon in *Escherichia coli*. *J. Bacteriol.* 177:3414-3419.
- Shingler, V., M. Bartilson, and T. Moore. 1993. Cloning and nucleotide sequence of the gene encoding the positive regulator (DmpR) of the phenol catabolic pathway encoded by pVI150 and identification of DmpR as a member of the NtrC family of transcriptional activators. *J. Bacteriol.* 175:1596-1604.
- Shingler, V., and T. Moore. 1994. Sensing of aromatic compounds by the DmpR transcriptional activator of phenol-catabolizing *Pseudomonas* sp. strain CF600. *J. Bacteriol.* 176:1555-1560.
- Simons, R.W., and N. Kleckner. 1988. Biological regulation by antisense RNA in prokaryotes. *Annu. Rev. Genet.* 22:567-600.
- Singh, H., J.H. LeBowitz, A.S. Baldwin, Jr., and P.A. Sharp. 1988. Molecular cloning of an enhancer binding protein: isolation by screening of an expression library with a recognition site DNA. *Cell.* 52:415-423.
- Stanier, R.Y., and L.N. Ornston. 1973. The β -ketoadipate pathway. *Adv. Microb. Physiol.* 9:89-151.
- Steinbüchel, A., N. Krüger, H. Valentin, A. Timm, A. Pries, E. Hustede, and H.G. Schlegel. 1992. Physiological and genetic analysis of polyhydroxyalkanoate biosynthetic pathways. In *Pseudomonas: Molecular biology and biotechnology*. E. Galli, S. Silver, and B. Witholt, editors. American Society for Microbiology, Washington, D.C., U.S.A. 315-327.

- Stock, J.B., A.J. Ninfa, and A.M. Stock. 1989. Protein phosphorylation and regulation of adaptive responses in bacteria. *Microbiol. Rev.* 53:450-490.
- Sulavik, M.C., L.F. Gambino, and P.F. Miller. 1995. The MarR repressor of the multiple antibiotic resistance (*mar*) operon in *Escherichia coli*: prototypic member of a family of bacterial regulatory proteins involved in sensing phenolic compounds. *Mol. Med.* 1:436-446.
- Tamaoka, J., D.M. Ha, and K. Komagata. 1987. Reclassification of *Pseudomonas acidovorans* den Dooren de Jong 1926 and *Pseudomonas testosteroni* Marcus and Talalay 1956 as *Comamonas acidovorans* comb. nov. and *Comamonas testosteroni* comb. nov., with an emended description of the genus *Comamonas*. *International Journal of Systematic Bacteriology*. 37:52-59.
- Tokito, M.K., and F. Daldal. 1992. *petR*, located upstream of the *fbCFBC* operon encoding the cytochrome *bc₁* complex, is homologous to bacterial response regulators and necessary for photosynthetic and respiratory growth of *Rhodobacter capsulatus*. *Mol. Microbiol.* 6:1645-1654.
- Tsoi, T.V., E.G. Plotnikova, J.R. Cole, W.F. Guerin, M. Bagdasarian, and J.M. Tiedje. 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.
- van der Meer, J.R., A.C.J. Frijters, J.H.J. Leveau, R.I.L. Eggen, A.J.B. Zehnder, and W.M. de Vos. 1991a. Characterization of the *Pseudomonas* sp. strain P51 gene *tcbR*, a LysR-type transcriptional activator of the *tcbCDEF* chlorocatechol oxidative operon, and analysis of the regulatory region. *J. Bacteriol.* 173:3700-3708.
- van der Meer, J.R., A.R.W. van Neerven, E.J. de Vries, W.M. de Vos, and A.J.B. Zehnder. 1991b. 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.
- van Pee, K.H. 1996. Biosynthesis of halogenated metabolites by bacteria. *Annu. Rev. Microbiol.* 50:375-399.
- Wang, Y., M. Rawlings, D.T. Gibson, D. Labbé, H. Bergeron, R. Brousseau, and P.C. Lau. 1995. Identification of a membrane protein and a truncated LysR-type regulator associated with the toluene degradation pathway in *Pseudomonas putida* F1. *Mol. Gen. Genet.* 246:570-579.
- Wheelis, M.L., N.J. Palleroni, and R.Y. Stanier. 1967. The metabolism of aromatic acids by *Pseudomonas testosteroni* and *P. acidovorans*. *Arch. Microbiol.* 59:302-314.
- Wilmes-Riesenberg, M.R., and B.L. Wanner. 1992. *TnphoA* and *TnphoA'* elements for making and switching fusions for study of transcription, translation, and cell surface localization. *J. Bacteriol.* 174:4558-4575.

- Wu, S., H. de Lencastre, and A. Tomasz. 1998. Genetic organization of the *mecA* region in methicillin-susceptible and methicillin-resistant strains of *Staphylococcus sciuri*. *J. Bacteriol.* 180:236-242.
- Wyndham, R.C. 1986. Evolved aniline catabolism in *Acinetobacter calcoaceticus* during continuous culture of river water. *Appl. Environ. Microbiol.* 51:781-789.
- Wyndham, R.C., C. Nakatsu, M. Peel, A. Cashore, J. Ng, and F. Szilagyi. 1994. Distribution of the catabolic transposon Tn5271 in a groundwater bioremediation system. *Appl. Environ. Microbiol.* 60:86-93.
- Wyndham, R.C., R.K. Singh, and N.A. Straus. 1988. Catabolic instability, plasmid gene deletion and recombination in *Alcaligenes* sp. BR60. *Arch. Microbiol.* 150:237-243.
- Wyndham, R.C., and N.A. Straus. 1988. Chlorobenzoate catabolism and interactions between *Alcaligenes* and *Pseudomonas* species from Bloody Run Creek. *Arch. Microbiol.* 150:230-236.
- Yanisch-Perron, C., J. Vieira, and J. Messing. 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene.* 33:103-119.
- You, I.S., and D. Ghosal. 1995. Genetic and molecular analysis of a regulatory region of the herbicide 2,4-dichlorophenoxyacetate catabolic plasmid pJP4. *Mol. Microbiol.* 16:321-331.
- Zylstra, G.J., W.R. McCombie, D.T. Gibson, and B.A. Finette. 1988. Toluene degradation by *Pseudomonas putida* F1: genetic organization of the *tod* operon. *Appl. Environ. Microbiol.* 54:1498-1503.

APPENDICES

Appendix A. Sequence of DNA coding for 16S rRNA from strain BR60 (pBRC60) and comparison to other sequences in the database. Work was performed at MIDI Labs (Newark, Delaware, USA) using primers that correspond to positions 5 and 531 of the *E. coli* 16S rRNA gene.

A) Sequence of PCR-generated product

```
TGGAGAGTTTGATCCTGGCTCAGATTGAACGCTGGCGGCATGCTTTACACATGCAAGTCGAACGGTAACAGG
TCTTCGGATGCTGACGAGTGGCGAACGGGTGAGTAATACATCGGAACGTGCCTAGTAGTGGGGGATAACTAC
TCGAAAGAGTAGCTAATACCGCATGAGATCTACGGATGAAAGCAGGGGACCTTCGGGCCTTGTGCTACTAGA
GCGGCTGATGGCAGATTAGGTAGTTGGTGGGGTAAAGGCTTACCAAGCCTGCGATCTGTAGCTGGTCTGAGA
GGACGACCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATTTTGGAC
AATGGGCGAAAGCCTGATCCAGCAATGCCGCGTGCAGGATGAAGGCCCTCGGGTTGTAAACTGCTTTTGTAC
GGAACGAAAAGCCTGGGGCTAATATCCCCGGGTCATGACGGTACCGTAAGAATAAGCACCGGCTAACTACGT
GCCAGCAGCCGCGGTA
```

B) Top ten matches (in % genetic distance) and UPGMA (unweighted paired group method with arithmetic mean) phylogenetic tree showing distance of strain BR60 to top ten matches.

```
Alignment: 520 C450 BR60 con
0.00 % 1528 Comamonas testosteroni
6.16 % 1524 Acidovorax konjaci
6.35 % 1528 Comamonas acidovorans
6.73 % 1524 Acidovorax avenae citrulli
6.92 % 1524 Acidovorax avenae cattleyae
6.93 % 1524 Acidovorax temperans
7.12 % 1524 Acidovorax avenae avenae
7.34 % 1528 Acidovorax delafieldii
7.91 % 1528 Acidovorax facilis
8.30 % 1526 Pseudomonas? spinosa
```

UPGMA Tree

Appendix B. Schematic of *cbaA-lacZ* constructs cloned into suicide delivery vector pUTCm. Indicated on the expression constructs are relevant restriction sites, in particular those at the ends (which are often part of the multiple cloning site of the pUC18-based vector from which they were removed). All constructs were cloned such that *lacZ* (not to scale) is transcribed in the opposite direction of *cat*. Transposition of DNA between the I and O end is mediated by *tnp**. The restriction map for pUTCm is based on de Lorenzo and Timmis (1994) and not all restriction sites are necessarily indicated. Loci on pUTCm are roughly to scale. See Chapter 2 (Material and Methods) for additional details.

Appendix C. Schematic diagram of *cbaR-lacZ* constructs and control *lacZ* construct cloned into suicide-delivery vector pUTCm or pUTCmMCS.3, which is identical to pUTCm but with a modified multiple cloning site. Indicated on the expression constructs are relevant restriction sites, in particular those at the ends (which are often part of the multiple cloning site of the pUC18-based vector from which they were removed). All constructs were cloned such that *lacZ* (not to scale) is transcribed in the opposite direction of *cat*. Transposition of DNA between the I and O end is mediated by *tmp**. The restriction map for pUTCm is based on de Lorenzo and Timmis (1994) and not all restriction sites may necessarily be indicated. Loci on pUTCm are roughly to scale. Unique restriction sites on pUTCm or pUTCmMCS.3 are indicated in bold. See Chapter 2 (Material and Methods) for additional details.

Appendix D. DNA sequence of junctions between '*lacZ*' and A) '*cbaA*' and B) '*cbaR*'; and C) the junction between the DNA coding for the six-histidine tag and the N-terminus of CbaR.

A) '*cbaA*' junction with '*lacZ*'. Shown is DNA upstream of '*cbaA*' until nucleotide 4564. Also indicated is the first half of the *ScaI* site (underlined), the second half of the *SrfI* site (bold), and the *BamHI* site at the extreme N-terminus of '*lacZ*'. Also indicated is the location of primer LACZ65, used to generate deletions in the '*cbaA*' promoter by PCR. See Chapter 2 (Materials and Methods) for additional details.

```

4564
CCCCTCATAA GGGTTTGACA CAAGTTGCTA AACCAACAAC ATCCGCAAAA TTTCAGGAGC AAGCT

ATG GCA TTG ACC AAG GAC AAC GAA GAT TTG GTG CGA GTG GGC CGT GGA ACC CCG
cbaA' →
ATG GGC GGC CTC ATG CGC GAG TGG GCG GAT CCC GTC GTT TTA CAA CGT CGT GAC
          (ScaI|SrfI) BamHI
                        'lacZ →
TGG GAA AAC CCT GGC GTT ACC CAA CTT AAT CGC CTT GCA GCA CAT CCC CCT TTC
      < --- --- --- --- --- ---
                LACZ65

```

B) '*cbaR*' junction with '*lacZ*'. Shown is DNA upstream of '*cbaR*' until nucleotide 4020. Also indicated is the first half of the *NruI* site (underlined), the second half of the *SrfI* site (bold), and the *BamHI* site at the extreme N-terminus of '*lacZ*'. See Chapter 2 (Materials and Methods) for additional details.

```

4020
TATGGCTTTT ACCGCCCCCA CCGAGTCTGG GGCCTACACG CATCAACGCT GGATACTGA

ATG CTT GCA AGA GAT CCT CGA AAA AGC CGG GTG ACC CTC AAT CTC GAT ACC TAC
cbaR' →
ACC CCG GCA TTT TTT ACC GCC ATC GCC AAC AAG CTC GGG GCG GAT CCC GTC GTT
          (NruI|SrfI) BamHI
                        'lacZ →

```

C) '*cbaR*' junction with DNA coding for N-terminal histidine tag of pQE30. Shown is the ribosome binding site (RBS) upstream of the start codon, the DNA coding for the histidine tag, and DNA at the N-terminus of '*cbaR*'. Also indicated is the *BamHI* site and the conceptual translation. See Chapter 3 (Materials and Methods) for additional details.

```

| RBS |
AGGAGAAATT AACT ATG AGA GGA TCG CAT CAC CAT CAC CAT CAC GGA TCC CTT GCA AGA
          M  R  G  S  H  H  H  H  H  H  G  S  L  A  R
          | →                histidine tag                ← | cbaR →
GAT CCT CGA AAA AGC CGG GTG ACC CTC AAT CTC GAT ACC TAC ACC CCG GCA TTT TTT
D  P  R  K  S  R  V  T  L  N  L  D  T  Y  T  P  A  F  F

```