Catechol 1,2-dioxygenase from *Rhodococcus rhodochrous* NCIMB 13259

Philip Strachan

Thesis submitted for the degree of Doctor of Philosophy

Division of Biochemistry and Molecular Biology
University of Glasgow
January 1997

To Mum and Dad

CONTENTS

Page

Page

LIST OF FIGURES

Page

Chapter 4

Chapter 5

Page

LIST OF TABLES

Page

ABREVIATIONS

The abbreviations recommended by the Biochemical Journal in its instructions to authors [*Biochem. J.* (1996) **313**, 1-15] have been used with the following additions:

NCIMB	National collections of Industrial and Marine Bacteria
psi	pounds per square inch
rpm	revolutions per minuite
TFA	trifluoroacetic acid
TEMED	*N,N,N',N'*-tetramethylenediamine
PEG	polyethyleneglycol
V'	apparent maximum velocity
K_m'	apparent michaelis constant
DTT	dithiothreitol
RBS	ribosome binding site
ε	molar absorption coefficient
A	absorbance
M_r	relative molecular mass
DTNB	dithionitrobenzoic acid

ACKNOWLEDGEMENTS

I wish to express my thanks to my supervisor, Professor Charles Fewson for his guidance and support throughout this project. I would also like to extend my gratitude to Dr Les Fixter, Dr Ian Hamilton, Dr D Gillooly and Mr Alan Scott for their encouragement and guidance.

My thanks also to Dr Jeff Keen (SERC sequencing unit, Department of Biochemistry and Molecular Biology, University of Leeds) for protein and peptide sequencing, Mr Graham Hamilton (Division of Molecular Genetics, University of Glasgow) for DNA sequencing, Dr Tino Krell (Division of Biochemistry and Molecular Biology, University of Glasgow) for electrospray mass spectroscopy and Mr Mike Beglan (Department of Chemistry, University of Glasgow) for atomic absorption spectroscopy.

SUMMARY

1. The Gram-positive bacterium *Rhodococcus rhodochrous* NCIMB13259 is able to grow on styrene, which it assimilates via a *meta*-cleavage pathway. This bacterium also contains a catechol 1,2-dioxygenase that carries out intradiol-cleavage of aromatic substrates.

2. A rapid and reliable method for the purification of catechol 1,2-dioxygenase was developed which involved chromatography on DEAE-Sephacel and phenyl-Sepharose. This procedure yielded at least 10 mg of protein from 15-20 g (wet weight) of cells. SDS-PAGE analysis of the final preparation showed that catechol 1,2-dioxygenase was virtually homogeneous. The enzyme was stable and could be stored at −20 °C with little loss in activity.

3. The enzyme was crystallised using 15% (w/v) polyethylene glycol 4 000, 0.33 M $CaCl_2$, 0.25 mM Tris/HCl, pH 7.5, using sitting drop vapour diffusion. A microseeding technique was then used to obtain single, diffraction quality rectangular plates, which reached a size of 0.5 mm down the longest axis. A preliminary X-ray crystallographic characterisation showed that the crystals are in space group C2 with unit cell dimensions (Å) a=111.9, b=78.1, c=134.6, β=100 °, with four subunits in the asymmetric unit.

4. A subunit M_r value of 39 800 was obtained with SDS-PAGE, which is approximately 20% more than the value of 31 558 which was obtained with electrospray mass spectroscopy. Similar overestimates by SDS-PAGE have also been obtained for other catechol 1,2-dioxygenases. The native M_r value of 120 000 obtained by gel-permeation chromatography indicated that catechol 1,2-dioxygenase is a tetramer, however, catechol 1,2-dioxygenase which had been chemically cross-linked, showed only two bands upon analysis with PAGE, indicating that the enzyme is a dimer, with a calculated native M_r value of 63 000. Considerable evidence accrued by other workers with several non-haem intradiol-cleavage enzymes, indicates that the overestimate could be due to the catechol 1,2-dioxygenase being more elliptical than the protein standards, causing it to move anomolously through gels.

5. The absorbance spectrum of catechol 1,2-dioxygenase had maxima at 220 and 280 nm. At higher concentrations of the enzyme another maximum at 426 nm was also evident. This is good evidence that the *Rhodococcus* enzyme is iron dependent, since in several non-haem iron intradiol-cleavage enzymes a broad absorption band at about 440 nm is due to an Fe^{3+} cofactor.

6. Iron quantitation using colorimetry and atomic absorption spectroscopy found significantly less than the expected one mole of iron per subunit (0.60-0.85 atoms per subunit).

7. The activity and stability of catechol 1,2-dioxygenase as a function of pH were investigated. Enzyme activity increased approximately four-fold from pH 5 and reached an optimum value at approximately pH 9. Incubation of the enzyme for 30 min in solutions which ranged from pH 5 to 9, resulted in no significant loss of activity. At pH 11.2 however, the enzyme was rapidly inactivated.

8. The apparent K_m and V values were determined with catechol, 3-methylcatechol, 4-methylcatechol and protocatechuate as the substrates. Catechol 1,2-dioxygenase oxygenated the first three substrates, but not protocatechuate. The results for the three catechols were broadly similar: the enzyme had an apparent K_m of 1-2 μM and apparent V values of 13-19 units (mg protein)$^{-1}$. Relative activity with different substrates showed that the *Rhodococcus* enzyme resembles the chlorocatechol 1,2-dioxygenases rather than the catechol 1,2-dioxygenases.

9. The first 24 amino acids from the *N*-terminal sequence of *Rhodococcus* catechol 1,2-dioxygenase showed substantial sequence identity with catechol 1,2-dioxygenase from *Arthrobacter*, another Gram-positive bacterium.

10. An internal sequence of 15 residues, obtained from a peptide which had been purified from a V8 protease digest of catechol 1,2-dioxygenase, showed 67% identity to a section of the *Arthrobacter* catechol 1,2-dioxygenase.

11. Three degenerate pools of oligonucleotides were designed on the basis of chemically determined amino acid sequence data and from amino acid sequence alignments with other catechol 1,2-dioxygenases and were synthesized.

12. An homologous oligonucleotide probe was made using the three pools of oligonucleotide primers in a hemi-nested PCR reaction.

13. The radiolabelled probe hybridised to a 3.5 kb genomic *Xma*I restriction fragment, which was then cloned and sequenced. A single open reading frame was detected, encoding a protein with a M_r value of 31 559, which is within one mass

unit of the value for catechol 1,2-dioxygenase obtained from electrospray mass spectroscopy.

14. Catechol 1,2-dioxygenase was expressed constitutively in *Escherichia coli* XL2-Blue pPDS to approximately the same levels as the wild-type expression in *Rhodococcus*, suggesting that its expression is under negative control. The enzyme was also overexpressed from the T7 promoter, to approximately 40% of the total cellular protein

15. A putative ribosome binding site (AGGAGG) was found 4 bp upstream from the initiation codon (ATG). In addition, putative –35 and –10 promoter sequences (TTGACA and CACAGT respectively) and two inverted repeat structures were discovered upstream of the coding region. The inverted repeats bear a strong resemblance to LysR binding motifs, which are normally involved in the positive regulation of bacterial catabolic operons.

16. Chemical analysis found 1.65 sulfhdryl groups per subunit. Since the derived amino acid sequence contains 2 Cys residues per subunit, it is assumed that there are no disulphide bonds in the structure of *Rhodococcus* catechol 1,2-dioxygenase.

17. An amino acid alignment of the *Rhodococcus* catechol 1,2-dioxygenase with other non-haem iron intradiol-cleavage enzymes showed that the *Rhodococcus* enzyme had 26-28% identity to protocatechuate 3,4-dioxygenases, 29-30% identity to the catechol 1,2-dioxygenase and 30-35% identity to the chlorocatechol 1,2-dioxygenases. The closest relationship was with the catechol 1,2-dioxygenase from the Gram-positive bacterium, *Arthrobacter*, which was 56% identical (71% similarity).

18. The *Rhodococcus* enzyme has the two His and two Tyr residues which are conserved in all the other non-haem iron intradiol-cleavage enzymes and which have been shown to chelate the active site iron. In addition, the retention of Pro and Gly in β-loops and at the end of putative secondary structural features, suggests that the *Rhodococcus* enzyme has the same fold as the other enzymes. Amino acid substitutions in the active site and the substantial differences both at the amino and carboxy termini, suggest a different substrate binding site architecture, in addition to a different mechanism for subunit association.

CHAPTER 1
Introduction

1.1 Introduction: enzyme oxygenation

Dioxygenases were discovered by Hayaishi *et al.* (1955) who showed that during the oxidation of catechol by catechol 1,2-dioxygenase, both atoms of molecular oxygen were incorporated into product. In the same year Mason *et al.* (1955) discovered a different type of oxygenase which incorporated only one atom of oxygen into its organic substrate and so this latter type of enzyme became known as a monooxygenase, although it is also referred to as a 'mixed function oxidase' or 'hydroxylase'. Both types of enzyme (Table 1.1a, b) are widely distributed in Nature and have important roles, such as in the biosynthesis, transformation and degradation of amino acids, lipids and vitamins, as well as the metabolism of xenobiotic compounds (Nozaki, 1979; Harayama *et al.*, 1992).

Reactions of oxygen with organic compounds at physiological temperatures are extremely slow, primarily because of the unusual electronic structure of oxygen which in the ground state has two unpaired valence electrons, in contrast to most organic compounds which have all their electrons paired. The direct reaction of a spin-paired, singlet molecule, with a triplet, diradical molecule to give singlet products is a spin-forbidden process. In order to overcome the kinetic inertia, both monooxygenases and dioxygenases have cofactors such as flavins, or transition metals such as copper or iron (Table 1.1a, b) (Hamilton, 1974; Malmström, 1982). Reduced flavins overcome the spin-restriction because they can form stable radical semiquinones capable of reacting with oxygen to form reactive peroxide intermediates, which are able to attack bound organic substrates. Another common reaction type involves complexing oxygen to transition metals in low oxidation states. The spin restriction is removed since transition metals have incompletely filled *d*-orbitals which can transfer electrons onto the oxygen to form spin-paired intermediates which can subsequently react with organic substrates (Harayama *et al.*, 1992). A third type of oxygen activation does not involve the metal prosthetic group directly. In these reactions the organic substrate is complexed to transition metals in high oxidation states which then react directly with oxygen bound at the active site (Malmström, 1982).

1.2 The role of oxygenases in bacterial aromatic catabolism

1.2.1 Dihydroxylated aromatic intermediates

Myriad aromatic compounds may be degraded aerobically by common soil bacteria such as *Pseudomonas* (Reineke & Knackmuss, 1988), *Acinetobacter* (Fewson, 1991) and *Rhodococcus* (Warhurst, 1993; Warhurst & Fewson, 1994). Many by-products of the chemicals industry are benzene derivatives and may be potentially harmful when released into the biosphere and so there has been considerable energy spent in constructing microorganisms which can degrade

Cofactors	Monooxygenases
FAD	Lysine monooxygenase
FAD	4-Hydroxylbenzoate hydroxylase
Cu	Dopamine β-hydroxylase
Biopterine and non-haem iron	Phenylalanine hydroxylase
Cytochrome P-450 (haemoprotein)	Liver microsomal drug hydroxylase
Non-haem iron	*Pseudomonas* fatty acid ω-hydroxylase

Cofactors	Dioxygenases
Fe^{2+}	Catechol 2,3-dioxygenase
Fe^{3+}	Catechol 1,2-dioxygenase
Haem	Tryptophan 2,3-dioxygenase
Cu	Quercitinase
Fe^{2+}	Lipoxygenase
Non-haem iron, [2F-2S] NAD(P)H, FAD	Benzoate 1,2- dioxygenase

Table 1.1 Some examples of oxygenases and their cofactors
(Nozaki, 1979; Harayama *et al.*, 1992)

them (Timmis *et al.*, 1994; Chakrabarty, 1996).

The biodegradation of aromatic compounds follows a common theme, in which aromatic substrates are first converted via one or two oxygenation reactions, into common dihydroxylated intermediates such as catechol and protocatechuic acid, which are broken down by subsequent reactions into primary metabolites (Figure 1.1). Introduction of single hydroxyl groups, monohydroxylation, is generally carried out by monooxygenases, while simultaneous introduction of two hydroxyl groups, is catalysed by dioxygenases (Figure 1.2). The majority of monooxygenases are typically single component flavoproteins, although multicomponent monooxygenases have also been found (Table 1.1a) (Harayama *et al.*, 1992). In contrast, dioxygenases such as benzoate 1,2-dioxygenase (Table 1.1b), which carry out the initial oxygenation of aromatic substrates, have two or more components: an hydroxylase component and an electron-transport component. All hydroxylase components contain an iron-sulphur centre and mononuclear non-haem iron. The iron-sulphur centre does not itself activate oxygen, but instead functions as a redox centre; receiving electrons from the electron-transport component and then reducing the non-haem iron to Fe^{2+} (Harayama *et al.*, 1992).

1.2.2 Non-haem iron aromatic ring-cleavage dioxygenases

Intradiol and extradiol ring-cleavage enzymes are commonly found at the beginning of the β-ketoadipate (Bruce *et al.*, 1989; Pettigrew *et al.*, 1991) and *meta*-pathways (Harayama *et al.*, 1989; Kukor & Olsen, 1991; Schreiner *et al.*, 1991) respectively. These multi-step pathways convert derivatives of catechol or protocatechuic acid, into primary metabolites such as succinate, pyruvate and acetyl-CoA. With one or two exceptions, the vast majority of the aromatic ring-cleavage enzymes utilise non-haem iron as the sole cofactor (Nozaki, 1979). Two basic modes of ring-cleavage are displayed by microbial enzymes (Figure 1.2): intradiol- or *ortho*-cleavage occurs when the carbon-carbon bond between adjacent hydroxyl groups is broken, giving rise to *cis,cis*-muconic acid derivatives; extradiol- or *meta*-cleavage occurs when the carbon-carbon bond adjacent to just one hydroxyl group is cleaved, giving rise to muconate semialdehydes (Nozaki, 1979). Dioxygenases catalysing intradiol-cleavage reactions such as catechol 1,2-dioxygenase, chlorocatechol 1,2-dioxygenase and protocatechuate 3,4-dioxygenase are red and contain ferric iron as the sole cofactor, whereas extradiol-cleaving enzymes such as catechol 2,3-dioxygenase and protocatechuate 4,5-dioxygenase contain iron in the ferrous oxidation state and are colourless (Nozaki, 1979).

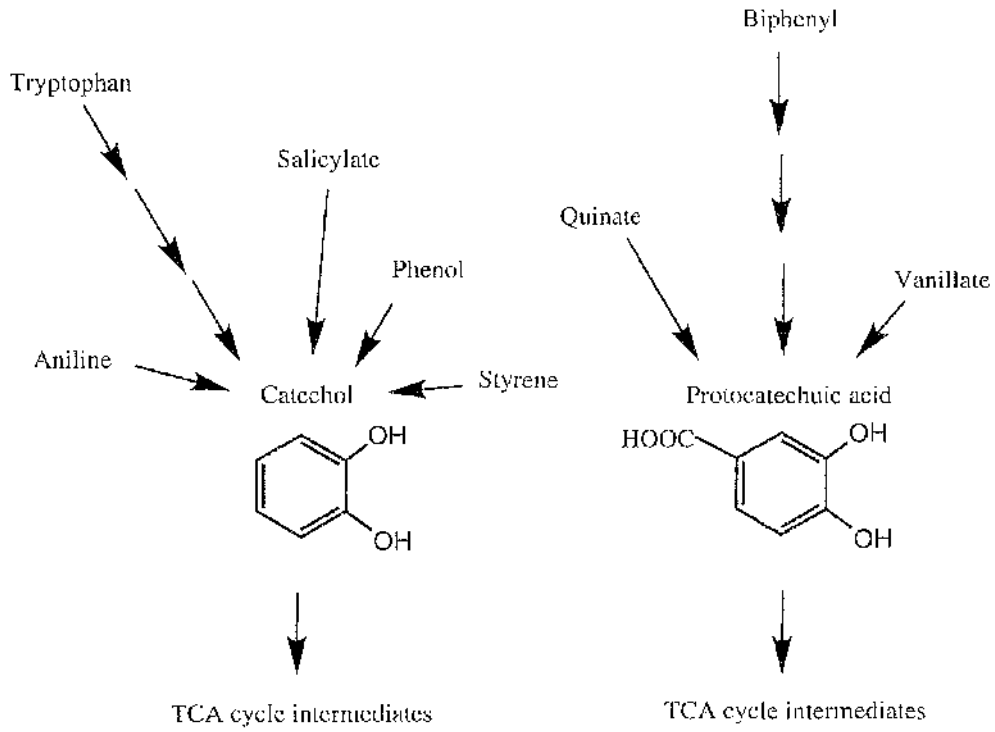

Figure 1.1 The metabolism of aromatic compounds via common dihydroxylated intermediates such as catechol and protocatechuic acid

Myriad aromatic compounds are converted into common dihydroxylated derivatives such as catechol and protocatechuic acid and then degraded by subsequent reactions into tricarboxylic acid cycle intermediates and other primary metabolites.

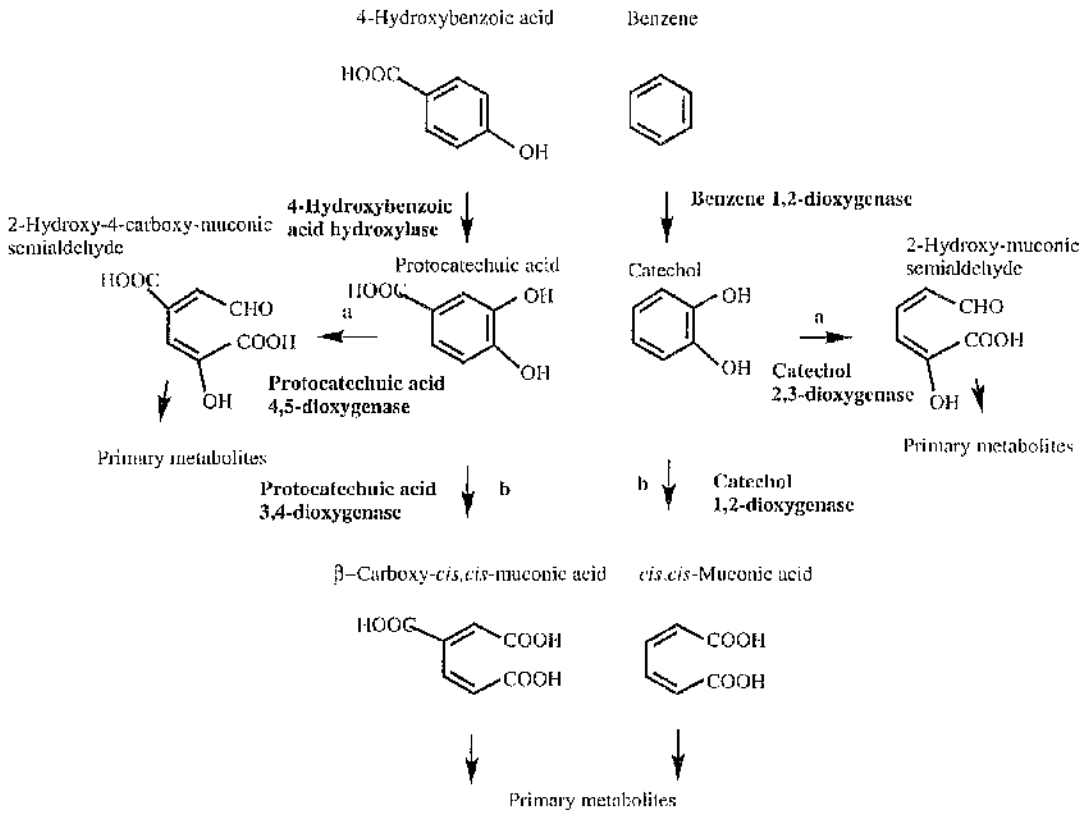

Figure 1.2 The role of oxygenases in bacterial aromatic catabolism

Aromatic compounds such as 4-hydroxybenzoic acid and benzene are first converted into dihydroxylated intermediates such as protocatechuic acid and catechol by the action of 4-hydroxybenzoic acid hydroxylase and benzene 1,2-dioxygenase. Activated compounds may then undergo extradiol ring-cleavage (a) by the action of catechol 2,3-dioxygenase or protocatechuate 4,5-dioxygenase, or intradiol ring-cleavage (b) by the action of catechol 1,2-dioxygenase or protocatechuate 3,4-dioxygenase (Nozaki, 1979; Harayama *et al.*, 1992).

1.3 General molecular characteristics of the non-haem iron intradiol-cleavage enzymes

Non-haem iron intradiol-cleavage enzymes have been isolated from a variety of microorganisms, including *Acinetobacter calcoaceticus* (Patel *et al.*, 1976) (Table 1.2), *Brevibacterium fuscum* (Whittaker *et al.*, 1984), *Moraxella* (Sterjiades & Pelmont, 1989), *Rhizobium leguminosarum* biovar *viceae* USDA 2370 (Chen & Lovell, 1990), *Pseudomonas putida* (Broderick & O'Halloran, 1991) (Table 1.2), *Pseudomonas acidovorans* (Hinteregger *et al.*, 1992), *Rhodococcus erythropolis* ICP (Maltseva *et al.*, 1994) and *Alcaligenes eutrophus* (Ghosal & You, 1988; Sauret-Ignaz *et al.*, 1996) (Table 1.2). The native M_r values of catechol 1,2-dioxygenases are generally in the 60 000-80 000 range and with one or two possible exceptions (Aoki *et al.*, 1984*a*, *b*), are dimers of either identical or non-identical subunits (Nakai *et al.*, 1990) (Table 1.2). Subunit M_r values derived from DNA sequence data are generally 30 000-34 000 (Neidle *et al.*, 1988; Nakai *et al.*, 1990) (Table 1.2). Protocatechuate 3,4-dioxygenases are larger and tend to contain equal numbers of non-identical subunits with the form $(\alpha\beta)_n$ (n = 3-12) (Nozaki, 1979), although the subunits are considerably smaller and have M_r values of 20 000-30 000 (Table 1.2). Chlorocatechol 1,2-dioxygenases have been purified from several Gram-negative strains (Dorn & Knackmuss, 1978*a*, *b*; Pieper *et al.*, 1988; Broderick & O'Halloran, 1991; Bhat *et al.*, 1993; Hinteregger *et al.*, 1992; Miguez *et al.*, 1992) and are all homodimers. Some of the genes encoding the chlorocatechol 1,2-dioxygenases have been cloned and encode subunits with M_r values of 28 000-30 000 (Frantz & Chakrabarty, 1987; van der Meer *et al.*, 1991*a*), which is somewhat smaller than the catechol 1,2-dioxygenases (Table 1.2). Iron-subunit stoichiometries vary: *A. calcoaceticus* catechol 1,2-dioxygenase contains two moles of iron per mole of enzyme, whereas the chlorocatechol 1,2-dioxygenase contains only one mole of iron per mole of enzyme (Table 1.2). The substrate specificities of the catechol 1,2-dioxygenases, chlorocatechol 1,2-dioxygenases and the protocatechuate 3,4-dioxygenases are mutually exclusive: the first two enzymes will oxygenate catechol and some of its derivatives but will not oxygenate protocatechuate, whereas protocatechuate 3,4-dioxygenase uses protocatechuate, but not catechol (Table 1.2). Another important distinguishing characteristic is in the product ratios when 3-methylcatechol or 3-methoxycatechol is the substrate: ratios of intradiol- to extradiol-cleavage for the catechol 1,2-dioxygenases from *Pseudomonas arvilla* C-1 and *Pseudomonas* B 13 with 3-methylcatechol and 3-methoxycatechol are approximately 17:1 and 5:1 (Fujiwara *et al.*, 1975; Dorn & Knackmuss, 1978*a*), moreover as the rate of intradiol-cleavage of 3-substituted

Organism	*Pseudomonas putida* pA27	*Acinetobacter calcoaceticus* BD413	*Pseudomonas arvilla* C-1	*Alcaligenes eutrophus* JMP134	*Pseudomonas putida* ATCC 23975 (formerly *P. aeruginosa*)
Enzyme	CCD	CTD	CTD	CCD	PCD
Enzyme activity[k]					
Catechol	100[a]	100[g]	100[j]	100[a]	0.4
3-Methylcatechol	354	12	8	167	0.4
4-Methylcatechol	273	18	90	-	0.2
Protocatechuate	0.0	0.0	0.0	-	100[j]
3-Chlorocatechol	1.31	-	3.6	124	-
4-Chlorocatechol	0.80	-	-	122	-
pI value	5.8[a]	-	-	-	5.2, 9.5[h]
pH optimum	7.5-7.9[a]	7-9[g]	7.5[d]	-	8[i]
$\varepsilon_{[\lambda max\ (nm)]}$ ($M^{-1}\ cm^{-1}$)	3 095[a] (based on Fe concentration)	2 860[g] (based on protein concentration)	-	-	-
λmax	430[a]	440[g]	440[d]	-	440[j]
M_r of subunits	28 988[b]	34 351[c]	α=30 000[d] β=34 187[i]	28 281[e]	α=22 500[f], β=25 000
Amino acid residues	260[b]	311[h]	β=310[i]	255[c]	200, 238[f]
Subunit composition	α_2Fe[a]	$(\alpha Fe)_2$[g]	ααFe, αβFe, ββFe[d]	-	$(\alpha\beta Fe)_{12}$[h]

Table 1.2 General molecular characteristics of some non-haem iron intradiol-cleavage enzymes

Abreviations: CTD, catechol 1,2-dioxygenase; CCD, chlorocatechol 1,2-dioxygenase; PCD, protocatechuate 3,4-dioxygenase.

a, Broderick & O'Halloran (1991); b, Frantz & Chakrabarty (1987); c Neidle *et al.* (1988); d, Nakai *et al.* (1990); e, Ghosal & You (1988); f, Frazee *et al.* (1993); g, Patel *et al.* (1976); h, Ohlendorf *et al.* (1994); i, Nakai *et al.* (1995); j, Nozaki (1979); k, relative to catechol for CCD and CTD and protocatechuate for PCD.

- = not recorded.

catechols becomes slower relative to the reaction with catechol, so the proportion of extradiol activity increases. This point is illustrated with the catechol 1,2-dioxygenase from *Frateuria* ANA-18, which catalyses the intradiol cleavage of 3-methylcatechol at a rate of less than 0.9% to that of catechol, but the ratio of intradiol to extradiol activity is 1.0:1.2 (Aoki *et al.*, 1984*b*). Chlorocatechol 1,2-dioxygenases by contrast, oxygenate 3-substituted catechols better than catechol and extradiol activity is typically less than 0.5% (Broderick & O'Halloran, 1991). The absorption spectra of all non-haem iron intradiol-cleavage enzymes are very similar to each other and concentrated solutions of the enzyme are distinctly red in colour, with a broad absorption between 390 and 690 nm, with a peak at about 440 nm (Nozaki, 1979) (Table 1.2). When the apoenzyme is prepared by anaerobic dialysis against a buffer solution containing both *o*-phenanthroline and sodium dithionite, both colour and activity disappear; reappearing after exposure of the solution to oxygen (Kojima *et al.*, 1967; Fujisawa *et al.*, 1972). Resonance Raman experiments have complemented these studies, since laser excitation within the visible absorption band results in a spectrum which is characteristic of Fe^{3+}-coordinated tyrosine ligands indicating that the red colour arises from tyrosine-Fe^{3+} charge transfer interactions (Tatsuno *et al.*, 1978; Que & Heistand, 1979).

1.4 Structure of protocatechuate 3,4-dioxygenase

The protocatechuate 3,4-dioxygenase from *P. putida* ATCC 23975 (formerly classified as *P. aeruginosa*) (Table 1.2), is the only non-haem iron intradiol-cleavage enzyme whose three-dimensional structure has been solved (Ohlendorf *et al.*, 1988; Ohlendorf *et al.*, 1994). Twelve αβ protomers form a rough truncated tetrahedron in which nearly all the protomeric contacts are provided by the β-subunits (Figure 1.3). These form a hollow inner shell bounding a central cavity of about 50 Å in diameter with the α-subunits clustered around the apices (Figure 1.3). The active site of the enzyme is located at the interface between the α- and β- subunits.

Within the active site, the ferric iron is coordinated by two His residues (one equatorial and one axial), two Tyr residues (one equatorial and one axial) and a hydroxyl group (equatorial) in a trigonal bipyramid arrangement (Figure 1.4). All of the coordinating ligands are derived from the β-subunit. The α-subunit does not bind iron and only two of the residues corresponding to the iron ligands on the β-subunit have been conserved (Ohlendorf *et al.*, 1988; Ohlendorf *et al.*, 1994).

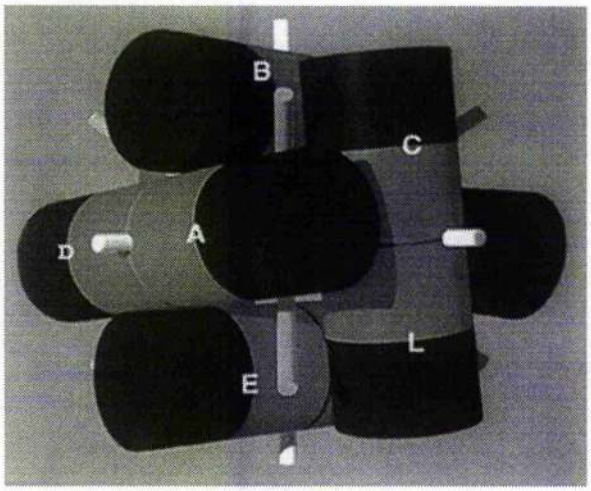

Figure 1.3 The packing of protocatechuate 3,4-dioxygenase protomers

Each protomer (A to L) is represented as a cylinder 60 Å long and 40 Å in diameter. The folded α (β) chains are shown in dark grey (light grey). The white rods indicate the three perpendicular 2-fold axes and the grey rods indicate the four 2-fold axes. The overall shape of the aggregate is a tetrahedron with three α-chains at each apex. The figure was reproduced from Ohlendorf *et al.* (1994).

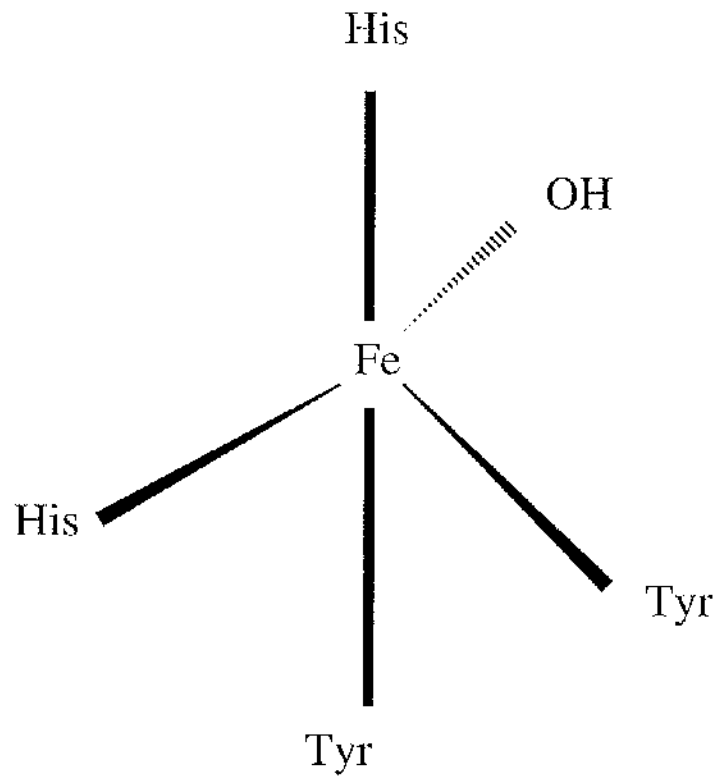

Figure 1.4 Trigonal bipyramid coordination of ferric iron within the active site of protocatechuate 3,4-dioxygenase

The ferric ion is coordinated to two His residues (one axial and one equatorial), two Tyr residues (one axial and one equatorial) and an equatorial hydroxyl ion (Ohlendorf *et al.*, 1994).

1.5 Enzyme-substrate complexes

Spectra of the non-haem iron intradiol-cleavage enzymes can be perturbed by the anaerobic addition of substrates, or by substrate and product analogues (Que *et al.*, 1980; Bull & Ballou, 1981; Que & Epstein, 1981; Walsh & Ballou, 1983). For both protocatechuate 3,4-dioxygenase and catechol 1,2-dioxygenase, this results in a steely grey solution, due to an additional, broad absorption band, centred at approximately 700 nm (Fujisawa *et al.*, 1972; Walsh & Ballou, 1983; Walsh *et al.*, 1983; Whittaker *et al.*, 1984). Resonance Raman spectroscopy has shown that this new feature is due to the substrate binding to the metal centre and that the endogenous tyrosine-Fe^{3+} interaction is retained (Que, 1980). The resonance Raman spectra of the enzyme-substrate complexes of catechol 1,2-dioxygenase and chlorocatechol 1,2-dioxygenase are very similar (Que & Heistand, 1979). In the former case, comparison of the enzyme-substrate complex with model complex spectra led to the conclusion that it is the ionised, dianion form of catechol which is bound to the active site, which by inference suggests that chlorocatechol 1,2-dioxygenase also binds its substrate in the dianion form. The similarity between the resonance Raman spectra indicates a similar mode of substrate binding in the two enzymes. Several differences are observed in comparison to the resonance Raman spectrum of the protocatechuate 3,4-dioxygenase-substrate complex. In addition, the chlorocatechol 1,2-dioxygenase and catechol 1,2-dioxygenase complexes with the inhibitor 4-nitrocatechol show similar visible spectra that are nearly identical with the visible spectrum of 4-nitrocatechol chelated to Fe^{3+} (Tyson, 1975). These enzyme-4-nitrocatechol spectra are distinct from the visible spectrum of the protocatechuate 3,4-dioxygenase-4-nitrocatechol complex, which resembles the free 4-nitrocatechol dianion. These results suggest that 4-nitrocatechol chelates the active site in the case of catechol- and chlorocatechol 1,2-dioxygenase but not in the case of protocatechuate 3,4-dioxygenase and so it appears that the iron centre in catechol 1,2-dioxygenase interacts with substrates and inhibitors in a way which is similar to that of chlorocatechol 1,2-dioxygenase, but distinct to that of protocatechuate 3,4-dioxygenase (Tyson, 1975). More recent evidence gleaned from nuclear magnetic resonance spectroscopy and electron spin resonance spectroscopy has shown that substrates are monodentate in the catechol 1,2-dioxygenase enzyme-substrate complex and chelated in the protocatechuate 3,4-dioxygenase enzyme-substrate complex (Que *et al.*, 1987; Orville & Lipscomb, 1989). It seems likely that the first substrate hydroxyl group coordinates to the iron by displacement of a hydroxyl group bound to the active site (Figure 1.5). It is not known for sure whether the second substrate hydroxyl is accommodated by expansion of the iron coordination number from five to six, or by displacement of one of the protein

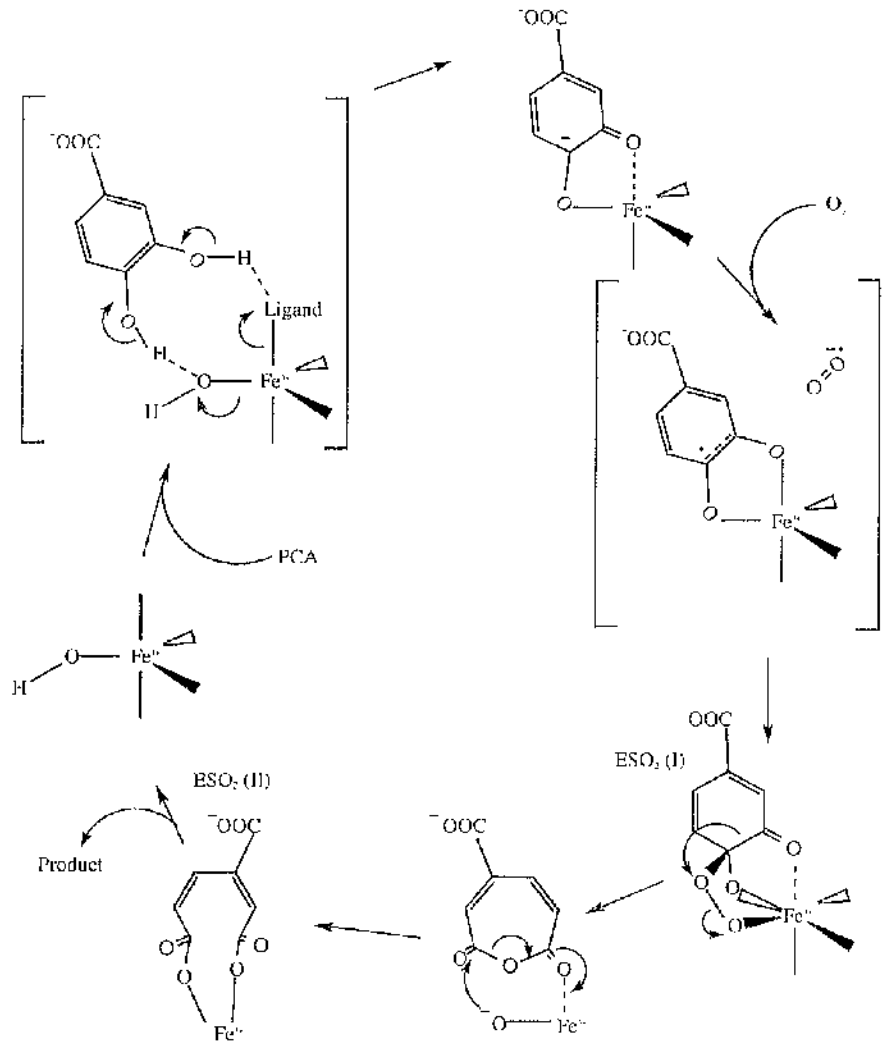

Figure 1.5 The intradiol-cleavage mechanism of protocatechuate 3,4-dioxygenase

Substrates are activated by the formation of a carbanion at C-4 which then reacts with oxygen to form the first oxygenated intermediate [ESO_2 (I)]. This intermediate then undergoes an acyl migration to form the second intermediate [ESO_2 (II)] and final product. The figure was reproduced from Ohlendorf *et al*, (1994).

PCA, protocatechuate.

ligands to the iron, but the available evidence indicates that the second hydroxyl group displaces one of the His ligands, losing a proton to it in the process (Figure 1.5) (True *et al.*, 1990).

1.6 Mechanistic implications

None of the methods reported, including circular dichroism, has revealed any perturbation of the native enzyme when it is mixed with buffer saturated with oxygen (Harayama *et al.*, 1992), but experiments done with protocatechuate 3,4-dioxygenase have shown that when enzyme-substrate complexes are mixed with buffer solutions containing oxygen, the absorbance at 470 nm decreases and then increases again as the substrate is converted into product. With the decline at 470 nm there is also a concomitant increase at approximately 520 nm (Fujisawa *et al.*, 1972). The new spectral species is distinct from those of the enzyme or enzyme-substrate species and is observed only in the presence of both oxygen and substrate, suggesting that the new species is an oxygenated intermediate (ESO_2). Rapid mixing techniques conducted with both catechol 1,2-dioxygenase and protocatechuate 3,4-dioxygenase have identified two sequential intermediates [ESO_2 (I) and ESO_2 (II)] during the breakdown of substrates, both of which exhibit some visible absorption due to the tyrosine-Fe^{3+} interaction (Bull *et al.*, 1981; Walsh *et al.*, 1983). In fact, at no stage during the reaction has a Fe^{2+} state ever been observed, in contrast to many haem enzymes (Ahktar & Wright 1991). Another feature of the intradiol-cleavage reaction which must be accounted for, is the fact that both atoms of molecular oxygen end up on carbons that were adjacent to one another in the original substrate (Hayaishi *et al.*, 1957). In the light of these observations, it has been suggested that instead of Fe^{2+} activating oxygen, Fe^{3+} promotes ketonisation at C-3 of the substrate, to form a carbanion at C-4 (Figure 1.5) (Que, 1980), which may then react with oxygen in two steps: in the first step, a superoxide radical and a resonance stabilised, semiquinone radical are produced (Malmström, 1982); both radicals or doublets then recombine, generating an α-hydroperoxy ketone, a compound which has long been mooted as a possible ESO_2 intermediate (Hamilton, 1974; Que *et al.*, 1977; Fraser & Hamilton, 1982) (Figure 1.5). It has been suggested that this intermediate then cyclises, forming a four-membered dioxetane ring; however, bond energy calculations have shown this would be extremely endothermic due to the considerable amount of bond strain involved (Hamilton, 1974). The decomposition of such compounds is also extremely exothermic and in model systems is accompanied by the emission of blue light. This scenario seems unlikely in dioxygenases, given the excessive stabilisation of intermediates required by the enzyme and the fact that enzymatic intradiol-cleavage reactions are not chemiluminescent. A different type of

mechanism and one that also has parallels in organic chemistry (Criegee, 1948), was proposed by Hamilton (1974). It is based on the observation that hydroperoxy ketones readily undergo acid-catalysed acyl migration. Such reactions are also accelerated by electron-donating substituents, which help to stabilise the incipient carbonium ion (Sawaki & Ogata, 1978). This satisfies other observations which were made during the rapid-reaction studies: that electron donating groups increased the rates of formation and decay of ESO_2 (I) complexes (Figure 1.5), whereas electron withdrawing groups such as fluorine had the opposite effect (Walsh & Ballou, 1983; Walsh *et al.*, 1983). The second intermediate [ESO_2 (II)] which was observed (Bull *et al.*, 1981; Walsh *et al.*, 1983), was interpreted as an enzyme-product complex (Figure 1.5) based on its resemblance to benzoate complexes of catechol 1,2-dioxygenase and enzymes product complexes of protocatechuate 3,4-dioxygenase (Que *et al.*, 1980; Que & Epstein, 1981). Product dissociation was shown to be rate-determining for both protocatechuate 3,4-dioxygenase and catechol 1,2-dioxygenase (Bull *et al.*, 1981; Walsh *et al.*, 1983).

1.7 Regulation of the β-ketoadipate pathway

The β-ketoadipate pathway, which is initiated by intradiol-cleavage, consists of two parallel branches for the catabolism of catechol and protocatechuate (*cat* and *pca* genes respectively) derived from benzoate and 4-hydroxybenzoate (Harwood & Parales, 1996). The *cat* branch involves a number of enzymes, including catechol 1,2-dioxygenase, muconate lactonising enzyme, muconolactone isomerase and hydrolase, encoded by *catA, catB, catC* and *catD* genes, respectively (Figure 1.6). In *P. putida.* and *A. calcoaceticus*, the *cat* genes are positively regulated by CatR and CatM respectively, which switch on gene expression when the inducer, *cis,cis*-muconate, an intermediate of the β-ketoadipate pathway, is present (Parsek *et al.*, 1992; Romero-Arroyo *et al.*, 1995) (Figure 1.6). *catA*, however, is expressed separately and is located several kbp away from the remainder of the pathway genes (Kukor *et al.*, 1988; Neidle *et al.*, 1988) and several studies suggest that in *A. calcoaceticus*, there exists another regulator, or other regulators, distinct from CatM, which is or are able to activate *catA* expression in response to both *cis,cis*-muconate and benzoate (Neidle *et al.*, 1987; Romero-Arroyo *et al.*, 1995). *catR* and *catM* are divergently transcribed from the operons which they regulate (Figure 1.7). The intergenic region, between *catR* and *catBC* in *P. putida*, has been particularly well studied and contains an activator binding site and a repressor binding site which contain inverted repeat sequence elements with TN$_{11}$A and GN$_{11}$A binding motifs respectively (Parsek *et al.*, 1994). CatR binds to the repressor binding site, both in the presence and

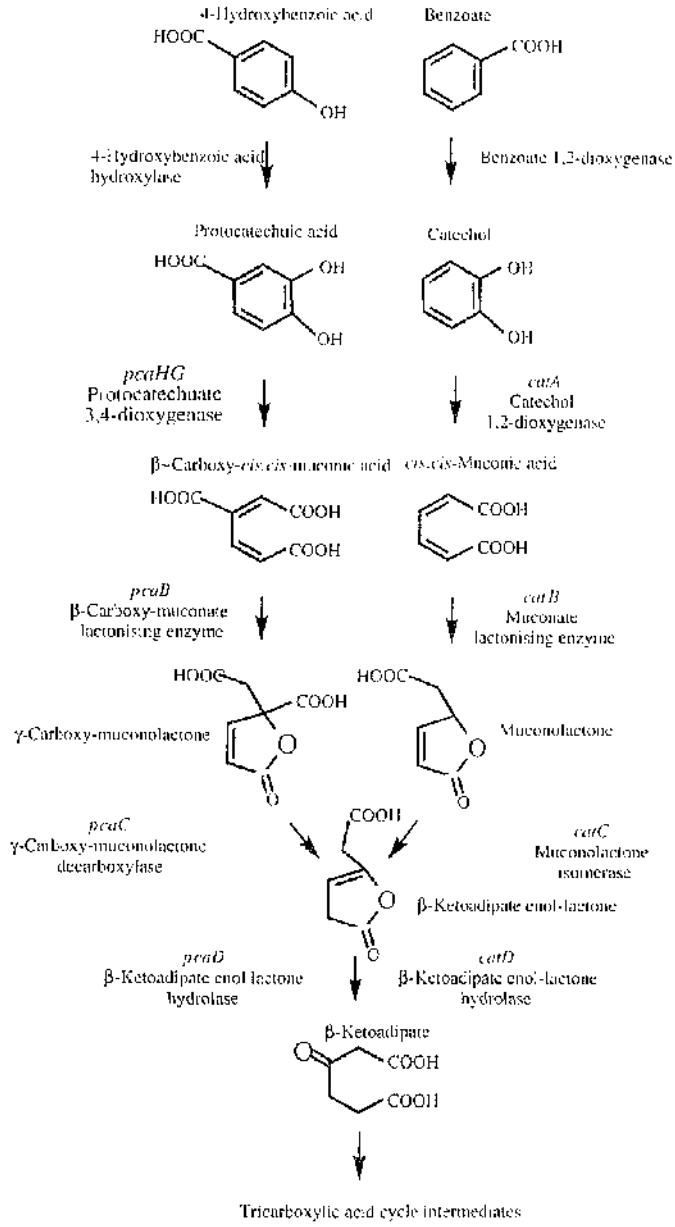

Figure 1.6 The two branches of the β-ketoadipate pathway

The β-ketoadipate pathway in *Pseudomonas putida* consists of two parallel branches, for the catabolism of catechol and protocatechuic acid (Harwood & Parales, 1996). Also shown are the two metabolic precursors for the pathway, 4-hydroxybenzoic acid and benzoate. The metabolic intermediate β-ketoadipate serves as the inducer molecule for the *pcaBDC* genes and protocatechuic acid induces the transcription of *pcaHG* (Romero-Steiner *et al.*, 1994).

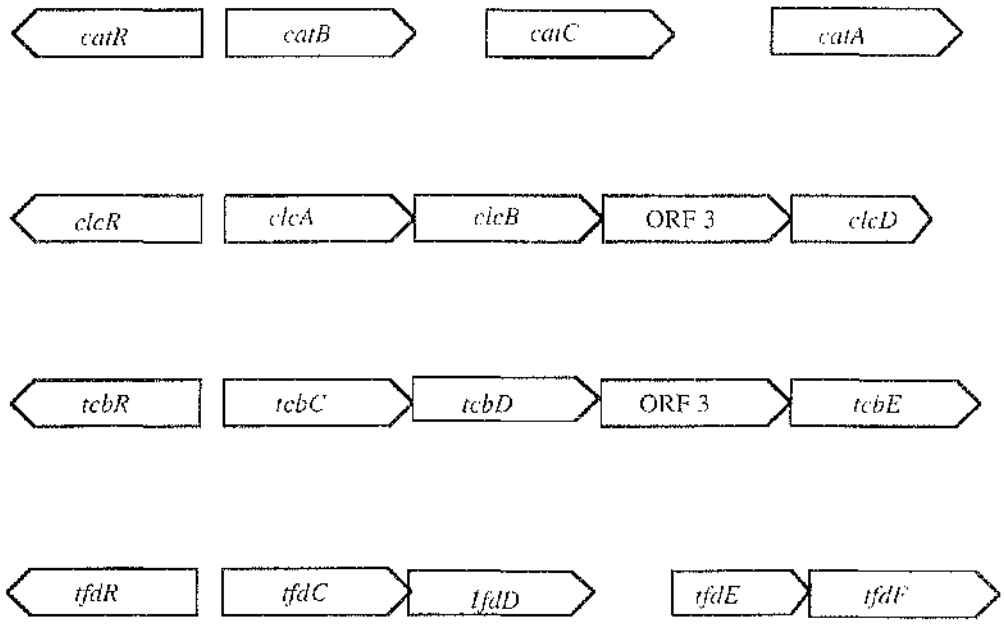

Figure 1.7 Organisation and regulation of the structural genes of different degradative operons in Gram-negative bacteria

The catechol degradative *catBCA* operon, 3-chlorocatechol degradative *clcABD* operon, trichlorocatechol degradative *tcbCDE* operon and the dichlorocatechol degradative *tfdCDE* operon, are positively regulated by divergently expressed regulator proteins (CatR, ClcR, TcbR, TfdR). The *clc*, *tcb* and *tfd* genes are cotranscribed, whereas *catBC* are transcribed separately from *catA* (Chakrabarty, 1996; Matrubutham & Harker, 1994; Coco *et al.*, 1993).

absence of inducer, thus autoregulating its own transcription, but when the inducer is present, the affinity for the activation binding site is increased more than twenty-fold (Rothmel *et al.*, 1991; Parsek *et al.*, 1994). CatM also binds upstream of both *catA* and *catBC* and a $TN_{11}A$ motif has been found in the *catM-catBC* intergenic region in the same relative position to those that have been observed for other LysR family members (Coco *et al.*, 1993; Schell, 1993). Differential binding activity of CatM in the presence of *cis,cis*-muconate has so far not been demonstrated, but CatM has been shown to autoregulate its own synthesis negatively (Romero-Arroyo *et al.*, 1995).

The transcriptional regulation of the genes required for the degradation of protocatechuate (the *pca* gene branch of the β-ketoadipate pathway) (Figure1.6) is less well understood than that of the *cat* branch; however, it is known that the *pca* genes are arranged in four physically distinct clusters (Hughes *et al.*, 1988; Frazee *et al.*, 1993) and are known to be induced by protocatechuate in the case of *pcaHG* (encoding protocatechuate 3,4-dioxygenase) and by β-ketoadipate in the case of the remaining genes. A regulatory protein, PcaR, has been implicated in the expression of *pcaBDC*, which are the genes encoding β-carboxymuconate lactonising enzyme, β-ketoadipate enol-lactone hydrolase and γ-carboxymuconolactone decarboxylase (Nichols & Harwood, 1995). However, unlike *catR* and other members of the *lysR* family, which are divergently transcribed within their regulatory system, the location of *pcaR* is somewhat removed from the defined members of the *pca* pathway; also, primary sequence comparisons of CatR and PcaR from *P. putida*, show that if they are related, they are only distantly so (Nichols & Harwood, 1995).

The microbial processes of chemotaxis and solute transport operate as early steps in benzoate and 4-hydroxybenzoate utilisation (Harwood *et al.*, 1984). In *P. putida*, chemotactic responses to benzoate and 4-hydroxybenzoate are mediated by β-ketoadipate. 4-Hydroxybenzoate uptake is also dependent upon a membrane bound protein encoded by *pcaK*, which is not only induced by β-ketoadipate but is also subject to repression by benzoate. The extra layer of regulatory control within the β-ketoadipate pathway helps to explain the observation that *P. putida* cells presented with a mixture of substrates, preferentially degrade benzoate before 4-hydroxybenzoate (Nichols &Harwood, 1995).

Of all synthetic chemicals, chlorinated aromatics are probably the most refractory to bacterial degradation (Ghosal *et al.*, 1985). The aerobic biodegradation of haloaromatics has been reviewed by Reineke & Knackmuss (1988) and it appears that in most instances, chlorocatechols are the central metabolites. With one possible exception, further assimilation is via the intradiol-

cleavage route (Kersten *et al.*, 1985). At least three different bacterial strains capable of degrading monochloro-, dichloro- and trichlorocatechols have been isolated. The degradative genes *clc* (Frantz & Chakrabarty, 1987), *tcb* (van der Meer *et al.*, 1991*a, b*) and *tfd* (Don *et al.*, 1985), which encode proteins for the breakdown of chlorocatechol (Figure 1.8), trichlorobenzene and 2,4-dichlorophenoxyacetate, respectively, have been characterised. While the catechol degradative genes are chromosomally encoded, the chlorocatechol degradative genes (*clc*, *tcb* and *tfd)* are plasmid encoded and cotranscribed (Chakrabarty, 1996) (Figure 1.7). The first two steps in the degradation of 3-chlorocatechol are carried out by chlorocatechol 1,2-dioxygenase and chloromuconate lactonising enzyme, which are encoded by *clcA* and *clcB* respectively (Figure 1.8). Entirely analogous reactions are carried by the *tfdC* and *tfdD* gene products and by the *tcbC* and *tcbD* gene products (Figure 1.7), during the degradation of dichlorcatechol and trichlorocatechol respectively (Harwood & Parales, 1996). All three of the chlorocatechol gene clusters, *clc*, *tcb* and *tfd* have there own positive regulators, which are divergently transcribed and which ensure the functionality of the pathway (Figure 1.7) (Matrubutham & Harker, 1994).

The precise mechanism of action of the LysR family of transcriptional regulatory proteins is not known, however, all contain a common *N*-terminal region which includes a helix-turn-helix DNA binding motif. A number of studies suggest that the inducer interacts with LysR and induces a conformational change upon the protein. The inducer-protein conformation may change the topology and binding characteristics of the promoter region DNA and/or allow the protein to interact with the RNA polymerase, resulting in transcriptional activation (Parsek *et al.*, 1994).

1.8 The genus *Rhodococcus* and *Rhodococcus rhodochrous* NCIMB 13259

Rhodococcus is a member of the nocardioform actinomycetes, alongside *Nocardia*, *Mycobacterium*, *Corynebacterium*, *Gordona* and *Tsukamurella* (Embley & Stackebrandt, 1994). Rhodococci are aerobic, Gram-positive and non-motile and have a variety of growth patterns, starting with a coccus or short rod stage, which may then form extensively branched hyphae (Finnerty, 1992). The colonies are usually red, orange or pink, due largely, it is thought, to the presence of carotenoid pigments (Takaichi *et al.*, 1990). Rhodococci have been isolated from a wide array of environments: from human sputum to oil-contaminated sea water and are thought to be a significant part of the soil microbial community (Warhurst, 1993). This diversity is perhaps a reflection of their enormous metabolic capabilities, degrading alkanes, halogenated aromatics and other substituted benzene derivatives, as well as heteroaromatics and polycyclic

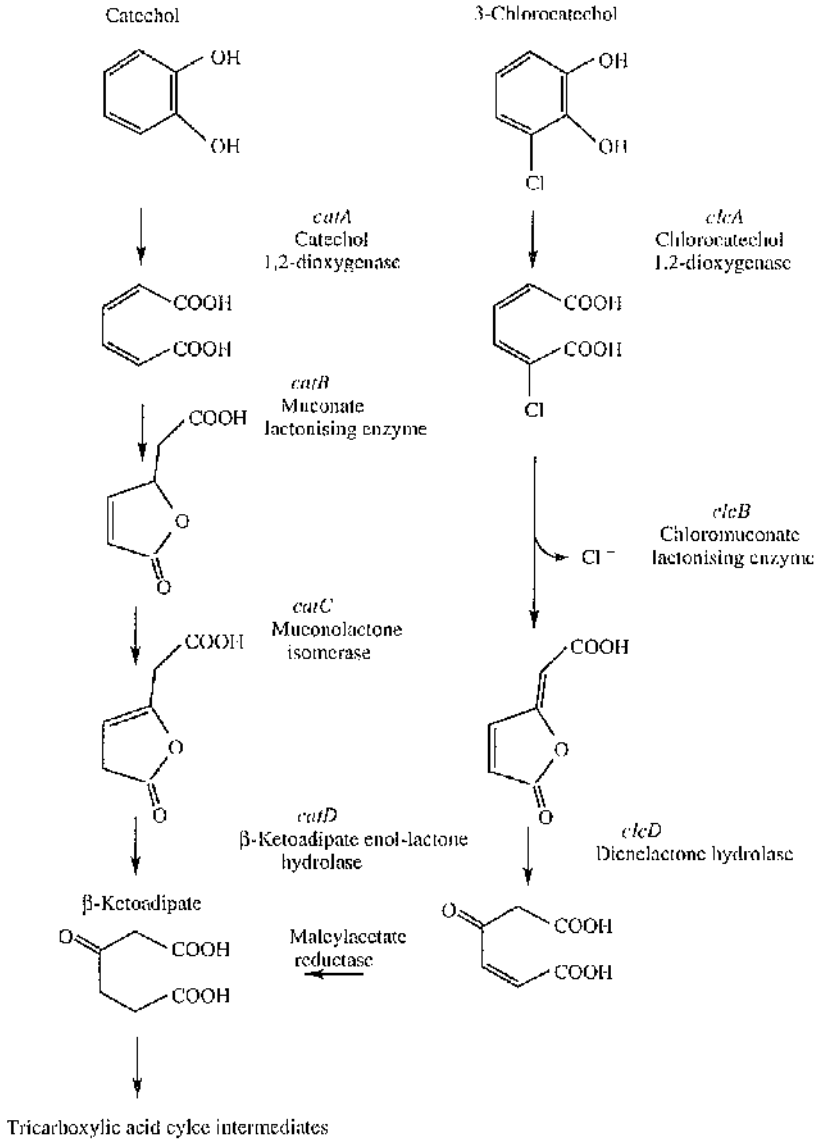

Figure 1.8 Comparison of β-ketoadipate and modified *ortho*-cleavage pathways in *Pseudomonas putida*

The catechol branch (Figure 1.6) of the chromosomally encoded β-ketoadipate pathway compared to the modified, plasmid encoded pathway for the degradation of 3-chlorocatechol. The chloromuconate lactonising enzyme catalyses dechlorination and formation of dienelactones. In the case of the β-ketoadipate pathway, the corresponding enol-lactone is generated by two steps catalysed by muconate lactonising enzyme and muconolactone isomerase. Structurally, the *clcB* and *catB* gene products are closely related. The hydrolases, encoded by *catD* and *clcD*, in contrast, show significant differences. The final step in the modified pathway, carried out by maleylacetate reductase, lacks any counterpart in the β-ketoadipate pathway (Harwood & Parales, 1996).

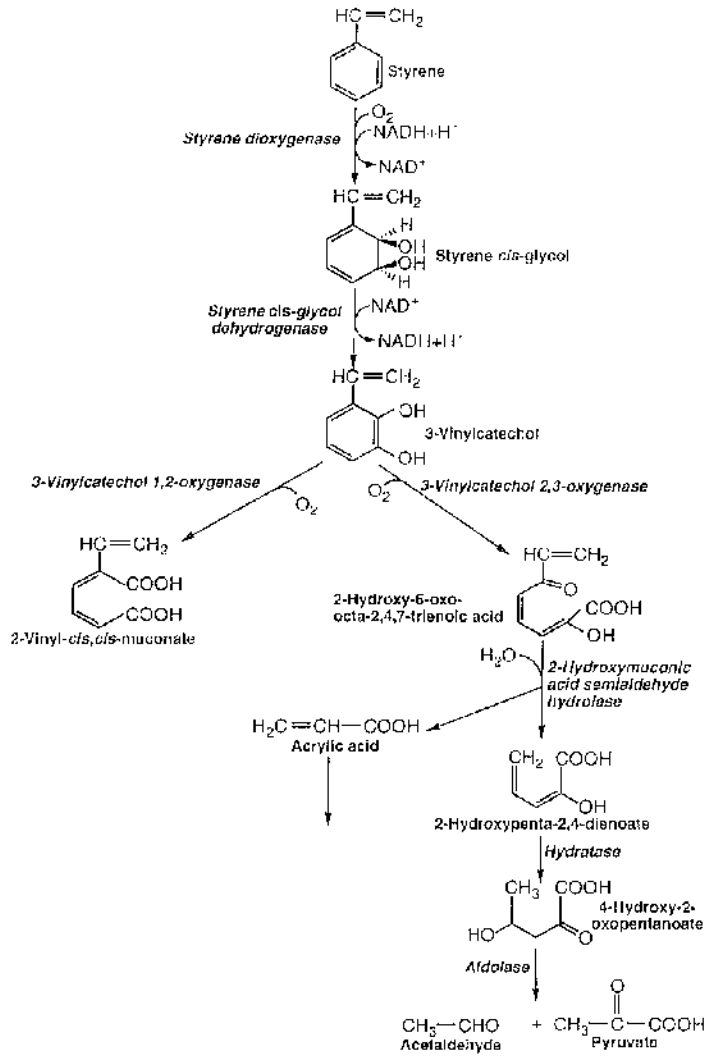

Figure 1.9 The pathway of styrene breakdown in *Rhodococcus rhodochrous* NCIMB 13259

The figure was taken from Warhurst (1993). Catechol 1,2-dioxygenase is referred to in this figure as '3-vinylcatechol 1,2-dioxygenase'.

hydrocarbons (Warhurst & Fewson, 1994). Unlike the pseudomonads and other Gram-negative bacteria, rhodococci do not generally exhibit catabolite repression and in some instances the presence of an easily degradable substrate such as glucose, actually accelerates the metabolism of other compounds such as aromatics (Appel *et al.*, 1984: Fuchs *et al.*, 1991). Because of these qualities and their known persistence in the environment, rhodococci have been found to be effective in bioremediation tests, for example, in the removal of pentachlorophenol from soil (Warhurst, 1993). These bacteria possess most of the pathways needed for degrading and modifying aromatic compounds including dioxygenase and monooxygenase ring attack and the cleavage of catechol and protocatechuate by intradiol- and extradiol-routes (Rast *et al.*, 1980; Warhurst, 1993). *Rhodococcus rhodochrous* NCIMB 13259 can grow on styrene, toluene, ethylbenzene and benzene as sole carbon sources and does this via an extradiol-cleavage pathway (Figure 1.9). This strain also contains a catechol 1,2-dioxygenase, but for reasons which are not clear, the remainder of the intradiol-cleavage pathway appears to be absent and *cis,cis*-muconic acid derivatives appear in the medium when the extradiol-cleaving catechol 2,3-dioxygenase is inhibited by 3-fluorocatechol (Warhurst, 1993; Warhurst *et al.*, 1994).

1.9 Aims

The principal aims of this work were:

- to design a purification procedure for catechol 1,2-dioxygenase from *Rhodococcus rhodochrous* NCIMB 13259,

- to determine its physical, chemical and kinetic properties,

- to clone and sequence the gene encoding catechol 1,2-dioxygenase.

Secondary objectives were:

- to over-express and crystallise catechol 1,2-dioxygenase,

- to characterise the crystals by X-ray crystallography.

CHAPTER 2
Materials and methods

2.1 Materials and general methods

2.1.1 Chemicals

Chemicals were AnalaR grade or better and were obtained from BDH Chemicals Ltd, Poole, Dorset, UK, except for those listed below or in the appropriate part of the text:

Aldrich Chemical Company Ltd, Gillingham, Dorset, UK: 4-methylcatechol, 3-methylcatechol.

Amersham International, Little Chalfont, Bucks., UK: Hybond-N hybridisation membranes.

Boeringer Corporation, Lewes, Sussex, UK: dithiothreitol, rabbit muscle fructose 1,6 bisphosphate aldolase.

Difco, Detroit, USA: nutrient agar.

Fisons Scientific Equipment, Loughborough, UK: boric acid, *N,N'*-methylene bis-acrylamide, sodium dodecyl sulphate and Tris-washed phenol.

FMC Bioproducts, Rockland, USA: SeaKem GTG agarose.

Gibco Life Technologies Ltd, Paisley, UK: 1 kb DNA ladder.

Merck, Darmstadt, Germany: nutrient broth, yeast extract, agar, bactotryptone.

New England Biolabs, MA, USA: *Vent* DNA polymerase.

Pharmacia Ltd, Milton Keynes, Bucks., UK: DEAE-Sephacel, phenyl-Sepharose, Sephadex G-25, low M_r standards for SDS-PAGE, containing rabbit muscle phosphorylase *b* (94 000) bovine serum albumin (67 000), ovalbumin (43 000), bovine erythrocyte carbonic anhydrase (30 000), soyabean trypsin inhibitor (20 100) and bovine milk α-lactalbumin (14 400).

Promega Corporation, Southampton, UK: all restriction enzymes, T4 DNA ligase, *Taq* DNA polymerase.

Sigma Chemical Company, Poole, Dorset, UK: heparin, ferrozine, Neocuprine, catechol, protocatechuic acid, lysozyme, RNase, low melting point agarose, Coomassie Brilliant Blue G250, ovalbumin.

Whatman International, Kent, UK: filter paper.

V8 protease was a gift from Prof H. G. Nimmo, Division of Biochemistry and Molecular Biology, University of Glasgow.

2.1.2 Glassware

Glassware was washed by soaking in Haemosol (10%), rinsed with tap water and then distilled water. Glassware for iron analysis was treated with a 1:1 (v/v) mixture of conc. H_2SO_4 and HNO_3, left for one hour and then rinsed with tap water and then Milli-Q water (Millipore, USA).

2.1.3 pH measurements

pH values were determined with an E.I.L. model 7010 pH monitor (E.I.L. Ltd, Cumbernauld, Strathclyde, UK) connected to a glass electrode (Probion Ltd, Glenrothes, Fife, UK).

2.1.4 Spectrophotometer

All spectrophotometric measurements were done with a Pye Unicam 8730 spectrophotometer (Cambridge, UK) in 1 ml plastic cuvettes, with a pathlength of 1 cm, except when measuring absorbances at wavelengths below 340 nm, when 1 ml quartz cuvettes were used.

2.1.5 Preparation of chromatography media

DEAE-Sephacel and phenyl-Sepharose were supplied pre-swollen. Both gels were degassed and then brought to 4 °C prior to pouring into columns, which were then equilibrated with 4-5 column volumes of the appropriate buffer before use. DEAE-Sephacel was cleaned with 2 M NaCl, and phenyl-Sepharose with distilled water. Both gel types were stored in 30% (v/v) isopropanol.

2.1.6 Sterilisation

Small volumes under 50 ml were filtered through Millex-GV sterile filters (0.22 μM pore, Millipore S.A., 67 Molsheim, France). All other solutions including growth media were sterilised by autoclaving at 120 °C (15 psi) for 75 min.

2.2 Microbiological methods

The bacterial strains and plasmids used during this project are listed in Table 2.1.

2.2.1 Culture media and supplements

All culture media were sterilised (Methods 2.1.6) prior to the addition of supplements. All quantities are per litre of distilled water.

Luria-Bertani medium (LB): 10 g NaCl, 5 g yeast extract, 10 g Bactotryptone

Nutrient broth: 13 g.

MMD: 26 g nutrient broth, 0.9g L-Glu, 2 g KH_2PO_4, 1 g $(NH_4)_2SO_4$, 0.4 g $MgSO_4.7H_2O$, adjusted to pH 7 with KOH (Barrowman & Fewson, 1985).

Chelated metals solution: to 500 ml of distilled water were added (in the order given and waiting for each to dissolve before adding the next); 50 g nitrilotriacetic acid; 125 ml of 5 M NaOH; the solution was then brought to pH 7 by the addition of 5 M HCl; 1.1 g $FeSO_4.7H_2O$, 50 mg $Na_2MoO_4.2H_2O$, 50 mg $MnSO_4.4H_2O$,

Strain/plasmid	Genotype	Source
Rhodococcus rhodochrous NCIMB 13259	wild type isolate	Prof C.A. Fewson, Division of Biochemistry and Molecular Biology, University of Glasgow
Escherichia coli DH5α	Φ80d*lac*ZΔM15, *recA1*, *endA1*, *gyrA96*, *thi-1*, *hsdR17* (r_k^-, m_k^+), *sup E44*, *rel* A1, *deoR*, Δ(*lacZYA-argF*) U169	Promega Corp, Madison, USA
Escherichia coli XL2-Blue	*recA1*, *endA1*, *gyr A96*, *thi-1*, *hsdR17*, *supE44*, *relA1*, *lac* [F' *proAB* $lacI^qZ\Delta M15$ Tn10 (Tet^r) Amy Cam^r]	Stratagene, La Jolla, USA
Escherichia coli JM 109 (DE3)	*recA1*, *supE44*, *endA1*, *hsdR17* (r_k^-, m_k^+), *gyr A96*, *rel A1*, *thi* , Δ(*lac-proAB*), λ_{DE3}, [F', *traD36*, *proAB*, $lacI^q$ *ZΔM15*]	Promega Corp, Madison, USA
pBluescript SK II-	General cloning vector, amp^r	Promega Corp, Madison, USA
pTB361	T7 expression vector, tet^r	Dr D. Gillooly, Division of Biochemistry and Molecular Biology, University of Glasgow

Table 2.1 Bacterial strains and plasmids

$ZnSO_4.7H_2O$, $CuSO_4.5H_2O$, $CoCl_2.6H_2O$, then made upto 1 l with water. Two ml of solution were added to each litre of growth medium (Beggs & Fewson, 1977).

Solid media were produced by the addition of 1.5% (w/v) agar and the hot solutions were then poured into petri-dishes and allowed to set overnight.
Antibiotics were filter sterilised (Methods 2.1.6) and added at the following concentrations: ampicillin 100 μg ml^{-1}, tetracycline 10 μg ml^{-1}.

2.2.2 Maintenance of bacterial cultures

All bacteria used in this project were stored at 4 °C. *Rhodococcus rhodochrous* NCIMB 13259 was maintained on nutrient agar plates and in 10-20 ml of NB. *Escherichia coli* strains were kept on LB agar plates containing the appropriate antibiotic, when required.

2.2.3 Growth of bacteria

2.2.3.1 *Escherichia coli*

Small cultures of up to 50 ml were prepared by inoculating LB (Methods 2.2.1), containing the appropriate antibiotic when required, with single colonies and incubating overnight at 37 °C on a rotary shaker (220 rpm). Larger culture volumes were prepared by inoculating LB, containing the appropriate antibiotic, with a portion of the overnight culture and grown again at 37 °C, on a rotary shaker.

2.2.3.2 *Rhodococcus rhodochrous*

R. rhodochrous NCIMB 13259 was grown at 30 °C on MMD medium (Methods 2.2.1). Small volumes of upto 20 ml were inoculated with single colonies and left without shaking for 2 d at 30 °C. A portion (2 ml) of the starter culture was then added to 200 ml of medium and grown at 30 °C in 500 ml conical flasks for 2 d on a rotary shaker. The culture was then transferred to 4 l of medium and grown in 10 l flasks which were mixed with magnetic stirrers and aerated at 4 l min^{-1}. After 24 h, 2.1 ml of benzyl alcohol were added to give a final concentration of 2 mM. A further 2.1 ml of benzyl alcohol were added 4 h later. After a total of 30 h the cells were harvested.

2.2.4 Harvesting bacteria

Bacteria were harvested by centrifuging at 5 000 rpm for 10 min in a Beckman JA-10 rotor and then washed in approximately 0.02 volumes of Z-1 buffer (1 g $(NH_4)_2SO_4$ l^{-1} and 2 g KH_2PO_4 l^{-1}, adjusted to pH 7 with KOH). The bacteria were then pelleted at 10 000 rpm for 10 min in a Beckman JA-14 rotor

and either stored at –20 °C or used immediately. Smaller culture volumes of 1-5ml were microfuged at 12 000 rpm for 1 min.

2.2.5 Cell breakage

Bacteria were suspended in 50 mM Tris/HCl, pH 7.5 to a final concentration of 20 g (wet weight) of cells /50 ml of buffer and broken by 4 passes through a French press (Aminco, SLM Instruments Inc., American instrument Co, Urbana, Illinois 61801, USA) at 117 MPa.

Cells used for analytical purposes, for example when monitoring protein expression, were resuspended in the requisite buffer and broken with a Lucas-Dawes (London, UK) ultrasonic probe, operated at 60 W for 6 x 10 s pulses with 20 s intervals for cooling. Cells were kept on ice water throughout the operation to prevent overheating.

2.3 Protein estimation

2.3.1 The absorption at 280 nm

During catechol 1,2-dioxygenase purification, the protein concentration of column fractions was determined by measuring the absorbance at 280 nm (Methods 2.1.4).

2.3.2 The Bradford method

Routine protein estimations were done using the Bradford (1976) method. Standard curves of the absorbance (595 nm) against protein concentration (0-100 μg ml^{-1}) were constructed. Both the experimental samples and the bovine serum albumin standards were made up to 100 μl with 0.15 M NaCl before the addition of 5 ml of the Bradford reagent [0.01% Coomassie Brilliant Blue G250, 5% (v/v) ethanol, 10% (v/v) orthophosphoric acid].

2.3.3 The absorption coefficient of catechol 1,2-dioxygenase

The concentration of purified catechol 1,2-dioxygenase (Methods 2.5) was determined spectrophotometrically at 280 nm (Methods 2.1.4) using the method of Gill & Von Hippel (1989). Knowing the primary sequence, the molar absorption coefficient (ε) for catechol 1,2-dioxygenase denatured in 6 M guanidinium/HCl ($\varepsilon_{M,\,Gdn/HCl}$) was found with the following equation:

$$\varepsilon_{M,\,Gdn/HCl} = a\varepsilon_{M,\,Tyr} + b\varepsilon_{M,\,Trp} + c\varepsilon_{M,\,Cys}$$

Where a, b,and c are the number of residues per subunit of Tyr, Trp and Cys respectively and $\varepsilon_{M,\,Tyr}$, $\varepsilon_{M,\,Trp}$ and $\varepsilon_{M,\,Cys}$ are the molar absorption coefficients of Tyr, Trp and Cys respectively as determined by Edelhoch (1967). After measuring the absorbance of catechol 1,2-dioxygenase at identical protein concentrations, both in water and in 6 M Gdn/HCl, the absorption coefficient of the native protein ($\varepsilon_{M,\,nat}$) was calculated by applying the following equation:

$$\varepsilon_{M,\,nat} = (Abs_{\,nat})\,(\varepsilon_{M,\,Gdn/HCl})/(Abs_{\,Gdn/HCl})$$

2.4 Catechol 1,2-dioxygenase activity determination

2.4.1 Assaying catechol 1,2-dioxygenase

Intradiol catechol 1,2-dioxygenase activity was assayed spectrophotometrically at 27 °C by measuring the increase in absorbance at 260 nm for the products from all substrates except protocatechuic acid whose product was monitored at 270 nm. Reactions were carried out in 50 mM sodium/potassium phosphate buffer, pH 7, using various amounts (1-20 µl) of enzyme and were initiated by the addition of 10 µl of 10 mM substrate (Warhurst, 1993).

Extradiol activity was monitored under the same conditions, except the increase in product absorbance was monitored at 375 nm.

One unit of enzyme activity is defined as the amount of enzyme needed to produce 1 µmol of product min^{-1}.

2.4.2 Determination of absorption coefficients of substrates and products

Due to the significant absorbance of catechols at 260 nm, it was necessary to apply a correction factor, since there is a negative contribution to the absorbance as substrate is consumed in the reaction. $\Delta\varepsilon_{260}$ (substrate-product) values were determined by enzymatically converting a known amount of catechol to product as described by Dorn & Knackmuss (1978*b*). This was done with 50 mM sodium/potassium phosphate buffer, pH 7, 2 µl of β-mercaptoethanol and 25 µg of purified catechol 1,2-dioxygenase (Methods 2.5). The spectrophotometer was set to zero absorbance and the reaction was started by the addition of 10 µl of substrate (100 µM). When the reaction was complete, the absorbance was recorded. The $\Delta\varepsilon_{260}$ (substrate-product) values in $mM^{-1}\,cm^{-1}$ for various catechols are as follows: catechol, 14.9; 3-methylcatechol, 15.9; 4-methylcatechol, 13.2. The experiment was also repeated with 100 mM His/HCl, pH 6 and 100 mM Tris/HCl, pH 9.2 with no significant change from the reported value. The $\Delta\varepsilon_{270}$ (substrate-product) value for protocatechuate was 9.1 $mM^{-1}\,cm^{-1}$ (Stanier & Ingraham, 1954).

2.4.3 Analysis of initial velocities and determination of kinetic coefficients

The rates of enzyme reactions were measured spectrophotometrically (Methods 2.1.4) with 7 to 8 different substrate concentrations (0.2- 5.0 μM) with air-saturated buffer (100 mM Tris/HCl, pH 9) at 27 °C. The data were analysed using the Direct Linear method and also by the Lineweaver-Burk and Hanes-Woolf linear methods using the Enzpack computer program (Williams, 1985), to give estimates of the apparent K_m and V values.

2.5 The purification of catechol 1,2-dioxygenase

All the following steps were done at 4 °C.

2.5.1 Buffer solutions used in the purification

Buffer A: 0.18 M NaCl, 50 mM Tris/HCl, pH 7.5.

Buffer B: 0.4 M NaCl, 50 mM Tris/HCl, pH 7.5.

Buffer C: NaCl 0.18 M, 15% $(NH_4)_2SO_4$, 50 mM Tris/HCl, pH 7.5.

Buffer D: 50 mM Tris/HCl, pH 7.5.

2.5.2 Preparation of the clarified extract

After disruption of the bacteria (Methods 2.2.5), cellular debris was pelleted at 12 000 rpm for 30 min in a Beckman JA-20 rotor. The supernate was retained and clarified further in a Beckman Ti.70 rotor at 45 000 rpm for 90 min and used immediately.

2.5.3 DEAE-Sephacel chromatography

The clarified extract (80 ml) obtained from 20 g of cells was loaded at 60 ml h^{-1} onto a DEAE-Sephacel column (2.6 x 10 cm) which had been equilibrated with Buffer A (Methods 2.5.1). The column was then washed at 60 ml h^{-1} for approximately 7 column volumes with Buffer A. Catechol 1,2-dioxygenase was then eluted with a 300 ml positive salt gradient of Buffer A-Buffer B (Methods 2.5.1). The flow rate was 20 ml h^{-1} and those fractions (5 ml) having 50% or more activity than that contained in the maximal fraction, were pooled.

2.5.4 Phenyl-Sepharose chromatography

The ion-exchange pool (27 ml) was brought to 15% saturation $(NH_4)_2SO_4$ by the addition of 8.2 g $(NH_4)_2SO_4$ $(100\ ml)^{-1}$ of eluate. The pool was then loaded at 60 ml h^{-1} onto a phenyl-Sepharose column (2.6 x 10 cm) which had been pre-equilibrated with Buffer C (Methods 2.5.1). The column was then washed for

approximately 1 column volume with the same buffer at 60 ml h^{-1} for 2 column volumes. Catechol 1,2-dioxygenase was then eluted with a negative $(NH_4)_2SO_4$ gradient of 300 ml (Buffer C-Buffer D) (Methods 2.5.1). The flow rate was 20 ml h^{-1} and all fractions (5 ml) having 50% or more of that in the maximal fraction were pooled.

2.5.5 Fast protein liquid chromatography

Protein required for crystallisation was purified further with a Mono Q 10/10 ion-exchange column (Pharmacia, UK) operated at 1 ml min^{-1} with a 30 min gradient [Buffer D (Methods 2.5.1), 0.2 M NaCl to Buffer D, 0.4 M NaCl]. Superose 12 gel-permeation columns were operated at or below 1 ml min^{-1}. A Pharmacia FPLC system was used throughout and the eluent monitored on-line at 280 nm.

2.5.6 Concentration of catechol 1,2-dioxygenase

(a) The $(NH_4)_2SO_4$ concentration of the pooled phenyl-Sepharose fractions (36 ml) was increased by adding 7.9 g of $(NH_4)_2SO_4$ (100 ml)$^{-1}$ of eluate. The pool was poured into a 20 ml syringe and concentrated in 1-2 ml of phenyl-Sepharose gel which was contained in the bottom of the syringe by a glass wool plug. Catechol 1,2-dioxygenase was eluted by gravity flow with 5-10 ml of Buffer A (Methods 2.5.1) and collected manually.

(b) Samples for crystallisation experiments were concentrated further in Centricon-30 ultra filtration units (Amicon).

2.5.7 Storage of catechol 1,2-dioxygenase

Protein which was used for crystallographic studies was kept at 4 °C, otherwise it was stored in Eppendorf tubes at –20 °C.

2.6 Crystallisation and X-ray crystallography

2.6.1 The crystallisation of catechol 1,2-dioxygenase

Crystals were grown using the sitting drop vapour diffusion technique (Blundell & Johnson, 1976), in 24-well, sitting drop trays (NBS Biologicals, UK). Each well was filled with 1 ml of precipitant solution (Jancarik & Kim, 1991); 4 µl was then transferred from each well to the corresponding reservoir where it was mixed with 4 µl of purified (Methods 2.5) protein (25 mg ml^{-1}). The plates were covered with 'crystal clear' tape (Manco inc., Westlake, Ohio, USA) and left to equilibrate with the well solution at 18 °C.

2.6.2 Microseeding

Protein was allowed to equilibrate with the well solution for 3-5 d. A crystal, obtained from a previous experiment, was washed briefly in 10 µl of precipitant solution and then transferred to 10 µl of fresh precipitant solution where it was crushed with a needle. The microcrystals were diluted 20-fold with precipitant and then introduced into each reservoir with a human hair (Stura & Wilson, 1992).

2.6.3 Characterisation by X-ray crystallography

X-ray crystallography was carried out on a Siemens area detector. Data were processed using the XDS program (Kabsch, 1987).

2.7 Polyacrylamide gel electrophoresis

2.7.1 SDS-PAGE

SDS-PAGE was performed essentially according to Laemmli (1970). The volumes quoted are for two gels.

(a) 12% (w/v) resolving gel

The following solution components: distilled water (3.3 ml), 30% (w/v) acrylamide
(4 ml), 1.5 M Tris/HCl, pH 8.8 (2.5 ml), 10% (w/v) SDS (0.1 ml), 10% (w/v) ammonium persulphate (0.1 ml) and TEMED (0.004 ml), were mixed, poured between glass plates and overlaid with several drops of distilled water. Once set, any unpolymerised liquid which remained was gently blotted from the gel surface.

(b) 5% (w/v) stacking gel

The following solutions: distilled water (3.4 ml), 30% (w/v) acrylamide (0.83 ml), 1M Tris/HCl, pH 6.8 (0.63 ml), 10% (w/v) SDS-PAGE (0.05 ml), 10% (w/v) ammonium persulphate (0.05 ml) and TEMED (0.005 ml) were mixed together and poured on top of the resolving gels. Ten-track Teflon combs were inserted and the gels left to polymerise.

(c) Sample preparation and electrophoresis

Samples (1-50 µg) were mixed with 2 volumes of loading buffer [60 mM Tris/HCl, pH 6.8, 1% SDS, 20% (v/v) glycerol, 50 mM dithiothreitol and 0.002% Bromophenol Blue] and boiled for 4 min. The reservoir buffer was 25 mM Tris, 250 mM glycine, pH 8.3, 0.1% SDS. Gels were run at 200 V on a BIO-RAD, Mini-PROTEAN II apparatus, until the Bromophenol Blue dye had reached the bottom of the gel.

2.7.2 Phosphate SDS-PAGE

Gels (5% acrylamide) were made from the following: distilled water (5.25 ml); 30% (w/v) acrylamide (2.25 ml); 0.2 M sodium phosphate, pH 7.2 (2 ml); 10% (w/v) SDS (0.1 ml); 10% (w/v) ammonium persulphate (0.1 ml); TEMED (0.005 ml). Ten-track Teflon combs were inserted and the gels left to polymerise.

(a) Sample preparation and electrophoresis

Proteins (5-10 µg) were mixed with 0.1 ml of sample buffer (1% (w/v) SDS; 0.01 M sodium phosphate, pH 7.2; 10% (v/v) glycerol; 0.002% Bromophenol Blue) and 1 M dithiothreitiol (5 µl) and boiled for 4 min. Samples (5-20 µl) were applied to the gel. Electrophoresis was performed in 0.1% (w/v) SDS, 0.05 M sodium phosphate reservoir buffer at 60 V until the dye layer was approximately 1 cm from the bottom of the gel.

2.7.3 Non-denaturing PAGE

The gels were cast as in Methods 2.7.1, but were aged overnight to remove unwanted free radicals.

(a) 8% (w/v) resolving gel

Resolving gels were cast from the following solutions: distilled water (4.6 ml), 30% (w/v) acrylamide (2.7 ml), 1.5 M Tris/HCl, pH 8.8 (2.5 ml), 10% (w/v) ammonium persulphate (0.1 ml) and TEMED (0.006 ml).

(b) 3% (w/v) stacking gel

Stacking gels consisted of the following solutions: distilled water (3.7 ml), 30% (w/v) acrylamide, 1.5 M Tris/HCl, pH 6.8 (0.52), 10% (w/v) ammonium persulphate (0.05 ml) and TEMED (0.004 ml)

(c)Sample preparation and electrophoresis

Samples (2-60 µg) were mixed with an equal volume of sample buffer [50% (v/v) glycerol, 0.002% Bromophenol Blue] and 10-25 µl applied to the gel. The reservoir buffer was 25 mM Tris, 250 mM glycine, pH 8.3. Electrophoresis was done at 60 V until the dye reached the bottom of the stacking gel and then increased to 120 V until the dye reached about three-quaters of the way down the gel. Gels were either stained for protein (Methods 2.7.4) or were cut up in strips (2 x 10 mm) and assayed for activity (Methods 2.4.1).

2.7.4 Staining and destaining of gels

All polyacrylamide gels were stained [0.25% Coomassie Brilliant Blue G250, 45% (v/v) methanol, 10% (v/v) glacial acetic acid] for 15 min at 55 °C and then destained [10% (v/v) acetic acid, 10% (v/v) methanol] for 2 h at the same temperature.

2.7.5 Gel Scanning

Gels were scanned with an LKB 2202 Ultrascan laser densitometer.

2.8 Protein chemistry

2.8.1 Lyophilisation

Frozen samples were placed in a desiccator attached to a Flexi-dry lyopohiliser (F.T.S. Systems Inc., Stone Ridge, NY, USA) and the desiccator evacuated with a Javac (Farnham, Surrey, UK) vacuum pump.

2.8.2 Desalting and buffer exchange

Sephadex G-25 gel filtration columns of the appropriate size, and pre-equilibrated with the requisite buffer were used for desalting and buffer exchange. In order to avoid excessive dilution, sample volumes were generally 10-20% of the total column volume. Columns were operated at 40-60 ml h^{-1}.

Where dilution was undesirable, samples were concentrated by centrifugation (12 000 g) in Centricon-30 ultra-filtration units (Amicon), diluted with the appropriate buffer, and the process repeated a further 5 times.

2.8.3 *Staphylococcus aureus* V8 protease digestion

Protein solutions were made 50 mM NH_4HCO_3, pH 7.8, 2 M guanidine/HCl by ultra-filtration (Method 2.8.2) and incubated for 15 min at 37 °C. The denatured protein was then diluted with 2 volumes of distilled water and digested with V8 protease for 5 h at 37 °C using a protease:protein ratio of 1: 50 (w/w). At the end of the incubation the peptides were snap-chilled over a mixture of methanol and $CO_{2(S)}$ and lyophilised (Methods 2.8.1).

2.8.4 Separation of peptides by HPLC

Peptidic digests (Methods 2.8.3) were dissolved in 0.1% trifluoroacetic acid (TFA) in water and separated by HPLC using a Waters bondpack μ18 column (Millipore, Milford, MA, USA) with the following gradient: 100% A (0.1% TFA)-50% B [0.1% TFA in 90% (v/v) acetonitrile]. The flow rate was 1 ml min^{-1} and the peptides were detected at both 220 nm and 280 nm with a Shimadzu photodiode array detector (Shimadzu, Japan).

2.8.5 *N*-Terminal and peptide sequencing

N-Terminal and peptide sequences were determined by Dr J. N. Keen, Department of Biochemistry, University of Leeds, Leeds, UK. Samples were subjected to automated liquid-pulse Edman degradation on an Applied biosystems 477A sequencer with on-line narrow bore HPLC and 610A data analysis.

2.8.6 Electrospray mass spectroscopy

Mass spectroscopy was carried out on a VG platform quadrupole mass spectrometer fitted with a pneumatically assisted electrospray source and controlled using the VG Mass Lynx software (VG biotech Ltd, Altrincham, Cheshire, UK). The carrier solvent [1:1 (v/v) acetonitrile/water] infusion was controlled at a flow rate of 10 µl min^{-1} using a Harvard syringe pump (Harvard Apparatus, South Natic, MA, USA). Capillary voltages were between 2.8-3.2 kV, extraction cone voltages 20-30 V and the focusing cone voltage offset by +10 V. The source temperature was set at 65 °C and the nebulizing gas flow at 10 l h^{-1}. Before use, the instrument was calibrated over the molecular weight range of interest with horse heart myoglobin. Samples were desalted by ultra-filtration (Methods 2.8.2) with HPLC grade water, diluted with an equal volume of 4% (v/v) formic acid in acetonitrile and then 10-20 µl samples were injected into the instrument. The raw data was processed using the MaxEnt deconvolution programme using 1 Da peak width and 1 Da/channel resolution.

2.8.7 Cross-linking of enzyme subunits

Purified catechol 1,2-dioxygenase (Methods 2.5) and aldolase were chemically cross-linked, in separate reactions, with the bis-imidoesters; dimethylpimelidate and dimethylsuberimidate. Due to the extremely short half-life of bis-imido esters in solution, a double strength stock solution of the cross-linking reagent was prepared immediately before use (0.2 M NaCl, 0.1 M triethanolamine/HCl, pH 8, 40 mM bis-imidoester, adjusted to pH 8 with 5 M NaOH). Proteins were placed in 0.2 M NaCl, 0.1 M triethanolamine/HCl, pH 8, by ultra-filtration (Methods 2.8.2) and the concentration adjusted to 2 mg ml^{-1}, so that upon addition of the cross-linking reagent the protein concentration was approximately 1 mg ml^{-1}. The reaction was carried out in a volume of 100 µl and was allowed to continue for 1 h at room temperature, whereupon the samples were analysed electrophoretically (Methods 2.7.2 & 2.7.5).

2.8.8 Iron analysis by colorimetry and atomic absorption spectrometry

Acid-washed glassware (Methods 2.1.2) was used throughout and all solutions were made with Milli-Q water. Iron standards were made from suitably diluted iron standard solution (BDH Laboratory supplies, Poole, UK).

2.8.8.1 Colorimetry

The colorimetric quantitation of iron was performed essentially according to Fish (1988). Catechol 1,2-dioxygenase was desalted on Sephadex G-25 (Methods 2.8.2) and its concentration adjusted to 2-4 mg ml^{-1}. The protein

solution (1 ml) was mixed thoroughly with 0.5 ml of solubilising reagent [0.6M HCl, 2.25% (w/v) $KMnO_4$] and digested for 2 h at 60 °C. Following digestion, 0.1 ml of chelating reagent (6.5 mM ferrozine, 13.1 mM Neocuprine, 2 M ascorbic acid, 5 M ammonium acetate) was added and the resulting solution left to stand for 1 h to allow the magenta colour to develop. The absorbance of the experimental samples was measured at 562 nm (Methods 2.1.4) and converted into µg Fe by using a standard curve constructed from the absorbance values of known iron standards. The absorbance of both the standards and the experimental samples were corrected with a reagent blank. Water was used as the blank correction to the standards and solubilising reagent to the experimental samples.

2.8.8.2 Atomic absorption spectroscopy

Atomic absorption spectroscopy was carried out by Mr M. Beglan, Department of Chemistry, University of Glasgow, on a Perkin Elmer Cetus Atomic absorption spectrophotometer with an oxy-acetylene flame and the amount of Fe determined by using known iron standards.

2.8.9 Determination of total sulfhdryl groups

The determinaton of total sufhdryl groups was performed as described by (Habeeb, 1972). Approximately 0.01-0.04 µl of protein solution (~1 mg ml^{-1}) was dissolved in 6 ml of solution containing 2% (w/v) SDS, 0.08 M sodium phosphate buffer, pH 8 and 0.5 mg EDTA ml^{-1} . To 3 ml of the solution was added 0.1 ml dithionitrobenzoic acid (DTNB) solution (40 mg DTNB in 10 ml of 0.1 M sodium phosphate buffer, pH 8). The colour was developed for 15 min and read at 410 nm (Methods 2.1.4), against protein solution in SDS, to give an apparent absorbance. A reagent blank was subtracted from the apparent absorbance to give the net absorbance, which was then combined with a molar absorption value of 13 600 M^{-1} cm^{-1} to give the total sulfhdryl content.

2.9 Molecular biological methods

2.9.1 Buffers and solutions

TE: 10 mM Tris/HCl, pH 8, 1 mM EDTA, pH 8.

10 x TBE: 108 g Tris base, 55 g boric acid, 40 ml 0.5 M EDTA, pH 8 and made upto 1 l with distilled water.

10 x TEA: 49 g Tris base, 11.4 ml acetic acid, 20 ml EDTA, pH 8 and made upto 1 l with distilled water.

STET: 10 mM Tris/HCl, pH 8, 50 mM EDTA, pH 8, 8% (w/v) sucrose, 0.5% Triton X-100.

20 x SSC: 3 M NaCl, 0.3 M sodium citrate.

Denaturing solution: 0.5 M NaOH, 1.5 M NaCl.
Neutralising solution: 1.5 M NaCl, 0.5 M Tris/HCl, pH 7.4.
Prehybridisation buffer: 6 x SSC, 50 µg heparin ml^{-1} .
Hybridisation solution: 6 x SSC, 50 µg heparin ml^{-1} , 0.5% SDS.
Loading buffer: 10 mM Tris/HCl, pH 7.2, 1 mM EDTA, 20% (w/v) ficol, 0.5% Bromophenol Blue, 0.05 mg ethidium bromide ml^{-1} .

2.9.2 Agarose gel electrophoresis

Samples were mixed with 0.2 volumes of loading buffer (Methods 2.9.1) and loaded onto 0.50-0.75% agarose gels, which were made with either 1 x TEA buffer or 1 x TBE buffer (Methods 2.9.1). Electrophoresis was performed in the requisite buffer at 80-120 V with a Hybaid Electro-4 system.

2.9.3 Purification and recovery of DNA from agarose gels

Samples were separated electrophoretically using 0.5-1.0% agarose/TEA gels (Methods 2.9.1, 2.9.2) and the requisite band excised with a scalpel under UV light. The fragment was then extracted with Spin-X centrifuge filters (Costar Inc, Mass, USA) according to the manufacturer's instructions.

2.9.4 DNA purification

Proteins and salts were removed with the Wizard clean-up kit (Promega, UK) according to the manufacturer's instructions.

2.9.5 DNA concentration

All nucleic acids were concentrated by ethanol precipitation. To 1 volume of the nucleic acid solution was added 0.1 volume of 1 M NaCl and 2 volumes of ethanol. The solution was vortexed briefly and then kept at –70 °C for 15 min, before being centrifuged at 12 000 rpm for 10 min in a Jouan MR14.11 centrifuge. Pellets were washed in 70% (v/v) ethanol (–20 °C), air-dried and then dissolved in distilled water.

2.9.6 Large scale plasmid purification

High quality plasmid DNA required for sequencing or PCR work, was purified from 50 ml overnight cultures (Methods 2.2.3.1) using Qiagen midi-preps (Qiagen Inc, Hilden, Germany) according to the manufacturer's instructions.

2.9.7 Small scale plasmid preparation

Plasmid DNA required for analytical purposes was obtained using a lysozyme boiling method. Overnight cultures (1ml) (Methods 2.2.3.1) were centrifuged for 5 min at 12 000 rpm on a bench top microfuge and the pellets resuspended in 350 µl STET (Methods 2.9.1). The suspension was boiled for 45 s with 25 µl lysozyme (10 mg ml^{-1}) and then microfuged for 15 min at 4 °C. The supernate was placed into a fresh Eppendorf tube and vortexed briefly with 40 µl sodium acetate (2.5 M) and 420 µl isopropanol and then kept at –20 °C for 15 min, before microfuging at 12 000 rpm at 4 °C. The supernate was then decanted off and the residue washed twice with 0.2 ml ethanol, before being air-dried for 5 min. The pellet was then dissolved in 10-30 µl of distilled water.

2.9.8 Purification of genomic DNA from *Rhodococcus rhodochrous*

Genomic DNA was purified essentially according to Saito & Miura (1966). Cultures (200 ml) (Methods 2.2.3.2) were centrifuged in a Beckman JA-20 rotor at 5 000 rpm for 10 min. The pellets were washed twice with 25 ml of TE buffer (Methods 2.9.1), resuspended in 6 ml TE buffer and incubated at 37 °C for 1 h with 12 mg of lysozyme. The lysate was then extracted using an equal volume of Tris-saturated phenol and the solution mixed gently for 0.5 h before centrifuging in a Beckman JA-20 rotor at 5 000 rpm for 10 min. Two further organic extractions were performed on the aqueous layer with phenol:chloroform:isoamylalcohol (Sigma, UK). Finally, the aqueous layer was retained and the DNA concentrated by ethanol precipitation (Methods 2.9.5).

2.9.9 Oligonucleotide synthesis

Oligonucleotides were made on an Applied Biosystems DNA synthesiser (Model 381A) by Dr V Math, Division of Biochemistry and Molecular Biology, University of Glasgow and supplied in a solution of 35% (v/v) NH_4OH. Before use, oligonucleotides were isolated by ethanol precipitation (Methods 2.9.5) and diluted in an equivalent volume of distilled water.

2.9.10 Polymerase chain reaction

The polymerase chain reaction was performed on a DNA Thermal cycler, Model 480 (Perkin Elmer Cetus, USA) in 0.5 ml micro centrifuge tubes. DNA polymerase was added after an initial 3 min denaturation step (95 °C) and all reaction mixtures were covered with 80 µl of light mineral oil (Sigma, UK) to prevent sample evaporation. Annealing temperatures (T_a) for each primer were estimated by subtracting 5 °C from the melting temperatures (T_m) (Bej *et al.*, 1991) which were calculated with the GeneJockey program (Taylor, 1991).

2.9.10.1 Amplification with degenerate oligonucleotide primers

PCR was performed with the following buffers in a reaction volume of 100 µl: 25 mM $MgCl_2$ (14 µl), 10x reaction buffer [(0.1 M Tris HCl, pH 9, 0.5 M KCl, 1% Triton X-100) 10 µl], 10 x dNTPs [(10 mM dATP, 10 mM dCTP, 10 mM dGTP, 10 mM dTTP) 8 µl], 1 µg of each primer (Methods 2.9.9), 4% (v/v) dimethylsulfoxide , 200 ng DNA (Methods 2.9.8) or 1 µl from a previous PCR experiment, 1 µl *Taq* DNA polymerase.

Thirty cycles were carried out as follows: denaturation at 95 °C for 1 min; annealing at 40-54 °C for 4 min, followed by a 45 s extension at 72 °C.

2.9.10.2 Amplification with homologous primers

Vent DNA polymerase was used because of the higher fidelity of this enzyme. Reaction conditions were the same as in Methods 2.9.10.1, except no added $MgCl_2$ was used, since the proprietary buffer already contained Mg^{2+} [(10 x ThermoPol buffer: 0.1 M KCl, 0.2 M Tris/HCl, pH 8.8, 0.1 M $(NH_4)_2SO_4$, 20 mM $MgSO_4$, 1% Triton X-100) 10 µl].

Twenty-five cycles were carried out as follows: denaturation at 95 °C for 1 min; annealing at 54 °C for 1.5 min, followed by a 45 s extension at 72 °C.

2.9.10.3 PCR for automated sequencing

PCR was carried out using the ABI PRISM Dye Terminator Cycle Sequencing Ready Reaction Kit using 200 ng of primer (Methods 2.9.9) and 1 µg of DNA (Methods 2.9.6). Thirty cycles were carried out as follows: denaturation at 95 °C for 1 min; annealing at 55 °C for 1 min, followed by a 2 min extension at 72 °C.

2.9.11 Restriction endonuclease digestion of DNA

Genomic DNA digestions were performed overnight on upto 10 µg of DNA (Methods 2.9.8) in 100-300 µl volumes. Upto 5 µg of plasmid DNA or PCR products were digested in 20 µl volumes for 3 h. All reactions were done using the appropriate number of enzyme units and under conditions recommended by the manufacturer. After digestion, restriction enzymes were inactivated where possible and when appropriate, by heating to 65 °C for 20 min.

2.9.12 Radioactive labelling

Specific fragments of DNA that were used as radioactive probes were first excised from the requisite vector DNA using the appropriate restriction enzymes and then purified by agarose gel electrophoresis (Methods 2.9.3). The DNA fragment (50 ng) was then denatured at 95 °C for 10 min and labelled with the

Random Primed DNA labelling kit (USB, Cleveland, USA) essentially according to the manufacturer's instructions. Unincorporated nucleotides were removed with Chroma Spin+TE-10 columns (Clontech Laboratories Inc, CA, USA). Before use, the purified probe was denatured by heating it to 80 °C.

2.9.13 Transferring DNA onto Hybond-N nylon membranes

2.9.13.1 Southern blot

DNA fragments (Methods 2.9.11) were transferred to Hybond-N membranes (Amersham, UK) using the capillary blot method of (Southern, 1975), as described by Sambrook *et al.* (1989). After electrophoresis of the DNA fragments, the agarose gel was soaked in denaturing solution (Methods 2.9.1) for 45 min to denature the DNA and then transferred to neutralising solution (Methods 2.9.1). After 30 min, the neutralising solution was replaced with fresh solution and the soaking continued for a further 15 min. A wick was made by wrapping a gel tray with a strip of Whatman 3MM filter paper and then placed in a glass casserole dish. The dish was filled with transfer solution (10 x SSC, Methods 2.9.1), until the level of the liquid almost reached the top of the wick. The gel was placed on top of the wick and surrounded with saran wrap to prevent short-circuiting the blotting process. A piece of Hybond-N membrane, the same size as the gel, was placed on top and any air bubbles removed with a glass rod. Three pieces of Whatman 3MM filter paper, cut to size and wetted with transfer solution, were placed on top of the membrane. A stack of paper towels was then placed on top of the 3MM paper and weighed down with about 0.5 kg. After about 15 h, the gel was removed, restained in ethidium bromide (10 µg ml^{-1}) and checked on a UV transilluminator to make sure that the transfer had been successful. The membrane was soaked for 5 min in 6 x SSC (Methods 2.9.1) and then allowed to air-dry.

2.9.13.2 Colony lifts

Recombinant bacteria (Methods 2.9.17) were transferred onto Hybond-N membranes according to Hybond-N (1995). Membrane filters were laid upon agar plates (pre-cooled for at least 1 h) and marked with a pin in order to ensure correct orientation of the colonies in subsequent manipulations. In order to lyse the bacteria and denature the DNA, the membrane filters were placed colony side uppermost on 10% (w/v) SDS-soaked Whatman 3MM filter paper for 3 min and then removed and placed on Whatman 3MM paper soaked in denaturing solution (Methods 2.9.1). After 7 min the membranes were removed and placed on Whatman 3MM filter paper soaked in neutralising solution (Methods 2.9.1). At the end of 3 min, the membranes were removed and laid on more filter papers

containing fresh neutralising solution. Finally, the membranes were washed vigorously in 2 x SSC (Methods 2.9.1) to remove proteinaceous material and allowed to air-dry.

2.9.14 Hybridisation of DNA with the radiolabelled probe

Hybridisations were done in heat-sealed bags in a total volume of 10-20 ml with 1-2 ng ml^{-1} of radiolabelled probe (2.9.12).

DNA (Methods 2.9.13) was cross-linked to Hybond-N membranes (Amersham, UK) by exposure to UV radiation for 2 min, after which the membranes were blocked by incubation in prehybridisation buffer (Methods 2.9.1) for 2 h at 68 °C. The prehybridisation buffer was then exchanged for hybridisation buffer (Methods 2.9.1). The temperature was maintained at 68 °C for 6 h, whereupon the membranes were washed at room temperature in 2 x SSC (Methods 2.9.1), 0.5% SDS for 5 min, followed by a subsequent wash with 2 x SSC, 0.1% SDS for 15 min. The filters were then given two further 1 h washes with 0.1 x SSC, 0.5% SDS, one at 37 °C followed by another at 68 °C. Finally, the filters were rinsed in 0.1 x SSC, dried and then autoradiographed (Methods 2.9.16) for 24 h at −70 °C.

2.9.15 Nucleic acid sequencing

2.9.15.1 Manual sequencing

Sequencing gels (6%) were made by mixing 63 g urea, 22.5 ml 40% (w/v) acrylamide, 15 ml 10 x TBE (Methods 2.9.1) and making upto 150 ml with distilled water. TEMED (125 µl) and 750 µl of 10% (w/v) ammonium persulphate were added, mixed well and the gel-mix poured between 45 mm gel-plates and allowed to set overnight. Purified plasmid DNA (Methods 2.9.6) was sequenced manually using the USB Sequensase 2.0 DNA Sequencing Kit (Amersham, USA) which was used according to the manufacturer's instructions. Just before the samples were loaded (3 µl/well), they were heated to about 80 °C and placed on ice to prevent renaturing. The samples were then loaded onto pre-heated gels and electrophoresed at 40 mA using a 45 mm standard thermoplate sequencer (IBI, Connecticut, USA), until the Bromophenol Blue dye reached the base of the gel. After electrophoresis, the gel was fixed for 30 min [10% (v/v) methanol, 10% (v/v) acetic acid], transferred to a sheet of 3MM Whatman filter paper and dried at 80 °C under a vacuum on a gel-drier. The dried gel was then autoradiographed (Methods 2.9.16).

2.9.15.2 Automated sequencing

PCR reaction products (Methods 2.9.10.3) were ethanol precipitated (Methods 2.9.5), redissolved in loading buffer (3 µl formamide) and then loaded onto 24 mm 4.75% (w/v) polyacrylamide gels, which were prepared and operated by Mr G. Hamilton, Division of Molecular Genetics, University of Glasgow, using an ABI DNA sequencer (model 373A). The results were analysed using the GeneJockey sequence processor (Taylor, 1991).

2.9.16 Autoradiography

Autoradiography was done by exposure of Hyperfilm MP (Amersham, England), which was then developed with a Kodak X-Omat.

2.9.17 Cloning

Digested plasmid DNA (Methods 2.9.11) was dephosphorylated in order to minimise the religation of plasmid DNA, 5' ends (a) and then ligated to insert DNA (b). Competent cells were then transformed with ligation products (c).

(a) Dephosphorylation

Calf intestinal alkaline phosphatase (Promega, UK) was added to digested plasmid DNA at 0.1 units $pmol^{-1}$ of 5' ends, where one unit is defined as the amount of enzyme required to catalyse the hydrolysis of 1 µmol of 4-nitrophenyl phosphate per min at 37 °C in 1 M diethanolamine, 10.9 mM paranitrophenyl phosphate, 0.5 mM $MgCl_2$, pH 9.8 (Promega, 1996). For those restriction enzymes which could not be heat deactivated and which exhibited star activity at high pH values, digests were first purified (Methods 2.9.4) prior to the addition of calf intestinal alkaline phosphatase. In all other instances, calf intestinal alkaline phosphatase was added directly to restriction digests and reactions carried out in volumes of 100 µl, under conditions recommended by the manufacturer. Upon completion of the reaction, the DNA was purified (Methods 2.9.4).

(b) Ligation

T4 DNA ligation of purified DNA moieties (Methods 2.9.4) was carried out in a total volume of 10 µl using 10 x T4 DNA ligase buffer (1 µl) and T4 DNA ligase (1 µl). Reactions were done at 15 °C overnight. Plasmid DNA: insert DNA molar ratios varied from 1:2 to 1:4, using upto 80 ng of plasmid DNA/reaction.

(c) Preparation of competent cells and their transformation

Apart from *E. coli* XL2-Blue ultra-competent cells which were used according to the manufacturer's instructions, all other cells were prepared as follows: LB (50 ml) was inoculated with 1 ml of overnight culture (Methods 2.2.3.1) and grown to an optical density at 550 nm of 0.4-0.6 and centrifuged in a

chilled Beckman rotor at 5 000 rpm for 5 min and resuspended in half volumes of ice cold 50 mM $CaCl_2$. The cells were left on ice for 30 min, pelleted and then resuspended in 4 ml of 50 mM $CaCl_2$. To transform the cells, 5 µl of ligation mix was added to 100 µl of competent cells and incubated for 30 min with occasional stirring. The cells were transferred to a 42 °C waterbath for 1 min and then immediately placed on ice for a further 2 min before resuspending in LB (1 ml). The cells were incubated for 1 h at 37 °C to allow expression of antibiotic resistance. Suitable samples were then plated out onto selective plates and incubated overnight at 37 °C (Methods 2.2.1).

CHAPTER 3
Purification and crystallisation of catechol 1,2-dioxygenase

3.1 Purification of catechol 1,2-dioxygenase

Catechol 1,2 dioxygenase activity in *Rhodococcus rhodochrous* NCIMB 13259 was first demonstrated by Warhurst (1993) but the enzyme was not isolated. In order to study catechol 1,2-dioxygenase in detail, a method for its purification had to be developed. Initial handling revealed catechol 1,2-dioxygenase to be stable in 50 mM Tris/HCl, pH 7.5 and so this buffer was used routinely throughout. The effectiveness of each potential purification step was monitored by recording purification factors and yields, the aim being to maximise both. Chromatographic and ammonium sulphate separations were also analysed by SDS-PAGE (Methods 2.7.1). All of the purification steps that were tried are described here, even if they were not included in the final method, which is described in Methods 2.5.

3.1.1 Growth of cells and the production of extracts

The apparent lack of catabolite repression in *R. rhodochrous* enabled Warhurst (1993) to grow the bacteria on NB before inducing catechol 1,2-dioxygenase with styrene vapour. At the outset of the present work, however, it was found to be much more convenient to induce catechol 1,2-dioxygenase, during growth on NB, with benzyl alcohol (Methods 2.2.3.2), since it is more soluble than styrene and less volatile and toxic. In order to achieve higher cell densities, NB was substituted by the much richer MMD medium (Methods 2.2.1), which gave twice the yield [8 g cells (wet weight) l^{-1}], without sacrificing the specific activity of catechol 1,2-dioxygenase, which was 0.2-0.4 units (mg protein)$^{-1}$, for both NB and MMD.

Cell harvesting and protein extraction were carried out according to Methods 2.2.4, 2.2.5 & 2.5.2.

3.1.2 Ammonium sulphate fractionation

This was one of the first purification methods to be investigated, not only because it can give cheap and effective separations, but also because it is a useful way of concentrating protein solutions. Optimal separations were achieved with 50-90% saturation $(NH_4)_2SO_4$ fractions, with 2-fold purification and yields of 60%, but the method was not included in the final preparation.

3.1.3 Ion-exchange chromatography in DEAE-Sephacel

Anion-exchange chromatography (Methods 2.5.3) was chosen as the first step in the purification. Preliminary experiments with DEAE-Sephacel resin using positive NaCl gradients showed that catechol 1,2-dioxygenase eluted between 0.2 and 0.4 M NaCl and so prior to loading, the gel resin was pre-equilibrated with

0.18 M NaCl. Unbound material was removed by washing with the same concentration of NaCl buffer until a level A_{280} baseline had been achieved. A shallow gradient of 0.18-0.4 M NaCl gradient was then used to elute catechol 1,2-dioxygenase. Figure 3.1 shows a typical elution profile of catechol 1,2-dioxygenase from a DEAE-Sephacel column.

3.1.4 Gel-permeation chromatography in Sephacryl S-200

The ion-exchange pool was concentrated to less than 0.5% of the column volume (Section 3.1.2) and loaded at 10 ml h^{-1} onto a Sephacryl S-200 column (1.6 x 72 cm) which had been pre-run at 10 ml min^{-1} overnight with 50 mM Tris/HCl, pH 7.5, 0.2 M NaCl. Fractions (2 ml) were collected and assayed for protein and catechol 1,2-dioxygenase activity (Methods 2.3.2, 2.4.1), but resulted in little if any purification and only 60% yields. It was also noted that when Mono Q FPLC (Methods 2.5.5) was done after the gel-permeation step, catechol 1,2-dioxygenase activity separated into two peaks. However, when the gel-permeation was discontinued as a purification step, Mono Q FPLC of purified catechol 1,2-dioxygenase revealed only a single catechol 1,2-dioxygenase activity peak. Gel-permeation chromatography was therefore not included in the final procedure.

3.1.5 Hydrophobic interaction chromatography in phenyl-Sepharose CL-4B

This method is based on the differential binding of proteins to hydrophobic side-chains in the presence of salts such as ammonium sulphate. A pilot study was undertaken to determine if and under what conditions catechol 1,2-dioxygenase would bind and it was found that when phenyl-Sepharose gel (0.5 ml) was equilibrated with 15% saturation $(NH_4)_2SO_4$ and catechol 1,2-dioxygenase applied, no activity could be detected in the eluate; however, when the buffer was 10% saturation $(NH_4)_2SO_4$, catechol 1,2-dioxygenase activity could be detected in the eluate. The experiment was repeated in the presence of 0.25 M NaCl with essentially the same result. On the basis of these findings, the ion-exchange pool was loaded and washed in the presence of 15% saturation $(NH_4)_2SO_4$ and eluted by a negative salt gradient. Figure 3.2 shows an elution profile of catechol 1,2-dioxygenase in phenyl-Sepharose and the method is described in Methods 2.5.4.

3.1.6 Concentration

Although chromatography in phenyl-Sepharose proved to be an excellent purification step, it generally resulted in excessive sample dilution. A concentration method was therefore sought which would bring about a rapid and efficient

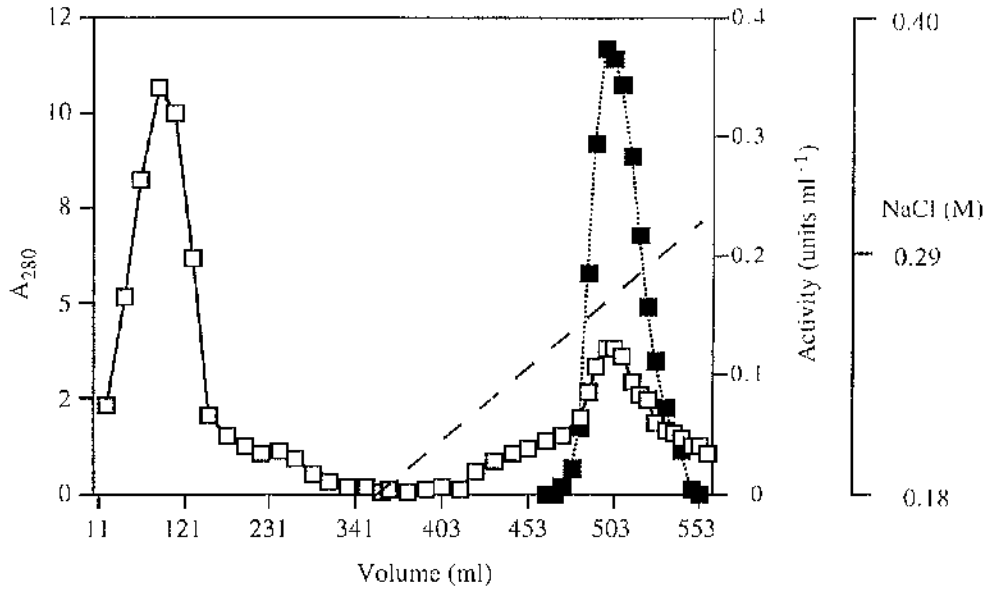

Figure 3.1 Chromatography of catechol 1,2-dioxygenase on DEAE-Sephacel

Extract (80 ml) was loaded at 20 ml h^{-1} onto a DEAE-Sephacel column (10 x 2.6 cm) and washed with 7 columns of Buffer A (Methods 2.5.3) at 60 ml h^{-1}. Catechol 1,2-dioxygenase was eluted with a positive gradient (300 ml) of Buffer A-Buffer B at a flow rate of 20 ml h^{-1} (Methods 2.5.3). Fractions were collected and assayed for A_{280} (□) and activity (■) (Methods 2.3.1 and 2.4.1 respectively).

The NaCl gradient (- - -) was calculated.

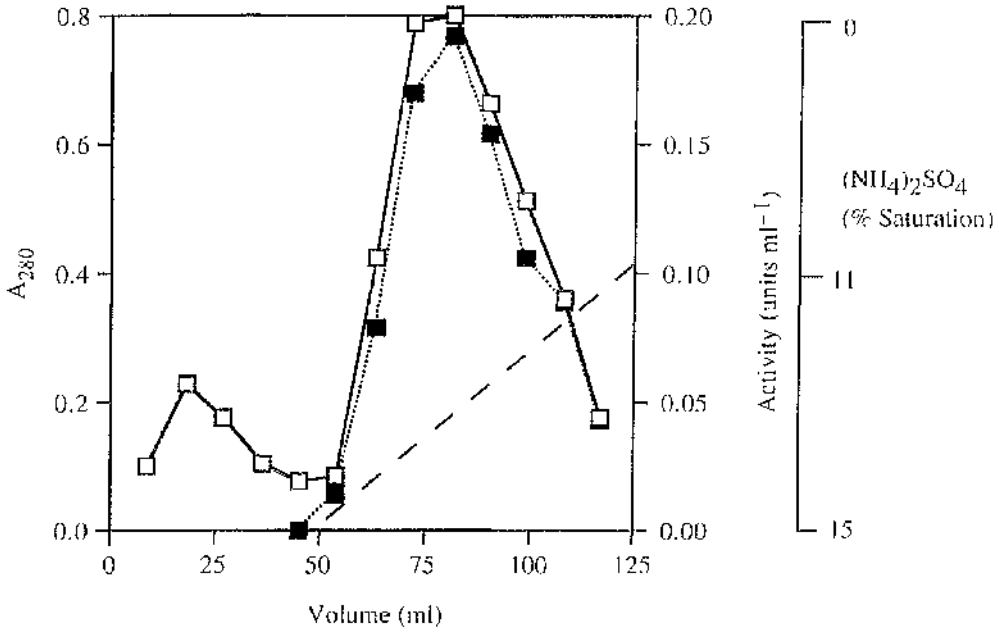

Figure 3.2 Chromatography of catechol 1,2-dioxygenase on phenyl-Sepharose

The DEAE-Sephacel pool (100 ml) was loaded at 60 ml h^{-1} onto a phenyl-Sepharose column (10 x 2.6 cm) which had been pre-equilibrated with Buffer C (Methods 2.5.4). The column was then washed at 20 ml h^{-1} for 50 ml before applying a negative gradient (300 ml) of Buffer C-Buffer A (Methods 2.5.4). Fractions (5 ml) were assayed for both A_{280} (□) and activity (■) (Methods 2.3.1 and 2.4.1 respectively).

The $(NH_4)_2SO_4$ gradient (---) was calculated.

reduction in the volume. A 200 ml Amicon ultra-filtration cell (Amicon Corp, Mass., USA) was tried initially, but this consistently resulted in up to 40% losses of activity. Dialysis was also tried, but this too resulted in large losses of activity. Finally, a two-step concentration procedure was developed which was very fast and resulted in much better recoveries of catechol 1,2-dioxygenase. A 10-20-fold reduction in volume was achieved in under 30 min using phenyl-Sepharose gel (Methods 2.5.6a) with complete recovery of catechol 1,2-dioxygenase. Further volume reduction was achieved using Centricon-30 ultra-filtration units (Amicon) (Methods 2.5.6b), which resulted in protein concentrations of 17-45 mg ml^{-1}.

3.1.7 Storage of purified catechol 1,2-dioxygenase

Catechol 1,2-dioxygenase tolerated a variety of different buffer conditions and dilution states for instance, relatively dilute, pooled phenyl-Sepharose fractions were kept at –20 °C and after one month displayed no loss in activity. Concentrated catechol 1,2-dioxygenase was also kept in 10-50 mM Tris/HCl, pH 7.5 for several weeks at 4 °C, again with no loss of activity.

3.1.8 The purification procedure

The results from one purification are summarised in Table 3.1 and are typical of four separate purifications (Table 3.2). From 15-20 g of cells, catechol 1,2-dioxygenase was purified more than 40-fold and gave at least 10 mg of protein, representing more than 40% of the starting enzyme (Table 3.1). The purification has been repeated four times on a larger scale (Table 3.2) with similar results, giving up to 60 mg of catechol 1,2-dioxygenase from 90-140 g (wet weight) of cells. In order to handle the larger amount of material, the ion-exchange bed-volume was increased to 12 x 5 cm and was operated throughout at 80 ml h^{-1} with the larger gradient volume of 1 l. The phenyl-Sepharose bed-volume was increased to 17 x 2.6 cm but was operated at the same flow rate as was used in the small scale preparation.

The purification was monitored by SDS-PAGE (Methods 2.7.1).and by this criterion catechol 1,2-dioxygenase was adjudged to be virtually homogeneous. Figure 3.3 shows a typical gel of the purification, resulting in a single band after phenyl-Sepharose chromatography, with minor, faster moving bands sometimes visible after overloading. Purified samples were also analysed on native PAGE gels (Methods 2.7.3) and showed a single band which co-migrated with catechol 1,2-dioxygenase activity when either the clarified extract or purified samples were run (Figure 3.4).

Step	Volume (ml)	Total activity (units)	Total protein (mg)	Activity (units mg^{-1})	Yield (%)	Purification (fold)
Extract	80	339	968	0.3	100	1
DEAE-Sephacel	100	243	59	4.3	72	13
Phenyl-Sepharose	36	149	N.D.	N.D.	N.D.	N.D.
After concentration on phenyl-Sepharose	9	146	10	14.3	43	42

Table 3.1 Purification of catechol 1,2-dioxygenase

The purification was carried out as described in Methods 2.5. Activity and protein concentration values were determined after each step (Methods 2.4.1 and 2.3.2). The protein concentration of the phenyl-Sepharose pool was not determined (N.D.).

Preparation number	Specific activity (units mg^{-1})	Yield (%)	Purification (fold)
1	6	36	25
2	9	32	43
3	20	42	50
4	14	43	42
5	14	46	28
6	11	76	37
7	9	53	43
8	14	46	28
Mean	12	47	37
SD	4.3	14	9

Table 3.2 Results from eight different purifications of catechol 1,2-dioxygenase from *Rhodococcus rhodochrous*

Purifications of catechol 1,2-dioxygenase were carried out as described in Methods 2.5. Activity and protein concentration values used in the calculation of specific acivity, yield and purification (fold), were determined after each step (Methods 2.4.1 and 2.3.2). Preparations 1 to 4 were done on a small scale, where approximately 10 mg of catechol 1,2-dioxygenase was obtained (Table 3.1). Preparations 5 to 8, however, were done at a later date and on a larger scale (Section 3.1.8).

The mean and percentage SD for all eight preparations were calculated and are shown beneath each column.

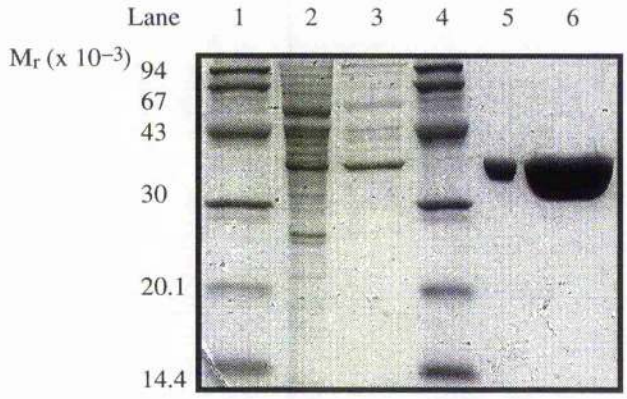

Figure 3.3 SDS-PAGE of samples taken during the purification of catechol 1,2-dioxygenase

Catechol 1,2-dioxygenase was purified according to Methods 2.5. Fractions were analysed using a 12% polyacrylamide separating gel with a 5% stacking gel and treated according to Methods 2.7.1. Lane 1, M_r markers; Lane 2, 10 μg of the clarified extract; Lane 3, 2 μg of the DEAE-Sephacel pool; Lane 4, M_r markers; Lane 5, 2 μg of the concentrated phenyl-Sepharose pool; Lane 6, 7 μg of the phenyl-Sepharose pool. Gels were stained and destained according to Methods 2.7.4.

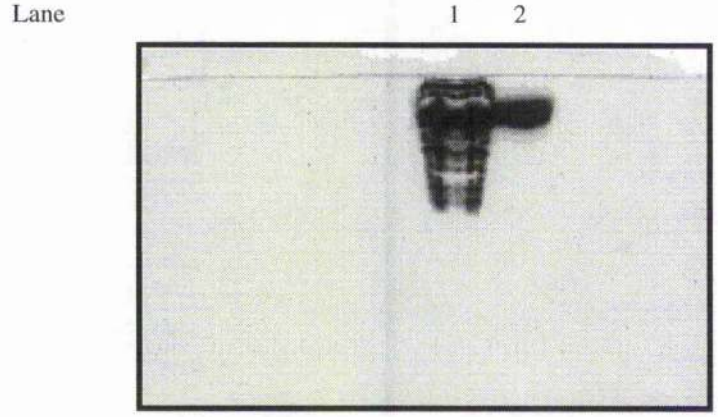

Figure 3.4 Native PAGE of the clarified extract and the purified catechol 1,2-dioxygenase

Clarified extract (Methods 2.5.3) and purified catechol 1,2-dioxygenase (Methods 2.5) were analysed using an 8% separating gel and 3% stacking gel under native conditions (Methods 2.7.3). Lane 1, 10 μg of clarified extract; Lane 2, 3μg of the concentrated phenyl-Sepharose pool. The gel was stained for protein according to Methods 2.7.4.

3.2 Crystallisation of catechol 1,2-dioxygenase

To date, the protocatechuate 3,4-dioxygenase from *Pseudomonas putida* ATCC 23975 (formerly classified as *Pseudomonas aeruginosa*) is the only non-haem iron intradiol-cleavage enzyme whose three-dimensional structure has been solved by X-ray crystallography (Ohlendorf *et al.*, 1988; Ohlendorf *et al.*, 1994). Protocatechuate 3,4-dioxygenase has also been crystallised from *Acinetobacter calcoaceticus* (Vetting *et al.*, 1993) and *Brevibacterium fuscum* (Earhart *et al.*, 1994*a*). The only other member of the non-haem iron intradiol-cleavage enzymes which has been crystallised and characterised is the catechol 1,2-dioxygenase ($\alpha\alpha$ isozyme) from *Pseudomonas arvilla* C-1 (Earhart *et al.*, 1994*b*). Although each of the aforementioned enzymes catalyse analogous reactions in the biodegradation of aromatic compounds, the structural basis for their divergent substrate specificities and quaternary structures (Table 1.2) is not known. As a first step towards addressing this conundrum, a preliminary crystallographic investigation of *Rhodococcus* catechol 1,2-dioxygenase was undertaken.

3.2.1 Identifying initial crystallisation conditions

Initial crystallisation conditions were identified with the sitting-drop vapour diffusion technique (Methods 2.6.1) using the fast screening method of Jancarik & Kim (1991). Soft crystals were obtained after 4 days with condition number 6 [0.2 M $MgCl_2$, 0.1 M Tris/HCl, pH 8.4, 30% (w/v) PEG 4 000] and condition number 28 from the screen [0.2 M sodium acetate, 0.1 M cacodylate, pH 6.8, 30% (w/v) PEG 8 000]. Since both protocatechuate 3,4-dioxygenase and catechol 1,2-dioxygenase have been crystallised in the presence of alkali earth metal ions (Table 3.3), crystal growth with condition number 6 was refined by varying the concentration of $MgCl_2$ and PEG 4 000. After several days, orange/brown, twinned, cuboidal crystals of about 0.1 mm were obtained using 15% (w/v) PEG 4 000, 0.35 mM $MgCl_2$ and Tris/HCl, pH 7.5 (Figure 3.5). The crystals stained readily with Coomassie Blue and disintegrated with the merest touch of a needle, indicating that they were protein and not salt. Crystal growth was also monitored in an analogous experiment, substituting $CaCl_2$ for the $MgCl_2$ used in the previous experiment. After 3-4 days much larger, orange/brown, twinned rectangular plates, approximately 0.5 mm down the longest axis, were obtained (Figure 3.6). In a further experiment to optimise growth conditions, crystals were grown from pH 6.5-8.2. Below pH 7.0, only microcrystals were observed, but between pH 7.0-7.4 the crystals increased in size and then appeared to remain the same, both in terms of size and quality, with each incremental rise in the pH.

Enzyme	Organism	Crystallisation conditions	Space group	Unit cell dimensions (Å)	Reference
PCD	*Pseudomonas putida* ATCC 23975 (formerly *Pseudomonas aeruginosa*)	20-75 mg protein ml^{-1}, 1.5-1.8 M $(NH_4)_2SO_4$, 50 mM Tris/HCl, pH 8.4	Monoclinic *C*2	Not given	Ohlendorf *et al.* (1994)
PCD	*Brevibacterium fuscum*	20-25 mg protein ml^{-1}, 50 mM Hepes, pH 7.0-7.8, 0.2 M $CaCl_2$, 5 mM β-mercaptoethanol, 20-26% (w/v) PEG 400	Triclinic *P*1	a=96.1 b=97.2 c=118.1	Earhart *et al.* (1994*a*)
PCD	*Acinetobacter calcoaceticus* BD413	10 mg protein ml^{-1}, 1.8 M $(NH_4)_2SO_4$, 20 mM β-mercaptoethanol 30 mM Tris/HCl, pH 7.3	Cubic *I*23	a=b=c=145.5	Vetting *et al.* (1993)
CTD	*Pseudomonas arvilla* C-1	15 mg protein ml^{-1}, 50 mM Tris/malate, pH 7.6, 0.2 M magnesium acetate, 4% (w/v) PEG 4 000	Orthorombic $C222_1$	a=62.7 b=71.5 c=71.5	Earhart *et al.* (1994*b*)

Table 3.3 The crystallisation of non-haem iron intradiol cleavage enzymes

Abreviations used are as follows; PCD, protocatechuate 3,4-dioxygenase; CTD, catechol 1,2-dioxygenase (αα isozyme).

Figure 3.5 Refinement of crystal growth conditions

Cuboidal, twinned crystals of catechol 1,2-dioxygenase of approximately 0.1 mm down the longest axis, appeared after several days. The crystals were grown using the sitting drop vapour diffusion method in 15% PEG 4 000, 0.35 M $MgCl_2$ and 25 mM Tris/HCl, pH 7.5 (Methods 2.6.1). The protein (4 µl of 25 mg ml^{-1}) was mixed with 4 µl of precipitant solution and left to equilibrate at 18 °C, over 1 ml of precipitant solution.

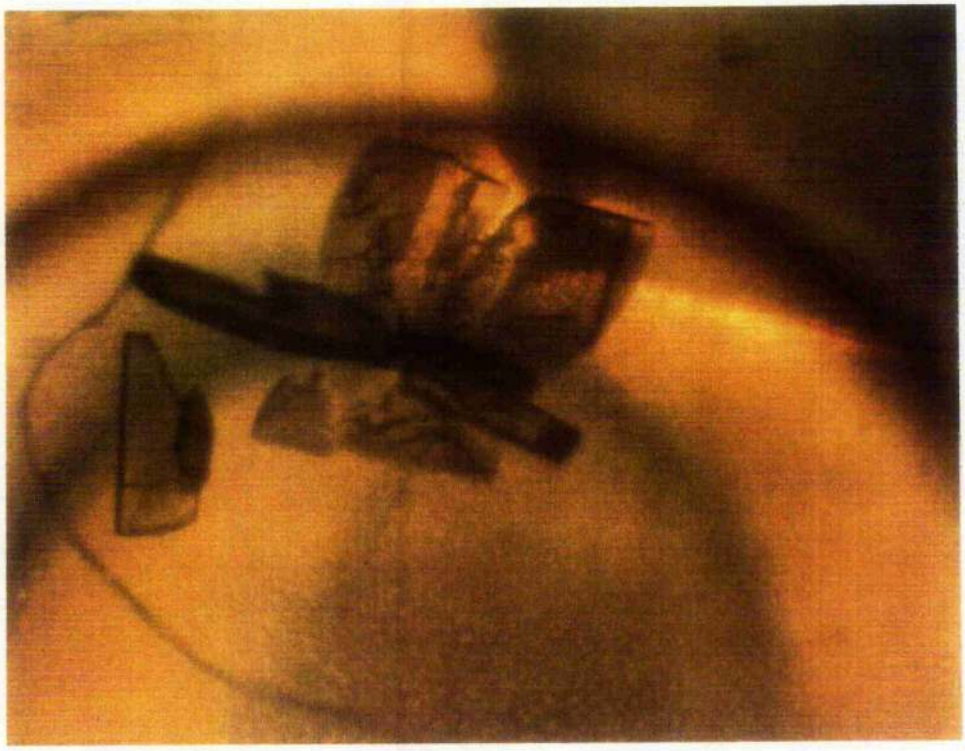

Figure 3.6 Larger crystals obtained with calcium chloride

Rectangular orange/brown plates of catechol 1,2-dioxygenase, approximately 0.5 mm down the longest axis, appeared after several days. The crystals were grown using the sitting drop vapour diffusion method in 15% PEG 4 000, 0.33 M $CaCl_2$ and 25 mM Tris/HCl, pH 7.5 (Methods 2.6.1). The protein (4 µl of 25 mg ml^{-1}) was mixed with 4 µl of precipitant solution and left to equilibrate at 18 °C, over 1 ml of precipitant solution.

3.2.2 Refinement of the crystal growth conditions

Since crystal growth may be adversely affected by trace impurities in the protein solution, catechol 1,2-dioxygenase was subjected to an extra chromatographic separation step to see if this would improve crystal formation. The purified catechol 1,2-dioxygenase (Methods 2.5) was applied to a Mono Q FPLC ion exchange column and eluted by increasing the salt gradient (Methods 2.5.5). In terms of the specific activity (units mg^{-1}), the protein which was pooled from the column appeared no different from the material which was loaded; however, sitting drop trays were prepared (Methods 2.6.1) with the desalted and concentrated pool (Methods 2.5.6) using 15% PEG 4 000, 0.33 M $CaCl_2$ 25 mM Tris/HCl, pH 7.5. Larger crystals (1 mm) were obtained after 1 wk, but they were also twinned, albeit less severely than before. Reducing agents such as dithiothreitol (1-5 mM) or β-mercaptoethanol (1-5 mM) were added to precipitant solutions, but led to no improvement in crystal quality. Finally, the Mono Q-purified catechol 1,2-dioxygenase was mixed with precipitant, but before any crystal growth, each well was seeded using a human hair (Methods 2.6.2). This resulted in single rectangular plates of 0.5 mm in the longest axis. The crystals were easily distinguishable from those crystals which had grown adventitiously, not only because they were not twinned, but also because they grew in a line which clearly followed the path of the hair (Figure 3.7). Crystals were also obtained from rhodococcal catechol 1,2-dioxygenase which had been cloned and expressed in *Escherichia coli* XL2-Blue pPDS (Section 5.7). In terms of size, quality and space group, the crystals obtained from *E. coli* XL2-Blue pPDS appeared identical to those obtained from *R. rhodochrous*, however, crystals obtained using catechol 1,2-dioxygenase which had been over-expressed in *Escherichia coli* JM109 (DE3) pPDSX4CTD (Section 5.7) produced only twinned needles. Cross-seeding, that is using seeds obtained from the wild-type protein, did not yield crystals.

3.2.3 Preliminary X-ray characterisation

A partial native data set consisting of 400 frames of 0.25 ° rotation were collected from a single crystal at room temperature (Methods 2.6.3). The monoclinic crystals are in space group C2 with unit cell dimensions (Å), a=111.9, b=78.1, c=134.6, β=100 °, with four subunits (two dimers) in the asymmetric unit (Mathews, 1968). Reduction of the diffraction data yielded 10 897 unique reflections to a resolution of 3.5 Å (70% complete) (Methods 2.6.3).

Figure 3.7 Improved crystal growth with seeding

After 7 days the crystals were seeded using human hair (Methods 2.6.2). Rectangular orange/brown plates of catechol 1,2-dioxygenase, approximately 0.5 mm down the longest axis, appeared after several days. The crystals were grown using the sitting drop vapour diffusion method in 15% PEG 4000, 0.33 M $CaCl_2$ and 25 mM Tris/HCl, pH 7.5 (Methods 2.6.1). The protein (4 µl of 25 mg ml^{-1}) was mixed with 4 µl of precipitant solution and left to equilibrate at 18 °C, over 1 ml of precipitant solution.

3.3 Discussion

3.3.1 Purification

Catechol 1,2-dioxygenase was first purified on a small scale, when upto 10 mg of purified protein was obtained. It was subsequently purified on a much larger scale. The combined SD values, for small and large scale purfications (Table 3.2), are unexpectedly large but are probably derived in part from the large experimental errors accrued during the very first small scale purifications (Table 3.2).

The specific activities and levels of expression of catechol 1,2-dioxygenase from *R. rhodochrous* (Table 3.1) are similar to those that have been reported for other bacterial catechol 1,2-dioxygenases (Patel *et al.*, 1976; Dorn & Knackmuss, 1978*a*; Aoki *et al.*, 1984*a*; Nakai *et al.*, 1990; Broderick & O'Halloran, 1991), and the degree of purification required to give homogeneous enzyme is approximately the same. A similar purification procedure to the one reported in this thesis, using ion-exchange and hydrophobic interaction HPLC, was used to purify both catechol 1,2-dioxygenase and chlorocatechol 1,2-dioxygenase from *Pseudomonas putida* B13 (Ngai *et al.*, 1990). These authors separated and partially purified catechol 1,2-dioxygenase and chlorocatechol 1,2-dioxygenase with ion-exchange chromatography and then purified each enzyme to homogeneity by hydrophobic interaction chromatography. Both enzymes were the last to be eluted from an $(NH_4)_2SO_4$ gradient, in contrast with the behaviour of catechol 1,2-dioxygenase from *R. rhodochrous*, which is intermediate in hydrophobicity (Figure 3.2). In many respects, catechol 1,2-dioxygenase from *R. rhodochrous* behaved in the same way as chlorocatechol 1,2-dioxygenase from *Rhodococcus erythropolis* 1CP (Maltseva *et al.*, 1994), which also eluted at above 0.2 M NaCl on ion-exchange chromatography and at 0.65 M $(NH_4)_2SO_4$ on hydrophobic interaction chromatography. Maltseva *et al.* (1994) also reported some modification of chlorocatechol 1,2-dioxygenase during its purification, since during the last Mono Q FPLC step chlorocatechol 1,2-dioxygenase separated into two peaks which were not present at the start of the purification but which they later designated chlorocatechol 1,2-dioxygenase-1 and chlorocatechol 1,2-dioxygenase-2. Catechol 1,2-dioxygenase from *R. rhodochrous* was also modified after gel-permeation chromatography (Section 3.1.4) and this was presumably due to an alteration in charge caused by deamidation of asparagine and glutamine residues.

An early attempt to purify the catechol 1,2-dioxygenase and chlorocatechol 1,2-dioxygenase from *P. putida* B13 resulted in only a 1.4-fold purification of chlorocatechol 1,2-dioxygenase. The low purification factor was thought to be because of its instability during chromatography and dialysis (Dorn &

Knackmuss, 1978*a*). The instability of chlorocatechol 1,2-dioxygenase was also noted by Broderick & O'Halloran (1991), who found that during the purification of chlorocatechol 1,2-dioxygenase from *Pseudomonas putida* pA27, the enzyme purified at pH 7.5 had a 40% lower specific activity but similar iron content to the enzyme purified at pH 8. Ngai & Ornston (1988) purified the same enzyme from *P. putida* pA27 and reported a significantly lower specific activity, later claimed to be due to the effects of ethylenediamine and β-mercaptoethanol in the purification buffer (Broderick & O'Halloran, 1991).

3.3.2 Crystallisation

Several factors may affect crystal quality, including the rate of crystal growth, the internal order of the initial nucleus and the protein purity (Stura & Wilson, 1992). Also, proteins which appear homogeneous by SDS-PAGE, may nevertheless contain trace impurites which can inhibit crystal growth and some workers have achieved much better crystallisation by subjecting an apparently pure protein to an extra purification step (Gourley D. & Emsley P. personal communication). Catechol 1,2-dioxygenase was therefore purified further with Mono Q FPLC (Methods 2.5.5) and resulted in some improvement in crystal quality. Although the number of single crystals was still extremely small, the few that were obtained were subsequently used in seeding trials and resulted in significant quantities of single, well formed rectangular plates (Figure 3.7).

Giege & Ducroix (1992) have stated that, in order to generate high-quality monocrystals suitable for X-ray crystallography, proteins not only have to be pure in terms of lack of contaminants, they also have to be conformationally pure. Thus, denatured proteins or proteins with stuctural microheterogeneities may adversely affect crystal growth more than do unrelated molecules, especially when the structural heterogeneities concern domains involved in crystal packing. Thus, the difficulty which was found in crystallising the overexpressed catechol 1,2-dioxygenase could have been due to either variation in secondary structure brought about by improper folding, or variation in primary structure brought about by proteolysis and/or errors in transcription or translation. Whether the known *N*-terminal modification (Section 5.7) of some polypeptides had any bearing on the crystallisation of the over-expressed catechol 1,2-dioxygenase would depend largely on whether the methionine residue is in a region involved in crystal packing and whether this residue is inside or outside the protein.

CHAPTER 4
Characterisation of catechol 1,2-dioxygenase

4.1 Introduction

Characterisation of the catechol 1,2-dioxygenase of *Rhodococcus rhodochrous* NCIMB 13259 has been limited to a preliminary kinetic comparison of the enzyme in partially purified extracts of bacteria that had been grown on styrene or benzyl alcohol (McIntosh, 1993). A detailed characterisation of the purified catechol 1,2-dioxygenase was therefore undertaken so that it could be compared with other non-haem iron intradiol-cleavage enzymes.

4.2 Physical and chemical characterisation

4.2.1 The subunit M_r of catechol 1,2-dioxygenase

The subunit M_r value was determined under denaturing conditions using SDS-PAGE (Methods 2.7.1). Gels were calibrated with M_r standards, stained, destained (Methods 2.7.4) and then scanned with a laser densitometer (Methods 2.7.5). Mobilities (arbitrary units) were then plotted as a function of log M_r. A representative plot of mobility against log M_r is shown in Figure 4.1. The mean M_r of catechol 1,2-dioxygenase was 39 800 (SD=635, n=5, where n equals five different batches of enzyme).

4.2.2 The native M_r of catechol 1,2-dioxygenase

The native M_r was measured on a Superose 12 FPLC column (Methods 2.5.5). A calibration curve was prepared by determining the elution volumes (V_e) of several standard proteins and then calculating the elution volume of each protein standard with respect to the elution volume of blue dextran (void volume, V_o). The calibration proteins and catechol 1,2-dioxygenase were run in at least two different groups and always in the presence of blue dextran. This was to ensure unambiguous peak assignment and hence the accurate determination of V_e/V_o for each protein. A representative calibration curve showing V_e/V_o against log M_r is shown in Figure 4.2. The estimated M_r value of catechol 1,2-dioxygenase was 120 400. Comparison of the subunit and native M_r suggests that catechol 1,2-dioxygenase is tetrameric.

4.2.3 Chemical cross-linking of the enzyme subunits

The quaternary structure of the enzyme was investigated by cross-linking the subunits with the reactive bis-imidoesters dimethylsuberimidate and dimethylpimelidate (Davies & Stark, 1970) (Methods 2.8.7) and the results analysed using sodium phosphate PAGE (Methods 2.7.2) (Figure 4.3). Gels were calibrated by using M_r standards prepared by cross-linking rabbit muscle fructose 1,6-bisphosphate aldolase, a known tetramer (Davies & Stark, 1970). A plot of mobility against log M_r is shown in Figure 4.4. Two bands of M_r 30 000

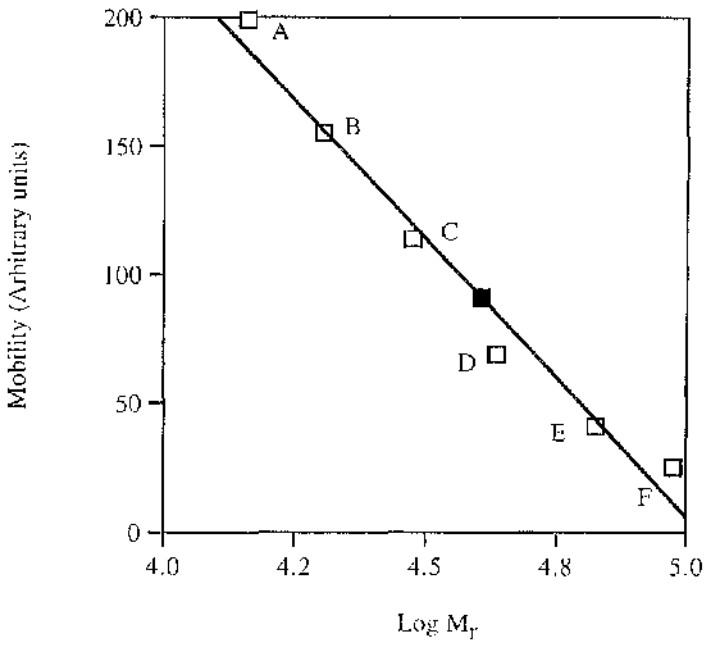

Figure 4.1 Estimation of the subunit M_r of catechol 1,2-dioxygenase using SDS-PAGE

The subunit M_r of purified of catechol 1,2-dioxygenase (Methods 2.5) was estimated on 12% SDS-polyacrylamide gels (Methods 2.7.1). The standard curve was generated with the following M_r standards:

	M_r standard	M_r
A	α-Lactalbumin	14 400
B	Soya bean trypsin inhibitor	20 100
C	Carbonic anhydrase	30 000
D	Ovalbumin	43 000
E	Bovine serum albumin	67 000
F	Phosphorylase *b*	94 000

■, Catechol 1,2-dioxygenase.

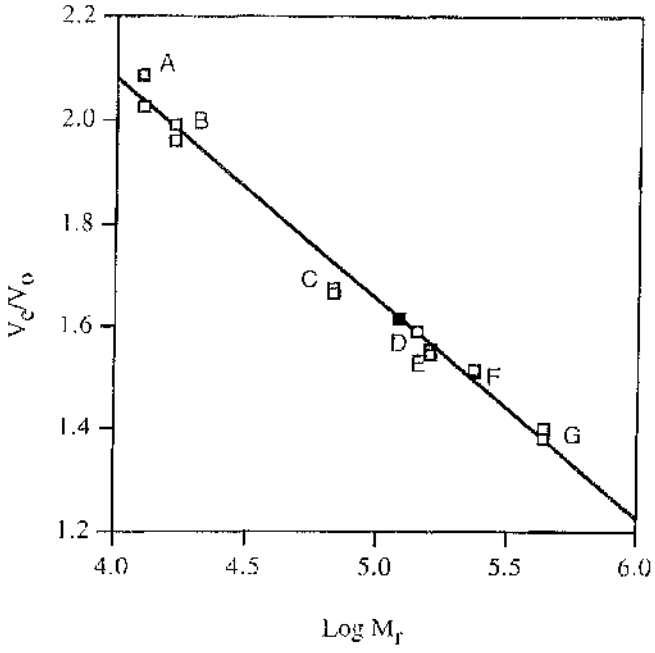

Figure 4.2 Estimation of the native M_r of catechol 1,2-dioxygenase using Superose 12 FPLC

The native M_r of purified catechol 1,2-dioxygenase (Methods 2.5) was estimated using Superose 12 FPLC equilibrated with 50 mM Tris/HCl, pH 8, 0.2 M NaCl, 0.5 mM DTT and operated at 0.5 ml min^{-1}. Protein standards were applied to the column in the presense of blue dextran (20 µg) and V_e/V_o values calculated for each protein in duplicate. Protein was measured on-line at 280 nm (Methods 2.5.5) and catechol 1,2-dioxygenase activity was assayed according to Methods 2.4.1. The standard curve was generated with the following M_r standards:

	M_r standard	M_r	Amount loaded (µg)
A	Cytochrome *c*	12 500	20
B	Myoglobin	16 700	20
C	Bovine serum albumin	67 000	50
D	Lactate dehydrogenase	140 000	25
E	Aldolase	158 000	50
F	Catalase	230 000	50
G	Ferritin	440 000	20

■, Catechol 1,2-dioxygenase.

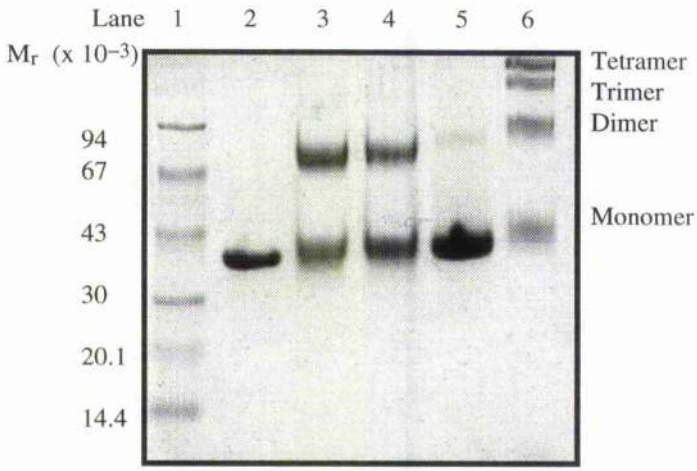

Figure 4.3 Sodium phosphate PAGE of catechol 1,2-dioxygenase cross-linked, in separate reactions, with dimethylpimelidate and dimethylsuberimidate

Catechol 1,2-dioxygenase was incubated for 1 h at room temperature with cross-linking reagent (Methods 2.8.7). Samples (3µg) were then analysed using sodium phosphate PAGE [15% (w/v) polyacrylamide] (Methods 2.7.2).

Lane	
1	M_r standards
2	Catechol 1,2-dioxygenase
3	Catechol 1,2-dioxygenase treated with dimethylpimelidate
4	Catechol 1,2-dioxygenase treated with dimethylsuberimidate
5	Aldolase
6	Aldolase treated with dimethylsuberimidate

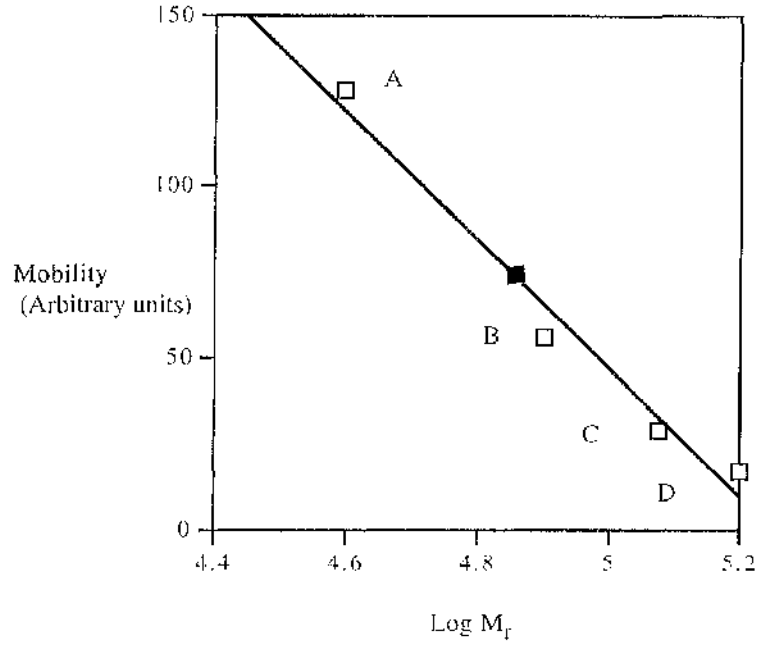

Figure 4.4 Sodium phosphate PAGE of cross-linked catechol 1,2-dioxygenase

Catechol 1,2-dioxygenase was cross-linked with dimethylsuberimidate and the products analysed using sodium phosphate PAGE (Methods 2.8.7 & 2.7.2). The gel (Figure 4.3) was calibrated using cross-linked rabbit muscle fructose 1,6-bisphosphate aldolase (Subunit M_r, 39 500) with A, B, C and D corresponding to the monomer, dimer, trimer and tetramer respectively. The M_r values of monomeric and dimeric catechol 1,2-dioxygenase were calculated by using the equation of the line ($y = -187.5x + 986.5$) and were 30 000 and 64 000 respectively.

■, Dimeric catechol 1,2-dioxygenase.

and 64 000 were present, corresponding to monomer and dimer, indicating a dimeric quaternary structure. The unmodified catechol 1,2-dioxygenase and aldolase subunits, migrated more slowly through the gel than their respective cross-linked monomeric species, with the effect being more pronounced with aldolase, presumably because of a reduction in the amount of bound SDS by the cross-linked species, or because of an effect on the shape caused by intramolecular cross-linking. Curiously, the monomeric M_r value obtained after cross-linking is very close to the subunit M_r value obtained with electrospray mass spectroscopy (Section 4.2.6).

4.2.4 Absorption spectra and the molar absorption coefficients

The absorption spectrum of catechol 1,2-dioxygenase (Figure 4.5) had maxima at approximately 220 nm and 280 nm. At higher concentrations of protein another maximum at 426 nm was also evident and this is a characteristic of all non-haem iron intradiol-cleavage enzymes (Nozaki, 1979). The molar absorption coefficient at 426 nm was 1850 M^{-1} cm^{-1} (1903, 1832) with respect to the subunit concentration (Methods 2.3.3). The calculated absorption coefficient at 280 nm (Methods 2.3.3) was 54 000 M^{-1} cm^{-1}

4.2.5 Iron quantitation and the iron:subunit stoichiometry

Iron was quantified by colorimetry and by atomic absorption spectroscopy (Methods 2.8.8). The subunit concentration was derived by combining the absorption coefficient and the M_r value obtained from the derived amino acid sequence (Figure 5.7). The results for three different enzyme preparations are summarised in Table 4.1.

4.2.6 Electrospray mass spectroscopy

Electrospray mass spectroscopy was carried out in the positive-ion detection mode (Methods 2.8.6). The deconvoluted electrospray mass spectrum (Methods 2.8.6) of catechol 1,2-dioxygenase, is shown in Figure 4.6. A major peak of M_r 31 558 and three subsidiary peaks of 31595, 31613 and 31638 are visible. Since the M_r value derived from the amino acid sequence is 31 559 (Section 5.6), the major band is attributed to the apoenzyme. The M_r differences of 37, 55 and 80, suggest that the three minor peaks may be adducts of K^+, Fe^{3+} and $2K^+$ respectively.

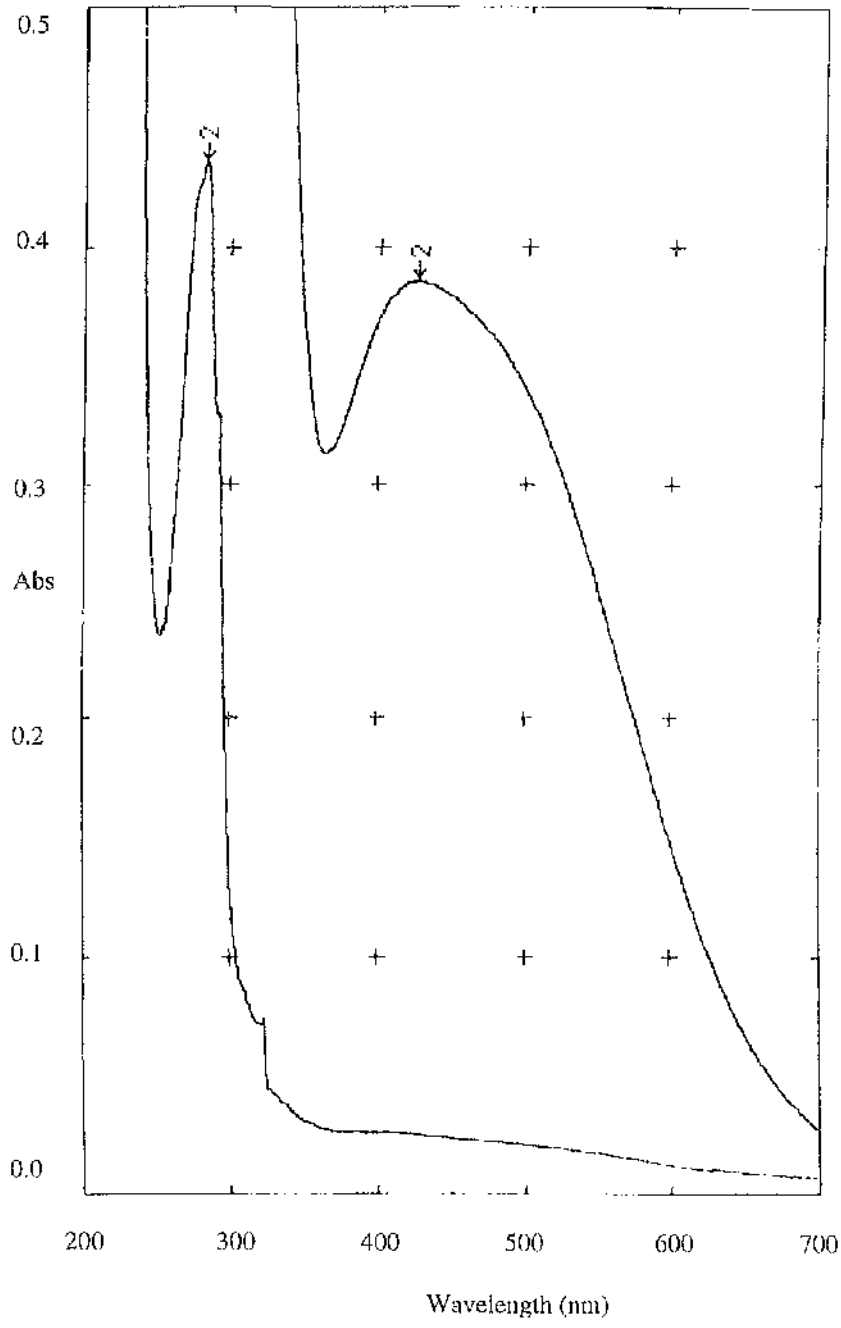

Figure 4.5 Absorption spectra of catechol 1,2-dioxygenase

The absorption spectrum of catechol 1,2-dioxygenase was recorded at two different concentrations of the enzyme (0.2 mg ml^{-1} and 6.2 mg ml^{-1} in Tris/HCl, pH 7.5, 0.18 M NaCl). On each occasion the buffered protein solution was scanned from 200-700 nm at 2 nm s^{-1} (Methods 2.1.4).

		Colorimetry		Atomic absorption spectroscopy	
Protein ($mg\ ml^{-1}$)	Specific activity ($units\ mg^{-1}$)	Fe content ($\mu g\ ml^{-1}$)	Fe/subunit stoichiometry	Fe content ($\mu g\ ml^{-1}$)	Fe/subunit stoichiometry
2.3	12.2	3.25 (3.3, 3.2)	0.75, 0.85	2.63	0.60
2.3	7.5	3.0 (3.2, 2.8)	0.65, 0.80	2.78	0.65
0.6[a]	10.0	N.D.	N.D.	0.50	0.50
0.6[b]	10.5	N.D.	N.D.	0.64	0.60

Table 4.1 Iron quantitation and the iron:subunit stoichiometry of catechol 1,2-dioxygenase

Iron concentration was determined by colorimetry and atomic absorption spectroscopy (Methods 2.8.8). Protein concentrations were determined using the molar absorption coefficient of catechol 1,2-dioxygenase (Methods 2.3.3).

a, Over-expressed from *Escherichia coli* (DE3) pPDSX4CTD (Section 5.7).

b, In this experiment *Rhodococcus rhodochrous* had been grown in medium which had been supplemented with a chelated metals solution (Methods 2.2.1).

Values in parenthesis are duplicate estimates. N.D., not determined.

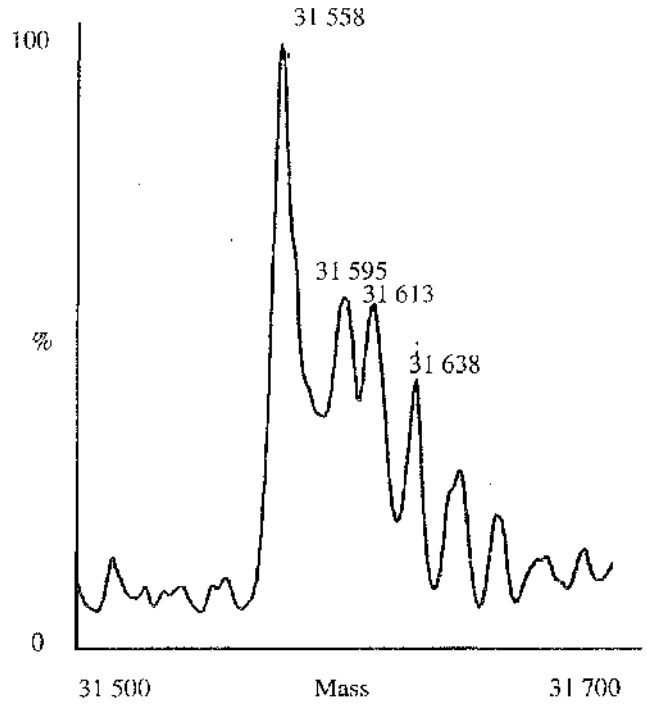

Figure 4.6 Electrospray mass spectrum of catechol 1,2-dioxygenase

Spectroscopy was carried out in the positive ion mode and the raw data were processed with the MaxEnt deconvolution software (Methods 2.8.6).

4.2.7 Isoelectric point

The isolelectric point was 4.8 and was determined from the primary amino acid sequence (Section 5.6), using the program ISOELECTRIC (Wisconsin Sequence Analysis Package, Genetics Computer Group, Madison, Wisconsin, USA).

4.2.8 *N*-Terminal sequence

The first 24 amino acids from the *N*-terminus (Methods 2.8.5) were determined and are shown in Figure 4.7 aligned with catechol 1,2-dioxygenase from another Gram-positive bacterium, *Arthrobacter* mA3 (Eck & Belter, 1993). The *N*-termini share 40% sequence identity.

4.2.9 Internal sequence data from peptidic digests

Peptides from V8 digests (Methods 2.8.3) were purified by HPLC (Methods 2.8.4) (Figure 4.8) and from one of them (V8/2) a sequence of 15 residues was obtained, which showed about 67% identity with catechol 1,2-dioxygenase from *Arthrobacter* mA3 (Eck & Belter, 1993) (Figure 4.7).

4.2.10 Sulfhdryl content

Total sulfhdryl concentration (Methods 2.8.9) of a 0.98 mg ml^{-1} solution of catechol 1,2-dioxygenase which was then diluted 6-fold in the presence of SDS, was 8.6 μM. The concentration of catechol 1,2-dioxygenase (Methods 2.3.3) was 5.2 μM, which corresponds to 1.65 sulfhdryl groups $(subunit)^{-1}$. Since the derived amino acid sequence (Figure 5.7) contains two cysteine residues, it is concluded that there are no disulfide bonds in the structure of catechol 1,2-dioxygenase.

4.3 Kinetic characterisation of catechol 1,2-dioxygenase

4.3.1 Assay reproducibility and linearity with respect to protein concentration

Assay reproducibility was determined by a measuring the activity (Methods 2.4.1) using 5 identical amounts of the purified enzyme. The SD as a percentage of the mean was found to be +/– 5.1.

The rate of catechol oxygenation was proportional to the protein concentration up to the maximum measured concentration of 2 μg ml^{-1}. All activity assays (Methods 2.4.1) contained less than this amount unless otherwise stated.

```
                1                                                   50

  N-terminus    TTTENPTAH ASGNAATDKF KS*Rv
Arthrobacter    KSTETEATAA ASGAGATARF RETKHVAAGT SKERVSALAG RVIKAINDTV

               51                                                  100

Arthrobacter    LEEKVTYDEY NALKAWLISV GETGEWPLFL DVWVEHSVEE VAKENRH..G

              101                                                  150

Arthrobacter    SKGTIEGIYY IPNAPTQNTF ATLPMRD.DK PGTPILLFQGQ VRNLAGEPLA

              151                                                  200
 V8/2 peptide                                                 ITTLKPAPYK
Arthrobacter    GAKIELWHAD DLGFYSQFAP GLPEWNLRGS IIADDQGNFQ INTMQPAPYQ

              201                                                  250
                IPADG
Arthrobacter    IPTDGACGAL IAAAGWHAWR PAHLHLKVSA PGHQLITTQL YFEGDEHVAD

              251                                  285

Arthrobacter    DIASAVKPEL VLAPTDRADG HGREVTYHFV LDPQ
```

Figure 4.7 Sequence alignment of the *N*-terminus and V8/2 peptide of catechol 1,2-dioxygenase from *Rhodococcus rhodochrous* with *Arthrobacter* mA3 catechol 1,2-dioxygenase

The V8/2 peptide was purified by HPLC from a V8 protease digest (Methods 2.8.3 & 2.8.4) (Figure 4.8). The *N*-terminus and V8/2 peptide were sequenced (Methods 2.8.5) and then aligned with *Arthrobacter* mA3 catechol 1,2-dioxygenase (Eck & Belter, 1993) using the PILEUP program (Wisconsin Sequence Analysis Package, Genetics Computer Group, Madison, Wisconsin, USA).

*, Unidentified amino acid.

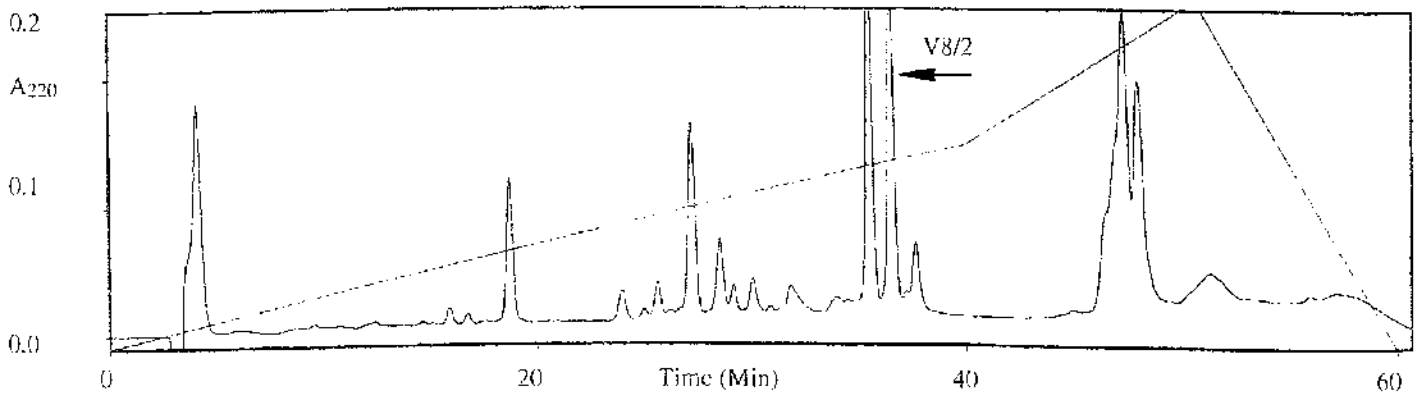

Figure 4.8 Separation of V8 protease digest by HPLC

Digests (Methods 2.8.3) were dissolved in 0.1% TFA in water and injected via a 100 µl loop onto a Waters bondpack µ18 column using a Shimadzu HPLC apparatus (Methods 2.8.4). Peptides were separated with a gradient consisting of 100% A (0.1% TFA)-50% B (0.1% TFA in 90% acetonitrile). The flow rate was 1 ml min^{-1} and the peptides were detected at 220 nm.

4.3.2 Activity with respect to ionic strength

Catechol 1,2-dioxygenase activity was measured with various concentrations of NaCl (0-0.4 M NaCl) in 100 mM Tris/HCl, pH 7.5 (Methods 2.4.1). No significant alteration in activity was observed.

4.3.3 Effects of pH value on the activity and stability of catechol 1,2-dioxygenase

The activity of catechol 1,2-dioxygenase as a function of pH was investigated by assaying enzyme activity (Methods 2.4.1) from pH 5.5 to 10.0, using solutions (100 mM) with overlapping buffering capacity. Activity increased approximately 4-fold from pH 5.0, reaching an maximal value at approximately pH 9.0 (Figure 4.9). Catechol 1,2-dioxygenase was also assayed at pH 11.0 in the presence of 100 mM Caps/NaOH buffer, but was rapidly inactivated.

No significant loss of activity was observed after incubating the enzyme at room temperature for 30 min in 100 mM Tris/HCl at pH 5, 7.5 and 9.0. However, the enzyme was rapidly inactivated in 100 mM Tris/NaOH, pH 11.2.

4.3.4 Determination of the apparent K_m and V values

The apparent K_m and V values (Methods 2.4.3) for various substrates are reported in Table 4.2. Because of the extremely high affinity which rhodococcal catechol 1,2-dioxygenase has for catechol and some of its derivatives, measuring the initial velocities with these substrates was extremely difficult. This was most apparent at sub-micromolar substrate concentrations and so preparing the assay mixture had to be done extremely quickly since linearity was only maintained for about 10 s. The situation was alleviated somewhat by the online software of the spectrophotometer which allowed the initial rate to be calculated with the 'zoom' function (Methods 2.1.4). Values of k_{cat} (s^{-1}) and k_{cat}/K_m' (M^{-1} s^{-1}) were determined (Table 4.3) using the kinetic constants obtained from the Direct Linear method (Methods 2.4.3).

4.4 Discussion

4.4.1 M_r determination and quaternary structure

The presence of a single band on SDS-PAGE (Figure 3.3) and the unambiguous *N*-terminal sequence data (Figure 4.7), suggest that the catechol 1,2-dioxygenase from *R. rhodochrous* consists of a single polypeptide species. The M_r of catechol 1,2-dioxygenase estimated from SDS-PAGE was 39 800 (Figure 4.1), approximately 20% more than the electrospray mass spectroscopy value of 31 558 (Figure 4.6). Ghosal & You (1988), Neidle *et al.* (1988), Kivisaar *et al.* (1991) and Eck & Belter (1993) noticed a similar lowering in the electrophoretic

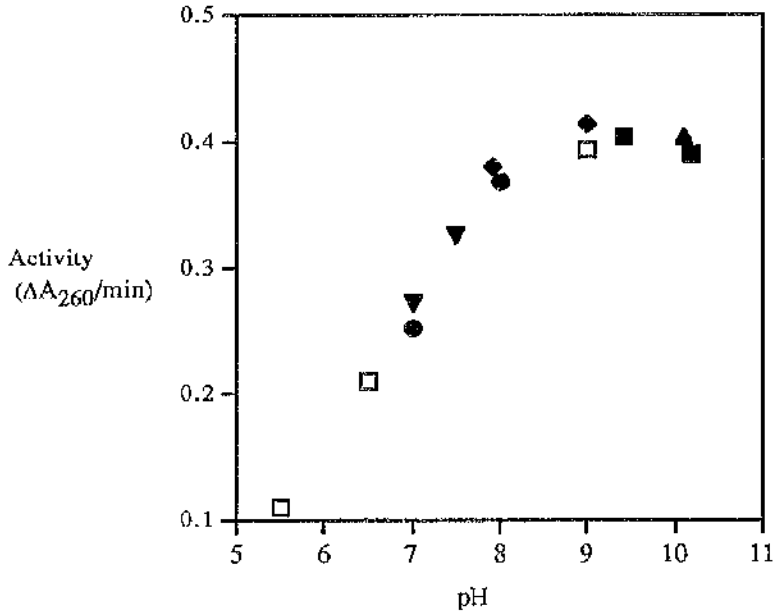

Figure 4.9 Activity of catechol 1,2-dioxygenase as a function of the pH

Catechol 1,2-dioxygenase activity was assayed according to Methods 2.4.1 in a reaction volume of 1 ml containing 100 mM buffers, 0.1 mM catechol and 1 μg of enzyme. The pH of the reaction mixture was determined just after each assay according to Methods 2.1.4

□, His/NaOH, pH 5.5-6.5 and pH 9.

▼, Mops/NaOH, pH 7-7.5.

●, Hepes/NaOH, pH 7-8.

◆, Taps/NaOH, pH7.9-9.

■, Gly/NaOH, pH 9.4-10.2.

▲, Caps/NaOH, pH 10.1.

	Direct linear		Lineweaver-Burk		Hanes-Woolf	
Substrate	Km'	V' (SD, n)	Km'	V' (SD, n)	Km'	V' (SD, n)
Catechol	1.1	19 (1.3, 4)	1.3	18 (6.5, 4)	1.1	19 (0.4, 4)
3-Methylcatechol	1.7	15 (4.2, 3)	1.9	13, 6.5, 3)	1.3	14 (4.9, 3)
4-Methylcatechol	1.2	13 (1.2, 4)	1.1	13 (1.5, 4)	1.4	14 (0.8, 4)
Protocatechuate		0*		0*		0*

Table 4.2 Comparison of the K_m' and V' values of catechol 1,2-dioxygenase obtained for catechol, 3-methylcatechol, 4-methylcatechol and protocatechuate

Initial velocities were measured in air-saturated buffers (Methods 2.4.3), using 7-8 different substrate concentrations ranging from 0.2 μM-5.0 μM. Results were analysed by the Direct linear, Lineweaver-Burk and Hanes-Woolf plotting methods using the Enzpack computer program (Williams, 1985). Values for K_m' and V' are in units of μM and units (mg protein)$^{-1}$ respectively (Methods 2.4.1).

*, No detectable activity.

Substrate	k_{cat} (s^{-1})	k_{cat}/K_m' (s^{-1} M^{-1}) (x 10^6)
Catechol	11.6	10.5
3-Methylcatechol	9.5	5.6
4-Methylcatechol	6.3	5.3
Protocatechuate	0.0	0.0

Table 4.3 Determination of k_{cat} and k_{cat}/K_m' values for catechol 1,2-dioxygenase

Values were determined using the apparent Michaelis constants calculated by the Direct Linear method (Methods 2.4.3) as reported in Table 4.2 Calculations were based on the approximation of one active site per subunit and a subunit M_r value of 31 559 (Section 5.6).

mobility of other catechol 1,2-dioxygenases during SDS-PAGE. These authors assumed that the consequent overestimate of M_r was due to the conformation of the protein and/or to the effects of the ferric ion in the enzyme. Certainly, some peculiarity in the secondary structure which prevented complete denaturation with SDS might be expected to reduce the charge/mass ratio due to the decreased amount of protein-bound detergent, hence resulting in an underestimate in the M_r value. The M_r value for the rhodococcal catechol 1,2-dioxygenase determined by electrospray mass spectroscopy has been confirmed by the M_r value of 31 559, obtained from the derived amino acid sequence (Figure 5.7) and is similar to the M_r of the catechol 1,2-dioxygenase from *Arthrobacter* mA3 which is 30 560 (Eck & Belter, 1993). The derived M_r values of other intradiol-cleavage enzymes are shown in Table 1.2. It can be seen that the chlorocatechol 1,2-dioxygenases (255-260 residues) are approximately 50 residues lighter than the catechol 1,2-dioxygenases (310-311), with the enzymes from *Arthrobacter* mA3 and *R. rhodochrous* falling between the two extremes.

The native M_r value of rhodococcal catechol 1,2-dioxygenase was determined to be 120 000 when measured on Superose 12 FPLC (Figure 4.2), which suggests that this enzyme is a tetramer. With three possible exceptions (Itoh, 1981; Aoki *et al.*, 1984*a*, *b*), all the catechol- and chlorocatechol 1,2-dioxygenases characterised to date have been dimers. Aoki *et al.* (1984*a*) noted that the catechol 1,2-dioxygenase they had isolated from *Rhodococcus erythropolis* AN-13 had a M_r value which varied with the salt concentration; when gel-permeation chromatography was done with 0.2 M NaCl at pH 8, they obtained the monomer, whereas when salt was omitted they obtained the tetramer. The same authors then measured the native enzyme using polyacrylamide gel gradients (Hedrick & Smith, 1968) and found that the enzyme migrated in the gel at pH 8.9 without polymerisation. Aoki *et al.* (1984*b*) found identical changes in oligomerisation with catechol 1,2-dioxygenase and chlorocatechol 1,2-dioxygenase from *Frateuria* ANA-18. In order to investigate the quaternary structure of rhodococcal catechol 1,2-dioxygenase directly in the present work, the subunits were chemically cross-linked with bis-imidoesters (Methods 2.8.7). Two bands were visible upon electrophoresis (Lanes 3 & 4, Figure 4.3), corresponding to the mono-mer and dimer, suggesting that catechol 1,2-dioxygenase is a dimer. With a subunit M_r of 31 559, this implies that the native M_r is 63 000. Both the gel-permeation chromatography and the cross-linking experiments were done at 0.2 M NaCl, pH 8 and at room temperature and so it is unlikely that the discrepancy was due to different oligomeric states. Bis-imido esters are known to react specifically with the amino groups of proteins (Ludwig & Hunter, 1968) and so it is possible, although hardly credible, that catechol 1,2-

dioxygenase is an isologous tetramer or 'dimer of dimers' (Hucho *et al.*, 1975; Hajdu *et al.*, 1976), with only a single available amino group per subunit. Since the cross-linking data constitute primary evidence of quaternary structure, it is more probable that under the stated conditions, catechol 1,2-dioxygenase either migrates anomolously in Superose 12 gels because it less globular than the standards used in the column calibration and/or the result is an experimental artifact of some kind. Dorn & Knackmuss (1978*a*) noticed similar discrepancies in their work on catechol 1,2-dioxygenase and chlorocatechol 1,2-dioxygenase from *Pseudomonas* B13. Initial attempts to measure the M_r from the sedimentation coefficients alone resulted in values of 66 000 and 52 000 for catechol 1,2-dioxygenase and chlorocatechol 1,2-dioxygenase respectively. Estimation of M_r by gel-permeation chromatography using Sephadex G-200, resulted in the considerably higher M_r values for catechol 1,2-dioxygenase and chlorocatechol 1,2-dioxygenase of 124 000 and 113 000 respectively. With the Stokes radii estimated from the gel-permeation chromatography and the sedimentation coefficient determined by means of density gradient centrifugation, the M_r and frictional coefficients for both catechol 1,2-dioxygenase and chlorocatechol 1,2-dioxygenase were found by applying the equations of Siegel & Monty (1966). The M_r values of catechol 1,2-dioxygenase and chlorocatechol 1,2-dioxygenase were 82 000 and 67 000 respectively and the corresponding frictional ratios were 1.45 and 1.49 respectively. Frictional ratios are a measure of ellipticity and the relatively high values obtained for catechol 1,2-dioxygenase and chlorocatechol 1,2-dioxygenase meant that gel-permeation elution volumes could not be correlated with log M_r values, since the frictional ratios differed too much from the calibration proteins. Dorn & Knackmuss (1978*a*) also estimated the M_r of catechol 1,2-dioxygenase as 48 000 by gradient gel electrophoresis. The underestimate was entirely predictable by the above criteria, since proteins come to a dead-stop only when they reach pores that are narrower than their lesser diameter (Margolis & Kendrick, 1968).

4.4.2 Iron:subunit stoichiometry

Iron quantitation by colorimetry and atomic absorption spectroscopy resulted in values which were significantly less than the expected one Fe atom per subunit stoichiometry (Table 4.1). The three protein samples that were assayed had different specific activities, but very similar protein:iron concentration ratios. Since catalysis is known to depend upon ferric iron and its amino acid ligands (Broderick & O'Halloran, 1991), the observed difference in specific activities is probably a reflection of some sample denaturation during purification. Observed Fe/subunit stoichiometries for a variety of non-haem iron intradiol cleavage

enzymes are reported in Table 1.2 Stoichiometries vary from 1:1 in the catechol 1,2-dioxygenase from *A. calcoaceticus* (Patel *et al.*, 1976) to 0.5 Fe/subunit in the *Pseudomonas* enzymes (Broderick & O'Halloran, 1991). The protocatechuate 3,4-dioxygenase from *Pseudomonas putida* ATCC 23975 (formerly classified as *Pseudomonas aeruginosa*) was originally thought to be an octomer of ($\alpha_2\beta_2$ Fe^{3+}) catalytic units (Yoshida *et al.*, 1976), but was later reconstituted with additional iron (Bull & Ballou, 1981). The treatment resulted in an increase in the visible absorption spectra, specific activity and the attainment of the theoretical $\alpha\beta Fe^{3+}$ stoichiometry. In the same vein, Broderick & O'Halloran (1991) tried several methods to increase the iron content and specific activity of chlorocatechol 1,2-dioxygenase: Fe^{2+} and Fe^{3+} were added to the assay buffer; Fe^{3+} was incubated with the enzyme, followed by gel-permeation chromatography; the enzyme was anaerobically incubated with Fe^{2+} and various reducing agents, followed by gel-permeation chromatography and air-oxidation; but none of the methods resulted in an increase in either iron content or specific activity. In the present work, an attempt was made to increase the content of iron by incorporating a chelated metals solution into the growth medium (Methods 2.2.1), but without any significant change in either iron content or specific activity (Table 4.1). Broderick & O'Halloran (1991) posited that alternate metal ions such as manganese or zinc occupied the second binding site in place of iron, but analysis of their enzyme revealed Zn/dimer and Mn/dimer ratios of less than 0.05 and 0.009, respectively. More recently, Maltseva *et al.* (1994) purified and characterised a chlorocatechol 1,2-dioxygenase from *R. erythropolis* 1CP with the unusual stoichiometry of FeMn/ 64 000 dimer. Certainly, Mn^{2+} is an analogue of Fe^{2+} and there is at least one known Mn^{2+} dependent extradiol dioxygenase isolated from *Arthrobacter globiformis* CM-2, known from the sequence data to be a member of the main Fe^{2+} dependent family of enzymes (Boldt *et al.*, 1995), but whether Mn^{2+} is catalytically active in an enzyme known to depend upon the ferric oxidation state remains to be seen, since during the initial stages of the purification, these workers used buffers supplemented with 2 mM $MnSO_4$ (Maltzeva *et al.*, 1994). Another possibility, is that the binding of Fe^{3+} to one subunit, reduces the metal ion affinity of the other iron-binding site, but this presupposes that the iron is incorporated after the protein has folded, since if it were incorporated during folding, all enzymes would have their full quota of iron. The protocatechuate 3,4-dioxygenase from *P. putida* is the only non-haem iron intradiol-cleavage enzyme whose 3-dimensional structure is known [$(\alpha\beta\ Fe^{3+})_{12}$] (Ohlendorf *et al.*, 1988; Ohlendorf *et al.*, 1994). The enzyme should have 12 Fe atoms per mole of enzyme; however, in no case have more than 8 atoms per mole been reported and this is the case for both the wild-type and the cloned proteins,

so it seems that despite the structural equivalence of the metal binding sites, the quaternary structure in solution leads to a static or dynamic inequivalence (Frazee *et al.*, 1993). Concentrated solutions of rhodococcal catechol 1,2-dioxygenase are red due to an absorption maxima at 426 nm (Figure 4.5), which is typical of all known non-haem iron intradiol-cleavage enzymes (Table 1.2). The mean molar absorption coefficient at this wavelength was 1 850 M^{-1} cm^{-1} with respect to the subunit concentration (Section 4.2.4), or 3 700 M^{-1} cm^{-1} with respect to the native protein concentration (assuming a M_r value of 63 000). Patel *et al.* (1976) reported a molar absorption coefficient of 2 860 M^{-1} cm^{-1} with respect to the native catechol 1,2-dioxygenase from *A. calcoaceticus*, but since this enzyme has 1 Fe atom per subunit (Figure 1.2) the difference in values is unexpectedly large, given that the Fe:subunit stoichiometry in the rhodococcal catechol 1,2-dioxygenase was closer to 0.5 Fe atoms per subunit (Table 4.1). In fact the molar absorption coefficient at 426 nm obtained for the rhodococcal enzyme is much closer to that reported for the chlorocatechol 1,2-dioxygenase from *P. putida* which has 0.5 Fe atoms per subunit (Table 1.2) (Broderick & O'Halloran, 1991).

Electrospray mass spectroscopy has recently emerged as a powerful tool, not only in the accurate determination of M_r values (Fenn *et al.*, 1989), but also in the analysis of oligomeric proteins containing labile Fe-S clusters (Petillot *et al.*, 1993). Since the Fe-S bonds are weaker than the other bonds involved in proteins, obtaining the intact holoenzyme requires milder operating conditions. During the present work, electrospray mass spectroscopy was operated in the positive ion mode by incorporating trifluoroacetic acid in the carrier solvent (Methods 2.8.6). Under the stated conditions, only the monomeric form of catechol 1,2-dioxygenase was seen and this was predominantly in the apoenzyme form (Figure 4.6). During their work on the homodimeric ferredoxin, a protein which contains 1 Fe-S cluster per subunit, Petillot *et al.* (1995) observed only the monomeric apoprotein when they carried out electrospray mass spectroscopy in the positive ion mode. Since this metalloenzyme has a sufficiently low pI, they were able to use electrospray mass spectroscopy in the negative ion mode at neutral pH. By varying the nature of the carrier solvent, they were able to get either dimeric holoenzyme and/or a mixture of the monomeric apoprotein and the dimeric holoenzyme. No signal corresponding to the dimeric apoprotein was observed and so these workers assumed that the loss of quaternary structure preceeded loss of the Fe-S clusters.

4.4.3 Kinetics

Rhodococcal catechol 1,2-dioxygenase was assayed with 0-0.4 M NaCl in 100 mM Tris/HCl, pH 7.5 according to (Methods 2.4.1). Neither activation nor

inhibition was observed, in contrast to the protocatechuate 3,4-dioxygenase from *Pseudomonas putida*, which was inhibited (competitively with protocatechuate) by chloride ions (Bull & Ballou, 1981). Also, in some of the earliest kinetics studies done on pseudomonad catechol 1,2-dioxygenase, some dependence upon the ionic strength, was observed (Hayaishi *et al.*, 1957), but again this was not experienced with the rhodococcal catechol 1,2-dioxygenase.

Catechol 1,2-dioxygenase was optimally active at pH 9.0 (Figure 4.9), similar to the catechol 1,2-dioxygenases that were isolated from *Rhizobium leguminosarum* biovar *viceae* USDA 2370 (Chen & Lovell, 1990) and *Rhizobium trifolii* (Chen *et al.*, 1985). In contrast, the catechol and chlorocatechol 1,2 dioxygenases from pseudomonads (Nakai *et al.*, 1990; Broderick & O'Halloran, 1991; Bhat *et al.*, 1993) and the enzymes from *A. calcoaceticus* (Patel *et al.*, 1976) and *Frateuria* ANA-18 (Aoki *et al.*, 1984*b*), are optimally active at pH 7.5-8.0. At pH 11.0, the rhodococcal enzyme was rapidly inactivated and so it was not possible to obtain an initial rate. The broad pH stability of rhodococcal catechol 1,2-dioxygenase contrasts with other catechol- and chlorocatechol 1,2-dioxygenases, which tend to be unstable at below about pH 7.

Steady-state kinetics experiments which were designed to measure catechol 1,2-dioxygenase activity with respect to different substrates, were hampered because of the difficulty in acquiring halo- and alkyl-substituted catechols. A more limited experiment did show, however, that rhodococcal catechol 1,2-dioxygense did not differ significantly in activity or binding affinity with catechol, 3- and 4-methylcatechol as substrates (Table 4.2). The very low K_m values obtained for the purified rhodococcal catechol 1,2-dioxygenase for each of these substrates substantiates the results obtained with partially purified extracts (McIntosh, 1993), although in some of the latter experiments, the maximum velocity with catechol was somewhat higher than with the 3- and 4-methyl-derivatives. The lack of activity with protocatechuate was no suprise, since as far as it is known, no chlorocatechol 1,2-dioxygenase or catechol 1,2-dioxygenase will oxygenate this substrate. Conversely, the pseudomonad protocatechuate 3,4-dioxygenase is typical in not degrading catechols to any great extent. The structural basis for this substrate exclusivity is not known, since the three-dimensional structure of only protocatechuate 3,4-dioxygenase is known (Ohlendorf *et al.*, 1988; Ohlendorf *et al.*, 1994).

Another aspect which McIntosh (1993) also addressed, again using partially purified extracts of *R. rhodochrous*, was whether catechol 1,2-dioxygenase activity obtained after induction on benzyl alcohol, is the same as that obtained after induction on styrene. The behaviour of each preparation during chromatography and the similarity of intradiol-cleavage activity towards

substituted catechols, suggests that they are the same (McIntosh, 1993). Since the catechol 1,2-dioxygenase purified (Methods 2.5) in this work was only ever obtained after induction on benzyl alcohol, a comparison of the purified enzymes could not be made and nor was it possible to establish whether the enzyme induced on benzyl alcohol would oxygenate 3-vinylcatechol, since this derivative of styrene (Figure 1.9) is not commercially available.

In order to measure the K_m and V values for the catechol- and chlorocatechol 1,2-dioxygenases from *Pseudomonas* B13, Dorn & Knackmuss (1978*b*) generated different concentrations of O_2 by bubbling defined O_2/N_2 mixtures through the assay prior to the addition of enzyme. It was found that the K_m values of chlorocatechol 1,2-dioxygenase for catechol at different O_2 concentrations were not significantly affected. It was concluded that O_2 did not affect the binding of organic substrates to chlorocatechol 1,2-dioxygenase and it was therefore assumed that the the K_m values for substituted catechols measured by using air-saturated solutions were the actual K_m values. On the other hand, the same authors found that only catechol 1,2-dioxygenase could approach maximum velocity in air-saturated solutions and then only when catechol or 4-methylcatechol were the substrates. For all other substrates tested, the maximum apparent velocity was scarcely half of V and with catechol as the organic substrate, the affinity of chlorocatechol 1,2-dioxygenase for O_2, was so low that half-maximum velocity was only just reached at 1×10^5 Pa of O_2. With this caveat in mind, the V' values of some catechol 1,2-dioxygenases, chlorocatechol 1,2-dioxygenases and protocatechuate 3,4-dioxygenases, normalised with respect to catechol or protocatechuate where appropriate, are compared in Table 4.4. The catechol 1,2-dioxygenases from *Acinetobacter calcoaceticus* and *Alcaligenes eutrophus* are much less active towards substituted catechols, with the latter enzyme less than 10% as active with 3-methylcatechol and showing scarcely any significant activity with 3-chlorocatechol. In contrast, the chlorocatechol 1,2-dioxygenases from *Pseudomonas putida* pAC27 and *A. eutrophus* for example, are distinguished by their ability to oxygenate substituted catechols, especially the 3- substituted ones such as 3-methylcatechol and 3-chlorocatechol. On the basis of relative activity, rhodococcal catechol 1,2-dioxygenase does not seem to discriminate between any of the substrates which were tested. To get a clearer impression of the substrate specificity of rhodococcal catechol 1,2-dioxygenase, the specificity constant (k_{cat}/K_m') (Fersht, 1985) for each substrate was calculated (Table 4.3). They demonstrate that if all the substrates were available to the enzyme simultaneously, catechol would be turned over faster than either 3- or 4-methylcatechol, since it has the highest specificity constant. If only the methyl derivatives were present, then neither would be turned over preferentially, since

Organism	*Rhodococcus rhodochrous*	*Pseudomonas putida* pAC27	*Acinetobacter calcoaceticus* BD413	*Pseudomonas putida* ATCC 23975 (formerly classified as *P. aeruginosa*)	*Alcaligenes eutrophus* JM134	*Alcaligenes eutrophus* JM134
Enzyme	CTD	CCD	CTD	PCD	CTD	CCD
Catechol	1.00	1.00	100	0.004	1.0	1.0
3-Methylcatechol	0.82	3.54	0.12	0.004	0.071	1.67
4-Methylcatechol	0.55	2.73	0 18	0.002	0.413	N.D.
Protocatechuic acid	0.0	0.0	0.0	1.0	N.D.	N.D.
3-Chlorocatechol	N.D.	1.31	N.D.	N.D.	0.002	1.24
4-Chlorocatechol	N.D.	0.80	N.D.	0.036	0.070	1.22

Table 4.4 Comparison of normalised V values for the oxygenation of substituted catechols

The V values are given relative to the V of catechol for the catechol 1,2-dioxygenase and chlorocatechol 1,2-dioxygenase and relative to the V of protocatechuate for protocatechuate 3,4-dioxygenase. All values are taken from Broderick & O'Halloran (1991) except those for the catechol 1,2-dioxygenase from *R. rhodochrous*, which are from this thesis and for the protocatechuate 3,4-dioxygenase from *P. putida*, which are taken from Nozaki (1979). Abreviations used are: CTD, catechol 1,2-dioxygenase; CCD, chlorocatechol 1,2-dioxygenase: PCD, protocatechuate 3,4-dioxygenase.

N.D., Not determined.

they both have approximately the same value. The k_{cat}/K_m' value has also been used as a measure of catalytic efficiency, since under some circumstances it can be reduced to an apparent second order rate constant for the reaction of free enzyme with substrate (Fersht, 1985). Given that the upper limit for a bimolecular reaction (10^8-10^9 M^{-1} s^{-1}) is diffusion controlled, the values of 10^6-10^7 M^{-1} s^{-1} obtained for rhodococcal catechol 1,2-dioxygenase are indicative of highly efficient catalysis.

Some of the non-haem iron intradiol-cleavage enzymes also cleave outside 3-substituted substrate hydroxyl groups under certain conditions (Fujiwara *et al.*, 1975). Rhodococcal catechol 1,2-dioxygenase was assayed for extradiol activity (Methods 2.4.1) with catechol, 3-methylcatechol and 4-methylcatechol, but no activity could be detected. Ratios of intradiol- to extradiol-cleavage for the catechol 1,2-dioxygenases from *P. arvilla* C-1 and *Pseudomonas* B13 with 3-methylcatechol and 3-methoxycatechol were approximately 17:1 and 5:1 (Fujiwara *et al.*, 1975; Dorn & Knackmuss, 1978*a*). Chlorocatechol 1,2-dioxygenases on the other hand, such as the one from *Brevibacterium fuscum* P-13, cleave all substrates between their hydroxyl groups. As the rate of intradiol-cleavage of 3-substituted catechols becomes slower relative to the reaction with catechol, so the proportion of extradiol activity increases. This point is illustrated with the catechol 1,2-dioxygenase from *Frateuria* ANA-18, which catalyses the intradiol-cleavage of 3-methylcatechol at a rate of less than 0.9% to that of catechol, with the ratio of intradiol to extradiol activity of 1.0:1.2 when 3-methylcatechol is the substrate (Aoki *et al.*, 1984*b*).

The tolerance towards substituted catechols such as 3-methylcatechol and the absence of any extradiol activity, have shown that rhodococcal catechol 1,2-dioxygenase is unlike other catechol 1,2-dioxygenases which have been characterised so far. At the same time, it is not sufficiently active with the substituted catechols to qualify as a chlorocatechol 1,2-dioxygenase, which are unequivocal in their preference towards substituted catechols. This combination of qualities is reflected in the subunit M_r which is heavier than the chlorocatechol 1,2-dioxygenases and lighter than the catechol 1,2-dioxygenases (Table 1.2). In this respect it is very similar to another Gram-positive enzyme from *Arthrobacter* mA3 (Eck & Belter, 1993) and with which it shares substantial sequence identity (Figure 4.7).

CHAPTER 5
Cloning, sequencing and expression of the gene encoding catechol 1,2-dioxygenase

5.1 Introduction

During the characterisation of catechol 1,2-dioxygenase (Chapter 4) from *Rhodococcus rhodochrous* NCIMB13259, *N*-terminal and peptide sequence data were obtained (Methods 2.8.3, 2.8.4 & 2.8.5) which showed substantial sequence identitity to the *Arthrobacter* mA3 catechol 1,2-dioxygenase (Eck & Belter, 1993) (Figure 4.7). Sequence alignments demonstrated that the rhodococcal *N*-terminal and peptide sequences were probably about 180 amino acids apart, representing approximately two-thirds of the *Arthrobacter* enzyme (Figure 4.7). It was hoped at the outset, therefore, that the amino acid sequence information would enable the design of oligonucleotide primers which could be used in polymerase chain reactions (PCR) to amplify an homologous probe from *R. rhodochrous* genomic DNA, thereby allowing the gene encoding catechol 1,2-dioxygenase (*cat*) to be cloned.

5.2 The design of three degenerate oligonucleotide primers

Two pools of degenerate oligonucleotide primers were designed using amino acid sequence data derived from the *N*-terminus of purified catechol 1,2-dioxygenase and a peptide purified from a *Staphylococcus aureus* V8 protease digest of catechol 1,2-dioxygenase (Sections 4.2.8 and 4.2.9). A third pool was designed on the basis of a strongly conserved hexapeptide which was discovered from multiple sequence alignments of protocatechuate 3,4-dioxygenase, catechol 1,2-dioxygenase and chlorocatechol 1,2-dioxygenase (residues 230-235, Figure 5.12). The amino acid sequences and the corresponding oligonucleotide primer pools are shown in Figure 5.1. Inosine was used in positions of 4-fold redundancy since it forms stable base pairs with all four bases (Martin *et al.*, 1985). One three-fold position was guessed, since a codon-usage table made from 40 *Rhodococcus* structural genes revealed a heavy codon bias towards cytosine at the third position (Table 5.1).

5.3 The optimisation of the polymerase chain reaction

In effect, two primer half-pairs (oligo1F-oligo1R and oligo1F-oligo2R) were made from the three pools of degenerate primers (Figure 5.2a). The products from each PCR were expected to be about 600 and 560 kb respectively. The two estimates were based on amino acid sequence alignment data shown in Figure 4.7 and the known difference in M_r of about 1 000 between the catechol 1,2-dioxygenases of *R. rhodochrous* (Section 4.2.6) and *Arthrobacter* mA3 (Eck & Belter, 1993).

During initial experiments to identify optimum PCR conditions for each primer half-pair, both the annealing temperature and the time of annealing were

(a) *N*-Terminal sequence

TTTENPTAHASG**NAATDKFK**SXR

*Eco*RI
Oligo1F 5'-GGGAATTCAACGCIGCIACIGACAAGTTCAA-3'
T T A T

(b) Conserved region (residues 230-235, Figure 5.12)

RPAH*H

*Bam*HI
Oligo1R 5'-GGGGATCCATGIAIATGIGCIGGIC-3'
G G

(c) V8/2 Peptide

ITTLK**PAPYKIPADG**

*Bam*HI
Oligo2R 5'-GGGGATCCGTCIGCIGGGATCTTGTA-3'
A A T A

Figure 5.1 The design of three degenerate oligonucleotide primers

Three pools of oligonucleotides, 1F, 1R and 2R, were designed from the highlighted parts of the amino acid sequences (a), (b) and (c) respectively. Oligonucleotides 1R and 2R were designed from the complementary strand and then reversed. In order to minimise the number of species in each oligonucleotide pool, positions of three- and four-fold redundancy were either represented by the neutral base inosine or were guessed using a codon-usage table (Table 5.1). Positions of two-fold degeneracy were represented by both bases. To facilitate cloning, each degenerate primer had a restriction site incorporated into the 5'-end. The melting temperature of each primer was estimated using the GeneJockey program (Taylor, 1991) and was approximately 58 °C in each case.

The V8/2 peptide was a product of a V8 protease digest (Methods 2.8.3) which had been purified by HPLC (Methods 2.8.4) (Figure 4.8). Both the *N*-terminal and the V8/2 peptide were sequenced according to Methods 2.8.5. The conserved region was taken from multiple sequence alignments (Figure 5.12).

X, Unassigned amino acid; *, Ile, Ala or Val.

Amino acid	Codon	Number	%	Amino acid	Codon	Number	%
Gly	GGG	228	16	Ile	ATA	36	5
	GGA	292	20		ATT	98	13
	GGT	320	22		ATC	593	82
	GGC	925	43				
				Thr	ACG	249	24
Glu	GAG	562	58		ACA	124	12
	GAA	414	42		ACT	88	8
					ACC	586	56
Asp	GAT	296	28				
	GAC	766	72	Trp	TGG	265	100
Val	GTG	427	32	END	TGA	39	72
	GTA	119	9		TAG	9	17
	GTT	147	11		TAA	6	11
	GTC	641	48				
				Cys	TGT	45	26
Ala	GCG	598	31		TGC	126	74
	GCA	329	17				
	GCT	189	10	Tyr	TAT	76	17
	GCC	792	42		TAC	363	83
				Leu	TTG	160	11
					TTA	20	1
Arg	AGG	58	5		CTG	558	38
	AGA	33	3		CTA	48	3
	CGG	271	25		CTT	103	7
	CGA	162	15		CTC	583	40
	CGT	166	15				
	CGC	404	37	Phe	TTT	64	12
					TTC	476	88
Ser	AGT	76	8				
	AGC	205	21	Gln	CAG	359	76
	TCG	293	30		CAA	113	24
	TCA	74	8				
	TCT	57	6	His	CAT	106	24
	TCC	259	27		CAC	340	76
Lys	AAG	311	79	Pro	CCG	405	44
	AAA	82	21		CCA	111	12
					CCT	83	9
Asn	AAT	82	20		CCC	318	35
	AAC	324	80				
Met	ATG	385	100				

Table 5.1 Codon-usage in the genus *Rhodococcus*

This codon-usage table was constructed from all of the *Rhodococcus* structural genes (a total of forty) contained within the Genbank and EMBL databases (1995) using the CodonFrequency program (Wisconsin Sequence Analysis Package, Genetics Computer Group, Madison, Wisconsin, USA). The overall G + C content was 65%.

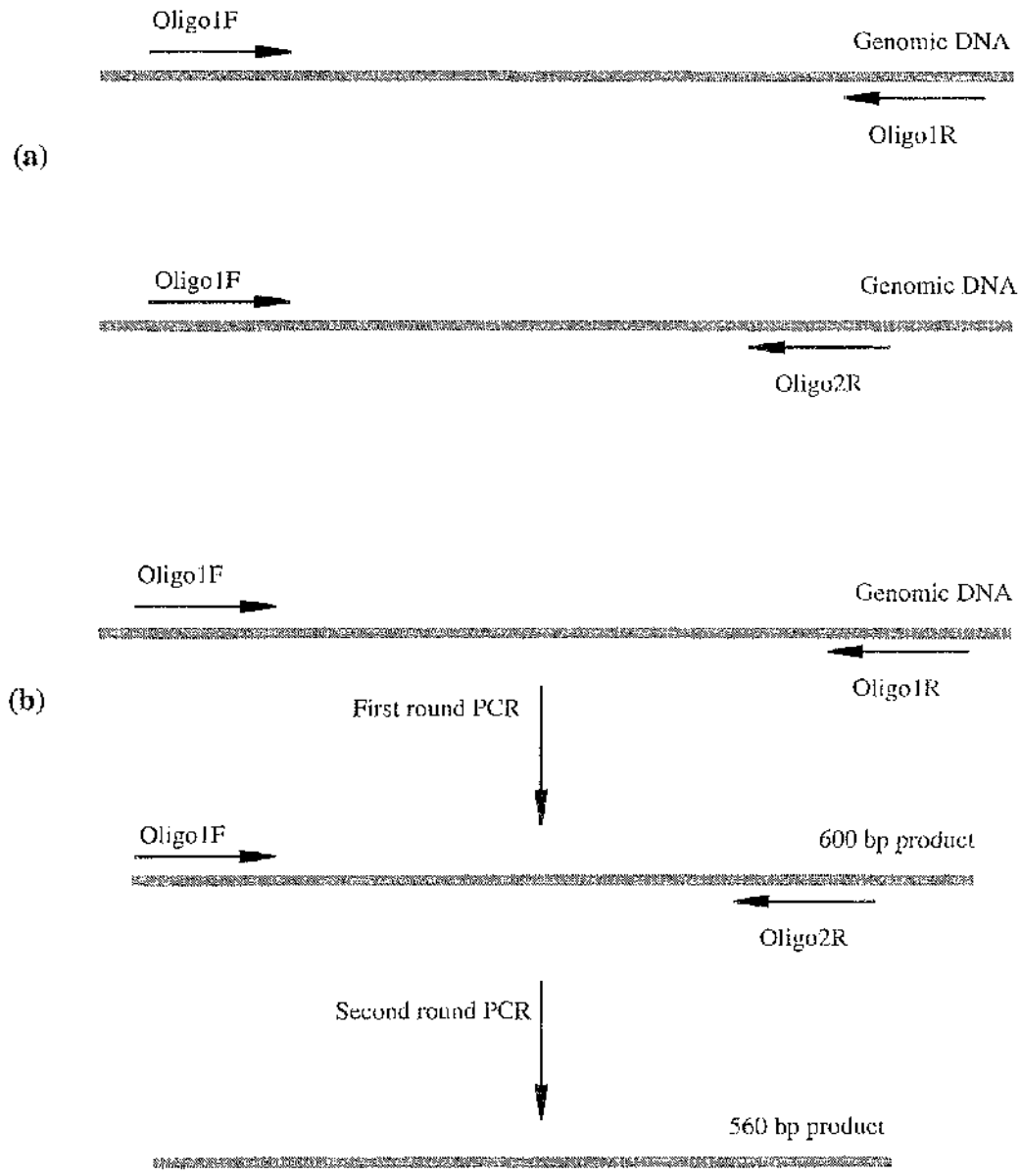

Figure 5.2 The use of three primers in two primer half-pair PCR reactions

(a) The three pools of degenerate primers (Figure 5.1) were used in two primer half-pair PCR reactions (Methods 2.9.10.1) consisting of an outer half-pair reaction (oligo1F-oligo1R) and an inner half-pair reaction (oligo1F-oligo2R). The products from the two reactions were expected to be 600 bp and 560 bp respectively.

(b) Hemi-nested PCR was carried out by linking both reactions. This was done by using 1% of the product from the outer reaction (first round PCR) to prime the inner reaction (second round PCR) (Methods 2.9.10.1).

varied, however, even when the annealing time was maintained at 40 °C for 3 min, well below the calculated annealing temperature of 53 °C for each primer (Methods 2.9.10), no products were visible upon analysis on agarose gels (Methods 2.9.2). It was not until the annealing time was extended to 4 min that a product of approximately 600 bp was obtained with the outer primer half-pair (1F-1R) (Lane 1, Figure 5.3), but still no products were obtained with the inner primer half-pair (1F-2R), even when the number of PCR cycles was increased to 35.

Another critical factor for successful PCR is the denaturing temperature. *Rhodococcus* DNA is extremely GC rich (G + C content 65%, see Table 5.1) so it was possible that the DNA was not completely denatured even though the temperature was maintained at 95 °C for 3 min prior to the addition of *Taq* DNA polymerase. Co-solvents such as glycerol and dimethylsulfoxide (DMSO) have been shown to greatly increase the amount of PCR amplification, presumably by increasing the solubility of the DNA template (Bej *et al.*, 1991) therefore the inner primer half-pair (1F-2R) PCR was done with 4% DMSO and resulted in some product formation (Lane 2, Figure 5.3).

The 600 and 560 bp products from the outer and inner half-pair PCR were purified by agarose gel electrophoresis (Methods 2.9.3) and cloned in *E. coli* DH5α (Methods 2.9.17). After restriction analysis upon agarose gels (Methods 2.9.2), transformants which contained the appropriate sized insert were purified using the Qiagen midi-prep (Methods 2.9.6) and sequenced manually (Methods 2.9.15.1), but none contained the requisite *Bam*HI-*Eco*RI insert.

5.4 A core region of the gene encoding catechol 1,2-dioxygenase was amplified using hemi-nested PCR

In order to increase specificity, the inner and outer half-pair PCR reactions were coupled by using a portion of the product from the outer half-pair reaction as a template for the inner reaction (Methods 2.9.10.1), in a hemi-nested PCR, whose overall strategy is illustrated in Figure 5.2b. The predicted 560 bp product of the hemi-nested PCR (Lane 4, Figure 5.3) was cloned in *E. coli* DH5α (Methods 2.9.17). Restriction analysis on agarose gels of 50 antibiotic resistant transformants (Methods 2.9.2) revealed eight recombinants, each containing a 560 bp insert. Three out of the four recombinant plasmids which were sequenced (Methods 2.9.15.1) contained a core region of the gene encoding catechol 1,2-dioxygenase. The identification was based upon known amino acid sequences and sequence similarities with other catechol 1,2-dioxygenases (Figure 4.7). Approximately two-thirds of the *Bam*HI-*Eco*RI insert from one of the recombinant plasmids, designated pP29, was sequenced (Methods 2.9.15.1) and

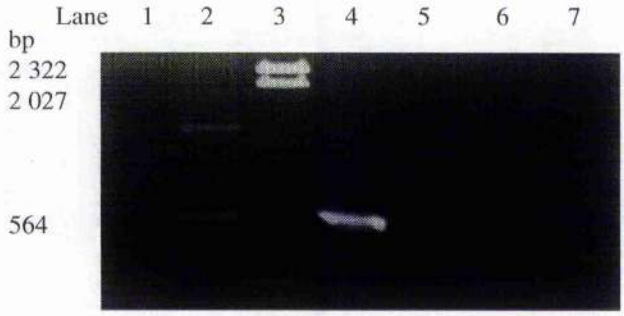

Figure 5.3 Agarose gel electrophoresis of PCR products

All PCR reactions were carried out as described in Methods 2.9.10.1 using an annealing temperature of 50 °C and were then analysed (20 µl per reaction) on a 1% agarose gel (Methods 2.9.2).

Lanes 1-4 were as follows: 1, outer half-pair product; 2, inner half-pair product; 3, three fragments from a *Hind* III restriction digest (Methods 2.9.11); 4, hemi-nested product; lanes 5-7 were single primer controls using oligos 1F, 1R and 2R respectively.

found to have a G + C content of about 65%. The derived amino acid sequence was very similar to the catechol 1,2-dioxygenase from *Arthrobacter* mA3 (Eck & Belter, 1993), in keeping with the amino acid sequence already obtained (Figure 4.7).

5.5 The use of the *Bam*HI-*Eco*RI insert as an homologous probe

The recombinant plasmid, pP29, was used as an homologous probe in Southern hybridisation experiments with *R. rhodochrous* DNA (Methods 2.9.13.1). The probe hybridised strongly to several fragments in each restriction digest, except for the *Sma*I digest where it hybridised to only a single fragment of approximately 3.5 kb (Lane 5, Figure 5.4). Since *Sma*I produces blunt ends which are difficult to ligate, *R. rhodochrous* genomic DNA was digested with the isoschizomer *Xma*I (Methods 2.9.11), to give fragments with suitable 5'-overhangs. The requisite 3.5 kb band was purified by agarose gel electrophoresis (Methods 2.9.3) and then cloned in *E. coli* XL2-Blue ultra-competent cells (Methods 2.9.17). Three colonies hybridised (Methods 2.9.13.2) strongly to the labelled probe (Figure 5.5). Restriction analysis of all three clones (Methods 2.9.2), showed that each contained the same size insert of approximately 3.5 kb. Automated sequencing reactions (Methods 2.9.15.2) using primers (2545 and 2708, Figure 5.6) designed from the *Bam*HI-*Eco*RI insert were carried out on one of the clones (pPDS2) and confirmed that the *Xma*I insert contained the gene encoding catechol 1,2-dioxygenase. At the same time, in-frame initiation and termination codons were discovered, suggesting that the entire gene had been cloned.

5.6 Sequencing the recombinant plasmid pPDS2

The partial nucleotide sequence of the *Xma*I fragment was obtained using the strategy shown in Figure 5.6. The sequence (Figure 5.7) contains within it an open reading frame which encodes a catechol 1,2-dioxygenase with a M_r value of 31 599, in agreement with the experimentally derived (Methods 2.8.6) value of 31 598 for the enzyme purified from *R. rhodochrous* (Figure 4.6). The derived amino acid sequence is also in agreement with the chemically determined (Methods 2.8.3, 2.8.4 and 2.8.5) amino acid sequence (Figure 4.7).

The codon-usage in *cat* is more strongly biased towards G and C in the third position and for Asp, Lys, Asn, Ile, Cys, Tyr, Phe, Gln and His there is a 100% major codon assignment (Table 5.2). Partial open reading frames were found both upstream and downstream of *cat*, but neither encoded regions of non-random codon-usage. Other potential coding regions were found in the remaining

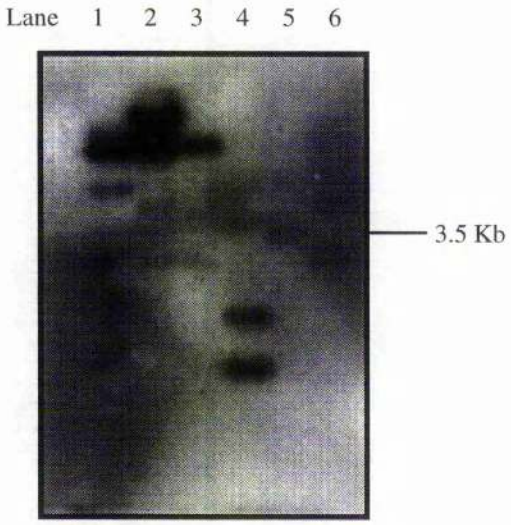

Figure 5.4 Southern hybridisation of *Rhodococcus rhodochrous* DNA with the *Eco*RI-*Bam*HI insert

A Southern blot of digested *Rhodococcus rhodochrous* DNA was hybridised to 10 ng of denatured, radiolabelled *Eco*RI-*Bam*HI insert for 8 h at 68 °C in a total volume of 10 ml (Methods 2.9.13.1 & 2.9.14). Filters were then autoradiographed at −70 °C for 24 h (Methods 2.9.16). Lanes 1- 6 were as follows: 1, *Bam*HI; 2, *Xho*I; 3, *Eco*RI; 4, *Pst*I; 5, *Sma*I; 6, *Sac*I.

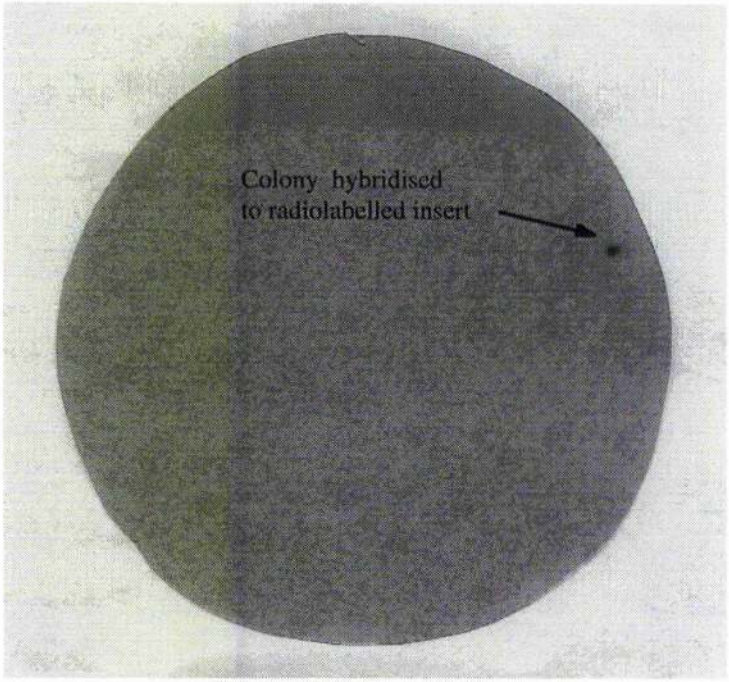

Figure 5.5 Colony hybridisation with the radiolabelled *Eco*RI-*Bam*HI insert

Hybridisation was carried out at 68 °C for 6 h using 20 ng of denatured radiolabelled insert in a total volume of 20 ml. Filters were then autoradiographed for 24 h at −70 °C (Methods 2.9.16). Three colonies hybridised to three separate filters but only one of the filters is shown.

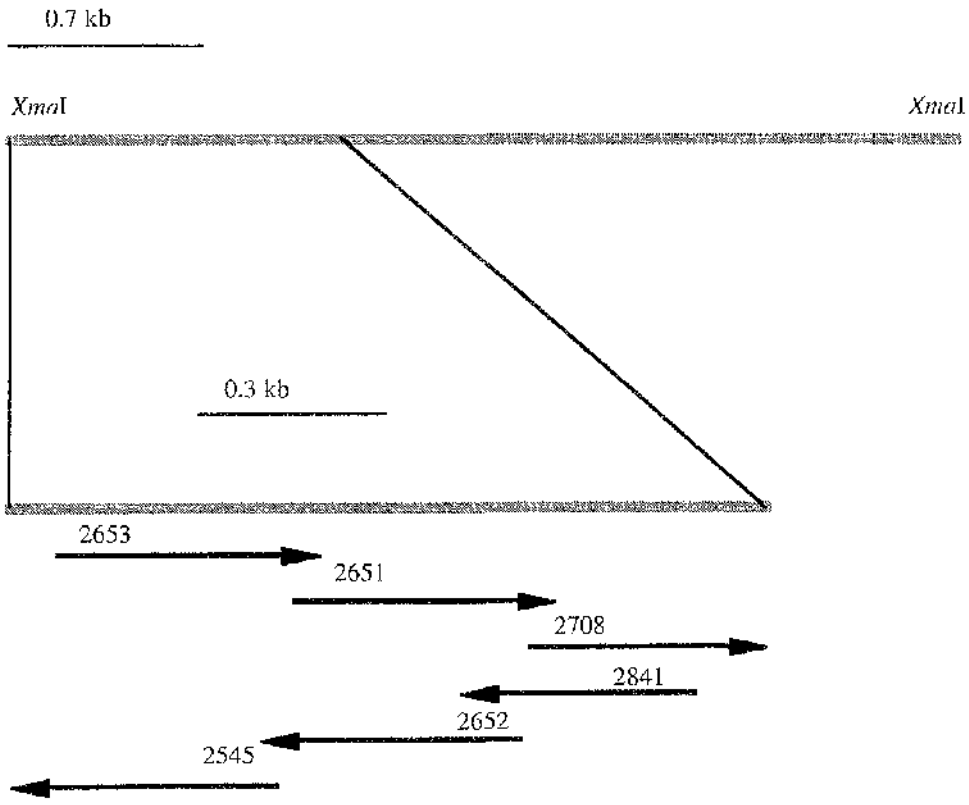

Figure 5.6 DNA sequencing strategy

The partial nucleotide sequence of the *Xma*I insert from the recombinant plasmid pPDS2 was obtained by automated sequencing (Methods 2.9.15.2). Each arrow represents an overlapping sequence determined with internal oligonucleotides (Methods 2.9.9). The sequence of each primer is shown, together with its position relative to the 5' *Xma*I restriction site.

Oligonucleotide primer	pPDS2 sequence	Position on *Xma*I fragment
2653	GATCACCACCCACAGCT	107-123
2651	ACGGCGAGTGGCCGCTGT	375-392
2708	GCTACTACTCGCAGTTCGCC	645-664
2652	GGGGCGAACTGCGAGTAG	666-649
2545	ACGTCCTTGTAGATGGCA	282-265
2841	CGCGACGACGAATGCCGGTGC	1089-1069

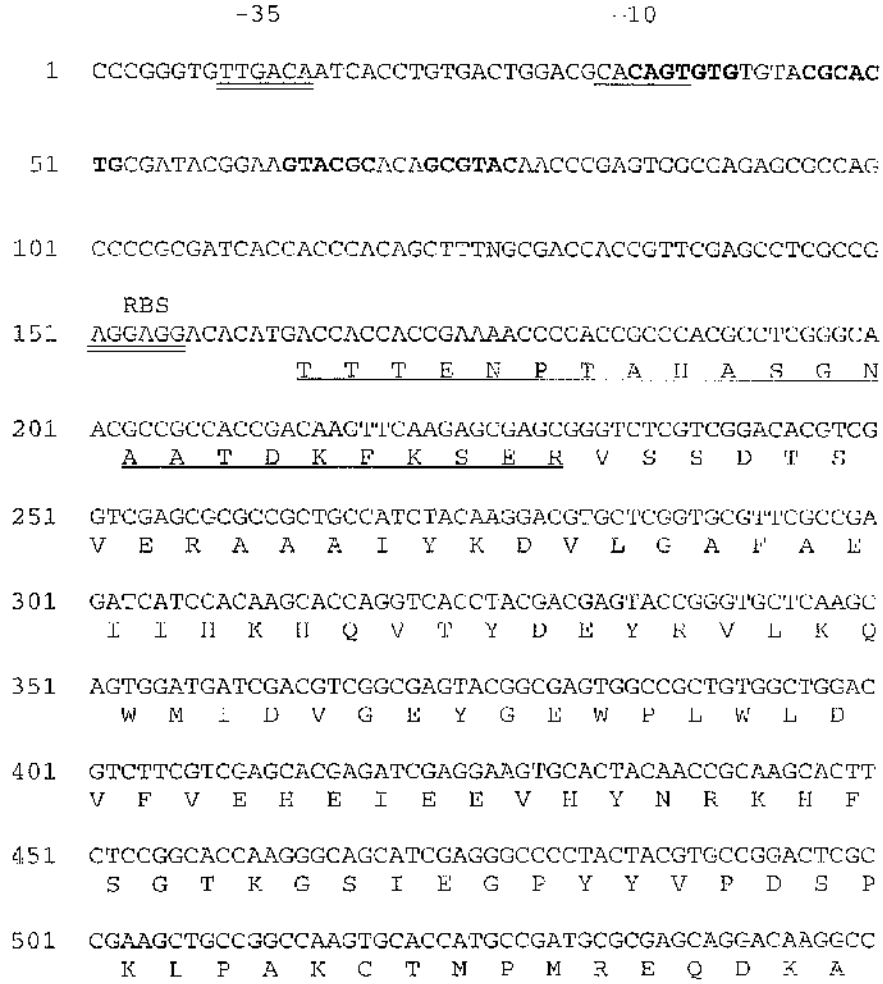

Figure 5.7 Partial nucleotide sequence of the *Xma*I insert of the recombinant plasmid pPDS2 and the deduced amino acid sequence of the gene encoding catechol 1,2-dioxygenase

DNA sequencing of the *Xma*I insert (Methods 2.9.15.2) from the recombinant plasmid was done according to the strategy shown in Figure 5.6. Regions of amino acid sequence confirmed from protein sequencing experiments (Sections 4.2.8 & 4.2.9) have been underlined. Putative −35 and −10 promoter sequences, together with a ribosome binding site have been double underlined. Two inverted repeat structures are in bold. Nucleotides are numbered from the first base in the *Xma*I site.

Abreviations used: RBS, ribosome binding site; *, stop codon.

```
551  GCCGAGGCGCTGATCTTCAAGGGCCAGGTCACCGACCTCGAGGGCAAGGG
       A  E  A  L  I  F  K  G  Q  V  T  D  L  E  G  K  G

601  CCTGCAGGGCGCCACCGTCNANCTGTGGCACGCCGACAAGGAGGGCTACT
       L  Q  G  A  T  V  X  L  W  H  A  D  K  E  G  Y  Y

651  ACTCGCAGTTCGCCCCCGGCATCCCCGAGTGGAACCTGCGCGGCACCATC
       S  Q  F  A  P  G  I  P  E  W  N  L  R  G  T  I

701  GAGTGCGACGACGAGGGCAACTTCGAGATCACCACGCTCAAGCCCGCGCC
       E  C  D  D  E  G  N  F  E  I  T  T  L  K  P  A  P

751  CTACAAGATCCCGGCCGACGGCCCCGACCGGCTGGTTCATCGACTCCTACG
       Y  K  I  P  A  D  G  P  T  G  W  F  I  D  S  Y  G

801  GCGGCCACCCGTGGCGTCCCGCCCACCTGCACCTCATGGTGAAGGCTCCG
       G  H  P  W  R  P  A  H  L  H  L  M  V  K  A  P

851  GGCAAGCGCGCCATCACCACCCAGCTCTACTTCCAGGGCGGCGAGTGGGT
       G  K  R  A  I  T  T  Q  L  Y  F  Q  G  G  E  W  V

901  CGAGGACGACGTCGCGACCGCGACCAAGCCCGAGCTGATCCTGGACCCGC
       E  D  D  V  A  T  A  T  K  P  E  L  I  L  D  P  Q

951  AGCCGAACGCCGACGGCGTCGACGAGGTGACCTACAACTTCCCGCTCGAC
       P  N  A  D  G  V  D  E  V  T  Y  N  F  A  L  D

1001 CCCGAGGCCTGAGCACTCCTGTTGAACCGGGTCCACGAGTGGTGACAAAT
     P  E  A  *

1051 GCCGCTTGTCCGCCAAGGGCACCGGCATTCGTCGTCGCGGAAGCCCGGCT

1101 GTTGCCCCGCATCACCCGCCCGAAATCCGATTTGTTGAACTTGTTCGATG

1151 TCCCGACCCCCGACCCTGAAATTCGTTCTCGATCCGATACGACGAATCGT

1201 CGAACTTTCCGCTCCTTGCGTTTCCCCAC
```

Amino acid	Codon	Number	%	Amino acid	Codon	Number	%
Gly	GGG	0	0	Ile	ATA	0	0
	GGA	0	0		ATT	0	0
	GGT	1	4		ATC	14	100
	GGC	22	96				
				Thr	ACG	2	10
Glu	GAG	23	92		ACA	0	0
	GAA	2	8		ACT	0	0
					ACC	18	90
Asp	GAT	0	0				
	GAC	20	100	Trp	TGG	8	100
Val	GTG	6	35	END	TGA	1	100
	GTA	0	0		TAG	0	0
	GTT	0	0		TAA	0	0
	GTC	11	65				
				Cys	TGT	0	0
Ala	GCG	6	23		TGC	2	100
	GCA	0	0				
	GCT	2	8	Tyr	TAT	0	0
	GCC	18	69		TAC	13	100
				Leu	TTG	0	0
					TTA	0	0
Arg	AGG	0	0		CTG	10	59
	AGA	0	0		CTA	0	0
	CGG	2	25		CTT	0	0
	CGA	0	0		CTC	7	41
	CGT	1	12				
	CGC	5	62	Phe	TTT	0	0
					TTC	10	100
Ser	AGT	0	0				
	AGC	2	20	Gln	CAG	9	100
	TCG	6	60		CAA	0	0
	TCA	0	0				
	TCT	0	0	His	CAT	0	0
	TCC	2	20		CAC	10	100
Lys	AAG	18	100	Pro	CCG	11	55
	AAA	0	0		CCA	0	0
					CCT	0	0
Asn	AAT	0	0		CCC	9	45
	AAC	7	100				
Met	ATG	4	100				

Table 5.2 Codon-usage in the gene encoding catechol 1,2-dioxygenase from *Rhodococcus rhodochrous* NCIMB 13259

Codon-usage in the gene encoding catechol 1,2-dioxygenase (Figure 5.7) was calculated using the CodonFrequency program (Wisconsin Sequence Analysis Package, Genetics Computer Group, Madison, Wisconsin, USA).

five reading frames, but all but one of them were less than 100 bp and none of them displayed a non-random codon-usage.

A putative ribosome binding site (AGGAGG) was found 4 bp upstream from the initiation codon (ATG) and had a calculated binding energy (ΔG) of 16.6 Kcal mol^{-1} (Strohl, 1992). In addition, putative −35 and −10 promotor sequences (TTGACA and CACAGT respectively) and two inverted repeat structures were discovered upstream of the coding region (Figures 5.7 & 5.8).

5.7 Over-expression of *Rhodococcus rhodochrous* catechol 1,2-dioxygenase

Expression analysis of the plasmid pPDS2 suggested that catechol 1,2-dioxygenase was being expressed from its own promoter in XL2-Blue cells (Lane 5, Figure 5.9) with a level of expression comparable to that obtained in *R. rhodochrous* (Lane 2, Figure 3.3). In order to increase expression, *cat* was cloned downstream of a strong promoter and ribosome binding site. The gene encoding catechol 1,2-dioxygenase was amplified with *Vent* DNA polymerase (Methods 2.9.10.2) from *R. rhodochrous* using a PCR technique (Figure 5.10). The amplicon was then cloned into the plasmid pTB361 which contains the T7 promoter (Table 2.1, Methods 2.9.17) producing the recombinant plasmid, pPDSX4CTD, which was sequenced (Methods 2.9.15.2) and found to be identical to the genomic sequence (Figure 5.7).

The expression analysis of pPDSX4CTD revealed that catechol 1,2-dioxygenase was expressed maximally after 4 h induction with isopropylthiogalactoside (IPTG) (Lane 8, Figure 5.11), to approximately 40% of the cellular protein as estimated by densitometry (Methods 2.7.5). The recombinant catechol 1,2-dioxygenase was purified twice (Methods 2.5) and on both occasions it behaved in the same way as the wild-type catechol 1,2-dioxygenase did during ion-exchange and hydrophobic interaction chromatography (Figures 3.1 & 3.2). Both preparations had an average specific activity [units (mg protein)$^{-1}$] of 11.5 (11, 12) (Methods 2.3.2 & 2.4.1) which is the same as the average specific activity of the wild-type protein (Table 3.2). Atomic absorption analysis (Methods 2.8.8.2) on one of the preparations showed it to have 0.5 atoms of Fe per subunit, which is slightly less iron than the wild-type protein contains (Table 4.1). *N*-Terminal amino acid analysis was also carried out on one of the preparations (Methods 2.8.5) and revealed that some of the polypeptide chains [(probably less than 25%), personal communication, J. N. Keen] contain a methionine residue which contrasts with the wild-type catechol 1,2-dioxygenase polypeptide, which does not have the *N*-terminal methionine residue (Figure 4.7).

	−10	−35
Escherichia coli	TAtAaT	TTGACa
Rhodococcus rhodochrous	CACAGT	TTGACA

Figure 5.8 Putative −10 and −35 promoter elements in *Rhodococcus rhodochrous*

Putative −10 and −35 promoter elements upstream of the gene encoding catechol 1,2-dioxygenase in *R. rhodochrous*, compared with consensus sequences from *E. coli* (Harley & Reynolds, 1987). Upper and lower cases refer to strongly and weakly conserved bases respectively.

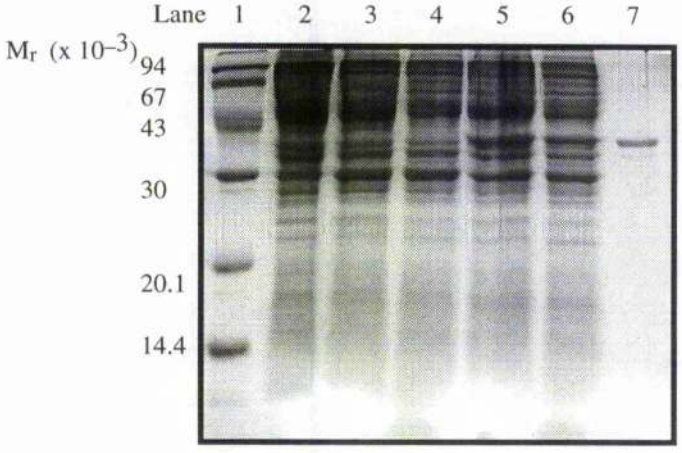

Figure 5.9 Expression of catechol 1,2-dioxygenase in *Escherichia coli* XL2-Blue pPDS2 and analysis with SDS-PAGE

E. coli was grown as described in Methods 2.2.3.1. Electrophoresis was carried out at 200 V using a 12% acrylamide separating gel and 5% stacking gel (Methods 2.7.1). Lanes 1-7 are as follows: 1, M_r markers; 2, *E. coli* XL2-Blue; 3, *E. coli* XL2-Blue SK II; 4, *E. coli* XL2-Blue SK II after 4 h induction on IPTG; 5, *E. coli* XL2-Blue pPDS2; 6, *E. coli* XL2-Blue pPDS2 aftr 4 h induction with IPTG; 7, Catechol 1,2-dioxygenase purified from *Rhodococcus rhodochrous* (Methods 2.5).

*Nde*I
CTDE1 5'-GGAATTCCATATGACCACCACCGAAAACCCCAC-3'

*Bgl*II
CTDE2 5'-GAAGATCTCAGGCCTCGGGGTCGAGCGCGAA-3'

Figure 5.10 The amplification of the gene encoding catechol 1,2-dioxygenase using PCR

The catechol 1,2-dioxygenase gene from *Rhodococcus rhodochrous* was amplified using PCR (Methods 2.9.10.2). The oligonucleotide CTDE1 had an *Nde*I restriction site to facilitate cloning and this was followed by the 5'-end of the gene. Similarly, the oligonucleotide CTDE2 had a *Bgl*II restriction site incorporated and this was followed by the 3'-end of the gene. The oligonucleotides were made according to Methods 2.9.9.

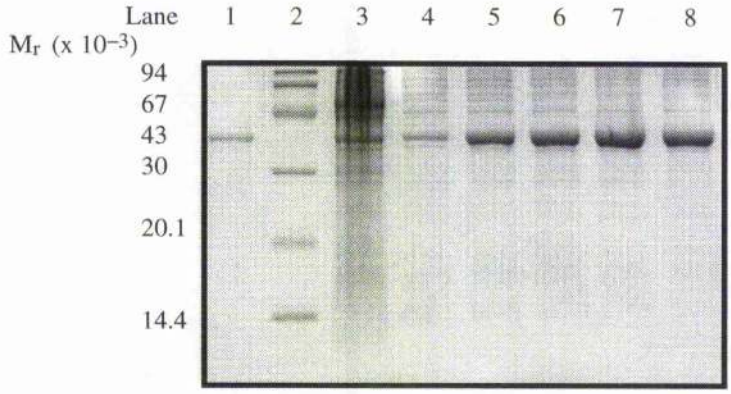

Figure 5.11 Over-expression of the gene encoding catechol 1,2-dioxygenase in *Escherichia coli* JM109(DE3) pPDSX4CTD

E. coli was grown as described in Methods 2.2.3.1. Electrophoresis was carried out at 200 V using a 12% acrylamide separating gel and 5% stacking gel (Methods 2.7.1). Lanes 1-8 are as follows: 1, Catechol 1,2-dioxygenase purified from *Rhodococcus rhodochrous* (Methods 2.5); 2, M_r markers; 3, *E. coli* JM109 (DE3); 4, *E. coli* JM109 (DE3) pTB361; 5-8, *E. coli* JM109 (DE3) pPDSX4CTD after 0, 1, 2, 4 h induction with IPTG. The specific activity of catechol 1,2-dioxygenase (Methods 2.3.2 & 2.4.1) reached a maximum of 4 units mg (protein)$^{-1}$ after 4 h induction with IPTG.

5.8 Discussion

5.8.1 DNA sequence

The expression of catechol 1,2-dioxygenase in *E. coli* pPDS2 in the absence of IPTG suggested that the gene encoding this enzyme was being expressed from its own promoter (Figure 5.9). Within the upstream region, more than 100 nt from the ATG coding initiation start site, were putative −10 and −35 promotor sequences, 18 nt apart (Figure 5.7). Both bear a strong resemblance to the consensus sequences for those *E. coli* promoters (Figure 5.8) recognised by RNA polymerases containing σ^{70}-like subunits (Harley & Reynolds, 1987). The DNA sequence upstream from the ATG start codon had a G+C content of 61%, which is slightly lower than the figure of 67% obtained for the coding region. In contrast, the average G+C content of *E. coli* promoters is 43% (Bibb *et al.*, 1985). *Streptomyces*, like *Rhodococcus,* are *Actinomycetes* (Embley & Stackebrandt, 1994) and their genes also have a higher G+C content than those isolated from *E. coli*. Strohl (1992) compiled a list of 139 streptomycete start sites and found that only 29 appeared to belong to the σ^{70}-like consensus group. This sub-section contained the only promoters which were active in *E. coli* and they also, not suprisingly, had the lowest G+C composition (~57%). Although there are *E. coli* promoter sequences that are recognised by specific σ-factors for special purposes, for example in the heat shock response and in the metabolism of nitrogen, there are comparatively few promoters that vary from the σ^{70} consensus sequences. It has been proposed that there are only two basic types of promoter in *E. coli*, those recognised by σ^{70} and which may be modulated by positive and negative regulators and those promoters recognised by σ^{54}, which are regulated only by activation. The first type of promoter may contain proximal regulatory elements which overlap the −35 and −10 promoter sequences, bringing any regulator into direct contact with the RNA polymerase. With the second type of promoter, however, the location of activator binding is less critical and may in some cases be remote from the binding of the RNA polymerase (Collado-Vides *et al.*, 1991). Catechol 1,2-dioxygenase is expressed constitutively in *E. coli* and is induced in *R. rhodochrous,* which suggests that its expression is under negative control. In contrast, the genes encoding most bacterial catechol- and chlorocatechol 1,2-dioxygenases appear to be positively regulated by the LysR family of DNA binding proteins (Frantz & Chakrabarty, 1987; Kasak *et al.*, 1993; Matrubutham & Harker, 1994) which recognise $TN_{11}A$ and $GN_{11}A$ motifs within inverted repeat structures straddling promoter regions (Goethals *et al.*, 1992). Both motifs, together with their inverted repeats, have also been found upstream of the rhodococcal gene, the first one ($GN_{11}A$) starting within the putative −10 region. This is a complete contrast to positively regulated genes, where the DNA binding

motifs normally overlap, to some extent, the −35 consensus site. The relative spatial difference between positive and negative regulatory sites within *E. coli* σ^{70} promoters was noted by Collado-Vides *et al.* (1991), who found that about 60% of the activator sites which were studied touched the −40 position, whereas none touched the −10 position. In contrast, 49% of repressible sites have operators overlapping the +1 site. The reason why the activation sites tend to be fixed around the −35 region is thought to stem from promoter recognition. Studies involving positively controlled regulons have shown that the −35 regions are often far from consensus, indicating weak promoter-binding activity (Collado-Vides *et al.*, 1991) suggesting that the role of the activator in these cases may be to replace the −35 element and provide a substitute contact point for RNA polymerase. There may however be nuances in the binding interactions across the LysR family of regulated promoters, with effete −35 sequences representing one extreme. For example, site-directed mutagenesis experiments within the positively regulated *cat*BC −35 promoter consensus had no effect on CatR binding but almost eliminated *in vivo* expression from the *cat*BC promoter regardless of whether the inducer was present or not (Parsek *et al.*, 1994), showing that the interaction of the RNA polymerase with the −35 region is an important event in the transcriptional activation of *cat*BC. Of course, the role of the inverted repeat structures within the putative promoter region of the rhodococcal gene is not known; however, the presence of the $GN_{11}A$ motif raises the intriguing possibility of whether, upstream of the rhodococcal *cat* gene, there is a divergently transcribed regulator protein, since in the *cat*BC operon this motif lies within a repressor binding site which mediates the autoregulation of CatR (Parsek *et al.*, 1994).

Rhodococcus catechol 1,2-dioxygenase was expressed in both *E. coli* XL2-Blue and *E. coli* JM109 (DE3) during this project (Figure 5.9 and Figure 5.11 respectively). In *E. coli* XL2-Blue, the level of expression was comparable to the wild-type expression in *R. rhodochrous*, whereas in *E. coli* JM109 (DE3), under control of the T7 promoter, catechol 1,2-dioxygenase amounted to approximately 40% of the soluble proteins, about twenty times more than the wild-type expression. Experiences with the expression of pseudomonad genes in *E. coli* strains have been mixed, with some workers reporting problems. For example, Frazee *et al.* (1993) attempted to express the *Pseudomonas putida* protocatechuate 3,4-dioxygenase in *E. coli* using in turn, the *lac*, *tac* and T7 expression systems. Expression was achieved, albeit at a very low level compared with the wild-type expression in *P. putida*. The failure was attributed to the comparatively high G+C content of *Pseudomonas* genes with respect to those in

E. coli and/or to a divergent pattern of codon-usage, imposing a block at the level of translation, although high levels of expression have been achieved with genes which have very different patterns of codon-usage to that of an optimal *E. coli* gene (Robinson *et al.*, 1984; Holm, 1986). Recent evidence shows that rare codons early in the first few codons of an open reading frame can affect both the quantity and quality of the synthesised protein. In particular, the clustering of such codons as CUA, AUA, CGA or CCC can be important in regulating translation efficiency (Kane, 1995). In the *cat* gene, by contrast, there is a strong bias toward a subset of major codons (Table 5.2). Although in unicellular organisms the codon-usage for most genes seems to reflect the average composition of the genome as a whole, there is a subset of genes, for example those encoding ribosomal proteins and elongation factors, for which the codon-usage is strongly biased toward a group of major codons (Andersson & Kurland, 1990). Major codon bias has been correlated with the expression level of the protein molecules and it has been observed that the higher the protein production level, the higher the tendency to use only a subset of codons in the gene (Andersson & Kurland, 1990).

The putative Shine-Dalgarno sequence (AGGAGG) in the rhodococcal gene (Figure 5.7) was identical to the proposed consensus sequence for *Streptomyces* genes (a/gGGAGG) (Strohl, 1992). McLaughlin *et al.* (1981) proposed that that the Shine-Dalgarno sequences of Gram-positive bacterial mRNAs are typically able to form strong complexes with the 16S mRNA, whereas *E coli* sequences are more variable. The data accumulated from the apparent Shine-Dalgarno sequences of 44 *Streptomyces* genes, however, suggest that the strength of *Streptomyces* Shine-Dalgarno sequences vary as much as those from *E. coli* (Gold *et al.*, 1981) with both types of bacteria tolerating relatively poor Shine-Dalgarno sequences (Strohl, 1992).

5.8.2 The derived amino acid sequence

The catechol 1,2-dioxygenase from *R. rhodochrous* was compared with nine other non-haem iron intradiol-cleavage enzymes (Figure 5.12). The protocatechuate 3,4-dioxygenases were represented only by the β-subunits, since the α-subunits from this enzyme contain only vestigial iron binding sites. The *Rhodococcus* sequence was found to have 26-28% identity to protocatechuate 3,4-dioxygenases, 29-30% identity to the catechol 1,2-dioxygenase from *Acinetobacter calcoaceticus* and *Pseudomonas putida* EST1001 respectively, and 30-35% identity to the chlorocatechol 1,2-dioxygenases. The closest alignment was with the catechol 1,2-dioxygenase from *Arthrobacter* mA3 (Eck & Belter, 1993) which was 56% identical (71% similarity) (Table 5.3). The high degree of

```
                1                                                      50
PCXB_PSEPU      .......... .......... .......... .......... ..PAQDNSRF
PCXB_ACICA      .......... .......... .......... .......... .MSQIIWGAY
PCXB_BURCE      .......... .......... .......... .......... ...MDSPTIL
CATA_ACICA      .....MEVKI FNTQDVQDFL RVASGLEQEG GNPRVKQIIH RVLSDLYKAI
PHEB_PSESP      .....MTVKI YDTPEVQDFL KIVAGLDQEG GNDRGKQIIH RILSDLYRTI
CLCA_PSEPU      .......... .......... .......... MDKRVAEVAG AIVEAVRKIL
TCBC_PSESP      .......... .......... .......... MNERVKQVAS ALVDAIQKTL
TFDC_ALCEU      .......... .......... .......... MNKRVKDVVD AIVAAVQRVL
CATA            MSTETEATAA ASGAGATARF RETKHVAAGT SKERVSALAG RVIKAINDTV
CAT             .TTTENPTAH ASGNAATDKF KSER.VSSDT SVERAAAIYK DVLGAFAEII
Consensus       ---------- ---------- ---------- ---------- ----------

                51                                                     100
PCXB_PSEPU      VIRDRNWHPK ALTPDYKTSI ARSPRQALVS IPQSISETTG PNFSHLGF..
PCXB_ACICA      AQSNTSDEPP AYRPGYKTSV LRSPKNRLIS IAETLSEVTA PHFSADKF..
PCXB_BURCE      TPRDWPSHPA YVHPDYRSSV KRGPTRPMIP LKERLRDQYA PVYGAEDL..
CATA_ACICA      EDLNITSDEY WAGVAYLNQL GANQEAGLLS PGLGFDHYLD MRMDAEDAAL
PHEB_PSESP      DDFDITAEQY WSAVSLLNAL GQASQFGLLS PGLGFDHYMD MRMDAADAEA
CLCA_PSEPU      LDKRVTEAEY RAGVDYLTEV AQTRETALL. ....LDVFLN STIFEGKAQR
TCBC_PSESP      TEQRVTEEEW RAGVGYMMKL AEAKEVAVL. ....LDAFFN HTIVDLKAQA
TFDC_ALCEU      DQKEVTEAEY RTAVHYLMQV AEQRETALL. ....CDVFFN STVAATKARI
CATA            LEEKVTYDEY NALKAWLISV GETGEWPLF. ....LDVWVE HSVEEVANEN
CAT             HKHQVTYDEY RVLKQWMIDV GEYGEWPLW. ....LDVFVE HEIEEVHYNR
Consensus       ---------- ----- ---- ---------- ---------- ----------
```

Figure 5.12 Sequence alignment of the catechol 1,2-dioxygenase from *Rhodococcus rhodochrous* with other non-haem iron intradiol-cleavage enzymes

The sequence alignment of the rhodococcal catechol 1,2-dioxygenase with other non-haem iron intradiol-cleavage enzymes, taken from the Swissprot and PIR databases, was performed with the PILEUP program (Wisconsin Sequence Analysis Package, Genetics Computer Group, Madison, Wisconsin, USA). A consensus sequence consisting of conserved residues only, is shown underneath each alignment, with the residues which ligate iron underlined. A conserved region (residues 230-235) is overlined.

Sequence	Organism	Enzyme	Reference
PCXB_PSEPU	*Pseudomonas putida* ATCC23975	PCD	Frazee *et al.* (1993)
PCXB_ACICA	*Acinetobacter calcoaceticus* BD413	PCD	Hartnett *et al.* (1990)
PCXB_BURCE	*Burkholderia cepacia*	PCD	Zylstra *et al.* (1989)
CATA_ACICA	*Acinetobacter calcoaceticus* BD413	CTD	Neidle *et al.* (1988)
PHEB_PSESP	*Pseudomonas putida* EST 1001	CTD	Kivisaar *et al.* (1991)
CLCA_PSEPU	*Pseudomonas putida* pAC27	CCD	Frantz & Chakrabarty (1987)
TCBC_PSESP	*Pseudomonas* p51	CCD	van der Meer *et al.* (1991*a*)
TFDC_ALCEU	*Alcaligenes eutrophus* JMP134	CCD	Ghosal & You (1988)
CATA	*Arthrobacter* mA3	CTD	Eck & Beller (1993)
CAT	*Rhodococcus rhodochrous*	CTD	This work

The abreviations are as follows: PCD, protocatechuate 3,4-dioxygenase (β-subunit); CTD, catechol 1,2-dioxygenase; CCD, chlorocatechol 1,2-dioxygenase.

```
                101                                                 150
PCXB_PSEPU      .......... ........GA HDHDLLLNFN NGGLPIGERT IVAGRVVDQY
PCXB_ACICA      .......... ........GP KDNDLILNYA KDGLPIGERV IVHGYVRDQF
PCXB_BURCE      .......... ........GP LDHDLTKNAV KNGEPLGERI VVTGRVLDEG
CATA_ACICA      GIENATPRTI EGPLYVAGAP ESVGYARMDD .GSDPNGHTL ILHGTIFDAD
PHEB_PSESP      KRTGGTPRTI EGPLYVAGAP EAEGFARMDD .DPDTDGETM WLHGQVRDTA
CLCA_PSEPU      SR..TSAPAI QGPYFLEGAP VVEGV..LKT YDTDDH.KPL IIRGTVRSDT
TCBC_PSESP      TR..GSRPAM QGPYFLEGAP VVAGA..LKT YEDDSH.HPL VIRGAVRTDD
TFDC_ALCEU      SE..GSTPAI EGPYYRDDAP LVDDR..LKT YDTDDH.KPL LIQGTVKAVD
CATA            RH..GSKGTI EGPYYIPNAP TQNTPATLPM RD.DEPGTPL LFQGQVRNLA
CAT             KHFSGTKGSI EGPYYVPDSP KLPAKCTMPM REQDKAAEAL IFKGQVTDLE
Consensus       ---------- ---------- ---------- ---------- ---G------

                151                                                 200
PCXB_PSEPU      GKPVPNTLVE MWQANAGGRY RHKNDRYLAP LDPNFGGVGR CLTDSDGYYS
PCXB_ACICA      GRPVKNALVE VWQANASGRY RHPNDQYIGA MDPNFGGCGR MLTDDNGYYV
PCXB_BURCE      GKPVRNTLVE VWQANAAGRY VHKVDQHDAP LDPNFLGAGR CMTDAEGRYR
CATA_ACICA      GKPLPNAKVE IWHANTKGFY SHFDP...TG EQQAFNMRRS IITDENGQYR
PHEB_PSESP      GKPIPGAKVE IWHCNSKGGY SFFDK...S. .QTPYNLRRT IIADNEGYYR
CLCA_PSEPU      GELLAGAVID VWHSTPDGLY SGIHD...NI ..PVDYYRGK LVTDSQGNYR
TCBC_PSESP      GAPAAGAVID VWHSTPDGKY SGIED...QI ..PTDMYRGK VVADAQGKYA
TFDC_ALCEU      GSVVEDVTID VWHSTPDGKY SGFED...DI ..PTDFYRGK LRVGTDGSFR
CATA            GEPLAGAKIE LWHADDLGFY SQFAP...GL ..PEWNLRGS IIADDQGNFQ
CAT             GKGLQGATVE LWHADKEGYY SQFAP...GI ..PEWNLRGT TRCDDEGNFE
Consensus       G--------- -W-----G-Y ---------- ---------- ------G---

                201                                                 250
PCXB_PSEPU      FRTIKPGPYP WRNGP..... ......NDWR PAHIHFGISG PSIATKLITQ
PCXB_ACICA      FRTIKPGPYP WRNRT..... ......NEWR PAHIHFSLIA DGWAQRLISQ
PCXB_BURCE      FLTIKPGAYP WGNHP..... ......NAWR PNHIHFSLFG DYFGSRLVTQ
CATA_ACICA      VRTILPAGYG CPPEGPTQQL LNQLGRHGNR PAHIHYFVSA DGHRK.LTTQ
PHEB_PSESP      ARSVIPSGYG VPEGAPTDQV LKLLGRHGER PAHIHYFISA PGHQH.LTTQ
CLCA_PSEPU      VRTTMPVPYQ IPYEGPTGRL LGHLGSHTWR PAHVHFKVRK DGFEP.LTTQ
TCBC_PSESP      VRTTMPAPYQ IPNKGPTGVL LEMMGSHTWR PAHVHFKVRK DGFAP.LTTQ
TFDC_ALCEU      VRTTMPVPYQ IPDQGPTGAL LETMGGHSWR PAHVHFKVKA PGYET.LTTQ
CATA            INTMQPAPYQ IPTDGACGAL IAAAGWHAWR PAHLHLKVSA PGHQL.ITTQ
CAT             ITTLKPAPYK TPADGPTGWF IDSYGGHPWR PAHLHLMVKA PGKRA.TTTQ
Consensus       -----P--Y- ---------- ---------R P-H-H-E----- ---------Q

                251                                                 300
PCXB_PSEPU      LYFEGDPLI. .PMCPIVKSI ANPEAVQQLI AKLDMNNANP MDCLAYRFDI
PCXB_ACICA      FYFEGDTLI. .DSCPILKTI PSEQQRRALI ALEDKSNFIE ADSRCYRFDI
PCXB_BURCE      MYFPGDPLL. .AYDPIPQGT P.EAARDRLI SRFSLDTTEE GHALGYEFDI
CATA_ACICA      INVAGDPYTY DDFAYATREG LVVDAVEHTD PEAIKANDVE GPFAEMVFDL
PHEB_PSESP      INLAGDPYTY DDFAFATRQD LAARGKRVEN HPAAQQYGVE GTVCEVTFNI
CLCA_PSEPU      YYFEGGKWVD DDCCHGVTPD LIT..PET.. ........IE DGVRVMTLDF
TCBC_PSESP      YYFEGGDWVD SDCCKGVAPD LVM..PTK.. ........TE GGAQVMDIDF
TFDC_ALCEU      YYFEGGDWTT DDCCNGVQSS LIT..PDI.. ........VE ECVRLMNINF
CATA            LYFEGDEHVA DDIASAVKPE LVLAPTER.. ........AD GHGREVTYHF
CAT             LYFQGGEWVE DDVATATKPE LILDPQPN.. ........AD GVD.EVTYNF
Consensus       ----G------ ---------- ---------- ---------- ----------

                301                 321
PCXB_PSEPU      VLRGQRKTHF ENC..... .
PCXB_ACICA      TLRG.RADLL RK..... .
PCXB_BURCE      VLRGRDATFM ER..... .
CATA_ACICA      KLTRLVDGVD KQVVDRPRLA V
PHEB_PSESP      ELSP...TAE EELQARP... .
CLCA_PSEPU      VIERFQAEQR KSATETVA.. .
TCBC_PSESP      VIERAREHV. .......... .
TFDC_ALCEU      VIEPARAQAG ANP..... .
CATA            VLDPQD.... .......... .
CAT             ALDPEA.... .......... .
Consensus       ---------- ---------- -
```

Sequence	Protein	Organism	Accession number	% Identity	% Similarity	Reference
PCXB_PSEPU	PCD	*Pseudomonas putida*	P00437	26	49	Frazee *et al.* (1993)
PCXB_ACICA	PCD	*Acinetobacter calcoaceticus*	P20372	28	51	Hartnett *et al.* (1990)
PCXB_BURCE	PCD	*Burkholderia cepacia*	P15110	27	47	Zylstra *at al.* (1989)
CATA_ACICA	CTD	*Acinetobacter calcoaceticus*	P07773	29	52	Neidle *et al.* (1988)
PHEB_PSESP	CTD	*Pseudomonas putida* EST1001	P31019	30	55	Kivisaar *et al.* (1991)
CLCA_PSEPU	CCD	*Pseudomonas putida* AC27	P11451	35	58	Frantz & Chakrabarty (1987)
TCBC_PSESP	CCD	*Pseudomonas putida* P51	P27098	30	57	van der Meer *et al.* (1991*a*)
TFDC_ALCEU	CCD	*Alcaligenes eutrophus* JMP134	P05403	34	56	Ghosal & You (1988)
CATA	CTD	*Arthrobacter* mA3	JT0613	56	71	Eck & Belter (1993)
CAT	CTD	*Rhodococcus rhodochrous*	_	100	100	This work

Table 5.3 Comparison of the catechol 1,2-dioxygenase from *Rhodococcus rhodochrous* with sequences obtained from the Swissprot and PIR databases

The catechol 1,2-dioxygenase from *Rhodococcus rhodochrous* was compared with other non-haem iron intradiol-cleavage enzymes taken from the Swissprot and PIR databases. Values for percentage similarity and identity were computed using the GAP program (Wisconsin Sequence Analysis Package, Genetics Computer Group, Madison, Wisconsin, USA). Abreviations used are: CTD, catechol 1,2-dioxygenase; CCD, chlorocatechol 1,2-dioxygenase; PCD, protocatechuate 3,4-dioxygenase.

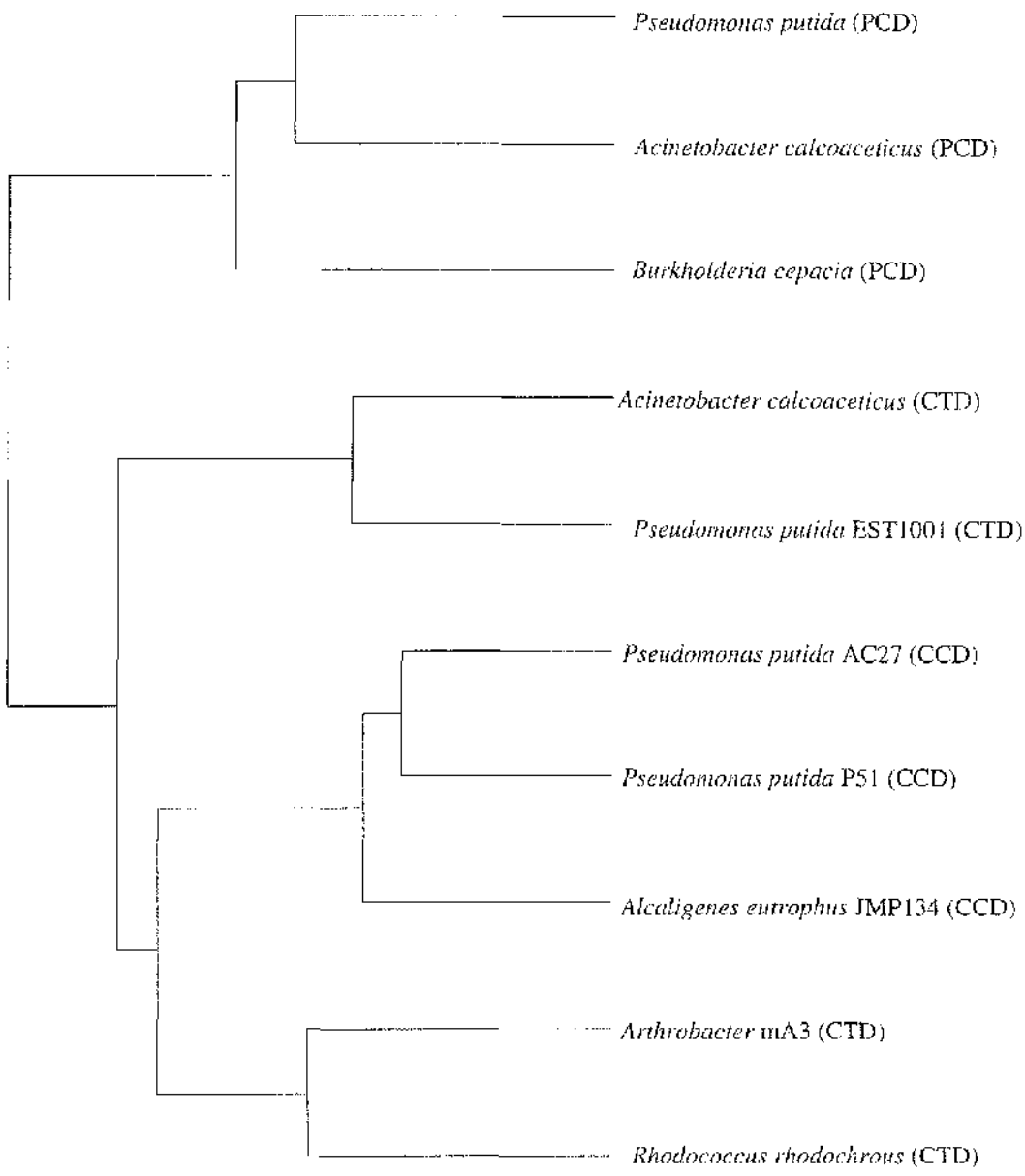

Figure 5.13 Dendrogram of amino acid sequence alignments

A dendrogram of the alignment of *Rhodococcus* catechol 1,2-dioxygenase with nine other non-haem iron intradiol-cleavage enzymes was created with the PILEUP program (Wisconsin Sequence Analysis Package, Genetics Computer Group, Madison, Wisconsin, USA). The enzymes shown here are the same as those detailed in Figures 5.12 and Table 5.3 Abreviations used are: CTD, catechol 1,2-dioxygenase; CCD, chlorocatechol 1,2-dioxygenase; PCD, protocatechuate 3,4-dioxygenase (β-subunit).

sequence similarity led Neidle *et al.* (1988) and Hartnett *et al.* (1990) to suggest that the non-haem iron intradiol-cleavage enzymes are evolutionarily homologous and probably evolved from a common ancestor by the acquisition of direct and inverted sequence repetitions of 6-10 base pairs. A dendrogram of the alignments shows four distinct sequence groups (Figure 5.13), highlighting the underlying substrate specificities (Section 1.3, Table 1.2). The *Arthrobacter* and *Rhodococcus* catechol 1,2-dioxygenases, clearly have more in common with the chlorocatechol 1,2-dioxygenases than with the *Pseudomonas* catechol 1,2-dioxygenases and this too has been supported by kinetics experiments with the *Rhodococcus* catechol 1,2-dioxygenase, which have shown that it will also degrade substituted catechols (Table 4.2).

Five Gly and two Pro are amongst the 14 amino acids which have been conserved in all the non-haem iron intradiol-cleavage enzymes (Figure 5.12). Their conservation suggests that the tertiary structures are highly conserved, since both Pro and Gly have special roles in maintaining structural integrity within proteins, where they are often found in β-turns, linking one secondary structural element with another (Richardson, 1981). In the protocatechuate 3,4-dioxygenase β-subunit, both Pro residues occur immediately after β-strands, and three of the Gly are involved in β-turns (Ohlendorf *et al.*, 1994). The α- and β-subunits of protocatechuate 3,4-dioxygenase have a common core structure consisting of an eight-stranded β-sheet, folded in half to yield two, four-stranded layers, forming the palm and back of a hand-like structure so that the fingers and palm of one hand grasp the the wrist and palm of the other (Ohlendorf *et al.*, 1994). The interface has residues which are concentrated primarily in the amino regions of both chains and in the carboxy region of the α-chain. Primary sequence alignments (Figure 5.12), show that the amino terminal region to be the least conserved, reflecting perhaps the different quaternary structures (Table 1.2). Intra-promoter charge-pair interactions derive in part from the difference in the pI between the α- and β-chains of protocatechuate 3,4-dioxygenase, which have been calculated as 5.2 and 9.5 respectively (Yoshida, 1976). The charge-pair interactions are all associated along the middle of the palms of both chains and may promote chain association at neutral pH. Whether an analogous arrangement of charge-pair interactions is available to the catechol- and chlorocatechol 1,2-dioxygenases, which are primarily homodimers with acidic pI values (Table 1.2), is not known.

The active site of protocatechuate 3,4-dioxygenase is located at the interface between the α- and β-subunits with several amino terminal residues from the α-subunit helping to form the active site (Figure 5.14) (Ohlendorf *et al.*, 1994). The conserved Tyr408, Tyr447, His460, His462 (corresponding to residues

Figure 5.14 The active site of protocatechuate 3,4-dioxygenase from *Pseudomonas putida*

The active site of protocatechuate 3,4-dioxygenase lies at the interface of the α- and β-subunits. Most of the residues, including His460, His462, Tyr408, Tyr447 (corresponding to residues 233, 235, 170 and 209 in Figure 5.12) which chelate Fe^{3+}, are contributed by the β-subunit (Ohlendorf *et al.*, 1994). Ribbons are used to portray the the β-subunit, whereas the α-subunit is shown only in stick form. All the residues which form the active site are labelled, including Tyr16 from the α-subunit. The hydroxyl ion which forms the fifth ligand to Fe^{3+} is not shown.

170, 209, 233 and 235 in Figure 5.12) chelate the active site Fe^{3+}. Immediately adjacent to the iron ligands are three hydrophilic residues; Tyr16 (in the α-subunit), Arg457 and Gln477 (Figure 5.14) (corresponding to residues 230 and 250 respectively in Figure 5.12). Tyr16 exists in the most common rotamer conformation, lying on the edge of the active site, where its terminal hydroxyl is unbonded to any other group (Ohlendorf *et al.*, 1994). By rotating the side-chain of Tyr16 its terminal hydroxyl group can nearly occupy the same position as the Fe-chelated water molecule, although this operation would cause severe steric clashes with one of the equatorial Tyr ligands (Tyr408 in Figure 5.14) (Ohlendorf *et al.*, 1994). Neidle *et al.* (1988) showed that Tyr16 is conserved in the chlorocatechol 1,2-dioxygenase from *Pseudomonas putida* pAC27 (corresponding to position 114 in Figure 5.12). This residue is retained in all other chlorocatechol 1,2-dioxygenases and in the two Gram-positive enzymes from *Arthrobacter* and *Rhodococcus*, but in the two Gram-negative catechol 1,2-dioxygenases, the equivalent position is substituted with a Leu residue (corresponding to position 114 in Figure 5.12).

The Arg457 and Gln477 (corresponding to residues 230 and 250 respectively in Figure 5.12), are conserved in every non-haem intradiol-cleavage enzyme and in the protocatechuate 3,4-dioxygenase their side-chains are hydrogen bonded with each other. Both residues are deep in the active site (Figure 5.14) and are probably buried upon substrate binding. Their position and conservation suggests that they may play an important catalytic role and Ohlendorf *et al.* (1994) proposed that the positively charged side-chain of Arg457 may stabilise the developing carbanion in the proposed mechanism (Section 1.6, Figure 1.5). The flat waist of the active site is formed by the residues Trp449 and Ile491 (corresponding to residues 211 and 266 in Figure 5.12), whose hydrophobic character matches that of the benzene ring of protocatechuate. In the catechol 1,2-dioxygenases from *Pseudomonas*, Trp449 is substituted with Cys or Val whilst in the chlorocatechol 1,2-dioxygenases and in the two enzymes from *Arthrobacter* and *Rhodococcus*, Trp449 is replaced by Ile. The other mid-site residue, Ile491, is also substituted in the catechol- and chlorocatechol 1,2-dioxygenases with Ala and Gly respectively. In general then, the bulky, hydrophobic mid-site residues of protocatechuate 3,4-dioxygenase are substituted by less bulky hydrophobic amino acids in the catechol- and chlorocatechol 1,2-dioxygenases, suggestive of a different substrate binding-site architecture.

CHAPTER 6
Conclusions

6.1 Conclusions

Many Gram-negative bacteria (Dorn & Knackmuss, 1978*a*; Aoki *et al.*, 1984*a*; Pieper *et al.*, 1988; Ngai *et al.*, 1990; Broderick & O'Halloran, 1991; Hinteregger *et al.*, 1992) and also some Gram-positive bacteria (Maltseva *et al.*, 1991), contain two types of 'catechol 1,2-dioxygenase': a chlorocatechol 1,2-dioxygenase which is characterised by the broad range of substituted catechols which it will oxygenate and a catechol 1,2-dioxygenase which oxygenates a more limited range of catechols (Section 1.3, Table 1.2). Chlorocatechol 1,2-dioxygenases are frequently plasmid encoded, whereas catechol 1,2-dioxygenases are chromosomally encoded (Section 1.7). When both are expressed simultaneously they may be separated by ion-exchange chromatography and then purified to homogeneity using hydrophobic interaction chromatography (Aoki *et al.*, 1984*a*; Ngai *et al.*, 1990; Hinteregger *et al.*, 1992). In this work, during the initial purification of the catechol 1,2-dioxygenase from *Rhodococcus rhodochrous* NCIMB13259, there was some evidence for more than one catechol 1,2-dioxygenase, since during Mono Q FPLC, both the intradiol-cleavage activity and the absorbance at 280 nm, resolved into two peaks (Section 3.1.4). However, this observation was later attributed to modification of one catechol 1,2-dioxygenase during gel-permeation chromatography, since when the gel-permeation was discontinued as a purification step, Mono Q FPLC of purified catechol 1,2-dioxygenase revealed only a single catechol 1,2-dioxygenase activity peak. Moreover, the *N*-terminal amino acid sequence analysis of the purified catechol 1,2-dioxygenase revealed only a single polypeptide (Section 4.2.8), which therefore showed that only one catechol 1,2-dioxygenase had been purified.

Warhurst (1993) showed that when *R. rhodochrous* NCIMB13259 was grown on styrene or nutrient broth and styrene, it contained *meta*-cleavage pathway enzymes and also catechol 1,2-dioxygenase activity (Section 1.8, Figure 1.9). In the present work, however, the catechol 1,2-dioxygenase from this strain was induced with benzyl alcohol, since this compound is more water soluble and less toxic than styrene. It is not known, therefore, whether the styrene- and benzyl alcohol-induced catechol 1,2-dioxygenase are the same enzyme. The possibility of multiple catechol 1,2-dioxygenases, however, was partly addressed by McIntosh (1993), who showed that partially purified extracts of *R. rhodochrous* which had been obtained after induction with benzyl alcohol or styrene were very similar in their intradiol-cleavage activity towards certain substituted catechols. Moreover, the intradiol-cleavage activity from both types of extract behaved in the same way during ion-exchange chromatography, suggesting that the styrene- and benzyl alcohol-induced catechol 1,2-dioxygenase are one and the same enzyme (McIntosh, 1993).

A common feature of the catechol- and chlorocatechol 1,2-dioxygenases is that they migrate much more slowly than expected during SDS-PAGE. Comparison of SDS-PAGE-derived M_r values with those derived from amino acid or DNA sequence data, has often resulted in overestimates as large as 10 to 20% (Ghosal & You, 1988; Neidle *et al.*, 1988; Kivisaar *et al.*, 1991; Eck & Belter, 1993). In the present work, for example, a value of 39 800 was obtained on SDS-PAGE (Section 4.2.1), whereas the values derived from electrospray mass spectroscopy and DNA sequence data were 31 598 and 31 599 respectively (Sections 4.2.6 & 5.6). Overestimates obtained both here and elsewhere, have been attributed to a conformational peculiarity and/or to the effects of the ferric ion in the enzyme (Ghosal & You, 1988; Neidle *et al.*, 1988: Kivisaar *et al.*, 1991). A less consistent phenomenon, but one that has nevertheless been reported, both in this and in other work (Dorn & Knackmuss, 1978*a*), is the anomalous behaviour of catechol 1,2-dioxygenase during gel-permeation chromatography, giving it a much larger apparent M_r value. In the present work, a M_r value of 120 000 was obtained (Section 4.2.2), which would have meant that this enzyme is a tetramer. This would be extremely rare, given that most catechol- and chlorocatechol 1,2-dioxygenase are dimers (Table 1.2, Section 4.4.1). By chemical cross-linking of the native enzyme, however, it was shown unequivocally that the catechol 1,2-dioxygenase, from *R. rhodochrous* is a dimer, with a calculated M_r value of 63 000 (Section 4.2.3). Dorn & Knackmuss (1978*a*) noticed similar discrepancies in their work on catechol 1,2-dioxygenase and chlorocatechol 1,2-dioxygenase from *Pseudomonas* B13 and were able to demonstrate that gel-permeation elution volumes could not be correlated with log M_r values, since the frictional ratios (which are a measure of ellipticity) of both catechol 1,2-dioxygenase and chlorocatechol 1,2-dioxygenase differed too much from the calibration proteins.

Since the rhodococcal catechol 1,2-dioxygenase is a dimer of identical subunits, iron:subunit stoichiometries of 1:1 were expected. Quantitation by colorimetry and atomic absorption spectroscopy however, resulted in significantly less iron than expected (Section 4.2.5, Table 4.1). The shortfall of iron in the non-haem iron intradiol-cleavage enzymes has been reported elsewhere (Nakai *et al.*, 1990; Broderick & O'Halloran, 1991) (Section 1.3) and some workers have attempted to increase the iron content of the non-haem iron intradiol-cleavage enzymes using various Fe^{2+} or Fe^{3+} reagents (Bull & Ballou, 1981; Broderick & O'Halloran, 1991). For protocatechuate 3,4-dioxygenase from *Pseudomonas putida* ATCC 23975 (formerly classified as *Pseudomonas aeruginosa*), this resulted in the attainment of the theoretical iron:subunit stoichiometry (Bull & Ballou, 1981), but for the chlorocatechol 1,2-dioxygenase from *Pseudomonas*

putida pAC27, there was no increase in iron content (Broderick & O'Halloran, 1991). The observed 'deficit' is unlikely therefore to be due to a iron-leakage during purification. Since the shortfall could be a reflection of a shortage of available iron in the growth medium, an attempt was made to increase the iron content of rhodococcal catechol 1,2-dioxygenase by incorporating a chelated metals solution into the growth medium, but this treatment had no significant change on the iron content and the iron:subunit stoichiometry of all proteins which had been purified from *R. rhodochrous* (as measured by atomic absorption spectroscopy), remained very similar (0.60, 0.60. and 0.65; Table 4.1). Since the concentration of free Fe^{3+} in aerobic solution at neutral pH is very low (1×10^{-17} M) (Guerinot, 1994) the observed deficit may be because demand outstrips supply i.e., that the intradiol-cleavage enzymes are synthesised faster than iron can be mobilised. In order to find out whether the marginally lower iron content found in the over-expressed rhodococcal catechol 1,2-dioxygenase (0.5 Fe per subunit, Table 4.1) is significant in this regard, it would be necessary to measure the iron content for several over-expressed proteins.

After the gene encoding the catechol 1,2-dioxygenase from *R. rhodochrous* had been cloned (Chapter 5), the derived amino acid sequence (Figure 5.7) was aligned with other non-haem iron intradiol-cleavage enzymes (Figure 5.12). The rhodococcal catechol 1,2-dioxygenase was more akin to the chlorocatechol 1,2-dioxygenases (30-35% identities) than to the catechol 1,2-dioxygenases (29-30% identities). The closest match was with the catechol 1,2-dioxygenase from *Arthrobacter* mA3 (56% identity), which is the only other catechol 1,2-dioxygenase to have been cloned from a Gram-positive bacterium (Eck & Belter, 1993). This strain of *Arthrobacter* can use 3-aminophenol as a sole source of carbon (Eck & Belter, 1993) but its catechol 1,2-dioxygenase has been neither purified nor characterised and so it is not possible to say how far the primary structure may be correlated with function. For the rhodococcal catechol 1,2-dioxygensase, however, the sequence data seem to support the kinetic data, in that there is scarcely any discrimination between the substrates catechol, 3-methylcatechol and 4-methylcatechol (Section 4.3.4, Tables 4.2 & 4.3). The absence of any extradiol activity, is also reminiscent to the Gram-negative chlorocatechol 1,2-dioxygenases (Fujiwara *et al.*, 1975; Dorn & Knackmuss, 1978*a*). In other respects, such as its stability in acidic solutions (Section 4.3.3), the rhodococcal enzyme is unlike the Gram-negative chlorocatechol 1,2-dioxygenases, which are extremely labile in acidic pH solutions (Dorn & Knackmuss, 1978*a*; Broderick & O'Halloran, 1991). Unfortunately, due to the lack of availability of 3-vinylcatechol, it was not possible to determine if this metabolite of styrene (Section 1.8, Figure 1.9) could be effectively oxygenated by

the catechol 1,2-dioxygenase which had been purified in this work (Chapter 3), but since there is no evidence that *R. rhodochrous* contains more than one catechol 1,2-dioxygenase, it seems likely that this compound would be oxygenated.

The degree of overall sequence identity and the conservation of Pro and Gly residues suggests that the non-haem intradiol-cleavage enzymes have similar secondary and tertiary structures as well (Figure 5.12). Using the known three-dimensional structure of protocatechuate 3,4-dioxygenase from *Pseudomonas putida* ATCC 23975 (formerly classified as *Pseudomonas aeruginosa*) (Ohlendorf *et al.*, 1988; Ohlendorf *et al.*, 1994) and sequence alignments with other non-haem intradiol-cleavage enzymes (Figures 5.14 & 5.12 respectively), the divergent pattern of substrate usage (Table 1.2), may be related to amino acid substitutions found within the putative active sites of catechol- and chlorocatechol 1,2-dioxygenases. For example, the residues Trp449 and Ile491 (corresponding to residues 211 and 266 in Figure 5.12) from the β-subunit, are purported to sandwich the aromatic ring of protocatechuate with Gly, Pro and Tyr residues from the *N*-terminus of the α-subunit (Ohlendorf *et al.*, 1994). The corresponding Trp residue is retained in every protocatechutae 3,4-dioxygenase (Figure 5.12), but in the Gram-negative catechol 1,2-dioxygenases it is substituted with a much smaller hydrophobic residue (Cys or Val) whilst in the chlorocatechol 1,2-dioxygenases and in the 2 enzymes from *Arthrobacter* and *Rhodococcus*, Trp449 is replaced by Ile (Figure 5.12). It is notable that in the protocatechuate 3,4-dioxygenase Trp449 (residue 211 in Figure 5.12) is followed by a β-loop (GPND) and then the remaining two His active site residues (Ohlendorf *et al.*, 1988; Ohlendorf *et al.*, 1994) (residues 233 & 235 in Figure 5.12), whereas in the catechol- and chlorocatechol 1,2-dioxygenases there is a sequence of nine residues in between a putative β-loop (residues 215-218) and the two His residues, which have no apparent equivalent in the protocatechuate 3,4-dioxygenases (Figure 5.12). The other mid-site amino acid, Ile491 (residue 266 in Figure 5.12), is substituted by Ala in both the Gram-positive and Gram-negative catechol 1,2-dioxygenases and by Gly in the chlorocatechol 1,2-dioxygenases. The pattern of substitution in and around the putative active sites seems to reflect the divergent pattern of substrate usage (Table 1.2), but the logic underwriting this is not known.

Putative LysR regulatory sequence motifs (Rothmel *et al.*, 1990) were found near a putative −10 promoter element and initiation site (Figures 5.7 & 5.8), which raises the question of how the expression of catechol 1,2-dioxygenase is regulated in *R. rhodochrous*. The fact that catechol 1,2-dioxygenase is expressed constitutively in *E. coli* in the absence of isopropylthiogalactoside suggests that

the enzyme is under negative control (Section 5.7). Regulatory sequences which overlap the initiation site are also consistent with negative regulation (Collado-Vides *et al.*, 1991), however, LysR proteins are positive regulators (Frantz & Chakrabarty, 1987; Kasak *et al.*, 1993; Matrubutham & Harker, 1994) of catechol- and chlorocatechol 1,2-dioxygenase in Gram-negative bacteria and normally recognise sequence motifs lying beyond the −35 promoter sequence, which raises the possibility that there may be a novel mechanism of transcriptional control.

6.2 Future work

The most pressing needs are to follow up the structural work and the molecular biology.

1. A complete native data set should be collected from the catechol 1,2-dioxygenase crystals. The X-ray crystallographic structure determination of catechol 1,2-dioxygenase could then be obtained in one of three ways:

a. Since the sequence data suggest that the non-haem iron intradiol-cleavage enzymes are homologous, a molecular replacement solution could be tried initially, using the coordinates of the β-subunit of protocatechuate 3,4-dioxygenase.

b. If the molecular replacement experiment is unsuccessful, then conventional heavy atom soak techniques would be initiated. The presence of two reactive Cys in each subunit would be a great advantage in the search for the appropriate heavy atoms, since mercury reacts preferentially with thiols.

c. Taking advantage of the anomalous X-ray signal of the ferric ion and determining the Fe^{3+} positions directly from a multiwavelength anomalous dispersion (MAD) experiment using the high energy, tuneable X-ray source on station 9.5 at CCLR, Daresbury, UK.

The structure of catechol 1,2-dioxygenase could then be compared with protocatechuate 3,4-dioxygenase. The mechanism of substrate specificity could then be probed by doing crystal soaks with substrate analogues and also by preparing a series of mutants by site-directed mutagenesis. Steady state kinetics experiments could then be carried out on each mutant using a broad range of substituted catechols, so that the structure could be correlated with function.

2. The entire DNA sequence downstream of the catechol 1,2-dioxygenase gene should be sequenced to see if it contains any other genes of the β-ketoadipate pathway. The evidence so far suggests that they do not exist, however, Warhurst (1993) postulated that since catechol 1,2-dioxygenase is often expressed separately from the remaining members of the pathway (Figure 1.7), the absence of expression may represent a fault in regulation.

A more long-term goal would be to investigate the mechanism of expression of catechol 1,2-dioxygenase, since this has not been investigated in Gram-positive bacteria and the evidence thus far suggests that there may be a novel mechanism of transcriptional control. Since the genes encoding catechol 1,2-dioxygenases frequently have contiguous, divergently transcribed regulator genes, the DNA upstream of the catechol 1,2-dioxygenase gene should be cloned to see if there are coding regions which bear any similarity to known regulatory gene products.

REFERENCES

Akhtar, M. & Wright, J. N. (1991) Natural Product Reports **8**, 527-551

Andersson, S. G. E. & Kurland, C. G. (1990) Microbiol. Rev. **54**, 198-210

Aoki, K., Konohana, T., Shinke, R. & Nishira, H. (1984*a*) Agric. Biol. Chem. **48**, 2087-2095

Aoki, K., Konohana, T., Shinke, R. & Nishira, H. (1984*b*) Agric. Biol. Chem. **48**, 2097-2104

Appel, M., Raabe, T. & Lingens, F. (1984) FEMS Microbiol. lett. **24**, 123-126

Barrowman, M. M. & Fewson, C. A. (1985) Curr. Microbiol. **12**, 235-240

Beggs, J. D. & Fewson, C. A. (1977) J. Gen. Microbiol. **103**, 127-140

Bej, A. K., Mahbubani, M. H. & Atlas, R. M. (1991) Critical Rev. Biochem. Mol. Biol. **26**, 301-334

Bhat, M. A., Ishida, T., Horiike, K., Vaidyanathan, C. S. & Nozaki, M. (1993) Arch. Biochem. Biophys **300**, 738-746

Bibb, M. J., Jansen, G. R. & Ward, J. M. (1985) **41**, 357-368

Blundell, T. L. & Johnson, L. N. (1976) in Protein Crystallography, pp 59-82, Academic Press, London, UK

Boldt, Y. R., Sadowsky, M. J., Ellis, L. B. M., Que, L. Jr. & Wackett, L. P. (1995) J. Bacteriol. **177**, 1225-1232

Bradford, M. M. (1976) Anal. Biochem. **72**, 248-254

Broderick, J. & O'Halloran, T. V. (1991) Biochemistry **30**, 7349-7358

Bruce, N. C., Cain, R. B., Pieper, D. H. & Engesser, K. -H. (1989) Biochem. J. **262**, 303-312

Bull, C. B. & Ballou, D. P. (1981) J. Biol. Chem. **256**, 12673-12680

Bull, C. B., Ballou, D. P. & Otsuka, S. (1981) J. Biol. Chem. **256**, 12681-12686

Burlage, R. S., Hooper, S. W. & Sayler, G. S. (1989) Appl. Environ. Microbiol. **55**, 1323-1328

Chakrabarty, A. M. (1996) ASM News **62**, 130-137

Chatterjee, D. K. (1991) J. Bacteriol. **143**, 639-646

Chen, Y. P. & Lovell, C. R. (1990) Appl. Environ. Microbiol. **56**, 1971-1973

Coco, W. M., Rothmel, R. K., Henikoff, S. & Chakrabarty, A. M. (1993) J. Bacteriol. **175**, 417-427

Collado-Vides, J., Magasanik, B. & Gralla, J. D. (1991) Microbiol. Rev. **55**, 371-394

Criegee, R. (1948) Liebig. Ann. Chem. **560**, 127

Davies, G. E. & Stark, G. R. (1970) Proc. Natl. Acad. Sci. **66**, 651-656

Don, R. H., Weightman, A. J., Knackmuss, H.-J. & Timmis, K. N. (1985) J. Bacteriol. **161**, 85-90

Dorn, E. & Knackmuss, H.-J. (1978*a*) Biochem. J. **175**, 73-84

Dorn, E. & Knackmuss, H.-J. (1978*b*) Biochem. J. **175**, 85-94

Ducroix, A. & Giege, R., (1992) in Crystallisation of Nucleic Acids and Proteins (Ducroix, A. & Giege, R., ed.), pp 73-97, Oxford University Press, New York, USA

Earhart, C. A., Radhakrishnan, R., Orville, A.M., Lipscomb, J. D. & Ohlendorf, D. H. (1994*a*) J. Mol. Biol. **236**, 374-376

Earhart, C. A., Hall, M. D., Michaudsoret, I., Que, L. & Ohlendorf, D. H. (1994*b*) J. Mol. Biol. **236**, 377-378

Eck, R. & Belter, J. (1993) Gene **123**, 87-92

Edelhoch, H. (1967) Biochemistry **6**, 1948-1954

Embley, T. M. & Stackebrandt, E. (1994) Annu. Rev. Microbiol. **48**, 257-289

Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F. & Whitehouse, C. M. (1989) Science **246,** 64-71

Fersht, A. (1985) Enzyme Structure and Mechanism,W. H. Freeman, New York, USA

Fewson, C. A. (1991) in The Biology of *Acinetobacter* (Towner, K. D. *et al.*, ed.), pp 351-390, Plenum Press, New York, USA

Finnerty, W. R. (1992) Annu. Rev. Microbiol. **46**, 193-218

Fish, W. W. (1988) Methods Enzymol. **158**, 357-364

Frantz, B. & Chakrabarty, A. M. (1987) Proc. Natl. Acad. Sci. USA **84**, 4460-4464

Fraser, M. S. & Hamilton, G. A. (1982) J. Am. Chem. Soc. **104**, 4203-4211

Frazee, R. W., Livingstone, D. M., LaPorte, D. C. & Lipscomb, J. D. (1993) J. Bacteriol. **175**, 6194-6202

Fuchs, K., Schreiner, A. & Lingens, F. (1991) J. Gen. Microbiol. **137**, 2033-2039

Fujisawa, H., Uyeda, M., Kojima, Y., Nozaki, M. & Hayaishi, O. (1972) J. Biol. Chem. **247**, 4414-4421

Fujiwara, M., Golovleva, L. A., Saeki, Y., Nozaki, M. & Hayaishi, O. (1975) J. Biol. Chem. **250,** 4848-4855

* Harayama, S., Rekik, M., Ngai, K. L. & Ornston, L. N. (1989) J. Bact. **171**, 6251-6258

Guerinot, M. L. (1994) Annu. Rev. Microbiol. **48**, 743-772

Ghosal, D. & You, I.-S. (1988) Mol. Gen. Genet. **211**, 113-120

Ghosal, D., You, I.-S., Chatterjee, D. K & Chakrabarty, A. M. (1985) Science **228**, 135-142

Gill, S. C. & von Hippel, P. H. (1989) Anal. Biochem. **182**, 319-326

Goethals, K., Montagu, M. V. & Holsters, M. (1992) Proc. Natl. Acad. Sci. USA **89**, 1646-1650

Gold, L., Pribnow, D., Schneider, T., Shinedling, S., Singer, B. S. & Stormo, G. (1981) Annu. Rev. Microbiol. **35**, 365-403

Gourley, D. & Emsley, P. (1995) Personal communication, Department of Chemistry, University of Glasgow

Habeeb, A. F. S. A. (1972) Methods Enzymol. **25**, 457-464

Hajdu, J., Bartha, F. & Friedrich, P. (1976) Eur. J. Biochem. **68**, 373-383

Hamilton, G. A. (1974) in Molecular Mechanisms of Oxygen Activation (Hayaishi, O., ed.), pp 405-445, Academic Press, New York, USA

Harayama, S., Kok, M. & Neidle, E. L. (1992) Annu. Rev. Microbiol. **46**, 565-601

Harayama, S. & Rekik, M. (1989) J. Biol. Chem. **264**, 15328-15333

Harley, C. B. & Reynolds, R. P. (1987) Nucl. Acids Res. **15**, 2343-2361

Hartnett, C., Neidle, E., Ngai, K. -L. & Ornston, L. N. (1990) J. Bacteriol. **172**, 956-966

Harwood, C. S. & Parales, R. E. (1996) Annu. Rev. Microbiol. **50**, 553-590

Harwood, C. S., Rivelli, M. & Ornston, L. N. (1984) J. Bacteriol. **160**, 622-628

Hayaishi, O., Katagiri, M. & Rothberg, S. (1955) J. Am. Chem. Soc. **77**, 5450

Hayaishi, O., Katagiri, M. & Rotherberg, S. (1957) J. Biol. Chem. **229**, 905-920

Hedrick, J. L. & Smith, A. J. (1968) Arch. Biochem. Biophys. **126**, 155-164

Hinteregger, C., Loidl, M. & Streichsbier, F. (1992) FEMS Microbiol. Lett. **97**, 261-266

Holm, L. (1986) Nucl. Acids Res. **14**, 3075-3087

Houghton, J. E., Brown, T. E., Appel, A. J., Hughes, J. & Ornston, L. N. (1995) J. Bacteriol. **77**, 401-412

Hucho, F., Mullner, H. & Sund, H. (1975) Eur. J. Biochem. **59**, 79-87

Hughes, E. J., Shapiro, M. J., Houghton, J. E. & Ornston, L. N. (1988) J. Gen. Microbiol. **134**, 2877-2887

Hybond-N (1995) Amersham International, Little Chalfont, Bucks, UK

Itoh, M. (1989) Agric. Biol. Chem. **45**, 2787-2797

Jancarik, J. & Kim, S.-H (1991) J. Appl. Cryst. **24**, 409-411

Kabsch, W. (1987) J. Appl. Cryst. **20**, 235-242

Kane, J. F. (1995) Curr. Op. Biotechnol. **6**, 494-500

Kasak, L., Horak, R., Nurk, A., Talvik, K. & Kivisaar, M. (1993) J. Bacteriol. **175**, 8038-8042

Kersten, P. J., Chapman, P. J. & Dagley, S. (1985) J. Bacteriol. **162**, 693-697

Kivisaar, M., Kasak, L. & Nurk, A. (1991) Gene **98**, 15-20

Kojima, Y., Fujisawa, H., Nakazawa, A., Nakazawa, T., Kanetsuna, F.,Tanuichi, H. Nozaki, M. & Hayaishi, O. (1967) J. Biol. Chem. **242**, 3271-3278

Kukor, J., Olsen, R. H. & Ballou, D. P. (1988) J. Bacteriol. **170**, 4458-4465

Kukor, J. & Olsen, R. H. (1991) J. Bacteriol. **173**, 4587-4594

Laemmli, U. K. (1970) Nature, London **227**, 680-685

Ludwig, M. L. & Hunter, M. J. (1968) Methods Enzymol. **11**, 595-604

Malmström, B. G. (1982) Annu. Rev. Biochem. **51**, 21-59

Maltseva, O. V., Solyanikova, I. P. & Golovleva, L. A. (1994) Eur. J. Biochem. **226**, 1053-1061

Margolis, J. & Kendrick, K. G. (1968) Anal. Biochem. **25**, 347-362

Martin, F. H., Castro, M. M., Aboul-ela, F. & Tinoco, I. Jr. (1985) Nucl. Acids Res. **13**, 8927

Mason, H. S., Fowlks, W. L. & Peterson, L. (1955) J. Am. Chem. Soc. **77**, 2914

Mathews, B. W. (1968) J. Mol. Biol. **33**, 491-497

Matrubutham, U. & Harker, A. R. (1994) J. Bacteriol. **176**, 2348-2353

McIntosh, K.J. (1993) The catechol oxygenases of *Rhodococcus rhodochrous* induced by growth on styrene and benzyl alcohol. B.Sc. dissertation, University of Glasgow

McLaughlin, J. R., Murray, C. L. & Rabinowitz, J. C. (1981) J . Biol. Chem. **246**, 11283-11291

van der Meer, J. R., Eggen, R. I. L., Zehnder, A. J. B. & de Vos, W. M. (1991*a*) J. Bacteriol. **173**, 2425-2434

van der Meer, J. R., van Neervan, A. R. W., de Vries, E. J., de Vos, W. M. & Zehnder, A. J. B. (1991*b*) J. Bacteriol. **173**, 6-15

Miguez, C. B., Greer, C. W. & Ingram, J. M. (1993) Can. J. Microbiol. **39**, 1-5

Murakimi, S., Nakanishi, Y., Shinke, R. & Aoki, K. (1991) Soil Biology & Biochem. (1991) **23**, 815-819

Nakai, C., Horiike, K., Kuramitsu, S., Kagamiyama, H. & Nozaki, M. (1990) J. Biol. Chem. **265**, 660-665

Nakai, C., Uyeyama, H., Kagamiyama, H., Nakazawa, T., Inouye, S., Kishi, F., Nakazawa, A. & Nozaki, M. (1995) Arch. Biochem. Biophys. **321**, 353-362

Neidle, E. L., Hartnett, C., Bonitz, S. & Ornston, N. L. (1988) J. Bacteriol. **170**, 4874-4879

Neidle, E. L. & Ornston, L. N. (1987) J. Bacteriol. **169**, 414-415

Ngai, K.-L., Neidle, E. L. & Ornston, L. N. (1990) Methods Enzymol. **188**, 122-126

Ngai, K.-L. & Ornston, N. L. (1988) J. Bacteriol. **170**, 2412-2413

Nichols, N. N. & Harwood, C. S. (1995) J. Bacteriol. **177**, 7033-7040

Nozaki, M. (1979) Top. Curr. Chem. **78**, 145-186

Ohlendorf, D. H., Orville, A. M. & Lipscomb, J. D. (1994) J. Mol. Biol. **244**, 586-608

Ohlendorf, D. H., Lipscomb, J. D. & Weber, P. C. (1988) Nature **336**, 403-405

Orville, A. M. & Lipscomb, J. D. (1989) J. Biol. Chem. **264**, 8791-8801

Parsek, M. R., Ye, P. & Chakrabarty, A. M. (1994) J. Biol. Chem. **269**, 11279-11284

Parsek, M. R., Shinabarger, D. L., Rothmel, R. K. & Chakrabarty, A. M. (1992) J. Bacteriol. **174**, 7798-7806

Patel, R. N., Hou, C. T., Felix, A. & Lilliard, M.O. (1976) J. Bacteriol. **127**, 536-544

Petillot, Y., Forest, E., Mathieu, I. & Moulis, J.-M. (1993) Biochem. J. **296**, 657-661

Petillot, Y., Golielli, M.-P., Forest, E. & Meyer, J. (1995) Biochem. Biophys. Res. Commun. **210**, 686-694

Pettigrew, C. A., Haigler, B. E. & Spain, J. C. (1991) Appl. Environ. Microbiol. **57**, 157-162

Pieper, U. H., Reineke, W., Engasser, K.-H., & Knackmuss, H.-J. (1988) Arch. Microbiol. **150**, 95-102

Promega (1996) Biological Research Products, Madison, Wisconsin, USA

Que, L. Jr., Lipscomb, J. D., Munck, E. & Wood, J. M. (1977) Biochim. Biophys. Acta. **17**, 1853-1860

Que, L. Jr. (1980) Struct. Bonding **40**, 39-72

Que, L. Jr. & Epstein, R. M. (1981) Biochemistry **20**, 2545-2549

Que, L. Jr. & Heistand, R. H. (1979) J. Am. Chem. Soc. **101**, 2219-2221

Que, L. Jr., Heistand, R. H., Mayer, R. H. & Roe, A. L. (1980) Biochemistry **19**, 2588-2593

Que, L. Jr., Lauffer, R. B., Lynch, J. B., Murch, B.P. & Pyrz, J. W. (1987) J. Am. Chem. Soc. **109**, 5381-5385

Rast, H. G., Engelhardt, G. & Wallnöfer, P.R. (1980) FEMS Microbiol. Lett. **7**, 1-6

Reineke, W. & Knackmuss, H.-J. (1988) Annu. Rev. Microbiol. **42**, 263-287

Richardson, J. S. (1981) Adv. Prot. Struct. **34**, 167-339

Robinson, M., Lilley, R., Little, S., Emtage, J. S., Yarranton, G., Stevens, P., Millican, A., Eason, R. & Brown, A. J. P. (1984) Nucl. Acids Res. **12**, 6663-6671

Romero-Arroyo, C. E., Schell, M. A., Gaines III, G.L. & Neidle, E. L. (1995) J. Bacteriol. **177**, 5891-5898

Romero-Steiner, S., Parales, R., Harwood, C. & Houghton, J. (1994) J. Bacteriol. **176**, 5771-5779

Rothmel, R. K., Shinabarger, D. L., Parsek, M. R., Aldrich, T. L. & Chakrabarty, A. M. (1991) J. Bacteriol. **173**, 4717-4724

Saure-Ignazi, G., Gagnon, J., Beguin, C., Barrelle, M., Markowicz, Y., Pelmonot, J. & Toussaint, A. (1996) Arch. Microbiol. **166**, 42-50

Saito, H. & Miura, K. -I. (1963) Biochim. Biophys. Acta. **72**, 619-629

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) Molecular Cloning: a Laboratory Manual, Cold Spring Harbor, New York, USA

Sawaki, Y. & Ogata, Y. (1978) J. Am. Chem. Soc. **100**, 856-860

Schell, M. A. (1993) Annu. Rev. Microbiol. **47**, 597-626

Schreiner, A., Fuchs, K., Lottspeich, F., Poth, H. & Lingens, F. (1991) J. Gen. Microbiol. (1991) **137**, 2042-2048

Siegel, L. M. & Monty, K. J. (1966) Biochim. Biophys. Acta. **112**, 346-362

Southern, E. (1975) J. Mol. Biol. **98**, 503-517

Stanier, R. Y. & Ingraham, J. L. (1954) J. Biol. Chem. **210**, 799-808

Sterjiades, R. & Pelmont, J. (1989) Appl. Environ. Microbiol. **55**, 349-347

Strohl, W. R. (1992) Nucl. Acids Res. **20**, 961-974

Stura, E. A. & Wilson, I. A. (1992) in Crystallisation of Nucleic Acids and Proteins (Ducroix, A. & Giege, R., ed.), pp 99-125, Oxford University Press, New York, USA

Takaichi, S., Ishidsu, J., Seki, T. & Fukada, S. (1990) Agric. Biol. Chem. **54**, 1931-1937

Tatsuno, Y., Saeki, Y., Iwakai, M., Yagi, T., Nozaki, M., Kitagawa,T. & Otsuka, S. (1978) J. Am. Soc. **100**, 4614-4615

Taylor, P. L. (1991) Biosoft GeneJockey sequence processor

Timmis, K. N., Steffan, R. J. & Unterman, R. (1994) Annu. Rev.Microbiol. **48**, 525-557

True, A. E., Orville, A. M., Pearce, L. L., Lipscomb, J. D. & Que, L. Jr. (1990) Biochemistry **29**, 10847-10854

Tyson, C. A. (1975) J. Biol. Chem. **250**, 1765-1770

Vetting, M.W., Earhart, C. A. & Ohlendorf, D. H. (1994) J. Mol. Biol. **236**, 372-373

Walsh, T. A. & Ballou, D. P. (1983) J. Biol. Chem. **258**, 14413-14421

Walsh, T. A., Ballou, D. P., Mayer, R. & Que, L. Jr. (1983) J. Biol. Chem. **258**, 14422-14427

Warhurst, A. M. (1993) Metabolism of styrene in *Rhodococcus rhodochrous*. PhD thesis, University of Glasgow

Warhurst, A. M., Clarke, K. F., Hill, R. A., Holt, R. A. & Fewson, C. A. (1994) Appl. Environ. Microbiol. **60**, 1137-1145

Whittaker, J. W., Lipscomb, J. D., Kent, T. A. & Munck, E. (1984) J. Biol. Chem. **259**, 4466-4475

Williams, P. A. (1985) Enzpack, Elsevier-BIOSOFT, Cambridge, UK

Yoshida, R., Hori, K., Fujiwara, M., Saeki, Y., Kagamiyama, H. & Nozaki, W. (1976) Biochemistry **15**, 4048-4053

Zylstra, G. J., Olsen, R. H. & Ballou, D. P. (1989) J. Bacteriol. **171**, 5907-5914