Benzyl alcohol dehydrogenase from
Acinetobacter calcoaceticus

David John Gillooly

Thesis submitted for the degree of Doctor of Philosophy

Division of Biochemistry and Molecular Biology
Institute of Biomedical and Life Sciences
University of Glasgow
April 1996

Dedicated to my Mum and Dad

Contents

Chapter 3. Cloning and sequencing of *xylB* (the gene encoding benzyl alcohol dehydrogenase) from A. calcoaceticus NCIB8250

List of Figures

Chapter 4.

Chapter 5.

List of Tables

Abbreviations

Abbreviations used are those recommended by The Biochemical Journal's instructions to authors [Biochem. J. (1996) **313**, 1-15] with the following additions:

BADH	Benzyl alcohol dehydrogenase
BSA	Bovine serum albumen
BZDHII	Benzaldehyde dehydrogenase II
DTT	Dithiothreitol
HLADH	Horse liver alcohol dehydrogenase
K'_m	Apparent Michaelis constant
NAD(H)	Nicotinamide adenine dinucleotide (oxidised form)
NCIB	National Collection of Industrial Bacteria, Aberdeen, U.K
OD	Optical density
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
p.s.i	Pounds per square inch
RNase	Ribonuclease A
r.p.m	Revolutions per minute
SDS	Sodium dodecyl (lauryl) sulphate
Tris	Tris (hydroxymethyl) aminomethane
V'_{max}	Apparent maximum velocity

Acknowledgements

I would like to express my warmest thanks to Professor Charles Fewson for his enthusiastic support and supervision throughout my project. I would also like to thank him, along with Professors Miles Houslay and Gordon Lindsay for making the facilities of the Division of Biochemistry and Molecular Biology available to me.

I would also like to thank Professor P. A. Williams for drawing our attention to the *xylW* sequence.

For help in the lab I would like to thank all the members of the lab, in particular Alan Scott, Phil Strachan, Alan Robertson, John Greene, Tino Krell Gillian Muir, Dr Les Fixter and Dr Ian Hamilton.

I would also like to thank my friends who have helped me throughout my time in Glasgow, especially John Christie, Jon Gilbert, Dave Gourlay (plus extra thanks for help with the molecular modelling) and Malcolm Horsburgh.

Summary

1. *Acinetobacter calcoaceticus* NCIB8250 can grow on benzyl alcohol as sole carbon and energy source. The benzyl alcohol pathway involves the NAD-dependent enzymes benzyl alcohol dehydrogenase and benzaldehyde dehydrogenase II. These two enzymes had previously been purified from *A. calcoaceticus* and *xylC*, the gene encoding benzaldehyde dehydrogenase II, had been sequenced. This thesis describes work which completed the cloning of *xylB*, the gene encoding benzyl alcohol dehydrogenase, the sequencing of *xylB* and further studies of this enzyme with particular emphasis on its substrate specificity and proton relay mechanism.

2. A clone from an *A. calcoaceticus* lambda genomic DNA library which in a previous study had been proposed to harbour the full lengths of both *xylC* and *xylB* was sub-cloned into pBluescript SK II and a series of sub-clones developed so that benzyl alcohol dehydrogenase could be expressed in *Escherichia coli* DH5α with and without benzaldehyde dehydrogenase II.

3. The nucleotide sequence of *xylB* was determined. The *xylB* structural gene is 1110 bp in length. The nucleotide sequence, together with the observation that both *xylC* and *xylB* are transcribed together from the *lac* promoter of pBluescript SK II after induction with IPTG, suggests that the two genes form part of an operon transcribed in the direction $xylC{\rightarrow}xylB$. A possible σ factor-independent transcription terminator sequence was identified downstream of *xylB*.

4. The nucleotide sequence of *xylB* indicated that benzyl alcohol dehydrogenase is a 370 amino acid protein with a relative molecular mass (M_r) of 38,923. Benzyl alcohol dehydrogenase could be purified from *E. coli* DH5α/pDG20 using the method previously described for purification of the enzyme from *A. calcoaceticus*. *N*-Terminal amino acid analysis confirmed that the purified enzyme was the *xylB* product. The M_r of purified benzyl alcohol dehydrogenase was determined to be 38,929 ± 8 by electrospray mass spectometry, thus confirming the nucleotide sequence.

5. A *xylB* clone was constructed using PCR so that a T7 RNA polymerase expression system could be used to over-express benzyl alcohol dehydrogenase. This system enabled benzyl alcohol dehydrogenase to be expressed at a level of at least 50% of the total soluble protein in *E. coli* JM109(DE3) and enabled the purification of greater than 10 mg quantities of enzyme from less than 10 g of cells.

6.   Alignment of the primary sequence of benzyl alcohol dehydrogenase with members of the group of NAD(P)-dependent long-chain zinc-dependent alcohol dehydrogenases showed benzyl alcohol dehydrogenase to be a member of this group of enzymes. In particular, the zinc ligands of horse liver alcohol dehydrogenase are all conserved in benzyl alcohol dehydrogenase and this suggests that the enzyme probably binds two zinc atoms per enzyme subunit one catalytic and one structural. The *A. calcoaceticus* enzyme shares 54% sequence identity with benzyl alcohol dehydrogenase encoded by the TOL-plasmid pWW0 of *P.seudomonas putida mt*-2 and greater than 30% sequence identity with several mammalian and plant zinc-dependent alcohol dehydrogenases. These enzymes all appear to share a common evolutionary ancestor.

7.   Apparent K_m and V_{max} values for benzyl alcohol oxidation were determined for benzyl alcohol dehydrogenase and horse liver alcohol dehydrogenase and also for ethanol oxidation by horse liver and yeast alcohol dehydrogenases. In general the microbial enzymes were the most efficient (higher "specificity constants") and these differences could be related to differences in the primary structures and to selection for the physiological roles of the enzymes. Also, primary structure differences could be correlated with differences in the substrate specificities of these enzymes.

8.   Residue His-51 of horse liver alcohol dehydrogenase has been proposed to act as a general base for catalysis in this enzyme by mediating transfer of a proton from the alcohol substrate through a proton relay system to the imidazole ring of this residue and then to the surface of the enzyme. His-51 is conserved in a large proportion of the zinc-dependent alcohol dehydrogenases, or else it is functionally replaced, and the proposed proton relay system is thought to be essential to zinc-dependent alcohol dehydrogenases in general. His-51 is replaced by Ile in benzyl alcohol dehydrogenase from *A. calcoaceticus*, whereas the other components of the proposed proton relay system are conserved. Reasons for the replacement of His-51 with an apparently non-functional replacement were investigated by "reintroducing" His-51 into benzyl alcohol dehydrogenase by site-directed mutagenesis. The mutant enzyme benzyl alcohol dehydrogenase-His51 was expressed and purified and was very similar to wild-type enzyme in terms of substrate specificity, pH dependence and efficiency of benzyl alcohol oxidation, even though the mutant enzyme enzyme had somewhat higher K'_m and V'_{max} values than the wild-type enzyme.

9. The histidine modifying reagent diethyl pyrocarbonate inactivated benzyl alcohol dehydrogenase. This inactivation could be protected against by NADH, NAD$^+$, and benzyl alcohol, with NADH and NAD$^+$ giving much more protection than benzyl alcohol. Carbethoxylated enzyme could be reactivated by treatment with hydroxylamine. It seems likely that inactivation was due to modification of an essential histidine residue, although it is possible that this residue is His-47, a residue thought to be necessary for coenzyme binding.

10. A possible role for Arg-53 in the proton transfer mechanism of benzyl alcohol dehydrogenase was investigated by site-directed mutagenesis. Ala-53 and His-53 mutant enzymes were created. The Ala-53 mutant could oxidise benzyl alcohol as efficiently as wild-type enzyme, indicating that Arg-53 does not act as a general base for catalysis. Benzyl alcohol dehydrogenase-His53 had a specificity coefficient for benzyl alcohol oxidation three orders of magnitude lower than the wild-type enzyme and could oxidise perillyl alcohol much more efficiently than it could benzyl alcohol.

11. *A. calcoaceticus* NCIB8250 can grow on L(+)-mandelate as sole carbon and energy source, using a pathway that converges with the benzyl alcohol utilising pathway. Attempts were made to clone the gene encoding L-mandelate dehydrogenase (*mdlB*), the first enzyme in the mandelate pathway. PCR using degenerate primers based on *N*-terminal protein sequence determined from L-mandelate dehydrogenase purified from *A. calcoaceticus*, random priming with *mdlB* from *P. putida* and radiolabelled degenerate oligonucleotides based on *N*-terminal amino acid sequence and a sequence motif found to be conserved among other 2-hydroxycarboxylic acid dehydrogenases were all unsuccessful as methods of generating a probe that could be used to screen a lambda genomic DNA library of *A. calcoaceticus*.

Chapter 1. Introduction

1.1 Hierarchy of catabolic pathways and the evolution of metabolic function

There exists a hierarchy of catabolic pathways. The peripheral catabolic pathways convert the vast array of compounds upon which bacteria can grow into a more restricted range of compounds which act as substrates for the secondary catabolic pathways. The secondary catabolic pathways, examples of which can be ubiquitous such as glycolysis, or more specialised and less widespread as in the case of the ß-ketoadipate pathway for the degradation of catechol and protocatechuate, in turn feed a small number of key intermediates into the central amphibolic pathways which are at the heart of metabolism. For example, by this method probably hundreds of aromatic compounds are converted into only about a dozen key substrates for oxygenative ring cleavage (Fewson, 1981; 1991). Experiments aimed at determining evolutionary relationships among catabolic enzymes tend to have concentrated on the secondary and central catabolic pathways, with workers shying away from the peripheral pathways because of the great effort that would be required to characterise such a plethora of pathways and their enzymes. However, the peripheral catabolic pathways may well provide evolutionary evidence that cannot be obtained from the other pathways, because under many growth conditions these pathways are not used by the organism and so are free to undergo more rapid evolutionary changes than the central and secondary pathways which are vital and hence more constrained against changes. Also, many peripheral catabolic pathways have important actual or potential industrial applications, for example, in the degradation of environmental pollutants and in the industrial biotransformation of commercially important materials (Fewson, 1988; Hummel & Kula, 1989; Nikolova & Ward, 1994a,b; Shin & Rogers, 1995). The optimisation and manipulation of peripheral pathways and their enzymes for use as biotransformation systems in industry would require the knowledge of how such pathways arose and how they operate in Nature.

1.2 Aromatic compounds in the environment

Peripheral catabolic pathways enable microorganisms to degrade, and consequently use as sources of carbon and energy, some of the enormous range of aromatic compounds that exist in the environment. Next to glucosyl residues, the benzene ring is the most widely distributed unit of chemical structure in Nature (Dagley, 1981). For example, lignin is a polymer of benzenoid units. Besides the large number of naturally occurring aromatic compounds, there is an increasing

number of aromatic compounds being produced industrially and then released into the environment. Some of these compounds, such as benzene, toluene diisocyanate and various pesticides are serious pollutants which when present in high enough amounts can cause irritant, fibrotic, asphyxiant or allergic responses (Lippman & Schlesinger, 1979). In addition the production of phenol from cumene and the synthesis of styrene cause pollution in the form of odours, and the release of catechol can result in highly coloured waste waters which are unsightly (Fewson, 1981). Xenobiotic compounds are completely synthetic chemical compounds that are not naturally occurring in the biosphere. Removal of these, and of industrially produced natural compounds, from the environment depends upon their ability to be accepted as substrates for microbial enzymes, and since many such compounds are aromatics the metabolism of aromatic compounds by microorganisms is of great interest.

1.3 Aromatic catabolism in microorganisms

The metabolism of many of the aromatic compounds that exist in the environment has been studied in a variety of microorganisms. A great deal of such work has been done with the filamentous fungi and yeasts such as *Aspergillus* spp., *Penicillium* spp., *Neurospora* spp., and *Rhodotorula* spp. (Cain, 1980; Durham, 1984), and gram-positive bacteria such as *Bacillus* spp. (Crawford & Olsen, 1979), *Nocardia* spp. (Rann & Cain, 1973) and *Rhodococcus* spp. (Rast *et al.*, 1980; Warhurst *et al.*, 1994). Possibly the most extensively studied microorganisms as regards their ability to metabolise aromatics are the Gram-negative soil bacteria *Pseudomonas* spp. and *Acinetobacter calcoaceticus*, which are able to grow on an extensive range of aromatic compounds as sole carbon and energy sources (Stanier *et al.*, 1966; Fewson, 1967; Baumann *et al.*, 1968).

The large number of compounds upon which bacteria can grow are converted into a much smaller range of intermediates by converging peripheral catabolic pathways (section 1.1). When aerobic bacteria degrade any of the multitude of aromatic compounds in the environment, they are converted into a small number of dihydric phenols and these key intermediates are then subjected to intra-diol (*ortho*) or extra-diol (*meta*) ring cleavage (Dagley, 1978; Fewson, 1981). For example, in *Acinetobacter* spp., only seven such compounds have been identified, and of these only catechol and protocatechuate (3,4-dihydroxybenzoate) are the point of convergence of a wide range of peripheral catabolic pathways (Fewson, 1991; Figs. 1.1 & 1.2). Ring cleavage in most of these cases is brought about by *ortho* fission of catechol by catechol 1,2-

Figure 1.1 The conversion of aromatic compounds into catachol by various strains of *Acinetobacter*. Based on Fewson (1991).

ethyl mandelate · methyl mandelate · L-tryptophan · D-mandelate · L-mandelate · 2-hydroxymandelate · L-formylkynurenine · phenylglyoxylate · 2-hydroxyphenylglyoxylate · kynurenate · benzyl alcohol · 2-hydroxybenzyl alcohol · benzaldehyde · L-kynurenine · 2-hydroxybenzaldehyde · acetylsalicylate · 4-hydroxyquinazoline · aniline · benzoate · 2-hydroxybenzoate · guaiacol glycerol ether · anthranilate · phenol · guaiacol · benzene · catechol

Figure 1.2 The conversion of aromatic compounds into protocatechuate by various strains of *Acinetobacter*. Based on Fewson (1991).

3-methoxy-4-hydroxymandelate · 4-hydroxymandelate · 3-methoxy-4-hydroxybenzyl alcohol · 3-methoxy-4-hydroxyphenylglyoxylate · 4-hydroxyphenylglyoxylate · 3,4-dihydroxymandelate · 4-hydroxybenzyl alcohol · 3-phenylpropionate · 3-methoxy-4-hydroxybenzaldehyde · cinnamate · 4-hydroxybenzaldehyde · 3,4-dihydroxyphenylglyoxylate · 4-hydroxycinnamate · 3-methoxy-4-hydroxybenzoate · quinate · 3,4-dihydroxybenzaldehyde · 4-hydroxybenzoate · 5-dehydroquinate · shikimate · 5-dehydroshikimate · protocatechuate

dioxygenase and of protocatechuate by protocatechuate 3,4-dioxygenase (Baumann *et al.*, 1968; Stanier & Ornston, 1973). The ß-ketoadipate pathway (Fig. 1.3) converts the products of *ortho*-cleavage into succinate and acetyl-CoA which are then fed into the Krebs cycle (Stanier & Ornston, 1973). In *Pseudomonas* spp., unsubstituted or haloaromatic compounds tend to be metabolised by *ortho*-cleavage pathways whereas alkylaromatics are not since many *ortho*-cleavage enzymes have low affinities for alkylcatechols (Rojo *et al.*, 1987). Some microorganisms have modified *ortho*-cleavage pathways in addition to the convential β-ketoadipate pathway and these have parallel enzymes with altered substrate specificities. Examples of such pathways enable the growth of *P. putida* on 3-chlorobenzoate (Coco *et al.*, 1993), *Pseudomonas* sp. strain P51 on 1,2,4-trichlorobenzene (van der Meer *et al.*, 1991) and *Alcaligenes eutrophus* on 2,4-dichlorophenoxyacetic acid (Don *et al.*, 1985).

Examples of *meta*-cleavage of peripheral pathway products are not as widely reported as examples of *ortho*-cleavage. *meta*-Cleavage pathways produce products such as pyruvate, acetaldehyde and formate which can then be fed into the central metabolic pathways (Stanier & Ornston, 1973; Dagley, 1978). However, despite the propensity of cases in which *ortho*-fission has been reported in strains of *Acinetobacters* Baumann *et al.* (1968) reported both the *ortho* and *meta* fission of catechol after growth of *Acinetobacter* strain 20 in their study on phenol and Sze & Dagley (1987) reported the *meta* cleavage of protocatechuate by protocatechuate 4,5-dioxygenase in a strain of *A. lwoffii* after growth on 4-hydroxymandelate. This latter study is the only reported case of *meta*-fission after growth on mandelic acid in any microorganism. Fig. 1.4 shows an example of a *meta*-cleavge pathway: that of the TOL-plasmids of *Pseudomonas* spp. (section 1.5). The *meta*-cleavage enzymes of these plasmids are able to tolerate alkyl substitued catechols and so enable their hosts to utilise a wider range of compounds for growth (Assinder & Williams, 1990). *meta*-Cleavage of catechols is usually brought about by catechol 2,3-dioxygenases. *Rhodococcus rhodochrous* strain CTM has two such enzymes, catechol 2,3-dioxygenase I and catechol 2,3-dioxygenase II, and it has been shown that these two enzymes can utilise the same substrates, with catechol 2,3-dioxygenase I showing a greater preference for substituted catechols except that catechol 2,3-dioxygenase II can also cleave 2,3-dihydroxybiphenyl and catechols derived from phenylcarboxylicacids (Schreiner *et al.*, 1991). *meta*-Cleavage of catechol by *Rhodococcus rhodochrous* has also been reported after growth on styrene (Warhurst *et al.*, 1994).

The genetic elements responsible for the catabolism of aromatic compounds are quite often harboured on plasmids, notably in pseudomonads. Some strains carry TOL-plasmids which convey the ability to grow on a variety of

6

Figure 1.3 β-Ketoadipate or *ortho* pathway.

Designations for structural genes associated with the ß-ketoadipate pathway in *Acinetobacter*. The protocatechuate 3,4-dioxygenase gene is indicated by *pcaG* and *pcaH* which respectively encode its nonidentical α and β subunits. Designations *catI* and *catJ* or *pcaI* and *pcaJ* represent the α and β subunits respectively of β-ketoadipate succinyl CoA transferase. Gene designations correspond to enzymes as follows: *pobA*, *p*-hydroxybenzoate hydroxylase; *pcaG* and *pcaH*, protocatechuate 3,4-dioxygenase; *pcaB*, ß-carboxy-*cis, cis* -muconate cycloisomerase; *pcaC*, γ-carboxymuconolactone decarboxylase; *pcaD* or *catD*, ß-ketoadipate enol-lactone hydrolase; *pcaI* and *pcaJ* or *catI* and *catJ*, ß-ketoadipate:succinyl-CoA transferase; *pca F* or *cat F*, ß-ketoadipyl CoA thiolase; *benABC*, reductive benzoate dioxygenase; *benD*, 1,2-dihydro-1,2-dihydroxybenzoate dehydrogenase; *cat A*, catechol 1,2-dioxygenase; *catB*, *cis,cis*-muconate cycloisomerase; *catC*, muconolactone isomerase.

Based on Ornston & Neidle (1991).

METABOLITES

GENES

GENES

benzoate

ben ABC

p- hydroxybenzoate

1,2-dihydro-1,2-dihydroxybenzoate

pob A

*ben*D

protocatechuate

catechol

pca HG O_2

O_2 *cat* A

β- carboxy-*cis, cis*-muconate

cis, cis-muconate

pca B

cat B

γ- carboxymuconolactone

muconolactone

pca C CO_2

cat C

β–ketoadipate enol-lactone

pca D

cat D

β- ketoadipate

pca IJ

cat IJ

β- ketoadipyl-CoA

pca F

cat F

succinyl-CoA + acetyl-CoA
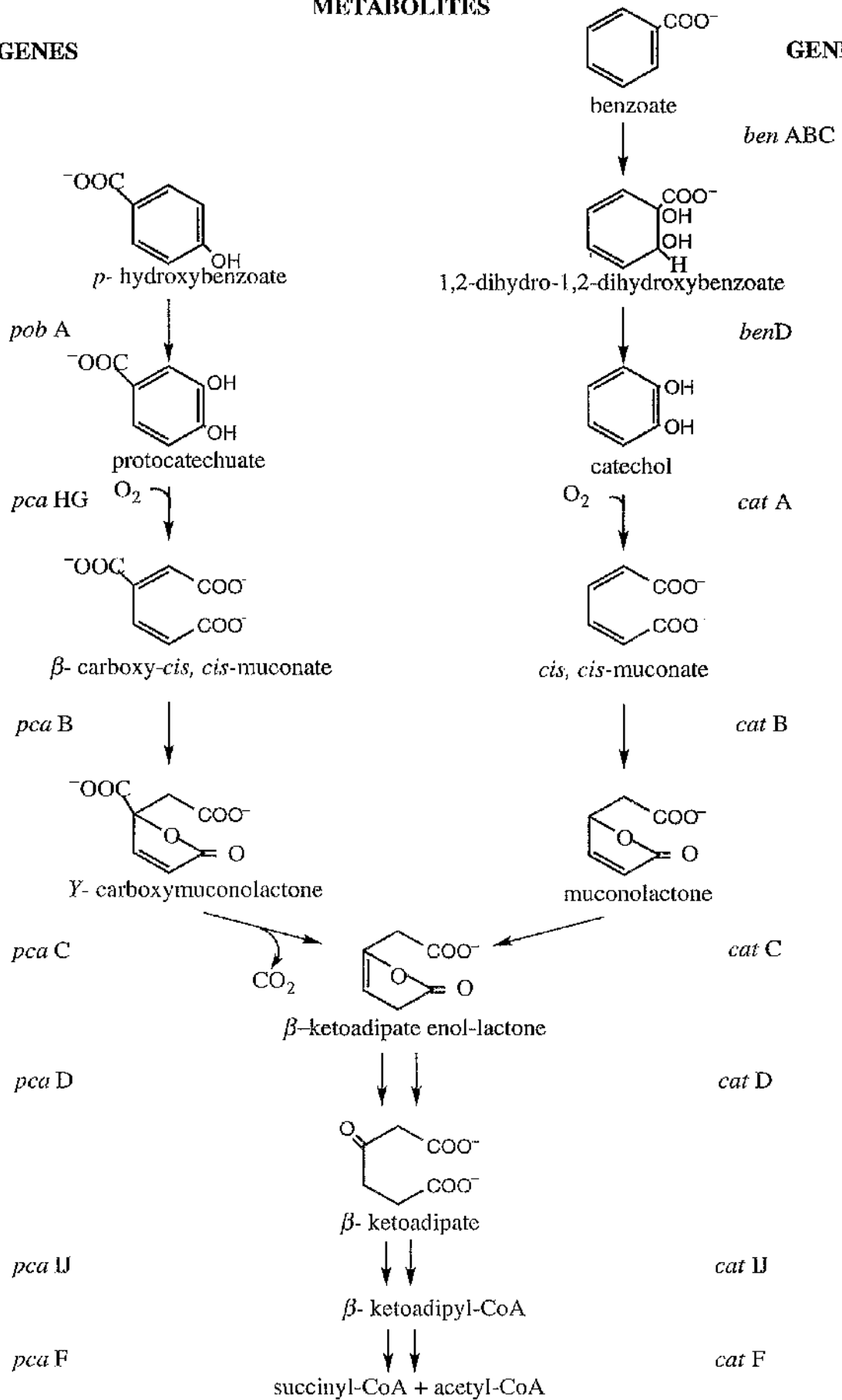
Figure 1.4 Complete pathway encoded by the TOL-plamids. The following hydrocarbons (1) serve as growth substrates: toluene ($R_1=R_2=H$); m -xylene ($R_1=CH_3$, $R_2=H$); p -xylene ($R_1=H$, $R_2=CH_3$); 3-ethyltoluene ($R_1=C_2H_5$, $R_2=H$); 1,2,4-trimethylbenzene (pseudocumene) ($R_1=R_2=CH_3$). Numbered compounds (toluene catabolism only) are: 2, benzyl alcohol; 3, benzaldehyde; 4, benzoate; 5, benzoate dihydrodiol (1,2-dihydrocyclohexa-3,5-diene carboxylate); 7, 2-hydroxymuconic semialdehyde; 8, 4-oxalocrotonate (enol); 9, 4-oxalocrotonate (keto); 10, 2-oxopentenoate (enol) or 2-hydroxypent-2,4-dienoate; 11, 4-hydroxy-2-oxovalerate. Enzyme abbreviations are: XO, xylene oxygenase; BADH, benzyl alcohol dehydrogenase; BZDH, benzaldehyde dehydrogenase; TO, toluate 1,2-dioxygenase; DHCDH, 1,2-dihydroxy-cyclohexa-3,5-diene carboxylate (benzoate dihydrodiol) dehydrogenase; C23O, catechol 2,3-oxygenase; HMSH, 2-hydroxymuconic-semialdehyde hydrolase; HMSD, 2-hydroxymuconic-semialdehyde dehydrogenase; 4OI, 4-oxalocrotonate isomerase; 4OD, 4-oxalocrotonate decarboxylase; OEH, 2-oxo-4-pentenoate (or 2-hydroxy-2,4-dienoate) hydratase; HOA, 4-hydroxy-2-oxovalerate aldolase. The genes encoding the protein subunit(s) are as designated. Their expression is regulated in two distinct operons, the upper-pathway operon and the *meta* - (or lower-) pathway operon.

Based on Assinder & Williams (1990).

"Upper"-pathway operon

COO^-

R_1

R_2

TO
(*xylD* or *xylXYZ*)

O_2 — NADH

NAD$^+$

^-OOC OH H

OH

R_1

NAD$^+$ R_2

DHCDH
(*xylL*)

NADH

OH

OH

R_1

R_2

C23O
(*xylE*)

O_2

HMSD (*xylG*)

OH

COO^- NAD$^+$

O

R_1

R_2

NADH

OH

COO^-

COO^-

R_2

4OI (*xylH*)

HMSH
(*xylF*)

R_1COO^-

OH

COO^-

R_2

CO_2

O

COO^-

COO^-

R_2

OEH
(*xylJ*)

4OD (*xylI*)

O

COO^-

HO R_2

HOA
(*xylK*)

$CH_3.CO.COO^-$

+

$R_2.CH_2.CHO$
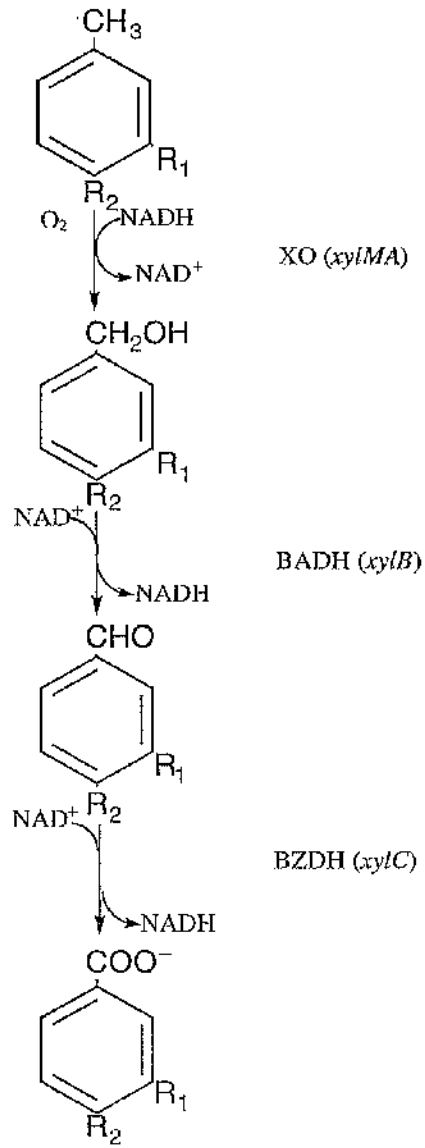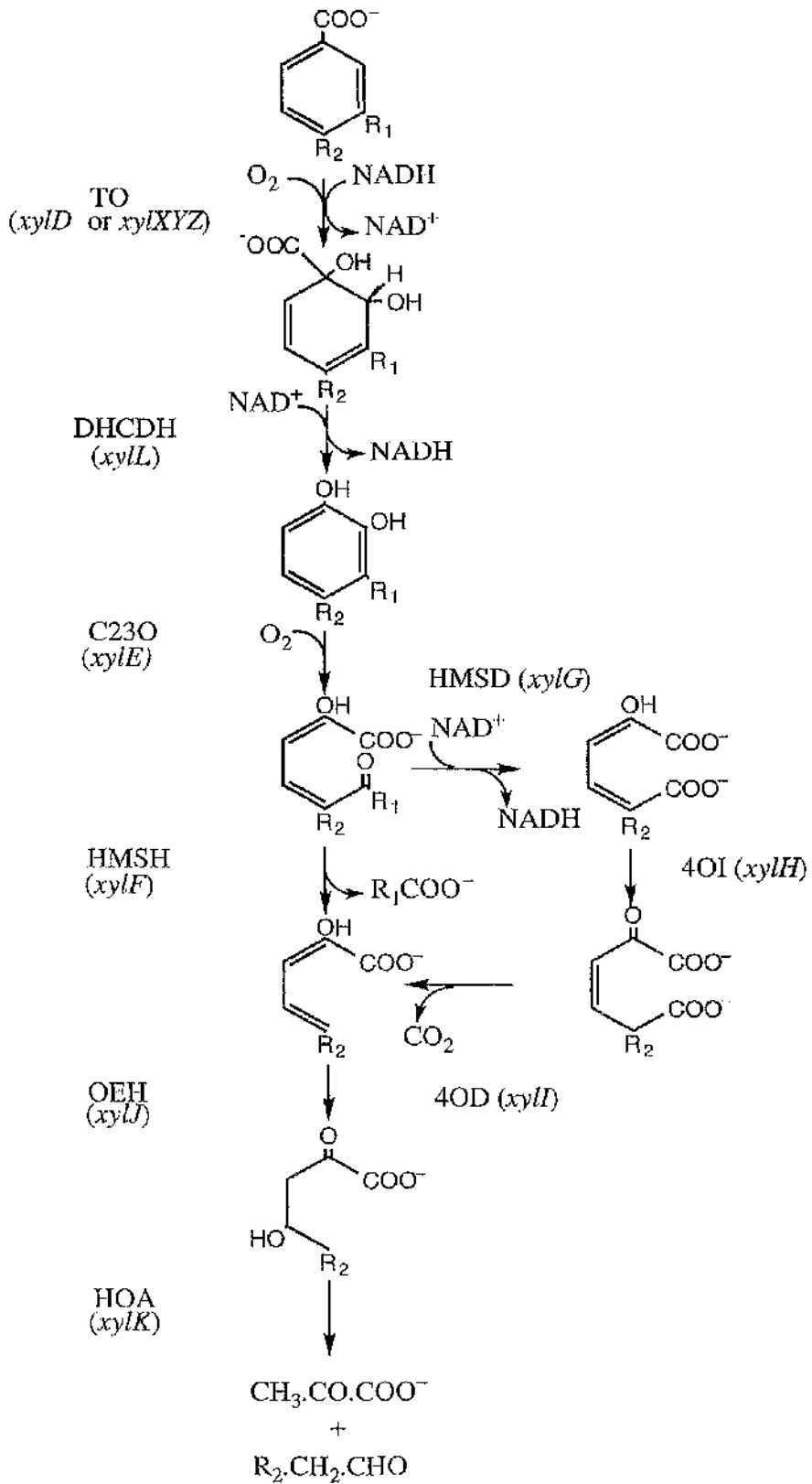
"Lower" (or *meta*)-pathway operon

alcohols, aldehydes and carboxylic acids (section 1.5; Assinder & Williams, 1990). Also, the ability of *P. putida* mt-2 to metabolise aromatic amines is conveyed by the plasmid pTDN1 (McClure & Venables, 1987) and the plasmid pJS1 encodes the enzymes for the dissimilation of 2,4-dinitrotoluene by *Pseudomonas* sp. strain DNT (Suen & Spain, 1993). In acinetobacters the mineralisation of 4-chlorobiphenyl has been shown to be plasmid-mediated (Shields *et al.*, 1985).

Only in relatively recent years have reports begun to emerge of biochemical pathways used by microorganisms for the anaerobic catabolism of aromatic compounds (Evans & Fuchs, 1988). Toluene and xylenes have been shown to be mineralized to benzoyl-CoA under anaerobic conditions by the denitrifying bacterium *Thauera* sp. strain K172 (Biegert & Fuchs, 1995). The initial step is postulated to be methylhydroxylation of toluene to benzyl alcohol with oxidation requiring dehydrogenase-type enzymes and water as oxygen source, as opposed to the aerobic enzymatic attack by oxygenases, which depends on molecular oxygen. Enzymes proposed to be involved in this pathway are an NAD^+-dependent benzyl alcohol dehydrogenase, an $NADP^+$-dependent benzaldehyde dehydrogenase, benzoate-CoA ligase (AMP-forming), and benzoyl-CoA reductase (dearomatizing). Other examples of anaerobic metabolism of aromatic compounds are the metabolism of 3,4,5-trihydroxybenzoate by *Eubacterium oxidoreducens*, in which formate or hydrogen is used as an electron donor for reductive attack on the aromatic ring (Haddock & Ferry, 1993), and the metabolism of 4-hydroxybenzoate by *Rhodopseudomonas palustris* in which a β-oxidation-like sequence of modification reactions leads to ring fission (Hutber & Ribbons, 1983).

1.4 The ß-ketoadipate or *ortho* pathway

The ß-ketoadipate or *ortho*-pathway, into which many products from aerobic peripheral catabolic pathways are fed, has been extensively studied in *Pseudomonas putida* and *Acinetobacter calcoaceticus*. In fact it was one of the first subjects of physiological investigation of enzyme induction in bacteria (Stanier, 1951). The genes encoding the enzymes in the pathway in both of these organisms have been cloned and sequenced and they provide useful information concerned with the processes of evolutionary divergence.

Figure 1.3 shows the two branches of the ß-ketoadipate pathway with the *pca* structural genes encoding the set of enzymes that convert protocatechuate to common metabolites, and the *cat* structural genes encoding analogous enzymes

which allow catechol to be utilized in a parallel metabolic sequence. The genes encoding enzymes for physiologically linked steps of the ß-ketoadipate pathway are linked in supraoperonic clusters in acinetobacters and pseudomonads (Ornston & Neidle, 1991). The enzymes encoded by *pca D, I, J* and *F* are identical in catalytic activity to those encoded by *cat D, I, J & F*. Within the *A. calcoaceticus* chromosome, the *pcaIJF* and *cat IJF* regions have experienced little evolutionary divergence from their presumably common ancestor. This is shown by the fact that their nucleotide sequences share over 98% identity (Kowalchuk *et al.*, 1994) and that these genes exhibit unusually high G+C content and differ from other *A. calcoaceticus* genes in their codon usage (Ornston & Neidle, 1991; Shanley *et al.*, 1994). These genes have probably managed to overcome the selective pressures to conform to the characteristic G+C content and codon usage of their host chromosome by their ability to exchange DNA sequence between themselves (Doten *et al.*, 1987). This is in contrast to *pca* D and *cat* D which have been shown to have diverged extensively (Hartnett & Ornston, 1994). The overall sequence identity between the two genes is 51%, but in regions where the selection for identical amino acid residues is not imposed, the identity of aligned segments is only 28.2%. In this case there has been selection for the avoidance of genetic exchange with *pca* D acquiring a complex DNA slippage structure in order to achieve the multiple mutations required for its extensive divergence from *cat* D (Hartnett & Ornston, 1994).

All of the enzymes in the ß-ketoadipate pathway of all the genera studied are inducible. However, the induction patterns of the enzymes differ between *Acinetobacter* and *Pseudomonas* spp. since gene rearrangements have resulted in divergent transcriptional controls (Ornston & Neidle, 1991). In *Acinetobacter*, the *catBC* sequence forms an operon with the downstream *catIJFD* region, whereas in *Pseudomonas* there is no direct counterpart to the *catIJFD* region with the physiological function of the region being fulfilled by the *pcaI, J, F* and *D* genes. In *Acinetobacter* all of the structural genes required for protocatechuate metabolism (Fig. 1.3) are located in the same (pca) operon, whereas in *P. putida* the gene organisation is fragmentary and in the case of the linked *pcaB* and *pcaD* genes the transcriptional order is reversed from that of *Acinetobacter*. Genetic analysis of transcriptional regulation is particularly useful when studying evolutionary divergence because the evolution of factors controlling the regulation of genes gives information concerning the selection pressures that dictated evolution in a particular organism. Isofunctional enzymes from these two species can all be traced to a common ancestor. The enzymes so far studied all share over 40% amino acid identity between the two genera. This is particularly remarkable because the GC content of the structural genes usually

differs by 20%. This difference in GC content is reflected in distinctive codon usage patterns (Ornston & Neidle, 1991).

1.5 The TOL-degradative pathway of *Pseudomonas putida*

In addition to the chromosomally encoded β-ketoadipate (*ortho*) pathway (section 1.4) some strains of *Pseudomonas* harbour plasmids containing genes required for *meta*-cleavage pathways and also genes required to degrade various aromatic compounds into metabolites for the *meta*-pathways.

The TOL-plasmids pWW0 of *Pseudomonas putida mt*- 2 and pWW53 of *P. putida* MT53 have been well characterised (Assinder & Williams, 1990; Williams & Murray, 1974). They are large extrachromosomal elements that convey the ability to grow on an array of about 20 substrates including hydrocarbons, alcohols, aldehydes and carboxylic acids such as toluene, (substituted-) xylenes and (substituted-) benzyl alcohols and benzaldehydes (Assinder & Williams, 1990). The 117 Kb pWW0 plasmid contains two catabolic operons. The "upper" operon encodes the enzymes that oxidize the hydrocarbons to their corresponding alcohol, aldehyde, and carboxylic acid derivatives (Fig. 1.4; Harayama *et al.*, 1986; 1989) and the "lower" operon encodes the *meta*-cleavage pathway enzymes which metabolise these compounds into Krebs cycle intermediates (Fig. 1.4; Harayama & Rekik, 1990). The upper pathway is composed of three enzymes and the operon order is promoter-*xylUWCMABN*, where *xylC* encodes benzaldehyde dehydrogenase, *xylM* and *xylA* encode the two different sub-units of xylene monoxygenase and *xylB* encodes benzyl alcohol dehydrogenase (Harayama *et al.*, 1989). The other three genes encode proteins which as yet have no known function: *XylN* codes for a 52 KDa protein which is processed to a 47 KDa protein (Harayama *et al.*, 1989), *xylU* encodes a small protein of 131 amino acid residues and *xylW* encodes a 348 amino acid protein that has a significant homology with the zinc-dependent long-chain alcohol dehydrogenases (section 3.4; Fig. 3.5). The *meta* operon contains the 13 genes involved in the *meta*-pathway in the order *xylXYZLTEGFJQKIH* (Harayama & Rekik, 1990). This operon spans over 10kb of DNA and is consequently one of the longest found in bacteria.

1.6 Mandelate and benzyl alcohol metabolism

The mandelate and benzyl alcohol pathways of *A. calcoaceticus* and *P. putida* are excellent examples of peripheral catabolic pathways (section 1.1) and have been well characterised (Fig. 1.5). In *A. calcoaceticus* the converging mandelate and benzyl alcohol pathways are chromosomally encoded and feed metabolites into the β-ketoadipate pathway (section 1.4). In *P. putida* the mandelate pathway genes are also chromosomally encoded and feed into the β-ketoadipate pathway whereas the enzymes for the degradation of benzyl alcohol are encoded by TOL-plasmids and feed metabolites into a *meta*-cleavage pathway (section 1.5). The ability to metabolise mandelate is quite wide-spread amongst a varied selection of Gram-negative and Gram-positive bacteria, as well as in filamentous fungi and yeast, and mandelate is degraded by at least as many different pathways as is glucose (Fewson, 1988). These pathways often enable growth on a wide variety of carbon sources because the enzymes involved have relaxed substrate specificities. For example, substitution on the 3-, 4-, and 5-positions of the benzene ring is tolerated and there is metabolism, or partial metabolism, of mandelate analogues (Fewson, 1988).

There are two enantiomers of mandelate. Each enantiomer is initially oxidised by a stereospecific mandelate dehydrogenase to phenylglyoxylate. Strains of *A. calcoaceticus*, *Aspergillus niger*, *Neurospora crassa* and *Rhodotorula graminis* that are able to grow on both enantiomers of mandelate have two stereospecific mandelate dehydrogenases (Fewson, 1988). Some strains of *Pseudomonas putida*, however, have been shown to possess a mandelate racemase which converts the D-enantiomer of mandelate into L-mandelate for oxidation by an L-mandelate dehydrogenase (Ransom *et al.*, 1988). An exception to the general pattern of mandelate metabolism in bacteria is that of *Pseudomonas convexa* (probably a strain of *P. putida*). This strain can grow on both mandelate enantiomers using a racemase but by following this with ring hydroxylation to 4-hydroxymandelate. 4-Hydroxymandelate is then oxidatively decarboxylated to 4-hydroxybenzaldehyde and feeds into the β-Ketoadipate pathway via protocatechuate, rather than catechol, which is the case in other bacteria (Fewson, 1988).

A. calcoaceticus strain NCIB8250 can only grow on the L-mandelate enantiomer, and strain EBF65/65 can only grow on the D-mandelate enantiomer and each of these strains possesses only the corresponding mandelate dehydrogenase. However, these strains give rise to mutants that can grow on the second enantiomer, and in every case this is due to the appearance of an extra dehydrogenase, specific for the other enantiomer (Hills & Fewson, 1983). The

Figure 1.5.

The mandelate and benzyl alcohol pathways of *Acinetobacter calcoaceticus*.

Reactions are catalysed by: 1, D-mandelate dehydrogenase; 2, L-mandelate dehydrogenase; 3, phenylglyoxylate decarboxylase; 4, benzyl alcohol dehydrogenase; 5, benzaldehyde dehydrogenase I; 6 benzaldehyde dehydrogenase II.

Adapted from MacKintosh & Fewson (1988a).

D-mandelate

L-mandelate

phenylglyoxylate

benzyl alcohol

benzaldehyde
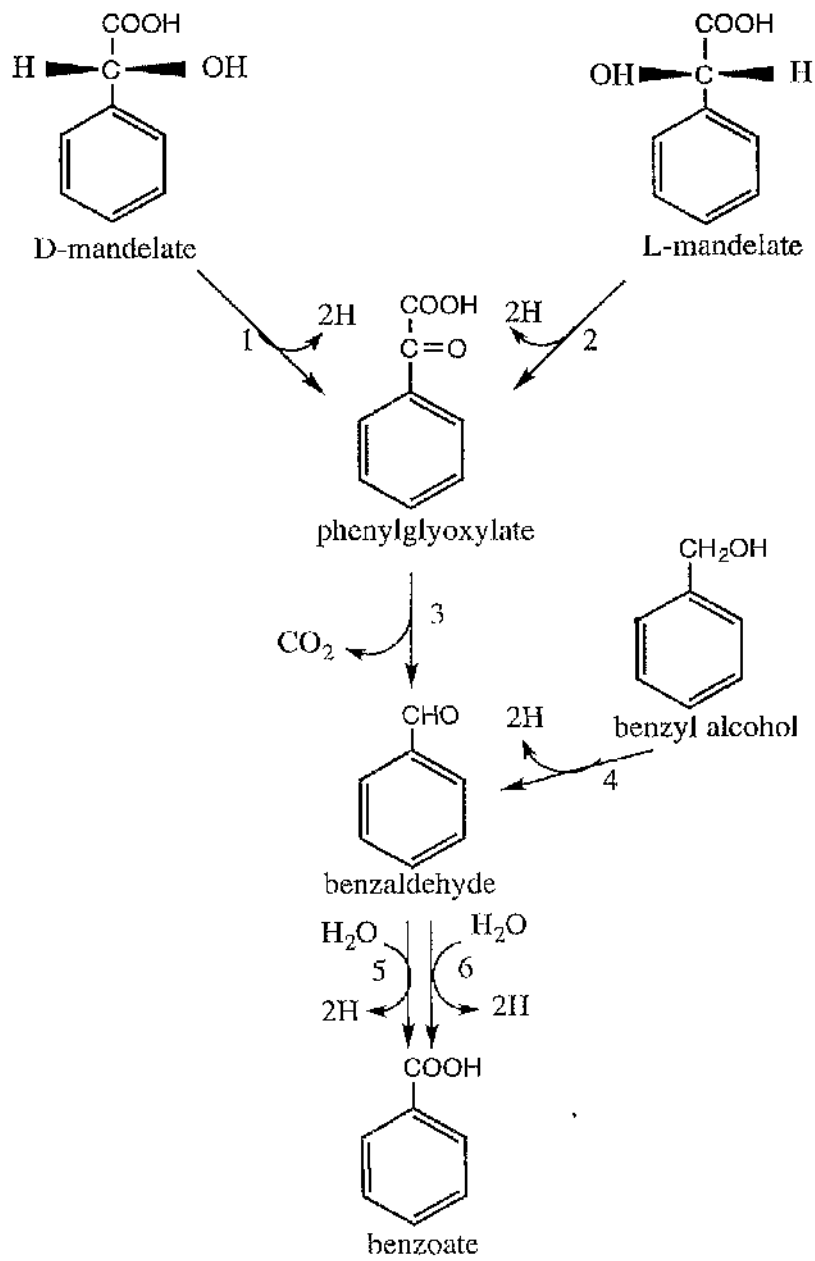
benzoate

appearance of these additional enzymes may be due to the expression of a previously cryptic or silent gene.

The mandelate and benzyl alcohol pathways of *A. calcoaceticus* converge at the point of benzaldehyde oxidation to benzoate. *A. calcoaceticus* has two benzaldehyde dehydrogenases (Livingstone *et al.*, 1972). Benzaldehyde dehydrogenase I (BZDH I) is associated with mandelate metabolism, whereas benzaldehyde dehydrogenase II (BZDH II) is associated with benzyl alcohol metabolism. These two enzymes differ in their heat stability and K^+-dependence. BZDH I is heat-stable, whereas BZDH II is heat-labile, and BZDH I is a K^+-activated enzyme, whereas BZDH II is K^+-independent. Both enzymes have been purified and shown to be NAD-dependent, soluble and tetrameric (MacKintosh & Fewson, 1987, 1988a,b; Chalmers & Fewson, 1989). The benzaldehyde dehydrogenases involved in benzyl alcohol metabolism in *P. putida* strains are encoded by TOL-plasmids and are distinct from two different chromosomally encoded NAD^+- and $NADP^+$-dependent benzaldehyde dehydrogenases implicated in the metabolism of mandelate (Inoue *et al.*, 1995; Tsou *et al.*, 1990).

1.7 Regulation of the mandelate and benzyl alcohol pathways

The mandelate pathway enzymes appear to be inducible in all of the organisms studied (Fewson, 1988). This is a common feature of peripheral pathways in that they are switched on or off as required. In the organisms in which induction of the mandelate enzymes has been examined the induction patterns appear to be different. In *P. putida* the mandelate enzymes are coordinately induced by D- or L-mandelate, or by phenylglyoxylate (Hegeman, 1966a, b). The L-mandelate dehydrogenase of *P. aeruginosa* is induced by L-mandelate, independently of the other enzymes. In this species phenylglyoxylate decarboxylase is induced by phenylglyoxylate and benzaldehyde dehydrogenase is induced by phenylglyoxylate or β-ketoadipate (Rosenberg, 1970). The mandelate enzymes of *A. calcoaceticus* are co-ordinately expressed and are induced by phenylglyoxylate, but not by mandelate (Livingstone & Fewson, 1972; Fewson *et al.*, 1978). Induction by product, rather than by substrate, is not uncommon in bacterial catabolism. In instances when compounds that can be formed endogenously by a microorganism are used to support growth (e.g. shikimate, tryptophan, histidine) this can be explained as a mechanism by which the organism prevents the catabolism of the compounds in question when they are formed as biosynthetic intermediates. In the case of mandelate metabolism, however, there is no competition between degradation and biosynthesis and so the

phenomenom of product induction may be an evolutionary relic, resulting from the original functions of the structural or regulatory genes now involved in mandelate metabolism. In such cases inducer may be formed by noninduced levels of enzymes being sufficient to produce small amounts of inducer, or by nonspecific conversion due to substrate ambiguity of unrelated enzymes.

In *A. calcoaceticus* benzyl alcohol dehydrogenase and benzaldehyde dehydrogenase II are also co-ordinately expressed but are induced by benzyl alcohol (Livingstone *et al.*, 1972). Since benzaldehyde dehydrogenase I is induced only in the mandelate pathway, the expression of the two isofunctional benzaldehyde dehydrogenases is quite separate. Interestingly, mandelate metabolism dominates benzyl alcohol metabolism, with the two enzymes specific for benzyl alcohol metabolism unable to be induced in the presence of phenylglyoxylate, although benzyl alcohol supports a faster growth rate and gives a higher molar growth yield than mandelate (Beggs *et al.*, 1976; Beggs & Fewson, 1977; Fewson, 1985). The induction of the benzyl alcohol pathway genes of the TOL-plasmids of *P. putida* has been studied at the molecular level. The substrates of the upper pathway, toluene/xylenes and (methyl)benzyl alcohol, induce transcription of both the "upper" and the "lower" operons, whereas substrates of the "lower" operon, benzoate/toluates, induce transcription of only the "lower" operon (Harayama & Timmis, 1989). Two regulatory proteins, XylR and XylS, interact with substrates and activate transcription from the operon promoters; Pu for the "upper" pathway and Pm for the lower pathway (Harayama & Timmis, 1989; Marqués *et al.*, 1994).

The mandelate enzymes of *A. calcoaceticus* are subject to some catabolite repression (e.g. by succinate) and to feedback repression (e.g. by benzoate) but the mechanisms are not known (Cook *et al.*, 1975; Fewson, 1988). There is no evidence for any feedback inhibition of the enzymes so it seems that regulation is at the level of gene expression. Catabolite repression (e.g. by succinate and glucose) has been shown to regulate the genes involved in benzyl alcohol metabolism in *P. putida* (pWW0) with the targets of this expression being the promoters Pu and Ps, the latter being the *xylS* promoter (Holtel *et al.*, 1994; Duetz *et al.*, 1994). The enzymes converting benzoate into catechol and the enzymes of the ß-ketoadipate pathway in *A. calcoaceticus* are regulated independently of the mandelate and benzyl alcohol pathways (Stanier & Ornston, 1973) although there may be some pathway cross-talk. In *P. putida* (pWW0) the XylS regulatory protein is mainly responsible for the regulation of expression of the *meta*-cleavage pathway (Duetz *et al.*, 1994).

In *Thauera* sp., the enzymes for the anaerobic mineralization of toluene (section 1.3) appear to be induced by benzyl alcohol or benzaldehyde (Biegert & Fuchs, 1995).

1.8 The enzymes of the mandelate and benzyl alcohol pathways

The enzymes of the mandelate and benzyl alcohol pathways of *A. calcoaceticus* and *Pseudomonas putida* have been purified and characterised. In the case of the mandelate and the TOL-plasmid encoded benzyl alcohol pathways of *P. putida* the genes for these enzymes have also been cloned (Tsou *et al.*, 1990; section 1.6).

Mandelate racemase is found only in a few strains of *Pseudomonas* putida. The enzyme is a tetramer of four identical sub-units (M_r 38,570) with an absolute requirement for a divalent metal ion (e.g. Mg^{2+} or Mn^{2+}) for activity (Ransom *et al.*, 1988; Kenyon & Hegeman, 1979).

All of the known mandelate dehydrogenases are stereospecific and can be classified into two distinct classes. Group I mandelate dehydrogenases are membrane-associated flavoproteins and catalyse essentially irreversible reactions. The L- and D-mandelate dehydrogenses of *A. calcoaceticus* (Allison *et al.*, 1985a,b; Hoey *et al.*, 1987), and the mandelate dehydrogenases of *P. putida* and *P. aeruginosa* (Tsou *et al.*, 1990) are examples of this type. The L-mandelate dehydrogenase of *A. calcoaceticus* contains non-covalently bound FMN as cofactor and has a M_r of 44,000, which is very similar to the L-mandelate dehydrogenase of *P. putida* (M_r 43,352; Hoey *et al.*, 1987; Tsou *et al.*, 1990). There is 62% positional identity between the 50 *N*-terminal amino acids of these two enzymes and at some of the other positions there are conservative replacements (Fig. 1.6; Fewson *et al.*, 1993). These two L-mandelate dehydrogenases are very similar to the L-lactate dehydrogenases of *A. calcoaceticus* and *Escherichia coli* in terms of location, solubilisation by both ionic- and non-ionic detergents, pH optimum and pI and K_m values as well as M_r, cofactors and *N*-terminal amino acid sequences (Fig. 1.6; Table 1.1; Hoey *et al.*, 1987; Fewson *et al.*, 1993). The L-mandelate dehydrogenase of the yeast *Rhodotorula graminis* appears to be homologous with these enzymes but, like the L-lactate dehydrogenases of yeasts such as *Saccharomyces cerevisiae*, contains an additional haem domain (Fewson *et al.*, 1993). D-Mandelate dehydrogenase from *A. calcoaceticus* is slightly larger (M_r 60,000) and contains FAD as cofactor (Allison *et al.*, 1985b) but is similar to the D-lactate dehydrogenases of *A. calcoaceticus* and *E. coli* in terms of location, solubilisation by ionic- and non-

Figure 1.6

A. *N*-Teminal amino acid alignments of L(+)-lactate and L(+)-mandelate dehydrogenases. 1, L(+)-Lactate dehydrogenase from *Saccharomyces cerevisiae* ; 2, L(+)-lacate dehydrogenase from *Hansenula anomola* ; 3, Glycolate oxidase from spinach; 4, L(+)-mandelate dehydrogenase from *Pseudomonas putida* ; 5, L(+)-mandelate dehydrogenase from *Acinetobacter calcoaceticus* ; 6, L(+)-lactate dehydrogenase from *A. calcoaceticus*. The numbering of the yeast lactate dehydrogenases is based on the mature protein sequences after the removal of the pre-sequences (containing haem).

B. *N*-Terminal amino acid sequence alignments of D(-)-lactate and D(-)-mandelate dehydrogenases. 1, D(-)-Lactate dehydrogenase from *Escherichia coli* ; 2, D(-)-mandelate dehydrogenase from *A. calcoaceticus* ; 3, D(-)-lactate dehydrogenase from *Acinetobacter calcoaceticus* .

Taken from Fewson *et al.* (1993).

```
1.116 - L P P L D N I T N L Y D F E Y L A S Q T L T K Q A W A Y Y S S G A N D E V T H R E N H N A Y H R I F F K P K I L V D V R -175
2.108 - K P E L S Q M I N L H D F E T I A R Q I L P P P A L A Y Y C S A A D D E V T L R E N H N A Y H R I F F N P K I L I D V K -167
3.  1   M S Q N E I T N V N E Y E A I A K Q K L P K M V Y D Y Y A S G A B D Q W T L A E N R N A F S R I L F R P R I L I D V T -  56
4.  1   M S Q N L F N V E D Y R K L A Q K R L P K M V Y D Y L E G G A E D E Y G V K H N R D V F Q W R F K P K R L V D V S -  56
5.  1     P H K M I N V E D Y Q K L A K S T L P K V I D Y L E G G A D D E K G L R H N R Q V E D Q K M F K P -  50
6.  1   M I I S S A N D Y R E A A R R L P P L F H Y I D I G G A Y A E Y T L K R N V E D L X W I A L W Q -  49

Common            A       L       Y       A       N
```

```
1   10        20        30        40        50        60
1. M S S M T T T D N K A F L N E - L A R L V G - - S S H L T - D P A K T A R Y R K G F R S G Q G D A L A V V F P G S L L E L W R V
2. M Q T N P I I K E - L H Q L F N - - P Q Q I L I - E D V D X S P Y L K G F X V G K G D A L A V V I P Y X L F Q L X G V
3. M Q S N I D S Q H T I Q Q L I N I V G Q H V L A X D V P X X L Y
```

Table 1.1 Comparison of the NAD(P)-independent lactate and mandelate dehydrogenases from *Acinetobacter calcoaceticus*

	L-Mandelate dehydrogenase	L-Lactate dehydrogenase	D-Mandelate dehydrogenase	D-Lactate dehydrogenase
Location	All the enzymes are integral parts of the cytoplasmic membranes			
Solubilization by nonionic detergents (% activity released by Triton X-100 or Lubrol)	43-72	55-76	87-91	83-98
Solubilization by ionic detergents (% activity released by cholate or deoxycholate)	4-42	13-32	78-91	63-87
Subunit M_r	44,000	40,000	60,000	63,000
Cofactor (noncovalently bound)	FMN	FMN	FAD	FAD
pH Optimum	7.5	7.5	8.0	7.7
pI Value	4.2	4.0	5.5	5.8
K_m Value (μM) for each substrate (all the enzymes are stereospecific)	186	83	385 (in addition, D-lactate is a substrate, K_m 4.8 mM)	N. D.

Based on Fewson (1992).

ionic detergents, pH optimum, pI value, K_m value, M_r, cofactors and N-terminal amino acid sequences (Fig. 1.6; Table 1.1; Hoey *et al.*, 1987; Fewson *et al.*, 1993). The amino acid sequence similarity relationships between all these enzymes have been confirmed by immunological cross-reactions, although a very large amount of protein had to be used to detect even a weak cross-reaction between L-mandelate dehydrogenase from *P. putida* and antiserum raised against SDS-denatured L-mandelate dehydrogenase from *A. calcoaceticus* (Fewson *et al.*, 1993). This suggests that the level of sequence identity observed at the N-termini of these two enzymes is not maintained throughout the whole of the two proteins.

Group II mandelate dehydrogenases are soluble, NAD(P)-linked enzymes catalysing reversible reactions. Examples are the D-mandelate dehydrogenases of *Rhodotorula graminis* (Baker & Fewson, 1989) and *Lactobacillus curvatus* (Hummel *et al.*, 1988).

Phenylglyoxylate decarboxylase has been purified from *A. calcoaceticus* and is a TPP-dependent tetramer of identical sub-units (M_r 58,000; Barrowman & Fewson, 1985) The gene for this enzyme has been cloned from *Pseudomonas putida* and the enzyme (subunit M_r 53,621) has been overexpressed and purified (Tsou *et al.*, 1990).

A. calcoaceticus NCIB8250 has two benzaldehyde dehydrogenases (section 1.7), BZDH I being involved in mandelate metabolism and BZDH II involved in benzyl alcohol metabolism. This gives a good chance to test the hypothesis of retrograde evolution and ideas of gene duplication since these two enzymes, and also benzyl alcohol dehydrogenase, share two common substrates (NAD^+ and benzaldehyde) and are in a peripheral catabolic pathway that is less constrained to evolutionary change than a central pathway (section 1.1). N-Terminal amino acid sequencing and immunological cross-reactivity studies showed that BADH is not homologous to the benzaldehyde dehydrogenases (Chalmers *et al.*, 1991). This rules out the idea that retrograde evolution may have occured in this particular pathway. However, the two BZDH's appear to have many features in common and so it is possible that they emerged from some gene doubling event. Both enzymes are tetramers with similar sub-unit M_r values (56,000 and 55,000 for BZDH I and BZDH II respectively), similar sensitivity to thiol-blocking agents, pH optima, substrate specificity, esterase activity and K_m values for benzaldehyde (Chalmers & Fewson, 1989). These two benzaldehyde dehydrogenases have apparent K_m and V_{max} values for benzaldehyde of a similar order to the values obtained for TOL-BZDH (pWW53 and pWW0; Chalmers *et al.*, 1990; Inoue *et al.*, 1995). Benzaldehyde dehydrogenase II cross-reacted with antisera raised against BZDH I and, more strongly, TOL-benzaldehyde dehydrogenase (pWW53) cross-reacted with antisera raised against BZDH II

(Chalmers *et al.*, 1991). In this latter case the reciprocal cross-reaction was also detected. However, a difference between the two BZDH's is the fact that BZDH I is activated by K⁺ and is heat stable whereas BZDH II is K⁺ independent and heat-labile (Chalmers & Fewson, 1989). This would suggest that there are fundamental differences between the structures of these two enzymes. *N*-Terminal amino acid sequences showed there to be sequence identity between the TOL-BZDH (pWW53) and BZDH II (37%), but little amino acid identity between these two and BZDH I (Chalmers *et al.*, 1991). It is therefore surprising that TOL-BZDH (pWW53) also has a requirement for K⁺ (Chalmers *et al.*, 1990). The nucleotide sequences of the genes encoding these enzymes that have not been determined so far need to be determined before the true extent of their evolutionary relationships can be determined.

1.9 Benzyl alcohol dehydrogenases

Benzyl alcohol dehydrogenase (BADH) from the benzyl alcohol pathway of *A. calcoaceticus* NCIB8250 catalyses the reversible oxidation of benzyl alcohol to benzaldehyde with the concomitant reduction of NAD⁺ to NADH (Fig. 1.5). Purification and preliminary characterisation of this enzyme has been carried out (MacKintosh & Fewson, 1988a, b). BADH has a relaxed substrate specificity which enables *A. calcoaceticus* to grow on a much wider range of compounds than would otherwise be possible. Table 1.2 shows the range of substrates oxidised by BADH and the rate of oxidation relative to that of benzyl alcohol and that BADH can also oxidise 2-, 3-, and 4-methylbenzyl alcohols, although the *meta* and *para* substitutions are greatly preferred (Chalmers *et al.*, 1990). It has been shown that *A. calcoaceticus* NCIB8250 can grow on benzyl alcohol, 2-hydroxybenzyl alcohol (salicyl alcohol), 4-hydroxybenzyl alcohol and 4-hydroxy-3-methoxybenzyl alcohol (vanillyl alcohol; Fewson, 1967). Despite being oxidised through BADH and BZDH II, 3-hydroxybenzyl alcohol, 3-methoxybenzyl alcohol, and 3- and 4-methylbenzyl alcohols are not utilised by *A. calcoaceticus* as growth substrates (Fewson, 1967). Oxygen-utilization patterns indicate that these compounds are not metabolized past the corresponding acid (Kennedy & Fewson, 1968). This is because the specificity of the ring-cleavage enzymes dictates which acids will be metabolized further. This restricted specificity of the ring-cleavage enzymes is probably also responsible for the inability of *A. calcoaceticus* to utilise cinnamyl and coniferyl alcohols as growth substrates (Fewson, 1967) despite the ability of BADH to oxidize these alcohols (MacKintosh & Fewson, 1988b). Cinnamyl alcohol and

Table 1.2
Substrate specificity of benzyl alcohol dehydrogenase from _A. calcoaceticus_

A. Relative activities of alcohol substrates

Substrates	Activity relative to the rate with benzyl alcohol (%)
Benzyl alcohol	100
Cinnamyl alcohol	64
Coniferyl alcohol	28
3,4-Dimethoxybenzyl alcohol	13
2-Furanmethanol	25
2-Hydroxybenzyl alcohol	40
3-Hydroxybenzyl alcohol	26
4-Hydroxybenzyl alcohol	67
4-Hydroxy-3-methoxybenzyl alcohol	39
4-Isopropylbenzyl alcohol	80
2-Methoxybenzyl alcohol	14
3-Methoxybenzyl alcohol	67
4-Methoxybenzyl alcohol	103
Pentafluorobenzyl alcohol	22
Perrilyl alcohol	80
2-Thiophenemethanol	50

Benzyl alcohol dehydrogenase from _A. calcoaceticus_ has also been shown to oxidise 2-, 3-, and 4-methylbenzyl alcohol.

B. Alcohols which are not oxidised

Aromatic alcohols:

2-Bromobenzyl alcohol, DL-1-phenylethanol, 2-phenylethanol

Other alcohols:

Allyl alcohol, butan-1-ol, cis -1, 2-cyclohexanediol, trans-1, 2-cyclohexanediol, cyclohexanol, decan-1-ol, ethanol, heptan-1-ol, hexan-1-ol, hexahydrobenzyl alcohol, inositol, mannitol, methanol, octan-1-ol, pentan-2-ol, propan-1-ol, propan-2-ol, 3-pyridinemethanol, 4-pyridinemethanol, sorbitol

Information taken from MacKintosh (1987); Chalmers _et al._(1990)

coniferyl alcohol (appendix VI and V respectively) are both intermediates of lignin biosynthesis and degradation and it is probably due to the degradation of lignin by fungi that acinetobacters encounter these compounds, and aromatic alcohols in general, in the environment (Cain, 1980). It is possible that the oxidation of these two alcohols through BADH and BZDH II could produce energy for *A. calcoaceticus* and the acids produced would then be available for metabolism by other micro-organisms. Table 1.2 shows that BADH is in general specific for aromatic alcohols, although the cyclohex-1-ene compound perillyl alcohol (appendix VII) is a good substrate. The benzene ring can be replaced by a thiophene or furan ring, but apparently not by a pyridine ring. The enzyme seems to have a preference for aromatic alcohols with small substituent groups in the *para* rather than the *ortho* position, i.e. away from the reactive carbinol group. This would support the idea that the active site may be a cleft structure, substituents in the *para* position on the aromatic ring being least sterically hindered for binding (MacKintosh & Fewson, 1988b). Thus the alcohol specificity of *A. calcoaceticus* BADH is similar to that of the bacterial aromatic alcohol dehydrogenases benzyl alcohol dehydrogenase and perillyl alcohol dehydrogenase of *Pseudomonas* spp. (Suhara *et al.*, 1969; Ballal *et al.*, 1966) and in particular the benzyl alcohol dehydrogenases of the TOL-plasmids pWW0 and pWW53 (Chalmers *et al.*, 1990; Shaw *et al.*, 1992) and benzyl alcohol dehydrogenase from *Thauera* sp. (Biegert *et al.*, 1995).

Benzyl alcohol dehydrogenase from *A. calcoaceticus* has been shown to be sensitive to the thiol-blocking agents iodoacetate, iodoacetamide, 4-chloromercuribenzoate and *N*-ethylmaleimide (MacKintosh & Fewson, 1988b). This inhibition could be protected against by the presence of substrate or cofactor, with NAD$^+$ being less protective than benzyl alcohol (MacKintosh, 1987). These results suggest that this inhibition is active-site directed. In addition to this, BADH encoded by the TOL-plasmid pWW0 has been shown to be sensitive to 4-chloromercuribenzoate, *N*-ethylmaleimide and to a lesser degree by iodoacetamide (Shaw & Harayama, 1990) and benzyl alcohol dehydrogenase from *Thauera* sp. has also been shown to be inhibited by such thiol-blocking agents (Biegert *et al.*, 1995). This inhibition by thiol-blocking reagents is a common feature of all the bacterial aromatic alcohol and aldehyde dehydrogenases and mammalian and yeast alcohol and aldehyde dehydrogenases that have been examined. Horse liver alcohol dehydrogenase, which also has benzyl alcohol dehydrogenase activity, is sensitive to iodoacetate (Sund & Theorell, 1963) and BADH purified from *P. putida* T-2 is sensitive to iodoacetate but is much more sensitive to 4-chloromercuribenzoate (Suhara *et al.*, 1969). Several bacterial aromatic alcohol and aromatic aldehyde dehydrogenases studied

are particularly sensitive to 4-chloromercuribenzoate (Ballal *et al.*, 1967, Suhara *et al.*, 1969; Kiyohara *et al.*, 1981; Jaeger *et al.*, 1981) and since 4-chloromercuribenzoate contains a benznene ring structure its potency could be an active-site-directed phenomenon.

The subunit molecular weight of benzyl alcohol dehydrogenase from *A. calcoaceticus* as calculated by SDS-PAGE was found to be 39,700 and the native molecular weight was found to be 155,000 as judged by gel filtration, indicating that the enzyme is a tetramer (MacKintosh & Fewson, 1988a). The latter was confirmed by cross-linking with dimethylsuberimidate (Chalmers & Fewson, 1989). BADH from TOL-plasmid pWW53 was shown to have a subunit M_r similar to that of *A. calcoaceticus* BADH (43,000), although in this case cross-linking studies indicated that the enzyme was a tetramer and gel filtration indicated that the enzyme was a dimer (Chalmers *et al.*, 1990). The gene encoding BADH (*xyl* B; section 1.6) from the TOL-plamid pWW0 has been cloned and sequenced (Shaw *et al.*, 1993). This particular benzyl alcohol dehydrogenase was shown to have a subunit M_r of 38,510 and to be a dimer binding two zinc atoms per sub-unit. TOL-benzyl alcohol dehydrogenase had 31% sequence identity with horse liver alcohol dehydrogenase and is a member of the long chain zinc-dependent group of alcohol dehydrogenases (section 1.10.1). Benzyl alcohol dehydrogenase purified from *Thauera* sp. has been shown to be a homotetramer with a subunit Mr of 40,000 and *N*-terminal amino acid sequence alignment indicated that this enzyme also belongs to the long-chain zinc-dependent alcohol dehydrogenase family (Biegert *et al.*, 1995).

MacKintosh & Fewson (1988b) showed BADH from *A. calcoaceticus* to be unaffected by the metal-ion chelators 10 mM EDTA, 1 mM 2, 2´-bipyridyl, 1 mM pyrazole and 1 mM 2-phenanthroline when purified and in crude extracts. This is unlike the alcohol dehydrogenaes of horse liver and *Saccharomyces cerevisiae* which have zinc-bound at the active-site and are sensitive to chelating agents (Sund & Theorell, 1963) but is similar to some other aromatic alcohol dehydrogenases (Suhara *et al.*, 1969; Jaeger *et al.*, 1981) and the benzyl alcohol dehydrogenase of *Thauera* sp. (Biegert *et al.*, 1995). Benzyl alcohol dehydrogenase encoded by the TOL-plasmid pWW0 was also not inhibited by 5 mM EDTA or 2,6-pyridine dicarboxylic acid, but could be inhibited by dialysis against 100 mM acetate buffer containing 5 mM benzyl alcohol and 5 mM *o*-phenanthroline (Shaw *et al.*, 1993).

Immunological cross-reactivity studies and *N*-terminal amino acid sequencing has been used to investigate the possible relationships between *A. calcoaceticus* BADH, TOL-BADH and horse liver alcohol dehydrogenase, the archetypal long chain zinc-dependent alcohol dehydrogenase (Chalmers *et al.*,

1991). Immunoprecipitation assays showed there to be a cross-reaction between *A. calcoaceticus* BADH and anti-(TOL-BADH) serum. Immunoblotting also gave cross-reactions between *A. calcoaceticus* BADH and antisera raised against TOL-BADH. The reciprocal cross-reactions were also observed. However, neither of the antisera raised against these enzymes cross-reacted with the alcohol dehydrogenases from horse liver or *Thermoanaerobium brockii*, this latter enzyme being another member of the family of long-chain zinc-dependent alcohol dehydrogenases (Peretz & Burstein, 1989). The *N*-terminal amino acid sequences of *A. calcoaceticus* BADH and TOL-BADH were shown to each have 36% sequence identity with each other and 26% sequence identity with the horse liver enzyme. *N*-Terminal amino acid alignments also showed there to be shared sequence identity between the two BADH enzymes and other members of the long-chain Zn^{2+}-dependent family of alcohol dehydrogenases. However, *A. calcoaceticus* BADH and TOL-BADH had only 8% and 16% *N*-terminal sequence identity respectively with the alcohol dehydrogenase from *Thermoanaerobium brockii*. This may explain why there were no immunological cross-reactions between the BADH enzymes and these other two alcohol dehydrogenases since it has been suggested that for cross-reactions to occur between denatured antigens there needs to be greater than about 40% sequence identity (Cassan *et al.*, 1986).

1.10.1 Alcohol dehydrogenases

Alcohol dehydrogenases such as benzyl alcohol dehydrogenase (section 1.9) enable the feeding of alcohols into central metabolism as carbon and energy sources for growth and utilise the oxidative reaction, resulting in the production aldehydes or ketones and reduced cofactor (e.g. NADH). In the case of some alcohol dehydrogenases this type of reaction enables the degradation of naturally occurring and xenobiotic aromatic compounds, some of which are pollutants (section 1.2). The reductive reaction of alcohol dehydrogenases produces alcohol and oxidised cofactor, for example in the formation of alcohols during fermentation by anaerobic bacteria and yeasts where the regeneration of NAD^+ is essential for other metabolic processes to continue (Clark, 1992). This is of use industrially in the production of alcoholic drinks and solvents. Alcohol dehydrogenases are involved in a remarkably wide range of metabolism. The interconversion of alcohols, aldehydes and ketones are important processes in both prokaryotes and eukaryotes. Substrates for alcohol oxidoreductases include

normal and branched-chain aliphatic and aromatic alcohols, both primary and secondary, and the corresponding aldehydes and ketones (Reid & Fewson, 1994).

The microbial alcohol dehydrogenases can be divided into three major categories (for review, see Reid & Fewson, 1994):

1. The NAD- or NADP-dependent dehydrogenases. This superfamily of enzymes is the best characterised of the three groups and can be further divided into three sub-groups. These are the group I long-chain (approximately 350 amino acid residues) zinc-dependent enzymes, e.g. alcohol dehydrogenases I, II, and III of *Saccharomyces cerevisiae* and the TOL-plasmid encoded benzyl alcohol dehydogenase of *P. putida*; the group II short-chain (approximately 250 residues) zinc-independent enzymes, e.g. ribitol dehydrogenase of *Klesbiella aerogenes* ; and the group III "iron-activated" enzymes. The enzymes of this latter category generally contain about 385 amino acids, e.g. alcohol dehydrogenase II of *Zymomonas mobilis* and alcohol dehydrogenase IV of *Saccharomyces cerevisiae*. However this group of enzymes contains enzymes of almost 900 residues in the case of the multifunctional alcohol dehydrogenases of *Escherichia coli* and *Clostridium acetobutylicum*.

2. NAD(P)-independent enzymes. These use pyrroloquinoline quinone, haem or cofactor F_{420} as cofactor. Examples of this group of enzymes are methanol dehydrogenase of *Paracoccus denitrificans* and the ethanol dehydrogenases of *Acetobacter* and *Gluconobacter* spp.

3. FAD-dependent alcohol oxidases. Theses enzymes catalyse the irreversible oxidation of alcohols, e.g. methanol oxidase of *Hansenula polymorpha*.

The group I NAD(P)-dependent long-chain zinc-dependent alcohol dehydrogenases do not exclusively contain enzymes from microbial sources. Besides fungal and bacterial enzymes, mammalian and plant enzymes can be classified as belonging to this group of enzymes (Jörnvall *et al.*, 1987b). The mammalian zinc-containing alcohol dehydrogenases constitute an enzyme family of multiple forms (Eklund *et al.*, 1990). Subunit types α, β_1, β_2, β_3, γ_1 and γ_2 in dimeric combinations constitute the isozymes of the human class I enzyme, whereas the human class II (π subunit), III (χ subunit) and IV (σ subunit) enzymes differ more considerably in primary structure and so constitute essentially separate and distinct enzymes (Jörnvall *et al.*, 1987a; Farrés *et al.*, 1994; Kedishvili *et al.*, 1995). The nine different subunit polypeptides are encoded by six different genes, two of which are polymorphic, and the active dimers formed from these subunits differ in substrate specificity and organ distribution (Bühler & von Wartburg, 1982; Wagner *et al.*, 1983; Ditlow *et al.*, 1984; Eklund *et al.*, 1990; Kedishvili *et al.*, 1995). The class I enzymes are

homologous to the alcohol dehydrogenases from horse liver (E or "ethanol active" chain) and rat class I enzymes. The rat appears to have a system of alcohol deydrogenases similar to that of the human (Julià et al., 1987; 1988). Long-chain zinc-dependent alcohol dehydrogenaes from plants such as maize and pea appear to resemble the human class II and III enzymes (Dennis et al., 1984; Borrás et al., 1989; Eklund et al., 1990). NADPH-Dependent aromatic alcohol dehydrogenases involved in lignification in vascular plants also belong to the of long-chain zinc-dependent alcohol dehydrogenase family, although they share less sequence identity with the mammalian enzymes than the maize and pea alcohol dehydrogenases (Grima-Pettenati et al., 1993).

1.10.2 Horse liver alcohol dehydrogenase

Amongst the group of zinc-dependent long-chain alcohol dehydrogenases (section 1.10.1) by far the best characterised is horse liver alcohol dehydrogenase. There is a large amount of information concerning the primary and tertiary structure of the apoenzyme and the holoenzyme (Jörnvall, 1970; Eklund et al., 1974; 1976; 1981; 1984; Ramaswamy et al., 1994). As a consequence of this extensive amount of background information, horse liver alcohol dehydrogenase is used as a paradigm for other members of the group.

Horse liver alcohol dehydrogenase is a 40 KDa dimer of identical subunits each containing 374 amino acid residues (Jörnvall, 1970). The tertiary structure is arranged into a catalytic domain and a coenzyme binding domain, divided by a long cleft that contains a wide and deep hydrophobic pocket. Each catalytic domain binds two zinc atoms. One of these zinc atoms is catalytic and is bound at the active site and the other is structural (Eklund et al., 1976).

There are structural differences between the apo- and holo-enzymes of the horse liver enzyme brought about by the binding of the cofactor NAD^+ (Eklund & Bränden, 1979; Eklund, 1988). The two coenzyme binding domains of the dimer form the centre of the enzyme molecule (Fig 1.7). When the coenzyme binds there is a conformational change in which the catalytic domains rotate towards each other and narrow the cleft between the two domains. This change provides one part of the substrate binding cleft. The conformational change increases the number of interactions between the coenzyme and the enzyme and brings Ser-48 and His-51 into positions to enable proton translocation. The binding of NADH is so tight that its release from the enzyme is the rate determining step in alcohol oxidation (Dalziel, 1975). This is an ordered reaction in which NAD^+ binds first and NADH is released last (Dalziel, 1975), although it has been suggested that the slow isomerization of the ternary complex is partially

Figure 1.7 Schematic drawing showing the conformational change of horse liver alcohol dehydrogenase induced by coenzyme binding

A, refers to the catalytic domain; B, refers to the coenzyme binding domain. The coenzyme parts covered by the domain are drawn with dashed lines.
Taken from Eklund (1988).
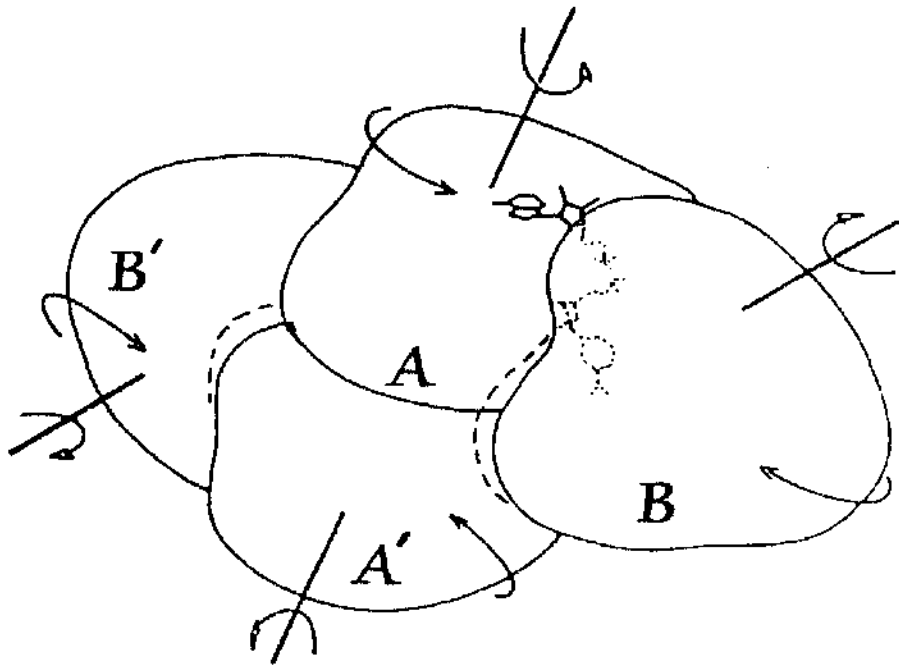
rate limiting in the turnover of ethanol and NAD$^+$ (Plapp *et al.*, 1986). Benzyl alcohol oxidation by horse liver alcohol dehydrogenase has also been shown to follow an ordered bi bi mechanism, although at high concentrations of benzyl alcohol enzyme-alcohol and abortive ternary complexes are formed, and benzaldehyde is slowly oxidised to benzoic acid (Shearer *et al.*, 1993). The reverse reaction for aldehyde reduction is random as regards binding of acetaldehyde and NADH but is ordered in terms of the release of the products, with desorption of coenzyme being the last step (Pettersson, 1987). Binding of NAD$^+$ below pH 10.0 is regulated by the protonation state of zinc-bound water in the free enzyme. This ionisation has a pK$_a$ of 9.2 (Andersson *et al.*, 1981). Holoenzyme formation perturbs this pK$_a$ down to 7.6 so that an alcholate ion intermediate can be formed when alcohol substrate is bound to the catalytic zinc atom (Kvassman *et al.*, 1981; Andersson *et al.*, 1981). The alcholate ion then undergoes hydride transfer to yield the aldehyde. Ligand binding and dissociation occur via the open enzyme conformation, whereas hydride transfer occurs within the closed conformation (Sartorius *et al.*, 1987). It has been proposed that the imidazole ring of His-51, with its pKa value close to 7.0, and which is exposed to bulk solvent, acts as the actual proton acceptor *in vivo*. However, the imidazole ring of His-51 is about 6 Å from the enzyme-bound alcohol substrate hydroxyl group and so direct proton transfer cannot occur (Eklund *et al.*, 1982). Rather, a proton relay system formed by hydrogen-bonding between the zinc bound alcohol substrate, the hydroxyl group of Ser-48, the 2'-hydroxyl of the coenzyme ribose ring and the imidazole ring of His-51 is thought to provide a link between the buried alcohol and the solvent at the surface of the enzyme (Fig. 1.8; Eklund *et al.*, 1982).

The catalytic zinc atom is essential for orienting the alcohol during the reaction and for stabilising the charged intermediate of the reaction. The zinc atom is bound to the ligands Cys-46, His-67 and Cys-174. A tetrahedral coordination of this zinc is completed by a water molecule in the apoenzyme and by the oxygen atom of the substrate in the holoenzyme (Eklund *et al.*, 1976; 1981; 1982). During the reduction of aldehyde, the zinc ion polarizes the substrate such that its first carbon atom increases its electrophilicity to enable hydride transfer to it (Eklund *et al.*, 1976). The structural zinc atom is also ligated in a tetrahedral type manner, this time by the sulphur atoms from the cysteine residues 97, 100, 103, and 111 (Eklund *et al.*, 1976).

Horse liver alcohol dehydrogenase is a class-A enzyme with respect to hydride transfer (Schneider-Bernlöhr *et al.*, 1986). Hydride transfer by dehydrogenases is always stereospecific in that some dehydrogenases transfer the

Figure 1.8 Hydrogen-bonding system proposed to be involved in the proton relay mechanism of horse liver alcohol dehydrogenase

Taken from Ramaswamy *et al.* (1994).
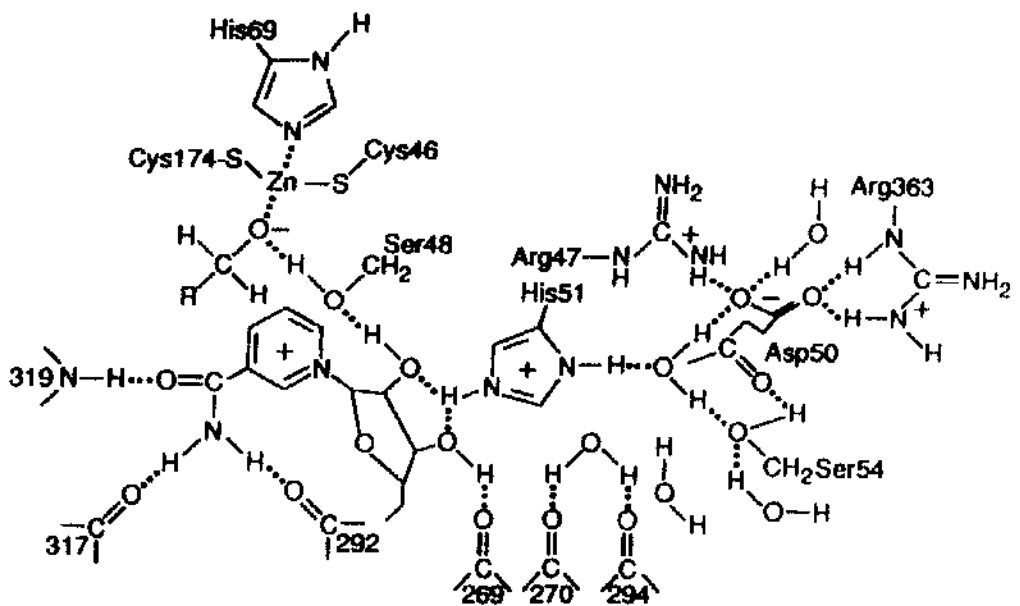
hydride ion to the *pro*-R or the *pro*-S positions of the nicotinamide ring. Class-A enzymes transfer the hydride ion to the *pro*-R position and class-B enzymes transfer the hydride ion to the *pro*-S position. Different sides of the nicotinamide ring are presented to the active site in these two classes of enzymes. This is because class-A enzymes bind the coenzyme with the nicotinamide ring in the *anti* conformation and class-B enzymes bind the coenzyme with the nicotinamide ring in the *syn* position, hence there is a 180° rotation of the nicotinamide ring around the glycosidic bond with respect to the adenosine ribose moiety.

1.10.3 Industrial biotransformations of alcohols and aldehydes by alcohol dehydrogenases

The advance in molecular biology over the last two decades that have enabled the cloning and sequencing of the genes encoding alcohol dehydrogenases such as horse liver alcohol dehydrogenase is likely also to advance the use of alcohol dehydrogenase enzymes in biotransformation systems. The high level expression of proteins from cloned genes will greatly facilitate the purification of enzymes in much larger quantities than before and protein engineering will probably enable the improvement of enzyme stability, activity and enable the alteration of substrate specificites. Genetic engineering to introduce novel proteins into whole cell systems may also prove to be a powerful tool. For example, a *Pseudomonas* sp. strain has been genetically manipulated such that it can now degrade chloro- and methylaromatics, many of which are pollutants, through a constructed *ortho* cleavage pathway (Pipke *et al.*, 1992). Also, a detailed knowledge of how alcohol dehydrogenases are regulated in the metabolic pathways in which they operate will facilitate their use in whole cell biotransformation systems. The wide substrate specificity of many enzymes, coupled with the diversity of biochemical pathways derived from microbial, animal and plant systems enables the production of a diverse range of chemical compounds by biotransformations using whole cell and isolated enzyme systems. Furthermore, the advent of protein engineering further increases the potential for synthesising compounds that might otherwise be difficult and expensive to synthesise.

Alcohol oxidoreductases have a wide range of possible applications within biotransformation systems. Biotransformations involving the interconversion of alcohols, aldehydes and ketones by alcohol oxidoreductases have great potential for the commercial production of pure optically active compounds and also for the treatment of industrial effluents. For example,

benzaldehyde is converted by the yeasts *Saccharomyces cerevisiae* and *Candida utilis* to the optically active compound L-phenylacetyl carbinol (Long & Ward, 1989; Shin & Rogers, 1995). This is a key step in the manufacture of L-ephedrine, a pharmaceutical that has α- and β-adrenergic activity and is an important component in some asthma medicines. Also, benzaldehyde is commonly found in wastewaters produced by the pharmaceutical and cosmetic industries (Wisniewski et al., 1983). The biotransformation of benzaldehyde to the less toxic benzyl alcohol prior to release of this wastewater into the environment would have impact in reducing the pollution caused by these industries.

Two different approaches are used in biotransformations. These are the use of whole cell or isolated enzyme systems. Which system is used for a particular biotransformation largely depends upon the relative cost of whole cells compared with that of the purified enzyme(s) (Pugh et al., 1987). The relative advantages of the two systems depends upon a number of parameters such as enzyme specificity, permeability of cells to substrate/product and the number of steps involved in the transformation of substrate to product (Pugh et al., 1987). A major problem encountered with both systems was that it is generally impractical to re-use suspensions of biocatalysts (including whole cells and enzymes) because their recovery from reaction mixtures is difficult and uneconomical (Hartmeier, 1985). To a large extent this problem has been overcome by the immobilisation of biocatalysts (whole cells & isolated enzymes) on solid supports. This allows both reuse of biocatalysts and their application in continuous processes. Furthermore, the co-immobilisation of isolated enzymes and their cofactors has greatly improved the efficiency of isolated cell systems (Hartmeier, 1985).

Bacteria, actinomycetes and fungi have all been used in whole cell systems for the biotransformation of alcohols, aldehydes and ketones (for review, see Pugh et al., 1987). Whole cell systems can take the form of viable growing cells which require a carbon and energy source to effect the conversion or non-viable cells, such as dried cell powder, which require a source of redox power. Whole cell systems have the advantage that the cells are able to continually regenerate enzyme cofactors as part of normal metabolism and thus maintain redox balance within the cell without the need to add expensive cofactors to the reaction. The phenomenom of cometabolism can also be used in whole cell systems whereby the organism will use a simple organic compound such as D-glucose as a carbon source and then use the "energy" produced to transform other added compounds. In addition, more complex transformations which involve a series of different enzyme reactions and cofactors are ideally suited to whole cell systems which have the necessary metabolic pathways to effect the transformation. However, whole cell systems have the disadvantage that they

often catalyse unwanted side-reactions which decrease the final yield of the required product (Long *et al.*, 1989). This latter fact illustrates the need for knowledge of how any metabolic pathways are regulated, and how their products interact with other pathways and enzymes.

Saccharomyces cerevisiae and *Candida utilis* have been used to produce L-phenylacetyl carbinol from benzaldehyde (Long *et al.*, 1989; Shin & Rogers, 1995). This reaction is catalysed by the pyruvate decarboxylase complex which decarboxylates pyruvate to acetaldehyde which then forms phenylacetyl carbinol in the presence of benzaldehyde as cosubstrate. However, in whole cell systems the presence of alcohol dehydrogenase brings about the conversion of some benzaldehyde to benzyl alcohol and benzoic acid, so reducing the final yield of L-phenylacetyl carbinol. Also this particular system illustrates the potential problem of toxic side effects brought about by substrate and/or products in biotransformations (Long & Ward, 1989). During biotransformation using *Saccharomyces cerevisiae* increased benzaldehyde concentrations in fermentations reduced yeast growth rate and, eventually, cell viability. It also altered the cells' permeability barrier to benzaldehyde, thus exposing enzymes within the cell to the denaturing effects of high benzaldehyde concentrations. Alcohol dehydrogenase was particularly susceptible, so reducing the cells' capacity to regenerate NAD^+ required for glycolysis. The alteration in cell permeability also allowed the release of cofactors TPP and magnesium from the cell, so reducing their cellular concentration. Hence the advantages of the whole cell system were lost when higher concentrations of benzaldehyde were to be transformed. These problems were overcome to some extent when using *Candida utilis* by immobilising the yeast, maintaining a constant benzaldehyde level and pulse feeding glucose (Shin & Rogers, 1995).

A whole-cell system has been used for the continuous biotransformation of benzaldehyde to benzyl alcohol by *Rhodotorula mucilaginosa* immobilised in an ultrafiltration cell (Wisniewski *et al.*, 1983) and benzaldehyde and derivatives of benzaldehyde have also been shown to be transformed to benzyl alcohol by *Saccharomyces cerevisiae* immobilised on various hydrophobic and hydrophilic polymer supports in aqueous and aqueous-organic two phase systems (Nikolova & Ward, 1994a, b). This has potential for the detoxification of industrial waste water.

The methanotrophs *Methylosinus trichosporium* and *Methylococcus capsulatus* have been identified as a major source of oxidised products and so resting cells have been identified as transforming alkanes (C_2-C_{16}) to alcohols and aldehydes, *m*-cresol to *m*- and *p*-hydroxybenzaldehyde, toluene to benzyl alcohol and methanol to formaldehyde (Pugh *et al.*, 1987).

Isolated redox enzyme systems have a number of advantages over whole cell systems for biotransformations. There is no requirement for a carbon and energy source and greater control can be achieved over product yield, process control and product separation (Bowen & Pugh, 1985). The use of isolated enzymes also allows much greater reaction specificity, particulary for the production of chiral compounds, and they also allow better control of the direction (oxidation/reduction) of the reaction (Hummel & Kula, 1989; Pugh *et al.*, 1987; Keinan *et al.*, 1986). Disadvantages of the use of purified enzymes are that they have a specific requirement for added cofactors and need a recycling system for these cofactors. Purified enzymes are also often more unstable than the cells that they were derived from and the availability and cost of the enzymes, or if necessary their purification, can be a problem.

Three major systems of cofactor recycling have been developed (Bowen & Pugh, 1985; Pugh *et al.*, 1987). Chemical systems have been used; sodium dithionite for NADH recycling and flavin mononucleotide and phenazine methosulphate for NAD^+ recycling. The second system developed is that of enzymatic recycling with sacrificial substrates. This technique involves the same dehydrogenase that is performing the transformation using a sacrificial second substrate that has been added to the reaction but in the reverse reaction direction. By utilising this substrate cofactor is regenerated. This system has been used in the horse liver alcohol dehydrogenase catalysed reduction of 2-(4-imidazolyl){1-H^3}acetaldehyde to (S)-{1-H^3}histaminol with NADH recycled using either ethanol or cyclohexanol. A second enzyme or multienzyme complex can also be used to utilise the sacrificial substrate and so recycle the cofactor. Thirdly, electrochemical methods can be used to regenerate cofactor. This involves an electrode being inserted into the reaction vessel and this producing the correct electrochemical gradient in order to regenerate reduced or oxidised cofactor.

Horse liver alcohol dehydrogenase (section 1.10.2) is the most widely used enzyme in biotransformations (Pugh *et al.*, 1987). The enzyme is extremely useful because of its ability to operate stereospecifically on a wide range of structurally varied substrates. The alcohol dehydrogenase of *Thermoanaerobium brockii* has been used for the asymmetric reduction of aliphatic acyclic ketones (C_4-C_{10} substrates). This enzyme can stereoselect between different sizes of substrates. Smaller substrates are reduced to R alcohols whereas larger ketones are reduced to the S enantiomer (Keinan *et al.*, 1986). High molecular weight aliphatic and aromatic alcohols have been shown to be oxidised by alcohol oxidase from *Pichia pastoris* (Murray & Duff, 1990). Many of these alcohols have a low solubility in water and biotransformation can be achieved using a biphasic system in which alcohol oxidase is in an aqueous phase and the substrate

is in an organic phase. The enzyme receives a steady stream of substrate and releases product into the organic phase, so reducing product inhibition and facilitating the recovery of the biocatalyst.

1.11 The genus *Acinetobacter*

The work in this thesis has followed on from much work in Glasgow on the bacterium *Acinetobacter calcoaceticus* NCIB8250. Acinetobacters are Gram-negative, non-motile, oxidase-negative, catalase-positive bacteria which lack pigmentation and have a G+C content of 38-45 mol % (Towner *et al.*, 1991). For a long time members of this genus have been incorrectly classified under at least 15 different generic names but are now classified as members of superfamily II of the gamma subdivision of the Proteobacteria and in the family Moraxellaceae as based on DNA/rRNA hybridisation (Rossau *et al.*, 1991; Olsen *et al.*, 1994). Acinetobacters are related to pseudomonads, with the two sharing the feature of metabolic versatility and having similar sets of metabolic pathways. Acinetobacters are ubiquitous organisms in the environment and are of importance because of their implication in food spoilage and because they are one of the main causes of nosocomial infections. This latter problem is exacerbated by the fact that many *Acinetobacter* strains are resistant to a variety of commonly used antibiotics and may form a reservoir of antibiotic resistance genes in hospital environments (Towner *et al.*, 1991).

A striking feature of strains of *Acinetobacter* is their ability to grow on a wide range of aromatic compounds as sole carbon sources (Fewson, 1991). Some individual strains can use more than 30 different aromatic compounds as sole carbon source (e.g. Fewson, 1967).

1.12 Work immediately preceeding this thesis

Some of the work undertaken in this thesis was a continuation of work begun by Dr A. G. S. Robertson in this laboratory (all unpublished results). This had resulted in the cloning and sequencing of *xylC* from *A. calcoaceticus* NCIB8250, the gene encoding benzaldehyde dehydrogenase II, and also the cloning and sequencing of part of *xylB*, the gene encoding benzyl alcohol dehydrogenase (section 3.1). The nucleotide sequence of *xylC* is shown in Fig. 1.9. The structural gene was found to be 1452 bp in length, encoding a 51,654 Da protein consisting of 484 amino acids per subunit. The *xylC* genewas shown to

Figure 1.9 Nucleotide sequence of *xylC* from *A. calcoaceticus* and the derived amino acid sequence of benzaldehyde dehydrogenase II

The full nucleotide sequence of *xylC* , and the start of *xylB* , are shown. Putative Shine-Dalgarno sequences are shown in bold print (nucleotides 320-323 & 1803-1807). Symbols correspond to: * xylC stop codon (1780-1782); # xylB translational start codon (1816-1818). Single letter symbols for the derived amino acid sequences of benzaldehyde dehydrogenase II and benzyl alcohol dehydrogenase are also shown.

Data from A. G. S. Robertson, unpublished results.

```
      10                     30                        50
       .                      .        .        .       ..
aattctaaacaagagaatcgccttcggttatgccgcaaattatgggggtg


              70                     90
       .        .        .        .        .
aacaacagctggatggaatcaaactggatcagaacatggaaaagcaacaa


      110                    130                       150
       .        .        .        .        .
gcgcgattggagtttcagcaaatgctgaatgcaaaaactcaattaagtgc


              170                    190
       .        .        .        .        .
acagttgctttcagacacacgaacagaaagcggatttaagaaagattacg


      210                    230                       250
       .        .        .        .        .
gtctaaattttcgtgtgtttтacattttttaatttaaagcaggtgtaaca


              270                    290
       .        .        .        .        .
cattgattgcgtcattтttacaaggatgcagttggaacagtaaagcaatt


      310                    330                       350
       .        .        .        .        .
gtctaaataaattttgaagaggtcatcatgtctatattcacaaaagagct
                               M  S  I  F  T  K  E  L    8


              370                    390
       .        .        .        .
ttgggataaaaaattatttaacggttcatggcagtcagcccaagacacct
W  D  K  K  L  F  N  G  S  W  Q  S  A  Q  D  T  Y   25
```

```
            410                 430                 450
             .                   .         .         .         .
acagtgtgattgaggttgcaacaggtcaagtgcttggtgaaatcggttat
  S   V   I   E   V   A   T   G   Q   V   L   G   E   I   G   Y     41


                  470                 490
             .         .         .         .         .
gccactgccgctgatgtggtgagtgctgcacaacaagccaaagccgctca
  A   T   A   A   D   V   V   S   A   A   Q   Q   A   K   A   A   Q     58


        510                 530                 550
         .         .         .         .         .
acaacagtggtgggctttaaattatcaagagcgtcaggcagtgtttgagc
  Q   Q   W   W   A   L   N   Y   Q   E   R   Q   A   V   F   E   R     75


                  570                 590
             .         .         .         .         .
gagctgctgctttactgacagaaaatcaagcggaagtgattgagtggctg
   A   A   A   L   L   T   E   N   Q   A   E   V   I   E   W   L     91


        610                 630                 650
         .         .         .         .         .
gttaaggaaagtggctcattacaactgaaagcaggctttgaagtgagtat
  V   K   E   S   G   S   L   Q   L   K   A   G   F   E   V   S   I   108


                  670                 690
             .         .         .         .         .
tgcgattcaagtgcttaaacattgtattgcttcacctacaaatgaacagg
  A   I   Q   V   L   K   H   C   I   A   S   P   T   N   E   Q   G 125


        710                 730                 750
         .         .         .         .         .
gcacattattacctacgcaaaatggcaaactaagtatagcgaaacgcttg
   T   L   L   P   T   Q   N   G   K   L   S   I   A   K   R   L   141


                  770                 790
             .         .         .         .         .
cccttaggtgttgttggggtcatttctccatttaactttccactgtactt
  P   L   G   V   V   G   V   I   S   P   F   N   F   P   L   Y   L   158
```

```
        810                   830                    850
         .         .          .          .          .
agctttacgtgctgtggcacctgctttggcctttggtaatgcggtggtgt
 A   L   R   A   V   A   P   A   L   A   F   G   N   A   V   V   L   175


                  870                   890
         .         .          .          .          .
taaaacctgatgagcgtacagcagtgtgcagtggctatgtgattgcacgg
  K   P   D   E   R   T   A   V   C   S   G   Y   V   I   A   R   191


        910                   930                    950
         .         .          .          .          .
atttttgaacttgcaggattaccaaagggggctgctgcatgtattgccagg
 I   F   E   L   A   G   L   P   K   G   L   L   H   V   L   P   G   208


                  970                   990
         .         .          .          .          .
tggggtagaagttggtgaagcactgactttggatcaaaatattgccagta
 G   V   E   V   G   E   A   L   T   L   D   Q   N   I   A   S   I   225


        1010                  1030                   1050
         .         .          .          .          .
tccaatttacgggatcaacccaagttggacggattgttggggcgaatgcg
  Q   F   T   G   S   T   Q   V   G   R   I   V   G   A   N   A   241


                  1070                  1090
         .         .          .          .          .
gcaaagaccttaaagaaagtttctctagaactgggtgggaaaaactcact
 A   K   T   L   K   K   V   S   L   E   L   G   G   K   N   S   L   258


        1110                  1130                   1150
         .         .          .          .          .
gattattttggatgatgcggatattgagttagcagcagaaaatattgcat
  I   I   L   D   D   A   D   I   E   L   A   A   E   N   I   A   W   275


                  1170                  1190
         .         .          .          .          .
ggggtgcattttacattctgggcagatttgtatgacttctggcaagatt
  G   A   F   L   H   S   G   Q   I   C   M   T   S   G   K   I   291


        1210                  1230                   1250
         .         .          .          .          .
ttgattcaccaaaaaatttatcaacaggtcaaacagcgtgtaattgaaaa
 L   I   H   Q   K   I   Y   Q   Q   V   K   Q   R   V   I   E   K   308
```

```
                1270                    1290
                  .         .         .         .         .
agtgcagaatttttgtggtagggaatccttgtgacaagaatgtcaccattg
 V  Q  N  F  V  V  G  N  P  C  D  K  N  V  T  I  G 325
        1310                 1330                 1350
        .         .         .         .         .
ggccattgatcaatgctaaacaagcgcaacgtgtagaacaattggtcagt
  P  L  I  N  A  K  Q  A  Q  R  V  E  Q  L  V  S  341


                1370                    1390
                  .         .         .         .         .
gctgcggttaatgaaggcgcaacattggaaatcggcgggcatgccgatgg
 A  A  V  N  E  G  A  T  L  E  I  G  G  H  A  D  G  358


          1410                 1430                 1450
          .         .         .         .         .
ggtattttttcaaccgactgtactgacagatgtgactgcaaacaattcga
 V  F  F  Q  P  T  V  L  T  D  V  T  A  N  N  S  I 375


                1470                    1490
                  .         .         .         .         .
tttttagtgaagagattttttggtcctgttgcggtattgatcccgtttagc
  F  S  E  E  I  F  G  P  V  A  V  L  I  P  F  S  391


          1510                 1530                 1550
          .         .         .         .         .
tcagatgaacaggccattgaattagccaatgacggcgattatgggctttc
 S  D  E  Q  A  I  E  L  A  N  D  G  D  Y  G  L  S  408


                1570                    1590
                  .         .         .         .         .
agcaggcattattacctccaatgtgggacgaggtatgcaactcggtgcgc
 A  G  I  I  T  S  N  V  G  R  G  M  Q  L  G  A  Q 425


          1610                 1630                 1650
          .         .         .         .         .
agctcaaagtgggcttgctacatatcaacgatcagaccgtaaatgatgaa
  L  K  V  G  L  L  H  I  N  D  Q  T  V  N  D  E  441
```

```
        1670                    1690
         .         .         .         .         .
acggtcaatccatttggtggttttggttcatcaggtaatggtacgcgtat
T  V  N  P  F  G  G  F  G  S  S  G  N  G  T  R  I  458


        1710                    1730                    1750
         .         .         .         .         .
tggcggccctgccaatgccgatgaatttacccagtggcaatggattacag
 G  G  P  A  N  A  D  E  F  T  Q  W  Q  W  I  T  V  475


        1770                    1790
         .         .         .         .         .
tacaggctcaagcgccacattatccttttaaacaattaaataaatcatg
 Q  A  Q  A  P  H  Y  P  F  *


        1810                    1830
         .         .         .
acaggacgttataccatgagtgaattaaaa
    #  S  E  L  K
```

have a codon usage pattern typical of *Acinetobacter* structural genes and reflected the G+C content of the organism. Benzaldehyde dehydrogenase II was expressed in *E. coli* and the recombinant enzyme purified. This enzyme has 46% amino acid sequence identity with TOL-benzaldehyde dehydrogenase.

The nomenclature used in this thesis for *A. calcoaceticus* structural genes is based on that used previously for *P. putida* structural genes. In *P. putida* (pWW0), benzaldehyde dehydrogenase and benzyl alcohol dehydrogenase are involved in the metabolism of xylenes and the structural genes encoding these enzymes had been termed *xylC* and *xylB* respectively (Harayama *et al.*, 1989). In light of the similarity between the genes from *A. calcoaceticus* and *P. putida*, the genes encoding benzaldehyde dehydrogenase II and benzyl alcohol dehydrogenase in *A. calcoaceticus* have also been termed *xylC* and *xylB* in our laboratory. In *P. putida* the gene encoding L-mandelate dehydrogenase is named *mdlB* (Tsou *et al.*, 1990) and because of the lack of any genetic information on the mandelate pathway in *A. calcoaceticus*, the gene encoding L-mandelate dehydrogenase in this organism has been tentatively termed *mdlB* also.

1.13 Aims and scope of this thesis

The original aim of this work was to clone and sequence *mdlB*, the gene encoding L(+)-mandelate dehydrogenase from *Acinetobacter calcoaceticus* NCIB8250. Previous studies which had shown that the enzymes of the mandelate pathway are co-ordinately induced and expressed suggested that the genes for the enzymes of the mandelate pathway are organised in the form of an operon (Livingstone & Fewson, 1972; Fewson *et al.*, 1978). Mandelate metabolism in *A. calcoaceticus* has also been shown to dominate benzyl alcohol metabolism (section 1.7; Beggs *et al.*, 1976; Beggs & Fewson, 1977). Therefore cloning of the gene encoding L-mandelate dehydrogenase would probably lead to the cloning of the other genes encoding the mandelate pathway enzymes. This would have complemented work ongoing in the laboratory which was trying to clone the genes encoding the enzymes of the benzyl alcohol pathway in *A. calcoaceticus*, in particular allowing a comparison of the genes encoding the two NAD-dependent benzaldehyde dehydrogenases (section 1.8). *A. calcoaceticus* strain NCIB8250 can grow on only the L-mandelate enantiomer, and strain EBF65/65 can only grow on only the D-mandelate enantiomer and each of these strains posseses only the corresponding mandelate dehydrogenase. However, these strains give rise to spontaneous mutants that can grow on the second enantiomer. In every case this is due to the appearance of an extra mandelate dehydrogenase, specific for the other

enantiomer (Hills & Fewson, 1983), possibly due to the expression of a previously cryptic or silent gene (section 1.6). Cloning of the gene for L-mandelate dehydrogenase from strain NCIB8250 would enable the probing of other *Acinetobacter* strains for silent L-mandelate dehydrogenase genes. Cloning of these genes would then enable the determination of the mechanisms by which these genes are silenced and switched on. Unfortunately, attempts to clone the gene encoding L-mandelate dehydrogenase by several different techniques were unsuccessful (chapter 8).

Once several attempts to clone the gene encoding L-mandelate dehydrogenase had failed, the project concentrated on completing work that had been started on cloning the genes for the benzyl alcohol pathway in *A. calcoaceticus* NCIB8250 (section 1.12). The bulk of the work described in this thesis involved the cloning of a DNA fragment harbouring the full length of *xylC* and *xylB*, the genes encoding benzaldehyde dehydrogenase II and benzyl alcohol dehydrogenase respectively, and the further study of *xylB* and its product. The chief aims were:

- to determine the nucleotide sequence of the *xylB* gene and the derived amino acid sequence of benzyl alcohol dehydrogenase and to compare these sequences with those of other similar genes and enzymes;

- to determine if other open reading frames were present in the vicinity downstream of *xylB*;

- to develop a system which allowed the high level expression of benzyl alcohol dehydrogenase in *Escherichia coli*;

- to determine the function of certain amino acid residues in the catalytic mechanism, and how they influence the substrate specificity of benzyl alcohol dehydrogenase, by using a combination of site-directed mutagenesis, chemical modification and kinetic studies; in particular, to investigate reasons for the lack of a conserved histidine residue at position 51, a residue that is thought to act as a general base in catalysis in zinc-dependent alcohol dehydrogenases, and to test the possible role for arginine-53 as a base in catalysis.

A detailed knowledge of factors influencing the substrate specificity and the catalytic mechanism of benzyl alcohol dehydrogenase should facilitate its application and manipulation, and that of other long-chain zinc-dependent alcohol dehydrogenases, in biotransformation systems. Also, comparisons of the primary sequence and kinetic properties of benzyl alcohol dehydrogenase with those of other zinc-dependent alcohol dehydrogenases provides information concerning the evolutionary origins of this enzyme and the way in which enzyme function may have evolved in a diverse range of microorganisms, plants and mammals, and this is discussed in chapter 9.

Chapter 2 Materials and methods

2.1 Materials

All reagents used were of the best quality commercially available and with the exception of those listed below were obtained from BDH Chemicals Ltd, Poole, Dorset, UK.

Ampicillin, tetracycline, heparin, Ficoll, bicine, polyvinylpyrollidone (PVP), lysozyme, RNase, pronase, low melting point agarose, 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal), diethylpyrocarbonate, horse liver alcohol dehydrogenase, yeast alcohol dehydrogenase and Coomassie Brilliant Blue G250 were all obtained from Sigma (London) Chemical Co., Poole, Dorset, UK.

Glucose, urea, boric acid, EDTA, magnesium sulphate, ammonium persulphate, acrylamide, N,N´-methylene *bis*-acrylamide, sodium dodecyl sulphate and tris-washed phenol were all obtained from Fisons Scientific Equipment, Loughborough, Leics., UK.

Coniferyl alcohol, cinnamyl alcohol and perillyl alcohol were all obtained from Aldrich Chemical Co., Gillingham, Dorset, U.K.

2-Methylbenzyl alcohol, 3-methylbenzyl alcohol and 4-methylbenzyl alcohol were all obtained from Lancaster Synthesis, Morecambe, Lancs., U.K.

Bactotryptone and yeast extract were obtained from Merck, Darmstadt, Germany.

Bactoagar was obtained from Difco, Detroit, USA.

Isopropyl-β-D-thiogalactopyranoside (IPTG) and nicotinamide adenine dinucleotide (NAD$^+$) were obtained from Boehringer Corporation (London) Ltd, Lewes, Sussex, UK.

SeaKem GTG agarose was supplied by FMC Bioproducts, Rockland, USA.

Filter paper was obtained from Whatman International, Maidstone, Kent, UK.

Other molecular biological reagents were obtained as described in the relevant part of the text.

2.2 Bacterial strains and plasmids

Strain / Plasmid	Phenotype	Source /Reference
Acinetobacter calcoaceticus NCIB8250	Wild type isolate	National Collection of Industrial Bacteria, Torry Research Station, Aberdeen, Scotland.
Pseudomonas putida NCIB9494	Wild type isolate	As above
Escherichia coli JM109	*rec* A1, *sup* E44, *end* A1, *hsd* R17, *gyr* A96, *rel* A1, *thi* Δ(lac-proAB)	Promega Corp, Madison, USA.
E. coli JM109 (DE3)	As above plus λ_{DE3}	As above
E. coli DH5α	F⁻, Φ80d, *lac* ZΔM15, Δ(lac ZYA-arg F), U169*deo* R, *rec* A1, *end* A1, *hsd* R17 (r⁻k, m⁺k), *sup* E44, λ⁻, *thi* -1, *gyr* A96, *rel* A1.	As above
E. coli KW251	F⁻, *sup* E44, *sup* F58, *gal* K2, *gal* T22, *met* B1, *hsd* R2, *mcr* B1, *mcr* A-, *arg* A81:*Tn* 10, *rec* D1014	As above
pBM	*mdlB* from *P. putida* NCIB9494 cloned into the *EcoRI* and *PstI* sites of pKK223-3	Dr J. A. Gerlt, Department of Chemistry & Biochemistry, University of Maryland, Maryland 20742.
pUC18	General cloning plasmid Amp^r	Yanisch-Perron *et al.* (1985)
pBluescript SK II-	General cloning plasmid Amp^r	Stratagene Short *et al.* (1988).
pTB361	T7 expression plasmid tet^r . European patent application no. 92301456.8	M. J. Horsburgh, Division of Biochemistry & Molecular Biology, University of Glasgow.

2.3 Growth media

2.3.1 Luria-Bertaini medium(LB)

LB contained per litre:

        10g bactotryptone

        5g yeast extract

        10g sodium chloride

The above ingredients were dissolved in distilled water and the pH value adjusted to 7.5 with 5 M NaOH before making up to the final volume.

2.3.2 L-Agar

L-Agar was made by adding 15 g bacto-agar per litre of LB. Hot agar was allowed to cool to 50°C prior to the addition of antibiotic supplements.

2.3.3 H-top agar

H-top agar was used for the growth of bacteriophage. Molten H-top agar containing bacteriophage infected cells was poured onto the surface of hardened L-agar and contained per litre:

        10g bactotryptone

        8g sodium chloride

        8g bacto-agar.

2.3.4 Sterilisation of growth media

Growth media were sterilised by autoclaving at 120°C (15 lb in^{-2}) for 20 minutes.

2.3.5 Antibiotic supplements

Ampicillin was added to media to a final concentration of 100 µg/ml. A stock solution of 25 mg/ml was made by dissolving solid ampicillin in distilled water and sterilised by filtration through a Millipore 0.22 µm filter.

Tetracycline was added to media to a final concentration of 12.5µg/ml. A stock solution of 12.5 mg/ml was prepared by dissolving solid tetracycline in 50% v/v ethanol and sterilised by filtration through a Millipore 0.22 µm filter.

Antibiotic stock solutions were stored at −20°C.

2.4 Restriction endonuclease (RE) digestion of DNA

Restriction endonucleases were used as directed in the manufacturers' instructions. Typically an enzyme would be provided with a 10 x working concentration reaction buffer and DNA would be digested in a 50 µl volume containing 1 unit enzyme/µg DNA for 2 hours at 37°C.

Restriction digests of genomic DNA preparations, e.g. for Southern blotting, used larger amounts of enzyme; typically 5 units/µg DNA and were incubated for longer periods; typically 5 hours (Sambrook *et al.*, 1989).

2.5 Ligations

DNA was ligated using T4 DNA ligase (Promega Corp, USA). Ligations were carried out in a 10 µl total volume with a 1:3 molar ratio of plasmid: insert DNA (total DNA 100ng), 1 µl 10 x T4 DNA ligase buffer, 1 µl T4 DNA ligase (3 units) and made up to 10 µl with sterile H_2O. The digested DNA and H_2O were mixed in a microcentrifuge tube and the DNA "ends" melted at 45°C for 5 minutes. After chilling on ice, buffer and enzyme were added and the reaction incubated overnight at 16°C.

2.6 Calf intestinal alkaline phosphatase (CIAP) treatment

Digested plasmid DNA was treated with CIAP (Promega corp, USA) in order to prevent religation to itself. Reactions were carried out in a 100 µl volume with 10 µl 10 x reaction buffer and 1 µl CIAP (1 unit) at 37°C for 30 minutes.

2.7 Promega Wizard DNA clean-up system

After digestion with restriction endonucleases or CIAP treatment, samples were cleaned using the Promega Wizard DNA clean-up system according to the manufacturers' instructions.

2.8 Ethanol precipitation of DNA

DNA was concentrated by precipitation with ethanol. A 0.1 sample volume of 3M sodium acetate, pH 5.2 and 2.5 volumes of cold (–20°C) absolute ethanol were added to the samples to be concentrated. After mixing, samples were incubated at –20°C for 30 minutes and centrifuged at 12,000 r.p.m in a Jouan MR14.11 refrigerated centrifuge for a further 30 minutes. The supernatants were discarded and the pellets washed with cold (–20°C) 70% ethanol by centrifugation at 12,000 r.p.m in a Jouan MR14.11 refrigerated centrifuge for 10 minutes. After removal of the supernatants, the pellets were air-dried for 10 minutes and finally resuspended in a suitable volume of sterile distilled water.

2.9 Agarose gel electrophoresis

Samples to be run in an agarose gel were first mixed with an appropriate volume of 6 x STOP MIX. STOP MIX consists of 2 volumes of loading buffer (10 mM tris/HCl, pH 7.2, 1 mM EDTA, 20% ficol, 0.5% Bromophenol Blue, 0.05 mg ethidium bromide/ml) to 1 volume 0.4 M EDTA. Samples were typically run in a 0.75% agarose/TBE gel on a Hybaid Electro-4 electrophoresis system. Agarose gels used to purify samples for random-priming (section 2.22.1) were run in 40 mM Tris-acetate, pH8.0, 5 mM sodium acetate buffer.

2.10 Purification of DNA by agarose gel electrophoresis

The sample to be purified was run on a 0.75-1.00% low melting point agarose/TBE gel containing ethidium bromide (0.5 mg/ml). The tank buffer also contained 0.5 mg ethidium bromide/ml. Whilst viewing on a UV transilluminator, the required DNA band was excised using a clean scalpel. The gel slice was melted in an

equal volume of TBE at 65°C and extracted twice with phenol and once with chloroform (section 2.11).

2.11 Phenol/chloroform extraction

DNA samples to be purified by phenol extraction were mixed with an equal volume of TE saturated phenol (Fisons, UK). After vortexing, each sample was spun in a microfuge for 10 minutes. The upper aqueous layer was transferred to fresh microfuge tubes and the lower layer extracted again with TE. The two aqueous layers were pooled and extacted with phenol/TE. The aqueous layer was transferred to a fresh microcentrifuge tube and an equal volume of chloroform was added. After vortexing, the sample was spun in a microcentrifuge for 5 minutes and the aqueous layer transferred to a clean tube. DNA was then recovered by ethanol precipitation (section 2.8).

2.12 Preparation of competent cells of *Escherichia coli*

A culture of *E. coli* was grown overnight by inoculating a single colony of the required strain from an agar plate into 10 ml LB (section 2.3.1). The culture was incubated at 37°C with shaking (100 r.p.m). This culture (1 ml) was used to inoculate 50 ml LB which was incubated with shaking at 37°C until the optical density at 550 nm was approximately 0.4. The culture was chilled on ice for 10 minutes and harvested by centrifugation at 5000 r.p.m in a Beckman JA-20 rotor. The cells were gently resuspended in 0.5 volumes cold sterile 50 mM $CaCl_2$ and incubated on ice for 60 minutes. After harvesting again as previously, the cells were gently resuspended in 3 ml cold 50 mM $CaCl_2$ and 200 µl samples used for transformation.

2.13 Transformation of competent cells of *E. coli*

Before transformation with DNA from ligation reactions, reactions were diluted 5-fold with TNE (10 mM Tris/HCl, pH7.5, 1 mM EDTA). Typically, half of a ligation reaction was used to transform a 200 µl sample of competent cells (section 2.12). Competent cells and DNA were gently mixed in a microcentrifuge tube and incubated on ice for 30 minutes. The cells were heat-shocked at 42°C for 45 seconds, chilled on ice, and then made up to 1 ml with LB. Expression of antibiotic resistance

was allowed for 1 hour at 37°C and then 200 µl samples were spread onto selective agar-plates and incubated overnight at 37°C.

2.14 Preparation of genomic DNA from *Acinetobacter calcoaceticus* and *Pseudomonas putida*

A. calcoaceticus and *P. putida* genomic DNA was prepared as previously described for *P. putida* (Tsou *et al.*, 1990). A 50 ml overnight LB culture was harvested in a Beckman JA-20 rotor and the bacterial pellet resuspended in 10 ml 50 mM tris/HCl, pH8.0, 20 mM EDTA. after the addition of 1 ml self-digested pronase (5 mg/ml) and 0.1 ml 10% SDS the cells were allowed to lyse for 6 hours at 37°C. After 5 hours lysis, 50 µg RNase/ml was added (stock solution 10 mg/ml). The lysate was extracted three times with 10 ml TE-saturated phenol and the resulting aqueous layer was then extracted three times with water-saturated ether. The DNA was then spooled onto a sealed glass pasteur pipette after the addition of 0.1 volumes 3M sodium acetate, pH 5.2, and 2 volumes of cold absolute ethanol. Alternatively, the aqueous phase resulting from ether extraction was dialysed against 10 mM Tris/HCl, pH7.5, 1 mM EDTA, 100 mM NaCl for 24 hours and then against TE for a further 24 hours.

2.15 Small scale preparations of plasmid DNA

Small scale preparations of plasmid DNA for the screening of recombinant plasmids were performed as described by Birnboim & Doly (1979). Overnight 1ml cultures of the selected colonies following transformation were harvested in a microcentrifuge and the growth medium removed using an aspirator. The bacterial pellets were resuspended in 100 µl 50 mM glucose, 25 mM Tris/HCl, pH8.0, 10 mM EDTA; 25 µl (10 mg/ml) lysozyme was added and the samples incubated at room temperature for 5 minutes. The cells were then lysed by the addition of 200 µl 0.2 M NaOH, 1% SDS. The samples were mixed gently by inversion and incubated on ice for 5 minutes. The lysates were neutralised by the addition of 150 µl ice-cold 5 M potassium acetate, pH 4.8, gently mixed by vortexing, incubated on ice for 5 minutes and centrifuged at 12,000 r.p.m in a Jouan MR14.11 refrigerated centrifuge. The supernatants were transferred to fresh microcentrifuge tubes and the DNA recovered by ethanol precipitation (section 2.8). DNA pellets were finally resuspended in 20 µl sterile distilled water. RNA was removed from these preparations following

digestion with restriction endonucleases by the addition of 1 μl (10 mg/ml) DNase-free RNase just prior to loading on an agarose gel.

2.16 Large scale preparations of plasmid DNA

Large scale preparations of plasmid DNA for making stock solutions of plasmids and DNA sequencing templates were made using the Qiagen plasmid midi-kit (Qiagen Inc, Hilden, Germany) exactly as instructed in the manufacturer's protocol. Each preparation consistently yielded 50 μg of plasmid DNA.

2.17 Preparation of lambda (λ) bacteriophage DNA

2.17.1 Preparation of lambda liquid culture lysates

A genomic DNA library of *Acinetobacter calcoaceticus* had been made previously by Dr A. Robertson in this laboratory using the Lambda GEM-11 *Bam* H I cloning system (Promega Corporation, Madison, USA). Lysates from this library were used to make fresh lysates for the isolation of λ clones for further sub-cloning.

E.coli KW251 was used as host strain. An overnight culture was set up from a glycerol stock by inoculating a loopful of bacteria into 50 ml LB containing 500 μl 20% maltose and 500 μl 1M MgSO$_4$. This culture was then grown overnight at 37°C. Cells were harvested at 6,000 r.p.m in a Beckman JA-20 rotor for 15 minutes at 4°C. The cells were resuspended in 0.01 M MgSO$_4$ such that they had an optical density of approximately 2.0 at 600 nm.

E.coli cells prepared in this way were infected by incubating 300 μl of cell suspension with 100 μl of a λ liquid lysate at 37°C for 30 minutes. The infected cells were then inoculated into 100 ml LB containing 1 ml 20% maltose and 1 ml 1M MgSO$_4$ and incubated at 37°C with shaking for 8 hours. Complete lysis was ensured by the addition of 2 ml chloroform after 8 hours and the incubation continued for a further 15 minutes. After centrifugation at 10,000 r.p.m in a Beckman JA-20 rotor the now cleared liquid lysate was stored at 4°C.

2.17.2 Calculating phage titre of liquid lysates

To calculate the number of plaque forming units (p.f.u) per ml of λ liquid lysates, serial dilutions of lysates were made in SM buffer (50 mM tris/HCl, pH 7.5, 100 mM NaCl, 8 mM MgSO$_4$, 0.01% gelatin) and 100 μl of each dilution incubated with 300 μl *E.coli* KW251 cells in 0.01 M MgSO$_4$ (section 2.17.1) at 37°C for 30 minutes. The infected cells were added to 3 ml cooled molten H-top agar (section 2.3.3), the mixture poured onto hardened L-agar plates, allowed to harden, and then incubated inverted overnight at 37°C.

2.17.3 Preparation of lambda DNA

Lambda DNA was prepared using the Qiagen Lambda DNA midi-kit (Qiagen Inc, Hilden, Germany) exactly as described in the manufacturer's protocol. Lambda DNA liquid lysates were prepared as described in section 2.17.1.

2.18 Preparation of oligonucleotides

Oligonucleotides were made by Dr. V. Math, Division of Biochemistry and Molecular Biology, Institute of Biomedical and Life Sciences, University of Glasgow on an Applied Biosystems Model 280A DNA synthesiser using phosphoramidate chemistry.

Oligonucleotides were supplied in a solution of 35% (w/v) NH$_4$OH and were isolated by ethanol precipitation (section 2.8) and then finally resuspended in a suitable volume of sterile distilled water (typically 100 μl). The concentration of the purified oligonucleotide was determined by measuring the absorbance at 260 nm. Typically 20 μl of oligonucleotide was diluted in 3 ml of water and the absorbance measured in a 1 cm quartz cuvette, assuming that an absorbance of 1.0 corresponds to a concentration of 33 μg/ml for a single-stranded oligonucleotide.

2.19 DNA sequencing using Sequenase version 2.0

Sequenase version 2.0 (USB, Cleveland, USA) was used for double stranded sequencing of plasmid DNA using the dideoxy chain termination method (Sanger *et al.*, 1977). Template DNA was prepared using the Qiagen plasmid midi-kit (section 2.16).

2.19.1 Template denaturation and primer annealing

Approximately 4 μg template DNA was denatured in 0.2 M filter-sterilised NaOH, 0.2 mM EDTA at 37°C for 30 minutes. The mixture was then neutralised by adding 0.1 volumes 3 M sodium acetate, pH 5.2 and the DNA recovered by ethanol precipitation (section 2.8). The DNA pellet was finally redissolved in 7 μl sterile distilled water and 2 μl Sequenase reaction buffer and 1 μl (3-4 pmoles) primer added. The mixture was boiled for 1 minute in a boiling water bath and the primers allowed to anneal at 37°C for 30 minutes.

2.19.2 Labelling reactions

Prior to use in labelling reactions dGTP labeling mix was diluted 5-fold with distilled water and Sequenase was diluted 1:8 with ice-cold enzyme dilution buffer. Labelling reactions were set up as follows:

Template-primer	10.0 μl
DTT(0.1 M)	1.0 μl
Diluted dGTP labeling mix	2.0 μl
$[\alpha$-$^{35}S]$ dATP	0.5 μl
Diluted Sequenase version 2.0	2.0 μl

Labelling reactions were incubated at room temperature for 2 minutes. $[\alpha$-$^{35}S]$ dATP (10 μCi/μl; 1000 Ci/mmol) was supplied by Amersham, UK. For sequencing very close to the primer, 1 μl manganese buffer was included in the labelling reaction.

2.19.3 Termination reactions

For each labelling reaction, 4 tubes were set up such that each contained either 2.5 μl of ddATP, or ddCTP, or ddGTP, or ddTTP termination mix. These tubes were pre-warmed at 37°C for at least 1 minute. Once the labelling reaction was complete, 3.5 μl was transferred to each of the tubes containing termination mix and the reactions incubated at 37°C for 3 minutes. Finally, 4 μl of stop solution was added to each reaction and the samples heated to 75-80°C for 2 minutes before loading 3 μl onto a 6% polyacrylamide gel (section 2.19.4).

2.19.4 Polyacrylamide gel electrophoresis

Polyacrylamide sequencing gels were run using the IBI STS 45 sequencing apparatus at 60 W constant power. Sequencing gels were pre-run for 30 minutes in TBE prior to loading. To obtain longer DNA sequences, 0.4-1.2 mm wedge gels were used. Gels were fixed in 10% (v/v) acetic acid, 10% (v/v) methanol for 30 minutes following electrophoresis and were vacuum dried onto a sheet of Whatman 3MM paper using a Flowgen Proebel gel drier. The dried gel was autoradiographed using Fuji RX film at room temperature.

2.20 Southern blotting of DNA

DNA was transferred to Hybond N nylon membranes (Amersham, UK) using the capillary blot method (Southern, 1975). Restriction digests of genomic DNA (section 2.4) were separated on a 0.75% agarose/TBE gel (section 2.9). The gel was soaked in 0.25 M HCl until the dyes in the gel had changed colour, plus an additional 10 minutes. This step "nicks" large DNA fragments in order to aid blotting. The gel was rinsed in distilled water and the DNA denatured by soaking in 1.5 M NaCl, 0.5 M NaOH for 30 minutes with gentle shaking. After rinsing in distilled water the DNA was neutralised by soaking in 1.5 M NaCl, 0.5 M tris/HCl, pH7.2, 0.001 M EDTA for 15 minutes, again with gentle shaking. This neutralisation step was repeated with fresh solution.

A glass dish was filled with blotting buffer (20 x SSC; 3 M NaCl, 0.3 M-tri-sodium citrate) and a platform made using a gel tray. A wick was made from 3 sheets of Whatman 3MM filter paper, saturated with blotting buffer. The gel was placed on top of the wick and surrounded with Saran Wrap to prevent the blotting buffer by-passing the gel and "short-circuiting" the blotting process. A piece of Hybond-N membrane cut to the exact size of the gel was placed on top of the gel and any air bubbles squeezed out using a glass rod. Three pieces of Whatmann 3MM paper cut to size and wetted with blotting buffer were placed on top of the membrane and a stack of absorbent paper towels (approximately 5 cm high) placed on top of the 3MM paper with a weight on top. Transfer was allowed to proceed for 16 hours.

After transfer was complete, the gel was restained in ethidium bromide and viewed on a UV transilluminator to check that transfer had been successful. The membrane was washed in 6 x SSC for 5 minutes to remove any bound agarose, wrapped between two sheets of Whatman 3MM paper and the DNA fixed to the membrane by baking at 80°C for 2 hours. Membranes were stored in air-sealed bags at room temperature.

2.21.1 Hybridisation of radiolabelled probes to Southern blots

Membranes were pre-washed in a heat-sealed plastic bag with 100 ml pre-washing solution (5 x SSC, 0.5% SDS, 1 mM EDTA) for 2 hours at 42°C with shaking. The filter was transferred to a clean plastic bag containing 40 ml pre-hybridisation solution (5 x SSC, 0.5% SDS, 0.05% sodium pyrophosphate, 200 μg heparin/ml or 5 x Denhardt's solution; 50 x Denhardt's solution is 1% bovine serum albumin, 1% Ficoll, 1% polyvinylpyrrolidone). Heparin was used as the blocking agent when using end-labelled oligonucleotides and Denhardt's solution was used when using random-primed probes. Pre-hybridisation was carried out at 65°C for at least 4 hours. Probes were prepared as described in section 2.22 and were hybridised in 20 ml hybridisation solution (5 x SSC, 0.5% SDS, 0.05% sodium pyrophosphate, and if necessary 200 μg heparin/ml) at the required temperature for 3-4 hours, in the case of oligonucleotide probes, or overnight in the case of random-primed probes. When hybridisation was complete the membrane was removed and was washed in 200 ml 5 x SSC, 0.05% SDS at increasingly elevated temperatures according to the level of stringency required. After washing, the membrane was wrapped in Saran wrap and autoradiographed using Fuji RX film at −70°C with an intensifying screen.

2.21.2 Estimation of oligonucleotide melting temperatures for hybridisation to Southern blots

The melting temperatures (T_m) of oligonucleotides in hybridisation solutions were calculated as described by Sambrook et al. (1989), using the following formula:

$T_m = 81.5°C + 16.6(\log_{10}[Na^+]) + 0.41(\text{fraction G+C of oligo}) - 0.63(\%$ formamide in hybridisation solution) $- (600/l)$

where l = length of the hybrid in base pairs.

2.22 ^{32}P Labelling of DNA

2.22.1 Random primed labelling of DNA

Random primed labelling of DNA was performed using the USB Random Primed DNA Labelling Kit (USB, Cleveland, USA). DNA to be labelled was purified by digestion with restriction endonucleases (section 2.4) followed by agarose gel electrophoresis (section 2.10). The purified DNA denatured by heating to 95°C for 10 minutes followed by chilling on ice. The labeling reaction was set up as follows:

Denatured DNA (25 ng)+ water	9 μl
dATP, dGTP, dTTP mixture (0.5 mM)	3 μl
Reaction mixture (random hexanucleotides in 10 x buffer)	2 μl
[α-^{32}P] dCTP	5 μl (50 μCi)
Klenow DNA polymerase enzyme	1 μl (2 units)

The reaction was incubated for 30 minutes at 37°C and was terminated by the addition of 2 μl 0.2 M EDTA, pH 8.0.

Alternatively, the agarose gel fragment containg the DNA to be labelled could be used directly in the labelling reaction (section 2.9). The gel slice was diluted in 0.33 volumes of water and boiled for 7 minutes, after which the sample was kept at 37°C until required. A total reaction volume of 50 μl was used and the volumes of the dNTPs and the reaction mixture were scaled up accordingly.

[α-^{32}P] dCTP (3000 Ci/mmol) was obtained from Amersham, UK. Unincorporated nucleotides were removed using a Chroma Spin+TE-10 column (section 2.22.3). Finally the probe was denatured by boiling for 10 minutes and added directly to the bag containing the membrane and hybridisation solution.

2.22.2 5´-End labelling of DNA using T4 polynucleotide kinase (T4 PNK)

Oligonucleotide probes were labelled using T4 PNK (Promega Corp. Madison, USA). Labelling reactions were set up as follows:

Oligonucleotide in distilled water	8-10 pmoles
	(approx 100 ng 32mer)
10x Polynucleotide Kinase Buffer	1.0 µl
[γ-^{32}P] ATP (3000Ci/mmol)	5.0 µl (50 µCi)
Sterile distilled water	vol to 9 µl
T4 Polynucleotide kinase	1.0 µl (8 units)

The reaction was incubated at 37°C for 45 minutes and terminated by the addition of 2 µl 0.5 M EDTA. [γ.-^{32}P] ATP was supplied by Amersham, UK. Unincorporated nucleotides were removed using a Chroma Spin+TE-10 column (section 2.22.3) and the probe added directly to the bag containing the membrane and hybridising solution.

2.22.3 Removal of unincorporated nucleotides

Unincorporated nucleotides were removed using a Chroma Spin+TE-10 column (Clontech Laboratories Inc, USA). TE buffer was removed from the column by centrifugation at 2000 r.p.m in a Beckman TJ-6 benchtop centrifuge for 5 minutes. The sample was applied to the top of the column and then centrifuged again at 2000 r.p.m for 5 minutes. The sample when spun through the column into a microcentrifuge tube was free from unincorporated nucleotides. The level of incorporation of radiolabel into the sample was determined by spotting 1 µl of the sample onto a strip of Whatman 3MM paper before and after the removal of unincorporated nucleotides and then by measuring the counts in a Scintillation counter. Typical specific activities of probes were 10^9 c.p.m/µg DNA for random priming and 10^8 c.p.m/µg DNA for end-labelled oligonucleotides.

2.23 Polymerase chain reactions

The polymerase chain reaction (PCR) was used to try to amplify from genomic DNA using degenerate oligonucleotide primers and for site-directed mutagenesis.

2.23.1 Amplification of genomic DNA using degenerate oligonucleotide primers.

Each 100 µl reaction volume contained: approximately 1 µl target genomic DNA, 1 µg of each primer, all four dNTPs at a final concentration of 0.2 mM, 10 µl 10 x *Taq* polymerase buffer, 6 µl 25 mM magnesium chloride, 2.5 units (0.5 µl) *Taq* DNA polymerase (Promega corp. USA) and sterile distilled water up to a final volume of 100 µl. The reaction was overlaid with an equal volume of sterile paraffin oil.

Reaction mixtures containing only one primer, no target DNA and no enzyme were used as controls. The reaction profile was set up as follows using a Perkin Elmer Cetus DNA Thermal Cycler:

> 3 minutes at 94°C (melting)
>
> Cycle x 30:
>
>> 1.5 minutes at 94°C (melting)
>>
>> 2.0 minutes at 45-60°C (annealing)
>>
>> 3.0 minutes at 72°C (polymerisation).
>>
>> 7.0 minutes at 72°C (final polymerisation).

The exact annealing temperature used was dependent upon the calculated melting temperature of the primers and the desired stringency of the polymerisation. Melting temperatures of primers were calculated using the Biosoft GeneJockey sequence processor (Taylor, 1991). The exact times of the various cycle parameters were also varied in order to optimise PCR conditions.

On completion of the cycles the samples were cooled immediately to 4°C and 10 µl visualised on a 2% agarose/TBE gel (section 2.9).

2.23.2 Site-directed mutagenesis

When using PCR to amplify genes that would later be used to express protein, *Vent* DNA polymerase (New England Biolabs, Beverly, USA) was used as this enzyme has $3' \rightarrow 5'$ exonuclease (proofreading) activity (Ling *et al.*, 1991).

Site-directed mutagenesis was performed using the method described by Higuchi *et al.* (1988). Oligonucleotide primers were designed complementary to each strand and incorporating the required mutagenic changes, as were "outer" primers to amplify from each end of the gene. These "outer" primers incorporated *Nde* I and *Bgl* II restriction sites to aid future cloning. For the first round of mutagenic PCR, each 100 µl reaction contained: 100 ng plasmid template, all four dNTPs to a final total concentration of 1.2 mM (300 µM of each), 2-6 mM MgSO₄, 10 µl of 10 x *Vent* DNA polymerase buffer, 1 µg of each primer, 1 µl *Vent* DNA polymerase (2 units)

and sterile distilled water up to 100 μl. Plasmid DNA template was restricted prior to use in PCR and the required amount of this reaction mix was added to the PCR reaction. Reactions were overlayed with an equal volume of liquid paraffin.

Reactions containing (i) no template and (ii) both "outer" primers but no mutagenic primers, were set up as controls. The reaction profile was set up as follows using a Perkin Elmer Cetus DNA Thermal Cycler:

<div align="center">

3 minutes at 94°C (melting)

Cycles x30

1.0 minute at 94°C (melting)

1.0 minute at 55°C (annealing)

1.5 minutes at 72°C (polymerisation)

</div>

On completion of the cycles, the samples were cooled immediately to 4°C and 20 μl run on an agarose gel (section 2.9). If the reaction was successful, 40 μl of each reaction was run on a 1% low melting point agarose (section 2.10) and the required band cut out and purified by phenol/chloroform extraction (section 2.11).

For the second round of mutagenic PCR, each 100 μl reaction contained 100 ng of both products from the first round of PCR and both of the "outer" primers. All of the other components were the same as in the first round of PCR. In this round of PCR the annealing temperature was lowered to 50°C and only 20 cycles were performed. On completion of the cycles the reactions were cooled to 4°C and 20 μl run on an agarose gel (section 2.9). The mutagenic PCR product was now cleaned up using the Promega Wizard PCR clean up system (section 2.23.3).

2.23.3 Promega Wizard PCR clean up system

Direct clean up of PCR products was performed using the Promega Wizard DNA clean up system (Promega Corporation, Madison, USA) exactly as described in the manufacturer's instructions.

2.24 Growth of *E. coli* cells expressing benzyl alcohol dehydrogenase (and benzaldehyde dehydrogenase)

E. coli cells transformed with plasmids harbouring the *xylB* (and *xylC*) gene were grown in LB supplemented with the appropriate antibiotic at 30°C. Cells transformed with pDG30, pDG40, pDG50 and pDG60 required a larger inoculum than those transformed with pDG12 and pDG20; 8% (v/v) as opposed to 2%.

Cells were grown to an optical density of approximately 0.5 at 600 nm (approximately 4 hours; LKB Ultraspec II) and were then induced with 0.5 mM IPTG (stock solution 100 mM). For benzyl alcohol dehydrogenase preparations the incubation was continued for a further 4 hours. Large culture volumes (>1 litre) were grown on stirring magnet mixers with a steady flow of filter sterilised air passed into the flask and were harvested at 6,000 r.p.m for 15 minutes at 4°C in a MSE Mistral 6L centrifuge. Small culture volumes (<1 litre) were grown on shaking platforms and 1 ml volumes were harvested in a microfuge for 5 minutes.

2.25 Ultrasonic disruption of bacteria

Pelleted cells (1 ml) were resuspended in 500 ml 25 mM Tris/HCl, pH 8.0, 2 mM DTT in a microcentrifuge tube. The tube was placed in an ice-water slurry and sonicated at 60 W for 6 x 10s periods with 20 s intervals for cooling, using the micro-tip of a Dawe Soniprobe apparatus (type 1130A, Dawe Instruments Ltd, London, UK). The crude extract was then spun in a microfuge for 5 minutes and the supernatants kept on ice until required for assays.

2.26 Enzyme assays

Spectrophotometric assays were performed at 27°C in a LKB Ultrospec II 4050 spectrophotometer. Plastic 3 ml or 1.5 ml cuvettes with a 1cm light path were used. For kinetic studies, enzyme, substrates and inhibitors were added using Hamilton glass syringes. One unit of enzyme activity is defined as 1 μmol substrate converted per minute. Specific activities are given as units/mg protein. The absorption coefficient of NADH at 340nm is 6.3×10^3 M^{-1} cm^{-1} (Boehringer, 1976).

2.26.1 Benzyl alcohol dehydrogenase (benzyl alcohol oxidation)

Benzyl alcohol dehydrogenase was assayed according to the method of MacKintosh & Fewson (1988a). The reaction was measured in a 3.0 ml reaction mixture containing 2.0 ml 150 mM-Bicine/540 mM-hydrazine, pH adjusted to 9.2 with NaOH (assay concentration 100 mM-Bicine/ 360 mM-hydrazine), 0.1 ml 60 mM-NAD$^+$, pH 7.0 (assay concentration 2 mM), distilled water, enzyme, and the reaction was initiated with 0.1 ml 6 mM benzyl alcohol (assay concentration 200 μM). The rate of reduction of NAD$^+$ was followed at 340 nm.

2.26.2 Benzyl alcohol dehydrogenase (oxidation of various alcohols)

For the majority of alcohols benzyl alcohol dehydrogenase activity was assayed as described in the previous section 2.26.1 except that the various alcohols replaced benzyl alcohol. The ability of benzyl alcohol dehydrogenase to oxidise coniferyl alcohol was followed by monitoring the production of coniferaldehyde at 400nm. Coniferaldehyde absorbs strongly at 400nm with an absorption coefficient of 22.4×10^3 M^{-1} cm^{-1} at pH 9.2 (Jaeger *et al.*, 1981). In this case hydrazine was omitted from the reaction buffer.

2.26.3 Benzaldehyde dehydrogenase II (benzaldehyde oxidation)

Benzaldehyde dehydrogenase II was assayed according to the method of MacKintosh & Fewson (1988a). The reaction was measured in a 3.0 ml reaction mixture containing 2.0 ml 150 mM-Bicine/NaOH buffer, pH 9.5 (assay concentration 100 mM), 0.1 ml 60 mM NAD$^+$, pH 7.0 (assay concentration 2 mM), distilled water, enzyme, and the reaction was initiated by the addition of 0.1 ml 0.3 mM benzaldehyde (0.01 mM assay concentration). The rate of reduction of NAD$^+$ was followed at 340 nm.

2.26.4 Horse liver and baker's yeast alcohol dehydrogenases

Horse liver and baker's yeast alcohol dehydrogenase were assayed as described by Plapp (1970). Stock solutions of horse liver alcohol dehydrogenase were made by dissolving crystalline enzyme in 10 mM Na$_4$P$_2$O$_7$, 20 mM glycine, pH 9.0 at a concentration of 10 mg protein/ml (15 units/ml). Trace ethanol was removed from this solution by two passages down a Sephadex G-25 gel filtration column equilibrated with 10 mM Na$_4$P$_2$O$_7$, 20 mM glycine, pH 9.0. Fractions (1 ml) were collected and peak fractions, as judged by A$_{280}$, were pooled. Samples of horse liver alcohol dehydrogenase for assay were diluted as necessary in 1 mg/ml bovine serum albumen (dissolved in 10 mM Na$_4$P$_2$O$_7$, 20 mM glycine). Horse liver alcohol dehydrogenase concentrations were determined by UV absorption assuming that A$_{280}$ = 0.455/cm for 1 mg/ml protein (Hennecke & Plapp, 1983).

Stock solutions of baker's yeast alcohol dehydrogenase were made by dissolving crystalline enzyme in distilled water at a concentration of 10 mg protein/ml (3700 units/ml). Samples for assay were diluted in 50 mM KH$_2$PO$_4$, pH 7.5.

For both horse liver alcohol dehydrogenase and baker's yeast alcohol dehydrogenase the reaction was measured in a 1 ml reaction mixture containing 500 µl reaction buffer {170 mM $Na_4P_2O_7$ (assay concentration 85 mM), 13 mM semicarbazide HCl (assay concentration 6.5 mM), 36 mM glycine (assay concentration 18 mM), pH 9.0}, 100 µl 17.5 mM NAD^+ (assay concentration 1.75 mM), distilled water, enzyme, and the reaction was initiated by the addition of 100 µl 5.5 M ethanol (assay concentration 550 mM). The rate of reduction of NAD^+ was followed at 340 nm and 25°C.

2.26.5 Standardisation of substrate concentration

The concentrations of stock substrate solutions were determined as follows:

(a) **Alcohols.** The benzyl alcohol dehydrogenase or horse liver alcohol dehydrogenase assay described previously (section 2.26.1) was used with an excess of NAD^+ (2mM) but limiting the alcohol concentration. Assuming a 1:1 alcohol utilisation:NADH production then the complete A_{340} corresponds to the amount of alcohol initially added.

(b) **Aldehydes.** The benzaldehyde dehydrogenase II assay described previously (section 2.26.2) was used with an excess of NAD^+ (2mM) but limiting the aldehyde concentration. Assuming a 1:1 alcohol utilisation:NADH production then the complete A_{340} corresponds to the amount of aldehyde initially added.

2.26.6 Analysis of initial velocities and determination of kinetic coefficients

(a) **K'_m and V'_{max} determinations.** To determine K'_m and V'_{max} values for various substrates, initial velocities were measured over a range of concentrations of the substrate in the presence of fixed excess of the second substrate (NAD^+; 2mM). The resulting data were analysed using the Enzpack computer program (Williams, 1985). This program determines the kinetic coefficients by the Direct-Linear method (Eisenthal & Cornish-Bowden, 1974; Cornish-Bowden & Eisenthal, 1974) and also by the Lineweaver-Burk, Hanes-Woolf and Eadie-Hofstee linear methods. Values quoted in this thesis were determined by the Direct-Linear method, but these values were checked against the regression coefficients given for the three linear methods.

(b) Absorption coefficients and correction factors. Some aromatic substrates and the corresponding aldehydes absorb light at 340nm. In such cases it is necessary to make corrections to any assay results in order to ontain a true value of the amount of NADH produced. It has been shown previously that benzaldehyde, benzyl alcohol, 2-, 3- and 4-methylbenzyl alcohol, 2-, 3-, and 4-methylbenzaldehyde, cinnamyl alcohol, perillaldehyde and perillyl alcohol do not absorb significantly (i.e less than 1% of the ε_{340} value of NADH in benzyl alcohol dehydrogenase assay buffer) but that cinnamaldehyde has an absorption coefficient of 2.00×10^3 M^{-1} cm^{-1} (MacKintosh & Fewson, 1988b).

Measured changes in absorbance when assaying the oxidation of cinnamyl alcohol were therefore corrected using the correction factor calculated by using the following formula:

$$-\text{Correction factor} = \frac{\varepsilon \text{NADH}}{\varepsilon \text{NADH} - \varepsilon \text{substrate} + \varepsilon \text{product}}$$

2.27 Purification of benzyl alcohol dehydrogenase

Benzyl alcohol dehydrogenase (BADH) was purified from *E. coli* DH5α/pDG20 and *E. coli* JM109(DE3)/pDG30 as described by MacKintosh & Fewson (1988a) for the enzyme from *A. calcoaceticus* NCIB 8250. All steps were carried out at 0-4°C.

Step 1 Preparation of extracts

Cells (3g) of *E. coli* DH5α/pDG20 or *E. coli* JM109 (DE3)/pDG30 (or mutants) grown as described in section 2.24 were resuspended in 15-45 ml buffer A (75 mM-potassium phosphate buffer, pH 7.5, 2 mM DTT) and the cells disrupted by 4 passages through the French pressure cell (4-3398A; American Instruments Company, Silver Spring, Maryland, USA) at a pressure of 98 MPa (14,300 $lb.in^{-2}$). If necessary the homogenate was diluted to 45 ml and centrifuged at 40,000 r.p.m (64,000 g) in a Beckman L5-65 ultracentrifuge (T.70 rotor). The supernatant fraction was divided into three and stored temporarily on ice or for longer at –20°C.

Step 2 Chromatography on DEAE-Sephacel

One third (15 ml) of the extract prepared in step 1 was applied to a DEAE-Sephacel column (3.8 cm x 2.6 cm for small scale preps; 12 cm x 5 cm for large scale preps) pre-equilibrated in Buffer A. The column was washed in buffer A for 1 hour (2-3 volumes) and then BADH eluted in Buffer B (110 mM-potassium phosphate buffer, pH 7.5, 2 mM DTT). For large scale preps the fow was reversed for elution. The flow rate throughout was 50 ml/h. The fractions containing BADH activity were pooled (approx. 150 ml) and dialysed overnight against 3 x 1000 ml Buffer C (10 mM-potassium phosphate buffer, pH 6.0, 2 mM DTT, 5 mM $MgCl_2$).

Step 3 Chromatography on Blue Sepharose CL-6B

The dialysed pool (approximately 160 ml) was applied to a Blue Sepharose CL-6B column (9.5 cm x 2.6 cm) pre-equilibrated with Buffer C. Once loading was complete the pump was stopped for 30 minutes and then the column was washed with Buffer C for 2 hours (3-4 volumes). The buffer was then changed to Buffer C containing 0.12 mM NAD^+ and the flow direction reversed to elute BADH. The flow rate throughout was 90 ml/h. The fractions containing BADH activity were pooled (approximately 125 ml), concentrated down to 5-10 ml using carboxymethylcellulose and further concentrated down to 0.5 ml using a Centricon-30 microconcentrator (Amicon Ltd., Stonehouse, UK).

BADH preparations to be used for diethylpyrocarbonate inactivation studies (section 2.29) needed to have any bound NAD^+ removed and so the Blue Sepharose CL-6B pool was dialysed overnight against 3 x 1000 ml buffer D (50 mM-potassium phosphate buffer, pH 7.5, 2 mM DTT) before concentrating as described above.

2.28 Protein estimations

Protein concentrations were determined by the method of Lowry *et al.* (1951), with bovine serum albumin as a standard.

2.29 SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE was carried out using the discontinuous tris/glycine buffer system described by Laemmli (1970) and all gels were prepared, run, fixed and stained as described by Sambrook *et al.* (1989). Resolving gels were 10% (w/v)

polyacrylamide with stacking gels 5% (w/v) polyacrylamide. Protein samples were mixed with sample buffer (Sambrook *et al.*, 1989) and DTT (final concentration 50 mM) and were boiled for 5 minutes prior to loading.

Molecular weight markers used for SDS-PAGE were the Pharmacia LMW Electrophoresis Calibration kit (Pharmacia Ltd, Milton Keynes, UK) containing phosphorylase *b* (subunit MW 94,000), bovine serum albumin (67,000), ovalbumin (43,000), carbonic anhydrase (30,000), soyabean trypsin inhibitor (20,100) and α-lactalbumin (14,400).

To determine the level of expression of benzyl alcohol dehydrogenase (% total soluble protein), stained SDS-polyacrylamide gels were scanned using an LKB 2202 Ultrascan Laser Densitometer. The level of expression was calculated automatically by the scanner.

2.30 Protein sequencing

The *N*-terminal sequence of recombinant BADH purified from *E.coli* (section 2.27) was determined by Dr J.N. Keen, Department of Biochemistry and Molecular Biology, University of Leeds, UK.

A sample of BADH was electroblotted from a 15% polyacrylamide gel, the polypeptide band excised from the blot and washed with 0.1% (v/v) trimethylamine in 50% (v/v) aqueous methanol before placing in the Blott™ reaction cartridge of the sequencer. The polypeptide was subjected to automated liquid-pulse Edman degradation on an Applied Biosystems 477A sequencer with on-line 420A narrow-bore HPLC and 610A data analysis system according to manufacturer's instructions.

2.31 Electrospray mass spectrometry

Electrospray mass spectrometry was performed with the help of Mr T. Krell, Division of Biochemistry & Molecular Biology, Institute of Biomedical and Life Sciences, University of Glasgow on a VG platform quadropole mass spectrometer fitted with a pneumatically assisted electrospray source and controlled using the VG Mass-Lynx software (VG Biotech. Ltd., Altrincham, Cheshire, UK). The infusion of carrier solvent (1:1 (v/v) acetonitrile:water) was controlled at a flow rate of 10 μl/min using a Harvard syringe pump (Harvard Apparatus, South Natic, MA, USA). Capillary voltages were between 2.8 and 3.2 kV, extraction cone voltages 20-30 kV and the focusing cone voltage offset by +10 V. Source temperature was set at 65°C and the nebulising gas flow at 10 l/h. Lens stack voltages

were adjusted to give maximum ion currents. Before use the instrument was calibrated with horse heart myoglobin (Sigma).

Protein samples had to be extensively desalted prior to injection into the electrospray mass spectrometer. This was done by centrifugation in a Centricon-30 (Amicon, Stonehouse, UK) microconcentrator using HPLC-grade water. Samples were then diluted in an equal volume of 4 % (w/v) formic acid in acetonitrile and 10-20 µl aliquots were injected directly into the carrier stream.

2.32 Modification of histidine residues with diethyl pyrocarbonate

Diethyl pyrocarbonate was used to selectively modify histidyl residues in benzyl alcohol dehydrogenase as described by Hennecke & Plapp (1983) for horse liver alcohol dehydrogenase. Diethyl pyrocarbonate reacts with histidine residues to form the N -carbethoxyhistidine derivative (section 6.1), the production of which can be followed spectrophotometrically because it absorbs strongly between 230-250 nm (Miles, 1977). The number of modified histidine residues can be calculated from the increase in absorbance at 242 nm between carbethoxylated and unmodified enzyme since the absorption coefficient of N -carbethoxyhistidine at 242 nm is 3200 M^{-1} cm^{-1} (Ovádi et $al.$, 1967). Any histidine modifications can then be correlated against a loss of enzyme activity in order to assess the involvement of any histidine residues in enzyme catalysis.

2.32.1 Preparation of stock solutions of diethyl pyrocarbonate

Diethyl pyrocarbonate was freshly diluted with anhydrous ethanol for experiments with benzyl alcohol dehydrogenase or with anhydrous acetonitrile for experiments with horse liver alcohol dehydrogenase. Approximately 1.5 µl of diethyl pyrocarbonate was diluted into 1 ml of ethanol and the concentration of the stock solution was determined by reaction with imidazole (Miles, 1977). Diluted diethyl pyrocarbonate (10 µl) was added to 3 ml 10 mM imidazole, pH 7.5 in a 1 cm quartz cuvette and the complete increase in absorbance at 230 nm due to the formation of N -carbethoxyimidazole was recorded. This correlates to the concentration of the diethyl pyrocarbonate added assuming that the reaction has come to completion and that the absorption coefficient of N -carbethoxyimidazole is 3.0×10^3 M^{-1}. cm^{-1} (Melchior & Fahrney, 1970).

2.32.2 The carbethoxylation reaction

Purified benzyl alcohol dehydrogenase (section 2.27) was diluted to 1 mg/ml with 100 mM potassium phosphate buffer, pH 6.0. Stock diethyl pyrocarbonate was added to a 1 ml aliquot of 1 mg/ml benzyl alcohol dehydrogenase in a 1.5 ml quartz cuvette to give a final concentration of 1 mM. The cuvette was incubated at 25°C and the increase in absorbance at 242 nm monitored for the production of N -carbethoxyhistidine in a Philips PU 8700 UV/VIS scanning spectrophotometer. Control samples were run containing an equivalent volume of ethanol in benzyl alcohol dehydrogenase experiments, or acetonitrile in horse liver alcohol dehydrogenase experiments replacing diethyl pyrocarbonate.

At various time intervals 5 μl aliquots were removed from the cuvette, diluted 1:100 with 50 mM potassium phosphate buffer, pH 7.5 and then 5 μl aliquots of this were immediately assayed for benzyl alcohol dehydrogenase activity as described previously (section 2.26.1).

Substrate protection experiments were carried out as described above except that substrate was incubated with the enzyme for 5 minutes prior to the addition of diethyl pyrocarbonate.

The carbethoxylation of horse liver alcohol dehydrogenase was used as a positive control experiment. The enzyme was acetimidylated prior to modification with diethyl pyrocarbonate by dissolving 30 mg crystalline enzyme in 3 ml 0.5 M triethanolamine hydrochloride buffer, pH 8.0 and at 25°C adding 4 x 0.05 volumes of 2.1 M ethylacetimidate hydrochloride at 60 minute intervals. The enzyme was then passed down a Sephadex G-25 gel filtration column equilibrated with 33 mM NaH_2PO_4, pH 8.0, 0.25 mM EDTA and the pooled fractions were dialysed against 5 x 250 ml 33 mM NaH_2PO_4, pH 8.0, 0.25 mM EDTA.

Stock diethyl pyrocarbonate was added to a 1 ml aliquot of 1 mg/ml horse liver alcohol dehydrogenase in a 1.5 ml quartz cuvette to give a final concentration of 1 mM. The cuvette was incubated at 25°C and the increase in absorbance at 242 nm monitored for the production of N -carbethoxyhistidine in a Philips PU 8700 UV/VIS scanning spectrophotometer. At various time intervals 5 μl aliquots were removed from the cuvette, diluted 1:10 with 46 mM NaH_2PO_4, pH 8.0, 0.25 mM EDTA, 1 mg/ml BSA, and then 5 μl aliquots of this were immediately assayed for alcohol dehydrogenase activity as described previously (section 2.26.4).

2.32.3 Decarbethoxylation of modified benzyl alcohol dehydrogenase

The modification of histidine residues with diethyl pyrocarbonate can be reversed by hydrolysis with hydroxylamine (Miles, 1977). Aliquots of carbethoxylated benzyl alcohol dehydrogenase (100 µl) were diluted two-fold in 1 M hydroxylamine (final concentration 0.5 M) and incubated at 27°C for 23 hours. Aliquots of this decarboxylated enzyme were then diluted 1:50 with 50 mM potassium phosphate buffer, pH 7.5 and 5 µl assayed for benzyl alcohol dehydrogenase activity as described previuosly (section 2.26.1) to determine whether decardethoxylation had occured.

Chapter 3. Cloning and sequencing of *xylB* (the gene encoding benzyl alcohol dehydrogenase) from *A. calcoaceticus* NCIB 8250

3.1 Subcloning *xylB* from a lambda genomic DNA library of *A. calcoaceticus* NCIB 8250 and the creation of benzyl alcohol dehydrogenase expression clones

Previously in this laboratory a lambda (λ) genomic DNA library from *A. calcoaceticus* NCIB 8250 had been made Dr A. G. S. Robertson (unpublished results). The library was made by partially digesting *A. calcoaceticus* NCIB 8250 genomic DNA with *Sau* 3A1 and ligating fragments in the 14-20 kb range into the vector λGEM-11 (Promega Corp.) digested with *Bam* HI. Probes had also been made to screen this library for the genes encoding the enzymes benzaldehyde dehydrogenase II (BZDH II; gene nomenclature, *xylC*) and benzyl alcohol dehydrogenase (BADH; *xylB*) from the benzyl alcohol pathway (section 1.7; Fig. 1.5) by using the polymerase chain reaction (PCR) to amplify the *N*-termini encoding regions of these genes using degenerate oligonucleotide primers designed using the *N*-terminal amino acid sequences determined from the purified enzymes (Chalmers *et al.*, 1991). By this method, 120 bp from *xylC* and 96 bp from *xylB* were amplified and these DNA fragments used to identify various clones in the λ chromosomal library harbouring these genes. Restriction mapping using Southern blots probed with these gene fragments suggested that *xylB* and *xylC* are contiguous on the chromosome (A. G. S. Robertson, unpublished results; Reid, 1993).

The clone λAR201 gave positive signals when plaque lifts were hybridised with both the *xylC* and *xylB* probes. *XylC* was further sub-cloned and the DNA sequence fully determined. The derived amino acid sequence for BZDH II is very similar to that of other (chiefly eukaryotic) aldehyde dehydrogenases (A. G. S. Robertson, unpublished results; section 1.12). This sequence, together with that of the start of the *xylB* gene, suggested that the two genes are not just contiguous on the chromosome but in fact are part of an operon transcribed in the direction *xylC-xylB*. However, the λAR201 clone contained only the first 429 bp of *xylB*. A longer clone, λAR1101 (Fig. 3.1), had also shown positive signals when plaque lifts were hybridised with both of the *xylC* and *xylB* probes (Reid, 1993) and it was from this starting point that the present work began.

Fig. 3.1 shows the relationship between the *A. calcoaceticus* NCIB 8250 genomic DNA map and the various sub-clones that were constructed. A 4.4 kb *EcoRI-XhoI* restriction fragment (section 2.4), which according to the genomic restriction map should harbour both *xylC* and *xylB*, was sub-cloned from λAR1101 into the general cloning vector pBluescript SK II- (section 2.2) creating the 7.36 kb plasmid pDG12. This plasmid was shown to express both

Figure 3.1 Construction of *xylB* and *xylC* subclones

Genomic map

λAR201

λAR1101

pDG12

pDG20

pDG30

2 kb

xylC (BZDH II)

xylB (BADH)
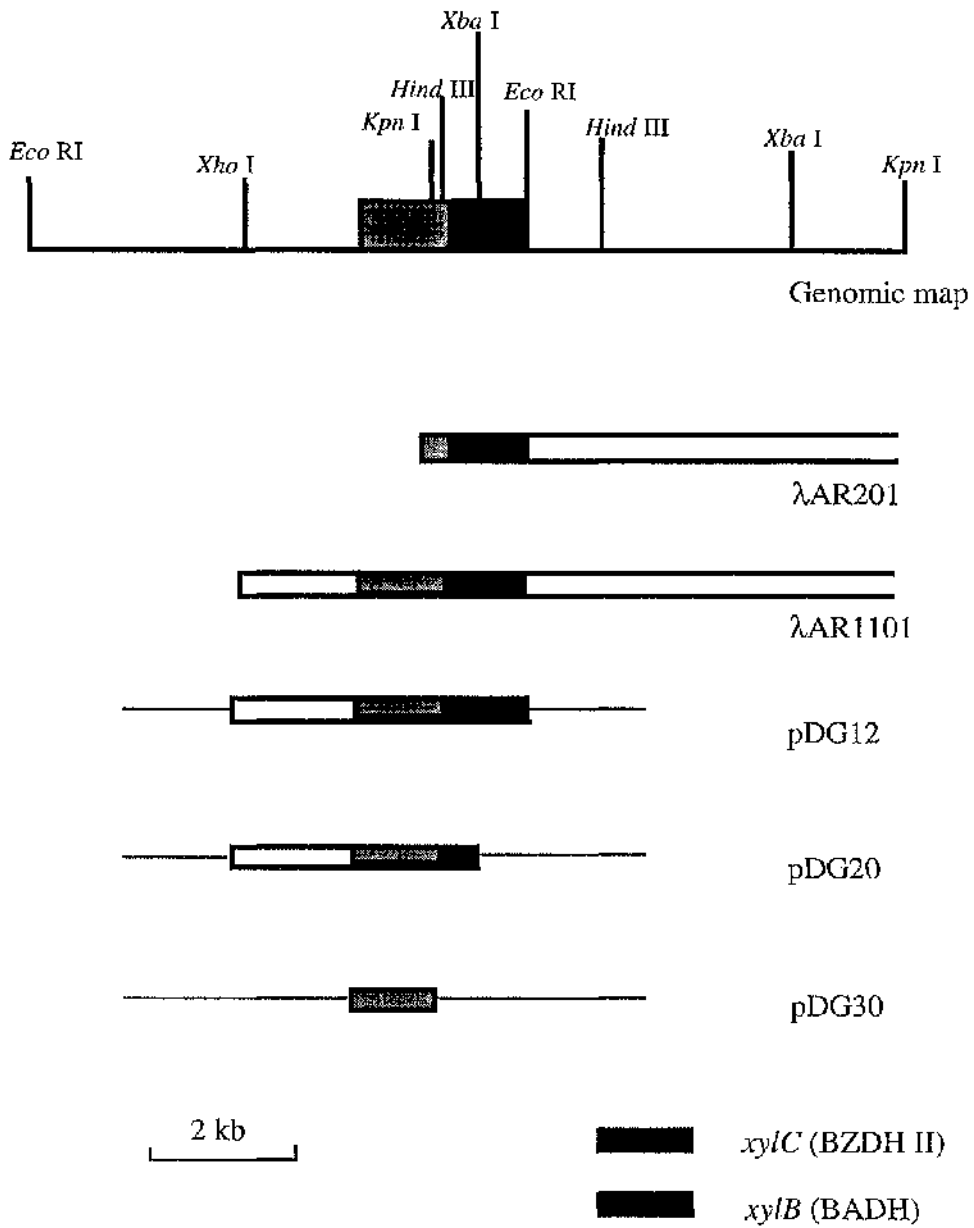
Oligonucleotides used to create pDG30:

5´-GAATTC<u>CATATG</u>AGTGAATTAAAAGATATT-3´
         *Nde*I   (oligo 2166)

5´-GA<u>AGATCT</u>AAGATTATGCAATTTTAATAATCG-3´
    *Bgl*II   (oligo 2168)

benzaldehyde dehydrogenase II and benzyl alcohol dehydrogenase when expressed in *E. coli* DH5α as shown by enzyme assays of crude extracts (sections 2.26 & 4.1). To aid the purification of benzyl alcohol dehydrogenase, a further sub-clone was constructed from pDG12 in order to express benzyl alcohol dehydrogenase in the absence of benzaldehyde dehydrogenase II. The 1.1 kb *XbaI* restriction fragment between the *XbaI* restriction site in the *xylC* gene (Fig. 3.1) and the *XbaI* site in the multiple cloning sequence of the pBluescript SK II- vector was removed (section 2.4) and the remaining plasmid re-ligated (section 2.5) to create the plasmid pDG20. This clone harboured *xylB*, truncated *xylC* (the first 746 bp were removed), and a further 1.47 kb downstream of *xylB*. This clone was able to express benzyl alcohol dehydrogenase in *E. coli* DH5α and the recombinant benzyl alcohol dehydrogenase could be purified (sections 2.27 & 4.2).

Once the nucleotide sequence of *xylB* had been determined (section 3.2) the clone pDG30 was generated with a T7 expression vector using the polymerase chain reaction (section 2.23). This was to create a high over-expression system for benzyl alcohol dehydrogenase (section 4.1) and also to create a clone containing *xylB* with no further up- or down-stream sequence in order to facilitate site-directed mutagenesis studies (sections 2.23.2, 5.1 & 7.1). Two oligonucleotides were designed based on the nucleotide sequence at the ends of *xylB* to be used as PCR primers (Oligos 2166 & 2168; Fig. 3.1). *NdeI* or *BglII* restriction sites were added to the primers' respective 5'-ends (Fig. 3.1) to facilitate cloning of the PCR product into the T7 expression vector pTB361 (section 2.2; Fig 3.2). This 4.45 kb plasmid carries the bacteriophage T7 promoter upstream from its multiple cloning site. The *NdeI* restriction site includes an ATG translational start codon such that when inserts are ligated into the *NdeI* restriction site in the vector they are located the correct distance from the vector's Shine-Dalgarno sequence to enable expression. Expression is achieved by transforming an *E. coli* strain that has had the gene encoding T7 RNA polymerase inserted into its chromosome under *lac* control (DE3 strains) with the expression vector. This means that when IPTG is added to the growth medium, T7 RNA polymerase is produced which then transcribes the DNA ligated downstream of the T7 promoter in the expression vector. By this means the availability of RNA polymerase is not a limiting factor on expression and so greater protein yields can be obtained (Studier & Moffat, 1986).

To create pDG30, PCR was performed using *Vent* DNA polymerase to amplify *xylB* using oligos 2166 and 2168 with pDG20/*XbaI* as template (section 2.23.1). *Vent* polymerase was used for this amplification rather than *Taq* polymerase, which had been used to create probes for these genes, since this

Figure 3.2 Restriction map of the T7 expression vector pTB361

Plasmid (section 2.2) and map were a gift from M. J. Horsburgh, Division of Biochemistry and Molecular Biology, Institute of Biomedical and Life Sciences, University of Glasgow.

4453 / EcoRI
XbaI
NdeI < KpnI < BgIII
NaiI < XhoI
BcaI < XbaI < SalI
SphI < BamHI
500 < ClaI
oriV
1000
oriV
1500
bla.part
PatI < 2000 < SalI
BclI
2500
tetR
3000
AatII
SalI
EcoRV
3500
SacII
tetA
BclI
PvuII
4000
StuI
StyI
cer

pTB361
4453
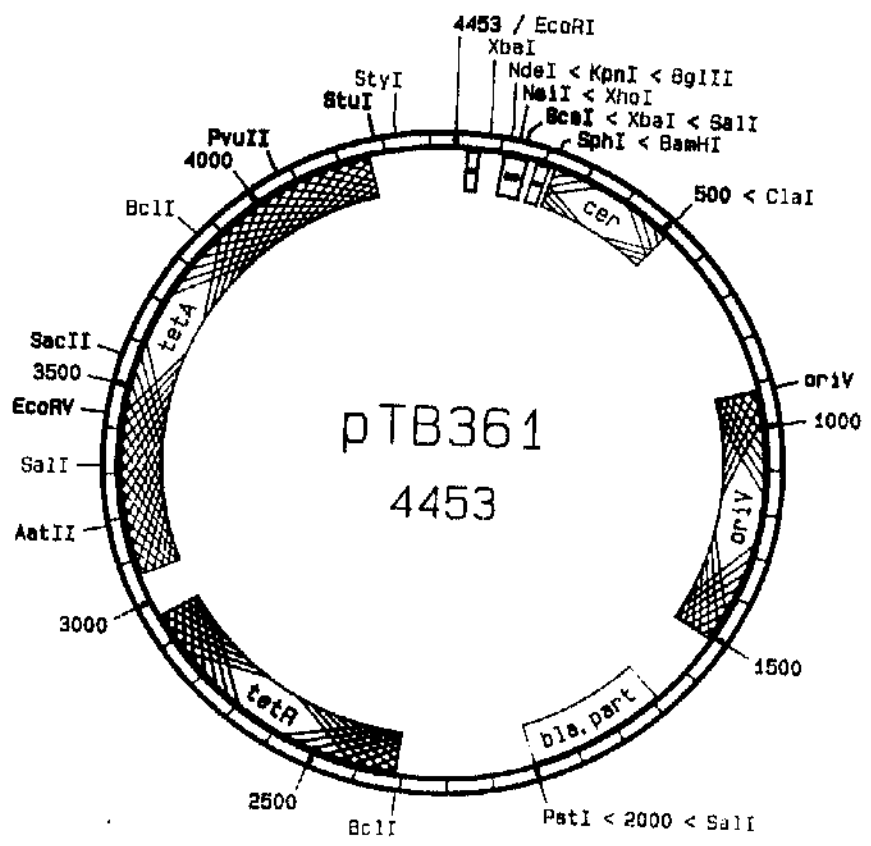
enzyme has 3′→5′exonuclease ("proof-reading") activity (Ling *et al..*, 1991). This means that fewer PCR errors that might alter the expressed protein would be generated. *XylB* generated by PCR with *Vent* polymerase was digested with *NdeI* and *BglII* (section 2.4) and ligated into pTB361 (section 2.5) to create pDG30. This clone was sequenced with the primers used to sequence pDG20 (sections 2.19 & 3.2). In two separate clones generated, no PCR errors had been introduced.

3.2 Strategy for sequencing *xyl B* in the clone pDG20

The nucleotide sequence of *xylB* was determined in the construct pDG20 by primer walking (Fig. 3.3) as described in section 2.19. Initially two oligonucleotides, 1965 and 1966, were designed on the basis of the sequence previously determined for *xylC* and the start of *xylB* (A. G. S. Robertson, unpublished results; section 1.12). Oligo 1965 was an 18-mer based on nucleotides 45 bp from the end of *xylC* and oligo 1966 was based on nucleotides 219 bp into the *xylB* gene. The sequence determined from these primers was found to be identical to that determined previously from the clone λAR201.

Typical sequencing gels produced 250-300 bp of new sequence from each primer. Wedged (0.4-1.2 mm) sequencing gels were found to be effective at increasing the length of sequence obtained since non-wedge gels would routinely produce only 200 bp of sequence from each primer (2.19.4). The addition of manganese to the labelling reactions (section 2.19.2) was very effective in enabling the determination of sequence very close to the primer. Its addition enabled sequencing as close as 7 bp from the end of the primer.

Sequencing primers were designed every 200 bp of new sequence. This allowed an overlap of about 50 bp between the sequences determined by adjacent primers. *XylB* was sequenced on both strands, except for the region at the start that had been sequenced previously (section 3.1). This region had already been sequenced on both strands and so had now been sequenced on both strands and twice on one of the two strands. A further 203 bp were sequenced downstream from the end of *xylB* in order to check for any further genes in this putative operon or for any further open reading frames.

Figure 3.3 Strategy for sequencing *xylB* in vector pDG20

Arrows indicate the length and direction of sequence determined
using each oligonucleotide.

■■■	*xylC* (BZDH II)
▨▨▨	*xylB* (BADH)
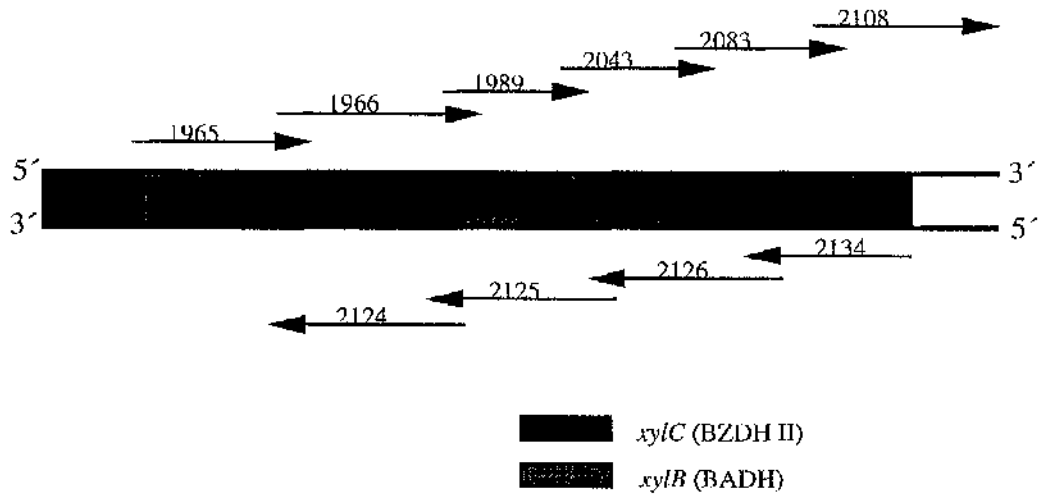
Sequences of the *xylB* sequencing oligonucleotides

Forward direction:

5´-TGGCAATGGATTACAGTA-3´	Oligo 1965
5´-GGCCCTAACGTAACCGAG-3´	Oligo 1966
5´-GAACTATTAGGACCATTA-3´	Oligo 1989
5´-ATTGGCAAAACAATTGGG-3´	Oligo 2043
5´-TGCACAATTTGATGTAAA-3´	Oligo 2088
5´-AAGGGTATTTGATGTAAA-3´	Oligo 2108

Reverse direction:

5´-AAAGCACTCAGACCAACA-3´	Oligo 2124
5´-GTTTTAATATCTCAGGTC-3´	Oligo 2125
5´-GTCAAAGGGAAATTTCCC-3´	Oligo 2126
5´-AGACTAAACCCAAAATTT-3´	Oligo 2134

3.3 The *xylB* nucleotide sequence and the derived amino acid sequence for benzyl alcohol dehydrogenase

Fig. 3.4 shows the complete nucleotide sequence for *xylB* and the derived amino acid sequence for benzyl alcohol dehydrogenase. The *xylB* structural gene is 1110 bp in length (excluding translational start and stop codons) and encodes a 370 amino acid protein with a deduced relative molecular mass (M_r) of 38,923. This value is in good agreement with that obtained for the purified benzyl alcohol dehydrogenase from *A. calcoaceticus* (39,700; MacKintosh & Fewson, 1988a). A putative Shine-Dalgarno sequence was identified upstream from the translational start codon and a possible σ factor-independent transcription terminator sequence was identified downstream of the translational stop codon using the Brendel & Trifonov method (1984) on the GCG UNIX system (Fig. 3.4). This would suggest that there are no more genes within this putative operon downstream of *xylB*. In addition, no further open reading frames could be found in the 200 bp that were sequenced downstream of *xylB*. In the work prior to this thesis (section 3.1) 326 bp were sequenced upstream of *xylC* in λAR201 (Fig. 3.1). Within these nucleotides no open reading frame could be found that could encode another protein. However, no putative promoter sequences could be identified in order to verify that *xylC* is indeed the start of an operon.

Table 3.1 shows the codon usage of the *xylB* structural gene and compares this to that of other *Acinetobacter* structural genes (White *et al.*, 1991). It can be seen that the codon usage in *xylB* is very similar to that of the other *Acinetobacter* genes. The codon usage table shows that acinetobacters have a definite preference for A or T in the wobble position, a fact that is reflected in the G+C content of the organism. *XylB* has a G+C content of 42% which is typical of *Acinetobacter* genes (section 1.11; Towner *et al.*, 1991). *XylC* has also been shown to have a codon usage pattern and a G+C content (44%) typical of acinetobacters (section 1.12; A. G. S. Robertson, unpublished results). These features suggest that the benzyl alcohol pathway genes are not a recent acquisition by *A. calcoaceticus* NCIB 8250.

The *xylB* structural gene from *A. calcoaceticus* has 60% sequence identity with *xylB* from the *P. putida* TOL-plasmid pWW0 as determined by the Bestfit program on the GCG/UNIX system. The TOL-plasmid *xylB* has a G+C content of 51%. This value is low since pseudomonads have an average G+C content of 58-69 moles % (Palleroni, 1975), and is reflected in the codon usage table for *xylB* from *P. putida* (Table 3.2) which shows that there is not the strong preference for G or C in the wobble position that might be expected in

Figure 3.4 Nucleotide sequence of *xylB* and the derived amino acid sequence of benzyl alcohol dehydrogenase

The nucleotide sequence of *xylB* (starting at nucleotide 1) is shown. The putative Shine-Dalgarno sequence is shown in bold print (nucleotides −15 to −11) and the putative transcription terminator sequence is underlined (1123 to 1173). Symbols correspond to: ‡ , *xylB* translational start codon (−3 to −1); #, *xylC* stop codon (−38 to −36); *, *xylB* stop codon (1111 to 1113). Single letter symbols for the derived amino acid sequence of benzyl alcohol dehydrogenase are also shown.

```
      -38                 -20                            1
       .         .         .         .         .         .
       taaacaattaaataaatcatgacaggacgttataccatgagtgaattaaa
       #                                          ‡  S  E  L  K
```

```
              20                  40                    60
               .         .         .         .         .
       agatattattgccgcagtgaccccatgtaaaggtgctgactttgagctgc
    5   D  I  I  A  A  V  T  P  C  K  G  A  D  F  E  L  Q
```

```
       0                   80                   100
               .         .         .         .         .
       aagctttaaaaatacgtcagccgcaaggtgatgaagtattggtaaaagtg
   22   A  L  K  I  R  Q  P  Q  G  D  E  V  L  V  K  V
```

```
             120                 140                   160
               .         .         .         .         .
       gttgctacaggaatgtgccacaccgatttgattgtacgtgatcaatatta
   38   V  A  T  G  M  C  H  T  D  L  I  V  R  D  Q  Y  Y
```

```
                         180                 200
               .         .         .         .         .
       tccggtaccgctgcctgccgtccttggacatgaagggtcaggaattattg
   55   P  V  P  L  P  A  V  L  G  H  E  G  S  G  I  I  E
```

```
             220                 240                   260
               .         .         .         .         .
       aggcgattggccctaacgtaaccgagcttcaagtcggcgatcatgtggtt
   72   A  I  G  P  N  V  T  E  L  Q  V  G  D  H  V  V
```

```
                         280                 300
               .         .         .         .         .
       ttaagctatggttattgtgggaaatgtacccaatgtaatactggtaatcc
   88   L  S  Y  G  Y  C  G  K  C  T  Q  C  N  T  G  N  P
```

```
             320                 340                   360
               .         .         .         .         .
       tgcctattgttcagagttttttggacgaaattttagtggggcagattcag
  105   A  Y  C  S  E  F  F  G  R  N  F  S  G  A  D  S  E
```

```
                    380                      400
                      .          .            .            .            .
       aagggaatcatgcgctttgtacccatgatcaaggcgtggtaaatgatcat
122     G   N   H   A   L   C   T   H   D   Q   G   V   V   N   D   H


                   420                      440                    460
                     .          .            .          .            .
       tttttgcccaatcctcctttgcaacctatgctttaagtcgtgaaaataa
138     F   F   A   Q   S   S   F   A   T   Y   A   L   S   R   E   N   N


                             480                      500
                               .          .            .          .
       cacggttaaagtgaccaaagacgtaccaattgaactattaggaccattag
155     T   V   K   V   T   K   D   V   P   I   E   L   L   G   P   L   G


                   520                      540                    560
                     .          .            .            .          .
       gttgtggtattcaaacaggtgctggcgcatgtataaatgctttaaaagtg
172      C   G   I   Q   T   G   A   G   A   C   I   N   A   L   K   V


                             580                      600
                               .          .            .          .
       acacctgccagtagtttggtgacttggggtgccggcgctgttggtctgag
188     T   P   A   S   S   L   V   T   W   G   A   G   A   V   G   L   S


                   620                      640                    660
                     .          .            .            .          .
       tgctttacttgcggcaaaagtgtgtggcgcttcgatcattatcgctgttg
205     A   L   L   A   A   K   V   C   G   A   S   I   I   I   A   V   D


                             680                      700
                               .          .            .          .
       atattgttgaatcacgtttggaattggcaaaacaattgggtgcaactcat
222      I   V   E   S   R   L   E   L   A   K   Q   L   G   A   T   H


                   720                      740                    760
                     .          .            .            .          .
       gtgatcaatagtaaaactcaagatcctgtggctgcgattaaagagattac
238     V   I   N   S   K   T   Q   D   P   V   A   A   I   K   E   I   T
```

```
                  780                    800
         .          .          .          .          .
     tgatggtggagtcaactttgcattggaatctactgggcgacctgagatat
255  D  G  G  V  N  F  A  L  E  S  T  G  R  P  E  I  L
```

```
                  820                    840                    860
         .          .          .          .          .
     taaaacaaggtgttgatgcgctgggtattttaggtaaaatcgccgtggta
272  K  Q  G  V  D  A  L  G  I  L  G  K  I  A  V  V
```

```
                  880                    900
         .          .          .          .          .
     ggtgcacctcaattaggcacaactgcacaatttgatgtaaatgatctgtt
288  G  A  P  Q  L  G  T  T  A  Q  F  D  V  N  D  L  L
```

```
                  920                    940                    960
         .          .          .          .          .
     attgggtgggaaaacgattttaggtgtcgttgaggggagtggttcgccca
305  L  G  G  K  T  I  L  G  V  V  E  G  S  G  S  P  K
```

```
                  980                    1000
         .          .          .          .          .
     agaaatttatccctgagttagtccgcttatatcaacaagggaaatttccc
322  K  F  I  P  E  L  V  R  L  Y  Q  Q  G  K  F  P
```

```
                  1020                   1040                   1060
         .          .          .          .          .
     tttgaccaacttgtaaaattctatgcatttgatgaaatcaatcaagctgc
338  F  D  Q  L  V  K  F  Y  A  F  D  E  I  N  Q  A  A
```

```
                  1080                   1100
         .          .          .          .          .
     gatagatagtcacaagggtattacactcaaaccgattattaaaattgcat
355  I  D  S  H  K  G  I  T  L  K  P  I  I  K  I  A  *
```

```
                  1120                   1140                   1160
         .          .          .          .          .
     aatcttgcatggcatgctcattgaggagcatgccattttgattataagtt
```

```
              1180                    1200
         .          .          .          .          .
ttttaagtttaaattttgggtttagtcttagattggaggtttagtgtttt
```

```
     1220                 1240                 1260
       .          .          .          .          .
aagctagttggatgactttagagctttcttcgcttttccagtaattttcc
```

```
              1280                    1300
         .          .          .          .          .
aaataataaacccagctcaaacaataaccacattgggatcgcaagcataa
```

```
1313

tc
```

Table 3.1 Codon usage in *xylB* and its comparison with codon usage in other *Acinetobacter* structural genes

Table was created using the CodonFrequency program on the GCG/UNIX system. Values in parentheses are taken from White *et al.* (1991) and correspond to the values for all 20 *Acinetobacter* structural genes that had been sequenced upto that time.

Amino acid	Codon	Total	%		Amino acid	Codon	Total	%	
Gly	GGG	8	29	(8)	Met	ATG	1	100	
	GGA	6	15	(14)					
	GGT	19	47	(55)	Ile	ATA	4	14	(3)
	GGC	7	17	(23)		ATT	18	64	(69)
						ATC	6	21	(28)
Glu	GAG	8	44	(22)					
	GAA	10	56	(78)	Thr	ACG	2	10	(17)
						ACA	5	24	(25)
Asp	GAT	16	84	(74)		ACT	7	33	(35)
	GAC	3	16	(26)		ACC	7	33	(23)
Val	GTG	11	32	(25)	Trp	TGG	1	100	
	GTA	10	29	(22)					
	GTT	8	24	(35)	Cys	TGT	9	90	(72)
	GTC	5	15	(18)		TGC	1	10	(28)
Ala	GCG	6	16	(18)	Tyr	TAT	8	100	(74)
	GCA	12	32	(38)		TAC	0	0	(28)
	GCT	12	32	(27)					
	GCC	7	19	(17)	Leu	TTG	8	23	(18)
						TTA	15	43	(35)
Arg	AGG	0	0	(1)		CTG	5	14	(11)
	AGA	0	0	(4)		CTA	1	3	(8)
	CGG	0	0	(3)		CTT	5	14	(19)
	CGA	2	29	(11)		CTC	1	3	(9)
	CGT	4	57	(62)					
	CGC	1	14	(19)	Phe	TTT	13	93	(71)
						TTC	1	7	(29)
Ser	AGT	9	47	(18)					
	AGC	1	5	(15)	Gln	CAG	1	6	(35)
	TCG	2	11	(8)		CAA	17	94	(65)
	TCA	4	21	(28)					
	TCT	1	5	(27)	His	CAT	6	75	(70)
	TCC	2	11	(5)		CAC	2	25	(30)
Lys	AAG	2	9	(24)	Pro	CCG	4	24	(16)
	AAA	20	91	(76)		CCA	3	18	(47)
						CCT	8	47	(32)
Asn	AAT	10	77	(63)		CCC	2	12	(5)
	AAC	3	23	(3)					
					Term	TGA	0	0	(30)
						TAG	0	0	(10)
						TAA	1	100	(60)

Table 3.2 Codon usage in TOL-*xylB* (pWW0) and its comparison with codon usage in *xylB* from *A. calcoaceticus*

Table was created using the CodonFrequency program on the GCG/UNIX system. Values in parentheses indicate the codon usage in *xylB* from *A. calcoaceticus* (Table 3.1).

Amino acid	Codon	Total	%	Amino acid	Codon	Total	%
Gly	GGG	16	42(29)	Met	ATG	6	100
	GGA	3	8(14)				
	GGT	12	32(55)	Ile	ATA	3	12(14)
	GGC	7	18(23)		ATT	11	44(64)
					ATC	11	44(21)
Glu	GAG	12	52(44)				
	GAA	11	48(56)	Thr	ACG	5	28(10)
					ACA	3	17(24)
Asp	GAT	9	56(84)		ACT	3	17(33)
	GAC	7	44(16)		ACC	7	39(33)
Val	GTG	8	22(32)	Trp	TGG	0	0
	GTA	8	22(29)				
	GTT	19	51(24)	Cys	TGT	7	78(90)
	GTC	2	5(15)		TGC	2	22(10)
Ala	GCG	9	24(16)	Tyr	TAT	7	100(100)
	GCA	10	26(32)		TAC	0	0 (0)
	GCT	11	29(32)				
	GCC	8	21(19)	Leu	TTG	5	15(23)
					TTA	1	3(43)
Arg	AGG	1	8 (0)		CTG	7	21(14)
	AGA	1	8 (0)		CTA	4	12 (3)
	CGG	3	23 (0)		CTT	13	39(14)
	CGA	1	8(29)		CTC	3	9 (3)
	CGT	3	23(57)				
	CGC	4	31(14)	Phe	TTT	12	71(93)
					TTC	5	29 (7)
Ser	AGT	3	15(47)				
	AGC	3	15 (5)	Gln	CAG	7	70 (6)
	TCG	5	25(11)		CAA	3	30(94)
	TCA	1	5(21)				
	TCT	4	20 (5)	His	CAT	5	62(75)
	TCC	4	20(11)		CAC	3	38(25)
Lys	AAG	8	57 (9)	Pro	CCG	10	43(24)
	AAA	6	43(91)		CCA	4	17(18)
					CCT	5	22(47)
Asn	AAT	6	46(77)		CCC	4	17(12)
	AAC	7	54(23)				
				Term	TGA	1	100 (0)
					TAG	0	0 (0)
					TAA	0	0(100)

Pseudomonas. Therefore, it is a possibility that *xylB* from the TOL-plasmid pWW0 was acquired from an *Acinetobacter* species (sections 9.1 and 9.3).

3.4 Comparison of the derived amino acid sequence of benzyl alcohol dehydrogenase with group I long-chain zinc-dependent alcohol dehydrogenases

The derived amino acid sequence of benzyl alcohol dehydrogenase was aligned against the primary sequences of group I long-chain zinc-dependent microbial alcohol dehydrogenases (including the product of the TOL plasmid gene *xylW*, which as yet has no known function) and also against the primary sequences of horse liver alcohol dehydrogenase, the archetypal long-chain zinc-dependent alcohol dehydrogenase (section 1.10.2), and those of a range of plant alcohol dehydrogenases from this enzyme family (e.g. Dennis *et al.*, 1984). Fig. 3.5 shows the alignment of the primary sequences and Fig. 3.6 expresses this alignment in terms of a dendrogram. The level of sequence identity between benzyl alcohol dehydrogenase from *A. calcoaceticus* and the other enzymes in the sequence alignment are recorded in Table 3.3.

There are 12 residues that are conserved across all of the sequences. These are Cys-46, Asp-49, His-67, Glu-68, Val-80 and the glycine residues 66, 71, 77, 86, 201, 204 and 235. In addition, Gly-261 is conserved in all but the cinnamyl alcohol dehydrogenases of the cider and poplar trees, Ala-211 is conderved in all of the microbial and plant enzymes apart from isoenzyme IV of *Kluyveromyces lactis* and Cys-103 is conserved in all of the enzymes except for alcohol dehydrogenase from *Clostridium beijerinckii*, secondary alcohol dehydrogenase from *Thermoanaerobium brockii*, fermentative alcohol dehydrogenase from *Alcaligenes eutrophus* and alcohol dehydrogenase from *Entamoeba histolytica* (section 9.4.1). The considerable identity between benzyl alcohol dehydrogenase from *A. calcoaceticus* and the enzymes in this group of alcohol dehydrogenases (Table 3.3), notably the conservation of residues proposed to be involved in the reaction mechanism of these enzymes, indicates that it is a member of the long-chain zinc-dependent alcohol dehydrogenases (section 9.2).

Crystallographic studies of horse liver alcohol dehydrogenase have shown that the enzyme binds two atoms of zinc per subunit (section 1.12; Eklund *et al.*, 1976). A catalytic zinc atom is bound at the active site to the ligands Cys-46, His-67 and Cys-174, and a structural zinc atom is bound to the cysteine residues 97, 100, 103, and 111. These residues are all conserved in benzyl alcohol

Figure 3.5 Comparison of the primary sequence of benzyl alcohol dehydrogenase from *A. calcoaceticus* with those of microbial, horse liver, human and various plant NAD(P)-dependent long-chain zinc-dependent alcohol dehydrogenases.

The comparison is based on sequences obtained from the SwissProt and GenEMBL databases through the UNIX GCG system, except ACBADH was obtained from this work and ECTH was obtained from Aronson *et al.* (1989). Abbreviations used are: ciderADH, *Eucalyptus gunii* (cider tree) cinnamyl-alcohol dehydrogenase; popADH, *Populus deltoides* (poplar tree) cinnamyl-alcohol dehydrogenase; BSADH, *Bacillus stearothermophilus* alcohol dehydrogenase; ZMADH, *Zymomonas mobilis* type I alcohol dehydrogenase; ANADHI, *Aspergillus nidulans*, type I alcohol dehydrogenase; ANADHIII, *A. nidulans* type III alcohol dehydrogenase; SADHI, *Saccharomyces cerevisiae* isoenzyme I; SADHII, *S. cerevisiae* isoenzyme II; SADHIII, *S. cerevisiae* isoenzyme III; KADHI, *Kluyveromyces lactis* isoenzyme I; KADHII, *K. lactis* isoenzyme II; KADHIII, *K. lactis* isoenzyme III; KADHIV, *K. lactis* isoenzyme IV; SPADH, *Schizosaccharomyces pombe* fermentative alcohol dehydrogenase; TBADH, *Thermoanaerobium brockii* secondary alcohol dehydrogenase; CBADH, *Clostridium beijerinckii* alcohol dehydrogenase; EADH, *Entamoeba histolytica* alcohol dehydrogenase; AEADH, *Alcaligenes eutrophus* fermentative alcohol dehydrogenase; ECTH, *Escherichia coli* L-threonine dehydrogenase; *xylW*, xylW gene product; HLADH, horse liver alcohol dehydrogenase; humADH, human class I β-subunit alcohol dehydrogenase; BarADH, *Hordeum vulgare* (barley) alcohol dehydrogenase I; MADH, *Zea mays* (maize) alcohol dehydrogenase I; riceADH, *Oryza sativa* (rice) alcohol dehydrogenase I; potADH, *Solanum tuberosum* (potato) alcohol dehydrogenase I; peaADH, *Pisum sativum* (garden pea) alcohol dehydrogenase I; ACBADH, *Acinetobacter calcoaceticus* benzyl alcohol dehydrogenase; TOL-BADH, *Pseudomonas putida* (pWW0) benzyl alcohol dehydrogenase; PADH, *Pseudomonas* sp. alcohol dehydrogenase (*terp* operon).

Sequences were aligned using the pileup program on the UNIX GCG system. Gap weight = 3.0. Gap weight length = 0.1. * corresponds to conserved residues. The sequences are numbered according to HLADH.

```
SwissProt or                                           1          10         20
GenEMBL code                                           |          |          |
p31655 cidcrADH    ...........  ...........  ......MGSL EKERTTTGWA ARDPSGVLSP
p31657 popADH      ...........  ...........  ......MGSL ETERKIVGWA ATDSTGHLAP
d90421 BSADH       ...........  ...........  .......... ...MKAAVVE QFK..KPLQV
p20368 ZMADH       ...........  ...........  .......... ...MKAAVIT K.D..HTIEV
m16196 ANADHI      ...........  ...........  .......MCI PTMQWAQVAE KVGG..PLVY
x02764 ANADHIII    ...........  ...........  ........MSV PEVQWAQVVE KAGT..PPVY
p00330 SADHI       ...........  ...........  .........SI PETQKGVIFY ESHG..KLEY
p00331 SADHII      ...........  ...........  .........SI PETQKAIIFY ESNG..KLEH
p20369 KADHI       ...........  ...........  ......MAASI PETQKGVIFY ENGG..ELQY
x64397 KADHII      ...........  ...........  ........MSI PETQKGVIFY ENGG..ELQY
x62766 KADHIII     .MLRLTSARS  IVSPLRKGAF  GSIRTLATSV PETQKGVIFY ENGG..KLEY
p07246 SADHIII     MLRTSTLFTK  RVQPSLFSRN  TLRLQSTAAT PKTQKGVIFY ENKG..NLHY
x62767 KADHIV      MFRLARAQTA  LANKASVSRS  FLRLNSSFAI PETQKGVIFY ENGG..KLEY
p00332 SPADH       ...........  ...........  .......MTI PDKQLAAVFH THGGPENVKF
p25984 CBADH       ...........  ...........  .......... .....MKGF AMLGINKLGW
p14941 TBADH       ...........  ...........  .......... .....MKGF AMLSIGKVGW
p35630 EADH        ...........  ...........  .......... ......MKGL AMLGIGRIGW
p14940 AEADH       ...........  ...........  .......... ..MTAMMKAA VFVEPGRIEL
       ECTH        ...........  ...........  .........M KALSKLKAE. .....EGIWM
u20269 xylW        ...........  ...........  .........M LAIRKTKAG. ....PGGLSV
p00327 HLADH       ...........  ...........  .......STA GKVIKCKAAV LWEEKKPFSI
p00325 humADH      ...........  ...........  .......STA GKVIKCKAAV LWEVKKPFSI
p05336 BarADH      ...........  ...........  ......MATA GKVIKCKAAV AWEAGKPLTM
p00333 MADH        ...........  ...........  ......MATA GKVIKCKAAV AWEAGKPLSI
p20306 riceADH     ...........  ...........  ......MATA GKVIKCKAAV AWEAGKPLSI
p14673 potADH      ...........  ...........  .....MSTTV GQVIRCKAAV AWEAGKPLVM
p12886 peaADH      ...........  ...........  .....MSNTV GQIIKCRAAV AWEAGKPLVI
       ACBADH      ...........  ...........  .......... SELKDIIAAV TPCKGADFEL
p39849 TOL-BADH    ...........  ...........  .......... ...MEIKAAI VRQKNGPFLL
m91440 PADH        ...........  ...........  .......... ..........
```

```
              30         40         50                    60
               |          |          |                     |
   ciderADH  YTYSLRNTCP EDLYIKVLSC G.VCHSDIHQ IK..NDLGMS .HYPMVPGHE
     popADH  YTYSLRDTGP EDVFIKVISC G.VCHTDIHQ IK..NDLGMS .HYPMVPGHE
      BSADH  KEVEKPKISY GEVLVRIKAC G.VCHTDLHA AH..GDWPVK PKLPLIPGHE
      ZMADH  KDTKLRPLKY GEALLEMEYC G.VCHTDLHV KN..GDFGDE T..GRITGHE
    ANADHII  KQIPVPKPGP DQILVKIRYS G.VCHTDLHA MM..GHWPIP VKMPLVGGHE
   ANADHIII  KQVPVPKPGP DEILVKMRYS G.VCHTDLHA MK..GDWPLP SKMPLIGGHE
      SADHI  KDIPVPKPKA NELLINVKYS G.VCHTDLHA WH..GDWPLP TKLPLVGGHE
     SADHII  KDIPVPKPKP NELLINVKYS G.VCHTDLHA WH..GDWPLP TKLPLVGGHE
      KADHI  KDIPVPKPKA NELLINVKYS G.VCHTDLHA WK..GDWPLP TKLPLVGGHE
     KADHII  KDIPVPKPKA NELLINVKYS G.VCHTDLHA WK..GDWPLP TKLPLVGGHE
    KADHIII  KDIPVPKPKP NEILINVKYS G.VCHTDLHA WK..GDWPLP TKLPLVGGHE
    SADHIII  KDIPVPEPKP NEILINVKYS G.VCHTDLHA WH..GDWPLP VKLPLVGGHE
     KADHIV  KDLPVPKPKA NEILINVKYS G.VCHTDLHA WK..GDWPLP VKLPLVGGHE
      SPADH  EEVPVAEPGQ DEVLVNIKYT G.VCHTDLHA LQ..GDWPLP AKMPLIGGHE
      CBADH  IEKERPVAGS YDAIVRPLAV S.PCTSDIHT VFEGALGDRK ...NMILGHE
      TBADH  IEKEKPAPGP FDAIVRPLAV A.PCTSDIHT VFEGAIGERH ...NMILGHE
       EADH  IEKKIPECGP LDALVRPLAL A.PCTSDTHT VWAGAIGDRH ...DMILGHE
      AEADH  ADKPIPDIGP NDALVRITTT T.ICGTDVH. ILKGEYPVAK ...GLTVGHE
       ECTH  TDVPVPELGH N.DLLIKIRK TAICGTDVHI YNWDEWSQKT IPVPMVVGHE
       xylW  DDVAAPKEAH STDVLVKVEA AGICGTDLLI YKWGEFA.KR MKLPTILGHE
      HLADH  EEVEVAPPKA HEVRIKMVAT G.ICRSDDHV V...SGTLVT P.LPVIAGHE
     humADH  EDVEVAPPKA YEVRIKMVAV G.ICRTDDHV V...SGNLVT P.LPVILGHE
     BarADH  EEVEVAPPQA MEVRVKILFT S.LCHTDVYF W...EAKGQI PMFPRIFGHE
       MADH  EEVEVAPPQA MEVRVKILFT S.LCHTDVYF W...EAKGQT PVFPRIFGHE
    riceADH  EEVEVA..KE MEVRVKILFT S.LCHTDVYF W...EAKGQT PVFPRIFGHE
     potADH  EEVDVAPPQK MEVRLKILYT S.LCHTDVYF W...EAKGQN PVFPRILGHE
     peaADH  EEVEVAPPQA GEVRLKILFT S.LCHTDVYF W...EAKGQT PLFPRIFGHE
     ACBADH  QALKIRQPQG DEVLVKVVAT G.MCHTDLIV R...DQYYPV P.LPAVLGHE
   TOL-BADH  EHVALNEPAE DQVLVRLVAT G.LCHTDLVC R...DQHYPV P.LPMVFGHE
       PADH  ......MPRR EMMKCLFVSE SCVCHTDLAI K...DGVLPF P.LPAVLGHE
                                    *    *                    ***
```

```
            70          80          90          100         110
            |           |           |           |           |
            |           |           |           :           |
 ciderADH   VVGEVLEVGS  EVTK..YRVG  DRVGTGIVVG  CCRSCSPCNS  DQEQYCNKKI
   popADH   VVGEVVEVGS  DVTR..FKVG  DVVGVGVIVG  SCKNCHPCKS  EIEQYCNKKI
   BSADH    GVGVIEEVGP  GVT..HLKVG  DRVGIPWLYS  ACGHCDYCLS  GQETLCERQQ
   ZMADH    GIGIVKQVGE  GVT..SLKAG  DRASVAWFFK  GCGHCEYCVS  GNETLCRNVE
  ANADHI    GAGIVVAKGE  LV..HEFEIG  DQAGIKWLNG  SCGECEFCRQ  SDDPLCARAQ
 ANADHIII   GAGVVVAKGE  LVKDEDFKIG  DRAGIKWLNG  SCLSCEMCMQ  ADEPLCPHAS
   SADHI    GAGVVVGMGE  NVK..GWKIG  DYAGIKWLNG  SCMACEYCEL  GNESNCPHAD
   SADHII   GAGVVVGMGE  NVK..GWKIG  DYAGIKWLNG  SCMACEYCEL  GNESNCPHAD
   KADHI    GAGVVVAMGE  NVK..GWKIG  DFAGIKWLNG  SCMSCEYCEL  SNESNCPEAD
   KADHII   GAGVVVAMGE  NVK..GWIIG  DFAGIKWLNG  SCMSCEYCEL  SNESNCPDAD
  KADHIII   GAGVVVAMGE  NVK..GWNIG  DFAGIKWLNG  SCMSCEYCEL  SNESNCPDAD
  SADHIII   GAGVVVKLGS  NVK..GWKVG  DLAGIKWLNG  SCMTCEFCES  GHESNCPDAD
   KADHIV   GAGIVVAKGE  NVK..NFEIG  DYAGIKWLNG  SCMSCELCEQ  GYESNCLQAD
   SPADH    GAGVVVKVGA  GVTR..LKIG  DRVGVKWMNS  SCGNCEYCMK  AEETICPHIQ
   CBADH    AVGEVVEVGS  EVK..DFKPG  DRV.IVPCTT  PDWRSLEVQA  GFQQHSNGM.
   TBADH    AVGEVVEVGS  EVK..DFKPG  DRV.VVPAIT  PDWRTSEVQR  GYHQHSGGM.
   EADH     AVGQIVKVGS  LVK..RLKVG  DKV.IVPAIT  PDWGEEESQR  GYPMHSGGM.
   AEADH    PVGIIEKLGS  AVT..GYREG  QRV.IAGAIC  PNFNSYAAQD  GVASQDGSY.
   ECTH     YVGEVVGIGQ  EVK..GFKIG  DRV.SGEGHI  TCGHCRNCRG  GRTHLCR...
   xylW     VSGVIEQVGS  DVK..GLRPG  MRV.SLESHL  PCGTCYTCRR  GWAHVCP...
   HLADH    AAGIVESIGE  GVT..TVRPG  DKV.IPLFTP  QCGKCRVCKH  PEGNFCLKND
   humADH   AAGIVESVGE  GVT..TVKPG  DKV.IPLFTP  QCGKCRVCKN  PESNYCLKND
   BarADH   AGGIVESVGE  GVT..DVAPG  DHV.LPVFTG  ECKECPHCKS  AESNMCDLLR
   MADH     AGGIIESVGE  GVT..DVAPG  DHV.LPVFTG  ECKECAHCKS  AESNMCDLLR
  riceADH   AGGIVESVGE  GVT..DLAPG  DHV.LPVFTG  ECKECAHCKS  AESNMCDLLR
   potADH   AAGIVESVGE  GVT..ELGPG  DHV.LPVFTG  ECKDCAHCKS  EESNMCSLLR
   peaADH   AGGIVESVGE  GVT..HLKPG  DHA.LPVFTG  ECGECPHCKS  EESNMCDLLR
   ACBADH   GSGIIEAIGP  NVT..ELQVG  DHV.VLSY.G  YCGKCTQCNT  GNPAYCSEF.
 TOL-BADH   GAGVVERVGS  AVK..KVQPG  DHV.VLTF.Y  TCGSCDACLS  GDPTSCANS.
   PADH     GSGIVEAVGP  GVK..HLKPG  DAV.VMTF.A  SCGHCASCEH  ERPSYCLDF.
            *           *           *           *
```

```
                120          130          140          150          160
                 |            |            |            |            |
  ciderADH   WNYNDVYTDG  ..........  ......KPT  QGGFAGEIVV  GERFVVKI..
    popADH   WSYNDVYTDG  ..........  ......KPT  QGGFAESMVV  HQKFVVRI..
     BSADH   NA......G   ..........  ......YSV  DGGYAEYCRA  AADYVVKI..
     ZMADH   NA......G   ..........  ......YTV  DGAMAEECIV  VADYSVKV..
    ANADHI   LS......G   ..........  ......YTV  DGTFQQYALG  KASHASKI..
  ANADHIII   LS......G   ..........  ......YTV  DGTFQQYTIG  KAALASKI..
     SADHI   LS......G   ..........  ......YTH  DGSFQQYATA  DAVQAAHI..
    SADHII   LS......G   ..........  ......YTH  DGSFQEYATA  DAVQAAHI..
     KADHI   LS......G   ..........  ......YTH  DGSFQQYATA  DAVQAAKI..
    KADHII   LS......G   ..........  ......YTH  DGSFQQYATA  DAVQAARI..
   KADHIII   LS......G   ..........  ......YTH  DGSFQQYRTA  DAVQAARI..
   SADHIII   LS......G   ..........  ......YTH  DGSFQQFATA  DAIQAAKI..
    KADHIV   LS......G   ..........  ......YTH  DGSFQQYRTA  DAVQAAQI..
     SPADH   LS......G   ..........  ......YTV  DGTFQHYCIA  NATHATII..
     CBADH   ......L...  ..........  .AGWKFSNFK  DGVFGEYFHV  NDAD.MNLAI
     TBADH   ......L...  ..........  .AGWKFSNVK  DGVFGEFFHV  NDAD.MNLAH
      EADH   ......L...  ..........  .GGWKFSNFK  DGVFSEVFHV  NEAD.ANLAL
     AEADH   ......LMAS  GQCGCHGYKA  TAGWRFGNMI  DGTQAEYVLV  PDAQ.ANLTP
      RCTH   .NTIGVGVNR  PG........  ..........  ..CFAEYLVI  PAFNAFKIPD
      xylW   .KTRYPGVDF  DG........  ..........  ..GFASFVVV  PESVCWPVPC
     HLADH   LSMPRGTM.Q  DGTSRFTCRG  KPTH..HFLG  TSTFSQYTVV  DEISVAKIDA
    humADII   LGNPRGTL.Q  DGTRRFTCRG  KPIH..HFLG  TSTFSQYTVV  DENAVAKIDA
     BarADH   INTDRGVMIG  DGKSRFSIGG  KPIY..HFVG  TSTFSEYTVM  HVGCVAKINP
      MADH   INTDRGVMIA  DGKSRFSING  KPIY..HFVG  TSTFSEYTVM  HVGCVAKINP
    riceADH   INTDRGVMIG  DGKSRFSING  KPTY..HFLG  TSTFSEYTVM  HVGCVAKINP
     potADH   INTDRGVMIN  DGQSRFSING  KPIY..HFVG  TSTFSEYTVV  HVGCVAKINP
     peaADH   INTDRGVMLN  DNKSRFSIKG  QPVH..HFVG  TSTFSEYTVV  HAGCVAKINP
    ACBADH   FGRNFSGADS  EGNHALCTHD  QGVVNDHFFA  QSSFATYALS  RENNTVKV..
  TOL-BADH   FGPNFMGRSV  TGECTIHDHQ  GAEVGASFFG  QSSFATYALS  YERNTVKV..
      PADH   CAQNYSAQRA  DGPVLL.SQG  DEVISGFFFG  QSSFSSMAMA  REHNLVKIDA
```

```
                     170         180         190         200
                      |           |           |           |
  ciderADH  .PDGLESEQA APLMCAGVTV YSPLVRFGLK QSGLRGGILG LGGVGHMGVK
    popADH  .PDGMSPEQA APLLCAGLTV YSPLKHFGLK QSGLRGGILG LGGVGHMGVK
     BSADH  .PDNLSFEEA APIFCAGVTT YKALKVTGAK PGEWVAIYGI .GGLGHVAVQ
     ZMADH  .PDGLDPAVA SSITCAGVTT YKAVKVSQIQ PGQWLAIYGL .GGLGNLALQ
    ANADHI  .PAGVPVDAA APVLCAGITV YKGLKEACVR PGQTVAIVGA GGGLGSLAQQ
  ANADHIII  .PDNVPLDAA APILCAGITV YKGLKESGAR PGQTVAIVGA GGGLGSLAQQ
     SADHI  .PQGTDLAEV APVLCAGITV YKALKSANLM AGHWVAISGA AGGLGSLAVQ
    SADHII  .PQGTDLAEV APTLCAGITV YKALKSANLR AGHWAAISGA AGGLGSLAVQ
     KADHI  .PVGTDLAEV APVLCAGVTV YKALKSANLK AGDWVAISGA AGGLGSLAVQ
    KADHII  .PKGTDLAEV APILCAGVTV YKALKSADLK AGDWVAISGA CGGLGSLAIQ
   KADHIII  .PKGTDLAEV APTLCAGVTV YKALKSANLK AGDWVAISGA AGGLGSLAVQ
   SADHIII  .QQGTDLAEV APILCAGVTV YKALKEADLK AGDWVAISGA AGGLGSLAVQ
    KADHIV  .PKGTDLAEI APILCAGVTV YKALKTADLK PGQWVAISGA AGGLGSLAVQ
     SPADH  .PRSVPLEVA APIMCAGITC YRALKESKVG PGEWICIPGA GGGLGHLAVQ
     CBADH  LPKDMPLENA VMITDMMTSG FHGA.ELADI QMGSSVVVIG IGAVGLMGIA
     TBADH  LPKEIPLEAA VMIPDMMTTG FHGA.ELADI ELGATVAVLG IGPVGLMAVA
      EADH  LPRDIKPEDA VMLSDMVTTG FHGA.ELANI KLGDTVCVIG IGPVGLMSVA
     AEADH  IPDGLTDEQV LMCPDIMSTG FKGA.ENANI RIGHTVAVFA QGPIGLCATA
      ECTH  NISDDLAAIF DPFGNAVHTA L.SFDLV... ..GEDVLVSG AGPIGIMAAA
      xylW  GIPPLQATMM RPFGLAVHAS LEGSGVS... ..GLNVLVSG CGPIGLMNIA
     HLADH  ...ASPLEKV CLIGCGFSTG YGSAVKVAKV TQGSTCAVFG LGGVGLSVIM
    humADH  ...ASPLEKV CLIGCGFSTG YGSAVNVAKV TPGSTCAVFG LGGVGLSAVM
     BarADH  ...EAPLDKV CVLSCGISTG LGASINVAKP PKGSTVAIFG LGAVGLAAAE
      MADH  ...QAPLDKV CVLSCGYSTG LGASINVAKP PKGSTVAVFG LGAVGLAAAE
   riceADH  ...AAPLDKV CVLSCGISTG LGATINVAK. PKGSTVAIFG LGAVGLAAAE
    potADH  ...LAPLDKV CVLSCGISTG LGATLNVAKP TKGSSVAIFG LGAVGLAAAE
    peaADH  ...DAPLDKV CILSCGICTG LGATINVAKP KPGSSVAIFG LGAVGLAAAE
    ACBADH  .TKDVPIELL GPLGCGIQTG AGACINALKV TPASSLVTWG AGAVGLSALL
   TOL-BADH  .TKDVPLELL GPLGCGIQTG AGSVLNALNP PAGSAIAIFG AGAVGLSAVM
      PADH  LVDDAPIELL GPLGCGVQTG AGAVMISLDV RAGRSFLVLG GGAVGLSAVM
                                                            *    *
```

```
                210        220        230        240        250
                 |          |          |          |          |
  ciderADH   IAKAMGHHVT VISSSDKKRT EALEHLGADA YL......V SSDENGMKEA
    popADH   IAKAMGHHVT VISSSDKKRE EAMEHLGADE YL......V SSDVESMQKA
     BSADH   YAKAM.GLNV VAVDLGDEKL ELAKQLGADL VVN.PKHDDA AQWIKEKVG.
     ZMADH   YAKNVFNAKV IAIDVNDEQL AFAKELGADM VIN.PKNEDA AKIIQEKVG.
    ANADHI   YAKAM.GIRV VAVDGGDEKR AMCESLGTET YVDFTKSKDL VADVRHG.RG
  ANADHIII   YAKAM.GLRT IAIDSGDEKK AMCEQLGAEV FIDFSKSADV VADVKAATPG
     SADHI   YAKAM.GYRV LGIDGGEGKE ELFRSIGGEV FIDFTKEKDI VGAVLKATNG
    SADHII   YAKAM.GYRV LGIDGGPGKE ELFTSLGGEV FIDFTKEKDI VSAVVKATNG
     KADHI   YAKAM.GYRV LGIDAGEEKA KLFKDLGGEY FIDFTKSKNI PEEVIEATKG
    KADHII   YAKAM.GYRV LGIDTGAEKA KLFKELGGEY FVDYAVSKDL IKEIVDATNG
   KADHIII   YAKAM.GYRV VGIDGGEEKG KLVKQLGGEA FVDFTKTKDM VAEIQEITNG
   SADHIII   YATAM.GYRV LGIDAGEEKE KLFKKLGGEV FIDFTKTKNM VSDIQEATKG
    KADHIV   YRKAM.GLRV LGIDGGDGKE ELFKQCGGEV FIDFRKSKDM VADIQEATNG
     SPADH   YAKAM.AMRV VAIDTGDDKA ELVKSFGAEV FLDFKKEADM IEAVKACTNG
     CBADH   GAKLRGAGRI IGVGSRPICV EAAKFYGATD ILNYKN.GHI VDQVMKLTNG
     TBADH   GAKLRGAGRI IAVGSRPVCV DAAKYYGATD IVNYKD.GPI ESQIMNLTEG
      EADH   GANHLGAGRI FAVGSRKHCC DIALEYGATD IINYKN.GDI VEQILKATDG
      AEADH   GARLCGATTI IAIDGNDHRL EIARKMGADV VLNFRN.CDV VDEVMKLTGG
      ECTH   VAKHVGARNV VITDVNEYRL ELARKMGITR AVNVAK.ENL NDVMAELGMT
      xylW   AAKALGASKV IATDIHPLRL TAAAKMGADE CINATE.ESV YKTVRSLLGE
     HLADH   GCKAAGAARI IGVDINKDKF AKAKEVGATE CVNPQDYKKP IQEVLTEMSN
    humADH   GCKAAGAARI IAVDINKDKF AKAKELGATE CINPQDYKKP IQEVLKEMTD
     BarADH   GARIAGASRI IGVDLNAVRF EEARKFGCTE FVNPKDHTKP VQQVLADMTN
      MADH   GARIAGASRI IGVDLNPSRF EEARKFGCTE FVNPKDHNKP VQEVLAEMTN
   riceADH   GARTRGASRI IGIDLNANRF EEARKFGCTE FVNPKDHDKP VQQVLAEMTN
    potADH   GARIAGASRI IGVDLNASRF EQAKKFGVTE FVNPKDYSKP VQEVIAEMTD
    peaADH   GARISGASRI IGVDLVSSRF ELAKKFGVNE FVNPKEHDKP VQQVIAEMTN
     ACBADH   AAKVCGASII IAVDIVESRL ELAKQLGATH VINSKT.QDP VAAIKEITDG
  TOL-BADH   AAVVAGCTTI IAVDVKENRL ELASELGATH IINPAA.NDP IEAIKEIFAD
      PADH   AAKLRGCSRI IVSEPSAAKR EQALALGATE VIDPLN.ENL VERVQQITEG
                                   *
```

```
            260         270         280         290         300
             |           |           |           |           |
  ciderADH  TDSLDYIFDT  IPVVHPLEPY  LALLKLDGKL  ILTGV.INA.  ..PLQFISPM
    popADH  ADQLDYIIDT  VPVVHPLEPY  LSLLKLDGKL  ILMGV.INA.  ..PLQFVTPM
     BSADH  ..GVHATVVT  AVSKAAFESA  YKSIRRGGAC  VLVGLPPE..  ..EIPIPIFD
     ZMADH  ..GAHATVVT  AVAKSAFNSA  VEAIRAGGRV  VAVGLPPE..  ..KMDLSIPR
   ANADHII  CLGAHAVILL  AVSEKPFQQA  TEYVRSRGTI  VAIGLPPDA.  ..YLKAPVIN
  ANADHIII  GLGAHAVILL  AVAEKPFQQA  TEYVRSHGSV  VAIGLPANA.  ..FLKAPVFT
     SADHI  ..GAHGVINV  SVSEAAIEAS  TRYVRANGTT  VLVGMPAGA.  ..KCCSDVFN
    SADHII  ..GAHGIINV  SVSEAAIEAS  TRYCRANGTV  VLVGLPAGA.  ..KCSSDVFN
     KADHI  ..GAHGVINV  SVSEFAIEQS  TNYVRSNGTV  VLVGLPRDA.  ..KCKSDVFN
    KADHII  ..GAHGVINV  SVSEFAIEQS  TNYVRSNGTV  VLVGLPRDA.  ..KCKSDVFT
   KADHIII  ..GPHGVINV  SVSEAAMNAS  TQFVRPTGTV  VLVGLPAGA.  ..VIKSEVFS
   SADHIII  ..GPHGVINV  SVSEAAISLS  TEYVRPCGTV  VLVGLPANA.  ..YVKSEVFS
    KADHIV  ..GPHGVINV  SVSEAAVSMS  TEYLRPTGLV  VLVGLPADA.  ..YVKSEVFS
     SPADH  ..GAHGTLVL  STSPKSYEQA  AGFARPGSTM  VTVSMPAGA.  ..KLGADIFW
     CBADH  E.GVDRVIMA  GGGSETLSQA  VSMVKPGGII  SNINYHGSGD  ALLIPRVEWG
     TBADH  K.GVDAAIIA  GGNADIMATA  VKIVKPGGTI  ANVNYFGEGE  VLPVPRLEWG
      EADH  K.GVDKVVIA  GGDVHTFAQA  VKMIKPGSDI  GNVNYLGEGD  NIDIPRSEWG
     AEADH  R.GVDASIEA  LGTQATFEQS  LRVLKPGGTL  SSLGVYSSD.  .LTIPLSAFA
      ECTH  E.GFDVGLEM  SGAPPAFRTM  LDTMNHGGRI  AMLGIPPSD.  ........MS
      xylW  R.GVDVAIDY  SGQAAALKEL  IQSITHGGEL  RLMGVPSHD.  ........IL
     HLADH  G.GVDFSFEV  IGRLDTMVTA  LSCCQEAYGV  SVIVGVPPDS  Q.NLSMNPML
    humADH  G.GVDFSFEV  IGRLDTMMAS  LLCCHEACGT  SVIVGVPPAS  Q.NLSINPML
     BarADH  G.GVDRSVEC  TGNVNAMIQA  FECVHDGWGV  AVLVGVPHKD  A.EFKTHPMN
      MADH  G.GVDRSVEC  TGNINAMIQA  FRCVHDGWGV  AVLVGVPHKD  A.EFKTHPMN
   riceADH  G.GVDRSVEC  TGNINAMIQA  FECVHDGWGI  AVLVGVPHKD  A.EFKTHPMN
    potADH  G.GVDRSVEC  TGHIDAMISA  FECVHDGWGV  AVLVGVPHKE  A.VFKTHPMN
    peaADH  G.GVDRAVEC  TGSIQAMISA  FECVHDGWGV  AVLVGVPSKD  D.AFKTHPMN
    ACBADH  ..GVNFALES  TGRPEILKQG  VDALGILGKI  AVVGAPQ.LG  T.TAQFDVND
  TOL-BADH  ..GVPYVLET  SGLPAVLTQA  ILSSAIGGEI  GIVGAPP.MG  A.TVPVDIN.
      PADH  Q.GCDYALEC  TGLVSVMEQA  IDSMAMRGQL  AVVGVPPKID  A.TAAVSPLA
```

```
                310         320         330         340         350
                 |           |           |           |           |
 ciderADH   VMLGRKSTTG  SFIGSMKETE  E......MLE  FCKEKGLT..  ..SQIEVIKM
   popADH   VMLGRKSITG  SFIGSMKETE  E......MLE  FCKEKGVA..  ..SMIEVIKM
    BSADH   TVLNGVKIIG  SIVGTRKDLQ  E......ALQ  FAAEGKVK..  ..TIVEVQPL
    ZMADH   LVLDGIEVLG  SLVGTREDLK  E......AFQ  FAAEGKVK..  ..PKVTKRKV
   ANADHI   TVVRMITIKG  SYVGNRQDGV  E......ALD  FFARGLIK..  ..APFKTAPL
 ANADHIII   TVVRMINIKG  SYVGNRQDGV  E......ALD  FFARGLIK..  ..APFKKAPL
    SADHI   QVVKSISIVG  SYVGNRADTR  E......ALD  FFARGLIK..  ..SPIKVVGL
   SADHII   HVVKSISIVG  SYVGNRADTR  E......ALD  FFARGLVK..  ..SPIKVVGL
    KADHI   QVVKSISIVG  SYVGNRADTR  E......AID  FFSRGLVK..  ..APIHVVGL
   KADHII   QVVKSVSIVG  SYVGNRADTR  E......ALD  FFARGLVH..  ..APIKIVGL
  KADHIII   HVVKSININKG SYVGNRADTR  E......AIN  FFANGHVH..  ..SPIKVVGL
  SADHIII   HVVKSININKG SYVGNRADTR  E......ALD  FFSRGLIK..  ..SPIKIVGL
   KADHIV   HVVKSISIKG  SYVGNRADTR  E......ATD  LFTRGLVK..  ..SPIKIIGL
    SPADH   LTVKMLKICG  SHVGNRIDSI  E......ALE  YVSRGLVK..  ..PYYKVQPF
    CBADH   CGMAHKTIKG  GLCPGGRLRA  E.....MLRD  MVVYNRVDLS  KLVTHVYHGF
    TBADH   CGMAHKTIKG  GLCPGGRLRM  E.....RLID  LVFYKRVDPS  KLVTHVFRGF
     EADH   VGMGHKHIHG  GLTPGGRVRM  E.....KLAS  LISTGXLDTS  KLITHRFEGL
    AEADH   AGLGDHKINT  ALCPGGKERM  R.....RLIN  VIESGRVDLG  ALVTHQYR.L
     ECTH   IDWTKVIFKG  LFIKGIYGRE  MFETWYKMAA  LIQSG.LDLS  PLITHRF.SI
     xylW   LNLEEWLFKG  LIVRGLHGRR  LYETWERSTS  LLVTGRVDLS  SLVSHRL.PL
    HLADH   LL.SGRTWKG  AIFGGFKSKD  SVP...KLVA  DFMAKKFALD  PLITHVL.PF
   humADH   LL.TGRTWKG  AVYGGFKSKE  GIP...KLVA  DFMAKKFSLD  ALITHVL.PF
   BarADH   FL.NERTLKG  TFFGNFKPRT  DLP...NVVE  MYMKKELEVE  KFITHSV.PF
     MADH   FL.NERTLKG  TFFGNYKPRT  DLP...NVVE  LYMKKELEVE  KFITHSV.PF
  riceADH   FL.NERTLKG  TFFGNYKPRT  DLP...NVVE  LYMKKELEVE  KFITHSV.PF
   potADH   LL.NERTLKG  TFFGNYKPRS  DIP...SVVE  KYMNKELELE  KFITHTL.PF
   peaADH   FL.NERTLKG  TFYGNYKPRT  DLP...NVVE  KYMKGELELE  KFITHTV.PF
   ACBADH   LLLGGKTILG  VVEGSGSPKK  FIP...ELVR  LYQQGKFPFD  QLVKF.Y.AF
 TOL-BADH   FLLFNRXLRG  IVEGQSISDI  FIP...RLVE  LYRQGKFPFD  KLIKF.Y.PF
     PADH   LIQKGLKLMG  VIEGDSCRVY  SST...SCTR  SSRLGVFRSP  R.........
```

```
                     360              370
                      |                |
  ciderADH   DYVNTALERL EKNDVRYRFV VDVVGSKLD. ...
    popADH   DYINTHSRGL RKMM...... .......... ...
     BSADH   ENINDVFDRM LKGQINGRVV LKVD...... ...
     ZMADH   EEINQIFDEM EHGKFTGRMV VDFTHH.... ...
    ANADHI   KDLPKIYELM EQGRIAGRYV LEMPE..... ...
  ANADHIII   QDLPQIFELM GQGKIAGRYV LEIPE..... ...
     SADHI   STLPEIYEKM EKGQIVGRYV VDTSK..... ...
    SADHII   SSLPEIYEKM EKGQIAGRYV VDTSK..... ...
     KADHI   SELPSIYEKM EKGAIVGRYV VDTSK..... ...
    KADHII   SELADVYDKM VKGEIVGRYV VDTSK..... ...
   KADHIII   SELPKVYELM EQGKILGRYV VDTSN..... ...
   SADHIII   SELPKVYDLM EKGKILGRYV VDTSK..... ...
    KADHIV   SELPEAYALM EQGKILGRFV VDTYK..... ...
     SPADH   STLPDVYRLM HENKIAGRIV LDLSK..... ...
     CBADH   DHIEEALLLM KDKPKDLIKA VVIL...... ...
     TBADH   DNIEKAFMLM KDKPKDLIKP VVILA..... ...
      EADH   EKVEDALMLM KNKPADLIKP VVRIHYDDED TLH
     AEADH   DDIVAAYDLF ANQRDGVLKI AIKPH..... ...
      ECTH   DDFQKGFDAM RS..GQSGKV ILSWD..... ...
      xylW   SAAEDAFEMA LN..GQSLKI LFEPNGSCI. ...
     HLADH   EKINEGFDLL RS..GESIRT ILTF...... ...
    humADH   EKINEGFDLL HS..GKSIRT VLTF...... ...
     BarADH  SEINTAFDLM AK..GEGIRC IIRMDN.... ...
      MADH   AEINKAFDLM AK..GEGIRC IIRMEN.... ...
   riceADH   SEINTAFDLM IIK..GEATRC IIRMEN.... ...
    potADH   AEINKAFDLM LK..GEGLRC IITMED.... ...
    peaADH   SEINKAFDYM LK..GESIRC IIKMEE.... ...
    ACBADH   DEINQAAIDS HK..GITLKP IIKIA..... ...
  TOL-BADH   DEINRAAEDS EK..GVTLKP VLRIG..... ...
      PADH   .......... .......... .......... ...
```

Figure 3.6 Dendrogram indicating the primary sequence alignment of the group I long-chain zinc-dependent alcohol dehydrogenases

Lineup was created using the pileup program on the UNIX GCG system as described for Fig. 3.5. Graphics were created using the Wisconsin Package Interface (WPI) and the GCG/MacX system. Abbreviations used are as described for Fig. 3.5.

cider ADH
popADH
BSADH
ZMADH
ANADHI
ANADHIII
SADHIII
KADHIV
KADHIII
SADHI
SADHII
KADHI
KADHII
SPADH
CBADH
TBADH
EADH
AEADH
ECTH
xylW
HLADH
humADH
barADH
MADH
riceADH
potADH
peaADH
ACBADH
TOL-BADH
PADH
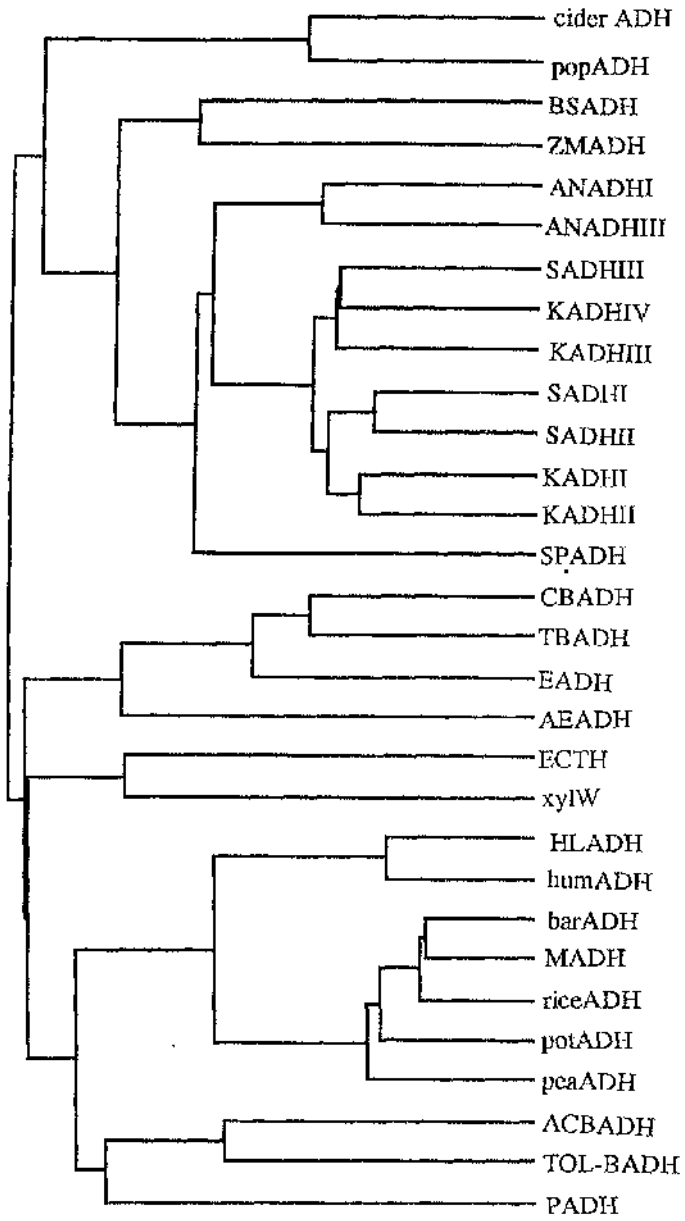
Table 3.3 Table showing the percentage sequence identites and similarities of members of the group I long-chain zinc-dependent alcohol dehydrogenases with benzyl alcohol dehydrogenase from *A. calcoaceticus*

Values were determined using the Bestfit program on the GCG UNIX system. Abbreviations are as described for Fig. 3.5.

Enzyme	% identity with ACBADH	% similarity with ACBADH
ciderADH	23%	48%
popADH	22	46
BSADH	32	56
ZMADH	29	51
ANADHI	25	50
ANADHIII	25	48
SADHI	24	49
SADHII	25	49
SADHIII	24	49
KADHI	24	49
KADHII	24	51
KADHIII	26	48
KADHIV	24	49
SPADH	27	51
CBADH	27	54
TBADH	27	53
EADH	26	50
AEADH	23	48
ECTH	25	54
xyIW	28	56
HLADH	32	54
humADH	35	55
BarADH	30	57
MADH	31	57
riceADH	31	58
potADH	33	56
peaADH	33	58
TOL-BADH	54	71
PADH	42	63

dehydrogenase from *A. calcoaceticus* (Fig. 3.5). Residue 223 (horse liver alcohol dehydrogenase numbering) has previously been implicated in determining coenzyme specificity (Fan *et al.*, 1991). NAD^+ utilising enzymes have an aspartate residue in this position whilst NADP utilising enzymes have a glycine. *A. calcoaceticus* benzyl alcohol dehydrogenase has an aspartate in this position and has been shown to have a specific requirement for NAD^+ as electron acceptor (MacKintosh & Fewson, 1988a).

The proposed proton relay system of horse liver alcohol dehydrogenase involves residues Ser-48 and His-51 (section 1.12; Eklund *et al.*, 1982). Across the group of long-chain zinc-dependent alcohol dehydrogenases residue 48 is conserved as either a serine or a threonine (Fig. 3.5; Reid & Fewson, 1994). Site-directed mutagenesis has shown that the OH group from either of these residues is necessary for catalytic activity and that serine and threonine are able to be interchanged in this position without the loss of catalytic activity (Creaser *et al.*, 1990; Sakoda & Imanaka, 1992). *A. calcoaceticus* benzyl alcohol dehydrogenase is shown here to have a threonine in this position. His-51, or rather the imidazole ring of His-51, has been proposed to be involved in proton transfer from the buried alcohol in the active site and the solvent at the surface of the enzyme (Eklund *et al.*, 1982). This residue is conserved in the majority of the enzymes lined up in Fig 3.5. These include the *A. calcoaceticus* and TOL-benzyl alcohol dehydrogenases, and also the plant alcohol dehydrogenases from barley, maize, rice, potato and peas. These latter plant enzymes have a tyrosine in this position. A tyrosine residue would be capable of performing the hydrogen bonding role proposed to be carried out by the His-51 residue. Benzyl alcohol dehydrogenase from *A. calcoaceticus* and TOL-benzyl alcohol dehydrogenase have an isoleucine and a valine, respectively, in this position. This suggests that the proton relay system may be rearranged in these two enzymes (section 9.4.4).

Fig. 3.6 expresses the sequence alignment in graphical form (Needleman & Wunsch, 1970). Distance along the vertical axis is proportional to the sequence identity between clustered sequences. The clustering of sequences along the horizontal axis indicates which sequences are most related to each other. It can be seen that benzyl alcohol dehydrogenase from *A. calcoaceticus* has most sequence identity with the TOL-benzyl alcohol dehydrogenase from *P. putida* (pWW0) and that these two enzymes are more related to the alcohol dehydrogenases of horse liver, barley, maize, rice, potato and pea, rather than the other microbial enzymes. TOL-benzyl alcohol dehydrogenase and *A. calcoaceticus* benzyl alcohol dehydrogenase share 54% identity as determined by the Bestfit program on the GCG UNIX system. The *A. calcoaceticus* enzyme also shares at least 30% sequence identity with the horse liver , barley, maize, rice,

potato and pea enzymes (Table 3.3). This raises interesting questions concerning the evolutionary origins of these enzymes (section 9.3).

Chapter 4. Heterologous expression studies with cloned benzyl alcohol dehydrogenase (and benzaldehyde dehydrogenase II) and the subsequent purification of benzyl alcohol dehydrogenase

4.1 The heterologous expression of cloned benzyl alcohol dehydrogenase and benzaldehyde dehydrogenase II

The clones pDG12 and pDG20 were initially used to express benzyl alcohol dehydrogenase and, in the case of pDG12, benzaldehyde dehydrogenase II also. Table 4.1 shows the specific activities of these enzymes in extracts of *E. coli* DH5α transformed with pDG12 or pDG20 and compares them with the specific activities of the enzymes in the wild-type organism. The levels of expression of both benzyl alcohol dehydrogenase and benzaldehyde dehydrogenase II 10-fold higher from pDG12 than in the wild type organism and the level of expression of benzyl alcohol dehydrogenase from pDG20 was 20 times that of *A. calcoaceticus*. The increased expression of the two recombinant enzymes above the level of expression in the wild-type organism could not be detected by SDS-PAGE (results not shown). Benzyl alcohol dehydrogenase could be purified from *E. coli* DH5α/pDG20 (sections 2.27 & 4.2). Benzaldehyde dehydrogenase II had previously been purified from *E. coli* JM109 transformed with a clone expressing the enzyme to a similar level as pDG12 (A. G. S. Robertson, unpublished results).

The clone pDG30 was constructed (section 3.1) to obtain a very high over-expression of benzyl alcohol deydrogenase using the T7 expression system (Studier & Moffat, 1986). This clone enabled the purification of greater than 10 mg quantities of benzyl alcohol dehydrogenase from less than 10 g of cells (section 4.2). Section 3.1 described the construction of pDG30 by ligating PCR generated *xylB* into the T7 expression vector pTB361. When transformed into a bacterial strain lysogenised with λ_{DE3} (i.e a strain that carried the T7 RNA polymerase gene on its chromosome under *lac* control) the high level expression of *xylB*, now situated downstream of the T7 promoter, could be achieved. Fig 4.1 shows a time course expression of benzyl alcohol dehydrogenase in *E. coli* JM109 (DE3)/pDG30. The increase in expression of the enzyme with time can clearly be seen. *XylB* was expressed prior to the addition of IPTG and so this system is "leaky". This is due to low-level constitutive *lac* expression and consequently the presence of T7 RNA polymerase prior to the addition of IPTG. Once expression had been induced the level of expression was sufficient to retard the growth of the cells. This was clearly visible by comparing the growth of *E. coli* JM109 (DE3)/pTB361 4 hours post-induction to that of *E. coli* JM109 (DE3)/pDG30 4 hours post-induction.

Table 4.1 shows that the specific activity of benzyl alcohol dehydrogenase expressed from *E. coli* JM109 (DE3) was about 30-fold higher than that of *A. calcoaceticus* . This value seems a little low considering that the

Table 4.1 Specific activities of benzyl alcohol dehydrogenase and benzaldehyde deydrogenase II expressed from various clones in _E. coli_

Enzyme	Specific activities (U/mg protein) of enzymes expressed in _A. calcoaceticus_ and from various clones in _E. coli_			
	A.calcoaceticus NCIB8250	pDG12	pDG20	pDG30
BADH	0.8	7.0	15.7	25.1
BZDH II	0.1	1.3	0.0	N. D.

pDG12 and pDG20 were expressed in _E. coli_ DH5α. pDG30 was expressed in _E. coli_ JM109 (DE3). Specific activities were calculated from extracts of bacteria grown to an OD_{600} of 0.4, induced with 0.5 mM IPTG and then allowed to express the enzymes for 18 hours (section 2.25). Enzymes expressed in _A. calcoaceticus_ were induced with 5 mM benzyl alcohol. Samples (1 ml) were removed from the cultures, pelleted by centrifugation, resuspended in 25 mM Tris/HCl, pH7.5, 2 mM DTT and sonicated for 6 x 10s (sections 2.25). The homogenate was then centrifuged and enzyme activities determined as described in section 2.26. Protein concentrations were determined in the supernatants as described in section 2.28 by the method of Lowry _et al._. (1951).

protein was now very prominent in crude extracts analysed by SDS-PAGE (Fig. 4.1). Laser scanning densitrometry of these gels (section 2.29) showed that benzyl alcohol dehydrogenase constituted approximately 56% of the total soluble protein in these crude extracts. It is possible that in these extracts the presence of NADH oxidase cause the underestimation of enzyme activities. These particular extracts were centrifuged at about 12,000 r.p.m for 5 minutes in a microfuge whereas the crude extracts made for protein purification were centrifuged at 40,000 r.p.m for 2 hours (sections 2.25 & 2.27). Longer and harder spins will pellet more of the membranes and consequently should reduce the NADH oxidase activity of crude extracts. Extracts used for enzyme purification had much higher specific activities. For example, the extract used in the purification of benzyl alcohol dehydrogenase described in section 4.2 had a specific activity of 303 U/mg protein, although some of this increase may be a consequence of better expression that was achieved when cells were grown in larger flasks with better aeration (sections 2.24 & 4.2).

4.2 Purification of benzyl alcohol dehydrogenase

Benzyl alcohol dehydrogenase was purified from both *E. coli* DH5α/pDG20 and *E. coli* JM109 (DE3)/pDG30 using the method of MacKintosh & Fewson (1988a) developed for purification from *A. calcoaceticus* NCIB8250 (section 2.27). Fig. 4.2 shows an example of a purification from *E. coli* JM109 (DE3)/pDG30, showing that benzyl alcohol dehydrogenase could be purified equally as well from an *E. coli* background using this method as from the wild-type organism. The purification from *A. calcoaceticus* requires a further step in fact, because both benzyl alcohol dehydrogenase and benzaldehyde dehydrogenase II co-purify on the Blue-Sepharose CL-6B column and needed a Matrex Gel Red A column to be separated (MacKintosh & Fewson, 1988a). The specific activity of benzyl alcohol dehydrogenase purified from *E. coli* JM109 (DE3)/pDG30 (Fig. 4.2) was 328 units/mg protein. This value is higher than was previously obtained for benzyl alcohol dehydrogenase purified directly from *A. calcoaceticus* (238 units/mg protein; MacKintosh & Fewson, 1988a).

Figure 4.1 Time course following the expression of benzyl alcohol dehydrogenase in *E. coli* JM109 (DE3)/pDG30

A 50ml LB/tetracylcline culture of *E. coli* JM109 (DE3)/pDG30 was grown in 50 ml LB/tetracycline until the OD_{600} was approximately 0.4. Expression of benzyl alcohol dehydrogenase was then induced with 0.5mM IPTG and 1ml samples removed at various times. Samples were retained for assay and also for analysis by SDS-PAGE.

Samples for SDS-PAGE were pelleted by centrifugation in a microfuge for 5 min, resuspended in 100 μl gel loading buffer including 50mM DTT, boiled for 5 minutes and then 15 μl loaded onto a 5% SDS-PAGE stacking gel, 10% resolving gel (section 2.29).

Tracks: A and I, M_r markers (1 μg of each protein); B, *E. coli* JM109 (DE3)/pTB361 4 hours after induction with IPTG; C-H, *E. coli* JM109 (DE3)/pDG30 0, 1, 2, 3, 4 and 18 hours after induction with IPTG respectively. Tracks contain 75-150 μg total protein.

M_r x 10^{-3} A B C D E F G H I

94

67

43

30

20.1

14.4
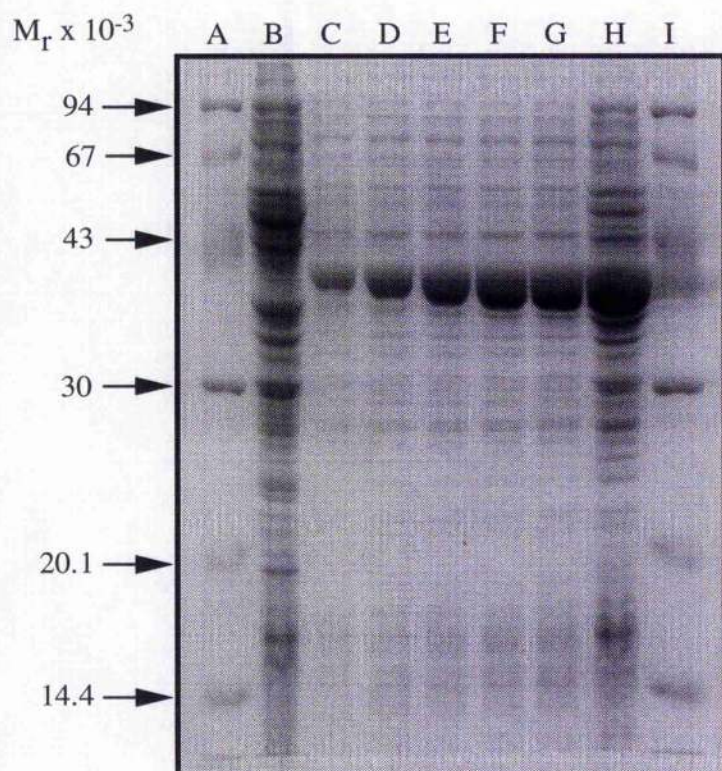
Figure 4.2 Purification of benzyl alcohol dehydrogenase from *E. coli* JM109 (DE3)/pDG30

Benzyl alcohol dehydrogenase was purified as described in section 2.27. Purification was monitored on a 10% (w/v) polyacrylamide slab gel containing 0.1% (w/v) SDS.

Tracks: A and E, M_r markers (1 μg of each); B, extract (13 μg protein); C, dialysed ion-exchange elution pool (6 μg protein); D, concentrated Blue-Sepharose CL-6B elution pool (3 μg protein).
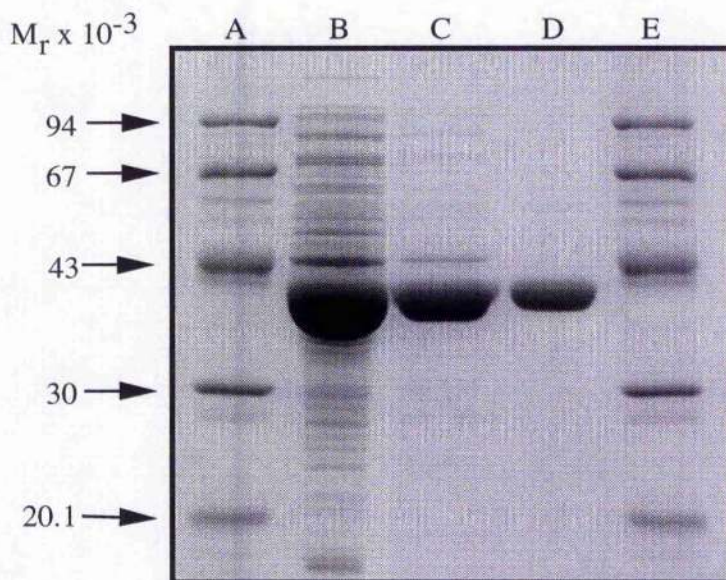
For benzyl alcohol dehydrogenase purifications 9 litres of cells were typically grown (section 2.24). Approximately 10 g of cells could be harvested 4 hours after induction with IPTG. This method of growth enabled expression of benzyl alcohol dehydrogenase at a level as high as 84% of total soluble protein as judged by laser scanning densitrometry (section 2.29; Fig. 4.2). This produced a crude extract with a specific activity of 303 U/mg total protein. These results show what a powerful tool the T7 expression system is for expressing large quantities of recombinant proteins.

Purification of benzyl alcohol dehydrogenase from *E. coli* JM109 (DE3)/pDG30 was carried out on two different scales. Small scale preps used 1g of cells to give a typical yield of 3-4 mg of pure protein. This was enough protein for kinetic studies but for protein modifications (section 2.32) much larger quantities were needed. Purifications were scaled up by using 10g of cells and increasing the volume of the DEAE-Sephacel ion-exchange column from 20 ml up to 250 ml (section 2.27). This latter step was necessary because the smaller volume was unable to bind all of the applied enzyme. Some problems were also encountered when trying to load larger quantities of protein onto the Blue-Sepharose CL-6B affinity column, although the capacity of the column was not exceeded (Pharmacia, 1992). When the original purification was devised, several factors were found to influence the binding of benzyl alcohol dehydrogenase to the affinity column. These were salt concentration, the prescence of Mg^{2+} ions, pH and temperature (MacKintosh, 1987). In addition, only a low concentration of NAD^+ is necessary to elute the enzyme from the column (0.12 mM). These factors suggest that the enzyme does not bind particularly strongly to the column. This was not a problem when small quantities of protein were loaded onto the column but when larger quantities were loaded some of the enzyme eluted into the loading and washing fractions. Larger purifications typically yielded 20-25 mg of pure enzyme but this could be increased by splitting the ion-exchange elution fraction into two samples and loading them separately onto a newly-regenerated Blue-Sepharose CL-6B column.

4.3 *N*-Terminal amino acid analysis of purified benzyl alcohol dehydrogenase

N-Terminal amino acid sequencing of purified benzyl alcohol dehydrogenase from *E. coli* DH5α/pDG20 was carried out using Edman degradation as described in section 2.30. This provided confirmation that the purified protein was indeed the product of *xylB*. Table 4.2 shows the ten amino acids sequenced: S E L K D I I A A V and T. These exactly matched the amino

Table 4.2 *N*-Terminal amino acid sequence analysis of purified benzyl alcohol dehydrogenase expressed in *E. coli* DH5α/pDG20

Cycle	Amino acid	Yield (pmol)
1	Ser	8.4
2	Glu	4.5
3	Leu	6.4
4	Lys	3.2
5	Asp	NQ
6	Ile	5.2
7	Ile	6.6
8	Ala	1.7
9	Ala	2.3
10	Val	3.5

NQ= not quantified.

Sequence was determined as described in section 2.30.

acids predicted from the gene sequence and those sequenced previously from the enzyme purified from *A. calcoaceticus* (Chalmers *et al.*, 1991). N-Terminal amino acid analysis confirmed that the N-terminal methionine residue is cleaved off in the mature protein.

4.4 Subunit M_r of benzyl alcohol dehydrogenase

Table 4.3 compares the subunit M_r values for benzyl alcohol dehydrogenase as determined by various methods. The M_r value obtained from the derived amino acid sequence was calculated by the GeneJockey sequence processor (Taylor, 1991).

Electrospray mass spectrometry is a very powerful technique for the accurate determination of molecular weight. Electrospray mass spectrometry was carried out as described in section 2.31 for benzyl alcohol dehydrogenase purified from *E. coli* DH5α/pDG20. The raw data output from this analysis is shown in Fig. 4.3. The value obtained, within the error range given, was the same as that calculated from the derived amino acid sequence (Table 4.3). This provides additional confirmation that the nucleotide sequence is correct. These values are in reasonable agreement with that obtained by SDS-PAGE of benzyl alcohol dehydrogenase purified from *A. calcoaceticus* (MacKintosh & Fewson, 1988a).

When the nucleotide sequence of *xylC* was determined prior to this study the derived amino acid sequence of benzaldehyde dehydrogenase II had a calculated M_r of 51,654 and this was confirmed by electrospray mass spectrometry of the purified enzyme (A. G. S. Robertson, unpublished results). However, the subunit M_r as determined by SDS-PAGE was 55,000. These results indicate that Mr determinations by SDS-PAGE should be treated with caution.

Table 4.3 Subunit M_r of benzyl alcohol dehydrogenase as determined by various methods

Method of benzyl alcohol dehydrogenase expression		
A.calcoaceticus NCIB8250	Heterologous expression in *E. coli*	
SDS-PAGE	Electrospray mass spectrometry	Derived from the amino acid sequence
39,700[a]	38,929 ± 8	38,923[b]

[a] MacKintosh & Fewson (1988a). Heterologous expression was *E. coli* DH5α/pDG20. Electrospray mass spectrometry was performed as described in section 2.31. [b] Calculated by the GeneJockey sequence processor (Taylor, 1991).

Figure 4.3 Electrospray mass spectrometer output for benzyl alcohol dehydrogenase purified from *E. coli* DH5α/pDG20

Sample was prepared, and the electrospray run, as described in section 2.31
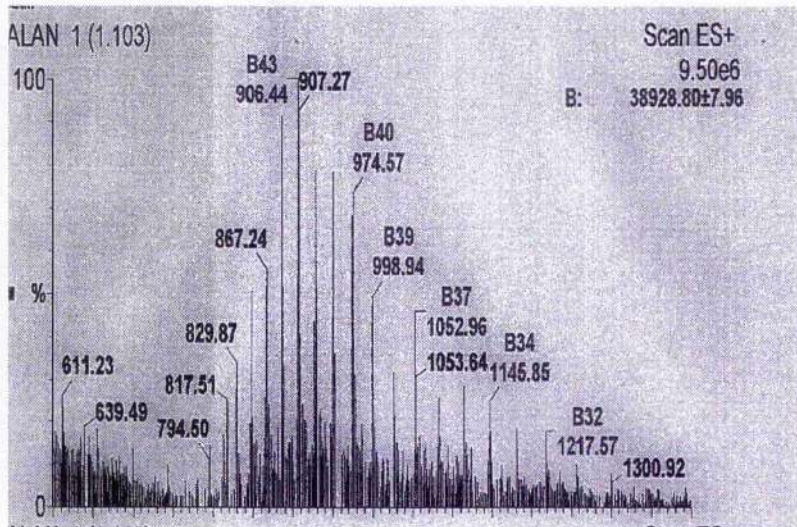
Chapter 5. Site-directed mutagenesis of *xylB* to put a "conserved" histidine residue into benzyl alcohol dehydrogenase

5.1 Site-directed mutagenesis of *xylB* to create benzyl alcohol dehydrogenase-His51

Site-directed mutagenesis was used to replace residue Ile-48 of benzyl alcohol dehydrogenase from *A. calcoaceticus* with a histidine residue. This residue corresponds to His-51 of horse liver alcohol dehydrogenase (Fig. 3.5), a residue that has been proposed as contributing to the proton relay system of the enzyme (Eklund *et al.*, 1982). This His residue is conserved across the majority of the group I long-chain zinc-dependent alcohol dehydrogenases (Fig. 3.5). Notable exceptions to this are the benzyl alcohol dehydrogenases of *A. calcoaceticus* and the TOL-plasmid of *P. putida*, the *xylW* product which as yet has no known function, and also the alcohol dehydrogenases of the plants barley, maize, rice, potato and peas. However, the other components of the proposed proton relay system (Eklund *et al.*, 1982) are still conserved in these enzymes. The plant enzymes have a tyrosine in the His-51 position, a residue which could fulfill the proposed role of the His-51 residue, whereas benzyl alcohol dehydrogenase from *A. calcoaceticus* and TOL-benzyl alcohol dehydrogenase have an isoleucine and a valine respectively in this position. These residues would not be able to hydrogen bond in the manner proposed by Eklund *et al.* (1982) for His-51. In order to study possible reasons for the lack of this His residue the effects of putting this "conserved" His residue into benzyl alcohol dehydrogenase of *A. calcoaceticus* were studied.

Since horse liver alcohol dehydrogenase has been used as a paradigm for the study of other long-chain zinc-dependent alcohol dehydrogenases (section 1.10.2) workers have tended to number residues in zinc-dependent alcohol dehydrogenases according to the numbering of horse liver alcohol dehydrogenase. For this reason, unless otherwise stated, in this thesis benzyl alcohol dehydrogenase residues will be numbered according to the numbering of horse liver alcohol dehydrogenase using the alignment in Fig. 3.5. Hence benzyl alcohol dehydrogenase mutants will be named with the mutated residue numbered according to the analogous horse liver enzyme residue.

Site-directed mutagenesis was performed using the method of Higuchi *at al.* (1988). This method is outlined in Fig. 5.1 and involves the use of PCR with *Vent* polymerase as described in section 2.23.2. The primers used to make the required mutagenic changes are shown in Fig. 5.2. These primers were used in conjunction with the "outer" primers (oligos 2166 and 2168) that were used to create pDG30 (Fig. 3.1). For the primary PCR, oligos 2166 and 2254 were used for the "left" or "N-terminal" reaction and oligos 2253 and 2168 were used for the "right" or "C-terminal" reaction (Fig. 5.1). These two reactions yielded products

Figure 5.1 Site-directed mutagenesis strategy

PCR was performed as described in section 2.23.2.
Mutagenic oligonucleotides are shown in Fig. 5.2.
Outer primers are shown in Fig. 3.1. The strategy
is based on Higuchi *et al.*(1988).
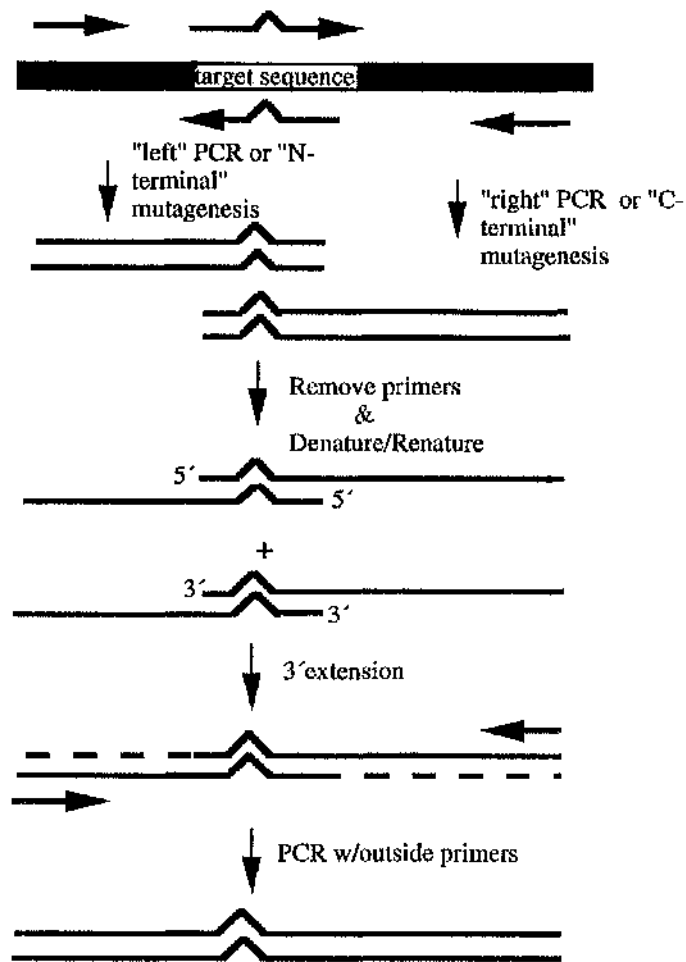
Figure 5.2 Design of mutagenic oligonucleotides to create benzyl alcohol dehydrogenase-His51 mutant (pDG40)

Residue Ile-51 was changed to His51 by the PCR method described in section 2.23.2 and Fig. 5.1 using the oligonucleotides shown below. The mutated *xyl* B product was cloned into pTB361 (section 2.2) to create pDG40.

Wild-type sequence   5´-TGCCACACCGATTTG**ATT**GTACGTGATCAATA-3´
                              C   H   T   D   L   I   V   R   D   Q   Y

Forward direction   5´-TGCCACACCGATTTG**CAT**GTACGTGATCAATA-3´
(Oligo 2253)   C   H   T   D   L   H   V   R   D   Q   Y

Reverse direction   5´-TATTGATCACGTACA**TG**CAAATCGGTGTGGCA-3´
(Oligo 2254)

of 170bp and 996bp respectively (Fig. 5.3) using pDG30/*XbaI* as template. The products of these reactions were purified using a low melting point agarose gel (section 2.10) followed by phenol/chloroform extraction (section 2.11) and were then used as the template for secondary PCR with the two "outer" primers. Secondary PCR produced a single product of 1137bp (Fig. 5.3; *xylB* plus restriction sites and overhangs) which could then be digested with *NdeI* and *BglII* (section 2.4) and cloned into pTB361 (section 2.2) to create pDG40.

The clone pDG40 was sequenced with the primers used to sequence pDG20 (section 2.19, Fig 3.2). Fig. 5.4 shows that the two necessary base changes were made in order to change residue Ile-51 to a His residue. In addition, the whole of the mutated *xylB* gene was sequenced and this confirmed that no PCR errors had been introduced.

5.2 Expression and purification of benzyl alcohol dehydrogenase-His51 (BADH-His51).

The clone harbouring mutated *xylB*, pDG40, was used to transform *E.coli* JM109 (DE3). BADH-His51 could be expressed at a level comparable to that of *E. coli* JM109 (DE3)/pDG30 as judged by laser scanning densitrometry of SDS-PAGE (section 2.29). The mutant enzyme was purified by the same procedure as the wild-type enzyme (section 2.27; Fig. 5.5).

5.3 Kinetic coefficients of BADH-His51 and their comparison with wild-type benzyl alcohol dehydrogenase

The kinetic coefficients of wild-type benzyl alcohol dehydrogenase and BADH-His51 for the oxidation of benzyl alcohol are recorded in Table 5.1. A range of substrate concentrations from 10 μM to 400 μM was used to give initial velocities which were then analysed to determine these coefficients (section 2.26.6). The turnover number, k_{cat}, was calculated from V'_{max} using the assumption that there is one active site per subunit and using the subunit M_r determined from the derived amino acid sequence (38,923). This value could then be used to calculate the "specificity constant", k_{cat}/K'_m This constant combines the rate of coversion of substrate to product with the binding of substrate to enzyme and thus determines the specificity towards competing substrates. A high k_{cat}/k_m' value suggests that there is a high degree of complementarity between the enzyme and the transition-state analogue of the particular substrate (Fersht, 1985).

Figure 5.3 Site-directed mutagenesis using PCR with *Vent* polymerase to create benzyl alcohol dehydrogenase-His51

Panel A. Primary PCR reactions were run as described in section 2.23.2 and then 20 μl of each 100 μl reaction was run in a 0.75% agarose gel (section 2.9). pDG30/Xba I (100ng) was used a template in each reaction.
Tracks: 1 and 7, λ/*Hind* III markers (250ng); 2, pDG30/*Xba* I (100ng); 3, no template control (oligos 2166 & 2168); 4, template control (i.e. both "outer" primers but no mutagenic primers); 5, "left" (Fig. 5.1) or "N-terminal" primary reaction (oligos 2166 & 2254); 6, "right" (Fig. 5.1) or "C-terminal" primary reaction (oligos 2253 & 2168).

Panel B. Secondary PCR reactions were run as described in section 2.23.2 and then 20 μl of each 100 μl reaction was run in a 0.75% agarose gel (section 2.9). Tracks: 1, λ/Hind III markers (250ng); 2, positive control (i.e. "outer" primers used to amplify non-mutagenised template); 3, mutagenesis secondary reaction (oligos 2166 & 2168).

Marker sizes indicated are in base pairs.

A.

B.

125
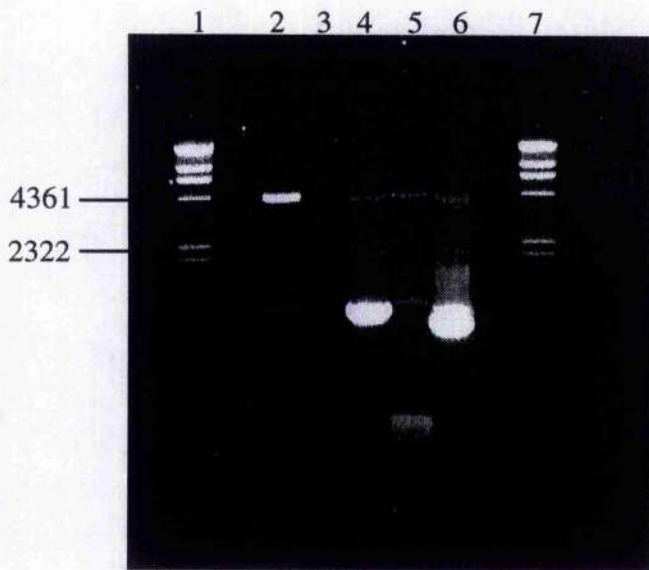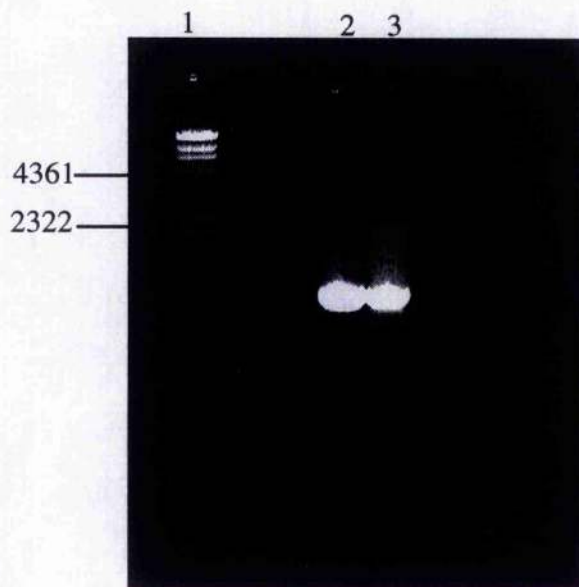
Figure 5.4 DNA sequence autoradiographs of pDG20 and pDG40 showing the mutagenic changes made to create benzyl alcohol dehydrogenase-His51

Nucleotide sequence was determined as described in section 2.19.

Panel A: Wild-type *xyl* B sequence (pDG20).
Panel B: Mutagenised *xyl* B sequence (pDG40).
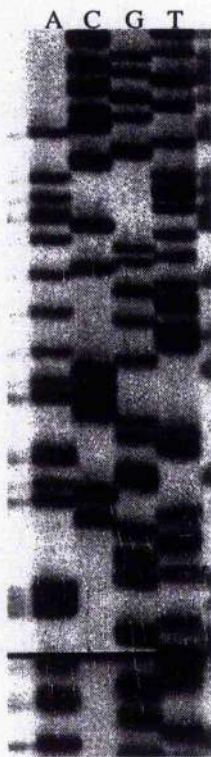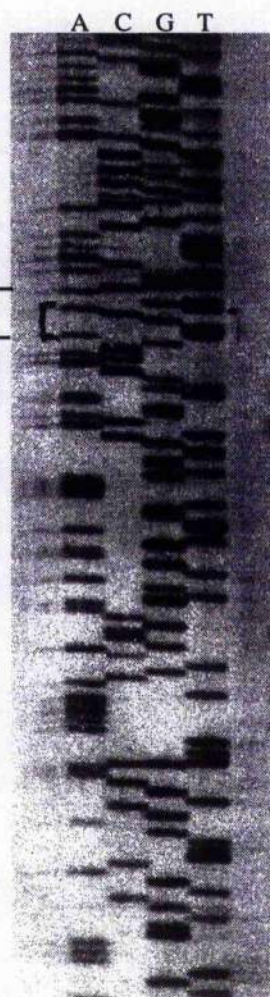
Figure 5.5 Purification of benzyl alcohol dehydrogenase-His51 mutant (BADH-His51) from *E. coli* JM109 (DE3)/pDG40

BADH-His51 was purified as described in section 2.27. Purification was monitored on a 10% (w/v) polyacrylamide slab gel containing 0.1% (w/v) SDS. Tracks: A and E, M_r markers (1 µg of each); B, crude extract (18 µg protein); C, dialysed ion-exchange elution pool (0.5 µg protein); D, concentrated Blue-Sepharose CL-6B elution pool (2.5 µg protein).
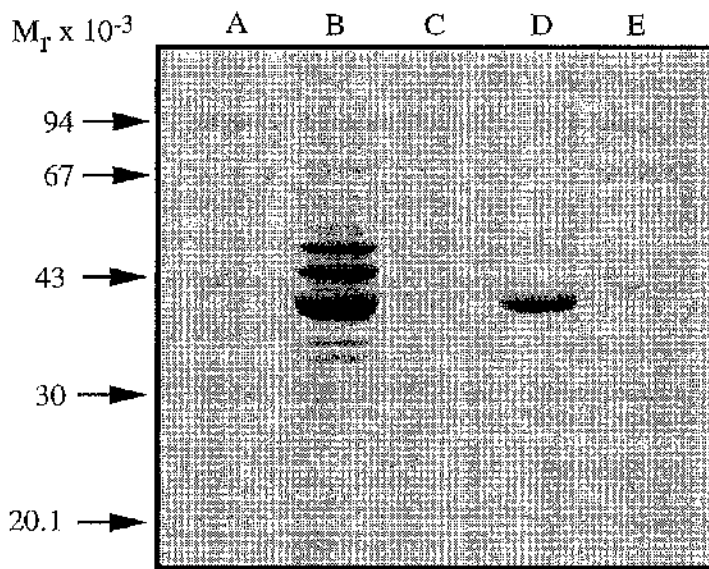
Table 5.1 Comparison of the kinetic coefficients of wild-type benzyl alcohol dehydrogenase and benzyl alcohol dehydrogenase-His51

	Kinetic coefficients for benzyl alcohol			
Enzyme	**Apparent K_m (μM)**	**Apparent V_{max} (Units/mg protein)**	**k_{cat} (s^{-1})**	**Specificity constant k_{cat}/K'_m (s^{-1} μM^{-1})**
BADH	19 ($\pm$4)	81 (84,78)	53	2.79
BADH-His51	44 (53,35)	211 (190,231)	137	3.11

Values were determined as described in section 2.26. Values in parentheses correspond to standard error mean values (n=3 separate experiments) or the results of entirely separate duplicate experiments. Within each experiment all assays were done in triplicate.

The kinetic coefficients values determined for recombinant wild-type benzyl alcohol dehydrogenase (Table 5.1) are in general agreement with those previously determined for BADH purified directly from *A. calcoaceticus* (MacKintosh & Fewson, 1988b). All of the kinetic coefficients determined for BADH-His51 are somewhat greater than the values obtained for the wild-type enzyme. K'_m, V'_{max} and k_{cat} are all about 2.5 times greater but consequently the k_{cat}/K'_m values are similar. These results suggest that the insertion of a histidine residue into benzyl alcohol dehydrogenase at position 51, and so reconstituting the proton relay system of the long-chain zinc-dependent alcohol dehydrogenases, has had little effect on the overall effectiveness of the enzyme at oxidising benzyl alcohol.

5.4 The activity of benzyl alcohol dehydrogenase and benzyl alcohol dehyrogenase-His51 with various alcohol substrates relative to benzyl alcohol

Table 5.2 records the activities of benzyl alcohol dehydrogenase wild-type and the His-51 mutant with a range of substrates relative to their activities with benzyl alcohol. MacKintosh & Fewson (1988b) showed that benzyl alcohol dehydrogenase of *A. calcoaceticus* has a wide substrate range, primarily against aromatic alcohols (section 1.9; Table 1.2). It is possible that benzyl alcohol is not the "natural" or "original" substrate for the enzyme and the selection against His-48 in the active site as part of the proton relay system was necessary to allow oxidation of this substrate. The substrates utilised in Table 5.2 are all possible natural substrates for the enzyme but the results suggest that the inclusion of His-48 has not altered the enzyme's ability to oxidise any of them. In addition the mutant enzyme had not gained the ability to oxidise the short chain aliphatic alcohols methanol and ethanol.

5.5 Effect of pH on the activity of wild-type benzyl alcohol dehydrogenase and benzyl alcohol dehydrogenase-His51

The effect of pH on the activity of both wild-type benzyl alcohol dehydrogenase and the His-51 mutant is shown in Fig. 5.6. An increase of 1.0 pH unit from pH9.2 to pH10.2 completely abolished activity by both of the enzymes, whereas a decrease of the same amount from pH9.2 down to pH8.2 resulted in a 32-40% loss of activity by both of the enzymes. Since the two enzymes behaved in a similar manner it seems likely that the insertion of His-51 has not

Table 5.2 Relative activities of benzyl alcohol deydrogenase and benzyl alcohol dehydrogenase-His51 with various alcohol substrates compared to benzyl alcohol

Alcohol substrate	% activity relative to benzyl alcohol	
	BADH	BADH-His51
Benzyl alcohol (I)	100%	100%
2-Methylbenzyl alcohol (II)	34(36,32)	39(40,37)
3-Methylbenzyl alcohol (III)	82(93,71)	74(73,74)
4-Methylbenzyl alcohol (IV)	97(100,93)	99(93,105)
Perillyl alcohol (VII)	55(64,46)	50(47,53)
Cinnamyl alcohol (VI)	48(49,46)	47(45,48)
Coniferyl alcohol (V)	8(9,6)	12(9,15)

In addition, neither enzyme was found to oxidise methanol or ethanol at detectable rates.

Assays were performed as described in section 2.26. Substrate concentration was fixed at 200 μM. Values in parentheses are for two separate experiments. Within each experiment each assay was carried out in triplicate.

Specific activities with benzyl alcohol = 320 and 190 units/mg protein for BADH and BADH-His51 respectively.

Roman numerals refer to structures of these compounds in appendix.

Figure 5.6 Effect of pH on the activity of benzyl alcohol dehydrogenase-His51 compared with the effect on wild-type benzyl alcohol dehydrogenase

Benzyl alcohol dehydrogenase and the -His51 mutant enzyme were purified from *E. coli* JM109 (DE3)/pDG30 or pDG40, respectively, as described in section 2.27. Enzyme activity was determined as described in section 2.26. The assay buffer used at pH 8.2 and pH 9.2 was 150mM Bicine, 540mM hydrazine adjusted to the required pH with sodium hydroxide. At pH 10.2 50mM CAPS/NaOH and 0.1M glycine/NaOH were both used. Neither enzyme was found to be active in either of the latter two buffers at pH 10.2. All enzyme assays were carried out in triplicate.
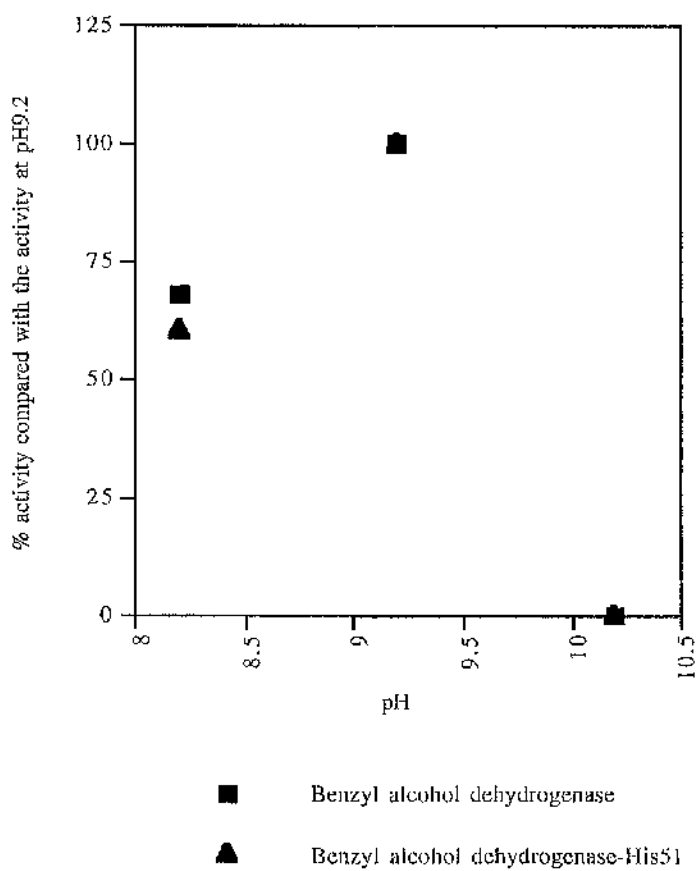
% activity compared with the activity at pH9.2

pH

■ Benzyl alcohol dehydrogenase

▲ Benzyl alcohol dehydrogenase-His51

significantly altered the pK_a of any proton release group involved in benzyl alcohol oxidation.

5.6 Comparison of wild-type benzyl alcohol dehydrogenase and BADH-His51 kinetic coefficients with those of baker's yeast and horse liver alcohol dehydrogenase

The kinetic coefficients of baker's yeast and horse liver alcohol dehydrogenases for the oxidation of ethanol are recorded in Table 5.3. A range of ethanol concentrations from 0.25 mM to 25 mM for yeast alcohol dehydrogenase, and from 0.2 mM to 4 mM for horse liver alcohol dehydrogenase, were used to obtain initial velocities which were then analysed as described in section 2.26.6 to determine the coefficient values. For the horse liver enzyme kinetic coefficients were also determined for benzyl alcohol oxidation (Table 5.3). A range of benzyl alcohol concentrations from 10 μM to 600 μM were used in the horse liver alcohol dehydrogenase assay described in section 2.26.4 and initial velocities analysed as described in section 2.26.6. Yeast alcohol dehydrogenase was found not to oxidise benzyl alcohol at detectable rates.

Yeast alcohol dehydrogenase has a K'_m for ethanol 3.6 times larger than does horse liver alcohol dehydrogenase but has a V'_{max} greater than 60 times that of the liver enzyme. Consequently yeast alcohol dehydrogenase has a "specificity constant" for ethanol oxidation which is almost 16 times higher than the value for horse liver alcohol dehydrogenase. These "specificity constants" indicate that the yeast enzyme is much more efficient at oxidising ethanol than the horse liver enzyme and that the active site of the yeast enzyme has a structure more similar to that of ethanol than that of the horse liver enzyme (section 9.4.3). Horse liver alcohol dehydrogenase was found to have a K'_m for benzyl alcohol oxidation similar to that of benzyl alcohol dehydrogenase, although the V'_{max} value is much lower (Table 5.3). The "specificity constant" for benzyl alcohol oxidation by horse liver alcohol dehydrogenase in 5.5 times higher its "specificity constant" for ethanol oxidation (Table 5.3). This suggests that the active site of the horse liver enzyme has a structure that more closely resembles the structure of benzyl alcohol than the structure of ethanol (section 9.3.4).

If the "specificty constants" for yeast and horse liver alcohol dehydrogenases are compared to those of wild-type benzyl alcohol dehydrogenase and BADH-His51 (Table 5.3), the latter of which now has the proton relay system of yeast and horse liver alcohol dehydrogenases restored, it is apparent that both the benzyl alcohol dehydrogenases are very much more "efficient" alcohol

Table 5.3 Comparison of kinetic coefficients of the alcohol dehydrogenases from baker's yeast and horse liver with those of benzyl alcohol dehydrogenase and the mutant benzyl alcohol dehydrogenase-His51

Enzyme	Substrate	Kinetic coefficients			
		Apparent K_m (μM)	Apparent V_{max} (units/mg protein)	k_{cat} (s^{-1})	Specificity constant k_{cat}/K'_m (s^{-1} mM^{-1})
Yeast ADH	Ethanol	4700	124	75.00	15.96
	Benzyl alcohol	NA	NA	NA	NA
HLADH	Ethanol	1300	2	1.34	1.03
	Benzyl alcohol	15	0.12	0.08	5.53
BADH[a]	Benzyl alcohol	19	81	53.00	2.79×10^3
BADH-His51[a]	Benzyl alcohol	44	211	137.00	3.11×10^3

NA = No detectable activity.

Values were determined as described in section 2.26. Within each experiment all assays were performed in triplicate. Values of k_{cat} were calculated with the assumption of one active site per subunit and subunit molecular weights of 36,250 for yeast alcohol dehydrogenase (Brändén et al., 1975) and 40,000 for horse liver alcohol dehydrogenase (HLADH; Jörnvall, 1970).

[a]Values taken from Table 5.1.

dehydrogenases (when oxidising their principle substrates) than yeast or horse liver alcohol dehydrogenase. The differences in "specificity constants" arise because of the very high K'_m values for the yeast and horse liver alcohol dehydrogenases, and also much lower V'_{max} values in the case of the mammalian enzyme. These differences in kinetic coefficients may be reflected in the respective physiological roles of the enzymes. Benzyl alcohol dehydrogenase has the role of feeding benzyl alcohol into central metabolism as a carbon and energy source for growth of the microorganism. Concentrations of benzyl alcohol encountered in the environment will be low so it is necessary for benzyl alcohol dehydrogenase to have a low K'_m for benzyl alcohol to enable high growth rates. In contrast fermentative baker's yeast alcohol dehydrogenase converts acetaldehyde that accumulates during anaerobic metabolism into ethanol in order to regenerate NAD^+ to enable other cellular processes to continue (Clark, 1992) and horse liver alcohol dehydrogenase removes high levels of alcohols circulating in the blood stream. In these latter two cases the alcohol dehydrogenases involved are required to work only at raised alcohol levels, and for these reasons it is likely that these enzymes have relatively high K'_m values for ethanol (section 9.4.3).

Chapter 6. Is a histidine residue involved in the catalytic mechanism of benzyl alcohol dehydrogenase from *A. calcoaceticus*? Inactivation studies with diethyl pyrocarbonate

6.1 Modification of benzyl alcohol dehydrogenase from *A. calcoaceticus* with diethyl pyrocarbonate (DEPC)

Although benzyl alcohol dehydrogenase from *A. calcoaceticus* does not have the histidine residue (His-51 in horse liver alcohol dehydrogenase) that is conserved in the proton relay system (sections 1.12 & 3.4; Fig. 3.5) of the majority of other group I zinc-dependent alcohol dehydrogenases it was still a possibilty that in this enzyme the proton relay system might be rearranged due to differences in the three-dimensional structure between benzyl alcohol dehydrogenase and the other enzymes in such a way that a different histidine residue is involved. This would not necessarily be apparent from primary sequence alignments.

To investigate the possible involvement of a histidine residue in the catalytic mechanism of benzyl alcohol dehydrogenase, histidine residues were modified with diethyl pyrocarbonate and modifications correlated with any enzyme inactivation. Diethyl pyrocarbonate reacts with histidine residues such that there is a substitution at one of the nitrogen atoms in the imidazole ring to yield an *N*-carbethoxyhistidyl derivative with an increased absorbance between 230 and 250 nm (equation below; Miles, 1977). It is therefore possible to calculate the number of histidine residues modified per enzyme subunit from the increase in absorbance at 242 nm.

230-250nm

Histidyl residue   *N*-carbethoxyhistidyl derivative
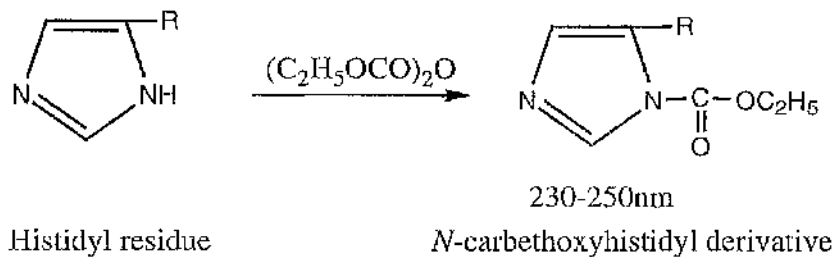
Further evidence that a histidine residue is a catalytic group, or is located at the substrate binding site, can be obtained from substrate protection experiments in which a substrate is incubated with the enzyme during reaction with diethyl pyrocarbonate. If the histidine residue is a catalytic group, or is located at the active site, then the presence of the substrate should prevent diethyl pyrocarbonate from modifying the histidine residue and consequently prevent inactivation of the enzyme.

Inactivation experiments using diethyl pyrocarbonate were carried out with benzyl alcohol dehydrogenase purified from *E. coli* JM109(DE3)/pDG30 (section 2.27), as described by Hennecke & Plapp (1983) for the inactivation of

horse liver alcohol dehydrogenase (section 2.32). The inactivation of horse liver alcohol dehydrogenase with diethyl pyrocarbonate was repeated as a positive control experiment (section 6.2). A diethyl pyrocarbonate concentration of 1 mM was used for the inactivation of benzyl alcohol dehydrogenase as well as horse liver alcohol dehydrogenase. NADH, NAD^+ and benzyl alcohol were substrates used to protect against carbethoxylation. These substrates were all used at a concentration of 200 µM; at least 5 times their respective K_m values (MacKintosh & Fewson, 1988a). Benzaldehyde, which is also a substrate for the enzyme was not used since this compound can have toxic effects on enzymes (MacKintosh & Fewson, 1988a, b) and so could confuse any interpretations of results.

6.2 Modification of horse liver alcohol dehydrogenase with diethyl pyrocarbonate

The modification of horse liver alcohol dehydrogenase with diethyl pyrocarbonate as described by Hennecke and Plapp (1983) was used as a positive control experiment. Horse liver alcohol dehydrogenase was acetimidylated prior to modification with diethyl pyrocarbonate (section 2.32.2). Acetimidylation prevents the modification of lysine residues by diethyl pyrocarbonate and was necessary since modification of lysine-228 with imido esters or isocyanantes has been shown to activate horse liver alcohol dehydrogenase (Plapp, 1970; Zoltobrocki et al., 1974; Sogin & Plapp, 1975) and consequently complicate inactivation experiments (Morris & McKinley-McKee, 1972). This residue has been replaced by an arginine residue in benzyl alcohol dehydrogenase (Fig. 3.5).

Acetimidylated horse liver alcohol dehydrogenase was incubated with 1 mM diethyl pyrocarbonate, the absorbance at 242 nm monitored and samples were removed at various times for the measurement of residual enzyme activity (section 2.32.2). Fig. 6.1 plots the increase in absorbance at 242 nm and the decrease in enzyme activity against time respectively. It can be seen that an increase in absorbance at 242 nm corresponding to the carbethoxylation of histidine residues corresponds with a loss of alcohol dehydrogenase activity.

The number of histidine residues modified per enzyme subunit was calculated by assuming the subunit M_r of horse liver alcohol dehydrognase to be 40,000 (Jörnvall, 1970). Fig. 6.2 plots the percentage of inital enzyme activity remaining at various time points against the number of histidine residues modified. Complete enzyme activity was lost when three histidine residues had been modified. A total of 4.5 histidine residues could be modified, after which the

140

Figure 6.1

a. Increase in absorbance at 242 nm with time during the incubation of horse liver alcohol dehydrogenase with diethylpyrocarbonate

Acetimidylated horse liver alcohol dehydrogenase was incubated with 1 mM diethylpyrocarbonate and the absorbance at 242 nm monitored as described in section 2.32.2.

b Inhibition of activity during the incubation of horse liver alcohol dehydrogenase with diethylpyrocarbonate

During incubation of acetimidylated horse liver alchohol dehydrogenase with 1 mM diethylpyrocarbonate (Panel A) samples were removed at various times (section 2.32.2) for assay as described in section 2.26.4.

All values are taken from duplicate experiments. Standard error mean (SEM) values were less than 5%.

a.

b.

Figure 6.2 Correlation between the percentage of initial alcohol dehydrogenase activity remaining and the number of histidine residues modified per enzyme subunit during the incubation of horse liver alcohol dehydrogenase with diethylpyrocarbonate

Acetimidylated horse liver alcohol dehydrogenase was incubated with 1 mM diethylpyrocarbonate, the absorbance at 242 nm monitored and samples removed (section 2.32.2) at various times for the measurement of residual activity (section 2.26.4). These are the same data as in Fig. 6.1. The number of histidine residues modified per enzyme subunit was calculated by taking the absorption coefficient of N-carbethoxyhistidine to be 3200 M^{-1} cm^{-1} at 242 nm (Ovádi *et al.*, 1967) and the subunit M_r of horse liver alcohol dehydrogenase to be 40,000 (Hennecke & Plapp, 1983).

Histidine residues modified
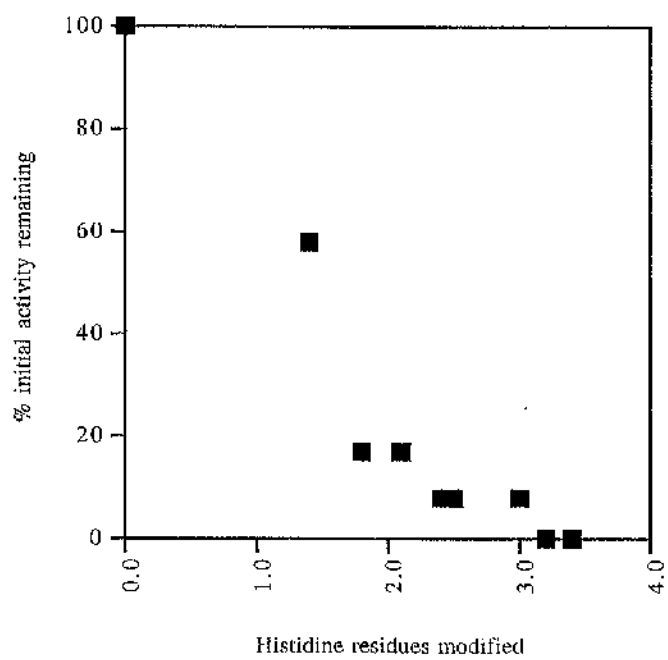
enzyme began to denature and precipitate. All of these results are in excellent agreement with those of Hennecke and Plapp (1983).

6.3 Correlation of histidine modifications with inactivation of benzyl alcohol dehydrogenase

Fig. 6.3a shows that when incubated with 1 mM diethyl pyrocarbonate (section 2.32.2) the absorbance of benzyl alcohol dehydrogenase at 242 nm increased dramatically. This suggests that histidine residues were being modified to N-carbethoxyhistidine derivatives. A maximum increase of 0.35 absorbance units was obtained, which corresponds to the modification of 4.2 histidines per enzyme subunit (out of a possible 8). This calculation used the assumption that the subunit M_r of benzyl alcohol dehydrogenase was 38,923 (Table 4.3) and from this value µM concentrations of enzyme subunits were calculated. The results correspond well with those of Hennecke & Plapp (1983) in which 4.8 histidine residues per subunit of horse liver alcohol dehydrogenase could be modified.

It is apparent from Fig. 6.3a that when benzyl alcohol dehydrogenase was incubated with diethyl pyrocarbonate in the presence of NADH, NAD$^+$ or benzyl alcohol the number of histidines modified was reduced. A maximum of 2.7 histidines could be modified in the presence of any of the substrates. These results suggest that the substrates are protecting against the reaction of one or more histidine residues with diethyl pyrocarbonate (section 9.5.1).

During the incubation of benzyl alcohol dehydrogenase with 1 mM diethyl pyrocarbonate samples were removed at various times for the determination of residual enzyme activity (section 2.32.2). Fig. 6.3b plots the percentage of the initial benzyl alcohol dehydrogenase activity remaining at each time point. It can be seen that benzyl alcohol dehydrogenase activity decreases as the absorbance at 242 nm increases (Fig. 6.3a) with a maximum loss of 75% of enzyme activity. This suggests that the decrease in enzyme activity is due to the modification of a particular histidine residue or several histidine residues (section 9.5.1). When benzyl alcohol dehydrogenase was incubated with diethyl pyrocarbonate in the presence of substrates, the amount of enzyme activity lost over the time of the incubation decreased (Fig. 6.3b). NADH gave the greatest protection against inactivation, followed by NAD$^+$. Benzyl alcohol provided the least protection; in fact after five minutes incubation, the average extent of inactivation in the presence of benzyl alcohol was almost identical to incubations without any substrates present at all. The weaker substrate protection provided by benzyl alcohol suggests that *direct* proton transfer to a histidine residue, as is the
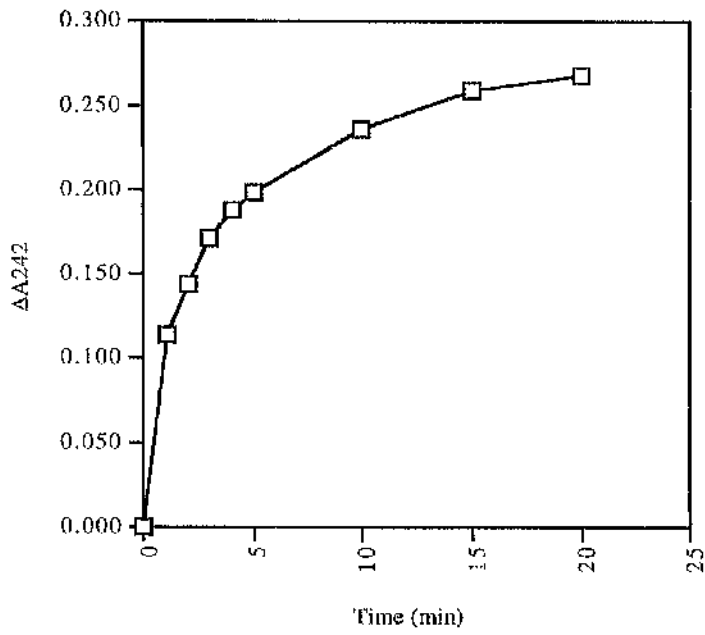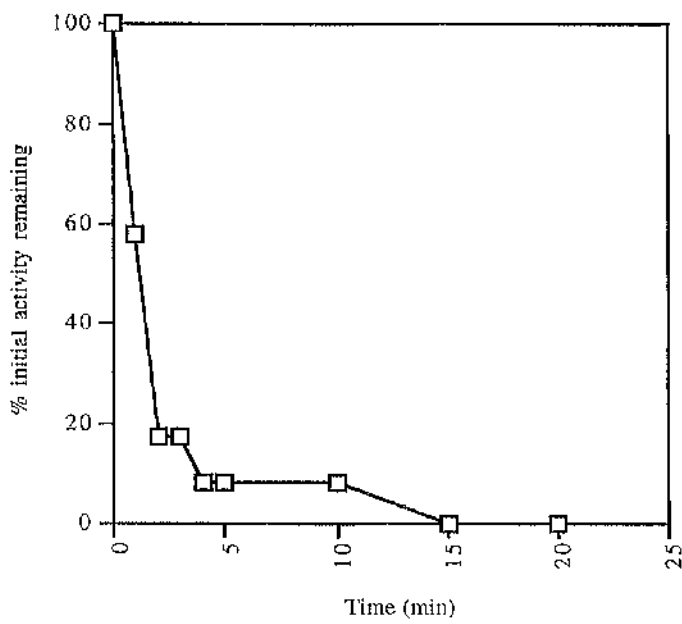
Figure 6.3

a Increase in absorbance at 242 nm with time during the incubation of benzyl alcohol dehydrogenase with diethylpyrocarbonate in the presence and absence of substrates

Benzyl alcohol dehydrogenase was incubated with 1 mM diethylpyrocarbonate in the presence and absence of various substrates and the absorbance at 242 nm was monitored as described in section 2.32.2.

b. Inhibition of activity during incubation of benzyl alcohol dehydrogenase with diethylpyrocarbonate in the presence and absence of substrates

During incubation of benzyl alcohol dehydrogenase with 1 mM diethylpyrocarbonate in the presence and absence of various substrates (Panel A) samples were removed at various times (section 2.32.2) for the measurement of residual activity (section 2.26.1).

All values are averages of the following number of separate experiments: no substrate, 2; NADH, 4; benzyl alcohol, 3; NAD^+, 2. SEM values were less than 10%.

a.

b.

Legend:

—□— 1 mM DEPC	---○--- 1 mM DEPC + 0.2 mM benzyl alcohol
·····◇····· 1 mM DEPC + 0.2 mM NADH	---△--- 1 mM DEPC + 0.2 mM NAD$^+$
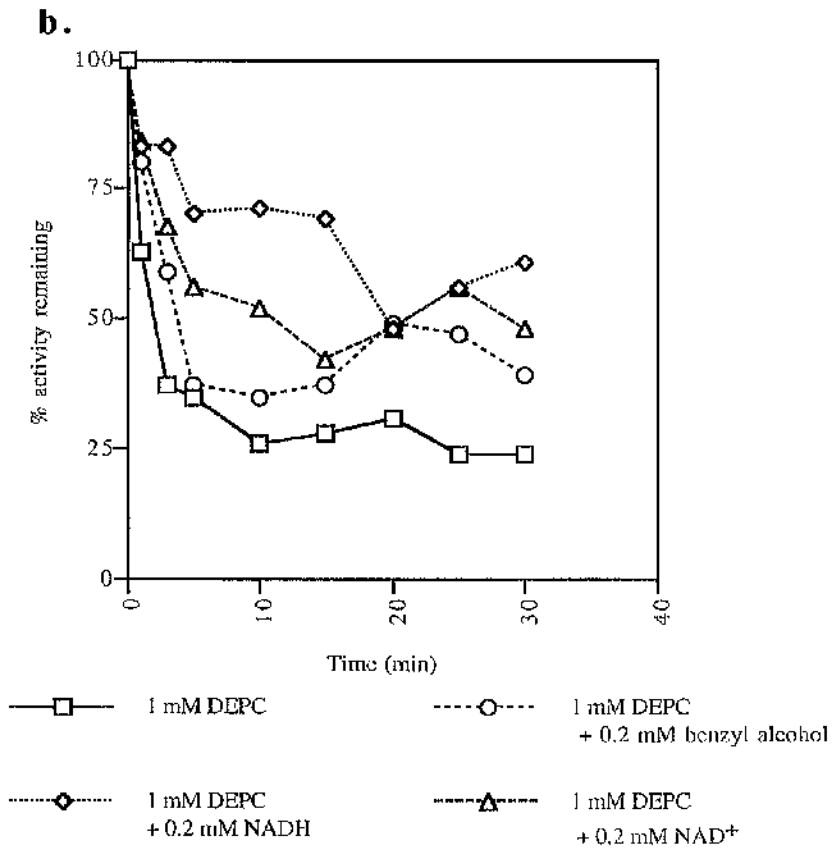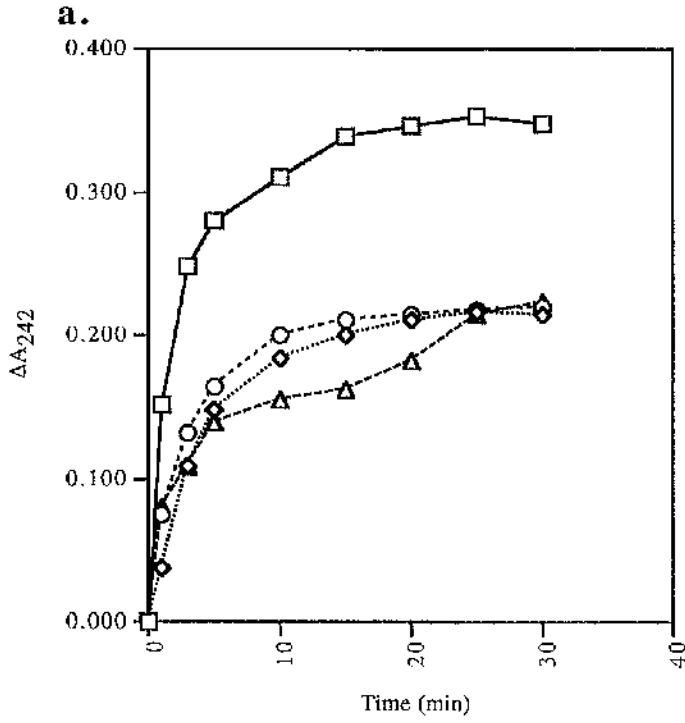
case for lactate dehydrogenase (Holbrook & Ingram, 1973; Holbrook *et al.*, 1975), does not occur in benzyl alcohol dehydrogenase (section 9.5.1). These results suggest that an active site or a coenzyme binding site histidine residue (or residues) is being modifed by diethyl pyrocarbonate and that this modification consequently inactivates the enzyme, although it is possible that carbethoxylation of His-47 and/or His-67, residues involved in coenzyme binding and catalytic zinc ligation respectively, is responsible for this inactivation and not carbethoxylation of a histidine residue involved in proton transfer in benzyl alcohol dehydrogenase (section 9.5.1).

Fig. 6.4 plots the percentage of benzyl alcohol dehydrogenase activity remaining at various times during the incubation of benzyl alcohol dehydrogenase with diethyl pyrocarbonate, against the number of histidine residues calculated as being modified from the increase in absorbance at 242 nm (Fig. 6.3a). The data for all of the conditions follow a linear pattern, with activity decreasing as the number of histidine residues modified increases. The results show that some activity is lost after the modification of one histidine residue, although it seems that two histidine residues need to be modified for there to be a loss of more than 50% of enzyme activity. In contrast to this, when two hisidine residues of horse liver alcohol dehydrogenase had been modified, at least 80% of enzyme activity had been lost (Fig. 6.4). The modification of three histidine residues in horse liver alcohol dehydrogenase resulted in the complete inactivation of the enzyme, whereas complete inhibition of benzyl alcohol dehydrogenase could not be obtained (Fig. 6.4).

6.4 Decarbethoxylation of modified benzyl alcohol dehydogenase

Benzyl alcohol dehydrogenase that had been modified with diethyl pyrocarbonate was incubated with 0.5 M hydroxylamine for 23 hours before reassaying for the return of enzyme activity (section 2.32.2). Hydroxylamine removes the carbethoxy group group from N-carbethoxyhistidine residues (Miles, 1977). After incubation with hydroxylamine 73% ± 4% (replicate of four experiments) of the initial benzyl alcohol dehydrogenase activity had returned. This is further evidence that enzyme activity had been lost due to the modification of a histidine residue(s) (section 9.5.1).

Figure 6.4 Correlation between the percentage of initial benzyl alcohol dehydrogenase activity remaining and the number of histidine residues modified per enzyme subunit during the incubation of benzyl alcohol dehydrogenase with diethylpyrocarbonate in the presence and absence of substrates

Benzyl alcohol dehydrogenase was incubated with 1 mM diethylpyrocarbonate in the presence and absence of various substrates, the absorbance at 242 nm monitored and samples removed at various times (section 2.32.2) for the measurement of residual activity (section 2.26.1). These are the same data as in Fig. 6.3. The number of histidine residues modified per enzyme subunit was calculated by taking the absorption coefficient of N-carbethoxyhistidine to be 3200 M^{-1} cm^{-1} at 242 nm (Ovádi et $al.$, 1967) and the subunit M_r of benzyl alcohol dehydrogenase to be 38,923 (Table 4.3).

% activity remaining

Histidines Modified
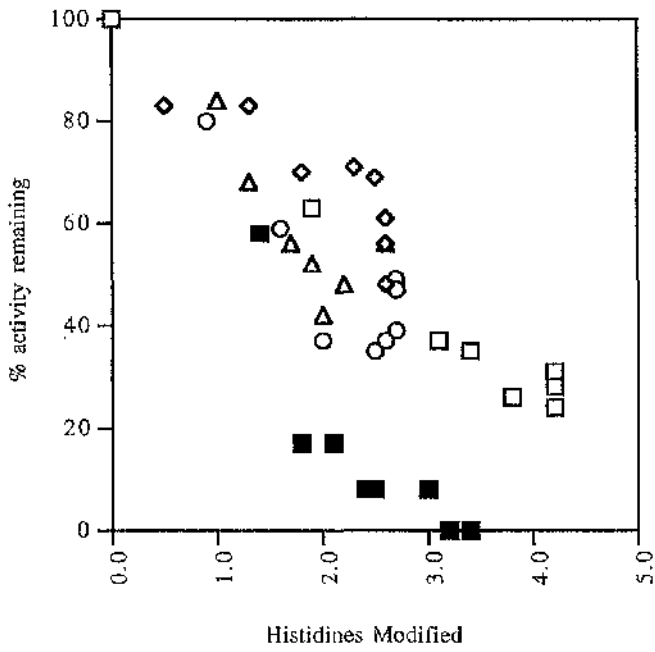
□	1 mM DEPC
◇	1 mM DEPC + 0.2 mM NADH
○	1 mM DEPC + 0.2 mM benzyl alcohol
△	1 mM DEPC + 0.2 mM NAD^+
■	HLADH data (1 mM DEPC)

Chapter 7 Site-directed mutagenesis of *xylB* to investigate a possible role for residue arginine-53 in the reaction mechanism of benzyl alcohol dehydrogenase

7.1 Site-directed mutagenesis of *xylB* to create benzyl alcohol dehydrogenase-His53 and -Ala53 (BADH-His53 and BADH-Ala53)

Since benzyl alcohol dehydrogenase from *A. calcoaceticus* lacks the conserved His-51 residue proposed to act as a general base in the proton relay system of long-chain zinc-dependent alcohol dehydrogenases (section 1.10.2; Eklund *et al.*, 1982) it seemed possible that Arg-50 (position 53 with respect to horse liver alcohol dehydrogenase numbering; Fig 3.5) might possibly serve a similar role in a rearranged proton relay system in benzyl alcohol dehydrogenase since, unlike plant enzymes which have a tyrosine replacement in this position (Fig. 3.5), the isoleucine replacement cannot function in this role. The role of Arg-53 as a possible replacement for the conserved histidine residue was investigated using site-directed mutagenesis. Theoretically if an arginine was involved with proton translocation then changing this residue to an alanine should abolish activity, whereas changing it to a histidine should not, although the pK_a of the proton translocating group might be altered and so the pH optimum of the enzyme might be altered also (Sakoda & Imanaka, 1992). Mutants of benzyl alcohol dehydrogenase in which Arg-53 was changed to a His or an Ala residue were therefore made.

Site-directed mutagenesis was performed by the method of Higuchi *et al.* (1988) as previously described (section 2.23.2; Fig. 5.1). Fig. 7.1 shows the design of the oligonucleotides required to make the necessary changes to *xylB*. These primers were used in conjunction with the "outer" primers (oligos 2166 and 2168) used to create pDG30 (Fig. 3.1). Primary PCRs using pDG30/*XbaI* as template yielded the required 176bp and 990bp products for the "left" or "N-terminal" and "right" or "C-terminal" reactions respectively (Fig. 7.2). These products were purified from any other spurious products by low melting point agarose gel electrophoresis (section 2.10) followed by phenol/chloroform extraction (section 2.11). The purified primary products were then used as template for secondary PCRs using the "outer" primers. The products of secondary PCR needed to be purified by low melting point agarose gel also, because of the presence of another spurious product in each reaction (Fig. 7.2). The yield of "Ala-53" secondary PCR product was low, probably because of poor recovery during purification of one of the primary reactions. This secondary product was used as template for a further amplification using the "outer" primers and the product purified. The final mutagenised *xylB* products were digested with *NdeI* and *BglII* and cloned into pTB361 (section 2.2) to create pDG50 (His-53 mutant) and pDG60 (Ala-53 mutant).

Figure 7.1 Design of mutagenic oligonucleotides used to create benzyl alcohol dehydrogenase-His53 and -Ala53 mutants (pDG50 and pDG60)

Residue Arg-53 was changed to His-53 or Ala-53 using the PCR method described in section 2.23.2 and Fig. 5.1. The mutagenic *xyl* B products were cloned into pTB361 (section 2.2) to create pDG50 and pDG60.

Wild-type sequence   5´-ACCGATTTGATTGTA**CGT**GATCAATATTATCC-3´
                            T   D   L   I   V   R   D   Q   Y   Y   P

a. Benzyl alcohol dehydrogenase-His53 mutant

Forward direction   5´-ACCGATTTGATTGTAC**A**TGATCAATATTATCC-3´
(Oligo 2445)   T   D   L   I   V   H   D   Q   Y   Y   P

Reverse direction   5´-GGATAATATTGATCA**T**GTACAATCAAATCGGT-3´
(Oligo 2446)

b. Benzyl alcohol dehydrogenase-Ala53 mutant

Forward direction   5´-ACCGATTTGATTGTA**GC**TGATCAATATTATCC-3´
(Oligo 2447)   T   D   L   I   V   A   D   Q   Y   Y   P

Reverse direction   5´-GGATAATATTGATCA**GC**TACAATCAAATCGGT-3´
(Oligo 2448)

Figure 7.2 Site-directed mutagenesis using PCR with *Vent* polymerase to create benzyl alcohol dehydrogenase-His53 and -Ala53 mutants

Panel A. Primary PCR was performed as described in section 2.23.2 and then 20 μl of each 100 μl reaction was run in a 0.75% agarose gel (section 2.9). pDG30/*XbaI* (100ng) was used as template in each reaction.

Tracks: 1, λ/*Hind* III markers (250ng); 2, pDG30/*Xba* 1 (100ng); 3, no template control (oligos 2166 & 2168); 4, template control (i.e. both "outer" primers with no mutagenic primers); 5, His53 "left" (Fig. 5.1) or "N-terminal" primary reaction (oligos 2166 & 2446); 6, His53 "right" (Fig. 5.1) or "C-terminal" primary reaction (oligos 2445 & 2168); 7, Ala53 "left" (Fig. 5.1) or "N-terminal" primary reaction (oligos 2166 & 2448); 8, Ala53 "right" (Fig. 5.1) or "C-terminal" primary reaction (oligos 2447 & 2168).

Panel B. Secondary PCR was performed as described in section 2.23.2 and then 20 μl of each 100 μl reaction was run in a 0.75% agarose gel (section 2.9). Tracks: 1, λ/*Hind* III markers (250ng); 2, positive control (i.e. "outer" primers used to amplify unmutagenised template); 3, His53 mutagenesis secondary reaction; 4, Ala53 mutagenesis secondary reaction (oligos 2166 & 2168).

A

B
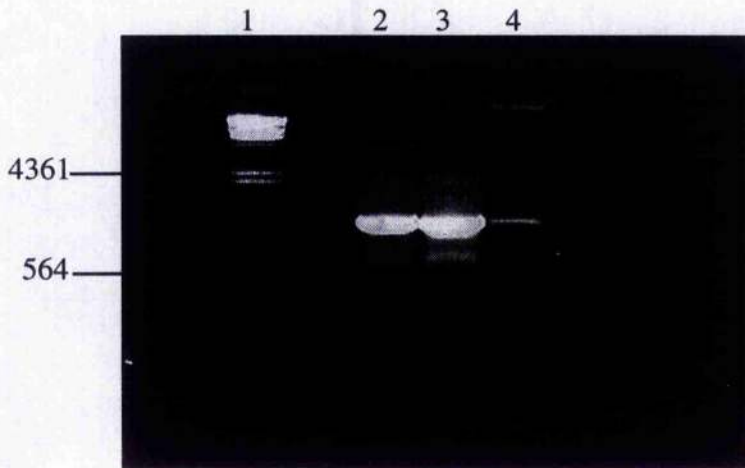
The clones pDG50 and pDG60 were sequenced using the oligonucleotides used to sequence pDG20 (section 2.19; Fig. 3.2). Fig. 7.3 shows that the correct bases had been altered in order to make the required amino acid alterations. The entire lengths of the two *xylB* mutant genes were sequenced and no PCR errors had been introduced into either clone.

7.2 Expression of the benzyl alcohol dehydrogenase mutants BADH-His53 and BADH-Ala53 in *E. coli* JM109 (DE3) and the purification of BADH-His53 from E. coli JM109 (DE3)/pDG50

E. coli JM109 (DE3) was transformed with pDG50 or pDG60 and the mutant enzymes expressed. These two enzymes, as well as wild-type benzyl alcohol dehydrogenase and BADH-His51, were expressed in a single experiment and the amounts of expression of all four enzymes was comparable as judged by laser scanning densitometry of SDS-PAGE (Fig. 7.4; section 2.29).

Table 7.1 records the specific activities of the samples analysed in Fig. 7.4. The specific activity of BADH-Ala53 was found to be almost identical to that of wild-type benzyl alcohol dehydrogenase (13.5 & 14.5 units/mg protein respectively). However, the specific activity of BADH-His53 was appreciably lower than that of the wild-type enzyme despite a similar level of expression as judged by the protein band in SDS-PAGE. The low level of activity of this enzyme meant that the determination of any kinetic coefficients in crude extract was difficult because of interference from NADH oxidase. For this reason and because it looked as though this enzyme might have interesting properties BADH-His53 was purified from *E. coli* JM109 (DE3)/pDG50 (Fig. 7.5) by the same method as wild-type benzyl alcohol dehydrogenase and BADH-His51 (section 2.27).

7.3 Kinetic coefficients of benzyl alcohol dehydrogenase-Ala53 (BADH-Ala53) and their comparison with wild-type benzyl alcohol dehydrogenase

Kinetic coefficients were determined as described in sections 2.26.6 and 5.3 by analysis of initial velocities over the substrate range 10 μM to 400 μM. In crude extracts in which there is a large over-expression it is possible to determine the K'_m of a very active enzyme. The K'_m values for benzyl alcohol oxidation determined from crude extracts for wild-type benzyl alcohol dehydrogenase and BADH-Ala53 were comparable with each other and with the

Figure 7.3 DNA sequencing autoradiographs of pDG20, pDG50 and pDG60 showing the mutagenic changes made to create benzyl alcohol dehydrogenase-His53 and -Ala53

Nucleotide sequences were determined as described in section 2.19.

Panel A: wild-type *xylB* sequence (pDG20).

Panel B: mutagenised *xylB* sequence to create benzyl alcohol dehydrogenase-His53 (pDG50).

Panel C: mutagenised *xylB* sequence to create benzyl alcohol dehydrogenase-Ala53 (pDG60).

A.
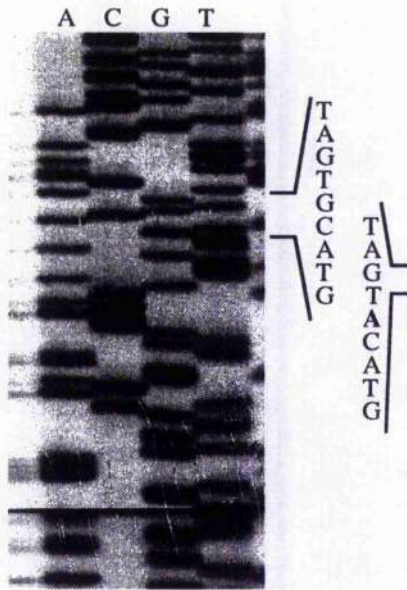
B
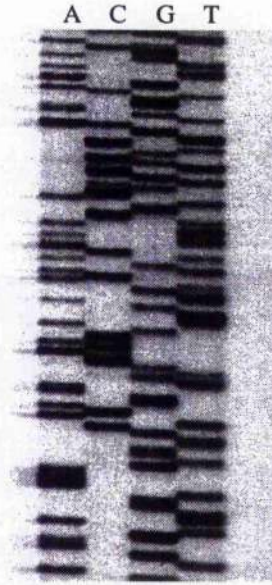
C
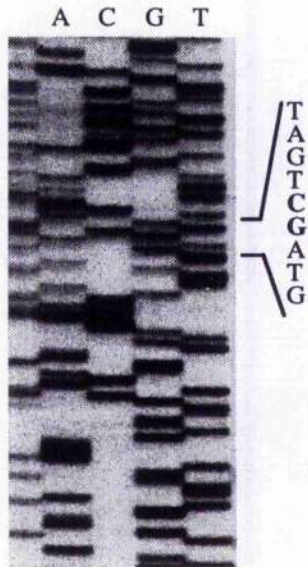
Figure 7.4 Expression of benzyl alcohol dehydrogenase wild-type and mutant enzyme clones in *E. coli* JM109 (DE3)

E. coli JM109 (DE3) transformed with the plasmids pDG30, pDG40, pDG50 or pDG60 was grown in 50ml LB+tetracylcline until the OD_{600} was 0.4. The expression of the enzymes was then induced by the addition of 0.5 mM IPTG (final concentration). After 3 hours 1 ml samples were removed for assay and for analysis by SDS-PAGE.

Samples for SDS-PAGE were resuspended in 100 μl loading buffer (incl. 50 mM DTT) and 5 μl of each was loaded onto a 5% (w/v) stacking and 10% (w/v) resolving polyacrylamide gel containing 0.1% (w/v) SDS (section 2.29).

Tracks: A and F, M_r markers (1 μg of each protein); B, C, D, and E, *E. coli* JM109(DE3)/pDG30, pDG40, pDG50 and pDG60 respectively (approximately 3 μg of total protein).

The % total protein constituted by BADH and mutant BADHs in each lane was: B (wild-type), 56%; C (BADH-His51), 62%; D (BADH-His53), 70% and E (BADH-Ala53), 66%; as judged by laser scanning densitrometry (section 2.29).
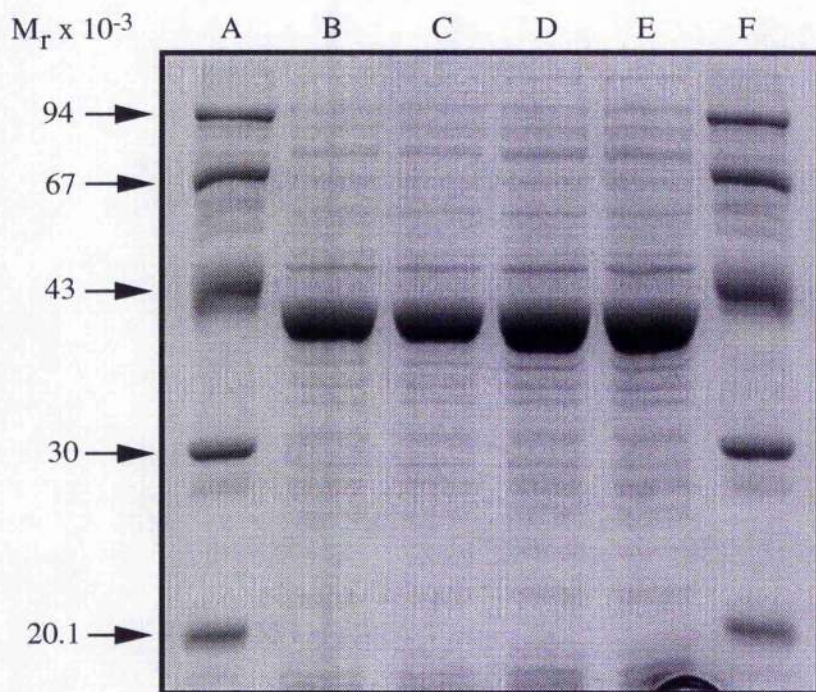
Table 7.1 Specific activities of benzyl alcohol dehydrogenase and the benzyl alcohol dehydrogenase-His51, -His53 and -Ala53 mutants when expressed in *E. coli* JM109 (DE3)/pDG30, pDG40, pDG50 or pDG60

Cultures (50ml) of *E. coli* JM109 (DE3)/pDG30, pDG40, pDG50 or pDG60 were grown to an OD_{600} of 0.4 when the expression of benzyl alcohol dehydrogenase and mutants was induced by the addition of 0.5 mM IPTG (final concentration). Expression was allowed to continue for 3 hours when 1ml samples were removed for assay and SDS-PAGE (Fig. 7.1). Samples for assay were resuspended in 0.5 ml 25mM Tris/HCl, 2mM DTT, sonicated for 6 x 10s at 50W (section 2.25) and spun in a microfuge for 5 minutes. Enzyme activity was determined as described in section 2.26 and protein concentrations determined by the method of Lowry *et al.* (1951; section 2.28).

Enzyme	Specific activity in the extract (Units/mg protein)
BADH	14.5
BADH-His51	26.4
BADH-His53	5.2
BADH-Ala53	13.5

Figure 7.5 Purification of benzyl alcohol dehydrogenase-His53 (BADH-His53) from *E. coli* JM109 (DE3)/pDG50

BADH-His53 was purified as described in section 2.27. Purification was monitored on a 10% (w/v) polyacrylamide slab gel containing 0.1% (w/v) SDS. Tracks: A and C, M_r markers (1 µg of each); B, concentrated Blue-Sepharose CL-6B elution pool (3 µg protein).
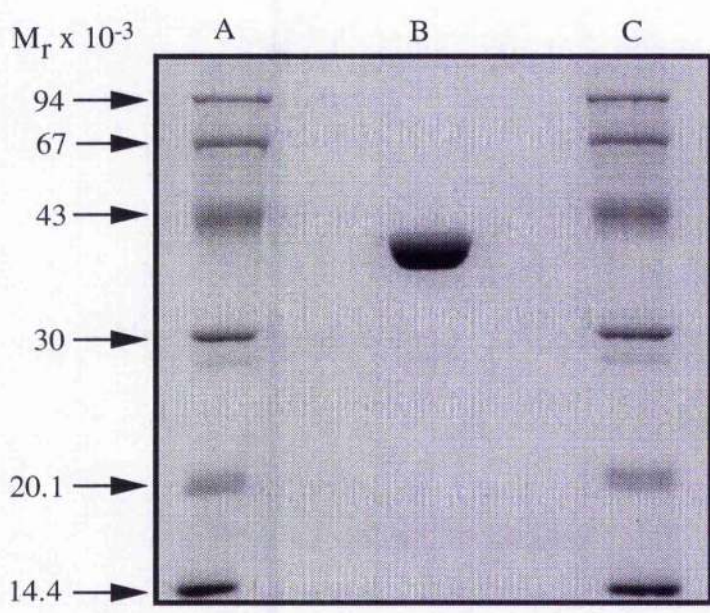
value obtained for the purified wild-type enzyme (Table. 7.2). The V'_{max} values obtained for the two crude extracts were also comparable. These results indicate that BADH-Ala53 has similar kinetic properties to the wild-type enzyme and suggests that Arg-53 is not involved in any proton transfer mechanism of benzyl alcohol dehydrogenase.

7.4 Comparison of kinetic coefficients and substrate specificity of benzyl alcohol dehydrogenase-His53 (BADH-His53) with those of the wild-type enzyme

The kinetic coefficients for benzyl alcohol oxidation by purified BADH-His53 were determined over the substrate range 0.8 mM to 13.3 μM. The K'_m of this mutant enzyme was 130-fold higher than that of the wild-type enzyme and the V'_{max} was 10-fold lower (Table 7.3). Consequently the k_{cat} of BADH-His53 was 10-fold lower than the wild-type enzyme and the "specificity constant" k_{cat}/K'_m almost 1400 times lower.

It is apparent from these results that BADH-His53 has a severely reduced ability to oxidise benzyl alcohol. However, BADH-Ala53 was shown to have very similar kinetic properties to wild-type benzyl alcohol dehydrogenase (section 7.3). Taken together these results suggest that Arg-53 cannot be involved as part of the proton relay system of the enzyme. The reduced ability of BADH-His53 to oxidise benzyl alcohol was probably due to steric effects brought about by a change in the secondary and/or tertiary structure of the enzyme.

The activity of BADH-His53 with various alcohol substrates relative to its activity with benzyl alcohol is recorded in Table 7.4. The relative activities followed a similar pattern to the wild-type enzyme, with 2-methylbenzyl alcohol a slightly poorer substrate and 4-methylbenzyl alcohol a slightly better substrate for BADH-His53. However, perillyl alcohol was a much better substrate for BADH-His53 than was benzyl alcohol.

The kinetic coefficients for perillyl alcohol oxidation by BADH-His53 and wild-type benzyl alcohol dehydrogenase were determined (Table 7.3). The K'_m of BADH-His53 for perillyl alcohol was only 17-fold higher than the wild-type enzyme compared to the 130-fold increase in the K'_m or benzyl alcohol, and the V'_{max} value for perillyl alcohol was greater for BADH-His53 than for wild-type benzyl alcohol dehydrogenase. Perillyl alcohol appears to be a "better" substrate for benzyl alcohol dehydrogenase than is benzyl alcohol, although *A. calcoaceticus* cannot use this substrate for growth. Substrate inhibition above 200 μM perillyl alcohol was observed during these experiments. Perillyl alcohol is the

Table 7.2 Comparison of the kinetic coefficients for benzyl alcohol oxidation of benzyl alcohol dehydrogenase-Ala53 with those of wild-type benzyl alcohol dehydrogenase

Enzyme	Apparent K_m (μM)	Apparent V_{max} (U/mg protein)
BADH (crude extract)	34	29
BADH-Ala53 (crude extract)	27	21

In addition, BADH-Ala53 was not able to oxidise methanol or ethanol at detectable rates.

Kinetic parameters were determined as described in section 2.26 from crude extracts prepared as described for Table 7.1.

Table 7.3 Kinetic coefficients of benzyl alcohol dehydrogenase-His53 compared with those of wild-type benzyl alcohol dehydrogenase

Enzyme	Substrate	Apparent K_m (μM)	Apparent V_{max} (U/mg protein)	k_{cat} (s^{-1})	Specificity constant k_{cat}/K'_m (s^{-1} μM^{-1})
BADH	Benzyl alcohol	19	81	53	2.79
	Perillyl alcohol	7	80	52	7.43
BADH-His53	Benzyl alcohol	2500	9	6	0.002
	Perillyl alcohol	119	130	84	0.710

In addition, BADH-His53 could not oxidise methanol or ethanol at detectable rates.

Kinetic parameters were determined as described in section 2.26. BADH and BADH-His53 were purified as described (sections 2.27, 3.2 and 8.2).

Table 7.4 Comparison of relative activities of wild-type benzyl alcohol dehydrogenase and benzyl alcohol dehydrogenase-His53 with various alcohol substrates

Substrate	% activity relative to benzyl alcohol	
	BADH-His53[a]	BADH[b]
Benzyl alcohol (I)	100%	100%
2-Methylbenzyl alcohol (II)	17	34
3-Methylbenzyl alcohol (III)	80	82
4-Methylbenzyl alcohol (IV)	113	97
Perillyl alcohol (VII)	280	55
Cinnamyl alcohol (VI)	53	48
Coniferyl alcohol (V)	6	8

In addition, BADH-His53 was not able to oxidise methanol or ethanol at detectable rates.

Substrate concentrations were fixed at 200 µM. Enzyme activities were determined as described in section 2.26. Within individual experiments all assays were performed in triplicate.

[a]Values are from a single individual experiment. Specific activity with benzyl alcohol = 6 units/mg protein.

[b]Values are the average of two entirely separate experiments (Table 5.2). Specific activity with benzyl alcohol = 320 units/mg protein.

Roman numerals refer to structures of these compounds in appendix.

only non-aromatic compound found to be oxidised by the enzyme and it appears that the changes in tertiary structure in BADH-His53 have altered the shape of the active site such that entry of this compound into the active site is not as restricted as the aromatic substrates. However, BADH-His53 had not gained the ability to oxidise the short chain aliphatic alcohols methanol and ethanol (section 9.5.2).

Chapter 8. Attempts to clone *mdlB*, the gene encoding L(+)-mandelate dehydrogenase, from *A. calcoaceticus* NCIB8250

8.1 Introduction

A. *calcoaceticus* NCIB8250 can grow on L(+)-mandelate as sole carbon and energy source. The mandelate pathway consists of three steps which converge with the pathway for benzyl alcohol utilisation at the point of benzaldehyde oxidation, although each pathway has a separate benzaldehyde dehydrogenase for this step (section 1.6; Fig. 1.5). The enzymes of the mandelate pathway had previously been purified from A. *calcoaceticus* (section 1.8 and references therein). Attempts were made to clone the gene encoding L-mandelate dehydrogenase (*mdlB*), the first enzyme in the mandelate pathway, by the techniques described in this chapter.

8.2 Use of the polymerase chain reaction to create a homologous probe for *mdlB*

Degenerate oligonucleotides were designed based on the N-terminal protein sequence of purified L(+)-mandelate dehydrogenase from A. *calcoaceticus* NCIB8250 (Fig. 8.1; Fewson *et al.*, 1993). These oligonucleotides were then used as PCR primers to amplify the N-terminal encoding region of the *mdlB* gene from A. *calcoaceticus* genomic DNA (section 2.23.1). It was intended that the gene fragment could then be used as a homologous probe to screen a lambda genomic DNA library of A. *calcoaceticus*, made prior to this study (section 3.1), in order to identify clones harbouring *mdlB*. This technique had previously been used in this laboratory to generate homologous probes for *xylC* and *xylB*, the genes from A. *calcoaceticus* encoding benzaldehyde dehydrogenase II and benzyl alcohol dehydrogenase respectively (section 3.1; A. G. S. Robertson, unpublished results; Reid, 1993).

Oligonucleotides 1381 and 1390 (Fig. 8.1) were designed so that they would amplify 138 bp from the N-terminal encoding region of *mdlB* during PCR with *Taq* polymerase using A. *calcoaceticus* genomic DNA as template. The theoretical melting temperatures of oligos 1381 and 1390 are 56°C and 60°C respectively as calculated by the GeneJockey sequence processor (Taylor, 1991) for the melting temperature of double stranded DNA of these particular sequences. PCR cycle parameters were set as follows: 1 minute melting at 94°C, 1 minute annealing at 50°C and 2 minutes polymerisation at 72°C for 35 cycles (section 2.23.1). A product of approximately the size required was obtained (Fig. 8.2). This product was purified from a low melting point agarose gel (section 2.10) and then used as template for a second PCR reaction. Again, the resulting

Figure 8.1 Design of degenerate oligonucleotides based on the *N*-terminal protein sequence of L-mandelate dehydrogenase in order to amplify part of *mdlB* using PCR

(a) *N*-Terminal sequence of L-mandelate dehydrogenase

```
                   _1450_ ▶▬
P H K M I N V E D Y Q K L A K S T
          ▬▬▬▬▶
            1381


L P K V I Y D Y L E G G A D D E K


                         ◀▬ _1449_
G L H H N R Q V F D Q K W F K P
          ◀▬▬▬▬
               1390
```
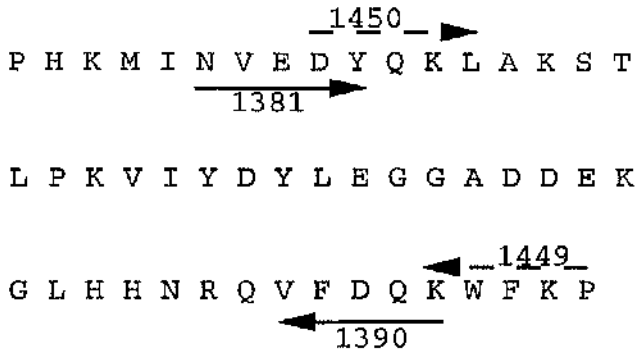
(b) Oligonucleotide sequences

BamHI

Oligo 1390 5´- CG**GGATCC**TTCTGATCAAAAAC-3´
                          T   G   G   G

Degeneracy=32
                                       C
                                       T

***Eco*RI**

Oligo1381 5´- CG**GAATTC**AACGTCGAAGACTA-3´
                         T   T   G   T

Degeneracy=32
                          A
                          G

***Bam*HI**

Oligo 1450 5´- CG**GGATCC**GAAGACTACCAAAAACT-3´
                         G   T   T   G   GT

Degeneracy=64

***Eco*RI**

Oligo 1449 5´- CG**GAATTC**AGGCTTAAACCACTTCT-3´
                       C   T   G   T   T
                       G

Degeneracy=64
                       T
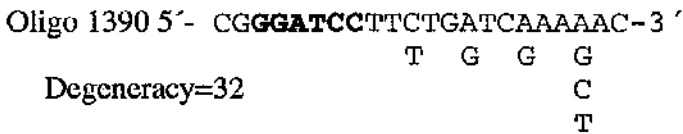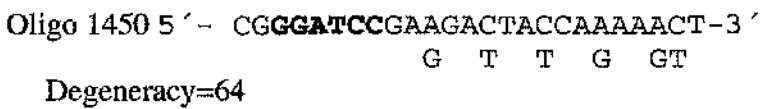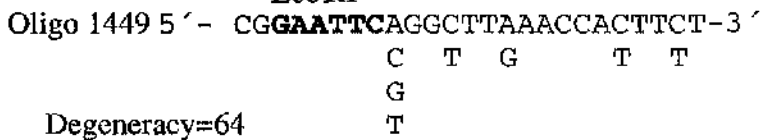
Figure 8.2 PCR using degenerate oligonucleotides (1381 and 1390) to amplify the *N*-terminal coding region of *mdlB* from *A. calcoaceticus* NCIB8250

PCR was run as described in section 2.23.1 using *A. calcoaceticus* NCIB8250 genomic DNA as template, oligonucleotides 1381 and 1390 as primers (Fig. 8.1) and *Taq* polymerase. A sample of each reaction (10 μl) was run in a 2% agarose gel (section 2.9). Tracks: A, no enzyme control; B, no target DNA control; C, no primer 1381 control; D, no primer 1390 control; E, reaction using 1 μg template DNA; F, reaction using 100 ng template DNA; G, reaction using 10 ng target DNA; H, pBR322/*HaeIII* markers.
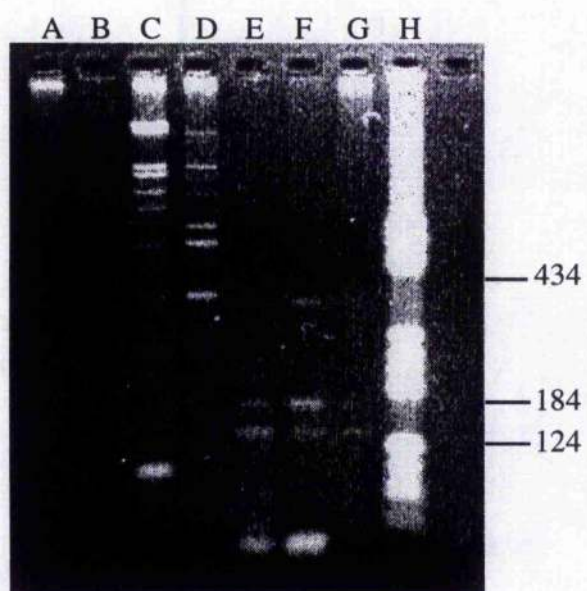
product of the required size was purified by low melting point agarose gel electrophoresis (section 2.10) and was then digested with *BamHI* and *EcoRI* (section 2.4) and ligated (section 2.5) into pUC18 treated with alkaline phosphatase (sections 2.2 & 2.6) and digested with *BamHI* and *EcoRI*. Clones identified as having inserts by the fact that they produced colonies on ampicillin-containing agar plates (section 2.3.5), and so must have harboured recombinant plasmids, often appeared to have no inserts when small scale plasmid preps (section 2.15) were digested with *BamHI* and *EcoRI* (section 2.4). A selection of putative recombinant clones were sequenced as described in section 2.19 and Fig 8.3 shows part of the sequences of inserts of these clones. Unfortunately, the product of the PCR reactions was found to be primer multimers and not the required *mdlB* fragment. These primer multimers are likely to have been shortened when restricted. Consequently they produced much smaller inserts than expected, which would not be seen when digested small scale plasmid preps were run in an agarose gel (section 2.9).

PCR cycle parameters were altered to try to alleviate the problem of primer multimers. Annealing temperatures of 45°C, 50°C and 55°C were all used. No products at all were obtained when the annealing temperature was raised to 55°C. Annealing times of 1, 1.5, 2 and 3 minutes were tried for each of the annealing temperatures but no products of the required size could be obtained. Also, at 50°C annealing temperature denaturation times of 1, 1.5 and 3 minutes at 94°C were used, but primer multimers were still formed. Increasing the number of cycles from 30 to 35 had no apparent effect under any reaction conditions.

A second set of oligonucleotides was designed based on the *N*-terminal protein sequence of L-mandelate dehydrogenase purified from *A. calcoaceticus* (oligos 1449 & 1450; Fig. 8.1) in case secondary structures in either of the oligos 1381 or 1390 were preventing polymerisation. Oligonucleotides 1449 and 1450 were based on sequence a few amino acids along from oligos 1381 and 1390 (Fig. 8.1). The theoretical melting temperatures of oligos 1449 and 1450 were 60°C and 62°C respectively, as calculated by the GeneJockey sequence processor (Taylor, 1991). The oligos 1449 and 1450 were designed so that they would amplify a 145 b.p fragment of *mdlB*. A product of approximately this size was obtained at an annealing temperature of 50°C (Fig 8.4). This product was purified from a low melting point agarose gel (section 2.10) and was then used as template for "hemi-nested" PCR using oligos 1390 and 1450 as primers. This technique is useful for confirming that a particular PCR product is the specific one required. However, "hemi-nested" PCR with oligos 1390 and 1450 failed to reamplify the 1449 and 1450 product. This showed that the primary PCR product was not the required *mdlB* fragment.

Figure 8.3 Nucleotide sequences of PCR products generated by amplification of *A.calcoaceticus* genomic DNA using oligonucleotides 1381 and 1390

Repeats of (or parts of) primer sequences are underlined.

Clone #1
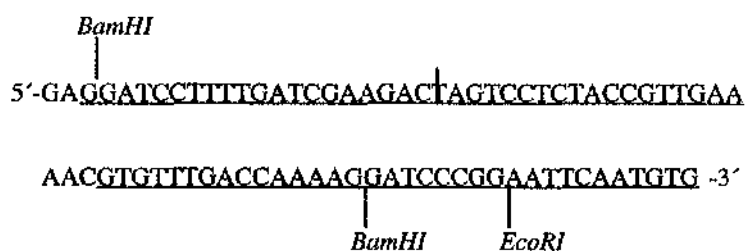

Clone #2
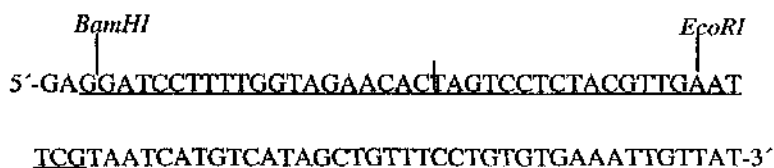
Clone #3
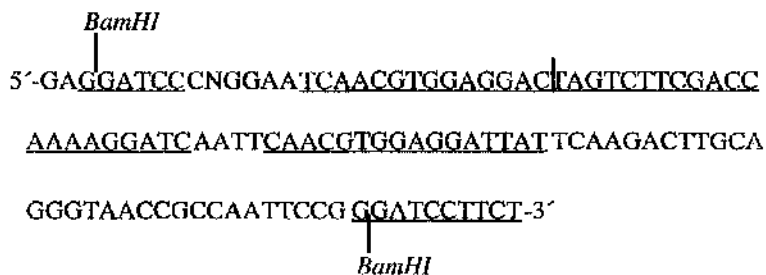
Figure 8.4 PCR using degenerate oligonucleotides (1449 & 1450) to amplify the *N*-terminal coding region of *mdlB* from *A. calcoaceticus* NCIB8250

PCR was run as described in section 2.23.1 using *A. calcoaceticus* NCIB8250 genomic DNA as template, oligonucleotides 1449 and 1450 as primers (Fig. 8.1) and *Taq* polymerase. A sample of each reaction (10 µl) was run in a 2% agarose gel (section 2.9). Tracks: A, pBR322/*HaeIII* markers; B, no enzyme control; C, no target DNA control; D, no primer 1449 control; E, no primer 1450 control; F, reaction using 10 ng template DNA; G, reaction using 100 ng template DNA; H, reaction using 1 µg target DNA; I and J, positive PCR control reactions using *Staphylococcus aureus* genomic DNA and DNA gyrase gyrB12 and gyrB13 primers (gift from Mr M. J. Horsburgh, Division of Biochemistry and Molecular Biology, Institute of Biomedical and Life Sciences, University of Glasgow).
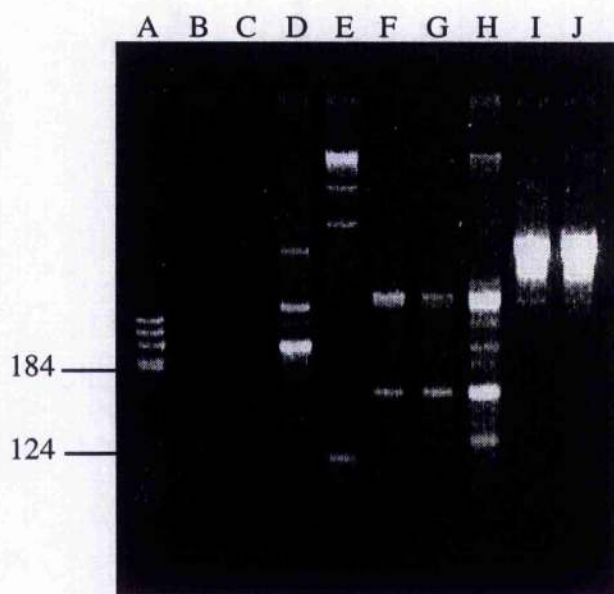
Two further combinations of primers were used to try to amplify a fragment of *mdlB* from *A. calcoaceticus* genomic DNA. Combinations of oligos 1390 and 1450, and oligos 1381 and 1449 were used together, but neither pair amplified a product of the required size.

Some additional techniques were also used to facilitate PCR using degenerate oligonucleotide primers. *A. calcoaceticus* genomic DNA digested with *KpnI* was used as template to ensure that inadequate denaturation of genomic DNA was not a problem. Also, 1% dimethylsulphoxide was used in the reaction mixtures as a means of reducing any secondary DNA structures that might prevent polymerisation. Neither of these techniques made any difference to the outcome of any of the experiments.

8.3 Use of *mdlB* from *Pseudomonas putida* NCIB9494 as a probe for *mdlB* from *A. calcoaceticus* NCIB8250

The gene encoding L(+)-mandelate dehydrogenase (*mdlB*) had previously been cloned and the nucleotide sequence determined (Tsou *et al.*, 1990). *N*-Terminal amino acid sequence alignments showed there to be 62% positional identity for the 50 residues at the start of their *N*-termini between L-mandelate dehydrogenase from *P. putida* and L-mandelate dehydrogenase purified from *A. calcoaceticus* (Fig. 1.6; Fewson *et al.*, 1993). These two enzymes also share some similar biochemical properties (section 1.8; Hoey *et al.*, 1987; Tsou *et al.*, 1990). For these reasons *mdlB* from *P. putida* was used to probe for *mdlB* from *A. calcoaceticus*.

The vector pBM (section 2.2) is a 5.8 k.b plasmid consisting of *mdlB* from *P. putida* cloned into the *EcoRI* and *PstI* restriction sites of the expression vector pKK223-3. pBM was digested with *EcoRI* and *PstI* (section 2.4) and the band corresponding to *mdlB* cut out from a 1.2% low melting point agarose/40 mM Tris-acetate, pH 8.0, 5 mM sodium acetate gel (section 2.9). This agarose slab could then be used directly in a random-priming DNA labelling reaction (section 2.22.1). Alternatively, the DNA could be purified from a low melting point agarose gel (section 2.10) prior to use in a random-priming DNA labelling reaction. It was found that DNA labelling was equally efficient irrespective of which labelling technique was used.

Initially, Southern blots were made of *A. calcoaceticus* genomic DNA digested with *BamHI*, *EcoRI*, *HindIII*, *KpnI*, *PstI*, *SacI*, *SalI*, *SmaI*, *XbaI* or *XhoI* (section 2.20). However, random-primed *mdlB* failed to hybridise (section 2.21) to any restriction fragments. As a positive control to ensure that the random-

priming labelling technique was working, and also to determine restriction fragment sizes for use when making further Southern blots, random-primed *mdlB* from *P. putida* was used to probe a Southern blot (sections 2.20 & 2.11) of *P. putida* genomic DNA (section 2.14) digested with various restriction enzymes (section 2.4; Fig. 8.5). The autoradiograph shows that this technique was working well.

EcoRI, *SalI* and *XhoI* digests of *P. putida* genomic DNA were chosen for use in further Southern blots since they produced a range of restriction fragments, with *SalI* and *XhoI* cutting within the gene (fragment lengths: *EcoRI*, 20 kb; *SalI*, 1.8 kb & 1.0 kb; *XhoI*, 12 kb & 5 kb). In addition to the three lanes of *P. putida* genomic DNA digested with *EcoRI*, *SalI* or *XhoI* the Southern blots also contained four lanes of *A. calcoaceticus* genomic DNA digested with either *EcoRI* or *HindIII*. These two enzymes were chosen because they restrict *A. calcoaceticus* genomic DNA frequently, producing a wide range of restriction fragment sizes. Two individual *A. calcoaceticus* genomic DNA preps (section 2.14) were also used. Fig. 8.6 shows a typical agarose gel used to make these Southern blots. Fig. 8.7 shows an autoradiograph of such a Southern blot probed with random-primed *mdlB* after one 15 minute wash at room temperature in 2 x SSC, 0.5% SDS, i.e at low stringency (section 2.11). From these experiments *mdlB* was shown to hybridise to *P. putida* genomic DNA but not to *A. calcoaceticus* genomic DNA.

The fact that *mdlB* from *P. putida* would not hybidise to *A. calcoaceticus* genomic DNA indicates that the two genes do not share a high enough degree of sequence identity. Since the two enzymes have been shown to have a high degree of *N*-terminal amino acid sequence identity it is possible that in this region the genes share a higher degree of sequence identity than along the rest of the genes. It is possible that a lack of sequence similarity further along the gene destabilises any hybridisation that may occur between the *N*-termini-encoding regions of the genes. For this reason a shorter *N*-terminal biased probe was created by digesting *mdlB* from *P. putida* with *SalI* (section 2.4) to produce a 321 b.p gene fragment. This DNA fragment was purified from a low melting point agarose gel (section 2.10) and labelled using random-priming (section 2.22.1). The radio-labelled 321 b.p DNA fragment was then used to probe a Southern blot, again containing three lanes of *P. putida* genomic DNA and four lanes of *A. calcoaceticus* genomic DNA. The blot was washed in 6 x SSC , 0.5% SDS for 15 minute at room temperature, i.e. at very low stringency, yet the probe failed to hybridise to *A. calcoaceticus* genomic DNA (Fig. 8.8). No visible bands were detectable in the *A. calcoaceticus* genomic DNA lanes even after the autoradiograph had been exposed for a week. These results suggest that *mdlB* of *A. calcoaceticus* and *P. putida* do not share a high degree of sequence similarity.

Figure 8.5 Southern blot of *Pseudomonas putida* genomic DNA probed with random-primed *mdlB* from *P. putida*

Each track contains approximately 5 µg *P. putida* NCIB9494 genomic DNA (prepared as described in section 2.14) digested with the following restriction enzymes (section 2.4): A, *BamHI*; B, *EcoRI*; C, *HindIII*; D, *KpnI*; E, *PstI*; F, *SacI*; G, *SalI*; H, *SmaI*; I, *XbaI*; J, *XhoI* and then probed with random-primed *mdlB* from *P. putida* (section 2.22.1). The autoradiograph was exposed for 24 hours.

A B C D E F G H I J
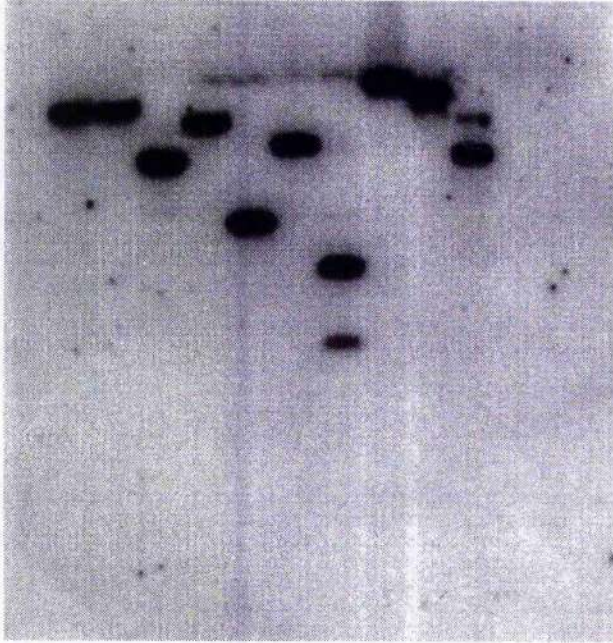
Figure 8.6 Typical agarose gel of restricted *A. calcoaceticus* NCIB8250 and *P. putida* NCIB9494 genomic DNA used for Southern blotting and then probing with random-primed *mdlB* or *mdlB* fragment

Tracks contain *A. calcoaceticus* NCIB8250 genomic DNA (either of two separate preps) or *P. putida* NCIB9494 genomic DNA (all prepared as described in section 2.14) digested with the stated restriction enzyme (section 2.4) and run on a 0.75% agarose gel (section 2.9). Tracks: A, λ/*HindIII*; B, *A. calcoaceticus* DNA prep 1/*EcoRI*; C, *A. calcoaceticus* DNA prep 1/*HindIII*; D, *A. calcoaceticus* DNA prep 2/*EcoRI*; E, *A. calcoaceticus* DNA prep 2/*HindIII*; F, *P. putida* DNA/*EcoRI*; G, *P. putida* DNA/*SalI*; H, *P. putida* DNA/*XhoI*.
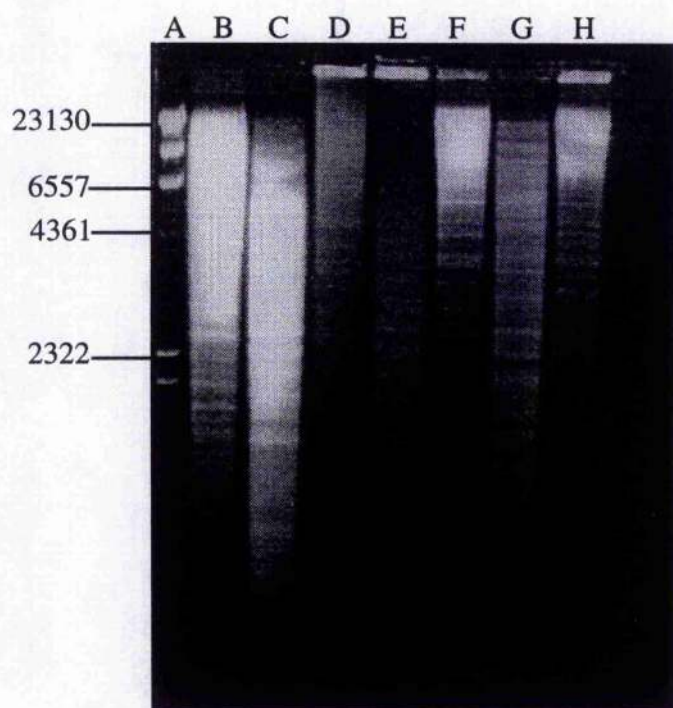
Figure 8.7 Southern blot of *A. calcoaceticus* and *P. putida* genomic DNA probed with random-primed *mdlB* from *P. putida*

Tracks contain *A. calcoaceticus* NCIB8250 genomic DNA (either of two separate preps) or *P. putida* NCIB9494 genomic DNA (all prepared as described in section 2.14) digested with the stated restriction enzyme (section 2.4) and probed with random-primed *mdlB* from *P. putida* (section 2.22.1). The autoradiograph was exposed for 24 hours. Tracks: A, λ/*HindIII* markers; B, *A. calcoaceticus* DNA prep 1/*EcoRI*; C, *A. calcoaceticus* DNA prep 1/*HindIII*; D, *A. calcoaceticus* DNA prep 2/*EcoRI*; E, *A. calcoaceticus* DNA prep 2/*HindIII*; F, *P. putida* DNA/*EcoRI*; G, *P. putida* DNA/*SalI*; H, *P. putida* DNA/*XhoI*.

A B C D E F G H
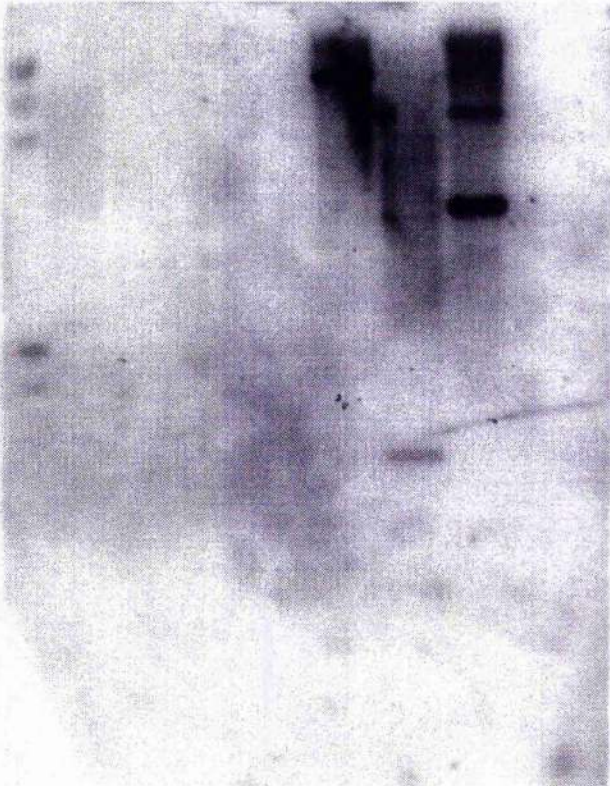
Figure 8.8 Southern blot of *A. calcoaceticus* NCIB8250 and *P. putida* NCIB9494 genomic DNA probed with a random-primed *N*-terminal coding region of *mdlB* from *P. putida*.

Tracks contain *A. calcoaceticus* NCIB8250 genomic DNA (either of two separate preps) or *P. putida* NCIB9494 genomic DNA (all prepared as described in section 2.14) digested with the stated restriction enzyme (section 2.4) and probed with a random-primed 321 bp *SalI* restriction fragment of *mdlB* (encoding the *N*-terminal region of L-mandelate dehydrogenase) from *P. putida* (sections 2.22.1 & 8.3). The autoradiograph was exposed for 8 days. Tracks: A, λ/*HindIII* markers; B, *A. calcoaceticus* DNA prep 1/*EcoRI*; C, *A. calcoaceticus* DNA prep 1/*HindIII*; D, *A. calcoaceticus* DNA prep 2/*EcoRI*; E, *A. calcoaceticus* DNA prep 2/*HindIII*; F, *P. putida* DNA/*EcoRI*; G, *P. putida* DNA/*SalI*; H, *P. putida* DNA/*XhoI*.

8.4.1 Use of degenerate oligonucleotides as probes for *mdlB* of *A. calcoaceticus*

Two approaches were used to design oligonucleotides for use as probes for *mdlB* from *A. calcoaceticus* (Fig. 8.9). A degenerate oligonucleotide (oligo 1532) was designed based on codons 2-12 of the *N*-terminal amino acid sequence of purified L-mandelate dehydrogenase from *A. calcoaceticus*. (Fewson *et al.*, 1993). Oligo 1532 was end labelled using T4 polynucleotide kinase (section 2.22.2) and was used to probe Southern blots (sections 2.20 & 2.21.1) of *A. calcoaceticus* genomic DNA digested with various restriction enzymes (section 8.2). The melting temperature (T_m) of oligonucleotides in hybridisations were calculated as described by Sambrook *et al.* (1989; section 2.21.2). T_m of oligo 1532 was calculated as 63°C. Hybridisation was carried out at 52°C and 54°C but at neither temperature, with non-stringent washes (two 15 minute washes at room temperature in 5 x SSC, 0.5% SDS), could any hybridisation signal be detected.

A degenerate oligonucleotide was also designed based on a protein motif identified as being conserved between the α-hydroxycarboxylic acid dehydrogenases: L-mandelate dehydrogenase from *P. putida*, flavocytochrome b_2 L-lactate dehydrogenase of the yeast *Saccharomyces cerevisiae* and glycolate oxidase of spinach (Fig. 8.9). The T_m of oligo 1530 was calculated as 46°C. The oligo was used to probe Southern blots of *A. calcoaceticus* genomic DNA, *P. putida* genomic DNA and serial dilutions of *mdlB* from *P. putida*. Hybridisation could be detected with as little as 2.5 ng of *mdlB* but not with either of the genomic DNA samples (Fig. 9.10) at 38°C, 39°C and 40°C.

8.4.2 Redesign of degenerate oligonucleotides with the inclusion of inosine residues

Oligo 1532 was highly degenerate (1536 degeneracies; Fig. 8.9). It was therefore possible that even if a small number of oligonucleotide species had hybridised to the Southern blot, such a small amount of the radiolabel might be hybridised that a signal would not be detected. For this reason the oligo was redesigned containing inosine residues to reduce the degeneracy of the probe (oligo 1599; Fig. 8.9). Inosine residues do not contribute to either increased or reduced hybridisation of an oligo (Sambrook *et al.*, 1989). They can be considered to be "neutral" residues and so if hybridisation occurs, more oligo species can hybridise than would be the case for degenerate oligonucleotides in which base mismatches are detrimental to hybridisation of an oligo. Oligo 1599 was used to

Figure 8.9 Design of degenerate oligonucleotides used to probe for *mdlB*

A. Degenerate oligonucleotides based upon the *N*-terminal protein sequence of mandelate dehydrogenase from *A. calcoaceticus*

```
MDH          H    K    M    I    N    V    E    D    Y    Q    K
Oligo 1532   CAC  AAA  ATG  ATA  AAC  GTA  GAA  GAC  TAC  CAA  AA
             T    G              C    T    C    G    T    T    G
                                 T         G
                                           T
                                        Degeneracy= 1536


Oligo 1599   CAC  AAI  ATG  ATI  AAI  GTI  GAI  GAI  TAC  CAA  AA
             T                                       T    G
                                        Degeneracy= 8
```

B. Degenerate oligonucleotides based upon a conserved protein motif from α-hydroxy carboxylic acid dehydrogenases

Degenerate oligonucleotides were designed against a motif conserved amongst L-mandelate dehydrogenase from *Pseudomonas putida* (L-MDH), flavocytochrome b_2 L-lactate dehydrogenase of the yeast *Saccharomyces cerevisiae* (L-LDH) and glycolate oxidase of spinach (GOX). Information taken from Tsou *et al.*, 1990; Smékal *et al.*, 1993. Final amino acid in each sequence is numbered.

```
L-LDH        W    Y    Q    L    Y    V    335
L-MDH        W    F    Q    L    Y    V    132
GOX          F    F    Q    L    Y    V    130

Motif 1.     W    F    Q    L    Y    V

Oligo 1530.  TGG  TTT  CAG  TTG  TAT  GT
                  C    A    C   A    C
                                T
                                C
                       Degeneracy= 64


Oligo 1600   TGG  TTI  CAI  ITI  TAT  GT
                                      C
                       Degeneracy=2
```

probe Southern blots containing *A. calcoaceticus* genomic DNA at hybridisation temperatures of 45°C, 50°C, 55°C and 60°C with non-stringent washes (two 15 minute washes at room temperature in 5 x SSC, 0.5% SDS). Faint hybridisation bands could be detected, particularly at 45°C and 50°C, after autoradiographs were exposed for four days. Unfortunately, a high level of background hybridisation meant that the hybridisation signal was not strong enough for the probe to be used to screen either plasmid or lambda libraries of *A. calcoaceticus*.

Oligo 1530 was also redesigned with the inclusion of inosine residues (oligo 1600; Fig. 8.9). In preliminary experiments in which oligo 1600 was used to probe Southern blots of *A. calcoaceticus* genomic DNA, *P. putida* genomic DNA and serial dilutions of *mdlB* from *P. putida* it was not possible to obtain stronger hybridisation bands than could be detected with oligo 1530.

8.5 Discussion and future work

8.5.1 Use of the polymerase chain reaction to create a homologous probe for *mdlB*

It was very surprising that attempts to amplify a fragment of *mdlB* by PCR using degenerate oligonucleotide primers designed using *N*-terminal amino acid sequence from L-mandelate dehydrogenase purified from *A. calcoaceticus* were unsuccessful (section 8.2), since this same technique had previously been used in the laboratory, and could be repeated in other work described in this thesis, to amplify fragments of *xylC* and *xylB* (section 3.1). A range of annealing temperatures, and annealing and polymerisation times were used to try to optimise amplification and so further attempts at optimisation are unlikely to yield better results. Redesign of oligos 1381, 1390, 1449 and 1450 so that they contain inosine residues instead of degeneracies might improve annealing of these oligos so that amplification of *mdlB* can be achieved. Inosine residues might be particularly beneficial at the 3'-end of oligonucleotides since it is at this end of the oligonucleotide that DNA polymerisation begins. Weak oligonucleotide-template DNA hybrids at the 3'-end of the oligonucleotide may result in poor DNA polymerase binding and consequently no (or very poor) amplification.

A problem of trying to amplify small products of about 100 bp or less was that these products are of a similar size to primer multimer products that can be obtained (section 8.1). Larger gene products could be obtained by using primers designed using "internal" amino acid sequence in conjunction with the

primers designed using *N*-terminal sequence. A conserved protein motif was identified in the three FMN-dependent α-hydroxy carboxylic acid dehydrogenases: L-mandelate dehydrogenase from *P. putida*, flavocytochrome b_2 L-lactate dehydrogenase from *S. cerevisiae* and glycolate oxidase from spinach (Fig. 8.9). This amino acid motif contains residue Tyr-129 of glycolate oxidase, a residue that is known to be in the active site of this enzyme (Lindqvist, 1989), and was used for the design of an oligonucleotide for use as a direct probe for *mdlB* from *A. calcoaceticus* (sections 8.3.1 and 8.3.2) and could also be used for the design of an "internal" PCR primer. There is another conserved protein motif in the same three α-hydroxy carboxylic acid dehydrogenases (Fig. 8.10). This motif contains Arg-257 from the active site of glycolate oxidase (Lindqvist, 1989). This motif could be used to design an "internal" PCR primer. Since these protein motifs contain active site and surrounding residues that are conserved across examples of bacterial, yeast and plant α-hydroxy carboxylic acid dehydrogenases it seems reasonable to suggest that these motifs will be conserved in L-mandelate dehydrogenase from *A. calcoaceticus*, another FMN-dependent α-hydroxy carboxylic acid dehydrogenase with similar properties to these enzymes (Fewson *et al.*, 1993). Use of PCR primers designed using the protien motifs described in Figs. 8.9 and 8.10, in conjunction with *N*-terminal PCR primers (all preferably containing inosine residues), should yield products of about 400 bp and 800 bp respectively.

An alternative method of designing "internal" PCR primers would be to try to obtain "internal" amino acid sequence information from the *A. calcoaceticus* L-mandelate dehydrogenase itself. This could be done by digesting the protein with various proteases and sequencing the peptides produced by this digestion.

8.5.2 Use of *mdlB* from *P. putida* as a probe for *mdlB* from *A. calcoaceticus*

Random-primed *mdlB* from *P. putida* failed to hybridise to Southern blots of *A. calcoaceticus* genomic DNA (section 9.2). Also, an *N*-terminal-encoding fragment of *mdlB* could not hybridise with *A. calcoaceticus* genomic DNA (section 9.2) despite the fact that there is 62% positional identity between the 50 *N*-terminal amino acids (Fig. 1.6; Fewson *et al.*, 1993). Acinetobacters have a typical G+C content of 38-45 mol % (Towner *et al.*, 1991). However, *mdlB* from *P. putida* has a G+C content of 55%. Since the *N*-terminal fragment of *mdlB* from *P. putida* did not hybridise with *A. calcoaceticus* genomic DNA in this

Figure 8.10 A conserved motif amongst hydroxy carboxylic acid dehydrogenases

The motif is conserved amongst the amino acid sequences of L-mandelate dehydrogenase from *Pseudomonas putida* (L-MDH), flavocytochrome b_2 L-lactate dehydrogenase of the yeast *Saccharomyces cerevisiae* (L-LDH) and glycolate oxidase of spinach (GOX). Information taken from Tsou *et al.* (1990); Smékal *et al.* (1993). Final amino acid in each sequence is numbered.

```
L-LDH       N   H   G   G   R   Q   L   D   457
L-MDH       N   H   G   G   R   Q   L   D   277
GOX         N   H   G   A   R   Q   L   D   259

Motif.      N   H   G   G   R   Q   L   D
```

region of strong identity it is possible that the G+C content of the two genes is sufficiently different to prevent hybridisation. Immunological cross-reaction studies had previously shown that a very large amount of protein needed to be used to detect even a weak cross-reaction between L-mandelate dehydrogenase from *P. putida* and antiserum raised against SDS-denatured L-mandelate dehydrogenase from *A. calcoaceticus* (Fewson *et al.*, 1993). Therefore it is also possible that the *A. calcoaceticus* and *P. putida* enzymes do not share a high degree of sequence identity across the whole protein sequence.

8.5.3 Use of degenerate oligonucleotides as probes for *mdlB* from *A. calcoaceticus*

Oligonucleotide 1599, designed using the *N*-terminal amino acid sequence of L-mandelate dehydrogenase from *A. calcoaceticus*, and oligonucleotides 1530 and 1600, designed using a protein motif conserved between several α-hydroxy carboxylic acid dehydrogenases (Fig. 8.9), produced weak hybridisation signals with Southern blots of *A. calcoaceticus* genomic DNA and *mdlB* from *P. putida* respectively (sections 9.3.1 and 9.3.2). Oligo 1599 could not produce a sufficiently strong hybridisation signal to be useful as a probe for either plasmid or lambda libraries of *A. calcoaceticus*. Hybridisation of this probe was carried out over at 45°C, 50°C, 55°C and 60°C with non-stringent washes. It is possible that smaller temperature changes, e.g. 2°C changes, may improve hybridisation further. Since oligos 1530 and 1600 could not hybridise to *P. putida* genomic DNA but could hybridise to larger quantities of *mdlB* from *P. putida* it is apparent that the technique was not sensitive enough to detect the very small quantities of the gene that would be present in total genomic DNA. Further optimisation of hybridisation using a large number of different temperatures needs to be carried out, particularly with oligo 1600 for which only preliminary experiments have been done.

It is essential that strong hybridisation under reasonably stringent conditions is achieved if oligonucleotides are to be used as probes for screening plasmid or lambda genomic DNA libraries, since stingent washes would be required during screens to remove background hybridisation that might interfere with the identification of colonies or plaques harbouring *xylB*.

8.5.4 Other strategies that could be used to clone *mdlB*

Complementation could be tried as a method of cloning *xylB* from *A. calcoaceticus*. *A. calcoaceticus* C1408 is a mutant strain of *A. calcoaceticus* NCIB8250 that is able to grow on phenylglyoxylate but not on mandelate (Livingstone & Fewson, 1972). Complementation of this mutant with *xylB* would therefore enable it to grow on L-mandelate. Transformation of *A. calcoaceticus* NCIB8250 has been achieved by electroporation using shuttle plasmids developed for transfer between *A. calcoaceticus* and *E. coli* (Ehrt *et al.*, 1994).

Another method that could be tried as a way of cloning *mdlB* is the construction of an expression library in *E. coli*. Colonies of this library could be screened for expression of L-mandelate dehydrogenase with antiserum previously raised against SDS-denatured L-mandelate dehydrogenase from *A. calcoaceticus* (Fewson *et al.*, 1993).

Chapter 9. Discussion

9.1 Genetic organisation of the benzyl alcohol pathway in *A. calcoaceticus* and nucleotide sequence of *xylB*, the gene encoding benzyl alcohol dehydrogenase

The results of this study and work previous to this study (sections 1.12, 3.1 & 3.3) have shown that there is a high degree of sequence identity between benzyl alcohol dehydrogenase and benzaldehyde dehydrogenase II from *A. calcoaceticus* and the corresponding enzymes encoded by the TOL-plasmid pWW0 of *P. putida* (54% & 46% identity respectively). It appears from the nucleotide sequence that the genes *xylB* and *xylC*, encoding benzyl alcohol dehydrogenase and benzaldehyde dehydrogenase II respectively, are part of an operon transcribed in the direction *xylC→xylB* (section 3.1; Fig. 3.4). This is supported by indirect evidence showing that the two enzymes are co-ordinately induced and expressed (Livingstone *et al.*, 1972) and they are also co-ordinately repressed during the metabolism of mandelate (Beggs *et al.*, 1976). In this study the enzymes could also be co-ordinately expressed in *E. coli* DH5α/pDG12 from the *lac* promoter of the vector pBluescript SK II after induction with IPTG (section 3.1; Fig. 3.1; Table 4.1). This indicates that *xylC* and *xylB* are both transcribed in the same mRNA transcript. However, the length of the RNA transcript needs to be determined by S1 nuclease mapping (Favaloro *et al.*, 1980) in order to determine where the exact start of transcription is, and whether the operon contains any further genes upstream of *xylC*. No further open reading frames could be found downstream of *xylB* and a putative σ factor-independent transcription termination sequence was identified downstream of the translational stop codon (section 3.3; Fig.3.4). These results suggest that there are no further genes in this putative operon downstream of *xylB*.

The organisation of the genes encoding the enzymes of the benzyl alcohol pathway in *A. calcoaceticus* is different from that of the TOL-plasmid pWW0 since the latter has an operon with additional genes. In the TOL-system, genes are transcribed in the order *xylUWCMABN* where *xylU* encodes a small protein consisting of 131 amino acid residues, but has no significant identity with any protein sequences in the data bases (GenBank accession no. u20269), and *xylW* encodes a 341 amino acid protein that is homologous with members of the group I zinc-dependent alcohol dehydrogenases (Fig. 3.5). In particular, the *xylW* product shares 37% sequence identity with threonine dehydrogenase of *E. coli* as calculated using the Bestfit program on the GCG UNIX system. As yet the *xylU* and *xylW* products have no known function. *xylC* encodes benzaldehyde dehydrogenase, *xylM* and *xylA* encode the two subunits of xylene monoxygenase, *xylB* encodes benzyl alcohol dehydrogenase and *xylN* encodes another

polypeptide of unknown function (section 1.5, Harayama *et al.*, 1989). The fact that *xylM* and *xylA* are not present in *A. calcoaceticus* is not surprising since this organism is not able to degrade toluenes and toluates (Fewson, 1967). Since the start and finish of the putative *A. calcoaceticus* operon have not been identified, it cannot be conclusively stated that this operon does not contain equivalent genes to *xylU*, *xylW* and *xylN*. If these genes are found not to be present in *A. calcoaceticus*, which appears to be likely, the study of metabolic differences between *A. calcoaceticus* NCIB8250 and *P. putida* mt-2 (pWW0) in relation to benzyl alcohol metabolism might help elucidate the function of these gene products.

Sequence analysis further upstream of *xylC* in *A. calcoaceticus* would also help to elucidate regulatory mechanisms involved in benzyl alcohol metabolism. In *P. putida* (pWW0) the regulatory protein XylR interacts with benzyl alcohol and toluene and stimulates transcription from the "upper" operon promoter, P_u, and also stimulates transcription of *xylS*, whose product (XylS) switches on the transcription of the *meta*-pathway genes (section 1.5; Harayama & Timmis, 1989). A similar system must exist in *A. calcoaceticus* by which the organism "senses" benzyl alcohol and consequently switches on the transcription of *xylC*, *xylB* and the β-ketoadipate genes. There must also be some pathway "cross-talk" between the mandelate pathway genes and the benzyl alcohol pathway genes, since the latter are repressed during metabolism of mandelate (Beggs *et al.*, 1976; Beggs & Fewson, 1977).

Codon usage in *xylB* from *A. calcoaceticus* was shown to be very typical of other acinetobacter structural genes (Table 3.1; White *et al.*, 1991). Also, the G+C content of *xylB* was 42%, which is typical of acinetobacters (average G+C 38-45 mol %; section 1.11; Towner *et al.*, 1991). In addition, *xylC* from *A. calcoaceticus* had previously been shown to have a G+C content (mol %) and a codon usage pattern typical of acinetobacters (section 3.1; A. G. S. Robertson, unpublished results). *xylB* from the TOL-plasmid pWW0 has a G+C content of 51%. This value is lower than is typical for pseudomonads, which have an average G+C content of 58-69 mol % (Palleroni, 1975). The two *xylB* genes share 60% sequence identity (section 3.3) with the differences in G+C content reflected in the respective codon usage patterns of the two genes (Tables 3.1 & 3.2). If TOL-*xylB* was more typical of *Pseudomonas* genes, then the codon usage would be expected to reflect the G+C content of the organism with a strong bias towards G or C in the wobble position. In some cases, for example the Cys and Phe codons, TOL-*xylB* is similar to *Acinetobacter* genes since it has a distinct preference for T and not C in the wobble position (Tables 3.1 & 3.2). These results suggest that the genes encoding the benzyl alcohol pathway are not a

recent acquisition by *A. calcoaceticus*. However, it remains a possibility that the TOL-plasmid is a recent acquisition by *P. putida*, after which codon usage has been altered to an extent to which expression of its genes is not hindered by the tRNA species present in the *P. putida* cellular pool. The *xyl* genes from pWW0 have been shown to be capable of transposing and/or recombining with other plasmids and into and out of chromosomes (Assinder & Williams, 1990). It therefore remains a possibility that *xylB* and *xylC* of pWW0 were recruited from acinetobacters and incorporated into their present operonic structure by further transposition.

9.2. Benzyl alcohol dehydrogenase from *A. calcoaceticus* is a member of the group of NAD(P)-dependent long-chain zinc-dependent alcohol dehydrogenases

The derived amino acid sequence of benzyl alcohol dehydrogenase from *A. calcoaceticus* (Fig. 3.4) has been shown to share a high degree of sequence identity with members of the group I zinc-dependent long-chain alcohol dehydrogenases (section 1.10.1) including horse liver alcohol dehydrogenase, the archetypal enzyme of this family (section 1.10.2).

Some of the previous work had suggested that benzyl alcohol dehydrogenase from *A. calcoaceticus* might not be a member of this group of zinc-dependent enzymes. MacKintosh and Fewson (1988b) had shown benzyl alcohol dehydrogenase to be unaffected by the metal-ion chelators 10 mM EDTA, 1 mM 2,2′-bipyridyl, 1 mM pyrazole and 1 mM 2-phenanthroline, unlike horse liver and yeast alcohol dehydrogenases (Sund & Theorell, 1963). However, it is possible that the three-dimensional structure of benzyl alcohol dehydrogenase is arranged such that zinc atoms are bound tightly, deep inside a tetrameric structure and so are inaccessable to metal chelating agents. The tetrameric benzyl alcohol dehydrogenase purified from the denitrifying bacterium *Thauera* sp., which shares many properties with the *A. calcoaceticus* enzyme has also been shown not to be inhibited by metal chelating agents (Biegert *et al.*, 1995). However, the dimeric benzyl alcohol dehydrogenase encoded by the TOL-plasmid pWW0 from *P. putida* could be inhibited by dialysis against 100 mM acetate buffer containing 5 mM benzyl alcohol and 5 mM 2-phenanthroline, although again this enzyme could not be inhibited by dialysis against 5 mM EDTA or 5 mM 2,6-pyridine dicarboxylic acid (Shaw *et al.*, 1993). This latter enzyme has been shown to bind two atoms of zinc per enzyme subunit (Shaw *et al.*, 1993) and shares 54% sequence identity with benzyl alcohol dehydrogenase from *A. calcoaceticus*

(Table 3.3). Also, it had previously been shown that antisera raised against benzyl alcohol dehydrogenase did not cross-react with the zinc-dependent alcohol dehydrogenases from horse liver and *Thermoanaerobium brockii* (Chalmers *et al.*, 1991). However, this latter study had shown there to be 36% sequence identity between the *N*-terminal amino acid sequences of benzyl alcohol dehydrogenase and horse liver alcohol dehydrogenase and this present study has shown that there is 32% sequence identity across the whole sequence of these two enzymes (Table 3.3). There appears to be a rule of thumb that immunological cross-reactions are detectable between denatured antigens that have greater than 40% sequence identity (Chalmers *et al.*, 1991; Fewson *et al.*, 1993), so there was probably insufficient sequence identity between benzyl alcohol dehydrogenase and horse liver alcohol dehydrogenase to produce an immunological cross-reaction.

Horse liver alcohol dehydrogenase has been shown by crystallographic studies to bind two zinc atoms per enzyme subunit, with a catalytic zinc atom bound at the active site to the ligands Cys-46, His-67 and Cys-174, and a structural zinc atom bound to the cysteine residues 97, 100, 103 and 111 (section 1.10.2; Eklund *et al.*, 1976). A tetrahedral coordination of the active site zinc is completed by a water molecule in the apoenzyme and by the hydroxyl group of the substrate in the holoenzyme (Eklund *et al.*, 1976; 1981; 1982; Ramaswamy *et al.*, 1994). It is interesting to note that the alcohol dehydrogenases from *C. beijerinckii* (CBADH), *T. brockii* (TBADH), *E. histolytica* (EADH) and *A. eutrophus* (AEADH) all appear not to contain a structural zinc atom since the Cys residues 97, 100, 103 and 111 are not conserved in their respective primary sequences (Fig. 3.5). The tertiary structures of these enzymes are likely to be altered with respect to the majority of the zinc-dependent alcohol dehydrogenases in order to compensate for the absence of a structural zinc atom. The four enzymes that appear not to bind a structural zinc atom, and also *E. coli* L-threonine dehydrogenase (ECTH) and the *xylW* product, also do not have the catalytic zinc ligand Cys-174, although the other two catalytic zinc ligands (Cys-46 and His-67) remain conserved in all of these enzymes (Fig. 3.5). Sheep liver sorbitol dehydrogenase has Glu at position 174 and the carboxyl group of this residue has been proposed to act as the third zinc ligand and alter the substrate specificity of the enzyme towards sorbitol and away from ethanol (Eklund *et al.*, 1985). The carboxyl group of Asp-174 in the enzymes CBADH, TBADH, EADH and AEADH could presumably also act as a zinc ligand. ECTH and the *xylW* product have Asn and Leu residues respectively at position 174, although Aronson *et al.* (1989) aligned the Asp residue at position 170 of *E. coli* L-threonine dehydrogenase (Fig. 3.5) with Cys-174 of horse liver alcohol dehydrogenase. The *xylW* product has a Glu residue at this position, so it seems

likely that the carboxyl groups of Asp- and Glu-170 (Fig. 3.5 numbering) could act as the third zinc ligand in ECTH and the *xylW* product.

The importance of the structural zinc atom in enzymes containing the conserved Cys residues (97, 100, 103 and 111) has been confirmed by site-directed mutagenesis in which the zinc ligands of human class I and class III alcohol dehydrogenase, both members of the group of long-chain zinc-dependent alcohol dehydrogenases, were replaced individually or collectively by alanine or serine residues; residues that are unable to ligate to zinc (Jeloková *et al.*, 1994). All of these mutations resulted in inactive, unstable enzymes that could not be easily purified. Also, the removal of the structural zinc from yeast alcohol dehydrogenase with dithiothreitol or EDTA resulted in an enzyme that was very sensitive to heat denaturation (Magonet *et al.*, 1992). All of the residues involved in the ligation of the catalytic and structural zinc atoms are conserved in benzyl alcohol dehydrogenase (Fig. 3.5). This confirms that benzyl alcohol dehydrogenase is indeed a member of the group 1 long-chain zinc-dependent alcohol dehydrogenases and that the enzyme probably binds two atoms of zinc per enzyme subunit in a manner analogous to horse liver alcohol dehydrogenase. It seems likely that the three dimensional structures of the benzyl alcohol dehydrogenases make it difficult to remove zinc atoms with metal chelating agents.

9.3 Evolutionary origins of benzyl alcohol dehydrogenase from *A. calcoaceticus*

When the multiple sequence alignment of the group I long-chain zinc-dependent alcohol dehydrogenases (Fig. 3.5) was expressed in the form of a dendrogram (Fig. 3.6) benzyl alcohol dehydrogenase of *A. calcoaceticus*, and the TOL-plasmid (pWW0) encoded benzyl alcohol dehydrogenase of *P. putida* mt-2, clustered with another alcohol dehydrogenase from *Pseudomonas* sp., but also with the alcohol dehydrogenases from horse liver and the plants barley, maize, rice, potato and pea, rather than other microbial alcohol dehydrogenases. This raises interesting questions concerning the evolutionary origins of these enzymes. Could it be that the plant and mammalian alcohol dehydrogenases and the benzyl alcohol dehydrogenases from *A. calcoaceticus* and *P. putida* (pWW0) all originated from a relatively recent common evolutionary ancestor?

Two major, but conflicting hypotheses have been proposed to explain the evolution of new proteins and metabolic pathways. The hypothesis of retrograde evolution proposes that enzymes are recruited in reverse order through

a metabolic pathway from a single parent enzyme through gene duplication and mutation, with the mutation accounting for the new-found substrate specificity (Horowitz, 1945). The benzyl alcohol pathway of *A. calcoaceticus* provides a good test for this hypothesis since benzyl alcohol dehydrogenase and benzaldehyde dehydrogenase II share two common substrates (benzaldehyde and NAD+) and catalyse sequential reactions. However, the retrograde hypothesis can be ruled out in this case since there is no significant sequence identity between *xylB* and *xylC* or between the amino acid sequences of the enzymes. This evidence is supported by previous work in which immunological cross reactivity studies indicated that benzyl alcohol dehydrogenase and benzaldehyde dehydrogenase II are not structurally homologous (Chalmers *et al.*, 1991).

The second hypothesis concerning the evolution of new proteins and metabolic pathways proposes that evolution occurs by way of gene recruitment (Jensen, 1976). In this proposal, copies of pre-existing genes are recruited, via gene duplication, mutation and translocation, and are then integrated into novel pathways performing analogous reactions to the original parent enzyme. This process would be aided by the relaxed substrate specificity of many enzymes. The origins of the parent genes in cases of evolution by gene recruitment is of great interest. It is possible that the parent gene could be recruited from another metabolic pathway within the same cell, but it is also a possibilty that horizontal gene transfer could occur. Horizontal gene transfer involves recruitment from another independent and fully developed cell (i.e. not from a parent to offspring), although this cell does not necessarily have to be of the same species (Amábile-Cuevas & Chicurel, 1993). In fact horizontal gene transfer has been proposed to occur between organisms as distantly related as mammals and bacteria (Smith *et al.*, 1992; Amábile-Cuevas & Chicurel, 1993).

The sequence data presented in this thesis suggest that benzyl alcohol dehydrogenase from *A. calcoaceticus* is more closely related to examples of mammalian and plant alcohol dehydrogenases than it is to bacterial enzymes from more closely related organisms. It is therefore a possibility that the gene encoding this enzyme was recruited from a eukaryotic source by horizontal gene transfer, although the fact that *xylB* is very typical of *Acinetobacter* genes in general (section 3.3) suggests that any transfer was a relatively distant one. For horizontal gene transfer to have occured there has to have contact between the two respective organisms. This can quite easily be accounted for in the case of *A. calcoaceticus* and plants, since *A. calcoaceticus* is found in soils (Towner *et al.*, 1991). Contact between mammalian species and *A. calcoaceticus* would be more difficult to explain since acinetobacters are strict aerobes and consequently not inhabitants of the rumen or intestines (Towner *et al.*, 1991). However, since plants are food for

many animals, the link between plants and mammals is obvious. Horizontal gene transfer between a eukaryote and a prokaryote has been proposed to have been responsible for the acquisition of a second glyceraldehyde-3-phosphate dehydrogenase (Gapdh) gene by an ancestor of *E. coli* (Smith *et al.*, 1992). *E. coli* has two *Gapdh* genes; *GapdhA* is more similar to eukaryotic sequences than it is to any other known prokaryotic sequence, whereas *GapdhB* reflects a conventional bacterial origin. The eukaryotic form of *Gapdh* has also been found in other species of enteric bacteria (Lawrence *et al.*, 1991; Nelson *et al.*, 1991) which suggests that horizontal transfer may have occured in the gut of a eukaryotic donor. Another example in which horizontal gene transfer has been implied to have occurred between a eukaryote and a prokaryote is the acquisition of the gene for glucose-6-phosphate isomerase (*Gpi*) by some strains of *E. coli*. The amino acid sequence of the *E. coli* enzyme and the plant *Clarkia ungulata* share 88% sequence identity (Froman *et al.*, 1989). The enzyme is not present in some strains of *E. coli* and other bacteria since glucose is primarily metabolised through the pentose phosphate shunt and it has been proposed that the lineage to *E. coli* lost the gene and then reacquired a copy from a plant (Smith *et al.*, 1992). There are also examples of possible horizontal gene transfer in the opposite direction, i.e. from a prokaryote to a eukaryote. A well documented example is that of *Agrobacterium tumefaciens* which "injects" a DNA fragment into the plant it infects (Zambryski *et al.*, 1989). These genes become incorporated into the genetic material of the plant and are expressed to produce proteins which the bacterium can then use as a carbon and nitrogen source. Considering these examples it is apparent that horizontal gene transfer between species as diverse as bacteria and plants can be implicated in evolutionary processes bringing about the acquisition of new enzymes and metabolic pathways and it is a distinct possibility that horizontal gene transfer from a eukaryote, most likely a plant, was responsible for the acquisition of *xylB* by *A. calcoaceticus*.

9.4 Comparison of the amino acid sequence of benzyl alcohol dehydrogenase with the primary structures of group I long-chain zinc-dependent alcohol dehydrogenases and the three-dimensional structure of horse liver alcohol dehydrogenase: relationship between the primary sequence, coenzyme binding, kinetic properties, substrate specificity and the catalytic mechanism of the enzyme

9.4.1 Conservation of amino acid residues amongst the group of long-chain zinc-dependent alcohol dehydrogenases

Aligning the primary sequences of a group of enzymes such as the long-chain zinc-dependent alcohol dehydrogenases, for which there is a crystal structure for typical members (Ramaswamy *et al.*, 1994; Hurley *et al.*, 1994) and for which there have been site-directed mutagenesis experiments on a number of residues, enables one to make predictions concerning the tertiary structures and catalytic mechanisms of other enzymes in this group, and consequently interpret kinetic data such as the substrate specificity of these enzymes. This is useful information for workers wishing to "engineer" the functions of an enzyme despite the absence of a three dimensional structure.

The alignment of the primary sequences of the long-chain zinc-dependent alcohol dehydrogenases (Fig. 3.5) shows that there are 12 residues conserved across all of the sequences. These are Cys-46, Asp-49, His-67, Glu-68, Val-80 and the glycine residues 66, 71, 77, 86, 201, 204 and 235. In addition, Gly-261 is conserved in all but the cinnamyl alcohol dehydrogenases of the cider and poplar trees. Ala-211 is conserved in all of the microbial and plant enzymes, apart from isoenzyme IV of *Kluyveromyces lactis,* and Cys-103 is conserved in all of the enzymes except for alcohol dehydrogenase from *Clostridium beijerinckii,* secondary alcohol dehydrogenase from *Thermoanaerobium brockii,* fermentative alcohol dehydrogenase from *Alcaligenes eutrophus* and alcohol dehydrogenase from *Entamoeba histolytica.*

There have not been any proposed catalytic roles for Val-80 and the glycine residues conserved across all of the long-chain zinc-dependent alcohol dehydrogenases but they all have structural roles, some of which involve coenzyme binding (Eklund *et al.*, 1976, 1981, 1982, 1984). Val-80, along with three of the conserved glycine residues (66, 71 and 77), His-67 and Glu-68 constitute the characteristic zinc-dependent alcohol dehydrogenase motif GHExxGxxxxxGxxV (Reid & Fewson, 1994).

The conserved residues Cys-46 and His-67 have been shown by protein crystallography techniques to be two of the four ligands involved in ligating the catalytic zinc atom in the active site of horse liver alcohol dehydrogenase (Eklund *et al.*, 1976; 1981; 1982; Ramaswamy *et al.*, 1994). The zinc atom is ligated to the sulphur atom of the cysteine residue and to one nitrogen atom of the histidine residue. The catalytic zinc has the function of binding the alcohol substrate in the correct orientation and facilitates formation of the alcoholate ion so that direct hydrogen transfer and subsequent rearrangement to aldehyde can occur (Eklund *et al.*, 1982; Kvassman *et al.*, 1981). Replacement of Cys-46 with a serine residue, a residue that would be unable to act as a zinc ligand, in the thermostable alcohol dehydrogenase of *Bacillus stearothermophilus* by site-directed mutagenesis confirmed the importance of this residue, since the Cys46Ser mutant enzyme had no alcohol dehydrogenase activity (Sakoda & Imanaka, 1992).

Asp-49 and Glu-68 are conserved across all of the long-chain zinc-dependent alcohol dehydrogenases; they are not inner sphere ligands to the catalytic zinc but are in the second sphere (Eklund *et al.*, 1976, 1981, 1984). The carboxyl group of Asp-49 interacts with the imidazole group of His-67, which is ligated directly to the zinc atom. The carboxyl group of Glu-68 has one oxygen close to the zinc with its other in a hydrogen bond to the guanidino group of Arg-369. It has been proposed that the negative charges provided by these residues (and also by Cys-46 and Cys-174) may act as a way of modulating the polarizing effect of the metal in order to facilitate hydride transfer to NAD^+ and to prevent substrates and products from binding too tightly (Kvassman *et al.*, 1981). Also, a recent report using theoretical molecular dynamic computations has suggested that it is possible that Glu-68 may intermittently coordinate the catalytic zinc, possibly as a means of facilitating the exchange of ligands in the active site by ligating to the zinc during "old" ligand dissociation (Ryde, 1995).

Site-directed mutagenesis experiments have been done using alcohol dehydrogenase I from *Saccharomyces cerevisiae* in order to evaluate the role of these conserved carboxyl groups by replacing Asp-49 and Glu-68 with their respective amide analogs, i.e. asparagine and glutamine (Ganzhorn & Plapp, 1988; Plapp *et al.*, 1990). These mutant enzymes displayed catalytic efficiencies (V/K_m) for ethanol oxidation and acetaldehyde reduction that were reduced by a factor of 1000 in the case of the Asp49Asn mutant and by a factor of 100 in the case of the Glu68Gln mutant. In addition, the dissociation constants of the mutant enzymes for coenzymes increased, turnover numbers decreased, and the pK values for ethanol oxidation were shifted from 7.7 in the wild-type to 6.2 and 6.8 in the Asp49Asn and Glu68Gln mutants respectively. It is possible that the

removal of the negative charges near the active site zinc creates an electrostatic environment in which conformational changes that require reorientation of groups around the zinc are hindered such that productive enzyme-substrate complexes are not formed as readily. Such conformational changes have been shown to occur in horse liver alcohol dehydrogenase upon coenzyme binding (Eklund *et al.*, 1981). These results are a good illustration of how residues not directly involved in binding substrates can contribute to catalysis.

9.4.2 Relationship between the primary structure of benzyl alcohol dehydrogenase from *A. calcoaceticus* and coenzyme binding

A combination of protein crystallography and site-directed mutagenesis techniques has been used to demonstrate the importance of various residues involved in coenzyme binding in long-chain zinc-dependent alcohol dehydrogenases. The coenzyme molecule binds in the large cleft region between the catalytic and coenzyme binding domains of the horse liver alcohol dehydrogenase subunit and interacts with residues from both domains (Fig. 1.7; Eklund *et al.*, 1981). The nicotinamide ring makes contact with Thr-178 and Val-203 (Eklund *et al.*, 1984). These residues are conserved in the benzyl alcohol dehydrogenase from *A. calcoaceticus* and the benzyl alcohol dehydrogenase encoded by the TOL-plasmid pWW0. Thr-178 is replaced by a serine in the alcohol dehydrogenase from *C. beijerinckii* and by an alanine in the *xylW* product, and Val-203 is replaced by a leucine or an isoleucine in many enzymes (Fig. 3.5). These are all conservative replacements that could perform the same function. Horse liver alcohol dehydrogenase is a class A enzyme with respect to hydride transfer (section 1.10.1; Schneider-Bernlöhr *et al.*, 1986), which means that coenzyme is bound with the nicotinamide ring in the *anti* position and transfer of hydride is to the *pro*-R position of the nicotinamide ring. Isotope studies have shown that benzyl alcohol dehydrogenase from *A. calcoaceticus* and benzyl alcohol dehydrogenase encoded by the TOL-plasmid pWW0 are also class A enzymes with respect to hydride transfer (Wales, 1992; Shaw & Harayama, 1990). From the crystal structure of horse liver alcohol dehydrogenase it has been predicted that binding of the *syn* conformer of the coenzyme is likely to be obstructed by the side chains of Thr-178 and Val-203 (Eklund *et al.*, 1984). Yeast alcohol dehydrogenase has conserved Thr-178, but Val-203 is conservatively replaced by a leucine residue (Fig. 3.5). The enzyme is extremely stereospecific for hydride transfer, making a stereochemical "error" every 7 billion turnovers (Weinhold *et al.*, 1991). However, the stereospecificity of the enzyme could be

reduced by replacing Leu-203 with an alanine and by making a double mutant in which Thr-178 was also replaced with a serine (Weinhold *et al.*, 1991). Therefore it appears that Thr-178 and Val-208 are responsible for determining the stereospecificity of hydride transfer in horse liver alcohol dehydrogenase and the benzyl alcohol dehydrogenases from *A. calcoaceticus* and *P. putida* (pWW0). The sequence alignment suggests that alcohol dehydrogenase from *C. beijerinckii* and the putative alcohol dehydrogenase encoded by *xylW* of the TOL-plasmid pWW0 could be class B enzymes in respect to hydride transfer, since they have serine and alanine residues respectively as replacements for Thr-178.

The adenine ring of the coenzyme molecule is sandwiched, in van der Waals contact, between the side chains of the two isoleucine residues 224 and 269, and the side chain of Arg-271 in horse liver alcohol dehydrogenase (Eklund *et al.*, 1984). These residues are not strictly conserved across the group of long-chain zinc-dependent alcohol dehydrogenases although there appear to be a number of conservative replacements (Fig. 3.5). However, both Ile-224 and Arg-271 are conserved in benzyl alcohol dehydrogenase from *A. calcoaceticus*, although Ile-269 is replaced by a threonine residue. A threonine residue in this latter position appears to be a common feature of plant alcohol dehydrogenases (Fig. 3.5). Replacements in these positions may be due to selection for altered rate constants since it has been shown that replacement of Ile-224 with glycine and Ile-269 with serine in horse liver alcohol dehydrogenase increased maximum velocities for ethanol oxidation, due to decreased affinity for coenzymes and hence faster release of NADH (Fan & Plapp, 1995).

The adenosine ribose binding site of horse liver alcohol dehydrogenase is at the surface of the coenzyme binding domain (Eklund *et al.*, 1984). It has been proposed that glycine residues 199, 201 and 202 allow the coenzyme to come close to the main protein chain (Eklund *et al.*, 1984). Gly-202 is replaced by an alanine or a proline residue in some microbial and plant sequences (Fig. 3.5). Alanine is found in this position in benzyl alcohol dehydrogenase from *A. calcoaceticus*. Glycines 199 and 201 have previously been proposed to be conserved in all zinc-dependent alcohol dehydrogenases (Jörnvall *et al*, 1987) and in the primary sequence alignment of Reid & Fewson (1994) this is also the case except that the fermentative alcohol dehydrogenase of *Alcaligenes eutrophus* has an alanine residue in position 199. However, the inclusion of additional primary sequences in Fig. 3.5 has altered the alignment such that in many of the microbial sequences this glycine residue is now positioned at position 198 with respect to horse liver alcohol dehydrogenase. Reid & Fewson (1994) also proposed Gly-192 to be conserved in all zinc-dependent alcohol dehydrogenases. Again, in Fig. 3.5 this residue is now aligned at position 191 in many of the microbial enzymes.

These residues align as might be expected if alcohol dehydrogenase from the *terp* operon of *Pseudomonas* species (PADH) is omitted. This enzyme appears to be more distantly related to the rest of the group of zinc-dependent alcohol dehydrogenases, although it has 42% sequence identity with benzyl alcohol dehydrogenase from *A. calcoaceticus* (Table 3.3). Interestingly, benzyl alcohol dehydrogenase from *A. calcoaceticus* has an alanine residue in position 192. This change could be brought about by a point mutation since the alanine codon in this case is GCC (Fig. 3.4), whereas a change to GGC would code for a glycine residue. Introduction of a gap in the alignment of these particular microbial sequences would realign both of these previously conserved glycine residues. It is highly likely that these two residues are conserved across all of these primary sequences, especially considering their proposed role in coenzyme binding. It is necessary to be aware of such "slippages" in alignments when aligning the primary sequences of a large number of potentially distantly related sequences, as is the case in Fig. 3.5 which aligns 30 mammalian, plant, yeast and bacterial sequences.

The carboxylate group of Asp-223 of horse liver alcohol dehydrogenase hydrogen bonds with the 2'- and 3'-hydroxyl groups of adenosine ribose and has also been proposed to be involved in determining the specificity of the enzyme for NAD(H) rather than NADP(H) (Eklund *et al.*, 1984). Enzymes that have this aspartate residue conserved tend to have a specific requirement for NAD(H) rather than NADP(H), for example horse liver and yeast alcohol dehydrogenases (Bränden *et al.*, 1975), whereas enzymes that have the smaller glycine or serine residue in this position tend to have the reverse requirement. The latter is illustrated in Fig. 3.5 by the alcohol dehydrogenases from *E. histolytica* (EADH) and *T. brockii* (TBADH), and the cinnamyl alcohol dehydrogenase from cider tree (ciderADH), all of which have an alanine or a serine in place of Asp-223 and all of which have been shown to be NADP(H)-dependent enzymes (Lo & Reeves, 1978; Lamed & Zeikus, 1981; Grima-pettenati *et al.*, 1993). Site-directed mutagenesis experiments have been done to try to alter the specificity of alcohol dehydrogenases for NAD(H) or NADP(H). Fan *et al.* (1991) replaced Asp-223 of alcohol dehydrogenase fom *S. cerevisiae* with a glycine residue and this mutation enabled the enzyme to use NADP(H). Lauvergeat *et al.* (1995) replaced the serine residue corresponding to Asp-223 of horse liver alcohol dehydrogenase in a plant NADP(H)-dependent cinnamyl alcohol dehydrogenase with an aspartate residue and found that this mutant enzyme was now able to utilise NAD(H) as coenzyme. However, in both of these studies the mutant enzyme had not become completely specific for the "new" coenzyme, although the mutants could not utilise the "old" coenzyme as efficiently as the wild-type enzymes. Neither mutant could utilise

the "new" coenzyme as efficiently as the wild-types could utilise the "old" coenzymes. These results suggest that the residues at position 223 might greatly influence specificity for coenzyme, but that they are not its sole determinants. Previous findings in which benzyl alcohol dehydrogenase from *A. calcoaceticus* was shown to have a specific requirement for NAD(H) as electron acceptor (MacKintosh & Fewson, 1988a) are supported here by the finding of an aspartate residue in the position analogous to 223 in the horse liver enzyme (Fig. 3.5).

The final residue proposed to be involved in the binding of the adenosine ribose moiety of the coenzyme in horse liver alcohol dehydrogenase is Lys-228. This residue has been shown by crystallography to hydrogen bond to the 3´-hydroxyl group of adenosine ribose (Eklund *et al.*, 1981; 1984). Modification of this residue with imido esters or isocyanates has been shown to activate the enzyme (Plapp, 1970; Zoltobrocki *et al.*, 1974; Sogin & Plapp, 1975). Site-directed mutagenesis in which Lys-228 was replaced by a glycine residue resulted in a mutant enzyme with an increased specific activity and dramatically increased K_m values for NAD^+ and NADH (Cho *et al.*, 1995). The mutant enzyme was also unaffected by acetimidylation. Lys-228 is conservatively replaced by an arginine residue in several zinc-dependent alcohol dehydrogenases, including benzyl alcohol dehydrogenase from *A. calcoaceticus* (Fig. 3.5).

The pyrophosphate bridge of the NAD^+ coenzyme binds in the central part of the cleft between the active site and coenzyme binding domains of horse liver alcohol dehydrogenase. The guanidino nitrogens of Arg-47 and Arg-369 hydrogen bond to each phosphate of the pyrophosphate bridge and also to the acidic side chains of Asp-50 and Glu-68 respectively (Eklund *et al.*, 1981; 1984). A number of water molecules have also been identified in this region as interacting with the pyrophosphate oxygens and amino acid residues (Ramaswamy *et al.*, 1994). Arg-47 can be found only in the mammalian enzymes (Fig. 3.5), with the majority of the other enzymes, including benzyl alcohol dehydrogenase from *A. calcoaceticus* having a conservative histidine replacement. Several enzymes, interestingly, have a glycine or a threonine residue in this position. Arginine or lysine residues appear to be conserved in position 369, this being another case in which there is an alignment "slippage" after the inclusion of the *Pseudomonas* sp. alcohol dehydrogenase (PADH) in the sequence alignment. Arginine appears to bind coenzyme more strongly than histidine since in site-directed mutagenesis studies the replacement of Arg-47 in human β-chain alcohol dehydrogenase with histidine resulted in a mutant enzyme with a decreased affinity for coenzyme, but with an incresed V_{max} and enzyme turnover (Hurley *et al.*, 1990a, b). The reverse of this was observed when His-47 of yeast alcohol dehydrogenase I was replaced with an arginine residue (Gould & Plapp, 1990).

Crystallographic studies have shown that the 100-fold decrease in affinity for coenzyme in the human β subunit His-47 mutant was not due to large changes in the enzyme structure (Hurley *et al.*, 1994). It is therefore clear that position 47 contributes significantly to the strength of protein-coenzyme interactions. Characterisation of yeast and horse liver alcohol dehydrogenases and benzyl alcohol dehydrogenase from *A. calcoaceticus* in this study (section 5.6; Table 5.3) showed that the two enzymes with His-47 rather than Arg-47, i.e. the yeast and *A. calcoaceticus* enzymes (Fig. 3.5), have much higher V'_{max}, and consequently k_{cat}, values than the horse liver enzyme which has Arg-47. It is possible that this is a result of the His-47 enzymes having higher dissociation constants for NADH.

Binding of the nicotinamide ribose moiety of the coenzyme has important implications regarding the proposed catalytic mechanism of horse liver alcohol dehydrogenase, and consequently long-chain zinc-dependent alcohol dehydrogenases in general (section 9.4.4; Eklund *et al.*, 1982; Ramaswamy *et al.*, 1994). The nicotinamide ribose binds in a narrow cleft between the active site and coenzyme binding domains, with the cleft lined on the catalytic side by residues 47, 48 and 51, and on the coenzyme binding domain side by residues 268-269 and 293-294 (Eklund *et al.*, 1984). There appears not to be any conservation amongst the zinc-dependent alcohol dehydrogenases as regards positions 268, 269, 293 and 294 (Fig. 3.5), with glycine, proline, alanine, valine, leucine, methionine, tyrosine, phenylalanine, histidine, cysteine, serine and threonine residues all found in these positions. This suggests that these residues have only minor interactions with the coenzyme molecule.

Although coenzyme binding in horse liver alcohol dehydrogenase has general similarities to coenzyme binding to dehydrogenases in general, it differs considerably in details (Rossman *et al.*, 1975; Eklund *et al.*, 1981; 1984). Comparison of the primary structure of benzyl alcohol dehydrogenase from *A. calcoaceticus* with the primary and three dimensional structures of horse liver alcohol dehydrogenase, with the primary structures of other long-chain zinc-dependent alcohol dehydrogenases, and also information from site-directed mutagenesis experiments has shown that coenzyme binding in benzyl alcohol dehydrogenase must occur in a manner analagous to coenzyme binding in horse liver alcohol dehydrogenase and to zinc-dependent alcohol dehydrogenases in general. This suggests that similarities between the horse liver and *A. calcoaceticus* enzymes are not just similarities common to NAD-dependent dehydrogenases in general, and implies that the three dimensional structure of benzyl alcohol dehydrogenase shares strong similarities with that of horse liver

alcohol dehydrogenase and that the two enzymes may share a common evolutionary origin.

9.4.3 The relationship between the primary structure of benzyl alcohol dehydrogenase from *A. calcoaceticus* and the substrate specificity of the enzyme

Benzyl alcohol dehydrogenase from *A. calcoaceticus* and benzyl alcohol dehydrogenase encoded by the TOL-plasmid pWW0 of *P. putida* are enzymes that display activity against a wide range of alcohol substrates (MacKintosh & Fewson, 1988b; Shaw & Harayama, 1990; Chalmers *et al.*, 1990; Shaw *et al.*, 1992; sections 1.9 & 5.2; Tables 1.2 & 5.2). Both enzymes can use substrates with substituents in the *ortho*, *meta*, and *para* positions of the benzene ring. The *A. calcoaceticus* enzyme greatly "prefers" *meta* and *para* substitutions. In this study 4-methylbenzyl alcohol was as good a substrate for the enzyme as benzyl alcohol, with the activity towards 3-methylbenzyl alcohol slightly lower than that found with benzyl alcohol (Table 5.2). 2-Methylbenzyl alcohol is a relatively poor substrate for the enzyme, only producing rates of oxidation about one third of those found with benzyl alcohol (Table 5.2). These results are in general agreement with previous results for benzyl alcohol dehydrogenase from *A. calcoaceticus* (MacKintosh, 1987; MacKintosh & Fewson, 1988b; Chalmers *et al.*, 1990). An apparent difference between benzyl alcohol dehydrogenase from *A. calcoaceticus* and that encoded by the TOL-plasmid pWW0 is that the TOL-encoded enzyme has a very strong preference for 3-methyl- and 3-methoxybenzyl alcohols as substrates compared to benzyl alcohol and 4-methyl- and 4-methoxybenzyl alcohols. Also, the TOL-enzyme can oxidise 2-methylbenzyl alcohol more efficiently than it can oxidise unsubstituted benzyl alcohol (Table 9.1; MacKintosh & Fewson, 1988b; Chalmers *et al.*, 1990; Shaw *et al.*, 1992). Benzyl alcohol dehydrogenase purified from the TOL-plasmid pWW53 also shows a strong preference for 3-methylbenzyl alcohol, although this enzyme does not use this substrate as efficiently as the pWW0 enzyme, whereas 2-methylbenzyl alcohol is a relatively poor substrate for the pWW53 enzyme (Table 9.1; Chalmers *et al.*, 1990). These results suggest that there may be differences in the structures of the active sites of these enzymes which bring about these slight differences in substrate specificities. There are differences in quaternary structure between these enzymes. Benzyl alcohol dehydrogenase from *A. calcoaceticus* is a tetramer (MacKintosh & Fewson, 1988a), whereas benzyl alcohol dehydrogenase encoded by the TOL-plasmid pWW0 is a dimer (Shaw & Harayama, 1990).

Table 9.1 Comparison of the kinetic constants of benzyl alcohol dehydrogenases from *A. calcoaceticus*, and encoded by the TOL-plasmids pWW53 and pWW0 for benzyl alcohol and various substituted benzyl alcohols

Data taken from: [a]This study (Table 5.1)

[b]MacKintosh & Fewson (1988b)

[c]Chalmers *et al.* (1990)

[d]Shaw *et al.* (1992)

Apparent "specificity constants" in Chalmers *et al.* (1990) are given in the form of apparent V/K_m' with units of l. $min^{-1}(mg\ protein)^{-1}$. These "specificity constants" were recalculated to include k_{cat} using the M_r value for the *A. calcoaceticus* enzyme determined in this thesis (38,923; Table 4.3) and 43,000 for benzyl alcohol dehydrogenase encoded by pWW53 (Chalmers *et al.*, 1990).

Apparent "specificity constants" taken from MacKintosh & Fewson (1988b) were recalculated such that k_{cat} was recalculated using the M_r value for benzyl alcohol dehydrogenase from *A. calcoaceticus* determined in this thesis (38,923; Table 4.3).

Source of BADH	Substrate	Apparent K_m (μM)	Apparent V_{max} (Units/mg protein)	Apparent specificity constant k_{cat}/K_m' ($s^{-1}\,\mu M^{-1}$)
Recombinant *A. calcoaceticus* expressed by *E.coli* JM109(DE3) /pDG30[a]	Benzyl alcohol (BA)	19	81	2.79
A. calcoaceticus wild-type[b]	BA	26	196	4.88
	3-MethoxyBA	145	230	1.03
	4-MethoxyBA	57	174	1.98
A. calcoaceticus wild-type[c]	BA	121	354	1.90
	2-MethylBA	992	405	0.27
	3-MethylBA	146	498	2.21
	4-MethylBA	118	593	3.26
pWW53 encoded enzyme (*P. putida*)[c]	BA	233	96	0.30
	2-MethylBA	605	56	0.07
	3-MethylBA	82	84	0.73
	4-MethylBA	106	91	0.61
pWW0 encoded enzyme (*P. putida*)[d]	BA	220	330	1.00
	2-MethylBA	100	280	1.90
	3-MethylBA	9	200	14.00
	4-MethylBA	84	320	2.50
	3-MethoxyBA	13	190	9.90
	4-MethoxyBA	37	93	1.70

Chalmers *et al.* (1990) proposed that the pWW53 encoded enzyme was a tetramer, as judged by cross-linking with dimethylsuberimidate followed by SDS-PAGE, although gel filtration suggested that the enzyme was a dimer. It is possible that differences in quaternary structure could affect access by substrates to the active sites of these enzymes.

Despite their affinities for a wide range of (substituted) aromatic alcohols, neither benzyl alcohol dehydrogenase from *A. calcoaceticus* nor the TOL-enzymes are able to oxidise aliphatic alcohols, although the *A. calcoaceticus* enzyme can oxidise the cyclohex-1-ene compound perillyl alcohol (Appendix V; Tables 1.2 & 5.2; MacKintosh & Fewson, 1988b; Shaw & Harayama, 1990). This is in contrast to horse liver alcohol dehydrogenase which can oxidise straight and branched chain primary alcohols, secondary alcohols, cyclohexanol and benzyl alcohol, and is also in contrast to yeast alcohol dehydrogenase which can oxidise short chain primary alcohols and little else (Sund & Theorell, 1963; Brändén *et al.*, 1975). Yeast alcohol dehydrogenase has been reported to slowly oxidise benzyl alcohol and *para*-substituted benzyl alcohols (Klinman, 1976). However, most studies report that the yeast enzyme does not oxidise this substrate (e.g. Creaser *et al.*, 1990; Green *et al.*, 1993). These differences may be due to the fact that some studies used cloned yeast alcohol dehydrogenase I, whereas others use commercially purchased yeast alcohol dehydrogenase which is likely to be a mixture of the various baker's yeast isoenzymes. Commercially purchased enzyme is likely to vary between different manufacturers, with different purification procedures resulting in different proportions of the various isoenzymes. In this study no activity with benzyl alcohol could be detected using commercially produced yeast alcohol dehydrogenase (Table 5.3).

Kinetic coefficients for the oxidation of ethanol and benzyl alcohol were determined for both yeast and horse liver alcohol dehydrogenases (section 5.6; Table 5.3). These values are in general agreement with literature values (Light *et al.*, 1992; Shearer *et al.*, 1993; Weinhold & Benner, 1995). The kinetic coefficients show that the yeast enzyme has a higher apparent K_m value than the horse liver enzyme. However, the yeast enzyme has a much higher apparent V_{max} value than horse liver alcohol dehydrogenase and consequently has a 15-fold higher "specificity constant". Yeast alcohol dehydrogenase did not have any detectable activity against benzyl alcohol, whereas the horse liver enzyme had an apparent K_m value similar to that of benzyl alcohol dehydrogenase from *A. calcoaceticus* (Table 5.3). Benzyl alcohol dehydrogenase from *A. calcoaceticus* has an apparent V_{max} 645-fold higher for benzyl alcohol oxidation than the horse enzyme, and consequently has a "specificity constant" about 500-fold higher. These differences in kinetic coefficients are reflected in the physiological roles of

these enzymes. Yeast alcohol dehydrogenase is involved in reducing accumulating acetaldehyde to ethanol during anaerobic metabolism in order to regenerate NAD^+ so that other cellular processes can continue (Clark, 1992). Mammalian liver alcohol dehydrogenases have the proposed function of metabolising circulating alcohols that have been ingested or produced by bacteria in the intestines (Brändén *et al.*, 1975). The yeast enzyme is therefore only likely to need to operate at high concentrations of acetaldehyde and even higher concentrations of ethanol, so it has a high K'_m value, but may need to turn over these compounds quickly to keep metabolism going. The higher efficiency of this yeast enzyme compared to mammalian enzymes no doubt reduces the energy requirements for biosynthesis of the enzyme, which produces alcohol in stoichieometric amounts, therefore increasing the efficiency of fermentation. In the mammalian liver, the rate of alcohol metabolism probably is not as important as the enzymes having activity on the wider range of alcohols that the body may need to metabolise, so a broad substrate specificity was a more likely selective pressure here, with the enzyme still needing to operate mainly with increased alcohol levels. The bacterial benzyl alcohol dehydrogenases from *A. calcoaceticus* and *P. putida* have very low K'_m values coupled with high turnover rates (Table 9.1). Aromatic compounds such as benzyl alcohol and its analogues are likely to be present in the environment in low abundance (Fewson, 1988) so bacterial enzymes that feed these compounds into central metabolism for growth of the microorganism need to have low K_m values. The high turnover numbers of these enzymes probably enables the microorganisms to have high growth rates when using aromatic aldehydes as sole carbon sources. As was probably the case with the yeast enzyme, the high efficiency of these enzymes probably also results in lower energy requirements for biosynthesis of the enzymes, thus facilitating increased growth rates.

It is usually considered that the higher the "specificity constant", the closer the structure of the substrate transition-state intermediate is to the structure of the enzyme's active site (Fersht, 1985). These results would therefore suggest that the active site of yeast alcohol dehydrogenase has a structure with the closest resemblance to that of ethanol, with horse liver alcohol dehydrogenase having an active site that more closely resembles the size and structure of benzyl alcohol than ethanol. The high "specificity constant" of benzyl alcohol dehydrogenase from *A. calcoaceticus* indicates that this enzyme has an active site structure very similar to the benzyl alcohol transition state intermediate.

Molecular modelling of primary sequences of alcohol dehydrogenases with differing substrate specificities onto the three dimensional structure of horse liver alcohol dehydrogenase, coupled with site-directed mutagenesis of selected

residues has enabled the identification of residues that might be involved in determining substrate specificity. For this study a molecular model of the active site of horse liver alcohol dehydrogenase was constructed from the three-dimensional structure of the enzyme complexed with pentafluorobenzyl alcohol (Figure 9.1; Ramaswamy *et al.*, 1994). This model was then used to create a putative model of the active site of benzyl alcohol dehydrogenase from *A. calcoaceticus* (Figure 9.2).

Changes at position 57 (horse liver alcohol dehydrogenase numbering) have been proposed to alter substrate specificity. This residue is a leucine in horse liver and human ββ and γγ enzymes, methionine in the human and Rhesus monkey α subunits, phenylalanine and aspartate in the human π and χ subunits respectively, and tyrosine in the benzyl alcohol dehydrogenases from *A. calcoaceticus* and encoded by the TOL-plasmid pWW0 (Table 9.2; Figs. 9.1 & 9.2; Eklund *et al.*, 1990; Green *et al.*, 1993). Aspartate at position 57 in class III human alcohol dehydrogenases (χ subunit) has been implicated as being involved in glutathione-dependent formaldehyde dehydrogenase activity of this enzyme; activity which appears to mediated by factors additional to, or different from alcohol dehydrogenae activity (Estonius *et al.*, 1994). For yeast alcohol dehydrogenase isoenzymes, Fig. 3.5 aligns a proline and Reid & Fewson (1994) align a leucine in position 57. However, modelling studies of yeast alcohol dehydrogenase on the three-dimensional structure of horse liver alcohol dehydrogenase have placed a tryptophan (Trp-54 of yeast ADH) in this position (Plapp *et al.*, 1987). Weinhold & Benner (1995) replaced Trp-57 (horse liver alcohol dehydrogenase numbering) with a leucine residue as in horse liver and the human β subunit enzymes. This replacement would be expected to increase space within the active site of the enzyme and so enable the oxidation of substrates with increased steric volumes. The Trp57Leu mutant had a greatly increased ability to oxidise the longer chain primary alcohols butanol, pentanol and hexanol and also alcohols with branching at the 4 position; notably 4-methyl-1-pentanol and cinnamyl alcohol. The mutant enzyme did not have an increased ability to oxidise secondary alcohols and the "specificity constant" for ethanol oxidation was an order of magnitude lower than in the wild-type enzyme. This mutant had therefore gained substrate specificities associated with horse liver rather than yeast alcohol dehydrogenase. However, replacement of Trp-57 of yeast alcohol dehydrogenase with methionine, as can be found in human and Rhesus monkey α subunits, did not yield a mutant enzyme with an increased catalytic ability towards any substrate (Green *et al.*, 1993). These α subunit enzymes have a distinct preference for longer chain primary alcohols and benzyl alcohol than they have for ethanol (Wagner *et al.*, 1983; Light *et al.*, 1992). In the latter study the Trp57Met mutant

Figure 9.1 Molecular model of the three-dimensional structure of the active site of horse liver alcohol dehydrogenase complexed with pentafluorobenzyl alcohol

The model was constructed from the coordinates placed in the Brookhaven data base (1HLD; Ramaswamy *et al.*, 1994) using the program O (Jones *et al.*, 1991) in collaboration with Mr D. Gourlay, Unit of Protein Crystallography, University of Glasgow.
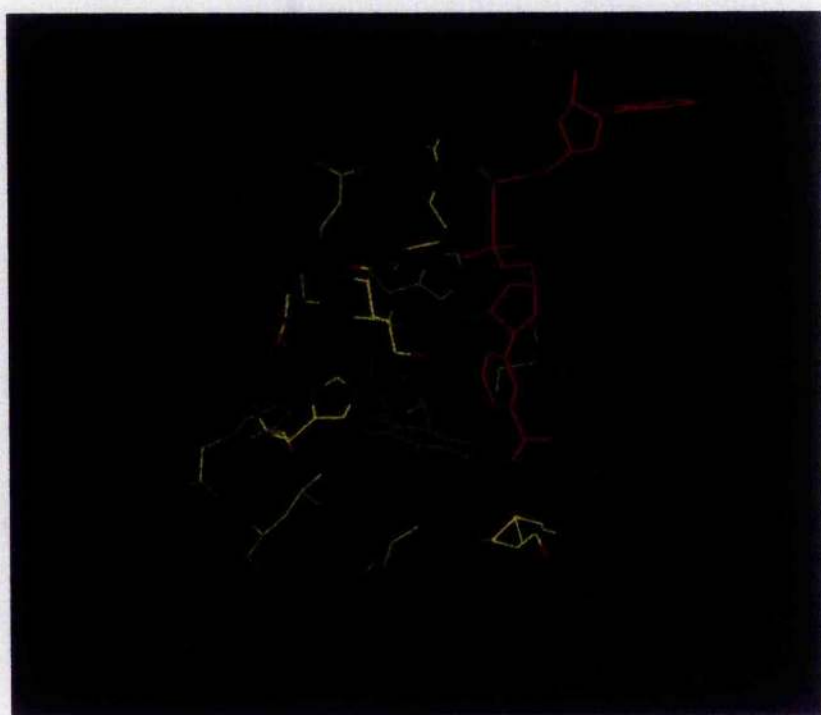
Figure 9.2 Putative molecular model of the active site of benzyl alcohol dehydrogenase from *A. calcoaceticus*

The model was constructed by using molecular replacement to replace amino acids in the three-dimensional structure of horse liver alcohol dehydrogenase (Fig. 9.1) with those judged to be in the equivalent position in benzyl alcohol dehydrogenase, using the program O (Jones *et al.*, 1991). The equivalence of amino acid positions between horse liver alcohol dehydrogenase and benzyl alcohol dehydrogenase was determined by sequence alignment between the primary structures of the two enzymes using the Bestfit program on the GCG/UNIX system. Amino acid residues that are different in the benzyl alcohol dehydrogenase model and the horse liver alcohol dehydrogenase model (Fig. 9.2) are shown in purple. Amino acid residues that are conserved between the two models remain yellow.

The model was constructed in collaboration with Mr D. Gourlay, Unit of Protein Crystallography, University of Glasgow.
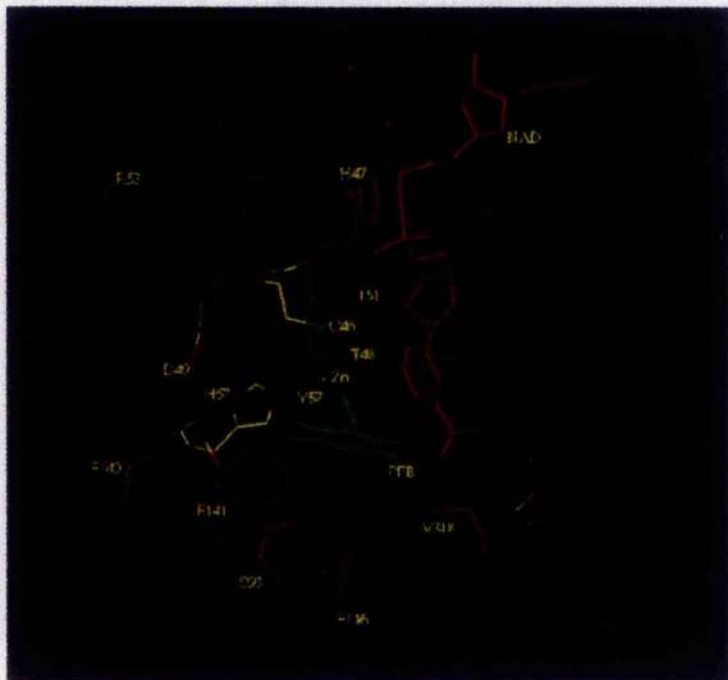
Table 9.2 Comparison of residues lining the substrate-binding cleft of horse liver alcohol dehydrogenase with human alcohol dehydrogenase isoenzymes, yeast alcohol dehydrogenase I, and the benzyl alcohol dehydrogenases from *A. calcoaceticus* and *P. putida* (pWW0)

Based on Eklund *et al.* (1990) with the addition of benzyl alcohol dehydrogenase sequences from Fig. 3.5 and human class IV σ subunit alcohol dehydrogenase from Kedishvili *et al.* (1995). Yeast ADH I sequence alignment is taken from Fig. 3.5 except that positions 57 and 93 are as in the model of Plapp *et al.* (1990; see text).

Residues differing from those of horse liver alcohol dehydrogenase are shown in bold. Active site zinc-liganding residues (conserved in all of the sequences) are not shown.

Source	Amino acid at position										
	(inner)				(middle)				(outer)		
	48	**93**	**140**	**141**	**57**	**116**	**294**	**318**	**110**	**306**	**309**
horse E	Ser	Phe	Phe	Leu	Leu	Leu	Val	Ile	Phe	Met	Leu
hum α	**Thr**	**Ala**	Phe	Leu	**Met**	**Val**	Val	Ile	**Tyr**	Met	Leu
hum β	**Thr**	Phe	Phe	Leu	Leu	Leu	Val	**Val**	**Tyr**	Met	Leu
hum γ	Ser	Phe	Phe	**Val**	Leu	Leu	Val	Ile	**Tyr**	Met	Leu
hum π	**Thr**	**Tyr**	Phe	**Phe**	**Phe**	Leu	Val	**Phe**	**Leu**	**Glu**	**Ile**
hum χ	**Thr**	**Tyr**	**Tyr**	**Met**	**Asp**	**Val**	Val	**Ala**	**Leu**	**Phe**	**Val**
hum σ	**Thr**	Phe	Phe	**Met**	**Met**	Ile	Val	**Val**	**Leu**	Met	**Phe**
yeast adh I	**Thr**	**Trp**	**Tyr**	**Thr**	**Trp**	Leu	**Pro**	**Tyr**	**Asn**	**Phe**	**Val**
AC-badh	**Thr**	**Tyr**	Phe	**Phe**	**Tyr**	**Phe**	**Pro**	**Val**	**Tyr**	**Asn**	Leu
Tol-badh	**Thr**	Phe	Phe	**Phe**	**Tyr**	**Phe**	**Pro**	**Val**	**Ser**	**Asn**	Leu

had decreased V_{max} and V_{max}/K values for all of the alcohols tested and although the K_m values for butanol and pentanol were decreased in the mutant, these decreases did not compensate enough for the decreases in V_{max}. It appears that in the α subunits of the human and monkey alcohol dehydrogenases the change from leucine to the larger methionine at position 57 has been compensated for by the replacement of Leu-116 with the smaller valine residue in this position (Eklund *et al.*, 1987; Light *et al.*, 1992). Therefore horse and the human β subunit alcohol dehydrogenases have a combination of Leu-57:Leu-116, and human and monkey α subunits have a combination of Met-57:Val-116. Yeast alcohol dehydrogenase has a leucine in position 116. Weinhold & Benner (1995) created a yeast mutant with the combination Leu-57:Leu-116 with horse enzyme type substrate specificities, so it is possible that the change from Trp-57 to Met-57, bringing about a combination of Met-57:Leu-116, would not bring about a sufficient increase in space in the active site to bring about the change in substrate specificities. Benzyl alcohol dehydrogenases from *A. calcoaceticus* and encoded by the TOL-plasmid pWW0 have a combination of Tyr-57:Phe-116 (Fig. 3.5). This latter combination must surely result in the narrowing of the active site in this region compared to horse liver alcohol dehydrogenase.

The inner part of the substrate-binding cleft in the active sites of the benzyl alcohol dehydrogenases from *A. calcoaceticus* and *P. putida* (pWW0) may be smaller than that in horse liver alcohol dehdrogenase. In the horse enzyme 13 amino acid residues line what is apparently a deep and hydrophobic cleft (Eklund *et al.*, 1981). The inner part of the substrate-binding cleft is lined with the catalytic zinc-liganding residues (Cys and His) and residues Ser-48, Phe-93, Phe-140 and Leu-141 (Table 9.2; Fig. 9.1). In the benzyl alcohol dehydrogenases Ser-48 and Leu-141 are replaced by larger residues, Thr and Phe respectively (Fig 9.2). The *A. calcoaceticus* enzyme also replaces Phe-93 with Tyr (Fig. 3.5; Table 9.2), although when the horse liver and *A. calcoaceticus* enzymes are aligned together independently of the other sequences, Ser-89 (benzyl alcohol dehydrogenase numbering) aligns with Phe-93 (Fig. 9.2). It therefore appears that in this inner region the benzyl alcohol dehydrogenases may have four aromatic residues clustered together forming an internal hydrophobic core (Fig. 9.2). This hydrophobic core may be necessary to facilitate interaction with the hydrophobic aromatic ring. In the middle part of the substrate-binding cleft the horse liver enzyme is lined by Leu-57, Leu-116, Val-294 and Ile-318 (Table 9.2; Figs. 9.1 & 9.2). The leucine residues 57 and 116 have been replaced by tyrosine and phenylalanine respectively in the benzyl alcohol dehydrogenases. Also, Ile has been replaced by valine, and Val is conserved when the sequences are aligned independently or is replaced by proline in the Fig. 3.5 alignment. The outer part of

the horse liver enzyme substrate-binding cleft is lined by Phe-110, Met-306 and Leu-309. Leu-309 is conserved in the benzyl alcohol dehydrogenases and Met-306 is replaced by asparagine residues. Phe-110 is replaced by tyrosine in benzyl alcohol dehydrogenase from *A. calcoaceticus* and serine in the TOL-encoded enzyme (Table 9.2). Position 110 appears to be the only position within the putative substrate-binding cleft that the two benzyl alcohol dehydrogenases have what could be seen as significant differences. The replacement of Phe-110 with serine in TOL-encoded benzyl alcohol dehydrogenase has been proposed to make the mouth of the active site pocket larger (Shaw *et al.*, 1993), and since this position may be the only significant difference between the putative substrate-binding clefts of the two benzyl alcohol dehydrogenases, with the *A. calcoaceticus* enzyme having tyrosine at the mouth of the pocket, it is likely that this difference is responsible for the differing substrate specificities of the *A. calcoaceticus* and *P. putida* (pWW0) enzymes (Table 9.1). The TOL-encoded enzyme has a preference for substituted benzyl alcohols. It is possible that tyrosine at the mouth of the *A. calcoaceticus* enzyme pocket restricts the entry of substituted benzyl alcohols, notably 2-substituted benzyl alcohols. It would be interesting to sequence the gene encoding benzyl alcohol dehydrogenase from the TOL-plasmid pWW53 since this enzyme has substrate specificities more similar to the *A. calcoaceticus* enzyme than the pWW0 enzyme (Table 9.1; Chalmers *et al.*, 1990), to see which residue the pWW53 enzyme has at position 110. The presence of Leu-110 in class IV (σ subunit) human alcohol dehydrogenase, rather than Tyr as in the class I enzymes (α, β and γ subunits), has been proposed to be responsible for the increased ability of this class to oxidise all-*trans*-retinol (Kedishvili *et al.*, 1995). The comparisons in Table 9.2 and Figs. 9.1 and 9.2 show that the putative substrate-binding sites of the benzyl alcohol dehydrogenases from *A. calcoaceticus* and *P. putida* (pWW0) have a number of larger residues than those found in the horse liver enzyme. These changes may result in a narrower substrate-binding cleft.

Amino acid changes at positions 48 and 93 have been shown to influence substrate specificity. All of the long-chain zinc-dependent alcohol deydrogenases have a serine or threonine residue in position 48 (Fig. 3.5) and this residue has been proposed to be involved in the proton transfer mechanism of these enzymes (section 9.4.4; Eklund *et al.*, 1982). Yeast alcohol dehydrogenase has a Thr in position 48 and a leucine in position 93 in the alignments in Fig. 3.5 and Reid & Fewson (1994). However, as in the case of Trp-57, the model of Plapp *et al.* (1990) placed a tryptophan at position 93 of yeast alcohol dehydrogenase I. Other workers have worked on the premise that the yeast enzyme has a tryptophan at position 93 relative to horse liver alcohol

dehydrogenase (Creaser *et al.*, 1990; Green *et al.*, 1993), and considering the presence of an aromatic residue in the majority of alcohol dehydrogenases at this position, this premise seems likely. Creaser *et al.* (1990) changed Trp-93 of yeast alcohol dehydrogenase 1 to Phe, as can be found in horse liver alcohol dehydrogenase and the human β and γ subunits, and also to Ala, as is found in the human α chain (Table 9.2). The Trp-93:Phe mutant was slightly more active with a range of alcohols but the Trp-93:Ala mutant only showed an increased activity against hexanol. In the same study Thr-48 was changed to serine, as can be found in the horse liver and human γ subunit enzymes. These latter two enzymes oxidise longer chain primary alcohols, secondary alcohols including cyclohexanol, and benzyl alcohol more efficiently than the human β, π and χ subunits, which have a serine in position 48 (Eklund *et al.*, 1987; 1990), although the human α chain subunit, with Ala at position 93, oxidises secondary alcohols with the greatest efficiency (Stone *et al.*, 1989). Changing Thr-48 to Ser in the yeast enzyme produced an enzyme with an increased catalytic efficiency towards aliphatic alcohols with a chain length up to eight. Creaser *et al.* (1990) also made double mutants with Ser-48:Phe-93 and Ser-48:Ala-93. The Ser-48:Phe-93 enzyme showed slightly improved relative activites with substrates larger than ethanol, whereas the Ser-48:Ala-93 mutant exhibited kinetic parameters not very much different than the wild-type. None of the mutants created in this particular study had any activity towards cyclohexanol or aromatic alcohols, although the double mutants could now oxidise both enantiomers of 2-octanol. Green *et al.* (1993) managed to invert the substrate specificity of yeast alcohol dehydrogenase by mutations at positions 48 and 93. Single substitution of Ser for Thr-48 produced an enzyme with the same pattern of activity (V/K) as wild-type with straight chain primary alcohols, but with increased activity against branched chained secondary alcohols and also weak activity against benzyl alcohol. A single change to alanine at position 93, and a double mutant with both changes (Ser-48:Ala-93), produced mutant enzymes with inverted substrate specificities for primary alcohols from that of the wild-type yeast alcohol dehydrogenase. These enzymes had 3- and 10-fold more activity towards hexanol, 350- and 540-fold less activity towards ethanol, and had gained weak activity towards branched chain alcohols, cyclohexanol and benzyl alcohol.

Comparing the sequences of isozymes such as those of human class I alcohol dehydrogenase and the class II, III and IV enzymes with their substrate specificities provides "natural" site-directed mutagenesis experiments in which it is possible to pinpoint residues that influence the various substrate specificities. The human αα alcohol dehydrogenase oxidises straight and branched chain secondary alcohols and cyclohexanol far more efficiently than the horse liver and

human ββ and γγ enzymes, whereas the α subunit enzyme oxidises ethanol less efficiently (Stone *et al.*, 1989). The human γ subunit has a Ser in position 48 and Phe in position 93 (Table 9.2). This combination is the same as horse liver alcohol dehydrogenase, and these two enzymes share similar substrate specificities (Stone *et al.*, 1989; Eklund *et al.*, 1990). The major difference in the putative substrate-binding clefts of these four enzymes (human α, β, γ, and horse enzymes) is the replacement of Phe-93 with Ala in the human α subunit (Table 9.3). From deduction it appears that the alanine replacement is responsible for the increased ability of the αα human enzyme to oxidise secondary alcohols more efficiently, probably because it increases space within the active site, thus facilitating binding of the more bulky straight and branched chain secondary alcohols and cyclohexanol. This supports the findings of Green *et al.* (1993) in which mutants of yeast alcohol dehydrogenase which had Ala-93 had similar substrate specificities to human αα alcohol dehydrogenase.

The human β and γ alcohol dehydrogenase subunits differ at position 48, with β having a Thr and γ a Ser residue (Table 9.2). The β subunit enzyme is much more active towards ethanol and methanol than the γ subunit enzyme, whereas the γ subunit enzyme is more reactive towards cyclohexanol, benzyl alcohol and the steroid 3β-hydroxy-5β-androstan-17-one than the β enzyme (Eklund *et al.*, 1987; Stone *et al.*, 1989; Höög *et al.*, 1992). Site-directed mutagenesis in which Thr-48 of the human β subunit was replaced by Ser gave the mutant enzyme γ subunit properties (Höög *et al.*, 1992). The mutant enzyme had a dramatically increased K_m for methanol, had dramatically increased "specificity constants" for benzyl alcohol and cyclohexanol, and the mutant had gained the ability to oxidise 3β-hydroxy-5β-androstan-17-one.

It is interesting, considering that a Thr-48:Ser yeast alcohol dehydrogenase mutant had gained some activity against benzyl alcohol (Green *et al.*, 1993) and that the human γ subunit (Ser-48) has greater activity against benzyl alcohol than the β subunit (Thr-48), that the presence of serine at position 48, rather than threonine, seems to facilitate activity against benzyl alcohol since both the benzyl alcohol dehydrogenases from *A. calcoaceticus* and *P. putida* (pWW0) have a threonine in position 48 (Table 9.2). The benzyl alcohol dehydrogenases also have an additional aromatic residue in the inner part of their putative substrate-binding clefts (Phe-141) that will narrow the active site further. The substrate specificity of benzyl alcohol dehydrogenase from *A. calcoaceticus*, in which substituent groups on the benzene ring are preferred if situated away from the reactive carbinol group, i.e. at the *para* rather than the *ortho* position on the aromatic ring (Tables 5.2 & 9.1), suggests that the active site of the enzyme may have a cleft structure. The high level of sequence identity between benzyl alcohol

dehydrogenase from *A. calcoaceticus* and horse liver, human and yeast alcohol dehydrogenases in the active site area, with regard to the substrate-binding cleft and catalytic zinc ligands, suggests that the active site of the *A. calcoaceticus* enzyme has a similar structure to the active sites of these enzymes. Despite the overall similarity between the active site structures of these enzymes, the enzymes have different substrate specificities due to amino acid replacements that slightly alter the size and shape of the active site substrate-binding cleft. For example, yeast and horse liver alcohol dehydrogenases have active sites with very similar structures, despite the enzymes sharing only 25% sequence identity (Jörnvall *et al.*, 1978), yet the yeast enzyme has an active site which contains a number of bulky residues (Table 9.2) that reduce the volume of the substrate-binding cleft. It has been described how this decrease in space in the active site has been shown to restrict the substrate specificity of the enzyme to short chain primary alcohols (Green *et al.*, 1993; Weinhold & Benner, 1995). The amino acid replacements thus far described in the putative substrate-binding site of benzyl alcohol dehydrogenase from *A. calcoaceticus* compared to that of the human, yeast, and in particular, horse liver alcohol dehydrogenases suggest that the *A. calcoaceticus* enzyme has a narrower substrate-binding cleft than the horse liver enzyme. The benzyl alcohol deyhdrogenases from *A. calcoaceticus* and *P. putida* (pWW0) show the highest sequence conservation within their putative substrate-binding clefts with the human class II (π subunit) alcohol dehydrogenase (Table 9.2). As would be expected the human class II enzyme has similar substrate specificities to the benzyl alcohol dehydrogenases (Ditlow *et al.*, 1984). In particular the human class II enzyme exhibits strong activity with benzyl alcohol and 4-hydroxy-3-methoxybenzyl alcohol (vanillyl alcohol), the enzyme has very high K_m values for ethanol and cyclohexanol (120 mM and 210 mM respectively) and will not oxidise methanol.

Considering the substrate specificities of the human and yeast alcohol dehydrogenases it seems peculiar that neither of the benzyl alcohol dehydrogenases shows even a small amount of activity with ethanol or longer chain aliphatic alcohols (Table 5.2; MacKintosh & Fewson, 1988b; Shaw & Harayama, 1990). The human class II enzyme has the same residues lining the inner substrate-binding cleft as benzyl alcohol dehydrogenase from *A. calcoaceticus* but displays weak activity against ethanol and strong activity against pentanol and octanol (Ditlow *et al.*, 1984). This suggests that residues outside the active site are influencing substrate specificity. Larger substrate-binding sites tend not to favour the oxidation of small primary alcohols such as ethanol. Yeast alcohol deydrogenase has the smallest substrate-binding site and has the strongest activity towards ethanol. Also, increasing the size of the pocket

by exchanging Trp-93 with Ala decreased activity towards ethanol. The human αα alcohol dehydrogenase with Ala-93 also has greater activity towards ethanol than the ββ alcohol dehydrogenase (Phe-93). Large substrate-binding pockets are likely to be unfavourable towards activity with ethanol since the additional space may allow the substrate to rotate into positions that cannot transfer hydrogen. However, the substrate-binding clefts of the benzyl alcohol dehydrogenases are likely to be more constricted than horse and human β and γ subunits, yet the benzyl alcohol dehydrogenases have less activity against ethanol. These findings would seem to suggest that it might be an overall change in shape, rather than volume, of the benzyl alcohol dehydrogenase active sites compared to those of the horse and human enzymes that restricts the benzyl alcohol dehydrogenases to aromatic substrates. Residues outside the active site may well have an important influence on this shape.

9.4.4 The relationship between the primary structure of benzyl alcohol dehydrogenase from *A. calcoaceticus* and the catalytic mechanism of the enzyme

The conserved serine or threonine residues at position 48 in the inner substrate-binding cleft of the long-chain zinc-dependent alcohol dehydrogenases have been implicated in the catalytic mechanism of these enzymes (Eklund *et al.*, 1982). The two residues appear to be interchangeable as regards to the catalytic mechanism (Plapp *et al.*, 1990; Sakoda & Imanaka, 1992), with the "choice" of residue depending upon the required substrate specificity of the enzyme (section 9.4.3). Crystallographic studies in which horse liver alcohol dehydrogenase was crystallised with NAD+ and dimethylsulfoxide (DMSO), *p*-bromobenzyl alcohol or 2,3,4,5,6-pentafluorobenzyl alcohol have shown that the hydroxyl groups of the substrates are ligated to the catalytic zinc and linked to the imidazole group of His-51 through a hydrogen-bonded system containing the hydroxyl groups of Ser-48 and the nicotinamide ribose (Fig. 1.8; section 9.4.2; Eklund *et al.*, 1981; 1982; Ramaswamy *et al.*, 1994). His-51 participates in a further hydrogen-bonded system in which it is linked by a water molecule to the carboxyl group of Asp-50 and the hydroxyl group of Ser-54 (Fig. 1.8; section 1.10.2; Ramaswamy *et al.*, 1994). The His residue is proposed to act as a general base during catalysis by abstracting a proton from the alcohol substrate, with proton transfer from the alcohol substrate to His-51, and then to the surface of the enzyme, mediated by this hydrogen-bonding system (proton relay system; Eklund *et al.*, 1982). His-51 is conserved in the majority of long-chain zinc-dependent alcohol dehydrogenases

(Fig. 3.5). The plant dehydrogenases tend to have a tyrosine in this position, as is the case with the human χ subunit (Fig. 3.5; Eklund *et al.*, 1990). A tyrosine residue could fulfill the role of His-51 by hydrogen bonding to the 2'-hydroxyl group of the nicotinamide ribose (Eklund *et al.*, 1990). However, the benzyl alcohol dehydrogenases from *A. calcoaceticus* and *P. putida* (pWW0) have isoleucine and valine residues respectively in position 51 (Fig. 3.5). Also, the *xylW* product has Leu, *Pseudomonas* sp. alcohol dehydrogenase from the *terp* operon has Ala, and human class II alcohol dehydrogenase (π subunit) has Ser or Thr (the gene is polymorphic) in this position (Fig. 3.5; Eklund *et al.*, 1990). These residues are either incapable of hydrogen-bonding in the manner of His-51 (or Tyr-51) or, as in the case of Ser, the distance that would be involved to form the hydrogen bond is too great. It is interesting to note that class II human alcohol dehydrogenase has His-51 replaced and that this human isoenzyme has a substrate-binding cleft very similar to the benzyl alcohol dehydrogenases, coupled with similar substrate specificities (section 9.4.3).

The importance of Ser- or Thr-48 and His-51 in the catalytic mechanism of long-chain zinc-dependent alcohol dehydrogenases has been confirmed by site-directed mutagenesis studies. Replacement of Thr-48 in yeast alcohol dehydrogenase with Ala or Cys, and replacement of Thr-48 of the thermostable alcohol dehydrogenase from *B. stearothermophilus* and the human class III alcohol dehydrogenase with Ala, completely abolished alcohol dehydrogenase activity (Plapp *et al.*, 1990; Sakoda & Imanaka, 1992; Estonius *et al.*, 1994). It is therefore apparent that the hydroxyl group of the Ser or Thr residue is necessary for hydrogen-bonding to the alcohol substrate and the nicotinamide ribose (Fig. 1.8). His-51 has been substituted with Gln or Glu in yeast alcohol dehydrogenase and with Gln in the human ββ alcohol dehydrogenase (Plapp *et al.*, 1990; Hurley *et al.*, 1990b; Ehrig *et al.*, 1991). These two residues are of an appropriate size to form the hydrogen bond with the nicotinamide ribose, thus enabling similar coenzyme binding as in the wild-type enzyme. However Gln would not be able to participate in base catalysis, whereas Glu could potentially accept a proton. In yeast alcohol dehydrogenase the catalytic efficiency of the mutants was greatly reduced compared to the wild-type and the pH dependency of the mutants was such that the pK value of 7.7 for wild-type was no longer visible. In the human alcohol dehydrogenase the catalytic efficiency of alcohol oxidation was reduced 6-fold in the His-51:Gln mutant. Activity of this mutant could be restored to wild-type levels in the presence of high concentrations of glycylglycine, i.e. in the presence of general base catalysts in the solvent. His-51 of the thermostable alcohol dehydrogenase from *B. stearothermophilus* has been mutated to Ala and Arg (Sakoda & Imanaka, 1992).

Ala would be unable to hydrogen-bond in the manner of His-51, and indeed the resulting mutant enzyme had no activity. An arginine residue could conceivably hydrogen bond and participate in proton abstraction. In this case the His-51:Arg mutant was fully active, but had an altered pH profile such that the enzyme had maximum activity at pH 9.0 rather than 7.8 as is the case in the wild-type. Hence the pK_a of the guanidino group of Arg-51 was now influencing the pH dependence of the enzyme rather than the pK_a of the imidazole ring of His-51. These results are all consistent with the proposal that His-51 acts as a general base catalyst during alcohol oxidation by these enzymes.

The components of the proton relay system of the long-chain zinc-dependent alcohol dehydrogenases are conserved in the benzyl alcohol dehydrogenases of *A. calcoaceticus* and *P. putida* (pWW0), with the exception of His-51 or a residue which could conceivably be a functional replacement (Fig. 9.2). Since the substrate-binding and coenzyme-binding clefts of the benzyl alcohol dehydrogenases appear to be similar to those of the other zinc-dependent alcohol dehydrogenases, and the overall structures of all these enzymes appear to be similar (sections 9.2.2 & 9.2.3), the fact that the proton relay system must be rearranged in the benzyl alcohol dehydrogenases is surprising. This situation also applies to the human class II (π subunit) alcohol dehydrogenase. Therefore, for some reason, in the evolution of benzyl alcohol and human class II alcohol dehydrogenases a catalytically important residue has been selected against. Reasons for the selection against His- or Tyr-51 were investigated by introducing His-51 into benzyl alcohol dehydrogenase from *A. calcoaceticus* (section 5.1).

Characterisation of the purified benzyl alcohol dehydrogenase-His51 mutant showed that the enzyme had somewhat greater K'_m and V'_{max} values than wild-type enzyme, although this resulted in similar "specificity constants" for the two enzymes (section 5.3; Table 5.1). Therefore the two enzymes oxidise benzyl alcohol with equal efficiency. The increased K'_m of the Ile-51:His mutant can probably be explained as being a result of the more bulky His restricting space within the substrate-binding cleft and so slightly hindering substrate binding. The increased V'_{max} may be expained by a perturbation of coenzyme binding. In most zinc-dependent alcohol dehydrogenases His-51 binds to the nicotinamide ribose of the coenzyme (section 9.4.2). Obviously this facet of coenzyme binding is altered in benzyl alcohol dehydrogenase, so the "reintroduced" His-51 may interact with the coenzyme, and/or restrict its binding, consequently increasing the dissociation constant of NADH. It would be interesting to determine the NADH dissociation constants of these enzymes to see if this is indeed the case.

Since the benzyl alcohol dehydrogenases of *A. calcoaceticus* and *P. putida* (pWW0) have been shown, in contrast to the other enzymes in the group of

long-chain zinc-dependent alcohol dehydrogenases, to be active almost exclusively against aromatic substrates, and have strong activity against various substituted aromatic compounds (Table 1.2; MacKintosh & Fewson, 1988b; Shaw & Harayama, 1990), it is reasonable to suggest that the replacement of His-51 with the less bulky isoleucine or valine residues was a result of selection for activity against substituted aromatics which are likely to be prevalent in the environment from the breakdown of lignin by fungi (Cain, 1980). This suggestion is supported by the fact that the human class II alcohol dehydrogenase isoenzyme has Ser-51 in place of His-51 and this enzyme has strong activity against benzyl alcohol and substituted benzyl alcohols (section 9.4.3; Ditlow et al., 1984). The oxidation of substituted benzyl alcohols has not been widely reported for other human alcohol dehydrogenase isoenzymes. The activity of benzyl alcohol dehydrogenase-His51 with various substituted benzyl alcohols, cinnamyl alcohol, coniferyl alcohol and the aliphatic compound perillyl alcohol relative to activity with benzyl alcohol was determined (section 5.4; Table 5.2). The activity of the Ile-51:His mutant with these compounds relative to activity with benzyl alcohol was virtually identical to the pattern of activity determined with these compounds for the wild-type enzyme (Table 5.2). It therefore seems that selection against His-51 was not due to the selection for activity against substituted benzyl alcohols. In addition, benzyl alcohol dehydrogenase-His51 had no activity with methanol or ethanol. Therefore the His-51 replacement did not alter the size or shape of the substrate-binding cleft such that it more closely resembled horse liver, or the human β and γ, alcohol dehydrogenase subunits, thus facilitating ethanol oxidation.

Benzyl alcohol dehydrogenase-His51 showed a pH dependence similar to that of wild-type benzyl alcohol dehydrogenase (section 5.5; Fig. 5.6). This would suggest that the insertion of His-51 has not significantly altered the pK_a of any proton release group involved in benzyl alcohol oxidation. Replacements at position 51 have been suggested to correlate with differences in activity profiles as a function of pH, with His-51 enzymes having high activity at pH 7.5 compared to pH 10.0 and enzymes with Tyr- or Ser-51 having low activity at pH 7.5 (Juliá et al., 1988). These factors seem to suggest that it is unlikely that the His-51 residue of the mutant enzyme is contributing to the mechanism of benzyl alcohol oxidation in the way that it does in horse liver alcohol dehydrogenase. Shaw et al. (1993) showed that for benzyl alcohol dehydrogenase encoded by the TOL-plasmid pWW0 , unlike the human $\beta\beta$ alcohol dehydrogenase His-51:Gln mutant, the concentration and pK_a of solvent proton acceptors had no effect on the catalytic efficiency of benzyl alcohol dehydrogenase. This suggests that benzyl alcohol dehydrognase has a proton transfer mechanism that is able to remove a

proton from the active site independently of the solvent. It is therefore unlikely that there is direct transfer of the proton to solvent or indirect hydrogen-bonding of a residue to the 2'-hydroxyl of the nicotinamide ribose and Thr-48 via a water molecule, as is the case for glyceraldehyde-3-phosphate dehydrogenase (Skarzynski *et al.*, 1987) and has been proposed to occur in human class II alcohol dehydrogenase (Eklund *et al.*, 1990). The latter proposal suggested that a water molecule might indirectly link the 2'-hydroxyl group of the nicotinamide ribose with Ser-51, thus reconstituting the proton relay system. However, studies with recombinant human class II alcohol dehydrogenase with Ser-51, Thr-51 or His-51, the latter introduced by site-directed mutagenesis, showed that the catalytic efficiency of these enzymes was not influenced by the presence of solvent proton acceptors (Davis *et al.*, 1994). Also, the "reintroduction" of His-51 did not alter the efficiency of ethanol oxidation (V_{max}/K_m). These results reflect those presented in this study for the *A. calcoaceticus* enzyme and suggest that it is unlikely that the proton relay sysytem is completed via a water molecule in human class II alcohol dehydrogenase or that Ser-51 is participating in the catalytic mechanism of this enzyme. It is not possible to reconstitute the horse liver alcohol dehydrogenase proton relay system in the benzyl alcohol dehydrogenases since Ile- or Val-51 are not capable of hydrogen-bonding because of their aliphatic side chains. This proton relay system, either via a water molecule or not, would have to be reconstituted to a position other than residue 51. The results of Shaw *et al.* (1993), Davis *et al.* (1994) and those presented here do not support the suggestion that Ser/Thr-48 and His/Tyr-51 as a proton relay system in alcohol dehydrogenases is essential in general. It appears that in the benzyl alcohol dehydrogenases of *A. calcoaceticus* and *P. putida* (pWW0), and human class II alcohol dehydrogenase there is a different mechanism for transferring a proton from the hydrophobic alcohol-binding site to solvent than in the majortiy of zinc-dependent alcohol dehydrogenases.

9.5 Investigation of residues that may act as a general base in proton transfer in benzyl alcohol dehydrogenase from *A. calcoaceticus*

9.5.1 A possible role for histidine residues in positions other than position 51?

The possibility remained that any proton transfer mechanism in the benzyl alcohol dehydrogenases involved a histidine residue. Benzyl alcohol dehydrogenase from *A. calcoaceticus* contains eight histidine residues per enzyme

subunit. These are found at positions 44 (47; horse liver ADH numbering), 64 (67), 85 (88), 124 (128), 129 (133), 137 (139), 237 (239) and 358 (363; Figs. 3.4 & 3.5). Residues 47 and 67 (horse liver numbering) have roles in coenzyme binding and as a catalytic zinc ligand (sections 9.2 & 9.4.2; Eklund *et al.*, 1976; 1984). None of the other histidine residues is conserved across the whole group of zinc-dependent alcohol dehydrogenases (Fig. 3.5): His-139 is conserved in horse liver, human, and the plant alcohol dehydrogenases, but not benzyl alcohol dehydrogenase from *P. putida* (pWW0); His-88 is conserved in the plant alcohol dehydrogenases and TOL-benzyl alcohol dehydrogenase; His-133 and His-239 are conserved in TOL-benzyl alcohol dehydrogenase; and His-128 and His-363 are unique to benzyl alcohol dehydrogenase from *A. calcoaceticus*. Conservation between the group of zinc-dependent alcohol dehydrogenases is not a requirement for the identification of a histidine residue in benzyl alcohol dehydrogenase from *A. calcoaceticus* that might fulfill the role His-51 plays in horse liver alcohol dehydrogenase. It could also be misleading to use the structure of horse liver alcohol dehydrogenase as a paradigm for identifying which of these histidine residues could be situated in a position that might enable it to act as part of a proton transfer mechanism since the *A. calcoaceticus* enzyme is likely to have structural rearrangements to account for a different proton transfer mechanism.

The histidine modifying agent diethyl pyrocarbonate was used to modify benzyl alcohol dehydrogenase from *A. calcoaceticus* (section 6.3). The enzyme was inactivated in the presence of 1 mM diethyl pyrocarbonate at pH 6.0 (Fig. 6.3b). Inactivation correlated with an increase in absorbance at 242 nm (Fig. 6.3), and carbethoxylated benzyl alcohol dehydrogenase could be reactivated by treatment with hydroxylamine (section 6.4). Inactivation must therefore have been due to the modification of histidine residues. Diethyl pyrocarbonate has previously been used to inactivate various dehydrogenases, including yeast and horse liver alcohol dehydrogenases (Table 9.3). Inactivation of horse liver alcohol dehydrogenase with diethylpyrocarbonate (Hennecke & Plapp, 1983) was repeated here as a positive control and to provide a direct comparison with carbethoxylation of benzyl alcohol dehydrogenase (section 6.2; Figs. 6.1 & 6.2). Horse liver alcohol dehydrogenase was inactivated more rapidly than benzyl alcohol dehydrogenase and could be totally inactivated at pH 8.0, whereas benzyl alcohol dehydrogenase retained about 25% of its original activity after complete carbethoxylation (Figs. 6.1b, 6.3b & 6.4). A thorough investigation of the pH dependence of benzyl alcohol dehydrogenase inactivation might result in complete activation, although carbethoxylated enzyme may be partially active. Complete inactivation of yeast alcohol dehydrogenase could be achieved using a large excess of diethtyl pyrocarbonate at pH 7.0 (Dickinson & Dickenson, 1977).

Table 9.3 Comparison of the carbethoxylation of various dehydrogenases

Dehydrogenase	Source	Inactivation by diethyl pyrocarbonate			Reactivation by hydroxylamine		
		pH of reaction	% residual activity[j]	Max no. His modified[j]	% original activity	NH_2OH conc (M)	Time
Alcohol[a]	Horse liver	6.5–10.0	0	4.8	70	0.5	24 h
Alcohol[b]	Yeast	7.0	5	2.5	78	NQ	7 d
Saccharopine[c]	Yeast	6.9	10	3.0	80	0.45	1 h
Octopine[d]	*Pecten maximus*	6.0	0	2.0	96	0.75	24 h
Lactate[e]	Pig heart	6.0	3	1.0	NQ	NQ	NQ
Malate[f]	*E. coli*	6.5	5	16.0	74	1.0	1 h
Malate[g]	Sorghum leaf	6.6	10	NQ	80	0.5	5 min
Dihydro-orotate[h]	Human	7.2	0	NQ	NQ	NQ	NQ
Benzyl alcohol[i]	*A. calcoaceticus*	6.0	25	4.2	73	0.5	23 h

NQ = not quoted

Data taken from: [a]Hennecke & Plapp (1983); [b]Dickenson & Dickinson (1975); [c]Fujioka *et al.* (1980); [d]Huc *et al.* (1971); [e]Holbrook & Ingram (1973); [f]Wright *et al.* (1995); [g]Lemaire *et al.* (1994); [h]Copeland *et al.* (1995); [i]This study.
[j]Percentage of residual activity as measured when the maximal number of histidine residues were modified.

However, large excesses of diethyl pyrocarbonate were avoided in the present work since this can lead to the formation of a disubstituted histidyl derivative leading to an overestimation of the number of modified histidine residues (Miles, 1977).

Substrate protection was observed during inactivation of benzyl alcohol dehydrogenase with diethyl pyrocarbonate (Figs. 6.3 & 6.4). NADH gave the greatest protection, followed by NAD$^+$, with benzyl alcohol providing the least protection. The maximum number of histidine residues modified was reduced from 4.2 per enzyme subunit in the absence of substrate, to 2.7 per enzyme subunit in the presence of any of the substrates. These results suggest inactivation of benzyl alcohol dehydrogenase with diethyl pyrocarbonate is due to the modification of an active site histidine residue(s).

It is apparent that modification of an active site histidine, or histidines, results in the inactivation of benzyl alcohol dehydrogenase from *A. calcoaceticus* but difficulty arises when trying to interpret which histidine residue is being modified. Two histidine residues are already thought to have important roles. His-67 is a catalytic zinc ligand (sections 9.2 & 9.4.1) and His-47 is involved in binding the pyrophosphate bridge of the coenzyme (section 9.4.2; Fig. 9.2). Modification of yeast alcohol dehydrogenase with diethyl pyrocarbonate indicated that one essential histidine residue was being modified which resulted in the inactivation of the enzyme (Dickenson & Dickinson, 1975; 1977). However, these latter studies could not determine which histidine residue was the proposed essential residue. Extensive studies in which horse liver alcohol dehydrogenase was modified with diethyl pyrocarbonate suggested that inactivation of the enzyme was due to the modification of one, but possibly two, histidine residues with carbethoxylation of His-51 proposed as the most likely reason for inactivation (Hennecke & Plapp, 1983). The horse liver enzyme has His-47 replaced by Arg (Fig. 9.1), but the modification of His-67 remained a possibility. In the present study the strong protection against inactivation provided by NAD$^+$ and NADH suggests that inactivation was caused by modification of a histidine residue involved with coenzyme binding, probably His-47. This could also explain the weaker protection afforded by benzyl alcohol, which would be bound at the active site and so hinder the access and binding of diethyl pyrocarbonate, but unlike coenzyme would not be bound to His-47. Carbethoxylation of His-47 may not completely prevent coenzyme binding. This may explain why the modified enzyme is partially active. Modification of His-67 was considered unlikely in the yeast and horse liver enzymes because of the bound zinc. This same argument could apply here. A histidine residue responsible for the direct transfer of a proton from substrate, as is the case for His-195 of lactate

dehydrogenase (Holbrook *et al.*, 1975), would probably be protected from modification more thoroughly by substrate than the protection given by benzyl alcohol in this study. Lactate dehydrogenase is inactivated by diethyl pyrocarbonate (Table 9.3) and this inactivation is protected against completely by substrate and substrate analogues, but is only slightly protected against by NADH (Holbrook & Ingram, 1973). All of the substrates used in this thesis protected against the modification of 1.5 histidines (Fig. 6.4). It is possible that besides protecting an essential histidine, the substrates could also protect a non-essential histidine residue, positioned in the vicinity of the active site, against carbethoxylation.

Results of carbethoxylation by diethyl pyrocarbonate can be difficult to interpret in terms of determining how many histidine residues are essential for activity and identifying which residues these are. The results presented here indicate that a histidine residue situated in, or close to, the coenzyme-binding site of benzyl alcohol dehydrogenase can be carbethoxylated by diethyl pyrocarbonate and that this modification inactivates the enzyme. It is likely that this residue is His-47, a residue responsible for binding the pyrophosphate moiety of the coenzyme (section 9.4.2). Comparison of these results with those for the inactivation of lactate dehydrogenase suggest that direct transfer of a proton from substrate to a histidine residue is unlikely in benzyl alcohol dehydrogenase. It remains a possibility that benzyl alcohol dehydrogenase from *A. calcoaceticus* has a proton relay system analagous to that proposed for horse liver alcohol dehydrogenase involving a histidine residue in a position other than 51, although the results presented here suggest that this is not the case.

9.5.2 A possible role for arginine-53 in proton transfer?

Since it seems likely that a residue other than a histidine is involved in proton transfer in benzyl alcohol dehydrogenase other residues that could potentially act as a general base in catalysis had to be considered and investigated. The benzyl alcohol dehydrogenases from *A. calcoaceticus* and *P. putida* (pWW0) both have an arginine residue at position 53 (Fig. 3.5). Interestingly, the alcohol dehydrogenase from the *terp* operon of *Pseudomonas* sp., an enzyme with Ala-51, also has a basic residue in this position (Lys-53). Replacement of His-51 of the thermostable alcohol dehydrogenase from *B. stearothermophilus* with Arg resulted in an enzyme that was fully active but which had an alkaline pH optimum (Sakoda & Imanaka, 1992). Also, the replacement of His-51 with other residues has been correlated with alkaline rather than acidic pH optima (Julià *et al.*, 1988).

Since benzyl alcohol dehydrogenase from *A. calcoaceticus* has a pH optimum of 9.2 (MacKintosh & Fewson, 1988a; Fig. 5.6) the possibility that Arg-53 is involved in proton transfer was investigated by site-directed mutagenesis.

Arg-53 was changed to Ala-53 or His-53 (section 7.1). The Ala-53 mutant enzyme was found to be fully active, with K'_m and V'_{max} values for benzyl alcohol oxidation in crude extracts very similar to those of the wild-type enzyme (section 7.3; Table 7.2). This result indicates that Arg-53 is not involved in proton transfer in benzyl alcohol dehydrogenase, since Ala-53 would not be able to act as a general base in catalysis. However, replacement of Arg-53 with His resulted in an enzyme with a "specificity constant" for benzyl alcohol oxidation three orders of magnitude lower than wild-type enzyme (section 7.4; Table 7.3). This would suggest that replacement of Arg-53 with a more bulky histidine residue had altered the protein structure around the active site area such that access or binding of benzyl alcohol was hindered. Since benzyl alcohol dehydrogenase-His53 appeared to have an active site with a shape and size altered from that of wild-type enzyme the substrate specificity of this enzyme was investigated.

Rates of activity relative to the rate obtained for the oxidation of benzyl alcohol for benzyl alcohol dehydrogenase-His53 with various substrates showed a pattern of activity similar to wild-type enzyme for cinnamyl and coniferyl alcohols and 3-methylbenzyl alcohol (section 7.4; Table 7.4). Relative activity with 2-methylbenzyl alcohol was slightly reduced, but was slightly increased with 4-methylbenzyl alcohol. However, the relative activity of benzyl alcohol dehydrogenase-His53 with perillyl alcohol was 2.8 times higher than the rate with benzyl alcohol. Determination of the kinetic coefficients for perillyl alcohol oxidation for benzyl alcohol dehydrogenase-His53 and the wild-type enzyme showed that the mutant enzyme has a "specificity constant" about 350-fold higher for perillyl alcohol oxidation than for benzyl alcohol oxidation (Table 7.3). Wild-type benzyl alcohol dehydrogenase has a "specificity constant" for perillyl alcohol oxidation about 2.6 times that for benzyl alcohol oxidation. These results suggest that replacement of Arg-53 with a histidine residue has altered the shape of the substrate-binding pocket such that the aliphatic cyclohex-1-ene compound perillyl alcohol (Appendix VII) can bind better than benzyl alcohol, although not as well as in the wild-type enzyme. Despite the fact that the mutant enzyme appears to have a more restricted substrate-binding pocket the enzyme could not oxidise the short chain aliphatic alcohols methanol and ethanol.

## 9.6 Final conclusions and future work

It appears that the benzyl alcohol dehydrogenases from *A. calcoaceticus* and *P. putida* (pWW0) constitute part of a group of long-chain zinc-dependent alcohol dehydrogenases that have a mechanism of proton transfer that is altered from that proposed for horse liver alcohol dehydrogenase and which was thought to be essential in general for this family of enzymes. It seems likely that this mechanism either involves direct transfer of a proton to a general base other than a histidine residue, or involves a proton relay system in which a basic residue acts as a general base for catalysis. This latter proposal cannot exclude the possibility that a histidine residue is acting as the general base. Future work on benzyl alcohol dehydrogenases should focus on the elucidation of the mechanism of proton transfer in these enzymes. The best way to approach this would be by determining the crystal structure of benzyl alcohol dehydrogenase, which if not providing an answer by itself should enable the identification of residues likely to be involved in proton transfer and so enable a rational approach to site-directed mutagenesis experiments. Crystallography studies would be facilitated by the expression system for benzyl alcohol dehydrogenase from *A. calcoacticus* developed in this study (section 4.1) which allowed a gross over-expression of the enzyme in *E. coli* and enabled the purification of tens of mg quantities of enzyme from less than 10 g of cells. Purification could be improved further by replacing the Blue-Sepharose CL-6B affinity column with one that binds benzyl alcohol dehydrogenase more stongly, thus enabling the full utilisation of the expression system (section 4.2). Once crystals of benzyl alcohol dehydrogenase have been grown, elucidation of the three-dimensional structure would be facilitated by the structures already published for horse liver alcohol dehydrogenase (Ramaswamy *et al.*, 1994) and human β alcohol dehydrogenase variants (Hurley *et al.*, 1994), since these structures would enable the use of molecular replacement techniques rather than having to obtain heavy atom derivatives which can be very time consuming.
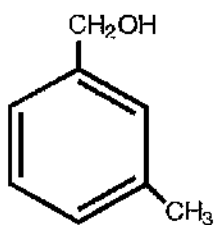
Appendix

Alcohol	Structure number
Benzyl alcohol	I
2-Methylbenzyl alcohol	II
3-Methylbenzyl alcohol	III
4-Methylbenzyl alcohol	IV
Coniferyl alcohol	V
Cinnamyl alcohol	VI
Perillyl alcohol	VII

CH₂OH

CH_2OH

I

CH_2OH CH₃

II

CH_2OH

CH_3

III

CH_2OH

CH_3

IV

CH_2OH

OCH_3

OH

V

CH_2OH

VI

CH_2OH
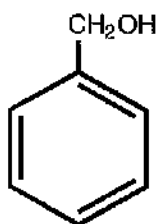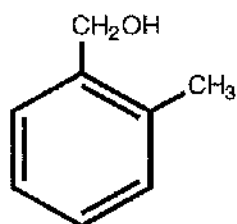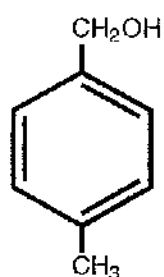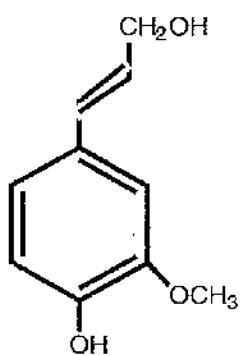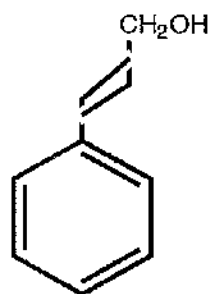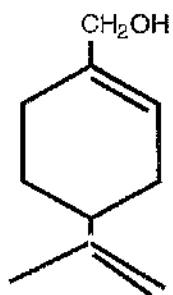
VII

References

Allison, N., O'Donnell, M. J. and Fewson, C. A. (1985a) Biochem. J. **227**, 753-757

Allison, N., O'Donnell, M. and Fewson, C. A. (1985b) Biochem. J. **231**, 407-416

Amábile-Cuevas, C. F. and Chicurel, M. E. (1993) American Scientist **81**, 332-341

Andersson, P., Kvassman, J., Lindström, A., Oldén, B. and Pettersson, G. (1981) Eur. J. Biochem. **113**, 425-433

Aronson, B. D., Somerville, R. L., Epperly, B. R. and Dekker, E. E. (1989) J. Biol. Chem. **264**, 5226-5232

Assinder, S. J. and Williams, P. A. (1990) Adv. Microbiol. Phys. **31**, 1-69

Baker, D. P. and Fewson, C. A. (1989) J. Gen. Microbiol. **135**, 2035-2044

Ballal, N. R., Bhattacharyya, P. K. and Rangachari, P. N. (1966) Biochem. Biophys. Res. Commun. **23**, 473-478

Ballal, N. R., Bhattacharyya, P. K. and Rangachari, P. N. (1967) Biochem. Biophys. Res. Commun. **29**, 275-280

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

Baumann, P., Doudoroff, M. and Stanier, R. Y. (1968) J. Bacteriol. **95**, 1520-1541

Beggs, J. D., Cook, A. M. and Fewson, C. A. (1976) J. Gen. Microbiol. **96**, 365-374

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

Biegert, T. and Fuchs, G. (1995) Arch. Microbiol. **163**, 407-417

Biegert, T., Altenschmidt, U., Eckerskorn, C. and Fuchs, G. (1995) Arch. Microbiol. **163**, 418-423

Birnboim, C. and Doly, J. (1979) Nucleic Acids Res **7**, 1513-1523

Boehringer. (1976) Biochemica Service, Issue 6b

Borrás, T., Persson, B. and Jörnvall, H. (1989) Biochemistry **28**, 6133-6139

Bowen, R. and Pugh, S. (1985) Chem. and Ind. **10**, 323-326

Bränden, C.-V., Jörnvall, H., Eklund, H. and Furugren, B. (1975) In: The enzymes, vol. 11 (Boyer, P. D., ed.), pp. 103-190, Academic Press, New York

Brendel, V. and Trifonov, E. N. (1984) Nucl. Acids. Res. **12**, 4411-4427

Bühler, R. and von Wartburg, J.-P. (1982) FEBS Letts. **144**, 135-139

Cain, R. (1980) In: Lignin biodegradation: Microbiology, chemistry, and potential applications., vol. 1, CRC Press, Florida

Cassan, M., Parsot, C., Cohen, G. N. and Patte, J.-C. (1986) J. Biol. Chem. **261**, 1052-1057

Chalmers, R. M. and Fewson, C. A. (1989) Biochem. J. **263**, 913-919

Chalmers, R. M., Scott, A. J., and Fewson, C. A. (1990) J. Gen. Microbiol. **134**, 637-643

Chalmers, R. M., Keen, J. N. and Fewson, C. A. (1991) Biochem. J. **273**, 99-107

Cho, S. H., Ryu, J. W. and Lee, K. M. (1995) Arch. Pharm. Res. **18**, 100-104
Clark, D. P. (1992) In: The evolution of metabolic function (Mortlock, R. P., Ed.) pp 105-114, CRC Press

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

Cook, A. M., Beggs, J. D and Fewson, C. A. (1975) J. Gen. Microbiol. **91**, 325-377

Copeland, R. A., Davis, J. P., Dowling, R. L., Lombardo, D., Murphy, K. B. and Patterson, T. A. (1995) Arch. Biochem. Biophys. **323**, 79-86

Cornish-Bowden, A. and Eisenthal, R. (1974) Biochem. J. **139**, 721-730

Crawford, R. L. and Olsen, P. E. (1979) FEMS Microbiol. Letts. **5**, 193-195

Creaser, E. H., Murali, C. and Britt, K. A. (1990) Prot. Engng. **3**, 523-526

Dagley, S. (1978) In: "The bacteria" Vol. VI Bacterial diversity (Ornston, L. N. and Sokatch, J. R., Eds.) pp 305-388, Academic Press, London and New York

Dagley, S. (1981) In: Microbial degradation of xenobiotics and recalcitrant compounds. (Leisinger, T., Cook, A. M., Hutter, R. and Nuesch, J., eds.), pp 181-188, Academic Press., London

Dalziel, K. (1963) J. Biol. Chem. **238**, 2850-2858

Dalziel, K. (1975) In: The enzymes, vol. XI (Boyer, P. D., ed.), pp. 1-60, Academic Press, New York

Davis, G. J., Carr, L. G., Hurley, T. D., Li, T.-K. and Bosron, W. F. (1994) Arch. Biochem. Biophys. **311**, 307-312

Dennis, E. S., Gerlach, W. L., Pryor, A. J., Bennetzen, J. L., Inglis, A., Llewellyn, D., Sachs, M. M., Ferl, R. J. and Peacock, W. J. (1984) Nucl. Acids. Res. **12**, 3983-4000

Dickenson, C. J. and Dickinson, F. M. (1975) Eur. J. Biochem. **52**, 595-603

Dickenson, C. J. and Dickinson, F. M. (1977) Biochem. J. **161**, 73-82

Ditlow, C. C., Holmquist, B., Morelock, M. M. and Vallee, B. L. (1984) Biochemistry **23**, 6363-6368

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

Doten, R. C., Ngai, K., Mitchell, D. J. and Ornston, L. N. (1987a) J. Bacteriol. **169**, 3168-3174

Doten, R. C., Gregg, L. A. and Ornston, L. N. (1987b) J. Bacteriol. **169**, 3175-3180

Durham, D. R. (1984) J. Bacteriol. **160**, 778-780

Ehrig, T., Hurley, T. D., Edenberg, H. J. and Bosron, W. F. (1991) Biochemistry **30**, 1062-1068

Ehrt, S., Ornston, L. N. and Hillen, W. (1994) J. Bacteriol. **176**, 3493-3499

Eisenthal, R. and Cornish-Bowden, A. (1974) Biochem. J. **139**, 715-720

Eklund, H., Nordstrom, B., Zeppezauer, E., Soderlund, G., Ohlsson, I., Boiwe, T. and Brändén, C. (1974) FEBS Letts. **44**, 200-204

Eklund, H., Nordstrom, B., Zeppezauer, E., Soderlund, G., Ohlsson, I., Boiwe, T., Soderberg, B., Tapia, O., Brändén, C. and Åkeson, Å. (1976) J. Mol. Biol. **102**, 27-59

Eklund, H. and Brädén, C. (1979) J. Biol. Chem. **254**, 3458-3461

Eklund, H., Samama, J., Wallen, L., Brändén, C., Åkeson, Å. and Jones, T. A. (1981) J. Mol. Biol. **146**, 561-587

Eklund, H., Plapp, B. V., Samama, J. and Brändén, C. (1982) J. Biol. Chem. **257**, 14349-14358

Eklund, H., Horjales, E., Jörnvall, H., Brändén, C.-I. and Jeffrey, J. (1985) Biochemistry **24**, 8005-8012

Eklund, H., Samama, J.-P. and Jones, T. A. (1984) Biochemistry **23**, 5982-5996

Eklund, H., Horjales, E., Vallee, B. L. and Jörnvall, H. (1987) Eur. J. Biochem. **167**, 185-193

Eklund, H. (1988) Biochem. Soc. Trans. **17**, 293-296

Eklund, H., Müller-Wille, P., Horjales, E., Futer, O., Holmquist, B., Vallee, B. L., Höög, J.-O., Kaiser, R. and Jörnvall, H. (1990) Eur. J. Biochem. **193**, 303-310

Estonius, M., Höög, J.-O., Danielsson, O. and Jörnvall, H. (1994) Biochemistry **33**, 15080-15085

Evans, W. C. and Fuchs, G. (1988) Annu. Rev. Microbiol. **42**, 289-317

Fan, F., Lorenzen, J. A. and Plapp, B. V. (1991) Biochemistry **30**, 6397-6401

Fan, F. and Plapp, B. V. (1995) Biochemistry **34**, 4709-4713

Farrés, J., Moreno, A., Crosas, B., Peralba, J. M., Allali-Hassani, A., Hjelmqvist, L., Jörnvall, H. and Parés, X. (1994) Eur. J. Biochem. **224**, 549-557

Favaloro, J., Treisman, R. and Kamen, R. (1980) Methods in Enzymology **65**, 718-749

Fersht, A. (1985) Enzyme structure and function, W. H. Freeman & Co., San Francisco

Fewson, C. A. (1967) J. Gen. Microbiol. **46**, 255-266

Fewson, C. A., Livingstone, A. and Moyes, H. M. (1978) J. Gen. Microbiol. **106**, 233-239

Fewson, C. A. (1981) In: Microbial degradation of xenobiotics and recalcitrant compounds (Leisinger, T., Cook, A. M., Hutter, R. and Nuesch, J., Eds.), pp 141-179, Academic Press

Fewson, C. A. (1985) J. Gen. Microbiol. **131**, 865-872

Fewson, C. A. (1988) FEMS Microbiol. Rev. **54**, 85-110

Fewson, C. A. (1991) In: The biology of *Acinetobacter* (Towner, K. J., Bergogne-Berezin, E. and Fewson, C. A., Eds.) pp 351-390, Plenum Press, New York

Fewson, C. A. (1992) In: The evolution of metabolic function (Mortlock, R. P. Ed.) pp 115-141, CRC Press

Fewson, C. A., Baker, D. P., Chalmers, R. M., Keen, J. N., Hamilton, I. D., Scott, A. J. and Yasin, M. (1993) J. Gen. Microbiol. **139**, 1345-1352

Fujioka, M., Takata, Y., Ogawa, H. and Okamoto, M. (1980) J. Biol. Chem. **255**, 937-942

Ganzhorn, A. J. and Plapp, B. V. (1988) J. Biol. Chem. **263**, 5446-5454

Gould, R. M. and Plapp, B. V. (1990) Biochemistry **29**, 5463-5468

Green, D. W., Sun, H.-W. and Plapp, B. V. (1993) J. Biol. Chem. **268**, 7792-7798

Grima-Pettenati, J., Feuillet, C., Goffner, D., Borderies, A. M. and Boudet, A. M. (1993) Plant Mol. Biol. **21**, 1085-1095

Haddock, J. D. and Ferry, J. G. (1993) J. Bacteriol. **175**, 669-673

Harayama, S., Leppik, R. A., Rekik, M., Mermod, N., Lehrbach, P. R., Reineke, W. and Timmis, K. N. (1986) J. Bacteriol. **167**, 455-461

Harayama, S., Rekik, M., Wubbolts, M., Rose, K., Leppik, R. A. and Timmis, K. N. (1989) J. Bacteriol. **171**, 5048-5055

Harayama, S. and Timmis, K. N. (1989) In: Genetics of bacterial diversity. (Chater, K. F. and Hopwood, D. A. eds.), pp 151-174, Academic Press, London

Harayama, S. and Rekik, M. (1990) Mol. Gen. Genet. **221**, 113-120

Hartmeier, W. (1985) Trends in Biotechnol. **3**, 149-153

Hartnett, G. B. and Ornston, L. N. (1994) Gene **142**, 23-29

Hegeman, G. D. (1966a) J. Bacteriol. **91**, 1140-1154

Hegeman, G. D. (1966b) J. Bacteriol. **91**, 1155-1160

Hennecke, M. and Plapp, B. V. (1983) Biochemistry **22**, 3721-3728

Higuchi, R., Krummel, B. and Saiki, R. K. (1988) Nucleic Acids Res. **208**, 7351-7383

Hills, C. A. and Fewson, C. A. (1983) Biochem. J. **209**, 379-386

Hoey, M. E., Allison, N., Scott, A. J. and Fewson, C. A. (1987) Biochem. J. **248**, 871-876

Holbrook, J. J. and Ingram, V. A. (1973) Biochem. J. **131**, 729-738

Holbrook, J. J., Liljas, A., Steindel, S. J. and Rossman, M. G. (1975) In: The enzymes, vol. XI (Boyer, P. D., ed.), pp. 191-292, Academic Press, New York

Holtel, A., Marques, S., Mohler, I., Jakubzik, U. and Timmis, K. N. (1994) J. Bacteriol. **176**, 1773-1776

Horowitz, N. H. (1945) Proc. Natl. Acad. Sci. USA **31**, 153-157

Höög, J.-O., Eklund, H. and Jörnvall, H. (1992) Eur. J. Biochem. **205**, 519-526

Huc, C., Olomucki, A., Lê-Thi-Lan., Pho, D. B. and van Thoai, N. (1971) Eur. J. Biochem. **21**, 161-169

Hummel, W., Schutte, H. and Kula, M.-R. (1988) Appl. Microbiol. Biotechnol. **28**, 433-439

Hummel, W. and Kula, M.-R. (1989) Eur. J. Biochem. **184**, 1-13

Hurley, T. D., Edenberg, H. J. and Bosron, W. F. (1990a) J. Biol. Chem. **265**, 16366-16372

Hurley, T. D., Ehrig, T., Edenberg, H. J., and Bosron, W. F. (1990b) In: Enzymology and molecular biology of carbonyl metabolism 3 (Weiner, H., ed.), pp. 271-275, Plenum Press, New York

Hurley, T. D., Bosron, W. F., Stone, C. L. and Amzel, L. M. (1994) J. Mol. Biol. **239**, 415-429

Hutber, G. N. and Ribbons, D. W. (1983) J. Gen. Microbiol. **129**, 2413-2420

Inoue, J., Shaw, J. P., Rekik, M. and Harayama, S. (1995) J. Bacteriol. **177**, 1196-1201

Jaeger, E., Eggeling, L. and Sahm, H. (1981) Curr. Microbiol. **6**, 333-336

Jeloková, D., Karlsson, C., Estonius, M., Jörnvall and Höög, J.-O. (1994) Eur. J. Biochem. **225**, 1015-1019

Jensen, R. A. (1976) Ann. Rev. Microbiol. **30**, 409-425

Jones, T. A., Zou, J.-Y., Cowan, S. W., and Kjeldgaard, M. (1991) Acta. Crystallogr. **A47**, 110-119

Jörnvall, H. (1970) Eur. J. Biochem. **16**, 25-40

Jörnvall, H., Eklund, H. and Brändén, C.-I. (1978) J. Biol. Chem. **253**, 8414-8419

Jörnvall, H., Höög, J.-O., Bahr-Lindström, H. V. and Vallee, B. L. (1987a) Proc. Natl. Acad. Sci. USA **84**, 2580-2584

Jörnvall, H., Persson, B. and Jeffrey, J. (1987b) Eur. J. Biochem. **167**, 195-201

Julià, P., Farrés, J. and Parés, X. (1987) Eur. J. Biochem. **162**, 179-189

Julià, P., Parés, X. and Jörnvall, H. (1988) Eur. J. Biochem. **172**, 73-83

Kedishvili, N. Y., Bosron, W. F., Stone, C. L., Hurley, T. D., Peggs, C. F., Thomasson, H. R., Popov, K. M., Carr, L. G., Edenberg, H. J. and Li, T.-K. (1995) J. Biol. Chem. **270**, 3625-3630

Keinan, E., Hafeli, E. K., Seth, k. k. and Lamed, R. (1986) J. Am. Chem. Soc. **108**, 162-169

Kennedy, S. I. T. and Fewson, C. A. (1968) J. Gen. Microbiol. **53**, 259-273

Kenyon, G. L. and Hegeman, G. D. (1979) Adv. Enzymol. **50**, 325-360

Kiyohara, H., Nagao, K. and Yano, K. (1981) J. Gen. Appl. Microbiol. **27**, 443-455

Klinman, J. P. (1976) Biochemistry **15**, 2018-2026

Kowalchuck, G. A., Hartnett, G. B., Benson, A., Houghton, J.E., Ngai, K. and Ornston, L. N. (1994) Gene **146**, 23-30

Kvassman, J., Larsson, A. and Pettersson, G. (1981) Eur. J. Biochem. **114**, 555-563

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

Lamed, R. J. and Zeikus, J. G. (1981) J. Bacteriol. **195**, 183-190

Lauvergeat, V., Kennedy, K., Feuillet, C., McKie, J. H., Gorrichon, L., Baltas, M., Boudet, A. M., Grimapettenati, J. and Douglas, K. T. (1995) Biochemistry **34**, 12426-12434

Lawrence, J. G., Ochman, H. and Hartl, D. L. (1991) J. Gen. Microbiol. **137**, 1911-1921

Lemaire, M., Schmitter, J.-M., Issakidis, E., Miginiac-Maslow, M., Gadal, P. and Decottignies, P. (1994) J. Biol. Chem. **269**, 27291-27296

Lindqvist, Y. (1989) J. Mol. Biol. **209**, 151-166

Ling, L. L., Keohavong, P., Dias, C. and Thilly, W. G. (1991) PCR Methods Applic. **1**, 63-69

Lippman, M. and Schlesinger, R. B. (1979) Chemical contamination in the human environment., Oxford University Press, New York

Livingstone, A., Fewson, C. A., Kennedy, S. I. T. and Zatman, L. J. (1972) Biochem. J. **130**, 927-935

Livingstone, A. and Fewson, C. A. (1972) Biochem. J. **130**, 937-946

Lo, H. S. and Reeves, R. E. (1978) Biochem. J. **171**, 225-230

Long, A. and Ward, O. P. (1989) Biotechnol. Bioeng. **34**, 933-941

Long, A., James, P. and Ward, O. P. (1989) Biotechnol. Bioeng. **33**, 657-660

Lowry, O. H., Rosebrough, N. J., Farr, A. L. and Randall, R. J. (1951) J. Biol. Chem. **193**, 265-275

MacKintosh, R. W. (1987) PhD. Thesis, University of Glasgow.

MacKintosh, R. W. and Fewson, C. A. (1988a) Biochem. J. **250**, 743-751

MacKintosh, R. W. and Fewson, C. A. (1988b) Biochem. J. **255**, 653-661

Magonet, E., Hayen, P., Delforge, D., Delaive, E. and Remacle, J. (1992) Biochem. J. **287**, 361-365

Marques, S., Holtel, A., Timmis, K. N. and Ramos, J. L. (1994) J. Bacteriol. **176**, 2517-2524

McClure, N. C. and Venables, W. A. (1986) J. Gen. Microbiol. **133**, 2073-2077

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

Miles, E. W. (1977) Meth. Enzymol. **47**, 431-443

Morris, D. L. and McKinley-McKee, J. S. (1972) Eur. J. Biochem. **29**, 515-520

Murray, W. D. and Duff, S. J. B. (1990) Appl. Microbiol. Biotechnol. **33**, 202-205

Needleman, S. B. and Wunsch, C. D. (1970) J. Mol. Biol. **48**, 443-453

Nelson, K., Whittam, T. S. and Selander, R. K. (1991) Proc. Natl. Acad. Sci. USA **88**, 6667-6671

Nikolova, P. and Ward, O. P. (1994a) Biocatalysis **9**, 329-341

Nikolova, P. and Ward, O. P. (1994b) J. Ind. Microbiol. **13**, 172-176

Olsen, G. J., Woese, C. R. and Overbeek, R. (1994) J. Bacteriol **176**, 1-6

Ornston, L. N. and Neidle, E. L. (1991) In: The biology of *Acinetobacter* (Towner, K. J., Bergogne-Berezin, E. and Fewson, C. A. Eds.), pp 201-237, Plenum Press, New York

Ovádi, T., Libor, S., and Elödi, P. (1967) Acta Biochim. Biophys. Acad. Sci. Hung. **2**, 455-458

Palleroni, N. J. (1975) In: Genetics and Biochemistry of *Pseudomonas* (Clarke, P. H. and Richmond, M. H., Eds.), pp 1-36, John Wiley and Sons, London

Peretz, M. and Burstein, Y. (1989) Biochemistry **28**, 6549-6555

Pettersson, G. (1987) Crit. Rev. Biochem. **21**, 349-389

Pharmacia. (1992) Blue Sepharose CL-6B Technical Bulletin. Edition AA.

Pipke, R., Wagner-Dobler, I., Timmis, K. N. and Dwyer, D. F. (1992) Appl. Environ. Microbiol. **58**, 1259-1265

Plapp, B. V. (1970) J. Biol. Chem. **45**, 1727-1735

Plapp, B. V., Sogin, D. C., Dworschack, R. T., Bohlken, D. P., Woenckhaus, C. and Jeck, R. (1986) Biochemistry **25**, 5396-5402

Plapp, B. V., Ganzhorn, A. J., Gould, R. M., Green, D. W., Jacobi, T., Warth, E. and Kratzer, D. A. (1990) In: Enzymology and molecular biology of carbonyl metabolism 3. (Weiner, H., ed.), pp 241-251, Plenum Press, New York.

Pugh, S. Y. R., Roberts, S. M. and Fewson, C. A. (1987) Report for the LGC Biotransformation Club.

Ramaswamy, S., Eklund, H. and Plapp, B. V. (1994) Biochemistry **33**, 5230-5237

Rann, D. L. and Cain, R. B. (1973) Biochem. Soc. Trans. **1**, 658-661

Ransom, S. C., Gerlt, J. A., Powers, V. M. and Kenyon, G. L. (1988) Biochemistry **27**, 540-545

Rast, H. G., Engelhardt, G. and Wallnofer, P. R. (1980) FEMS Microbiol. Letts. **7**, 1-6

Reid, M. F. (1993) MSc Thesis, University of Glasgow.

Reid, M. F. and Fewson, C. A. (1994) Crit. Rev. Microbiol. **20**, 13-56

Reiner, A. M. (1971) J. Bacteriol. **108**, 89-94

Rojo, F., Pieper, D. H., Engesser, K.-H., Knackmuss, H.-J. and Timmis, K. N. (1987) Science **238**, 1395-1398

Rosenberg, S. L. (1971) J. Bacteriol. **108**, 1257-1269

Rossau, R., Van Landschoot, A., Gillis, M. and De Ley, J. (1991) Int. J. Syst. Bacteriol. **41**, 310-319

Rossmann, M. G., Liljas, A., Brändén, C.-I. Banaszak, L. J. (1975) In: The enzymes, vol. XI (Boyer, P. D., ed.), pp. 62-102, Academic Press, New York

Ryde, U. (1995) Protein Science **4**, 1124-1132

Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) Cold Spring Harbor laboratory Press, Cold Spring Harbor, N. Y.

Sanger, F., Nicklen, S. and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. USA **74**, 5463-5467

Sartorius, C., Gerber, M., Zeppezauer, M. and Dunn, M. F. (1987) Biochemistry **26**, 871-882

Schneider-Bernlohr, H., Adolph, H. and Zeppezauer, M. (1986) J. Am. Chem. Soc. **108**, 5573-5576

Schreiner, A., Fuchs, K., Lottspeich, F., Poth, H. and Lingens, F. (1991) J. Gen. Microbiol. **137**, 2041-2048

Shanley, M. S., Harrison, A., Parales, R. E., Kowalchuk, G., Mitchell, D. J. and Ornston, L. N. (1994) Gene **138**, 59-65

Shaw, J. P. and Harayama, S. (1990) Eur. J. Biochem. **191**, 705-714

Shaw, J. P., Schwager, F. and Harayama, S. (1992) Biochem. J. **283**, 789-794

Shaw, J. P., Rekik, M., Schwager, F. and Harayama, S. (1993) J. Biol. Chem. **268**, 10842-10850

Shearer, G. L., Kim, K., Lee, K. M, Wang, C. K. and Plapp, B. V. (1993) Biochemistry **32**, 11186-11194

Shields, M. S., Hooper, S. W. and Sayler, G. S. (1985) J. Bacteriol. **163**, 882-

Shin, H. S. and Rogers, P. L. (1995) Appl. Microbiol. Biotechnol. **44**, 7-14

Short, J. M., Fernandez, J. M., Sorge, J. A. and Huse, W. D. (1988) Nucleic Acids Res **16**, 7583-7600

Skarzynski, T., Moody, P. C. E. and Wonacott, A. J. (1987) J. Mol. Biol. **193**, 171-187

Smith, M. W., Feng, D.-F. and Doolittle, R. F. (1992) Trends Biochem. Sci. **17**, 489-493

Sogin, D. C. and Plapp, B. V. (1974) J. Biol. Chem. **250**, 205-210

Southern, E. M. (1975) J. Mol. Biol. **98**, 503-507

Stanier, R. Y. (1951) Ann. Rev. Microbiol. **5**, 35-56

Stanier, R. Y., Palleroni, N. J. and Doudoroff, M. (1966) J. Gen. Microbiol. **43**, 159-271

Stanier, R. Y. and Ornston, L. N. (1973) Adv. Microbiol. Physiol. **9**, 89-151

Stone, C. L., Li, T.-K. and Bosron, W. F. (1989) J. Biol. Chem. **264**, 11112-11116

Studier, F. W. and Moffat, B. A. (1986) J. Mol. Biol. **189**, 113-130

Suen, W.-C. and Spain, J. C. (1993) J. Bacteriol. **175**, 1831-1837

Suhara, K., Takemori, S. and Katagiri, M. (1969) Arch. Biochem. Biophys. **130**, 422-499

Sund, H. and Theorell, H. (1963) In: The Enzymes (2nd edition), vol. 7 (Boyer, P. D., Lardy, H., and Myrbäck, K., eds.), pp. 25-83, Academic Press,, New York

Sze, I. S.-Y. and Dagley, S. (1987) J. Bacteriol. **169**, 3833-3835

Taylor, P. L. (1991) Biosoft GeneJockey Sequence Processor

Towner, K. J., Bergogne-Bérézin, E. and Fewson, C.A. (1991) In: The biology of *Acinetobacter* (Towner, K. J., Bergogne-Berezin, E. and Fewson, C. A. Eds.), pp 1-24, Plenum Press, New York

Tsou, A. Y., Ransom, S. C., Gerlt, J. A., Buechter, D. D., Babbitt, P. C. and Kenyon, G. L. (1990) Biochemistry **29**, 9856-9862

Wagner, F. W., Burger, A. R. and Vallee, B. L. (1983) Biochemistry **22**, 1857-1863

Wales, M. R. (1992) PhD. Thesis, University of Glasgow.

Warhurst, A. M., Clarke, K. F., Hill, R. A., Holt, R. A. and Fewson, C. A. (1994) Appl. Environ. Microbiol. **55**, 1323-1328

Weinhold, E. G., Glasfeld, A., Ellington, A. D. and Benner, S. A. (1991) Proc. Natl. Acad. Sci. USA **88**, 8420-8424

Weinhold, E. G. and Benner, S. A. (1995) Prot. Engng. **8**, 457-461

White, P. J., Hunter, I. S. and Fewson, C. A. (1991) In: The biology of *Acinetobacter* (Towner, K. J., Bergogne-Berezin, E. and Fewson, C. A., eds.), pp. 251-257, Plenum Press, New York

Williams, P. A. and Murray, K. (1974) J. Bacteriol. **120**, 416-423

Williams, P. A. (1985) Elsevevier-BIOSOFT, Cambridge

Wisniewski, J., Winnicki, T. and Majewska, K. (1983) Biotechnol. Bioeng. **25,** 1441-1452

Worsey, M. J. and Williams, P. A. (1975) J. Bacteriol. **124,** 7-13

Wright, S. K., Zhao, F. J., Rardin, J., Milbrandt, J., Helton, M. and Furumo, N. C. (1995) Arch. Biochem. Biophys. **321,** 289-296

Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) Gene **33,** 103-119

Zambryski, P., Tempe, J. and Schell, J. (1989) Cell **56,** 193-201

Zoltobrocki, M., Kim, J. C. and Plapp, B. V. (1974) Biochemistry **13,** 899-903