Purification and Characterisation of Benzaldehyde Dehydrogenase I from *Acinetobacter calcoaceticus* and the TOL Plasmid Encoded Benzaldehyde Dehydrogenase and Benzyl Alcohol Dehydrogenase from *Pseudomonas putida*

Ronald M. Chalmers

Thesis Submitted for the Degree of Doctor of Philosophy

Department of Biochemistry
University of Glasgow
1990

Contents

List of Figures

List of Tables

Page

Chapter 6

Abbreviations

The abbreviations recommended by the Biochemical Journal in its Instructions to Authors [Biochem. J. (1989) 257, 1-21] have been used with the following additions:

BADH	benzyl alcohol dehydrogenase
BZDH I	benzaldehyde dehydrogenase I
BZDH II	benzaldehyde dehydrogenase II
srDNA	simple repetitive DNA
DTT	dithiothreitol
f.p.l.c.	Fast Protein Liquid Chromatography (Pharmacia system)
K'_m	apparent Michaelis constant
N.C.I.B.	National Collection of Industrial Bacteria, Aberdeen, Scotland
PAGE	polyacrylamide-gel electrophoresis
PTH	phenylthiohydantoin
TEMED	*N,N,N',N'*-tetramethylenediamine
TOL-BADH	TOL plasmid-encoded benzyl alcohol dehydrogenase
TOL-BZDH	TOL plasmid-encoded benzaldehyde dehydrogenase
V'_{max}	apparent maximum velocity

Acknowledgements

I would like to thank the following people:

my supervisor Professor Charles A. Fewson for his tolerance and his willingness to share his great depth of knowledge and experience;

the late Professor R. M. S. Smellie and Professor M. D. Houslay for making the facilities of this department available to me;

Professor J. R. Coggins, Drs M. Anderson, A. M. Campbell, W. Cushley, L. M. Fixter, W. H. Holms, J. R. Kusel, D. P. Leader, J. G. Lindsay, E. J. Milner-White and H. G. Nimmo for many and various types of help such as discussion, advice and borrowing reagents;

Dr I. D. Hamilton for his generous help with immunoblotting;

Mrs. M. Cowan, Mr. P. Ferry and Mr. J. Jardine for technical assistance and for maintaining the laboratory;

Mr. Alan J. Scott for technical assistance and especially for his developmental work on the DEAE-Sephacel and Blue Sepharose chromatography steps in the purification of the TOL plasmid-encoded enzymes (Chapter 4). I would also like to thank him for help with the native M_r value determinations of the enzymes (Chapter 5) and for carrying out the thermal inactivation experiments for the TOL plasmid-encoded benzaldehyde dehydrogenase (Chapter 5).

I would also like to thank:

Dr R. W. MacKintosh for antiserum raised against BZDH II (code No. 314; Chapter 6);

Dr J. N. Keen (Dept. Biochemistry, University of Leeds) and Mr. B. Dunbar (Dept. Biochemistry, University of Aberdeen) for amino terminal sequence analysis;

Dr D. G. Campbell (Dept. Biochemistry, University of Dundee) for amino acid composition analysis (TOL plasmid-encoded enzymes; Chapter 5) when this departments' own analyser was out of commission;

all of my friends, past and present, in Lab. C24 for their company and diversions.

I would like to acknowledge the Science & Engineering Research Council for a postgraduate studentship for 36 of my 51 months work on this project and Bioflux for funding me for an additional 4 months, without which the most important results of the project would not have been obtained.

Finally I would like to thank the Science & Engineering Research Council, Beechams Pharmaceuticals, The Biochemical Society and Sir Reo Stakis for funding my attendance at the Fourth International Workshop on the Enzymology and Molecular Biology of Carbonyl Metabolism held in Gifu, Japan in the summer of 1988.

Summary

1 *Acinetobacter calcoaceticus* can grow on either mandelate or benzyl alcohol as sole sources of carbon and energy. L(+)-Mandelate is metabolised to benzoate with phenylglyoxylate and benzaldehyde as intermediates, whereas benzyl alcohol is converted to benzoate with benzaldehyde as the only intermediate. Although the intermediates of the mandelate and benzyl alcohol pathways converge at the level of benzaldehyde, the enzymes of the pathways are quite separate because of the presence of two different benzaldehyde dehydrogenases. This thesis is concerned primarily with the purification and characterisation of benzaldehyde dehydrogenase I which is induced during growth on mandelate. This was intended to complete the characterisation of all of the enzymes of the mandelate and benzyl alcohol pathways in *A. calcoaceticus*. The project was subsequently enlarged to include the purification and characterisation of the benzaldehyde dehydrogenase and the benzyl alcohol dehydrogenase encoded by the TOL plasmid pWW53 in *Pseudomonas putida* MT53, because this would allow a more broadly-based comparison of five aldehyde and alcohol dehydrogenases, including a comparison of chromosomal and plasmid encoded enzymes.

2 A procedure was developed for the purification of benzaldehyde dehydrogenase I using chromatography on DEAE-Sephacel, Phenyl Sepharose and f.p.l.c. Superose 12. The procedure yielded almost 2 mg of benzaldehyde dehydrogenase I from 20 g (wet weight) of cells. The enzyme was homogeneous as judged by isoelectric focusing and by denaturing and non-denaturing polyacrylamide-gel electrophoresis. A procedure was also developed for the purification of the TOL plasmid-encoded benzaldehyde and benzyl alcohol dehydrogenases from a single batch of *P. putida*. This involved chromatography on DEAE-Sephacel, Matrix gel red A, Blue Sepharose which separated the two enzymes, Phenyl Sepharose and finally Matrix gel green A. The final preparations were homogeneous as judged by denaturing and non-denaturing

polyacrylamide-gel electrophoresis. The procedure yielded 7 mg of the benzaldehyde dehydrogenase and 4 mg of the benzyl alcohol dehydrogenase from 28 g (wet weight) of cells.

3 The subunit M_r values of benzaldehyde dehydrogenase I, TOL-benzaldehyde dehydrogenase, and TOL-benzyl alcohol dehydrogenase were determined to be 56 000, 56 300 and 43 000, respectively, using SDS-polyacrylamide-gel electrophoresis. The apparent native M_r values were determined by gel-filtration chromatography, but different values were obtained depending upon whether f.p.l.c. Superose or Sephacryl S 300 columns were used. The ranges of apparent native M_r values obtained were as follows: benzaldehyde dehydrogenase I, 141 000 - 237 000; TOL-benzaldehyde dehydrogenase, 122 000 - 175 000 and TOL-benzyl alcohol dehydrogenase, 82 000 - 114 000.

4 Comparisons of the subunit and apparent native M_r values of the enzymes failed to indicate unambiguously the subunit structures of the enzymes and the investigation was extended using chemical cross-linking by dimethylsuberimidate. Cross-linked samples of the three *Acinetobacter* enzymes (benzaldehyde dehydrogenase I, benzaldehyde dehydrogenase II & benzyl alcohol dehydrogenase) and both of the *Pseudomonas* enzymes (TOL-benzaldehyde dehydrogenase & TOL-benzyl alcohol dehydrogenase) were prepared and then analysed using SDS-polyacrylamide-gel electrophoresis. All five enzymes appeared to behave as tetramers, although the results for TOL-benzaldehyde dehydrogenase were somewhat ambiguous.

5 The amino terminal sequences of the three *Acinetobacter* enzymes and both of the *Pseudomonas* enzymes were determined. The amino terminal sequences of benzyl alcohol dehydrogenase and TOL-benzyl alcohol dehydrogenase were clearly homologous to each other and also to members of the superfamily of long-chain, zinc-dependent alcohol/polyol dehydrogenases. The amino terminal sequences of benzaldehyde

dehydrogenase II and TOL-benzaldehyde dehydrogenase appeared to be homologous to each other; however, although the amino terminal sequence of benzaldehyde dehydrogenase I had some similarity with those of the other two aldehyde dehydrogenases the extent of the similarity was such that it was unclear whether this should be attributed to chance or to an evolutionary relationship.

6 The amino acid composition of the three *Acinetobacter* enzymes and both of the *Pseudomonas* enzymes were determined and pairwise comparisons were carried out among the aldehyde dehydrogenases and also between the alcohol dehydrogenases. Two different methods of comparison were used to indicate whether or not the amino acid compositions of the enzymes were similar enough to be consistent with close evolutionary relationships among the enzymes. The results using one method of pairwise comparison were consistent with relationships among the three aldehyde dehydrogenases and also between both of the alcohol dehydrogenases, however the results using the other method indicated that none of the enzymes were related. The amino acid compositions, together with the subunit M_r values were used to calculate the numbers of amino acid residues per subunit for each of the enzymes. These were: benzaldehyde dehydrogenase I, 525; benzaldehyde dehydrogenase II, 525; TOL-benzaldehyde dehydrogenase, 538; benzyl alcohol dehydrogenase, 381 and TOL-benzyl alcohol dehydrogenase, 412.

7 The rate of inactivation of each of the aldehyde dehydrogenases was determined at 60 ^{0}C. TOL-benzaldehyde dehydrogenase was the most stable and benzaldehyde dehydrogenase II the least stable. Benzaldehyde dehydrogenase II was considerably stabilised by the presence of benzaldehyde and perhaps slightly by NAD^+. Neither of these reagents had a great effect on the stability of either of the other two aldehyde dehydrogenases.

8 The absorption spectrum of benzaldehyde dehydrogenase I showed no evidence of a bound cofactor or prosthetic group.

9 Assays were developed for benzaldehyde dehydrogenase I, TOL-benzaldehyde dehydrogenase and TOL-benzyl alcohol dehydrogenase that involved monitoring the reduction of NAD^+ at 340 nm. The pH optima of the enzymes were as follows: benzaldehyde dehydrogenase I, 9.5; TOL-benzaldehyde dehydrogenase, 9.3 and TOL-benzyl alcohol dehydrogenase, 9.4.

The rates of the reactions with $NADP^+$ as cofactor, relative to those with NAD^+, were as follows: benzaldehyde dehydrogenase I, 4.9%; benzaldehyde dehydrogenase II, 4.3%; TOL-benzaldehyde dehydrogenase, 27%; benzyl alcohol dehydrogenase, 7.8% and TOL-benzyl alcohol dehydrogenase, 1.9%.

Benzaldehyde dehydrogenase I was progressively inhibited by benzaldehyde concentrations above approximately 4 μM.

Benzaldehyde dehydrogenase I was activated by K^+ (164-fold), Rb^+ (135-fold) and NH_4^+ (129-fold) but not by Na^+.

10 Both benzaldehyde dehydrogenase I and benzaldehyde dehydrogenase II had esterase activity with 4-nitrophenol acetate as substrate. The rates at pH 8.5 were each approximately 2% of the dehydrogenase activities and the enzymes were both activated by low concentrations of NAD^+.

11 The activity of benzaldehyde dehydrogenase I with fixed concentrations of various substrates was determined. The enzyme was more active when benzaldehyde was substituted at the *meta* and *para* positions of the aromatic ring than when it was substituted at the *ortho* position. Neither of the aliphatic aldehydes tested were substrates.

12 The apparent maximum velocity and the apparent Michaelis constant values were determined for benzaldehyde dehydrogenase I, benzaldehyde dehydrogenase II and

TOL-benzaldehyde dehydrogenase with NAD^+, benzaldehyde and the three possible monomethyl substituted benzaldehydes. The kinetic constants were also determined for benzyl alcohol dehydrogenase and TOL-benzyl alcohol dehydrogenase with NAD^+, benzyl alcohol and the three possible monomethyl substituted benzyl alcohols. The kinetic coefficients were used to calculate the apparent specificity constants (apparent maximum velocity/apparent Michaelis constant) of the enzymes for the substrates. The relative values of the apparent specificity constants of benzaldehyde dehydrogenase I, benzaldehyde dehydrogenase II and TOL-benzaldehyde dehydrogenase with benzaldehyde and the three monomethyl substituted benzaldehydes were broadly similar to each other. Benzaldehyde was the substrate with the highest apparent specificity constant for all three enzymes. Benzaldehyde dehydrogenase I was exceptional because it had a relatively low apparent specificity constant for 4-methylbenzaldehyde and TOL-benzaldehyde dehydrogenase differed from the other two enzymes because the apparent specificity constant for 2-methylbenzaldehyde was particularly low. The relative values of the apparent specificity constants for benzyl alcohol dehydrogenase and TOL-benzyl alcohol dehydrogenase with benzyl alcohol and the three possible monomethyl substituted benzyl alcohols were different. The substrates that gave the highest apparent specificity constants with benzyl alcohol dehydrogenase and TOL-benzyl alcohol dehydrogenase were 4-methylbenzyl alcohol and 3-methylbenzyl alcohol respectively.

13 Benzaldehyde dehydrogenase I was sensitive to inhibition by thiol-blocking reagents in the following order: 4-chloromercuribenzoate = *N*-ethylmaleimide > iodoacetamide > iodoacetate. The presence of NAD^+ or benzaldehyde afforded the enzyme slight protection from inhibition by *N*-ethylmaleimide and iodoacetate.

14 Antisera were raised in rabbits against the three *Acinetobacter* enzymes and both of the *Pseudomonas* enzymes. The extents of the cross-reactions among the various

antisera and each of the five enzyme antigens were determined by immunoprecipitation assays using native antigens and by immunoblotting using SDS-denatured antigens. Cross-reactions were observed among the three aldehyde dehydrogenases and also between the two alcohol dehydrogenases.

Chapter 1

Introduction

1.1 Fate of compounds in the environment

1.1.1 Dissimilation of synthetic compounds in the environment

By 1980 Chemical Abstracts had described more than 5 000 000 chemicals and at present the total annual world production of synthetic organic chemicals is more than $3x10^8$ tonnes (Jain & Slayer, 1987). Some of these chemicals are released in huge quantities into the environment. Accidental industrial spillage as well as both legal and illegal emissions have caused severe local contamination of top soil and ground water. However, a more serious problem is the global accumulation of pollutants in the oceans and the atmosphere. Many of these pollutants are 'new to biology' (or 'xenobiotics') in the sense that they contain unnatural structures to which organisms have not been exposed in the course of evolution (Fewson, 1988a). Some xenobiotics are readily degraded by microorganisms using pre-existing nonspecific catabolic pathways, modified pathways and fortuitous reactions. For example, *Pseudomonas* sp. strain B13 can grow on the xenobiotic 3-chlorobenzoate as the sole carbon source (Jeenes *et al.*, 1982); 3-chlorobenzoate is converted to 3-chlorocatechol and the chloromuconate cycloisomerase reaction of the subsequent *ortho* cleavage pathway causes the spontaneous elimination of the chlorine atom (Ramos & Timmis, 1987). Other xenobiotics are less readily degraded, such as the notoriously persistent insecticide DDT (1,1,1-trichloro-2,3-bis(4-chlorophenol)ethane). No microorganism able to mineralise it has ever been found, although a consortium able to co-metabolise it has been reported (Ramos & Timmis, 1987).

The accumulation of pollutants in the environment represents a major problem because many of the chemicals are toxic to many or all types of organisms. If the manufacture of toxic and recalcitrant compounds does not stop they will continue to accumulate and the inevitable result will be an environmental catastrophe much more serious than the one that we have already suffered over the last 150 years. The most

effective long term strategy would be for the chemical industries to use only waste-free synthetic routes and to search for biodegradable alternatives to recalcitrant products. These objectives are unlikely to be achieved because they are expensive compared with the present cost of waste disposal and because safer alternatives may simply not exist. The immediate prohibition of the manufacture of toxic and recalcitrant substances is not a plausible option because of the damage that this would cause to our industrial base and to our modern life style. We must therefore look for ways of removing persistent and damaging chemicals from the environment. The most logical approach is to examine the possibility of extending the degradative capabilities of microorganisms that are already responsible for the mineralisation of the vast majority of synthetic organic compounds. If suitable pathways for the degradation of industrial pollutants were to be constructed in microorganisms then many toxic wastes need never enter the environment; rather they could be made safe using a bioreactor.

Ramos & Timmis (1987) defined three basic experimental strategies for extending the degradative abilities of microorganisms:

i) "chemostat selection, which provides long term selective pressure for the evolution of organisms able to use a novel substrate, and which often involves progressive replacement of a mineralisable substrate by a recalcitrant analogue;

ii) *in vivo* genetic transfers, in which genes for critical enzymes of one organism are recruited into a pathway of another organism through experiments involving natural genetic processes such as transduction, transformation and especially conjugation. This approach is sometimes facilitated by the fact that the genetic information for recently-evolved pathways is frequently found to be located on transmissible plasmids or transposons (genetically promiscuous elements that readily move from one replication unit to another within the same cell). Thus, once a critical enzyme has evolved in one organism, its gene can be easily transferred to others and become recruited into related or unrelated existing or evolving pathways. Furthermore, once a new pathway has emerged in one organism, it can be readily transferred to others;

iii) *in vitro* evolution, in which cloned and well characterised genes are selectively introduced in a host organism in order to create a new pathway".

There are a number of important factors that determine the ability of microorganisms to mineralise any particular compound, for example, the transport of the compound across the cell membrane and the substrate and inducer specificities of the available enzymes. Many simultaneous mutations may therefore be required to establish a pathway for the mineralisation of a xenobiotic. It may be impossible to establish a new pathway *in vivo* by a series of single mutations because of the absence of a selection procedure for the incomplete pathway or because of the accumulation of toxic products from the incomplete pathway. Furthermore, if the number of steps in the putative pathway is large then the number of sequential or simultaneous mutations required may be so great that there is little more than a formal possibility of the pathway being established other than by *in vitro* evolution.

Once the pathway for the mineralisation of one or more pollutants has been assembled in a single organism, under the appropriate regulatory systems, it would be nice to imagine that the organism could be released into the environment where it would multiply and without further expense take care of the world's pollution crisis. Unfortunately there are two major complications that have often been overlooked. Firstly, the engineered pathway may not be stable in the absence of sustained selection and secondly, there are (or will be shortly) strict regulations controlling the release of viable engineered organisms into the environment. Thus, before we can realistically hope to solve the global pollution problem using microbes we must address two general questions: how can we deliberately engineer stable or unstable pathways and how can we assess the risks associated with the release of stable genetically engineered organisms? These questions may in part be answered by studying the evolutionary histories of natural catabolic pathways and it is the evolutionary histories of two such pathways that are the subject of this thesis.

This thesis is concerned with the evolutionary histories of the benzaldehyde dehydrogenases and the benzyl alcohol dehydrogenases involved in the degradative

pathways for mandelate and benzyl alcohol in *Acinetobacter calcoaceticus* and for toluene in *Pseudomonas putida.* The enzymes of these pathways can tolerate a number of different substitutions on the aromatic rings of their substrates (Chapter 5; Fewson, 1988b; Jeenes *et al.*, 1982; Murray *et al.*, 1972). The pathways therefore have the potential to assist in the degradation of a large variety of aromatic compounds, either produced within the cell or from exogenous sources. Unlike the central amphibolic pathways, these peripheral catabolic pathways are not strictly necessary for the viability of the cells in which they reside. They are therefore likely to be subject to more rapid evolution and also to be of a more recent origin than the central pathways. Together, these features make the pathways good model systems for studying metabolic evolution.

1.1.2 Dissimilation of aromatic compounds in the environment

The most abundant source of aromatic compounds in the environment is decaying plant material such as lignin. Approximately 1.5×10^{10} tonnes of carbon are fixed into wood each year and lignin represents between 18 and 35% of dry wood (Dagley, 1975). It is synthesised from monomeric alcohols, such as coumaryl, coniferyl and sinapyl alcohols, by the formation of carbon-carbon and ether linkages which are resistant to breakage both by chemical reagents and by enzymes (Dagley, 1975). Apparently it is the white-rot fungi which carry out the first stages of lignin degradation, however bacteria are involved in degrading the smaller aromatic products such as coniferyl alcohol, 4-hydroxy-3-methoxybenzaldehyde and 4-hydroxybenzoic acid (Cain, 1980). It is thought that the bacteria and the white-rot fungi may have a symbiotic relationship in which the bacteria supply vitamins to the fungi which in turn provide aromatic compounds which the bacteria use as sources of carbon and energy (Crawford, 1981).

The commonest low-M_r aromatic compounds in soil appear to be C_6-C_3 compounds (such as *p*-coumarate and ferulate) and C_6-C_1 compounds (such as protocatechuate, vanillate, vanillin and 4-hydroxybenzaldehyde). These latter compounds are possibly

produced by the β-oxidation of the phenylpropanoid monomeric units of lignin after depolymerisation (Fewson, 1988b). There is no good evidence that the breakdown of lignin in the soil produces significant amounts of mandelate, benzyl alcohol or toluene, nor is there any evidence of significant concentrations of free mandelate in soil or ground water (Fewson, 1988b). However, the mandelate, benzyl alcohol and toluene pathways have intermediates in common with other areas of metabolism and this together with the relaxed substrate specificity of many of the enzymes may help to capture the flux of carbon derived from the degradation of lignin.

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

In the past acinetobacters have been difficult to classify because of their diverse phenotypes and the lack of a sufficient number of unique phenotypic characters. Strains now known to be acinetobacters have previously been grouped into as many as 15 genera and they have most commonly been mistaken for *Bacterium antiratum*, *Herellea vaginicola*, *Mima polymorpha*, *Moraxella glucidolytica*, *Moraxella lwoffi*, *Micrococcus cerificans*, *Diplococcus mucosa* and *Achromobacter* (Juni, 1978). The discovery of an *Acinetobacter* strain highly competent for genetic transformation made possible the development of a transformation assay for the identification of strains of *Acinetobacter* (Juni, 1978). The genetics and physiology of *Acinetobacter* have been reviewed by Juni (1978), who described the genus as aerobic, Gram-negative, catalase-positive, oxidase-negative, non-motile coccobacilli with approximately 38 - 47 mol% G + C.

The bacterial strain N.C.I.B. 8250 has been classified as a strain of *A. calcoaceticus* (Fewson, 1967a,b; Baumann *et al.*, 1968). The organism is unable to grow on carbohydrates, however it is able to grow on a wide variety of other compounds; growth was tested on almost 450 compounds of which over 100 served as carbon + energy sources and almost 50 as nitrogen sources (Fewson, 1967a).

The mandelate and benzyl alcohol pathways of *A. calcoaceticus* N.C.I.B. 8250 have been investigated by Fewson and colleagues using both genetic and biochemical approaches. L(+)-Mandelate is metabolised to benzoate with phenylglyoxylate and benzaldehyde as intermediates. Benzyl alcohol is also converted to benzoate with benzaldehyde as the only intermediate (Figure 1.1; Kennedy & Fewson, 1968a,b). Although the intermediates of the mandelate and benzyl alcohol pathways converge at the level of benzaldehyde the enzymes of the pathways are quite separate because of the presence of two different benzaldehyde dehydrogenases (Livingstone *et al.*, 1972). Benzaldehyde dehydrogenase I (BZDH I) is induced by phenylglyoxylate, co-ordinately with the other enzymes of the mandelate pathway (Livingstone & Fewson, 1972). Benzyl alcohol dehydrogenase (BADH) and benzaldehyde dehydrogenase II (BZDH II) are co-ordinately induced by benzyl alcohol or benzaldehyde (Livingstone *et al.*, 1972).

The properties of the enzymes of the mandelate pathway in strain N.C.I.B. 8250 and other organisms have been reviewed by Fewson (1988b). In strain N.C.I.B. 8250 the mandelate and benzyl alcohol pathway enzymes have broad substrate specificities with respect to aromatic ring substitutions (Figure 1.2; Kennedy & Fewson, 1968a). The substituted benzoates derived from these compounds are converted either to catechol or to protocatechuate which are subjected to *ortho* ring cleavage and metabolised to succinate and acetyl CoA by the β-ketoadipate pathway (Figure 1.3; Kennedy & Fewson, 1968a). The intermediates of the catechol and the protocatechuate branches of the pathway converge at the level of β-ketoadipate enol-lactone. However, pairs of isofunctional enzymes are synthesised for the two subsequent reactions and the only enzyme common to both branches is β-ketoadipyl CoA thiolase, which catalyses the final step of the pathway. There is no evidence in strain N.C.I.B. 8250 that the genes for the mandelate, benzyl alcohol and β-ketoadipate pathways are other than chromosomal. However, a large plasmid encoding the β-ketoadipate pathway has been reported in *A. calcoaceticus* strain RJE74 (Winstanley *et al.*, 1987).

Pathways for the degradation of mandelate are widespread in Nature; they have been found in Gram-negative and Gram-positive bacteria as well as filamentous fungi and

yeast (Figure 1.4; Fewson, 1988b). At first this may seem surprising because C_6-C_2 compounds such as mandelate are not widespread in the natural soil and water environments of these organisms (Section 1.1.2; Fewson, 1988b). However, mandelate is periodically released into these environments by a number of higher plants and animals (Fewson, 1988b). Furthermore, in many organisms the mandelate pathway intersects other areas of metabolism and a wide variety of substrates may be channeled into the central amphibolic pathways after they have been converted to intermediates of the mandelate pathway (Figure 1.5; Fewson, 1988b).

The β-ketoadipate pathway is also widespread in Nature. It has been found in a variety of organisms with DNA ranging from 32 to 72 mol% G + C, including Gram-negative and Gram-positive bacteria and fungi (Ornston & Parke, 1977).

1.3 The TOL plasmid of *Pseudomonas putida*

There is currently great interest in the metabolic capabilities of strains of *P. putida* because they are able to degrade a wide variety of man-made and organic compounds some of which are recalcitrant and/or toxic. Many of the degradative pathways of *P. putida* are plasmid-encoded and they represent an important pool of genetic information for the species. For example, the TOL plasmid allows growth on toluene, the SAL plasmid on salicylate, the NAH plasmid on naphthalate, the CAM plasmid on camphor and the OCT plasmid allows growth on *n*-alkanes (Chakrabarty, 1976). *P. putida* also harbours many other plasmids of diverse types, such as those for drug and mercury resistance (Chakrabarty, 1976).

The TOL plasmid pWW0 was first described by Williams & Murray (1974) in *P. arvilla* mt-2 (now renamed *P. putida* mt-2), and it is the most thoroughly characterised of all catabolic plasmids. It has two catabolic operons which encode the enzymes for the oxidative catabolism of toluene, xylenes and other related aromatic compounds (Ramos *et al.*, 1987). The 'upper' pathway operon,

xylCAB, specifies the degradation of toluene, *m*-xylene and *p*-xylene to benzoate, *m*-toluate and *p*-toluate respectively (Figure 1.6; Burlage *et al.*, 1989). The 'lower' pathway operon, *xylDLEGFJIH*, encodes the enzymes for the degradation of benzoate and toluates to acetaldehyde and pyruvate by a bifurcating pathway that involves the *meta* fission of the aromatic ring (Figure 1.6; Burlage *et al.*, 1989). Different substrates are metabolised by different branches of the pathway; this is determined by the enzyme affinities of the two branches and serves to expand the substrate range of the pathway (Burlage *et al.*, 1989). For example, *m*-toluate is oxidised by the *xylF* branch, while *p*-toluate and benzoate are degraded by the *xylGHI* branch.

The expression of the TOL enzymes is regulated by the positively acting *xylR* and *xylS* gene products (Ramos *et al.*, 1987). The upper pathway is induced by a combination of the XylR protein, NtrA (a sigma factor) and upper pathway substrates. The lower pathway is induced by a combination of the XylS protein and lower pathway substrates. The lower pathway can also be induced by upper pathway substrates; the details of the mechanism are unclear but there is a requirement for both the XylR and the XylS proteins (Ramos *et al.*, 1987).

Williams & Worsey (1976) have shown that species of *Pseudomonas* with TOL plasmids are easily isolated from soil samples by selective enrichment on *m*-toluate minimal medium. They isolated 13 bacteria carrying plasmids superficially similar to pWW0 from nine different soil samples; eight were classified as *P. putida*, one as a fluorescent *Pseudomonas* sp. and four as nonfluorescent *Pseudomonas* sp. *P. putida* mt-2 was able to transfer the archetypal TOL plasmid, pWW0, into 11 of the cured isolates and of these eight were able to retransmit this foreign plasmid to their own cured derivatives (Williams & Worsey, 1976).

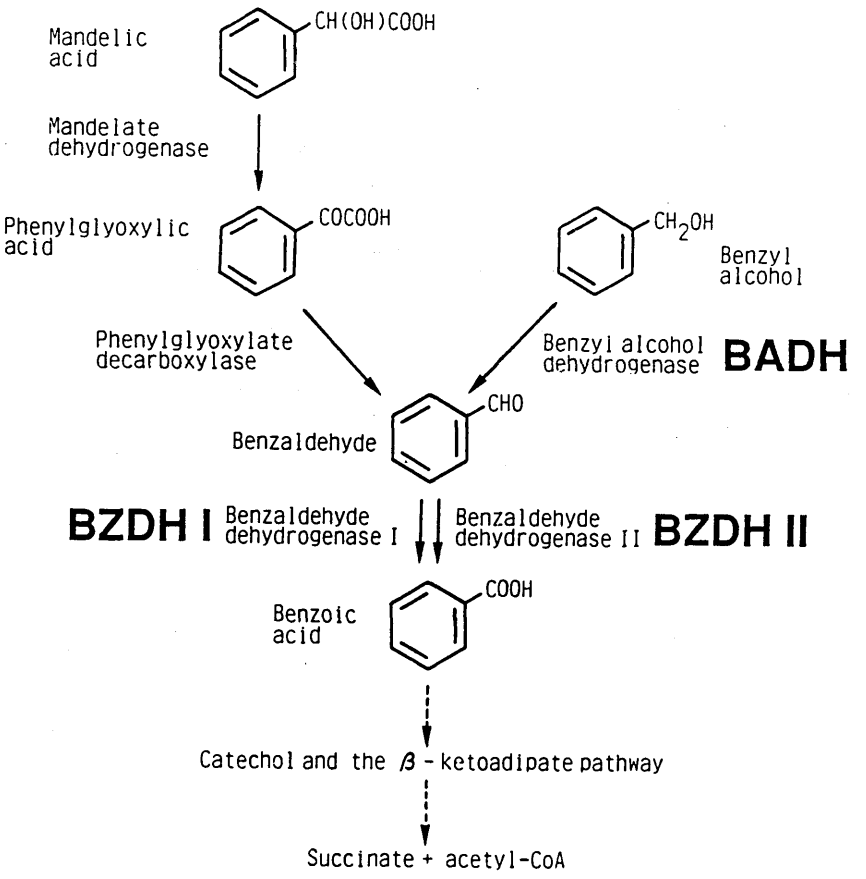
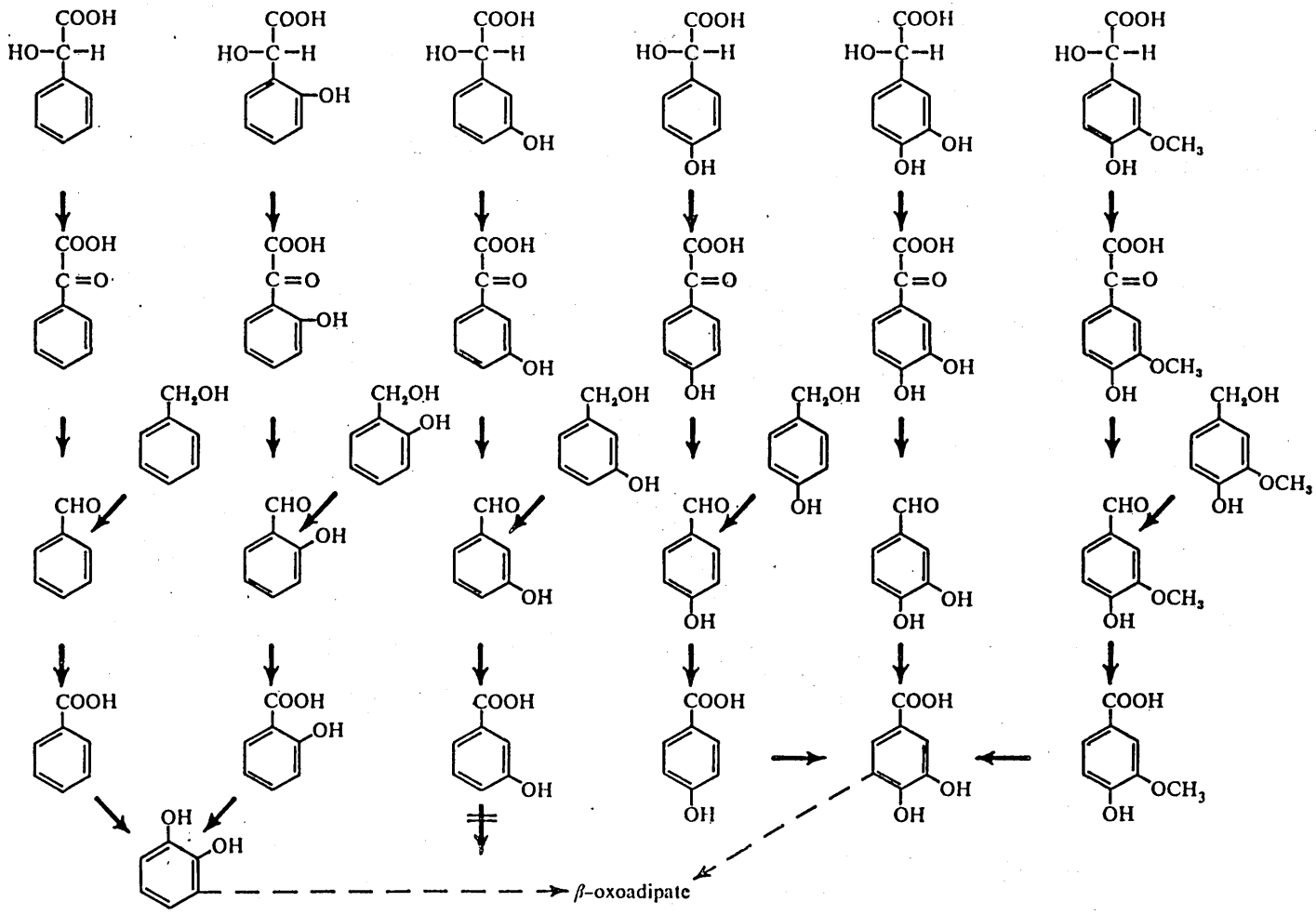


Figure 1.1 The mandelate and benzyl alcohol pathways of *Acinetobacter calcoaceticus* N.C.I.B. 8250

The induction patterns require the synthesis of isofunctional enzymes for the steps following the metabolic convergence of the pathways. L (+)-Mandelate dehydrogenase, phenylglyoxylate decarboxylase and BZDH I are co-ordinately induced by phenylglyoxylate. BADH and BZDH II are co-ordinately induced by benzyl alcohol or benzaldehyde.

Figure 1.2 The metabolism of mandelate and related compounds by *Acinetobacter calcoaceticus* N.C.I.B. 8250

The relaxed substrate specificities of the enzymes of the mandelate and benzyl alcohol pathways allow the conversion of 23 compounds to 6 benzoates with different ring substituents. From Kennedy & Fewson, 1968a.


β-oxoadipate

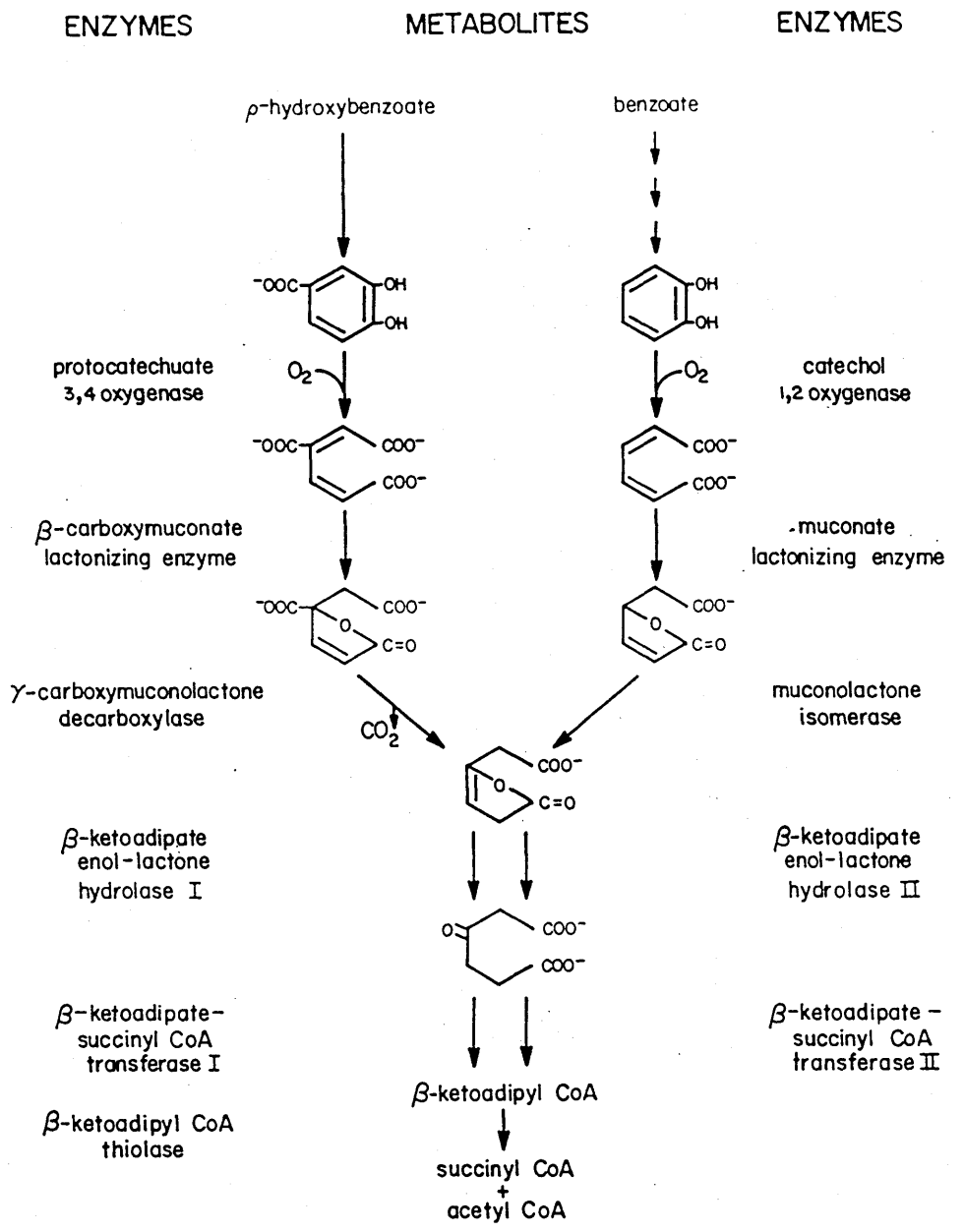


Figure 1.3 The β-ketoadipate pathway in *Acinetobacter calcoaceticus* and *Pseudomonas putida*

Analogous metabolic transformations convert protocatechuate and catechol to β-ketoadipate enol-lactone. In *Acinetobacter* the induction patterns require the synthesis of isofunctional enzymes for the steps following the metabolic convergence of the pathways. In *Pseudomonas* sequential enzyme induction does not require the synthesis of isofunctional enzymes. From Ornston & Yeh, 1982.

Figure 1.4 Metabolism of mandelate in a variety of organisms

Metabolism of mandelate in *Acinetobacter calcoaceticus* (⟶). *Pseudomonas putida* (⟶), *P. convexa* (- · ->), *Aspergillus niger*, *Neurospora crassa* and *Rhodotorula graminis* (---->). Enzymes: a, mandelate racemase; b', L(+)-mandelate dehydrogenase; b", D(-)-mandelate dehydrogenase; c, phenylglyoxylate decarboxylase; d, benzaldehyde dehydrogenase (there are several isofunctional enzymes); e, L-mandelate 4-hydroxylase; f, L-4-hydroxymandelate oxidase; g, benzoate 4-hydroxylase; h, 4-hydroxybenzoate 3-hydroxylase; i, benzoate 1,2-oxygenase; j, 3,5-cyclohexadiene-1,2-diol-1-carboxylate dehydrogenase; i+j = 'benzoate oxidase'. From Fewson, 1988b.

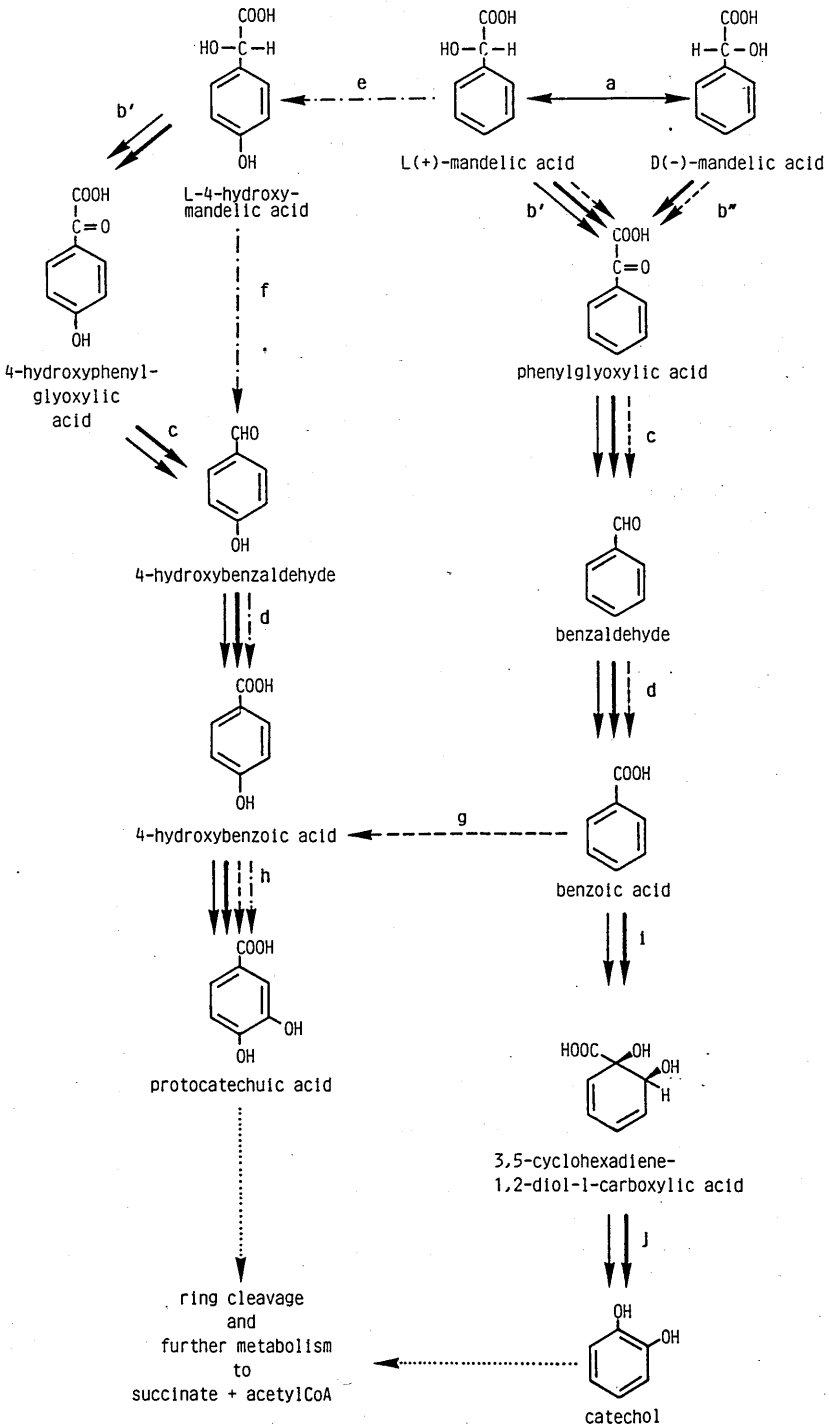
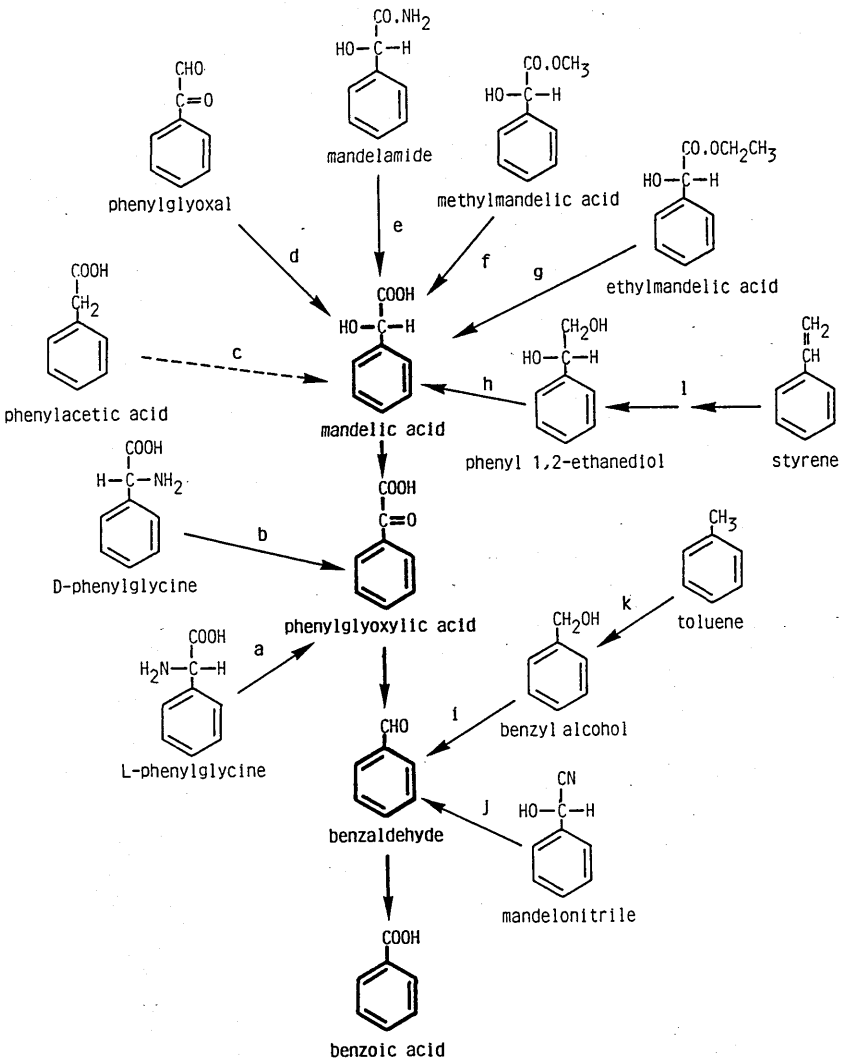

L(+)-mandelic acid
D(-)-mandelic acid
L-4-hydroxy-mandelic acid
4-hydroxyphenyl-glyoxylic acid
phenylglyoxylic acid
4-hydroxybenzaldehyde
benzaldehyde
4-hydroxybenzoic acid
benzoic acid
protocatechuic acid
3,5-cyclohexadiene-1,2-diol-1-carboxylic acid
catechol
ring cleavage and further metabolism to succinate + acetylCoA
a
b'
b''
c
d
e
f
g
h
i
j

Figure 1.5 The conversion of a variety of precursors to intermediates in the mandelate and benzyl alcohol pathways

(a) Example of the various conversions are: a and b, *Pseudomonas putida* and *Flavobacterium* sp.; c, *Penicillium chrysogenum*, mechanism uncertain; d, *P. putida*; i, *Acinetobacter calcoaceticus*, *P. putida*; k, *P. putida* containing TOL plasmid; l, occurs in mammalian tissues. From Fewson, 1988b.

(b) Tentative pathways for the formation of mandelate and 4-hydroxymandelate by the fungus *Aspergillus niger* (—>), the phytoplankton *Isochrysis galbana* and *Navicula incerta* (---->) and the red alga *Odonthalia floccosa* (---->). From Fewson, 1988b.

(a)


phenylglyoxal
mandelamide
methylmandelic acid
ethylmandelic acid
phenylacetic acid
mandelic acid
phenyl 1,2-ethanediol
styrene
D-phenylglycine
phenylglyoxylic acid
toluene
L-phenylglycine
benzyl alcohol
benzaldehyde
mandelonitrile
benzoic acid
a
b
c
d
e
f
g
h
i
j
k
l

Continued:

Figure 1.5 Continued:

(b)

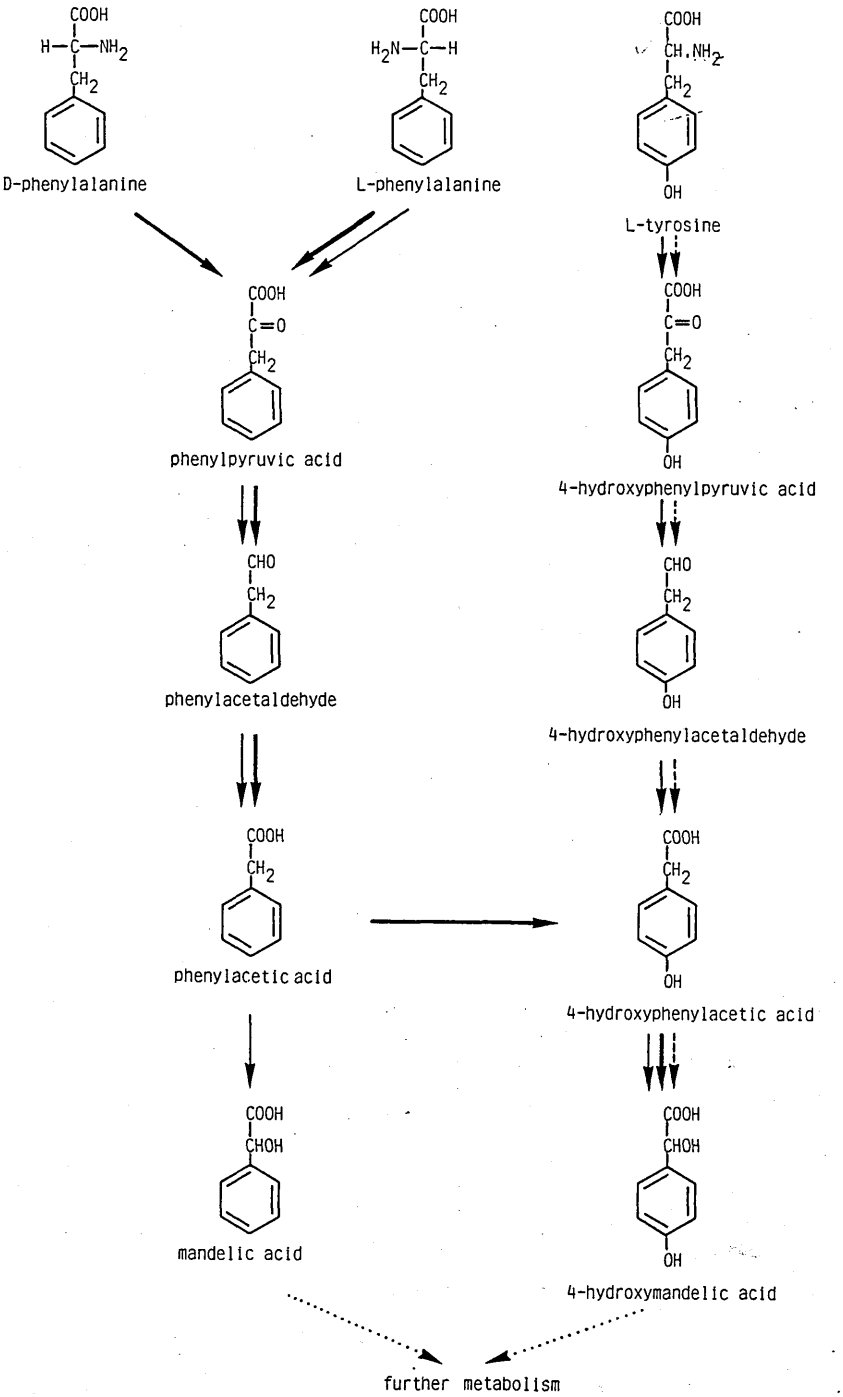
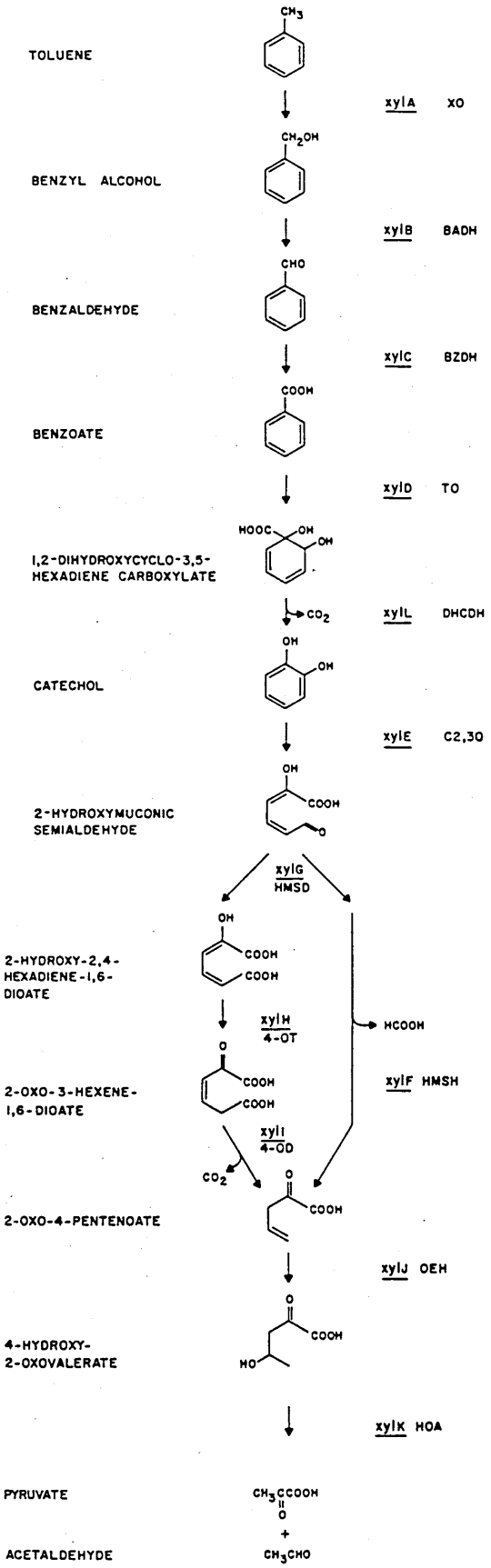


Figure 1.6 The degradative pathway for toluene encoded by the TOL plasmids pWW0 and pWW53

Chemical intermediates are listed on the left of the pathway while the genes and the abbreviations for the enzymes are on the right: XO, xylene oxidase; BADH, benzyl alcohol dehydrogenase; BZDH, benzaldehyde dehydrogenase; TO, toluate oxygenase; DHCDH, dihydroxycyclohexadiene carboxylate dehydrogenase; C2,3O, catechol 2,3-oxygenase; HMSD, 2-hydroxymuconic semialdehyde dehydrogenase; 4-OT, 4-oxalocrotonate isomerase; HMSH, 2-hydroxymuconic semialdehyde hydrolase; OEH, 2-oxopent-4-enoate hydratase; HOA, 2-oxyohydroxypent-4-enodate aldolase. Toluene, *m*-xylene and *p*-xylene are metabolised by the same suite of enzymes through a set of common intermediates. The methyl substituent of *m*-xylene is eliminated as acetic acid by the HMSH reaction. The methyl substituent of *p*-xylene is never eliminated and the products of the pathway are pyruvate and butyraldehyde instead of pyruvate and acetaldehyde which are produced when toluene is the substrate. From Burlage *et al.*, 1989.


TOLUENE
xylA XO
BENZYL ALCOHOL
xylB BADH
BENZALDEHYDE
xylC BZDH
BENZOATE
xylD TO
1,2-DIHYDROXYCYCLO-3,5-HEXADIENE CARBOXYLATE
xylL DHCDH
CATECHOL
xylE C2,3O
2-HYDROXYMUCONIC SEMIALDEHYDE
xylG HMSD
2-HYDROXY-2,4-HEXADIENE-1,6-DIOATE
xylH 4-OT
HCOOH
xylF HMSH
2-OXO-3-HEXENE-1,6-DIOATE
xylI 4-OD
2-OXO-4-PENTENOATE
xylJ OEH
4-HYDROXY-2-OXOVALERATE
xylK HOA
PYRUVATE
ACETALDEHYDE

Most studies of the molecular biology of the TOL plasmid have been carried out using pWW0 and very little work has been done on other TOL plasmids described in the literature, mainly because of their very large size (>200 kbp). The archetypal TOL plasmid pWW0 is relatively small, with only 117 kbp of which approximately 40 kbp are needed for the catabolic and the regulatory genes (Burlage *et al.*, 1989). Williams and colleagues have recently started a thorough characterisation of a second TOL plasmid carried by *P. putida* MT53 and designated pWW53 (Keil *et al.*, 1985). Like the archetypal TOL plasmid pWW0, pWW53 is relatively small; it is only about 110 kbp and the genes for the entire catabolic pathway have been cloned on a 35 kbp fragment. The two plasmids differ considerably in their restriction endonuclease digestion patterns and they belong to different incompatibility groups (Keil *et al.*, 1985). However, within the regions coding for the catabolic pathway there are strong similarities; there is considerable conservation of restriction sites and the order of the genes in both the upper and the lower pathway operons are the same on both plasmids, as are the relative positions of the regulatory genes (Keil *et al.*, 1987a,b). In both plasmids the direction of transcription is the same, however the distance between the two operons differs by about 11 kbp and on pWW0 the upstream operon is *xylCAB* whereas on pWW53 it is *xylDLEGF(J,K)IH* (Keil *et al.*, 1987a,b).

Very little is known about the TOL plasmid aside from the coding region. *P. putida* mt-2 constitutively expresses a pWW0-encoded pilus and the genes for replication and conjugal transfer have been mapped only roughly (Burlage *et al.*, 1989). Both pWW0 and pWW53 have an open reading frame at the start of the *xylCAB* operon that appears to code for a 14 000 Da polypeptide. No function has been assigned to this putative polypeptide although it is possible that *in vivo* it plays a role in the action of xylene oxygenase on toluenes (Keil *et al.*, 1987a,b). Finally, the role, if any, of plasmid genes in the transport of substrates remains to be worked out.

1.4 The evolution of metabolic pathways

The Enzyme Commission has defined just six basic types of enzymic reaction. However, there are almost 200 subgroups of enzymes which show great variation with respect to substrate specificity, kinetic constants, range of effectors and genetic regulation. There are more than 250 known alcohol and aldehyde dehydrogenases and these probably represent only a small fraction of the total. MacKintosh & Fewson (1987) have reviewed the great variety of alcohol and aldehyde dehydrogenases found in microbial systems. How has this enormous multiplicity of enzymes arisen? This thesis describes a preliminary attempt to investigate the evolutionary history of the benzaldehyde dehydrogenases and the benzyl alcohol dehydrogenases associated with the degradative pathways for mandelate and benzyl alcohol in *A. calcoaceticus* and toluene in *P. putida*. The remaining sections of this introduction will place this work in context by reviewing the hypotheses that have been proposed to account for the evolution of metabolic diversity.

1.4.1 The role of gene duplication in the evolution of metabolism

Gene duplication can potentially provide an organism with two functioning copies of a gene. Subsequent genome rearrangements may place one copy under different regulatory mechanisms that cause its expression or silencing under various metabolic circumstances. Gene duplication has at least three possible evolutionary advantages: (a) the copies may be regulated independently, one with each branch of a pathway, if the protein products act at a point after the convergence of the pathway, (b) one copy may be freed to accumulate mutations which at some time endow its protein product with a novel activity or allow a haploid organism heterozygous advantage, and (c) one copy may be recruited into an evolving pathway.

1.4.2 The hypothesis of retrograde evolution

The hypothesis of 'retrograde evolution' (Horowitz, 1945), which was further developed to include the concept of gene duplication (Lewis, 1951; Horowitz, 1965), is based on the idea that possession of part of a pathway is not in itself a selective advantage. It proposes that as some component of the primitive ecosystem was exhausted, an enzyme capable of producing it from a more simple substance was selected. In this way it can be envisaged that anabolic pathways evolved backwards from their complex end products down to meet what are now the central amphibolic pathways. Conversely, catabolic pathways could evolve from the central pathways to encompass progressively more complex substrates. Metabolic cycles could also have been created by the appearance of an enzyme that linked two branches of a pathway. These ideas can also be expressed as the concepts of vertical and horizontal evolution (Figure 1.7). Vertical evolution occurs when a duplicated gene accumulates mutations such that the product-binding site and catalytic activity of its enzyme product are changed so that it binds a novel substrate and converts it into the substrate of the ancestral enzyme (middle of Figure 1.7). Horizontal evolution is when the catalytic activity and/or substrate binding-site change so that a novel substrate is converted to the product of the ancestral enzyme (bottom of Figure 1.7). The concepts of vertical and horizontal evolution are intellectually satisfying because alone they are able to offer an explanation of the origin and diversity of metabolic pathways. Unfortunately their intellectual appeal does not vindicate their role in metabolic evolution and at present they are supported by only a little experimental evidence. The following examples are not exhaustive, but they probably illustrate the concepts of vertical and horizontal evolution better than most.

Glycolysis is one of the oldest and best preserved biochemical pathways and if retrograde evolution has been a potent force in the development of metabolic pathways then it is perhaps here that one would expect to find the evidence. In fact, however, there is little evidence of similarity between consecutive glycolytic enzymes (Fothergill-Gilmore, 1986); however, Lebioda & Stec (1988) have proposed that the two

consecutive glycolytic enzymes enolase and pyruvate kinase evolved by vertical evolution from a common ancestor. Both enzymes bind the same phosphoenolpyruvate ligand and require two Mg^{2+} cations per active site. The evidence for this putative relationship is based on the 3-dimensional structures of the two enzymes and it is discussed in Section 1.5a.

The *metB* and *metC* genes of *Escherichia coli* K-12 code for cystathionine-γ-synthase and β-cystathionase respectively (Figure 1.8). These enzymes specify two consecutive steps in the pathway for methionine biosynthesis and there is some evidence that they arose from a common ancestor by the process of vertical evolution. The genes have been cloned and sequenced and the amino acid sequences deduced (Duchange *et al.*, 1983; Belfaiza *et al.*, 1986). The MetB and MetC proteins have 386 and 395 amino acid residues respectively and when these are aligned with 20 short gaps, 126 amino acid residues appear to be conserved. Furthermore, 29 residues are acceptable replacements, which gives a total of 36% homology (Belfaiza *et al.*, 1986). The even distribution of the homologous regions throughout the proteins is consistent with the proteins sharing a common ancestral gene. If cystathionine-γ-synthase and β-cystathionase did indeed evolve from a common ancestral gene by a retrograde mechanism then by definition the ancestral gene must have coded for an ancestral β-cystathionase. The argument in favor of a retrograde mechanism is strengthened by the observation that like β-cystathionase, cystathionine-γ-synthase is also capable of catalysing a β-elimination reaction (Flavin, 1975). However, there is also the possibility that the ancestral enzyme was non-specific and that it catalysed some reaction that was functionally equivalent to the two reactions catalysed by the modern enzymes. This aspect of the argument is supported by the fact that in *Saccharomyces cerevisiae* the *MET25* gene product has substantial amino acid sequence homology with both cystathionine-γ synthase and β-cystathionase and does indeed specify a reaction that is functionally equivalent to those of the two enzymes (Figure 1.8; Parsot *et al.*, 1987).

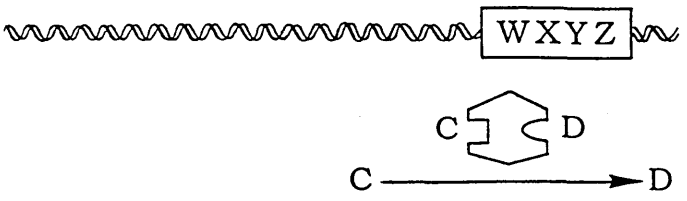
The amino acid sequence homologies between the carboxymuconolactone decarboxylases and the muconolactone isomerases from the β-ketoadipate pathways of *Acinetobacter* and *Pseudomonas* (Figure 1.3) are consistent with the two types of

enzyme having arisen from a common ancestral gene by the process of horizontal evolution (Ornston & Yeh, 1982). The enzymes specify chemically analogous interlactonic rearrangements; the major difference is that unlike muconolactone isomerase, carboxymuconolactone decarboxylase couples the migration of a double bond to a decarboxylation (Figure 1.9). Pairwise comparisons of the *N*-terminal amino acid sequences of the four enzymes (two from *Acinetobacter* and two from *Pseudomonas*) reveals identities of about 20%. This level of similarity could be due to chance or to convergent evolution, however when all four enzymes are compared simultaneously it is immediately obvious that they share a common ancestral gene. Furthermore, the possibility that the sequence identities are due to convergent evolution at the active sites is eliminated because of chemically dissimilar substitutions in regions where pairwise decarboxylase/isomerase comparisons show identity. For example, the aspartyl residue at position 8 of the *Pseudomonas* decarboxylase and *Acinetobacter* isomerase is substituted by lysine in position 8 of the *Acinetobacter* decarboxylase (Ornston & Yeh, 1982).

Whilst retrograde evolution no doubt played a part in metabolic evolution, it is naive to suppose that all of the enzymes of all metabolic pathways arose by vertical evolution following gene duplication. Many enzymes with the same type of activity have been observed to have similar features with respect to their structure and mechanism of catalysis; this supports the view that they share a common origin even though they now participate in quite separate metabolic pathways. For example, the stereospecificity of pyridine nucleotide hydrogen transfer is highly conserved; with few exceptions all of the enzymes carrying out a particular reaction transfer hydrogen from the same side of the dihydropyridine ring, in the case of aldehyde dehydrogenases it is from the A face (You *et al.*, 1978).

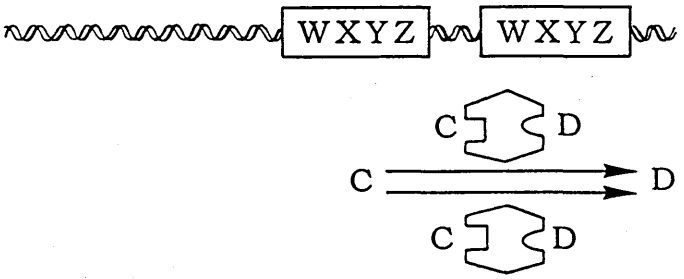
Figure 1.7 Vertical and horizontal evolution

Gene duplications followed by mutations may create new enzymic activities that cause the extension and branching of a pathway by vertical and horizontal evolution respectively. Block diagrams represent enzymes and their substrate and product binding sites. A, B, C, D: intermediates of a metabolic pathway; U, V, W, X, Y, Z: components of genes.



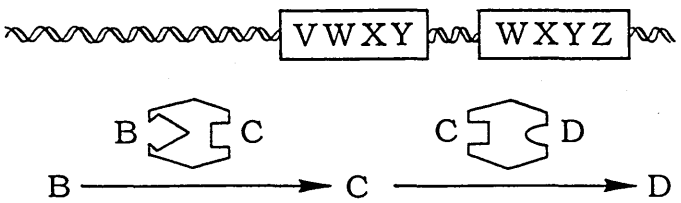
Gene

Pathway



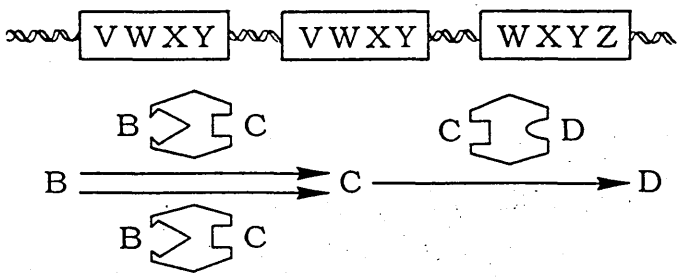
Gene duplication

Pathway with isofunctional enzymes



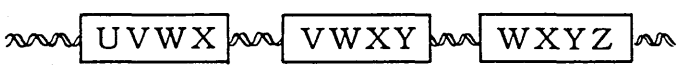
Mutation

Extended pathway

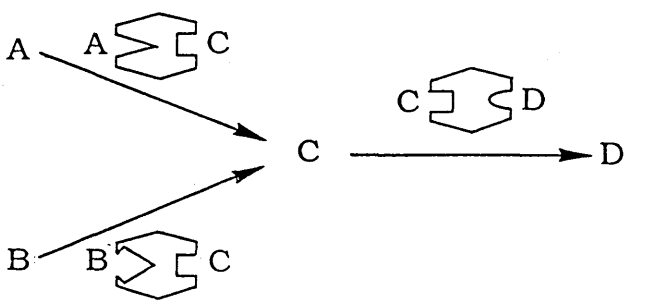


Second gene duplication

Pathway with a new pair of isofunctional enzymes



Second mutation



Branched pathway

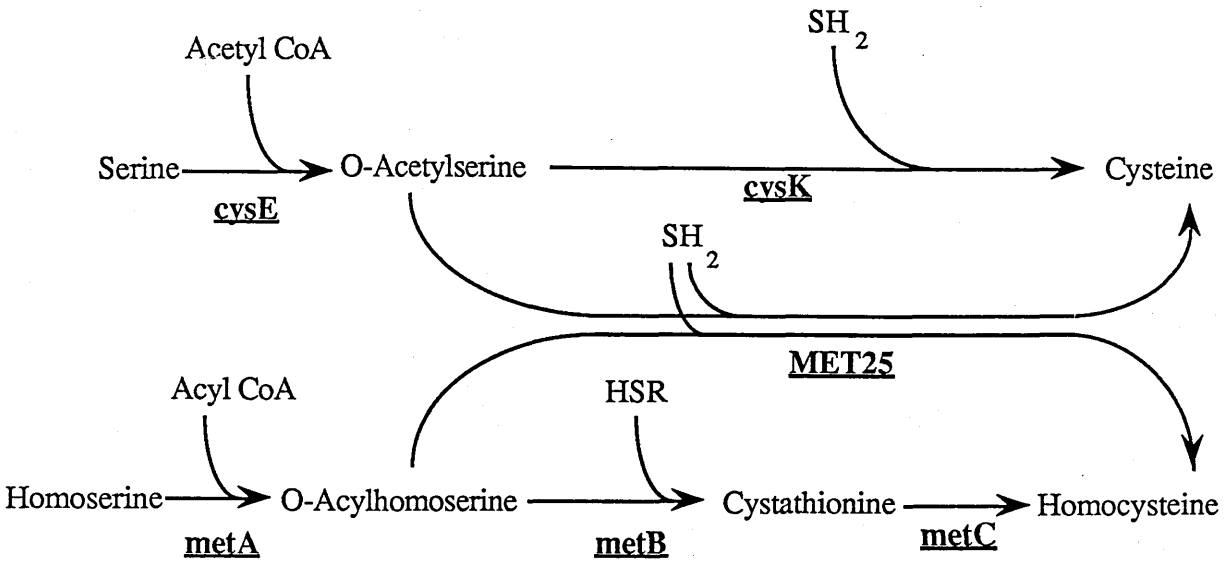


Figure 1.8 Some reactions involved in the cysteine and methionine biosynthetic pathways

The names of the *E. coli* genes encoding the relevant enzymes are indicated below the catalysed reactions; *cysE*, serine acetyltransferase; *cysK*, O-acetylserine sulfhydrylase; *metA*, homoserine transsuccinylase; *metB*, cystathionine-γ-synthase; *metC*, cystathionine-β-lyase. The reactions catalysed by the *S. cerevisiae MET25* gene product (O-acetylhomoserine, O-acetylserine sulfhydrylase) are shown. The radical R represents cysteine, and acyl can be either succinyl or acetyl. From Parsot *et al.*, 1987.

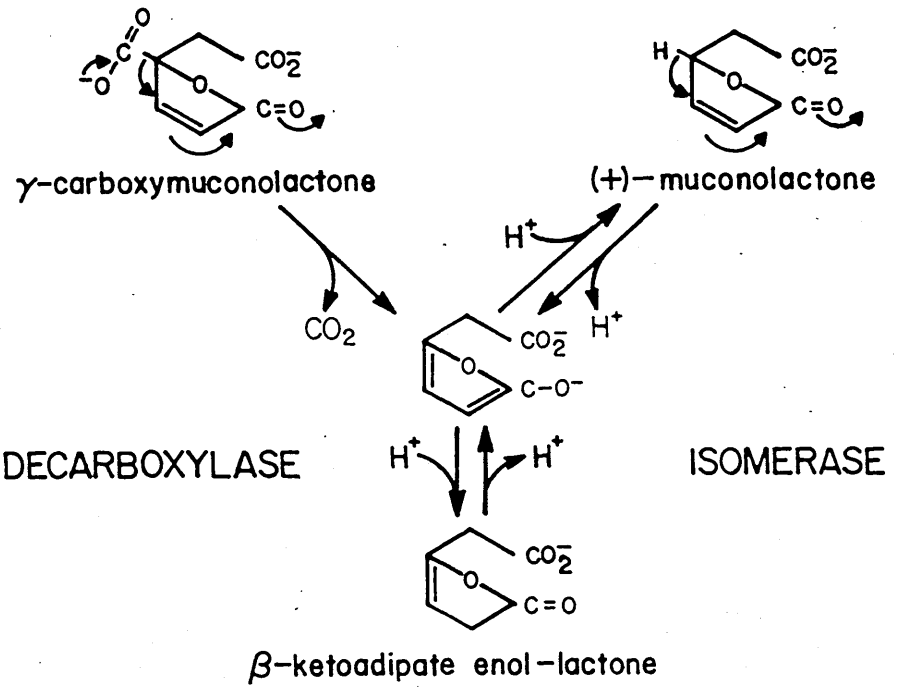


Figure 1.9 The reaction mechanisms of carboxymuconolactone decarboxylase and muconolactone isomerase

Carboxymuconolactone decarboxylase and muconolactone isomerase mediate chemically analogous interlactonic rearrangements. From Ornston & Yeh, 1982.

Furthermore, there are at least two major theoretical objections to the universitality of retrograde evolution. First, it is difficult to see how an enzyme could readily change to have a fundamentally different activity. Occasionally this problem may not be as great as it appears since superficially different enzymic reactions can be very similar at the electronic level, for example aldolases and hydrolyases (Jeffcoat & Dagley, 1973). Secondly, many metabolic intermediates are unstable or are produced only by the enzymes whose evolution is in question. In order for a pathway to extend past an unavailable intermediate, two novel enzymes must evolve from a single ancestor and become fixed simultaneously; the unavailable intermediate must be the product of one novel enzyme and the substrate of the other. This problem could sometimes be explained by postulating the existence of nonspecific enzymes such as the *MET25* gene product, *O*-acetylhomoserine, *O*-acetylserine sulfhydrylase (see above; Figure 1.8).

The hypothesis of retrograde evolution has enjoyed a popularity out of all proportion to the amount of data in support of it. The hypothesis has been cited in countless articles, despite the fact that there are very few good examples of consecutive enzymes that share a common ancestral gene. It is possible that retrograde evolution had an important role in establishing the metabolic diversity of primordial organisms. However, if this was the case then the evidence has been obscured by the influences of different selective pressures and genetic drift. Furthermore, modern organisms have available a large pool of enzymes which are likely to be recruited into an evolving pathway by far fewer mutational events than would be required for the evolution of a new enzyme by the retrograde mechanism.

1.4.3 Substrate ambiguity and evolution by gene recruitment

Metabolic pathways may evolve by recruiting duplicated copies of genes from other pathways. Subsequent point mutations or genetic rearrangements may fine-tune the activity and mode of regulation to suit a new metabolic function. For example, under suitable selection, *lac*$^-$ strains of *E. coli* give rise to revertants that express novel

β-galactosidases (Hall, 1984). The novel enzymes have different kinetic and physical properties and their gene is located far from the *lac* operon. They fall into several different classes, some of which represent multiple mutational events which have modified the hydrolytic activity of the enzymes towards lactose and/or the genetic regulation. Another, more striking, example of gene recruitment is that shown by mutants of *Salmonella typhimurium* defective in a subunit of isopropylmalate isomerase (LeuD). The *leuD* gene product is thought to be replaced by the *newD* gene product, but only in the presence of *supQ* mutations which remove the *newD* gene product's normal partner (Kemper, 1984).

In the examples given above, simple genetic lesions were repaired by the recruitment of cryptic genes or genes from other unidentified areas of metabolism. However, there is no reason why complete metabolic pathways should not evolve by a process of gene recruitment under the influence of suitable selection pressures. Indeed this may be the origin of all metabolic pathways in which the enzymes are not derived from a common ancestor by the process of retrograde evolution as described by Horowitz (Section 1.4.2).

Horowitz' hypothesis of retrograde evolution (Section 1.4.2) was based on the idea that part of a pathway is not in itself a selective advantage. This assumption is equally true for pathways created by gene recruitment, and it is therefore possible that such pathways have also evolved in a backwards fashion. The proposal that metabolic pathways evolved in a retrograde fashion by a process of gene duplication and recruitment avoids one of the major objections to Horowitz' hypothesis; namely, the prerequisite for the availability of each intermediate in the evolving pathway and the problem of unstable intermediates (Section 1.4.2). The substrate ambiguity of modern enzymes is to some extent masked by the specificity of regulatory mechanisms (Jensen, 1976). The basic requirement for an enzyme, or a number of enzymes, to be recruited into an evolving pathway is that they be expressed constitutively: this is often the condition of genes acquired by conjugal transfer, by the action of a phage or transposon and genes that have undergone duplication and transposition. A simple mutation, causing the loss of regulation, can therefore provide an enzyme, or a suite of enzymes, capable of

bridging the gap in an evolving pathway caused by an unavailable intermediate.

Jensen (1976) proposed that many pathways existed at a low level of expression due to fortuitous enzymic and non-enzymic reactions, and that novel metabolic pathways were established by the recruitment of suitable enzymes to amplify these reactions. The idea that metabolic pathways evolved in a retrograde fashion, both in the context used by Horowitz and also with respect to gene recruitment, will be destroyed if Jensen is correct. The evolution or recruitment of an enzyme for the last step in a pathway, existing at a low level of expression, will only have a small effect on the flux through the pathway because the enzyme will have an insignificant effect on the difference between the concentrations of its substrate and the preceding intermediate. In order for a novel enzyme to increase the flux through such a pathway it must significantly increase the difference between the concentrations of two consecutive intermediates. This is most likely if the enzyme specifies the first step of the pathway which draws flux from some large metabolic pool; the enzyme will cause an increase in the difference between the concentrations of its product and the next consecutive intermediate in the evolving pathway. Thus, a pathway existing at a low level of expression, due to fortuitous reactions, will only be effectively amplified if an enzyme is recruited for the first step of the pathway. In this scenario enzymes are likely to be recruited in the same order as the intermediates of the pathway and not in a retrograde fashion.

Jensen (1976) has pointed out that chorismate can be converted to phenylalanine by two non-enzymic reactions and a non-specific transamidase and that recruitment of two enzymes for the non-enzymic steps would have amplified this set of reactions. Stanier & Ornston (1973) have also pointed out that two of the six reactions in the catechol branch of the β-ketoadipate pathway (Figure 1.3) occur spontaneously and that three of the other reactions could be accomplished by non-specific enzymes from other areas of metabolism. Thus, the acquisition of one enzyme (muconate lactonising enzyme) could establish the whole pathway. It is interesting, and perhaps significant, that it is this very step that differentiates the β-ketoadipate pathway in prokaryotes and eukaryotes (Stanier & Ornston, 1973).

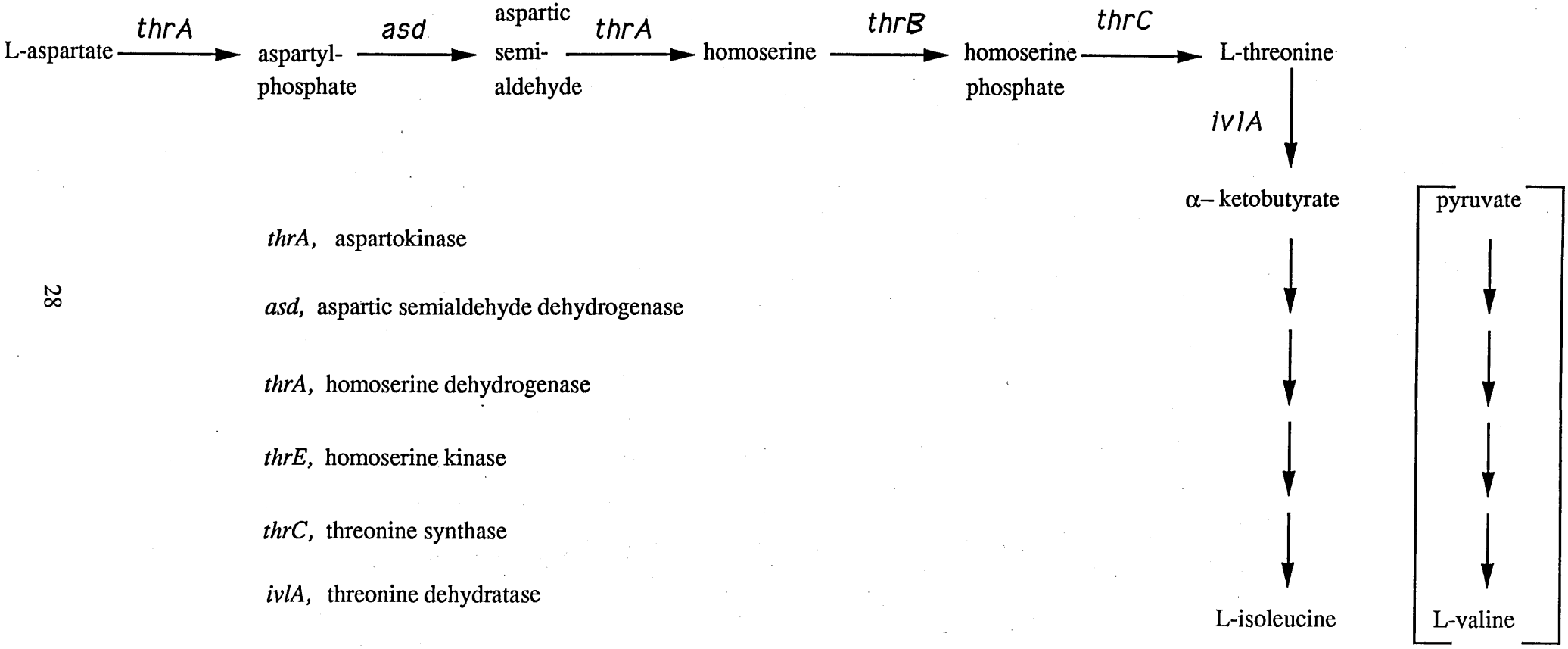
The pathways for the biosynthesis of amino acids are thought to be very ancient because with few exceptions the steps are conserved in all organisms in which they occur (Parsot *et al.*, 1987). Cohen and colleagues, working mainly with *E. coli* and *Bacillus subtilis*, have shown that enzyme recruitment has had an important role in establishing the biosynthetic pathways for threonine, isoleucine and tryptophan (Parsot *et al.*, 1987). Threonine is synthesised from homoserine by a two-step pathway using homoserine kinase and threonine synthase, which are encoded by the *thrB* and *thrC* genes respectively (Figure 1.10). Isoleucine is synthesised from threonine by a five step pathway (Figure 1.10) encoded by the genes *ilvA*, *ilvB*, *ilvC*, *ilvD* and *ilvE*. The only one of these genes unique to the isoleucine pathway is *ilvA*, which specifies the first step of the pathway, threonine dehydratase. All of the other genes are also involved in the biosynthesis of valine from pyruvate (Figure 1.10). Suppressor mutants, allowing the growth of isoleucine auxotrophs (*ilvA*$^-$ strains) on homoserine or threonine have been characterised and found to be constitutive for the *thrC* gene product, threonine synthase. Apparently threonine synthase also has a low level of threonine dehydratase activity. Furthermore, the amino acid sequences of threonine synthase and threonine dehydratase, both of which are pyridoxal phosphate dependent enzymes, are homologous. It therefore appears that the whole of the isoleucine biosynthetic pathway was established by recruitment; the first step was recruited from the threonine pathway and the remaining four enzymes were non-specific members of the valine biosynthetic pathway (Parsot *et al.*, 1987).

Three other pyridoxal phosphate dependent enzymes involved in amino acid metabolism have also been found to be homologous to threonine synthase and threonine dehydratase, these are the *tdc* gene product (which is the catabolic threonine dehydratase in *E. coli*), D-serine dehydratase, and the β subunit of tryptophan synthase (which catalyses the last step of the tryptophan biosynthetic pathway). In addition to their nominal activities threonine dehydratase and the β subunit of tryptophan synthase also have D-serine dehydratase activities (Parsot *et al.*, 1987).

Figure 1.10 The threonine and isoleucine biosynthetic pathways in *Escherichia coli*

The biosynthetic pathway leading to threonine and isoleucine are represented schematically. The four steps of the valine pathway which are catalysed by the same four enzymes involved in the isoleucine pathway are indicated in brackets.

Based on Parsot, 1986.



thrA, aspartokinase

asd, aspartic semialdehyde dehydrogenase

thrA, homoserine dehydrogenase

thrE, homoserine kinase

thrC, threonine synthase

ivlA, threonine dehydratase

The examples given above clearly demonstrate that substrate ambiguity and gene recruitment are important factors in the evolution of matabolic pathways. Recruitment is able to repair some simple genetic lesions in laboratory strains and the ancient relationships between the pyridoxal phosphate dependent enzymes of the amino acid pathways show that recruitment was also important in establishing the metabolic diversity of primative organisms. The proposal put forward by Jensen (1976), that substrate ambiguity provides sets of fortuitous reactions which can be amplified by recruitment, is exemplified by the isoleucine pathway which consists of four non-specific enzymes of the valine pathway and one enzyme recruited from the threonine pathway. The metabolic diversity of modern organisms and the substrate ambiguity of many of their enzymes makes gene recruitment a far more feasible route to new metabolic pathways, in terms of the number of mutations required, than retrograde evolution as described by Horowitz (Section 1.4.2). This is reflected in the ease with which examples of gene recruitment can be found and the lack of many good examples of vertical evolution.

1.4.4 Occurrence of gene duplications and deletions

In the preceding sections the evolution of metabolic pathways has been explained using the hypotheses of retrograde evolution and gene recruitment. It is likely that metabolic pathways have evolved by both of these mechanisms, although gene recruitment appears to be a more feasible mechanism in modern organisms which are able to draw on a large pool of efficient enzymes of diverse types. Whatever the balance between evolution by each of the mechanisms it is clear that they both rely on the availability of suitable duplicated genes.

Gene duplications are common in most, if not all, organisms and have occurred regularly during the evolutionary process. The best characterised set of duplicated genes is the globin family of mammals and specially man. The family consists of several isofunctional genes which are expressed at different stages of development, and a number

of pseudogenes (Lewin, 1987). Pseudogenes are areas of DNA, usually lying within the family, that have extensive homology with the active genes but are not expressed because of defective promoters, frame shift mutations, unstable mRNA or other reasons. Unlike active genes, in which mutations are fixed at a higher rate in the degenerate third codon position, pseudogenes have an equal number of fixed mutations in all three positions. By comparing the globin pseudogenes to active genes with respect to the number of mutations at each codon position, it is possible to construct a phylogenetic tree which shows that pseudogenes have arisen regularly by gene duplication and by the inactivation of functional genes (Lewin, 1987).

The most highly duplicated set of genes in all types of organisms are probably those for rRNA which are sometimes found in huge numbers; the salamander has 450, tandemly arranged (Ohno, 1970). Unlike the globin genes the copies are usually almost identical and this raises the question of how such homology can be maintained; in this case it may be that the copies are made each generation from a single master gene.

Families of genes seem to be less common in prokaryotes, although strains carrying duplicated genes are usually simple to isolate after a suitable selection procedure. For example, in some cases the resistance of *E. coli* to ampicillin is due to duplication events. The *ampC* gene (encoding β-lactamase) is amplified by the tandem duplication of a 9.8 kbp region located between 12 bp direct repeats (Sukhodolets, 1988).

When *Klebsiella aerogenes* is selected on xylitol in continuous culture, strains capable of utilising it quickly arise (Mortlock, 1984). Xylitol is not a natural sugar, but ribitol dehydrogenase has a little activity towards it and mutant strains expressing this enzyme constitutively readily take over the culture. If selection is continued other takeover events occur. These events have a variety of causes, such as improved transport of xylitol, superproduction of ribitol dehydrogenase and improvements in its activity (e.g. V_{max} and K_m). Superproduction of ribitol dehydrogenase is one of the most common events and in many cases it can be attributed to an increase in gene dosage by repeated duplication because when cultured on rich media such strains segregate into

groups with intermediate levels of ribitol dehydrogenase activity (Mortlock, 1984). Similarly, strains of *E. coli* selected for superproduction of β–galactosidase by growth in continuous culture under lactose limitation have been found to contain multiple duplications (Horiuchi *et al.*, 1963).

1.4.5 Mechanisms causing gene duplications and deletions

Given that gene duplications and deletions are relatively frequent events, at least in microorganisms, there are several important questions relating to their origins. Are they created by specific or gratuitous mechanisms? Are there specific mechanisms to eliminate duplications? Do these mechanisms benefit the fitness of an organism under normal circumstances?

Homologous sequences found at different locations in the genome are probably responsible for the initiation of most duplication and deletion events, although others may be due to the insertion of phage DNA or even to the duplication of the whole genome. A study of 250 deletions in the *lac* operon of *E. coli* demonstrated that the end points of deletions are not random (Albertini *et al.*, 1982). In this set of experiments two widely separated frameshift mutations were introduced into the *laci* gene so that only a deletion large enough to span both would restore the correct reading frame. After selection on lactose, revertants with unimpaired β–galactosidase activity were isolated in which a deletion had fused the start of *laci* onto the *lacz* gene. Analysis of the revertants demonstrated hot spots for deletion events which corresponded to repeated sequences, the most efficient of which were identical at 14 out of 17 nucleotide positions (Albertini *et al.*, 1982).

Repeated sequences may occasionally arise by chance; however, this becomes very improbable as the length of the repeat increases and a number of mechanisms have been proposed to explain the generation of lengthy repeats. Simple repetitive DNA (srDNA) is thought to be generated by slipped-strand mispairing due to the the action of repair mechanisms when nucleotide pairing slips out of true alignment (Levinson & Gutman,

1987). Once in existence, srDNA may expand by unequal crossing-over or be translocated by the action of a phage. A search of the EMBL data base using all the possible 15 bp probes consisting of mono-, di-, and trinucleotide motifs, produced 5-10 times more matches than when the probes consisted of random sequences (Tautz *et al.*, 1986). A high degree of cryptic simplicity was also found; that is sequences of biased nucleotide composition consisting of scrambled arrangements of repeated motifs (Tautz *et al.*, 1986). These simple sequences occur in coding and noncoding DNA but are far more common in eukaryotes than in prokaryotes.

In prokaryotes, gene duplications are very unstable particularly in *recA*$^+$ strains (Sukhodolets, 1988). Gene duplications are quickly lost during growth on nonselective media, presumably by unequal crossing over between sister chromosomes and by recombination and excision within the duplicated region. If a tandem duplication is deleted by the former mechanism one of the sister chromosomes will gain as many copies as are lost by the other. The elimination of duplications from the population shows that they are a selective disadvantage. This may be due to a general destabilisation of the genome or to a metabolic burden if the genes are expressed constitutively. Eukaryotic multigene families, such as the globin family, are quite stable even when their genes lie in close proximity to each other. This may be because of the complex nature of eukaryotic chromosomes which have sophisticated pairing mechanisms, are complexed with histones and contain introns which may help disguise the homology amongst closely related genes.

There is a variety of mechanisms that may lead to the stabilisation of duplications in prokaryotes. In *E. coli,* if duplicated *rrn* genes are flanked by direct repeats one copy may be excised by homologous recombination and this can reintegrate anywhere that the same repeated sequences are found. Such translocated duplications have an increased stability which can be used as a basis for their selection (Riley, 1984). Ornston & Yeh (1982) suggested an alternative mechanism for the stabilisation of duplicated genes, based on an investigation of the β-ketoadipate pathway (Figure 1.3) in *A. calcoaceticus*

and *P. putida*. The muconolactone isomerases and carboxymuconolactone isomerases of the two species are all homologous to each other. The three enol lactone hydrolases are also all homologous. The two enol lactone hydrolases of *A. calcoaceticus* contain sequences present in muconolactone isomerase; however, each enol lactone hydrolase incorporates a different set of sequences. Ornston & Yeh (1982) proposed that following the duplication that created the two enol lactone hydrolases, each gene incorporated different sequences from the muconolactone isomerase gene and that these masked their homology.

1.4.6 Conclusions concerning routes of metabolic evolution

We should no longer view the genome as a stable structure steadily accumulating point mutations that are neutral or that have been favoured by natural selection. Rather, it should be viewed as a dynamic structure regularly undergoing large rearrangements which occasionally survive long enough to be fine-tuned by smaller rearrangements and point mutations.

Gene duplications are important in the evolutionary process because they increase the scope for point mutations and subsequent natural selection. It may be more efficient to create new structures or enzymic activities by mutating a duplicated gene than by mutating a random sequence, because the gene copy already codes for a polypeptide capable of forming ordered three-dimensional structures. Repeated nucleotide sequences are important because they increase the frequency of duplications and translocations and perhaps played a part in the organisation of the prokaryotic genome into operons. They have also expanded greatly in higher eukaryotes to form a major component of the genome.

The twin goals of evolutionary molecular biology are to understand the types of changes that improve structure or function and to command a technology sufficiently advanced to implement similar changes, at least in microorganisms. These goals may be

best achieved by attempting to reconstruct the evolutionary events that produced modern organisms from their primitive ancestors. In the absence of a fossil record this can be attempted by detailed comparative studies at the level of nucleic acid sequence and three-dimensional protein structure. It is improbable that in the near future the type of evolutionary events described in this introduction will be related in any meaningful way to highly integrated multicellular organisms. However, attention is now focused on the simplest available evolving systems, single metabolic pathways.

1.5 Experimental methods for studying molecular evolutionary relationships

Until the development of nucleic acid sequencing, only indirect methods could be used to probe for homology between genes. Nucleic acid sequencing itself is not an infallible approach because it must be assumed that the differences between nucleic acid sequences are proportional to degree of relatedness, which is clearly only true if the sequences in question are evolving at the same rate and if parallel or convergent changes are rare. The success of indirect methods is dependent on how closely the property under study is related to the nucleic acid sequence. For example, protein sequences are a clear reflection of nucleic acid sequences whereas immunological methods are less accurate because they may depend on the three-dimensional structure of proteins and also on modifications such as the addition of carbohydrates or lipids (see also Section 6.4.2). A host of other indirect methods has been used to infer homology amongst genes: enzyme kinetics, co-purification, analysis such as proteolysis and subunit exchange, amino acid composition, stereochemistry of reactions or the chemistry of whole pathways, regulation of pathways and three-dimensional protein structure.

The comparison of amino acid compositions and amino acid sequences are discussed in Chapter 5 and they will therefore not be described in the remainder of this section. Immunological methods for comparing proteins are discussed in Chapter 6 and they also will be omitted from this section. With the exception of three-dimensional structural comparisons, which are beyond the scope of all but specialist laboratories, the remainder

of this section describes methods that were once widely used to indicate evolutionary homology but have disappeared from general use since the advent of rapid methods for protein and DNA sequencing.

(a) Three-dimensional structure

The only reported case of strong immunological cross-reactivity between proteins lacking sequence homology is for the intensely sweet proteins monellin and thaumatin (Ogata *et al.*, 1987). These proteins are a special case because they have a high specificity for the sweet receptor and therefore must share some structural features. Their common epitopes probably correspond to the three-dimensional structure of a sweet-receptor binding domain since they do not taste sweet when bound by antibodies nor when their structure is disrupted. This case highlights areas of growing awareness: that three-dimensional structures can converge in the absence of sequence similarities and that homology among three-dimensional structures can be conserved after considerable sequence divergence. For example, the serine proteases trypsin, chymotrypsin and elastase have almost identical conformation of the protein backbone yet 80% of their surface residues are different and there are even areas of difference in the interior of the molecules (Phillips *et al.*, 1970). Rossman & Argos (1977) addressed these problems in their attempt to establish a molecular taxonomy based on elements of protein structure. They asked three questions when comparing pairs of proteins: (a) can the proteins be superimposed so that the secondary structural elements are aligned, (b) are the elements in the same order on the polypeptide, and (c) how large are any insertions or deletions compared with the whole protein? This type of study could never replace nucleic acid or amino acid sequencing; however, it could be a very useful aid when investigating putative sequence homologies at the limit of statistical significance.

Enolase is the glycolytic enzyme that catalyses the dehydration of 2-phosphoglycerate to phosphoenolpyruvate. The larger of the two domains in the enolase subunit is a regular 8-fold barrel (Lebioda & Stec, 1988). This structure has also been found in 12

other functionally different enzymes including muconate lactonising enzyme (from the β-ketoadipate pathway of *P. putida*; Figure 1.3) and the other glycolytic enzymes pyruvate kinase and triose phosphate isomerase. Lebioda & Stec (1988) reasoned that if these proteins had diverged from a common ancestor then the natural superposition of elements of their three-dimensional structure (first strand on first strand, second strand on second strand *etc.*) should be better than the other seven superpositions (such as first strand on second strand and so forth). They made pairwise comparisons between enolase, pyruvate kinase, triose phosphate isomerase and muconate lactonising enzyme. The last enzyme was included as a control because its relationship with enolase is well characterised; it has 15% amino acid sequence homology with enolase and both enzymes share almost exactly the same folding patterns in both domains of their subunits. Lebioda & Stec (1988) found that the natural superpositions of enolase with pyruvate kinase and enolase with muconate lactonising enzyme were "obviously superior" and concluded that the two glycolytic enzymes share a common ancestor. They also concluded that when *P. putida* acquired the ability to prccess catechol, muconate lactonising enzyme evolved from enolase.

(b) Chemical mechanisms and regulation of pathways

At one time it was thought that the chemical mechanisms and genetic regulation of metabolic pathways were good indications of evolutionary homology, however as more data were collected it became apparent that this was not entirely true. A number of pathways have been worked out in a broad range of organisms, such as the β-ketoadipate pathway (Figure 1.3) that occurs in fungi, Gram-positive and Gram-negative bacteria (Ornston & Parke, 1977). The chemical steps and intermediates in this pathway are remarkably similar in different organisms with variations being the exception rather than the rule. However, this may be due to convergent evolution forced by a limited number of feasible routes to a given chemical conversion rather than to an evolutionary relationship. The manner in which two analogous pathways in different organisms are

genetically regulated may be an indication of whether the pathways are homologous. This notion is based on the idea that patterns of regulation are more free to evolve than the chemical steps of a pathway. Stanier and colleagues worked out the steps and patterns of regulation for the β-ketoadipate pathway in *Pseudomonas* and *Acinetobacter* (Figure 1.3; for review see Stanier & Ornston, 1973). The patterns of regulation in these two distantly related genera were quite different and on this basis they inferred that the pathways were analogous but not homologous (Canovas & Stanier, 1967). Subsequent work on the amino acid sequences of the enzymes disproved this idea by showing homologies amongst the enzymes from *Pseudomonas* and *Acinetobacter* (Ornston & Yeh, 1982).

(c) Enzyme kinetics, co-purification, proteolysis and subunit exchange

In the literature, enzyme kinetics, co-purification, proteolysis and subunit exchange experiments have been cited as methods that can be used to indicate evolutionary relationships. None of these methods is now widely used; either because they are too sensitive and often fail to show the relationship between highly homologous proteins or because there are too few unique character states to differentiate between homologous and analogous proteins.

Enzyme kinetics experiments are widely used to help assign the natural substrates of metabolic pathways and to determine enzyme reaction mechanisms. Unfortunately they are of little use in determining evolutionary relationships. Firstly, simple amino acid substitutions can have large effects on enzyme kinetics and closely related enzymes can therefore have very different values for their kinetic constants (Hartley, 1984). Secondly, it is often difficult to distinguish enzymes on the basis of the values of their kinetic constants because measurements are imprecise, vary amongst independent enzyme preparations and because of the relatively narrow range of values in which kinetic constants are generally found.

In some cases homologous proteins are co-purified in many types of chromatographic

systems; for example, the cystathionine-γ-synthase and β-cystathionase from *E. coli* (Belfaiza *et al.,* 1986). However, changes in the characters of surface amino acid residues can affect purification steps based on ionic and hydrophobic interactions. Furthermore, homologous proteins can have different multimeric forms which will affect purification steps based on native molecular size. For example, the muconolactone isomerases and carboxymuconolactone decarboxylases of *A. calcoaceticus* and *P. putida* are all homologous to each other (Section 1.4.2) and yet the isomerases are decamers or dodecamers of 11 000 Da subunits, while the decarboxylases are hexamers composed of 13 000 Da subunits (Ornston & Yeh, 1982).

Proteolytic enzymes can be used to generate peptide maps of purified enzymes. The technique is very sensitive because amino acid substitutions at or near the cleavage sites are enough to change the peptide map and only proteins with extensive homology can be compared.

The relationships amongst multimeric enzymes with the same activity can be studied using subunit exchange experiments, particularly if they are composed of non-identical subunits (Crawford, 1975; Sanderson, 1976). Experiments can be carried out with purified enzymes or simply by mixing cell-extracts from different strains that carry the appropriate genetic lesions. If two cell extracts are mixed, one from each of two strains carrying genetic lesions in the genes for each of two non-identical subunits, then enzyme activity is a convenient measure of subunit association. In order to interpret the results of such experiments it is necessary to establish whether the genetic lesion specifies the production of a defective subunit or whether it causes the complete loss of the subunit. If a defective subunit is produced then it may interfere with the association of perfect subunits when the cell extracts are mixed. An example of this type was given in Section 1.4.3: the defective LeuD subunit of isopropylmalate isomerase is only replaced by the NewD protein in the presence of *supQ* mutations which remove the normal partner of the NewD protein. Another factor complicating the interpretation of subunit exchange experiments is that the subunit interaction domains of homologous proteins are

likely to be among the most diverse regions of the molecules. This is because the accurate self association of protomers into oligomers is a selective force favouring the divergence of subunit interaction domains (Ornston & Yeh, 1982).

1.6 Aims and scope of this thesis

This thesis is concerned with the evolutionary histories of the benzaldehyde dehydrogenases and the benzyl alcohol dehydrogenases associated with the mandelate and benzyl alcohol pathways in *A. calcoaceticus* (Figure 1.1) and the toluene pathway in *P. putida* (Figure 1.6). The study of these dehydrogenases affords a good opportunity to test the hypothesis of retrograde evolution and also ideas about gene duplication. All five enzymes share two common substrates (NAD^+ and benzaldehyde) and participate in peripheral metabolic pathways that may be more likely to undergo rapid evolution than the central pathways, which are constrained because of the necessity of conserving features concerned with regulation. The relationships, if any, amongst these aldehyde and alcohol dehydrogenases are likely to be of a more recent origin than the relationships amongst the enzymes of the central pathways, and they may therefore be studied more easily. The fundamental questions addressed in this thesis are: (a) Were the isofunctional benzaldehyde dehydrogenases from the mandelate and benzyl alcohol pathways of *A. calcoaceticus* derived from a common ancestor? If so a subsequent question, beyond the scope of this thesis, would be: how were the genes stabilised against elimination by homologous recombination? (b) Are the aldehyde and alcohol dehydrogenases in *Acinetobacter* homologous or analogous to the ones in *Pseudomonas*? (c) Were the benzyl alcohol dehydrogenases derived from the benzaldehyde dehydrogenases by the mechanism of vertical evolution? The answers to these questions will further our understanding of the mechanisms by which metabolic diversity was established and in particular they may help us to define the origin of catabolic plasmids.

The original aim of the project was the purification and characterisation of BZDH I from *A. calcoaceticus* N.C.I.B. 8250 in order to compare it to BZDH II and BADH which had recently been purified and partially characterised (MacKintosh & Fewson, 1988a,b). This would complete the characterisation of all of the enzymes of the mandelate and benzyl alcohol pathways in *A. calcoaceticus* (Fewson, 1988b). The project was subsequently enlarged to include the purification and characterisation of the benzaldehyde dehydrogenase (TOL-BZDH) and the benzyl alcohol dehydrogenase (TOL-BADH), which are encoded by the TOL plasmid pWW53 in *P. putida* MT53, because this would allow a more broadly-based comparison of five aldehyde and alcohol dehydrogenases, including a comparison of chromosomal and plasmid encoded enzymes. It was envisaged that the physical, chemical, kinetic and immunological comparison of the five enzymes would establish whether or not there were any strong evolutionary relationships amongst the enzymes, although from the start it was realised that the exact nature of any relationships could only be established by sequencing the genes that encode the enzymes.

Chapter 2

Materials and Methods

2.1 Materials

Materials were generally the best grade commercially available and with the exception of those listed below were obtained from BDH Chemicals, Poole, Dorset, U.K.

2.1.1 Chemicals

2-Bromobenzaldehyde, 2-fluorobenzaldehyde, 3-fluorobenzaldehyde, 4-fluorobenzaldehyde, 4-methoxy-3-hydroxybenzaldehyde, 3-methoxy-4-hydroxybenzaldehyde, 2-methylbenzaldehyde, 3-methylbenzaldehyde, 4-methylbenzaldehyde and perillaldehyde were from Aldrich Chemical Co., Gillingham, Dorset, U.K.

$Na^{125}I$ in NaOH solution and D-[U-^{14}C] glucose were from Amersham International, Little Chalfont, Bucks., U.K.

Dithiothreitol, NAD^+ (free acid), $NADP^+$ (disodium salt), Tris (base), Bistris and rubidium chloride were from Boehringer Corp, Lewes, Sussex, U.K.

Nitrogen gas was from British Oxygen Co., Guildford, Surrey, U.K.

Urea was from BRL Ltd., Glasgow, U.K.

Glucose, sucrose and boric acid were from Formachem, Strathaven, U.K.

Acrylamide, *N,N'*-methylenebis-acrylamide, SDS and sodium chloride were from FSA Laboratory Supplies, Loughborough, U.K.

Pyronin Y was from G. T. Gurr Ltd., London, U.K.

Silver nitrate was from Johnson Mattey, Royston, Herts, U.K.

Cobalt sulphate was from Koch Light Laboratories, Haverhill, Sulfolk, U.K.

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

Acetic acid, formaldehyde, formic acid and nitric acid were from May and Baker, Dagenham, U.K.

Polybuffer PB74 was from Pharmacia LKB Biotechnology, Upsala, Sweden.

Hexan 1-al, octan 1-al and sodium bicarbonate were from Riedel-de Haen, Seelze, Hanover, F.D.R.

Bicine, Coomassie Brilliant Blue G250, Hepes, Mes, nitrophenol acetate, Tes, Tricine and Tween 20 were from Sigma Chemical Co., Pool, Dorset, U.K.

For the determination of kinetic parameters benzaldehyde was distilled and stored under nitrogen.

2.1.2 Chromatography Media

All chromatography media and prepacked columns were obtained from Pharmacia LKB Biotechnology, Upsala, Sweden except for Matrex gel red A, Matrex gel green A and Dye Matrex Screening Kit which were from Amicon, Danvers, MA, U.S.A.

2.1.3 Proteins and enzymes

All proteins were obtained from Sigma Chemical Co., Pool, Dorset, U.K. except for those listed below.

Bovine serum albumin (Fraction V) was from W. Smith Ltd., Edgware, Middlesex, U.K.

Combithek Calibration Kit for Gel Filtration Chromatography, alcohol dehydrogenase (horse liver) and glyceraldehyde-3-phosphate dehydrogenase (rabbit muscle) were from Boehringer Corp, Lewes, Sussex, U.K.

Gel Filtration Calibration Kit (high molecular weight) and Calibration Kit for SDS-PAGE were from Pharmacia LKB Biotechnology, Upsala, Sweden.

2.1.4 Miscellaneous materials

Centricon 30 microconcentrators and Centricon Centriprep concentrators were from Amicon, Danvers, MA, U.S.A.

Tryptone, yeast extract and Freund's Adjuvant (complete and incomplete) were from Difco Laboratories, Detroit, U.S.A.

Heat inactivated horse serum was from Gibco, Paisley, Strathclyde, U.K.

Nutrient broth and nutrient agar were from Oxoid, Basingstoke, Hampshire, U.K.

Nitrocellulose was from Schleicher & Schuell, Dassel, F.D.R.

Visking tubing was from Scientific Instruments Ltd., London, U.K.

Rabbit normal serum and donkey serum were from Scottish Antibody Production Unit, Carluke, Lanarkshire, U.K.

2.1.5 Bacteria

Acinetobacter calcoaceticus N.C.I.B. 8250 was originally obtained from the National Collection of Industrial Bacteria, Aberdeen, Scotland. *A. calcoaceticus* mutant strain C1005 (N.C.I.B 11339, previously HMM 5, constitutive for BZDH I as well as L-mandelate dehydrogenase and phenylglyoxylate decarboxylase) was derived from *A. calcoaceticus* N.C.I.B. 8250 (Fewson *et al.*, 1978).

Pseudomonas putida MT53, containing the TOL plasmid pWW53 was obtained from Professor P. A. Williams, Dept. of Biochemistry & Soil Science, University College of North Wales, Bangor, Gwynedd, Wales. The strain was isolated from soil by selective enrichment on *m*-toluate minimal medium (Keil *et al.*, 1985)

2.2 General methods

2.2.1 pH measurements

Routine measurements of pH were made with a Kent Electronic Instruments Ltd. (Chertsey, Surrey, U.K.), type 7010 meter. Fine measurements were made with a Radiometer (Copenhagen, Denmark) type M26 pH monitor fitted with a GK 2302 micro pH electrode.

2.2.2 Conductivity measurements

Conductivity measurements were made with a Radiometer (Copenhagen, Denmark) type CDM2e meter.

2.2.3 Glassware

Glassware was washed in a solution of approximately 1% (w/v) Haemo-sol (Meinecke & Co Inc., Baltimore, U.S.A.) and then rinsed thoroughly with glass-distilled water.

Glassware for amino acid compositions and amino acid sequencing was washed in a boiling solution of 10% (v/v) nitric acid for 30 min and then rinsed thoroughly with glass-distilled water.

2.2.4 Desalting by dialysis and gel filtration

Dialysis was carried out using Visking tubing which had been prepared by boiling three times in 1% (w/v) EDTA for 15 min and then three times in water also for 15 min. Visking tubing was stored in 20% (v/v) ethanol at 4 ^{0}C. Samples were dialysed

overnight against three changes of 100 volumes of buffer unless stated otherwise.

Samples were exchanged into different buffers by gel filtration through an f.p.l.c. desalting column (1.0x10 cm) at a flow rate of 1 ml/min.

2.2.5 Protein estimations

Protein concentrations were determined by the method of Lowry *et al.* (1951) using bovine serum albumin (Fraction V) as a standard. The protein concentrations of samples containing SDS were determined using the same method except that all of the solutions, apart from the Folin's reagent, contained 2% (w/v) SDS. The presence of protein in the column eluates was monitored using an LKB Uvicord monitor or a Pharmacia UV-1 monitor.

2.2.6 Lyophilisation

Samples were stored overnight in a deep freeze at approximately - 80 ^{0}C. They were quickly transferred to an LSL Secfroid Lyolab A lyophiliser and lyophilised.

2.2.7 Concentration of enzyme samples

Enzyme samples were concentrated using Centricon 30 microconcentrators and Centricon Centriprep concentrators according to the manufacturer's instructions.

2.2.8 Microfuge

Small samples were centrifuged in Eppendorf tubes at 15 000 g in an Eppendorf 3200 centrifuge.

2.2.9 Buffers and solutions

Buffers were prepared at room temperature by adjusting the pH values of approximately five-fourths strength solutions with the appropriate acid or base and then making to volume. All solutions were made using glass-distilled water.

2.3 Sterilisation

2.3.1 Moist heat

Media were sterilised by autoclaving at 109 ^{0}C for the appropriate time (C. A. Fewson, unpublished result) and the success of the sterilisation was verified using a Browne's tube (Albert Browne Ltd., Leister, U.K.)

2.3.2 Dry heat

Glass pipettes for inoculations were sealed in Kraft paper and heated in an oven at 160 ^{0}C for 1.75 h and the success of the sterilisation was again verified using a Browne's tube.

2.3.3 Filtration

Volatile or heat labile compounds, such as benzaldehyde or benzyl alcohol, were filtered through Millex-HV sterile filters (Millipore U.K. Watford, U.K.).

2.4 Growth, harvest, disruption and storage of bacteria

2.4.1 Maintenance of stock cultures

Stock cultures of *A. calcoaceticus* N.C.I.B. 8250 and C1005 were maintained at 4 ^{0}C in a complex medium as described by Allison *et al.* (1985). Stock cultures of *P. putida* MT53 were maintained at 4 ^{0}C on agar slopes containing a defined medium as described by Murray *et al.*(1972).

2.4.2 Growth of bacteria

(a) Small scale growth of *Acinetobacter* and *Pseudomonas*

Cultures of 50 ml and 500 ml volume were grown at 30 ^{0}C on a rotary shaker, in 250 ml and 2 l conical flasks, respectively. Cultures of up to 4 l were grown in round flasks with magnetic stirring and sterile air was supplied at a rate of 500 ml/min. Generally, *Acinetobacter* cultures were inoculated with 0.1% (v/v) of stock culture and *Pseudomonas* cultures with a loop of bacteria scraped from a slope culture.

(b) Large scale growth of *Acinetobacter*

Large quantities of strain C1005 were grown and harvested by a draw-and-fill procedure using a 10 l fermenter with a 20 l reservoir, both containing a complex medium plus 1.5 g DL-mandelate/l and 0.005% poly(propylene glycol) 2025 (Allison *et al.*, 1985; Hoey *et al.*, 1987). The wild type strain N.C.I.B. 8250 was grown in the same way except that 20 mM-benzyl alcohol replaced mandelate in the medium.

(c) Large scale growth of *P. putida* MT53

A small amount of bacteria scraped from a stock culture was used to inoculate 74 ml of Luria broth (10 g tryptone/l, 5 g yeast extract/l, 0.5 g NaCl/l) which was grown at 30 ^{0}C with shaking for 17 h. This culture was used to inoculate 3.7 l of Luria broth which was grown for 8 h as described in (a). A portion of the culture (3.3 l) was used to inoculate 6.7 l of three-halfs strength defined medium (Murray *et al.*, 1972). The three-halfs strength medium contained 1.5 g $(NH_4)_2SO_4$/l, 7.5 g KH_2PO_4/l, 0.15 g $MgSO_4.7H_2O$/l, 7.5 mg $FeSO_4.7H_2O$/l, 0.75 g nitrilotriacetic acid/l, 1.5 ml of stock salts solution (Bauchop & Elsden, 1960)/l, and the pH was adjusted to 7.0 with NaOH. The medium also contained 7.5 mM-3-methylbenzyl alcohol as carbon source. The culture was grown for 160 min in a 10 l fermenter under the conditions described by Hoey *et al.* (1987).

2.4.3 Harvest of bacteria

Cultures of *Acinetobacter* and *Pseudomonas* were harvested by centrifuging at 6 000 g for 20 min at 4 ^{0}C, washed by resuspension in ice-cold Z-1 buffer [2 g KH_2PO_4 + 1 g of $(NH_4)_2SO_4$ per litre, adjusted to pH 7.0 with NaOH] and then centrifuged at 12 000 g for 30 min at 4 ^{0}C (Hills & Fewson, 1983). Culture pellets were used immediately or stored at -20 ^{0}C.

2.4.4 Disruption of bacteria

Acinetobacter culture pellets were resuspended in ice-cold Z-1 buffer (Methods 2.4.3). *Pseudomonas* culture pellets were resuspended in ice-cold 100 mM-potassium phosphate buffer, pH 7.5.

(a) Ultrasonication

Samples of between 4.0 and 6.5 ml were placed in a 2 dram vial within a brass holder (Holms & Bennett, 1971) surrounded by an ice-water slurry. The cells were sonicated for four 30 s periods, alternating with three cooling periods of 30 s, using the 13 mm probe of the Dawe Soniprobe (type 1130A, Dawe Instruments Ltd., London, U.K.). The power delivered to the sample was approximately 80 W.

The homogenate was centrifuged at 15 000 g for 30 min at 4 ^{0}C and the supernatant was kept on ice or stored at -20 ^{0}C.

(b) French pressure cell disruption

Cell suspensions of up to 45 ml in volume were disrupted by passage through a French pressure cell at 0 - 10 ^{0}C (Allison *et al.*, 1985). The homogenate was centrifuged at 105 000 g for 150 min at 4 ^{0}C (Allison *et al.*, 1985) and the supernatant (extract) was saved.

2.5 Enzyme assays

All enzyme assays were carried out at 27 ^{0}C. One enzyme unit is defined as 1 µmol of substrate converted/min and the molar absorption coefficient of $NAD(P)^+$ was assumed to be 6.3×10^3 $M^{-1}.cm^{-1}$ (Boehringer, 1987).

2.5.1 Aldehyde dehydrogenase assays

BZDH I, BZDH II and TOL-BZDH activities were generally measured in 1 ml assay mixtures containing 0.15 M-glycine/KOH buffer, pH 9.5, 2 mM-$NAD(P)^+$, up to 50 µl of enzyme stock solution and the reaction was initiated by benzaldehyde (final concn. up

to 100 μM). The reduction of $NAD(P)^+$ was monitored at 340 nm and and the light path was 1 cm.

(a) High Range assay

Enzyme activities were routinely measured using an LKB Ultrospec II spectrophotometer with 100 μM-benzaldehyde. Occasionally a Pye Unicam SP8-100 spectrophotometer was also used. All three benzaldehyde dehydrogenases were inhibited by 100 μM-benzaldehyde and the rate of the reactions increased during the course of the assays as benzaldehyde was used up. Under these conditions the progress curve of the BZDH I reaction was not linear; however, the rate was proportional to the amount of enzyme added when the change in absorbance was $12x10^{-3}$ - $82x10^{-3}$/min.

(b) Low Range assay

When accurate measurements of enzyme activity were required assays were carried out using a Pye Unicam SP8-100 spectrophotometer and the concentration of benzaldehyde was lowered to between 3 and 12 μM. Within this range of benzaldehyde concentration none of the benzaldehyde dehydrogenases was significantly substrate inhibited and the progress curves of the reactions were linear until the substrate was almost completely used up. Under these conditions the rate of the BZDH I reaction was proportional to the amount of enzyme added when the change in absorbance was $1.5x10^{-3}$ - $9.5x10^{-3}$/min.

(c) Fluorimetric assays

In the steady-state kinetics experiments benzaldehyde dehydrogenase activities were measured in a Perkin-Elmer LS-5 Luminescence Spectrophotometer with an R 100A recorder, using the reaction mixture described in part (b) scaled up to 3 ml. The

reduction of NAD^+ was monitored at 340 nm (excitation)/460 nm (emission).

Reaction velocities were first calculated in arbitrary units; the chart speed was set at 30 mm/min with a full scale deflection of 5 mV and the velocity of the reaction was one arbitrary unit when the pen moved up 1 mm for every 5 cm of chart paper. The enzyme activity, expressed as arbitrary units per 100 ng of protein, was converted to specific activity (units/mg of protein) by dividing by 2.70. The conversion factor was calculated twice using two different methods and the figure quoted is an average of the two results (2.67 & 2.72). Firstly, the conversion factor was deduced by comparing the activities of an enzyme sample measured using the fluorometer and also using a Pye Unicam SP8-100 spectrophotometer. Secondly, the conversion factor was calculated from the change in fluorescence and the change in absorbance of an assay mixture after the complete reduction of NAD^+ in the presence of excess benzaldehyde and a large amount of enzyme.

2.5.2 Alcohol dehydrogenase assays

BADH and TOL-BADH activities were assayed in 1 ml reaction mixtures (MacKintosh & Fewson, 1988a) containing 0.1 M-Bicine/0.36 M-hydrazine (adjusted to pH 9.2 with NaOH), 2 mM-NAD^+, up to 50 µl of enzyme stock solution and the reaction was initiated by benzyl alcohol (final concentration 2 mM).

Reaction velocities were routinely measured using an LKB Ultrospec II spectrophotometer or a Pye Unicam SP8-100 spectrophotometer. Steady-state kinetics experiments were carried out using a Philips PU 8720 Scanning Spectrophotometer.

2.5.3 Standardisation of substrate concentrations

The concentrations of the substrate stock solutions were calculated from the ΔA_{340} after the complete enzymic oxidation or reduction of one substrate in the presence of

excess amount of the second substrate.

2.5.4 Analysis of reaction velocities using the Direct Linear method

Steady-state kinetics data and data from the ion-activation experiments were analysed by the Direct Linear method (Eisenthal & Cornish-Bowden, 1974) using the Enzpack computer program (Williams, 1985). Each of the sets of data used to calculate the apparent K_m values quoted in Chapter 5 had the form of a rectangular hyperbola.

2.5.5 Esterase assays

The esterase activities of BZDH I and BZDH II were measured in 1 ml assay mixtures containing 980 µl of 0.15 M-sodium pyrophosphate buffer (pH 8.5)/ 100 mM-KCl, 10 µl of 18 mM-4-nitrophenol acetate in acetone, various amounts of NAD^+ and the reaction was initiated by up to 10 µl of enzyme. The hydrolysis of 4-nitrophenol acetate was monitored at 400 nm and the molar absorption coefficient of 4-nitrophenol at pH 8.5 (14.69×10^3 $M^{-1}.cm^{-1}$) was calculated from the pK_a (7.15) and E_{IM}^{400} of the completely ionised species (1330/cm). For comparison, dehydrogenase activities were also measured in 990 µl of the same pyrophosphate buffer containing 100 mM-KCl, 2 mM-NAD^+, 100 µM-benzaldehyde and 10 µl of enzyme stock solution.

2.6 Purification procedures

2.6.1 Purification of BZDH I from *Acinetobacter calcoaceticus*

All steps were carried out at 4 ^{0}C except for step (d) which was at room temperature.

(a) Preparation of cell-free extract

Cells (20 g wet weight) of *A. calcoaceticus* C1005, grown on a rich medium (Methods 2.4.2b) and harvested as described in Methods 2.4.3, were resuspended in 60 ml of ice-cold Z-1 buffer (Methods 2.4.3) and disrupted by four passages through the French pressure cell (Methods 2.4.4b). The homogenate was centrifuged at 105 000 g for 150 min at 4 ^{0}C (Methods 2.4.4b). The resulting supernatant (extract) was dialysed (Methods 2.2.4) overnight at 4 ^{0}C against 2 l of 20 mM-potassium phosphate buffer, pH 7.5/2 mM-dithiothreitol (DTT) and used immediately or stored frozen at -20 ^{0}C. In either case the extract was centrifuged at 29 000 g for 20 min at 4 ^{0}C immediately before the first purification step.

(b) Ion-exchange chromatography on DEAE-Sephacel

Dialysed extract (50 ml) was applied at 20 ml/h to a DEAE-Sephacel column (2.6x9.0 cm) which had been pre-equilibrated with buffer A (25 mM-potassium phosphate, pH 7.5/2 mM-DTT). After washing with 10 column volumes of buffer A at 80 ml/h, BZDH I was eluted with a linear 25-200 mM-potassium phosphate buffer gradient (pH 7.5) containing 2 mM-DTT. The gradient volume was 500 ml and the flow rate was 20 ml/h. BZDH I was eluted after approximately 160 ml and fractions with peak activity were pooled.

(c) Hydrophobic-interaction chromatography on Phenyl Sepharose CL-4B/ DEAE-Sephacel

The pooled fractions (45 ml) from step (b) were applied to a Phenyl Sepharose column (2.6x9.8 cm) pre-equilibrated with buffer A. The column was washed with 50 ml of buffer A followed by 50 ml of buffer B (5 mM-potassium phosphate, pH 7.5/ 2 mM-DTT) and finally with 500 ml buffer B containing 30% (v/v) ethanediol. The flow rate was reduced from 50 ml/h to 20 ml/h and BZDH I was eluted from the column with 300 ml of buffer B containing 40% (v/v) ethanediol. The column eluate was passed through a DEAE-Sephacel column (1.6x3.0 cm) pre-equilibrated with buffer B. The DEAE-Sephacel column was washed at 20 ml/h with 20 ml buffer B and then with 60 ml of 50 mM-potassium phosphate buffer, pH 7.5/2 mM-DTT before BZDH I was finally eluted with buffer C (100 mM-potassium phosphate, pH 7.5/2 mM-DTT) at 6 ml/h. Fractions with peak activity were pooled.

(d) Gel-filtration chromatography through f.p.l.c. Superose 12

The pooled fractions (5.6 ml) from step (c) were concentrated to 180 µl (Methods 2.2.7) and applied to an f.p.l.c. Superose 12 column (1.0x30 cm) pre-equilibrated with buffer C. The flow rate was 0.3 ml/min throughout. Fractions containing BZDH I activity were checked for homogeneity by SDS-PAGE (Methods 2.7.1) before they were pooled.

(e) Storage of the purified enzyme

The pooled fractions (1.0 ml) from step (d) were mixed with an equal volume of ethanediol and stored at -20 ^{0}C.

2.6.2 Purification of TOL-BZDH and TOL-BADH

All steps were carried out at 4 ^{0}C and in addition some of the buffers used in step (d) were maintained at 0 ^{0}C in a bath containing an ice-water slurry.

(a) Preparation of cell free extract

Cells (28 g wet weight) of *P. putida* MT53, grown and harvested as described in Methods 2.4.2c and 2.4.3, were resuspended in 70 ml of phosphate buffer (Methods 2.4.4) and disrupted by 4 passages through the French pressure cell (Methods 2.4.4). The homogenate was centrifuged at 105 000 g for 150 min at 4 ^{0}C (Methods 2.4.4). The resulting supernatant (extract) was dialysed (Methods 2.2.4) overnight at 4 ^{0}C against 2 l of buffer D (50 mM-potassium phosphate, pH 7.5/2 mM-DTT) and used immediately.

(b) Ion-exchange chromatography on DEAE-Sephacel

Dialysed extract (85 ml) was applied at 40 ml/h to a DEAE-Sephacel column (5.0x4.0 cm) which had been pre-equilibrated with buffer D. After washing with 750 ml of buffer D at 80 ml/h, TOL-BZDH and TOL-BADH were eluted with a linear 50 - 350 mM-potassium phosphate buffer gradient (pH 7.5) containing 2 mM-DTT. The gradient volume was 1.0 l and the flow rate was 26 ml/h. Both TOL-BZDH and TOL-BADH were eluted after approximately 180 ml and fractions with peak activity were pooled. The pooled fractions (68 ml) were dialysed (Methods 2.2.4) for a total of 4 h against 8 changes (500 - 750 ml each) of buffer E (100 mM-potassium phosphate, pH 7.5/2 mM-DTT).

(c) Affinity chromatography on Matrex gel red A/Blue Sepharose CL-6B

The dialysed pooled fractions (68 ml) from step (b) were applied to a Matrex gel red column (2.6x8.0 cm) connected in series with a Blue Sepharose column (2.6x11.5 cm). Both columns were pre-equilibrated with buffer E and during application the flow (50 ml/h) was stopped for 20 min every 20 min. The flow rate was reduced to 25 ml/h and the columns were washed with 400 ml of buffer E. TOL-BADH passed through both columns and fractions with peak activity were pooled.

The Matrex gel red column was disconnected and the Blue Sepharose column was washed with 300 ml of buffer F (110 mM-potassium phosphate, pH 7.5/2 mM-DTT) at 85 ml/h. The flow through the column was then reversed and TOL-BZDH was eluted at 44 ml/h with buffer F plus 0.2 mM-NAD^+. As soon as enzyme activity was detected in the eluate the buffer was changed to buffer F plus 0.5 mM-NAD^+. Fractions with peak activity, containing homogeneous TOL-BZDH, were pooled.

(d) Hydrophobic interaction chromatography on Phenyl Sepharose CL-4B/DEAE-Sephacel

The pooled fractions (69 ml), containing TOL-BADH, from step (c) were applied to a Phenyl Sepharose column (2.6x11.0 cm) pre-equilibrated with buffer E. The column was washed with the following buffers in turn; 50 ml buffer E, 100 ml buffer G (5 mM-potassium phosphate, pH 7.5/2 mM-DTT), 200 ml of ice-cold buffer G containing 20% (v/v) ethanediol and finally with 250 ml of ice-cold buffer G containing 50% (v/v) ethanediol. The flow rate was reduced from 50 ml/h to 20 ml/h and TOL-BADH was eluted from the column with 250 ml of buffer G containing 65% (v/v) ethanediol. The eluate was passed through a DEAE-Sephacel column (1.6x3.5 cm) pre-equilibrated with buffer G. The DEAE-Sephacel column was washed with the following buffers in turn; 15 ml buffer G containing 50% (v/v) ethanediol, 15 ml of buffer G

containing 20% (v/v) ethanediol, 15 ml buffer G and finally with 35 ml of buffer D. The flow rate was reduced from 20 ml/h to 9.5 ml/h and TOL-BADH was eluted with 300 mM-potassium phosphate buffer, pH 7.5/2 mM-DTT. Fractions with peak activity were pooled and dialysed (Methods 2.2.4) for a total of 3 h against 6 changes (500 ml each) of 10 mM-potassium phosphate buffer, pH 6.5/2 mM-DTT/5 mM-$MgCl_2$.

(e) Affinity chromatography on Matrex gel green A

The dialysed pooled fractions (12.2 ml) from step (d) were applied at 80 ml/h to a Matrex gel green column (4.5x5.0 cm) pre-equilibrated with buffer H (20 mM-potassium phosphate, pH 6.5/2 mM-DTT/5 mM-$MgCl_2$). The column was washed with 15 ml of buffer H and the flow was then stopped for 30 min. The column was washed with a further 300 ml of buffer H before the flow rate was reduced to 30 ml/h and homogeneous TOL-BADH was eluted with 50 mM-potassium phosphate buffer, pH 6.5/2 mM-DTT/ 5 mM-$MgCl_2$/1 mM-NAD^+. Fractions with peak activity were pooled.

(f) Concentration and storage of the enzymes

The pooled fractions from steps (c) and (e), containing TOL-BZDH and TOL-BADH respectively, were concentrated by repeated use of two Centriprep concentrators (Methods 2.2.7). The concentrated samples were each mixed with an equal volume of ethanediol and stored at -20 ^{0}C.

2.6.3 Purification of BZDH II and BADH from *Acinetobacter calcoaceticus*

BZDH II and BADH were purified from the wild-type *A. calcoaceticus* N.C.I.B. 8250 essentially as described by MacKintosh & Fewson (1988a). Column washing was increased to ten column volumes at each stage in the purification and samples were

concentrated after the dye-ligand affinity steps using a small DEAE-Sephacel column rather than by vacuum dialysis.

2.6.4 Chromatofocusing

Purified BZDH I (90 µg; Methods 2.6.1) in 25 mM-Bistris/HCl buffer, pH 6.7 was applied to an f.p.l.c. Mono P chromatofocusing column (0.5x20 cm) pre-equilibrated with the same buffer. Protein was eluted with Polybuffer PB 74 diluted 1+9 with water and adjusted to pH 4.8 with HCl. The flow rate was 0.5 ml/min during sample application and 1.0 ml/min during elution. Fractions of 0.75 ml were collected and BZDH I was located in the eluate by enzyme assay (Methods 2.5.1a).

2.7 Polyacrylamide-gel electrophoresis (PAGE)

2.7.1 SDS-PAGE (discontinuous system)

SDS-PAGE was carried out as described by Laemmli (1970) except that 96 mM-Tris/HCl buffer replaced 125 mM-Tris/HCl buffer in the stacking gel. Slab gels (19.0x9.5x0.15 cm) consisted of a 5.3% (w/v) polyacrylamide stacking gel and a 12.5% (w/v) polyacrylamide separating gel. Gels were electrophoresed at 80 mA (constant current) and the apparatus was cooled by water circulating from an ice bath. Samples (40 µl, diluted with water if necessary) were prepared for electrophoresis by boiling for 2 min in a mixture with 25 µl 0.2 M-DTT and 10 µl of tracker dye mixture. The tracker dye mixture was prepared by adding 1.5 g SDS, 8 g sucrose and 100 µg pyronin Y to 10 ml 0.5 M-Tris/HCl (pH 6.8).

2.7.2 Phosphate-buffered SDS-PAGE

After the chemical cross-linking of enzyme subunits (Methods 2.9.1) quaternary structures were examined using a phosphate-buffered gel system in which 7.5% (w/v) polyacrylamide slab gels (19.0x9.5x0.15 cm) were prepared and used essentially as described by Shapiro & Maizel (1969).

The stock solutions for preparing the gels were as follows:

Solution A: 1 M-sodium phosphate, pH 6.5

Solution B: 20 % (w/v) SDS

Solution C: 28 % (w/v) acrylamide/0.735 % (w/v) *N,N'*-methylenebis - acrylamide

Solution D: 0.1 M-sodium phosphate, pH 6.5/0.1% (w/v) SDS

Slab gels were prepared with a degassed mixture containing 141 ml of water, 20 ml of Solution A, 1 ml of Solution B, 35 ml of Solution C, 100 µl of TEMED and polymerisation was initiated by the addition of 150 mg of ammonium persulphate.

Samples were prepared for electrophoresis by boiling for 5 min after the addition of an equal volume of tracker dye mixture. The tracker dye mixture (5 ml) contained 50 µl of Solution A, 0.25 ml of Solution B, 50 µl of 2-mercaptoethanol, 1 ml of glycerol and 0.1 ml of 0.5 % (w/v) bromophenol blue.

Gels were electrophoresed at 80 mA (constant current) using Solution D as reservoir buffer, until the tracker dye had reached the bottom of the gel. Another sample of tracker dye mixture was loaded onto the gel and electrophoresis was continued until the dye was approximately one third of the way down the gel.

2.7.3 Non-denaturing PAGE

Non-denaturing polyacrylamide-gels were prepared exactly as described in Methods 2.7.1 except that SDS was omitted from all of the solutions and the separating gel was only 7.5% (w/v) polyacrylamide. The samples were loaded directly onto the gel without any treatment or were first made 2% (v/v) in glycerol. Non-denaturing gels were electrophoresed at 20 mA (constant current) and the apparatus was cooled by water circulating from an ice bath.

2.7.4 Staining of the gels

(a) Coomassie blue stain

Gels were stained with Coomassie Brilliant Blue G250 using the reagents described by MacKintosh & Fewson (1988a).

(b) Silver stain

Gels were silver-stained by the method of Wray *et al.* (1981).

(c) Activity stain

Non-denaturing gels were activity-stained according to the method of MacKintosh & Fewson (1988a), except that the gels were first incubated in staining buffer (without staining reagents) for 30 min at 4 ^{0}C. Enzyme activity was detected as a purple precipitate of formazan.

2.7.5 Gel scanning

Coomassie-stained gels (Methods 2.7.4a) were scanned using an LKB 2202 Ultroscan laser densitometer.

2.7.6 Determination of the M_r values for enzyme subunits

The enzymes were each mixed with the calibration proteins and run on 12.5% (w/v) polyacrylamide gels (Methods 2.7.1) which were stained for protein (Methods 2.7.4a) and then scanned with an LKB Ultroscan laser densitometer. The calibration proteins were (M_r values in parentheses) phosphorylase *b* (94 000), bovine serum albumin (67 000), ovalbumin (43 000), carbonic anhydrase (30 000), trypsin inhibitor (20 100) and α-lactalbumin (14 400).

The distance between the start of the main gel and each calibration protein was plotted against ln (M_r) using the Cricket Graph V.1.1 program on an Apple Macintosh computer. The line of best fit was calculated by linear regression analysis and the equation of the line was used to find the apparent M_r values of the enzyme subunits. There was no need to convert the mobility of each of the calibration proteins into R_F values because the data from each track on the gels were plotted independently and no comparisons were made between tracks.

2.8 Immunological Methods

2.8.1 Production of antisera

The purified enzymes (Methods 2.6) were each mixed with 1.0 ml of Freund's complete adjuvant. The mixtures were made up to 2.0 ml with water and briefly sonicated to form an emulsion. The mixtures were subcutaneously injected into white

New Zealand rabbits at six sites. After six weeks the animals were boosted with further injections as described above except that Freund's incomplete adjuvant was used. After a further two weeks the animals were bled. The blood was kept overnight at 4 ^{0}C and the clotted material was removed. The antisera were stored at -20 ^{0}C.

2.8.2 Immunoprecipitation assay

Immunoprecipitation assays were carried out at 0 - 4 ^{0}C. The relative amounts of antiserum and antigen used were found by trial and error and depended upon the titre of the antiserum towards each antigen.

A serial dilution of antigen was prepared and an equal volume of each dilution was mixed with an equal volume of antiserum. The antiserum was either used undiluted or after dilution, depending upon the strength of the cross-reaction to be determined. After incubating for 1 h on ice the mixtures were microfuged (Methods 2.2.8) for 5 min. The supernatants, along with samples of the mixtures saved before centrifugation, were assayed for enzyme activity (Methods 2.5.1a & 2.5.2).

The amount of enzyme activity remaining in each mixture was expressed as a percentage relative to the activity of an enzyme sample that had been similarly treated except that the antiserum had been replaced by rabbit normal serum. The percentage of enzyme activity remaining in each mixture was plotted against the amount of antigen in the mixture (see Figure 6.2). The point of equivalence was the amount of antigen added when the percentage of activity first reached its minimum. The titre of an antiserum (μg of protein/ml of antiserum) is defined as the amount of antigen added at the point of equivalence multiplied by the dilution of the antiserum in the mixture.

2.8.3 *Staphylococcus aureus* immunoprecipitation assay

(a) Preparation of the *S. aureus* cells

S. aureus cells (Cowan I strain) were prepared as described by Kessler (1975). The cells were then washed three times by centrifugation (Methods 2.2.8) and resuspension to their original volume in 25 mM-potassium phosphate buffer, pH 7.5/2 mM-DTT.

(b) The assay

The *S. aureus* immunoprecipitation assay was carried out at 0 - 4 ^{0}C.

A small amount of enzyme was diluted to 50 µl with 25 mM-potassium phosphate buffer, pH 7.5/2 mM-DTT and mixed with an equal volume of undiluted antiserum. The amount of enzyme used was as small as possible, considering the need to assay the activity at the end of the procedure. The mixture was incubated for 1 h on ice and 100 µl of *S. aureus* cells were added. After a further 45 min on ice, with occasional mixing, the samples were microfuged (Methods 2.2.8) for 5 min and the supernatants were assayed for enzyme activity (Methods 2.5.1a & 2.5.2).

2.8.4 Quantitative immunoblotting

The following buffers were used in the iodination of protein A and in the nitrocellulose blotting:

Buffer A: 20 mM-Tris/HCl, pH 7.5/0.15 M-NaCl

Buffer B: 190 mM-glycine/25 mM-Tris (base)

Buffer C: Buffer B containing 0.02% (w/v) SDS

Buffer D: 10 mM-Tris/HCl, pH 7.2/0.15 M-NaCl/0.5% (w/v) Tween 20

(a) Preparation of [^{14}C]-India ink and iodination of protein A

[^{14}C]-India ink was prepared by mixing $1.9x10^5$ Bq of D-[U-^{14}C] glucose with 1 ml of India ink.

Up to 3 mg of protein A was dissolved in 0.5 ml of buffer A and incubated for 15 min in a mixture with $3.7x10^7$ Bq of Na^{125}I and 5 Iodobeads (Pierce Europe B.V., OUD-Beijerland, The Netherlands). A further 0.5 ml of buffer A was added and the mixture was gel filtered through a Pharmacia PD-10 prepacked column that had been pre-equilibrated with buffer A. Fractions of approximately 1 ml were collected and counted. Fractions containing the first peak of activity were pooled (approximately 3 ml) and diluted with buffer A so that 1 ml of the final solution contained approximately 100 µg of protein A (assuming no losses) and $1.2x10^6$ Bq of ^{125}I.

(b) Electroblotting of antigens from SDS-polyacrylamide gels

Purified proteins (0.05 - 0.4 µg) were subjected to SDS-PAGE (Methods 2.7.1) at 40 mA (constant current). One or two gels were placed in the Bio-Rad Trans-Blot cell and the proteins were electroblotted onto nitrocellulose membranes (Towbin *et al.*,1979) at 7.1 V/cm for 2 h .

(c) Application of proteins directly to nitrocellulose membranes

The purified proteins were made 2% (w/v) in SDS by the addition of the appropriate amount of 20% (w/v) SDS and the protein concentration was determined by the method of Lowry (Methods 2.2.5). The samples were then boiled for 2 min after they had been made 70 mM in DTT by the addition of the appropriate amount of 2 M-DTT. The protein solutions were next diluted with buffer B such that each contained 0.02% (w/v) SDS. The proteins were then ready for application to the nitrocellulose membranes, however

they were usually diluted with buffer C so that 200 µl of the solution contained 0.8 µg of protein.

Nitrocellulose membranes were wetted in buffer C and placed in the Bio-Rad Dot Blot SF apparatus. Up to 200 µl of protein solution (0.8 µg of protein) was placed in each well and if less than this amount was used the volume was made up to 200 µl with buffer C. A slight vacuum was applied to the apparatus so that the samples were drawn through the nitrocellulose membrane over a period of several minutes. The wells were then washed with 2x200 µl of buffer C and the membrane was quickly removed from the apparatus.

(d) Probing of the nitrocellulose membranes

Non-specific protein binding sites were blocked by incubating the membranes in buffer D for at least 6 h, at 4 ^{0}C. Antiserum (0.4 ml) in 40 ml of buffer D containing 5% (v/v) heat-inactivated horse serum was used to probe each membrane. After incubating overnight with gentle agitation at 23 ^{0}C, the membranes were washed five times in buffer D (15 min each) to remove any unbound antiserum. The membranes were next incubated for 2 h with gentle agitation, in 40 ml of buffer D containing ^{125}I-protein A (100 µg, 1.2x10^{6} Bq) . Washing was then repeated as described above. Finally, the membranes were dried at 30 ^{0}C between sheets of filter paper.

Reference spots of [^{14}C]-India ink were applied to the membranes which were autoradiographed for 24 - 72 h. After development, the reference spots were placed in register with those on the membranes thus allowing the location of the radioactive bands. Radioactive bands were cut out and counted using an LKB 1275 Minigamma. Blank sections of each membrane were also cut out and counted to provide the number of background counts which were subtracted from the number of counts in each band.

It was not necessary to autoradiograph membranes which had been loaded using the Bio-Dot SF apparatus [see part (c) above] because each well caused a visible indentation on the nitrocellulose membrane.

2.9 Chemical modifications of the enzymes

2.9.1 Cross-linking of enzyme subunits by dimethylsuberimidate

The method was based on that described by Lumsden & Coggins (1977). The following solutions were used:

Solution A: 50 mM-triethanolamine /HCl buffer, pH 8.0/0.1 M-NaCl/10 mM-$MgCl_2$

Solution B: as above but twice as concentrated

Solution C: 0.4 M-NaOH

Equal volumes of Solutions B and C were mixed and used to make a 0.1 M solution of dimethylsuberimidate. This was used immediately because the cross-linking reagent is unstable in the presence of water.

Protein samples (1 - 2 mg/ml) were dialysed (Methods 2.2.4) overnight against Solution A and then mixed with enough 0.1 M-dimethylsuberimidate to give a final concentration of 6 mM. After incubating for 30 min at room temperature the samples were treated again with a fresh solution of dimethylsuberimidate as described above. After incubating for a further 30 min the samples were analysed using a phosphate-buffered SDS-PAGE system (Methods 2.7.2).

2.9.2 Reduction and carboxymethylation

Proteins were dialysed (Methods 2.2.4) against 0.5% ammonium bicarbonate and lyophylised (Methods 2.2.6). They were then carboxymethylated using iodoacetate in the presence of DTT as described by Lumsden and Coggins (1978). Finally, the samples were dialysed and lyophylised as described above.

2.10 Amino acid compositions

Protein (800 µg) was dialysed (Methods 2.2.4) for two days against four changes of 5 mM-potassium phosphate buffer, pH 7.5 and divided equally among four acid washed (Methods 2.2.3) Pyrex test tubes, each containing 15 nmol of norleucine.

Three of the samples were dissolved in 1 ml of 6 M-HCl and the tubes were flushed with N_2 and sealed under vacuum. The samples were hydrolysed at 110 ^{0}C for 24, 48 and 72 h before the tubes were opened and the contents dried under vacuum.

The fourth protein sample was oxidised by incubating for 4 h in 1 ml of performic acid. Performic acid was prepared by mixing one volume of 30% (w/v) H_2O_2 and nine volumes of 98 - 100% formic acid (the mixture was first allowed to stand for 1 h to allow the generation of performic acid). After oxidation of the protein the mixture was dried under vacuum in the presence of solid NaOH and then twice resuspended in 1 ml of water and lyophylised (Methods 2.2.6). Finally, the oxidised protein was acid hydrolysed for 72 h as described above.

All four hydrolysed samples were twice resuspended in 100 µl of water and dried under vacuum. The amino acid mixtures prepared from the *Acinetobacter* enzymes were analysed using an LKB 4400 amino acid analyser, operated by Mr. J. Jardine according to the manufacturer's instructions. The amino acids were quantified by the spectrophotometric detection of the dansyl-amino acid derivatives. The amino acid mixtures prepared from the *Pseudomonas* enzymes were analysed using a Waters PICO TAG amino acid analyser, operated according to the manufacturers instructions by Dr D. G. Campbell in the Protein Phosphorylation Group, Dept. of Biochemistry, University of Dundee. The amino acids were quantified by the spectrophotometric detection of the phenylthiohydantoin (PTH)-amino acid derivatives.

The amino acid composition of each sample was normalised to one amino acid and the percentage composition of each amino acid was calculated. There was very little difference between the results obtained after normalising to glutamate or phenylalanine

and the results presented in this thesis were obtained after normalising to glutamate.

The amino acid compositions of the four samples were used to determine the amino acid composition of the protein by using the following rules:

(1) the values for methionine and cysteine were taken from the amounts of methionine sulphone and cysteic acid, respectively, in the oxidised sample,

(2) if the composition of an amino acid decreased with the time of acid hydrolysis (for example serine, threonine and tyrosine), the 'true' value was obtained by extrapolating back to zero time of acid hydrolysis,

(3) the values for most amino acids were an average of the values from the three unoxidised samples, however occasionally a value that appeared to be wild was arbitrarily eliminated

2.11 Protein sequencing

2.11.1 Electroelution of enzymes from non-denaturing polyacrylamide gels

Purified BZDH II and BADH (Methods 2.6.3) were electrophoresed on non-denaturing polyacrylamide gels (Methods 2.7.3). Each of the proteins was separated into two bands and these were electroeluted from the gels as described by Findlay *et al.* (1989) and lyophylised (Methods 2.2.6).

2.11.2 Preparation of samples

Purified proteins were carboxymethylated (Methods 2.9.2) and dialysed (Methods 2.2.4) for four days against eight changes of at least 1000 volumes. Proteins sequenced at Aberdeen University were dialysed against water while those sequenced at Leeds

University were dialysed against 0.1% (w/v) SDS.

Proteins that were eluted from polyacrylamide gels were sequenced without further treatment.

2.11.3 The sequencers

Some of the protein sequencing was carried out by Dr J. N. Keen in the Department of Biochemistry, University of Leeds, by automated solid-phase Edman degradation (Laursen, 1971) using the microsequence facility built by the Protein Sequence Unit. PTH-Amino acids were identified off line by reverse-phase (C_{18}) microbore HPLC. A full description of the sequencing protocol has been given by Findlay *et al.* (1989). Other sequences were determined by Mr. B. Dunbar at the Department of Biochemistry, University of Aberdeen using an Applied Biosystems Protein Sequencer, Model 470A/477A. PTH-Amino acids were identified off line by reverse-phase chromatography using a Waters Associates chromatography system with and Apex Cyano column (4.6 mm x 25 cm), and elution was carried out with a gradient of acetate/acetonitrile.

2.12 Safety

Bacterial cultures were killed by autoclaving before disposal and glassware was washed as described in Methods 2.2.3. Any bacterial spillage was swabbed with 10% (v/v) propan 1-ol.

All other precautions taken in the interests of safety are described in the University of Glasgow Safety Handbook (1987) and the Radiation Protection Service Health Physics Notes (1986).

CHAPTER 3

Purification of Benzaldehyde Dehydrogenase I from
Acinetobacter calcoaceticus C1005

3.1 Preliminary studies on maximising enzyme activity

A. calcoaceticus C1005 has previously been used as a source of L-mandelate dehydrogenase (Hoey *et al.*, 1987) and phenylglyoxylate decarboxylase (Barrowman & Fewson, 1985). These enzymes are co-ordinately induced along with BZDH I (Livingstone & Fewson, 1972) and optimised methods were already available for the growth, harvest and disruption of the cells (Methods 2.4.2, 2.4.3 &2.4.4). This section is therefore concerned only with the stability of BZDH I in extracts.

3.1.1 Effects of different buffers on the stability of enzyme activity

Samples of extract (Methods 2.4.4b) were dialysed overnight (Methods 2.2.4) against the following buffers at a concentration of 50-mM: potassium phosphate, pH 7.0, 7.5 and 8.5; maleate/KOH, pH 6.0 and 7.0; Tris/HCl, pH 7.5 and 8.5. After storage for five days at 4 ^{0}C, 90% of the activity remained in the potassium phosphate, pH 7.5 sample. The samples in maleate buffer were completely inactive and the samples in Tris buffer retained approximately 50% of their activity. Potassium phosphate buffer, pH 7.5 was therefore adopted as the buffer for the purification procedure.

3.1.2 The effect of DTT on the stability of enzyme activity

During some preliminary purification experiments there was a substantial loss of BZDH I activity. This loss could be partially reversed or almost completely avoided by making all buffers 2 mM in DTT.

3.2 Development of purification procedures

Various different techniques were used in preliminary experiments to establish a purification scheme for BZDH I. This section describes the behaviour of the enzyme in all of the purification steps that were attempted including the ones that do not appear in the final purification procedure (Methods 2.6.1).

3.2.1 Ammonium sulphate fractionation

BZDH I was precipitated from an extract (Methods 2.4.4b) between 35% and 45% ammonium sulphate saturation. The resuspended pellet was purified two-fold, with a recovery of 75%. Despite a poor degree of purification, specially for a first step, ammonium sulphate fractionation was considered to be potentially useful because of the reduction in sample volume that can be achieved.

3.2.2 Gel filtration through Sephacryl S300 and Sephadex G150

Samples of extract (Methods 2.4.4b) or resuspended ammonium sulphate pellet (Section 3.2.1) were gel filtered through columns of Sephacryl S300 or Sephadex G150. In each case the column buffer was 25 mM-potassium phosphate, pH 7.5/2 mM-DTT. The best purification achieved with extracts was only two-fold with a yield of 95%. When the sample had previously been fractionated by ammonium sulphate the best purification over the extract was only four-fold.

3.2.3 Ion-exchange chromatography on DEAE-Sephacel

BZDH I was found to bind tightly to DEAE-Sephacel when the extract (Methods 2.4.4b) and the ion-exchange material had been equilibrated with 25 mM-potassium

phosphate buffer, pH 7.5/2 mM-DTT. Elution could be achieved by increasing the concentration of the potassium phosphate buffer. The best method of elution with respect to purification fold, recovery and eluate volume was a linear gradient of 25 - 200 mM-potassium phosphate buffer, pH 7.5 containing 2 mM-DTT. This step was incorporated into the final purification procedure (Methods 2.6.1 & Section 3.3.1).

3.2.4 Isocratic elution from DEAE-Sephacel

Isocratic elution of an ion-exchange column is when the column is developed by starting buffer. The buffer must be such that protein can equilibrate rapidly between the the stationary and the mobile phase. As in true chromatographic systems the resolution of this method is best when the column is long and thin and the sample volume is small. BZDH I was isocratically eluted from a column of DEAE-Sephacel if the sample and the column were equilibrated with 70 - 80 mM-potassium phosphate buffer, pH 7.5/ 2 mM-DTT.

Experiments were carried out using a DEAE-Sephacel column (1.6 x 33 cm) and a slightly modified form of isocratic elution. A sample of partially purified BZDH I (Section 3.2.3) was applied to the column after both had been equilibrated with 25 mM-potassium phosphate buffer, pH 7.5/2 mM-DTT. At this ionic strength BZDH I was bound tightly to the column and the sample volume could therefore be relatively large because of the concentrating effect of the column. After extensive washing, elution was achieved by stepping the potassium phosphate concentration of the buffer from 25 to 70 mM. The first peak of protein was detected after the column had been developed with approximately one column volume of buffer and BZDH I was eluted after almost two column volumes.

The column was probably acting in the following manner. When the 70 mM-potassium phosphate buffer was applied to the column phosphate was taken up by the ion-exchange material in exchange for OH^- ions or together with H^+ ions. This

caused the pH on the column to rise. Although conditions were now less powerfully eluting (with respect to pH) some proteins were carried down the column because of the increased ionic strength. As elution continued the ion-exchange material was titrated back to the original pH and a front of equilibration passed down the column. When this emerged from the column a peak of protein was observed in the eluate. This contained proteins that moved down the column in the alkaline zone more slowly than the front of equilibration, but moved faster than this in the region behind the front. Proteins that emerged after the front (such as BZDH I) must have passed down the column in the equilibrated region and their elution was therefore isocratic.

Isocratic elution of the DEAE-Sephacel column gave a purification of almost two-fold over the starting material which had already been subjected to ion-exchange chromatography. This could have been a particularly useful step, specially in the latter stages of the purification, had it not been rendered impractical by the long washing time required and by the relatively large elution volume.

3.2.5 Affinity chromatography on immobilised triazine dye-ligands

Triazine dye-ligand chromatography uses immobilised synthetic dyes to bind proteins (Amicon, 1980). It has been assumed that some structural feature of the dyes resembles the $NAD(P)^+$ cofactor because dehydrogenase enzymes are specifically bound and can often be eluted by low concentrations of cofactor. In some cases non-specific ionic or hydrophobic interactions have been implicated in binding because some proteins are released by a relatively small rise in ionic strength or by a decrease in the polarity of the buffer (Amicon, 1980; Dean & Watson, 1979).

A set of experiments was carried out using an Amicon Dye Matrex Screening Kit which contained small columns of immobilised blue A, blue B, red A, green A and orange dyes. In general, the binding capacity of these materials is highest at low pH, low ionic strength and in the presence of metal ions such as Mg^{2+} (Amicon, 1980; Dean & Watson, 1979). Experiments were carried out in potassium phosphate buffers ranging

from 10 to 100 mM and from pH 6 to 7, both in the presence and in the absence of Mg^{2+}. The binding capacity of all of the columns for BZDH I was low except at low pH and low ionic strength. Unfortunately under these conditions little activity was recovered after elution by the cofactor NAD^+, even at concentrations of up to 10 mM. Elution was also attempted by increasing the pH and the ionic strength, however this gave a low yield and a poor degree of purification.

3.2.6 Hydroxyapatite chromatography on HA-Ultrogel

HA-Ultrogel is composed of crystals of calcium phosphate hydroxide encased in cross-linked agarose beads. It is clear that many proteins bind by ionic interaction, however in certain cases it may also act as an affinity matrix specific for proteins that bind calcium.

BZDH I bound to HA-Ultrogel when the buffer was 5 mM-potassium phosphate, pH 6.8. Elution was achieved by a linear gradient of 5 - 200 mM-potassium phosphate, pH 6.8. This was a good step with respect to the purification achieved, which varied according to the previous treatment of the sample. However, the large volume of eluate necessitated a subsequent concentration step which was time consuming and reduced the yield.

3.2.7 Hydrophobic-interaction chromatography on Phenyl Sepharose CL-4B

Phenyl Sepharose binds proteins by hydrophobic interactions. Binding is strongest at high ionic strength and elution can generally be achieved by lowering the ionic strength and/or including reagents that reduce the polarity of the buffer.

BZDH I bound to Phenyl Sepharose when the buffer was 5 mM-potassium phosphate, pH 7.5. BZDH I first appeared in the eluate when the ethanediol content of the buffer was increased to 30% (v/v) and most of the activity was eluted at 40% (v/v)

ethanediol. The volume of the eluate was large and the problem of concentrating it was compounded by the presence of ethanediol in the buffer. These difficulties were overcome if the sample was concentrated and freed from ethanediol by immediately passing the eluate from the Phenyl Sepharose column through a small DEAE-Sephacel column which captured BZDH I. The DEAE-Sephacel column was then washed free of ethanediol and eluted by an increase in ionic strength.

This combined purification step was successful with respect to the purification fold and the convenience of a small eluate volume. It was incorporated into the final purification protocol and a full description is given in Methods 2.6.1 and Section 3.3.2 .

3.3 Purification of BZDH I

A protocol, based on the preliminary purification studies (Section 3.2) was developed for the purification of BZDH I. It includes a gel-filtration step using an f.p.l.c. Superose 12 column that is not described in the preceding section. This step was carried out by following the recommendations of the column manufacturers (Pharmacia) and no developmental work was needed. The purification scheme is described in Methods 2.6.1 and is summarised in Table 3.1.

3.3.1 Ion-exchange chromatography on DEAE-Sephacel

A typical profile of the chromatography of an extract of *A. calcoaceticus* C1005 (Methods 2.6.1a) on DEAE-Sephacel is shown in Figure 3.1. BZDH I was eluted after approximately 160 ml of a 500 ml linear gradient of 25 - 200 mM-potassium phosphate buffer. There was a purification of seven-fold and a yield of 56%.

3.3.2 Hydrophobic-interaction chromatography on Phenyl Sepharose CL-4B

Chromatography on Phenyl Sepharose is the major purification step for BZDH I. The DEAE-Sephacel pool (Methods 2.6.1b) was applied to a column of Phenyl Sepharose and after washing the enzyme was eluted by buffer plus 40% (v/v) ethanediol. BZDH I was captured by passing the eluate through a column of DEAE-Sephacel and after it had been washed free of ethanediol BZDH I was eluted by a step in ionic strength. A typical elution profile for the small DEAE-Sephacel column is shown in Figure 3.2. BZDH I was purified a further 17-fold with a yield of 46%.

3.3.3 Gel-filtration chromatography through Superose 12

A typical elution profile of the chromatography of the concentrated Phenyl Sepharose pool (Methods 2.6.1c) on an f.p.l.c. Superose 12 column is shown in Figure 3.3. SDS-PAGE of the Phenyl Sepharose pool showed that there were only a few faint bands contaminating the BZDH I preparation (Figure 3.4). The contaminating proteins removed by chromatography on Superose 12 were so minor that there was apparently no increase in specific activity (Table 3.1), however there was a yield of 92%.

3.3.4 Storage of the purified enzyme

BZDH I was stable at 4 ^{0}C in potassium phosphate buffer, pH 7.5 (Sections 3.1.1 & 3.1.2). After the final purification step the enzyme was in 100 mM-potassium phosphate buffer, pH 7.5/2 mM-DTT and no activity was lost when it was mixed with an equal volume of ethanediol. The enzyme could then be stored at -20 ^{0}C without freezing and it was stable under these conditions for at least several months.

3.3.5 Overall purification

Details of a typical purification are given in Table 3.1. The specific activity of the purified enzyme was quite constant among independent preparations (25.5 units/mg of protein, SD = 2.5, n = 9) but the fold purification was dependent upon the specific activity of the extract and this varied from preparation to preparation.

3.3.6 Purity

The purification of BZDH I was monitored using SDS-PAGE (Methods 2.7.1; Figure 3.4a) with a protein stain (Methods 2.7.4a). The enzyme preparation was also checked for homogeneity, and specially for the presence of any contaminating BZDH II, using non-denaturing PAGE (Methods 2.7.3) with an activity stain (Methods 2.7.4c; Figure 3.4b) and also with a protein stain (Methods 2.7.4a; Figure 3.4c).

An activity stain of BZDH I in an extract and when purified is shown in Figure 3.4c, Tracks A and B respectively. Tracks C and D of the same gel show BZDH II in an extract and when purified (Methods 2.6.3), respectively. BZDH II produced two major bands in the cell-extract and when purified (Tracks C &D). This effect is probably due to minor charge heterogeneities since both of the major bands have the same amino terminal sequence (Section 5.1.4). Track D (purified BZDH II) also has a minor band corresponding in position to BZDH I. This was caused by carry-over from Track E; the band is absent from Track F which shows BZDH II at a higher loading than in Track D.

The purity of the BZDH I preparation was also confirmed by isoelectric focusing on an f.p.l.c. Mono P column (Section 5.1.3).

Table 3.1 Purification of BZDH I from *A. calcoaceticus* C1005

Details of the purification procedure are given in Methods 2.6.1. Enzyme assays were carried out with 100 µM-benzaldehyde (Methods 2.5.1a) and protein was measured as described in Methods 2.2.5.

	Volume (ml)	Total activity (units)	Total protein (mg)	Specific activity (units/mg of protein)	Yield (%)	Purification (fold)
Extract	50	197	895	0.22	100	1
DEAE-Sephacel	45	110	75	1.47	56	6.7
Phenyl Sepharose	5.6	51	2.0	25.5	26	116
Superose 12	1.0	47	1.85	25.4	24	116
Ethanediol added	2.0	47	1.85	25.4	24	116

Figure 3.1 Chromatography of an extract of *A. calcoaceticus* C1005 on DEAE-Sephacel

An extract (50 ml, Methods 2.6.1a) of *A. calcoaceticus* C1005 was applied to a column of DEAE-Sephacel (2.6 x 9.0 cm). After column washing, elution was carried out with a linear gradient of potassium phosphate buffer (Methods 2.6.1b). The flow rate was 19 ml/h and fractions of 6.4 ml were collected.

BZDH I activity, □; A_{280}, △; specific conductivity, ◇.

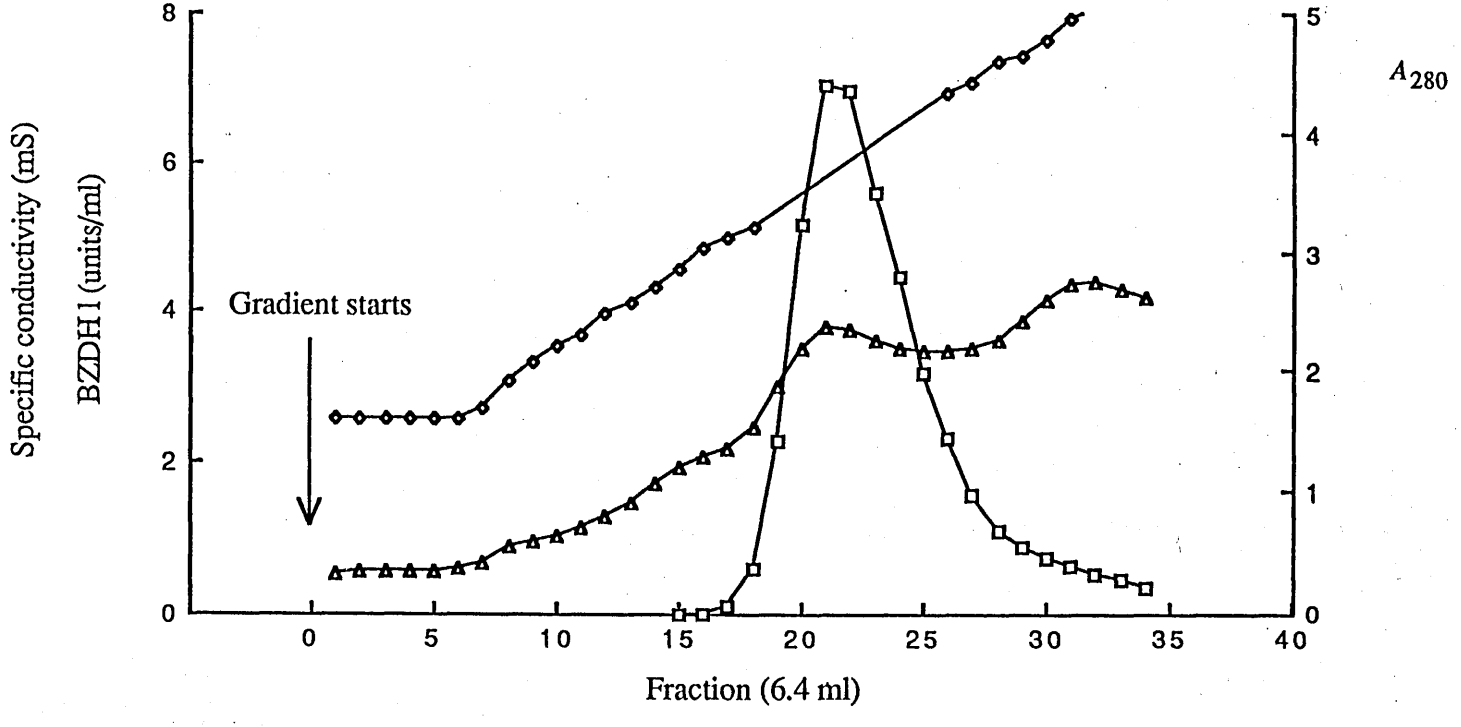

Specific conductivity (mS)
BZDH1 (units/ml)
A_{280}
Gradient starts
Fraction (6.4 ml)
0
2
4
6
8
0
1
2
3
4
5
0
5
10
15
20
25
30
35
40

Figure 3.2 Chromatography of a DEAE-Sephacel pool on DEAE-Sephacel following chromatography on Phenyl Sepharose CL-4B

A DEAE-Sephacel pool (45 ml, Methods 2.6.1b) was applied directly to a column of Phenyl Sepharose (2.6 x 9.8 cm). After washing with buffer plus 30% (v/v) ethanediol the column was eluted with buffer plus 40% (v/v) ethanediol through a second column of DEAE-Sephacel (1.6 x 3.0 cm) which captured BZDH I (Methods 2.6.1c). The DEAE-Sephacel column was washed free of ethanediol and eluted by a step in ionic strength (Methods 2.6.1c). The flow rate was 6.0 ml/h and fractions of 1.85 ml were collected.

BZDH I activity, □; A_{280}, △.

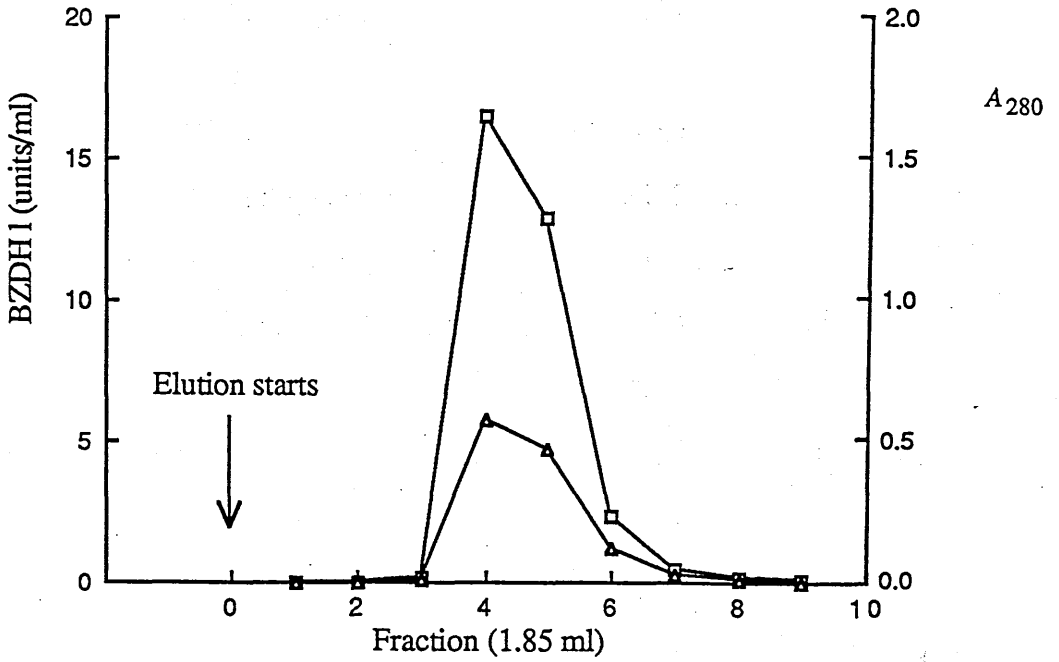

BZDH1 (units/ml)
20
15
10
5
0
2.0
1.5
1.0
0.5
0.0
A_{280}
Elution starts
0
2
4
6
8
10
Fraction (1.85 ml)

Figure 3.3 Chromatography of a concentrated Phenyl Sepharose pool on f.p.l.c. Superose 12

A Phenyl Sepharose pool (5.6 ml; Methods 2.6.1c) was concentrated to 0.18 ml (Methods 2.2.7) and injected onto a Superose 12 column (1.0 x 30 cm). The flow rate was 0.3 ml/min and fractions of 0.25 ml were collected (Methods 2.6.1d). The homogeneity of the fractions were checked by SDS-PAGE before they were combined.

BZDH I activity, □; A_{280}, △.

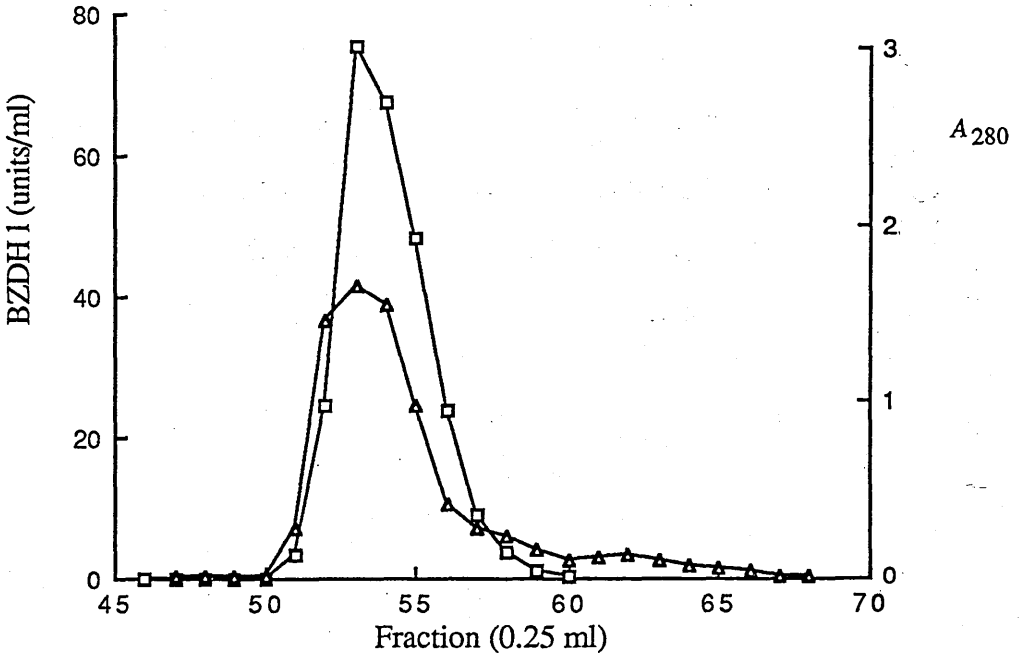
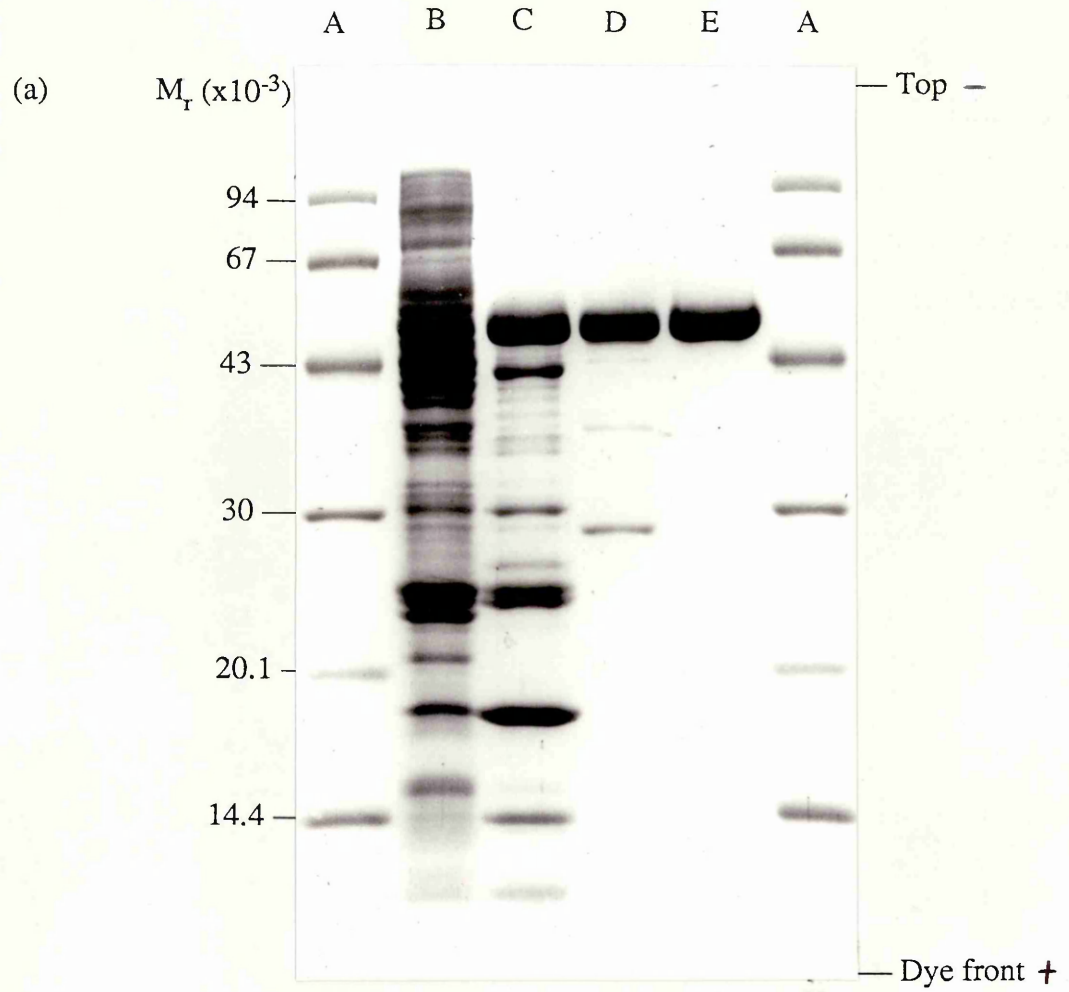

80
60
40
20
0
BZDH 1 (units/ml)
3
2
1
0
A_{280}
45
50
55
60
65
70
Fraction (0.25 ml)

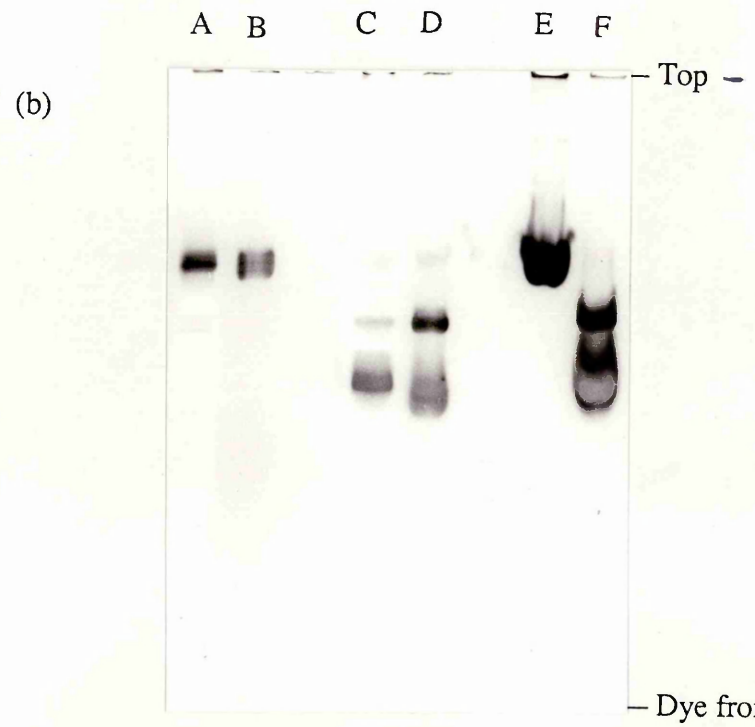
Figure 3.4 The purification of BZDH I as monitored by PAGE

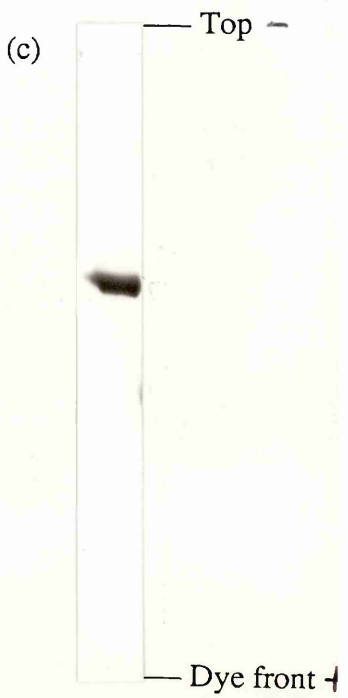
(a) A denaturing 12.5% (w/v) polyacrylamide slab gel containing 0.1% SDS was stained for protein. Tracks: A, M_r markers (approx. 1 μg of each); B, extract (42 μg of protein); C, DEAE-Sephacel pool (23 μg of protein); D, Phenyl Sepharose pool (7.3 μg of protein) and E, Superose 12 pool (8 μg of protein).

(b) A non-denaturing 7.5% (w/v) polyacrylamide slab gel was stained for activity; showing BZDH I together with BZDH II. Tracks: A, *A. calcoaceticus* C1005 extract (89 μg of protein); B, pure BZDH I (1.6 μg of protein); C, *A. calcoaceticus* N.C.I.B. 8250 extract (BZDH II induced; 180 μg of protein); D, pure BZDH II (14 μg of protein); E, pure BZDH I (40 μg of protein) and F, pure BZDH II (60 μg of protein).

(C) A non-denaturing 7.5% (w/v) polyacrylamide slab gel was stained for protein; 18 μg pure BZDH I.


A
B
C
D
E
A
(a)
M_r (x10^{-3})
94
67
43
30
20.1
14.4
Top
Dye front


A
B
C
D
E
F
(b)
Top
Dye front


(c)
Top
Dye front

3.3.7 Discussion

BZDH I was purified from *A. calcoaceticus* C1005 in three steps by ion-exchange, hydrophobic-interaction and gel-filtration chromatography (Methods 2.6.1). The procedure can be carried out in three days with approximately 116-fold purification and 24% yield, producing 1 - 2 mg of pure protein. BZDH I is stable both in extracts and after purification provided that DTT is included in the buffer (Section 3.1.2). The loss of about 76% of the activity during the purification was not due to the instability of BZDH I, but was due to the fact that not all of the activity recovered at each step of the purification went forward to subsequent step(s). For example, a substantial amount of activity was lost from the Phenyl Sepharose column during washing with 30% ethanediol (Methods 2.6.1c). Generally, approximately 80% of the activity was accounted for at the end of the purification procedure.

In the preliminary experiments to establish a purification procedure for BZDH I (Section 3.2) considerable effort was devoted to finding suitable conditions for chromatography on one of the triazine dye-ligand materials. These materials are capable of great resolution, especially if elution can be achieved with cofactor. Unfortunately it was not possible to find condition under which BZDH I was eluted by NAD^+ (Section 3.2.5). In retrospect, elution should also have been attempted with NADH.

The purification of BZDH I was monitored by SDS-PAGE (Section 3.3.6; Figure 3.4a). This is a sensitive and reproducible method but in one respect it is inappropriate for assessing the purity of BZDH I; it is unable to demonstrate that a BZDH I preparation is free from BZDH II because both proteins run in the same position. The validity of the immunological study (Chapter 6) depends particularly upon the homogeneity of BZDH I with respect to BZDH II. The two enzymes were easily separated by non-denaturing PAGE and the BZDH I preparation was shown to be free from BZDH II (Figure 3.4b,c). The relative purity of each of the five alcohol and aldehyde dehydrogenases characterised in this project is discussed further in Section 6.1.1.

CHAPTER 4

Purification of the Benzaldehyde Dehydrogenase and the Benzyl Alcohol Dehydrogenase Encoded by the TOL Plasmid pWW53 in *Pseudomonas putida* MT53

4.1 Preliminary studies on maximising enzyme activity

Extracts were prepared by ultrasonication of a cell suspension followed by centrifugation (Methods 2.4.4a).

4.1.1 Cell growth and enzyme induction

The enzymes of the toluate pathway, encoded by the TOL plasmid pWW53 in *Pseudomonas putida* MT53, are induced during growth in the presence of 3-methylbenzyl alcohol (*Keil et al.*, 1985). The highest specific activities for TOL-BZDH and TOL-BADH were achieved during growth on a defined medium (Murray *et al.*, 1985; Methods 2.4.2c) with 3-methylbenzyl alcohol as carbon source. The highest total activity was achieved by growth on Luria broth in the presence of 3-methylbenzyl alcohol (Methods 2.4.2c). A series of experiments was carried out to establish a practical compromise between the total activity and the specific activity of the two enzymes. The experiments were complicated and reiterative in nature and what follows is a summary of the conclusions that were drawn. All of the cultures were grown at 30 ^{0}C and most were 50 ml in volume (Methods 2.4.2a).

1 Growth on Luria broth plus 3-methylbenzyl alcohol gave higher total activity and specific activity than growth on yeast extract or tryptone plus 3-methylbenzyl alcohol.

2 3-Methylbenzyl alcohol above 5 mM was toxic to bacteria grown in defined or complex media and concentrations of less than 5 mM were less effective at inducing the enzymes.

3 After 6 h growth on Luria broth the specific activity of TOL-BZDH decreased steadily and that of TOL-BADH declined transiently and then increased. The apparent rise in TOL-BADH activity may have been due to the expression of a chromosomally encoded alcohol dehydrogenase.

4 Growth on full strength Luria broth give the best TOL-BZDH specific activity and growth on one quarter strength Luria broth gave the best TOL-BADH specific activity.

5 The addition of acetate or lactate or glucose to Luria broth reduced the specific activity of both enzymes.

6 The addition of acetate or lactate or glucose to defined media cultures increased the total activity and the specific activity of the enzymes when the inoculum was small but made little difference when the inoculum was large and had been grown on Luria broth.

7 The addition of potassium phosphate buffer to Luria broth cultures (intended to maintain the pH at neutrality during growth) did not improve the total activity or specific activity of either of the enzymes.

8 The most effective inocula for defined media cultures were grown on Luria broth in the absence of 3-methylbenzyl alcohol.

The best compromise between the total activity and the specific activity of TOL-BADH and TOL-BZDH was when a defined medium containing 7.5 mM-3-methylbenzyl alcohol was inoculated with 33% of its own volume of an 8 h Luria broth culture. These conditions are fully described in Methods 2.4.2c.

4.1.2 Extraction buffer

A bacterial pellet (Methods 2.4.2c & 2.4.3) was divided and resuspended in the following buffers at a concentration of 50 mM: Mes/KOH, pH 6.5; potassium phosphate, pH 7.5; Tes/KOH, pH 7.5; Hepes/KOH, pH 7.5; Tricine/KOH, pH 8.0; Bicine/HCl, pH 8.0 and glycylglycine/KOH, pH 8.0. The cell suspensions were disrupted by ultrasonication and an extract was prepared by centrifugation (Methods 2.4.4a). The extract prepared in potassium phosphate buffer, pH 7.5 had a slightly

higher specific activity for TOL-BZDH than the others. This buffer was therefore adopted as the cell extraction buffer (Methods 2.4.4).

4.1.3 Preparation of large amounts of cell-extract

It was impractical to prepare enough extract for the purification procedure by ultrasonic disruption (Methods 2.4.4a) and a method for French pressure disruption was therefore developed.

A culture pellet was resuspended in 100 mM-potassium phosphate buffer, pH 7.5 and passed repeatedly through a French pressure cell (Allison *et al.*, 1985). After each passage a sample was removed and centrifuged at 105 000 g for 150 min at 4 ^{0}C. The specific activity of TOL-BZDH and TOL-BADH increased with each passage up to a maximum after the fourth and final passage.

4.1.4 Effect of different buffers on the stability of enzyme activity

A French press extract (Methods 2.4.4b) was dialysed overnight (Methods 2.2.4) against the following buffers at a concentration of 50 mM: succinate/KOH, pH 6.0; potassium phosphate, pH 6.5 and 7.5; Tris/HCl, pH 7.5 and 8.5 and glycine/KOH, pH 9.0. The dialysed extracts were then stored at 4 ^{0}C and assayed for enzyme activity each day for three days. TOL-BADH was most stable in the potassium phosphate buffer, pH 6.5 and the succinate buffer, pH 6.0. TOL-BZDH was most stable in the potassium phosphate buffer, pH 7.5. Since the alcohol dehydrogenase activity was more stable than the aldehyde dehydrogenase activity the best buffer for the aldehyde dehydrogenase (potassium phosphate, pH 7.5) was adopted for the purification procedure. After three days (at 4 ^{0}C) in potassium phosphate buffer, pH 7.5, 18% of the TOL-BZDH activity and 52% of the TOL-BADH activity remained.

4.2 Development of purification procedures

Various techniques were used in preliminary experiments to establish a purification scheme for TOL-BZDH and TOL-BADH. This section describes the behaviour of the enzymes in all of the purification steps that were attempted, including the ones that do not appear in the final purification procedure (Methods 2.6.2). A brief description of some of the chromatography media is given in Section 3.2.

4.2.1 Ammonium sulphate fractionation

TOL-BZDH and TOL-BADH were precipitated from an extract (Methods 2.4.4b) between 20% and 50% ammonium sulphate saturation. When the pellet was resuspended, both enzymes were purified 1.4-fold with a recovery of 95%. Despite a poor degree of purification, specially for a first step, ammonium sulphate fractionation was considered to be potentially useful because of the reduction in sample volume that can be achieved.

4.2.2 Ion-exchange chromatography on DEAE-Sephacel

TOL-BZDH and TOL-BADH bound to DEAE-Sephacel when the extract (Methods 2.4.4b) and the material were equilibrated with 50 mM-potassium phosphate buffer, pH 7.5/2 mM-DTT. Elution was achieved with a linear gradient of 50 - 350 mM-potassium phosphate buffer, pH 7.5 containing 2 mM-DTT. The two activities separated slightly during elution but careful choice of fractions allowed most of both activities to be collected in a single pool. This step was incorporated into the final purification procedure (Methods 2.6.2b; Section 4.3.1).

4.2.3 Affinity Chromatography on Blue Sepharose CL-6B

Preliminary experiments involved changes in ionic strength, pH and the concentration of Mg^{2+} (see Section 3.2.4). This finally led to a procedure for the separation of TOL-BZDH and TOL-BADH.

After chromatography on DEAE-Sephacel (Section 4.2.2) the pool was dialysed (Methods 2.2.4) for a total of 4 h against eight changes of 100 mM-potassium phosphate buffer, pH 7.5/2 mM-DTT. When the dialysed pool was applied to a column of Blue Sepharose, equilibrated with the same buffer, TOL-BZDH bound to the column and TOL-BADH passed through in the wash. Most of the aldehyde dehydrogenase could then be eluted by buffer plus 0.2 mM-NAD^+. The portion that remained bound to the column could be recovered in a large volume by continued elution or more conveniently by a second step to column buffer plus 0.5 mM NAD^+. This step was incorporated into the final purification procedure (Methods 2.6.2c; Section 4.3.2).

4.2.4 Affinity chromatography on Matrex gel red A

When a dialysed DEAE-Sephacel pool (Sections 4.2.2 & 4.2.3) was applied to a column of Matrex gel red A that had been equilibrated with 100 mM-potassium phosphate buffer, pH 7.5/2 mM-DTT, TOL-BADH passed through the column in the wash. The column was capable of binding TOL-BZDH which could be eluted by buffer plus 0.5 mM-NAD^+, however the capacity of the column was low. If the capacity of the column for TOL-BZDH was exceeded, then a portion of the activity passed through the column during washing. This portion of the sample was slightly purified, presumably because the binding capacity for TOL-BZDH was saturated before the binding capacity for certain other proteins in the sample. Although the purification was only slight, several important contaminants were removed; in particular several that are not separated by chromatography on Blue Sepharose (Section 4.2.3). This step was incorporated into the

final purification procedure (Methods 2.6.2c; Section 4.3.2).

4.2.5 Hydrophobic-interaction chromatography on Phenyl Sepharose CL-4B

TOL-BZDH bound to Phenyl Sepharose when the sample and the column were equilibrated with 100 mM-potassium phosphate buffer, pH 7.5/2 mM-DTT. When the buffer concentration was reduced to 5 mM no activity was released. If the buffer was then made 50% (v/v) ethanediol up to 90% of the activity was recovered.

TOL-BADH also bound to Phenyl Sepharose under exactly the same conditions as described above except that the release of activity at 50% (v/v) ethanediol was temperature dependent. At room temperature (~ 20 ^{0}C) most of the activity was recovered quickly in a relatively small volume. At 0 - 4 ^{0}C the activity was released slowly and in a large volume. The elution volume (at 0 - 4 ^{0}C) was greatly reduced by increasing the ethanediol content of the buffer to 65% (v/v).

The column eluate was difficult to handle because of the relatively large volume and the difficulty associated with concentrating samples containing ethanediol. The eluate could be concentrated, and the ethanediol removed, by applying it directly to a small DEAE-Sephacel column. This method was adopted for the final purification procedure (Methods 2.6.2d; Section 4.3.3) and is also described in Section 3.2.6.

4.2.6 Affinity chromatography on Matrex gel green A

TOL-BADH bound to Matrex gel green A when both the sample and the column were equilibrated with 20 mM-potassium phosphate buffer, pH 6.5/5 mM-$MgCl_2$/2 mM-DTT. The enzyme did not bind at higher pH and the capacity of the column was reduced at higher ionic strength. Elution was possible by making the buffer 1.0 mM in NAD^+, but a higher yield and a smaller elution volume were achieved by changing to 50 mM-potassium phosphate buffer, pH 6.5/5 mM-$MgCl_2$/2 mM-DTT/1.0 mM-NAD^+. This

step was incorporated into the final purification procedure (Methods 2.6.2e; Section 4.3.4).

4.2.7 Ion-exchange chromatography on f.p.l.c Mono Q

TOL-BZDH and TOL-BADH bound to the Mono Q column when the samples and the column were equilibrated with 25 mM-potassium phosphate buffer, pH 7.5/ 2 mM-DTT. Elution was achieved by a linear gradient of 25 -350 mM-potassium phosphate buffer, pH 7.5 containing 2 mM-DTT. Both enzymes were eluted in the first half of the gradient.

4.3 Purification of TOL-BZDH and TOL-BADH

A protocol based on the preliminary purification studies (Section 4.2) was developed for TOL-BZDH and TOL-BADH. It is described in Methods 2.6.2 and is summarised in Table 4.1.

4.3.1 Ion-exchange chromatography on DEAE-Sephacel

A typical profile of the chromatography of an extract of *P. putida* MT53 (Methods 2.6.2a) on DEAE-Sephacel is shown in Figure 4.1. Both enzymes were eluted after approximately 200 ml of a 1.0 l linear gradient of 50 - 350 mM-potassium phosphate buffer. TOL-BZDH was purified 6.5-fold with a yield of 81% and TOL-BADH was purified 7.5-fold with a yield of 98%.

4.3.2 Affinity chromatography on Matrex gel red A/Blue Sepharose CL-6B

Chromatography on Matrex gel red/Blue Sepharose is a major purification step for TOL-BZDH. A dialysed DEAE-Sephacel pool (Methods 2.6.2b) was applied to a Matrex

gel red column which was connected in series with a Blue Sepharose column. Both enzymes passed through the Matrex gel red column which acted as a negative purification step. TOL-BADH also washed straight through the Blue Sepharose column and a typical elution profile is shown in Figure 4.2a. TOL-BZDH bound to the Blue Sepharose column which was then disconnected from the Matrex gel red column and washed with buffer. The flow through the Blue Sepharose column was reversed and homogeneous TOL-BZDH was eluted with buffer plus NAD^+. A typical elution profile is shown in Figure 4.2b. TOL-BADH was purified a further 2.4-fold with a yield of 64% and TOL-BZDH was purified a further 23-fold with a yield of 47%.

4.3.3 Hydrophobic-interaction chromatography on Phenyl Sepharose CL-4B

The wash pool, containing TOL-BADH (Methods 2.6.2c), from the Matrex gel red/Blue Sepharose column was applied to a Phenyl Sepharose column. After column washing TOL-BADH was eluted by buffer containing 65% (v/v) ethanediol and was captured on a DEAE-Sephacel column. After it had been washed free of ethanediol the DEAE-Sephacel column was eluted by a step in ionic strength. A typical elution profile for the DEAE-Sephacel column is shown in Figure 4.3. TOL-BADH was purified a further 7-fold with a yield of 75%.

4.3.4 Affinity chromatography on Matrex gel green A

A typical elution profile of the affinity chromatography of a dialysed Phenyl Sepharose elution pool (Methods 2.6.2d) is shown in Figure 4.4. TOL-BADH was eluted by a step in ionic strength and 1.0 mM-NAD^+. TOL-BADH was purified a further 1.6-fold with a yield of 81%.

4.3.5 Storage of the purified enzymes

After the final purification steps TOL-BZDH and TOL-BADH were in potassium phosphate buffers of pH 7.5 and 6.5 respectively. These buffers were optimal (among those tested) for the stability of the enzymes (Section 4.1.4), and only a little activity was lost when they were mixed with an equal volume of ethanediol (Table 4.1). In these mixtures the enzymes could be stored at -20 ^{0}C without freezing.

4.3.6 Overall purification

Details of a purification are given in Table 4.1. No details of the reproducibility of the method are available because it has only been carried out twice in the final format.

4.3.7 Purity

The purification of TOL-BZDH and TOL-BADH was monitored using SDS-PAGE (Methods 2.7.1) with a protein stain (Methods 2.7.4a; Figure 4.5a). The homogeneity of the enzymes was confirmed using non-denaturing PAGE (Methods 2.7.3) with an activity stain (Methods 2.4.7c; Figure 4.5b,c) and also with a protein stain (Methods 2.7.4a; Figure 4.5d).

Purified BZDH II and BADH (Methods 2.6.3), which are also shown on the gels in Figure 4.5, both produced two major bands on non-denaturing PAGE. This effect was probably due to minor charge heterogeneities since each of the major bands in each of the enzyme preparations have the same amino terminal sequences as the respective purified enzymes (Section 5.1.4). BZDH II also has a very minor third band on non-denaturing gels after protein staining (Figure 4.5d, Track B). This band does not appear after activity staining (Figure 4.5b, Track D) and probably represents either inactive enzyme or a minor contaminant.

Table 4.1 Purification of TOL-BZDH and TOL-BADH from *P. putida* MT53

Details of the purification are given in Methods 2.6.2. Enzyme activities and protein concentrations were measured as described in Methods 2.5.1a and 2.2.5, respectively.

ND, Not Determined

TOL-BZDH:

	Volume (ml)	Total activity (units)	Total protein (mg)	Specific activity (units/mg of protein)	Yield (%)	Purification (fold)
Extract	85	1262	2601	0.49	100	1
DEAE-Sephacel	68	1074	340	3.16	85	6.5
Dialysis	68	1010	ND	ND	80	ND
Matrex gel red/Blue	115	506	6.9	73.7	40	150
Concentration	4.1	506	6.9	73.7	40	150
Ethanediol added	8.2	415	6.9	60.2	33	123

TOL-BADH:

	Volume (ml)	Total activity (units)	Total protein (mg)	Specific activity (units/mg of protein)	Yield (%)	Purification (fold)
Extract	85	5372	2601	2.07	100	1
DEAE-Sephacel	68	5265	340	15.5	98	7.5
Dialysis	68	4910	ND	ND	91	ND
Matrex gel red/Blue	70	3367	92	36.8	63	17.8
Phenyl Sepharose	7.3	2520	9.8	257	47	133
Dialysis	12.2	1863	ND	ND	35	ND
Matrex gel green	103	2050	ND	ND	38	ND
Concentration	1.5	1762	4.0	446	33	215
Ethanediol added	3.0	1503	4.0	376	28	181

Figure 4.1 Chromatography of an extract of *P. putida* MT53 on DEAE-Sephacel

An extract (85 ml, Methods 2.6.2a) of *P. putida* MT53 was applied to column of DEAE-Sephacel (5.0 x 4.0 cm). After column washing, elution was carried out with a linear gradient of potassium phosphate buffer (Methods 2.6.2b). The flow rate was 26 ml/h and fractions of 5.7 ml were collected.

TOL-BZDH, ⊡; TOL-BADH, □; A_{280}, △; specific conductivity, ◇.

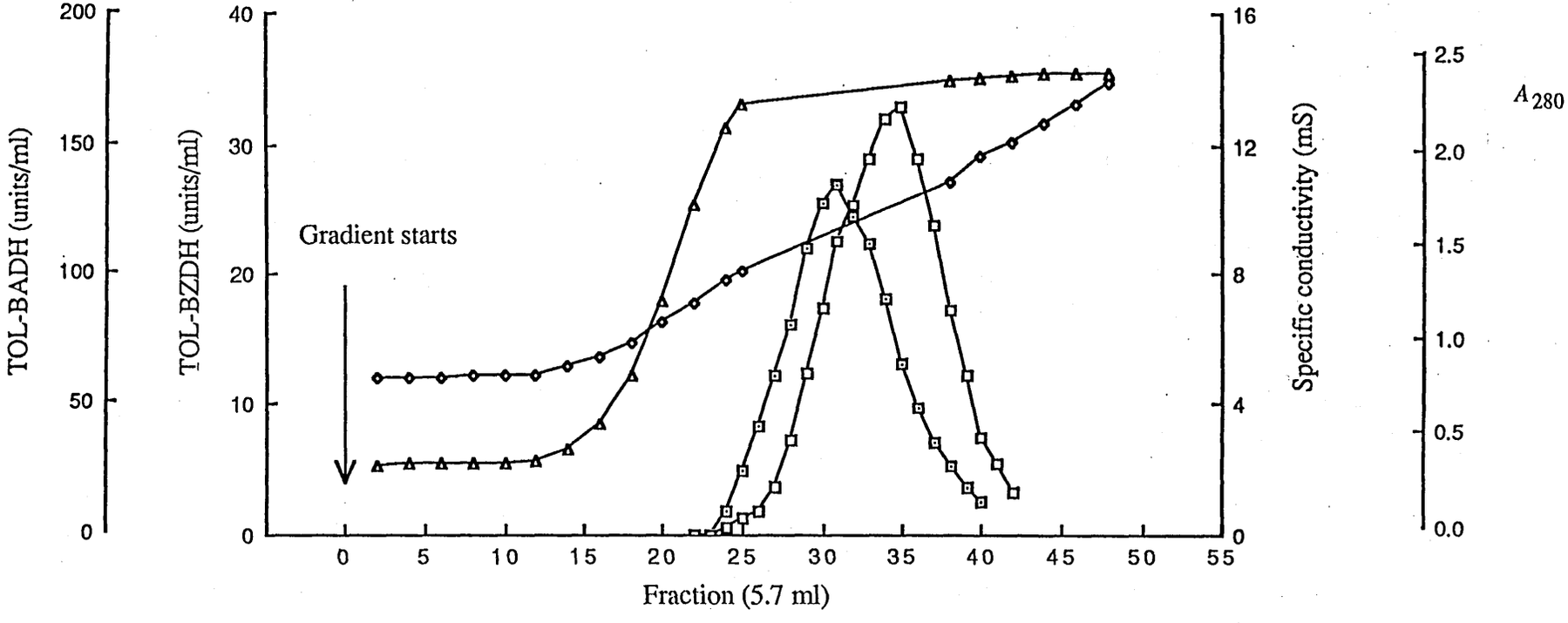

TOL-BADH (units/ml)
200
150
100
50
0
TOL-BZDH (units/ml)
40
30
20
10
0
Gradient starts
16
12
8
4
0
Specific conductivity (mS)
2.5
2.0
1.5
1.0
0.5
0.0
A_{280}
0 5 10 15 20 25 30 35 40 45 50 55
Fraction (5.7 ml)

Figure 4.2 Chromatography of a dialysed DEAE-Sephacel pool on Matrex gel red A/Blue Sepharose CL-6B

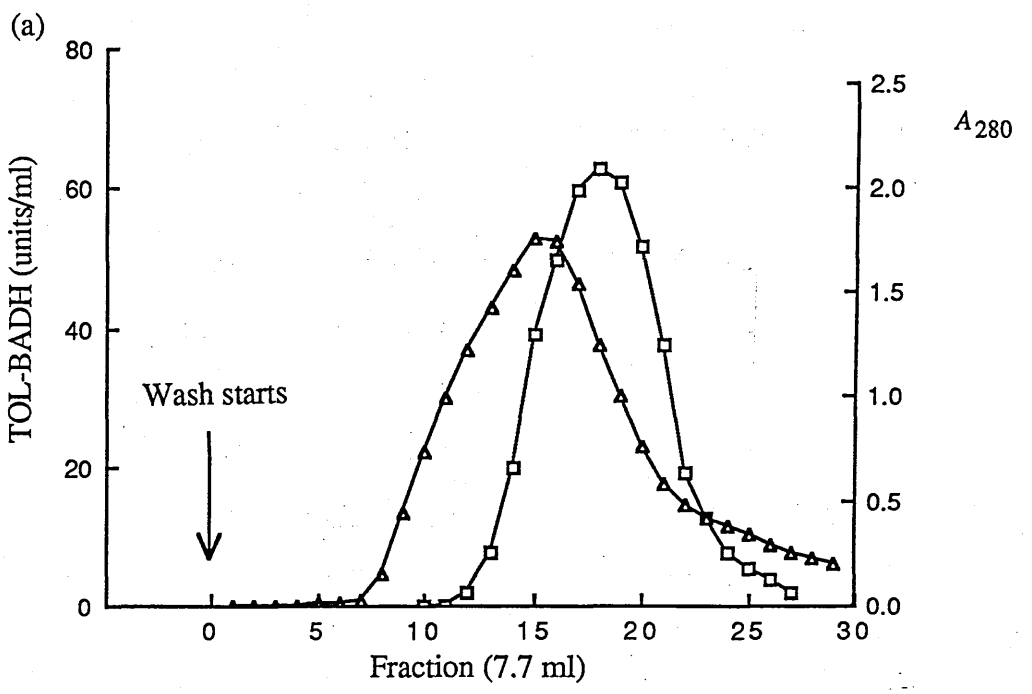
A dialysed DEAE-Sephacel pool (68 ml, Methods 2.6.2b) was applied to a Matrex gel red column (2.6 x 8.0 cm) connected in series with a Blue Sepharose column (2.6 x 11.5 cm).

(a) TOL-BADH washed straight through both columns (Methods 2.6.2c). The flow rate was 23 ml/h and fractions of 7.7 ml were collected.

TOL-BADH activity, □; A_{280}, △.

(b) TOL-BZDH bound to the Blue Sepharose column which was disconnected from the Matrex gel red column and washed with buffer. The flow through the Blue Sepharose column was reversed and homogeneous TOL-BZDH was eluted first with buffer plus 0.2 mM-NAD^+ and then with buffer plus 0.5 mM-NAD^+ (Methods 2.6.2c). The flow rate was 44 ml/h and fractions of 9.6 ml were collected.

TOL-BZDH activity, ⊡; A_{280}, △.


(a)
TOL-BADH (units/ml)
80
60
40
20
0
2.5
2.0
1.5
1.0
0.5
0.0
A_{280}
Wash starts
0
5
10
15
20
25
30
Fraction (7.7 ml)

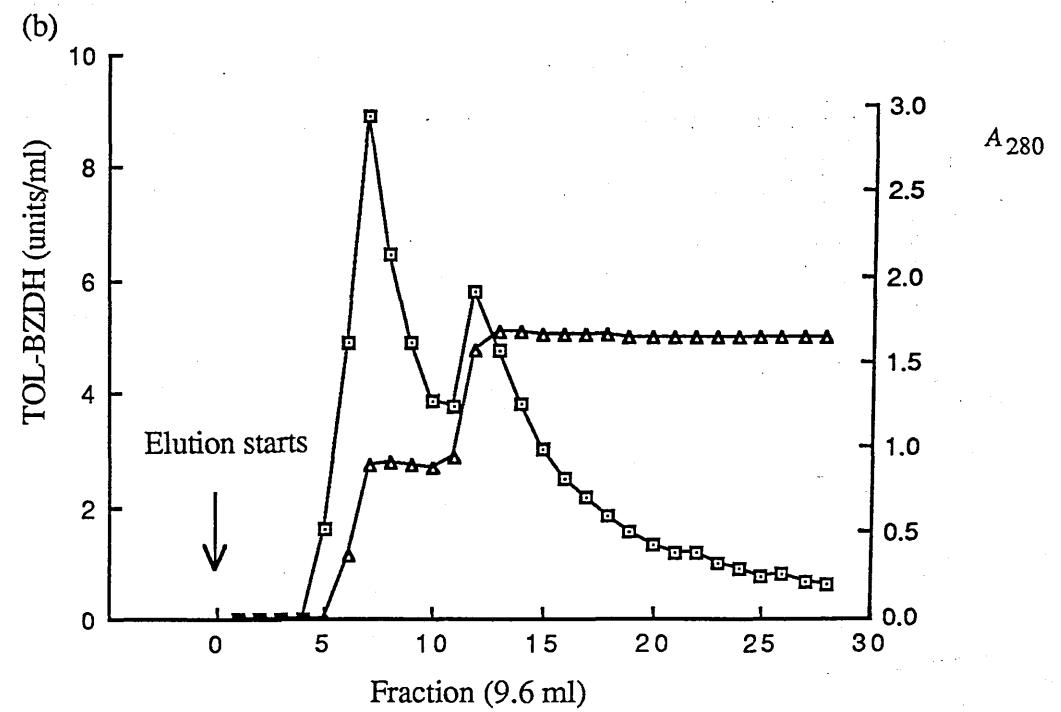

(b)
TOL-BZDH (units/ml)
10
8
6
4
2
0
3.0
2.5
2.0
1.5
1.0
0.5
0.0
A_{280}
Elution starts
0
5
10
15
20
25
30
Fraction (9.6 ml)

Figure 4.3 Chromatography of a Matrex gel red/Blue Sepharose wash pool on Phenyl Sepharose/DEAE-Sephacel

A wash pool (69 ml, Methods 2.6.2c) from a Matrex gel red/Blue Sepharose column, containing TOL-BADH activity, was applied directly to a column of Phenyl Sepharose (2.6 x 11.0 cm). After washing with buffer plus 50% (v/v) ethanediol the column was eluted with buffer plus 65% (v/v) ethanediol through a second DEAE-Sephacel column (1.6 x 3.5 cm) which captured TOL-BADH (Methods 2.6.2d). The DEAE-Sephacel column was washed free of ethanediol and eluted with a step in ionic strength (Methods 2.6.2d). The flow rate was 9.7 ml/h and fractions of 2.4 ml were collected.

TOL-BADH activity, □.

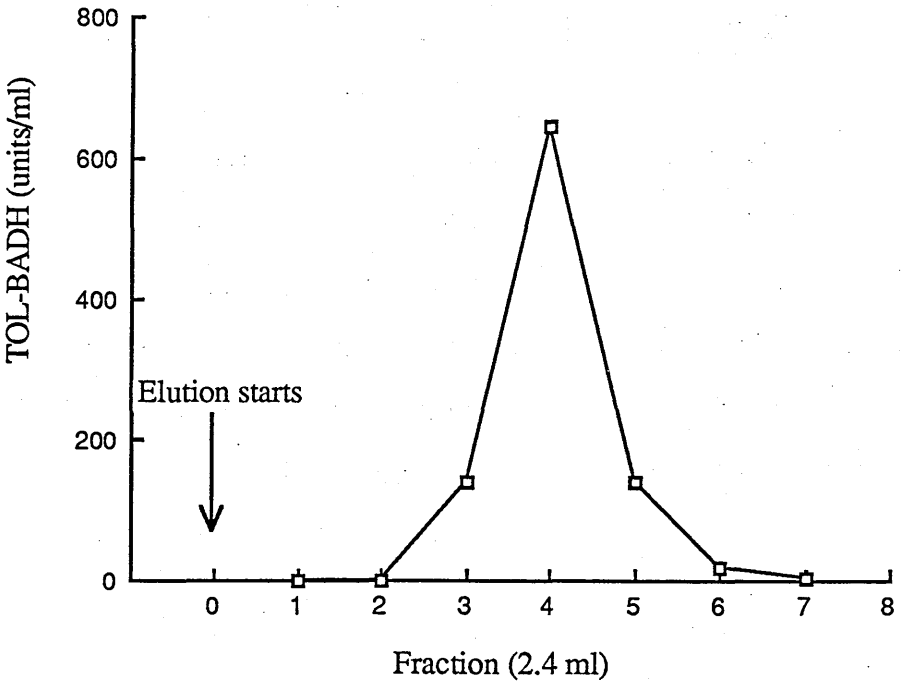

TOL-BADH (units/ml)
800
600
400
200
0
Elution starts
0
1
2
3
4
5
6
7
8
Fraction (2.4 ml)

Figure 4.4 Chromatography of a dialysed Phenyl Sepharose pool on Matrex gel green A

A dialysed Phenyl Sepharose pool (12.2 ml, Methods 2.6.2d) was applied to a column of Matrex gel green (4.5 x 5.0 cm). After column washing, TOL-BADH was eluted with buffer plus 1.0 mM-NAD^+ (Methods 2.6.2e). The flow rate was 32 ml/h and fractions of 7.9 ml were collected.

TOL-BADH activity, □.

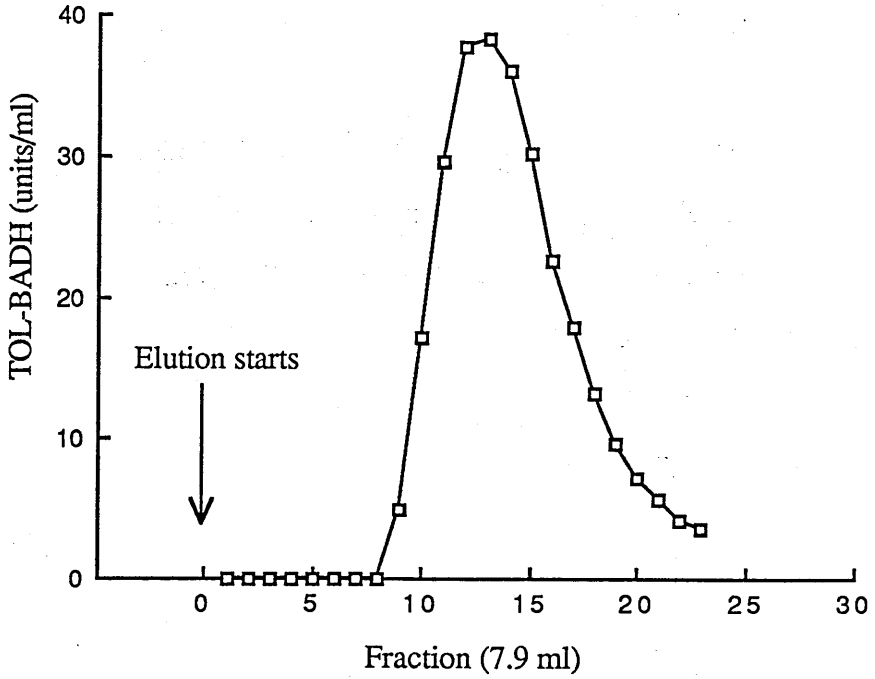

TOL-BADH (units/ml)
40
30
20
10
0
Elution starts
0
5
10
15
20
25
30
Fraction (7.9 ml)

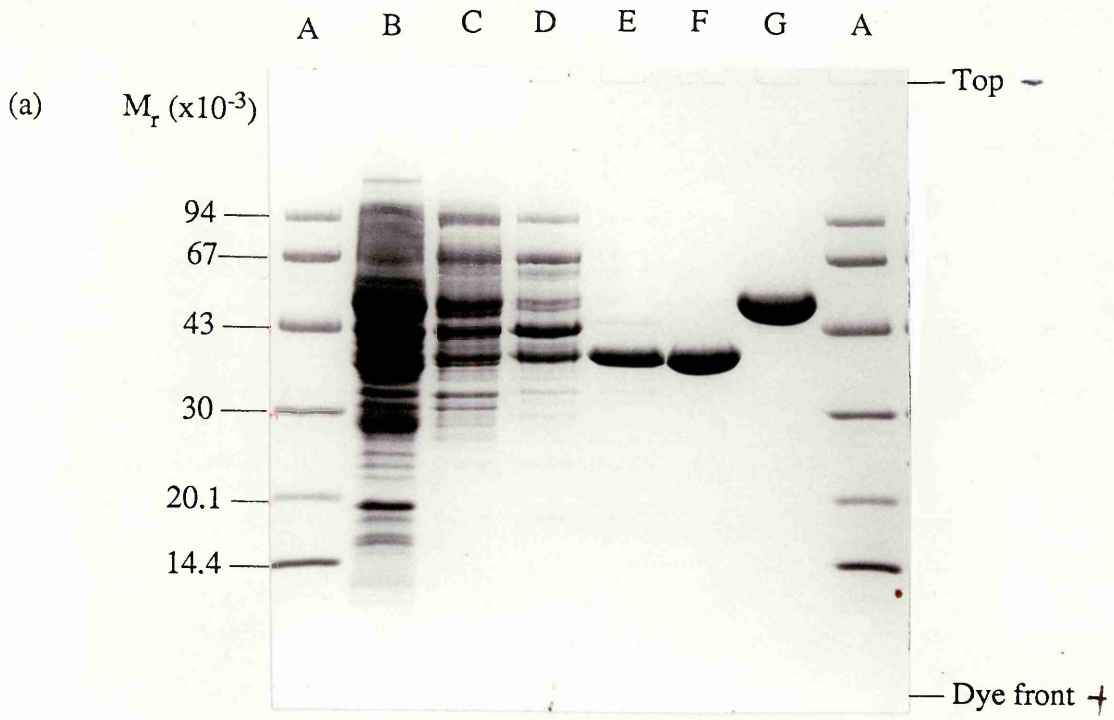
Figure 4.5 The purification of TOL-BZDH and TOL-BADH as monitored by PAGE

(a) A denaturing 12.5% (w/v) polyacrylamide slab gel containing 0.1% SDS was stained for protein. Tracks: A, M_r markers (approx. 1 μg of each); B, extract (122 μg of protein); C, DEAE-Sephacel pool (50 μg of protein); D, Matrex gel red/Blue Sepharose wash pool containing TOL-BADH (32 μg of protein); E, Phenyl Sepharose pool containing TOL-BADH (16 μg of protein); F, Matrex gel green pool containing homogeneous TOL-BADH (17 μg of protein); G, Blue Sepharose pool containing homogeneous TOL-BZDH (26 μg of protein).

(b) A non-denaturing 7.5% (w/v) polyacrylamide slab gel was stained for activity; showing TOL-BZDH together with BZDH I and BZDH II. Tracks: A, *P. putida* MT53 extract (900 μg of protein); B, pure TOL-BZDH (28 μg of protein); C, pure BZDH I (3.5 μg of protein); D, pure BZDH II (7.5 μg of protein).

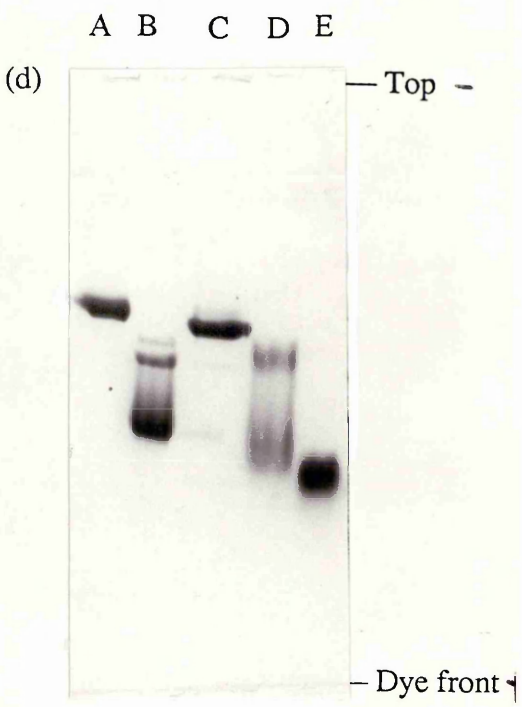
(c) A non-denaturing 7.5% (w/v) polyacrylamide slab gel was stained for activity; showing TOL-BADH together with BADH. Tracks: A, *P. putida* MT53 extract (900μg of protein); B, pure TOL-BADH (4.5 μg of protein); C, pure BADH (8.0 μg of protein).

(d) A non-denaturing 7.5% (w/v) polyacrylamide slab gel was stained for protein; showing TOL-BZDH and TOL-BADH together with BZDH I, BZDH II and BADH. Tracks: A, pure BZDH I (17.5 μg of protein); B, pure BZDH II (44.5 μg of protein); C, pure TOL-BZDH (28 μg of protein); D, pure BADH (16.5 μg of protein); E, pure TOL-BADH (27 μg of protein).


A B C D E F G A
(a)
M_r (x10^{-3})
94
67
43
30
20.1
14.4
Top
Dye front


A B C D
(b)


A B C
(c)


A B C D E
(d)
Top
Dye front

In Figure 4.5b, Track B (activity stain of purified TOL-BZDH) there is a faint smudge above the major band. This is carry-over from Track A which was heavily loaded because TOL-BZDH does not stain well in cell-extracts. This smudge was not seen in any of the other similar gels that have been run, nor is it visible after protein staining (Figure 4.5d, Track C).

Overall, the gels shown in Figure 4.5 (and others that are not shown) demonstrate that TOL-BZDH and TOL-BADH are homogeneous as judged by denaturing and non-denaturing PAGE with either protein staining or activity staining. The relative purity of each of the five alcohol and aldehyde dehydrogenases characterised in this project is discussed further in Section 6.1.1.

4.3.8 Discussion

TOL-BZDH and TOL-BADH were purified from the same batch of *P. putida* MT53 in a total of four steps by ion-exchange, hydrophobic-interaction and affinity chromatography (Methods 2.6.2). TOL-BZDH was purified 150-fold with a yield of 40% (~ 7 mg of protein) and TOL-BADH was purified 215-fold with a yield of 33% (~ 4 mg of protein). The whole procedure was carried out in four days.

The most important purification step for TOL-BZDH was on Matrex gel red A /Blue Sepharose (Sections 4.2.3, 4.2.4 & 4.3.2). The Matrex gel red column acted as a negative purification step, since once the capacity for TOL-BZDH had been saturated the remainder of the enzyme passed through. The portion of the activity that passed through the column was slightly purified, presumably because the binding capacity for TOL-BZDH became saturated before the binding capacity for other proteins. This step is potentially irreproducible because, for success, the capacity of the column must be small enough to let some TOL-BZDH pass through but large enough to remove the contaminants. There are a number of factors that can affect the capacity of the column; for example, the volume of material in the column may vary, the capacity may depend on

the age of the material and the frequency of regeneration and finally, different batches of the material vary with respect to the concentration of immobilised ligand. If the capacity of the column was too large, and all of the TOL-BZDH was bound, the activity could be recovered by eluting the column with 0.5 mM-NAD^+ (Section 4.2.4). If the capacity of the column was too small, and there was unsuccessful purification, the sample could be chromatographed on a miniature column of Matrex gel red A after it had been recovered from the Blue Sepharose column (Section 4.2.3).

CHAPTER 5

Physical, Chemical and Kinetic Characterisation of
Benzaldehyde Dehydrogenase I, TOL-Benzaldehyde Dehydrogenase and
TOL-Benzyl Alcohol Dehydrogenase

5.1 Physical and chemical characterisation of the purified enzymes

Some of the physical and chemical characteristics of BZDH I, TOL-BZDH and TOL-BADH were determined, largely to allow comparison of the enzymes with each other and with BZDH II and BADH. In some cases the characteristics of BZDH II and BADH were also determined, either to serve as controls or because characterisation had not previously been carried out.

The primary aim of this project was to purify and characterise BZDH I (Section 1.6) and therefore only the most important and relevant properties of TOL-BZDH and TOL-BADH have been measured.

5.1.1 M_r values under denaturing and non-denaturing conditions

The subunit M_r values of BZDH I, TOL-BZDH and TOL-BADH were determined under denaturing conditions using SDS-PAGE (Methods 2.7.6). A typical plot of the mobility of the calibration proteins against ln (M_r) is shown in Figure 5.1. The subunit M_r values were: BZDH I, 56 000 (SD = 363, n = 13); TOL-BZDH, 56 300 (SD = 600, n = 13) and TOL-BADH, 43 000 (SD = 680, n = 15).

The apparent native M_r values for BZDH I, TOL-BZDH, and TOL-BADH were determined by gel filtration through Sephacryl S300 and also through f.p.l.c. Superose 6 or Superose 12 columns. Determinations were made with the purified enzymes and in certain cases also with extracts. Standard curves of K_{av} against ln (M_r) were constructed for the calibration proteins (Figure 5.2) and the apparent native M_r values for the enzymes were calculated. The apparent native M_r value determinations for all of the enzymes are summarised in Table 5.1.

5.1.2 Chemical cross-linking of enzyme subunits

The quaternary structures of BZDH I, BZDH II, BADH, TOL-BZDH and TOL-BADH were investigated by cross-linking with dimethylsuberimidate (Methods 2.9.1). After cross-linking, the enzymes were examined using a phosphate-buffered SDS-PAGE system (Methods 2. 7.2). Four major bands appeared in all five enzyme preparations (Figure 5.3), corresponding in M_r value to monomer, dimer, trimer and tetramer (Figure 5.4).

Aldolase was included on each of the gels as a control because it is a well characterised protein and is known to form tetramers (Coggins, 1978; Boehringer, 1987). The mobilities of each of the four major bands were used to calibrate each of the cross-linking gels (Figure 5.3) and the M_r values for each of the major bands (corresponding to monomer, dimer, trimer and tetramer) in each of the enzyme preparations were calculated (Table 5.2).

5.1.3 Isoelectric point of BZDH I

The isoelectric point of BZDH I was determined in duplicate by chromatofocusing on an f.p.l.c. Mono P column (Methods 2.6.4). The isoelectric point was 5.55 (5.59 & 5.50) and the homogeneity of the purified enzyme was confirmed by the observation of a single peak in the A_{280} of the column eluent.

Figure 5.1 Typical plot of migration distance against ln (M_r) for the SDS-PAGE calibration proteins and for TOL-BADH

The distances migrated down the gel by the calibration proteins (approximately 1 μg of each) were plotted against ln (M_r) and the line of best fit was calculated using the Cricket Graph V 1.1 program on an Apple Macintosh computer. The calibration proteins were (M_r values in parentheses), **1** phosphorylase *b* (94 000), **2** bovine serum albumin (67 000), **3** ovalbumin (43 000), **4** carbonic anhydrase (30 000), **5** trypsin inhibitor (20 100) and **6** α-lactalbumin (14 400). The relative position of TOL-BADH is also shown (see Section 5.5.1).

Calibration proteins, □; TOL-BADH, ◇.

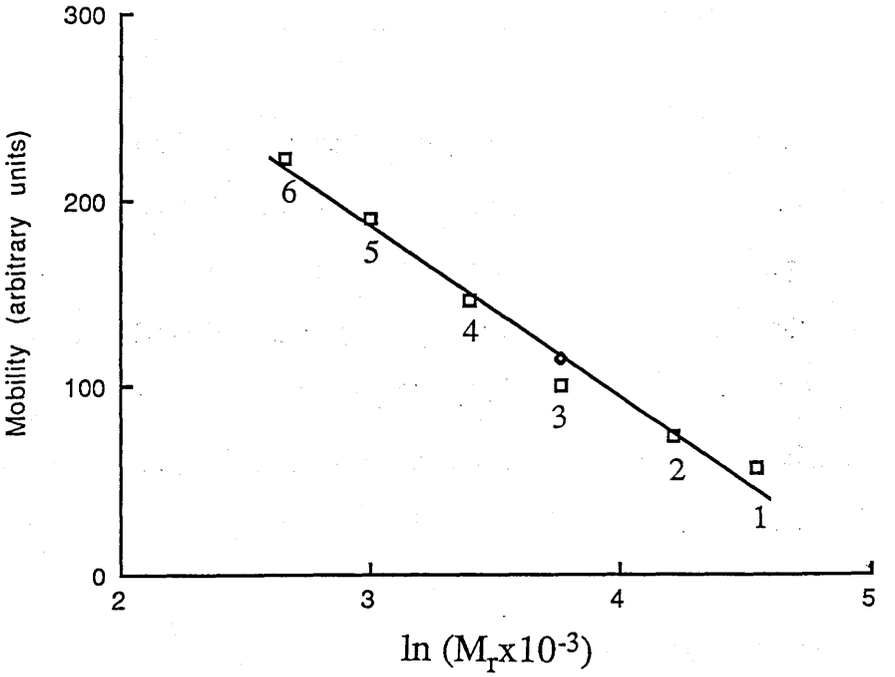

300
200
100
0
Mobility (arbitrary units)
6
5
4
3
2
1
2
3
4
5
ln (M_rx10^{-3})

Figure 5.2 Standard curves of K_{av} against ln (M_r) for the gel filtration calibration proteins

The apparent native M_r values for BZDH I, TOL-BZDH and TOL-BADH were determined by gel filtration through Sephacryl S300 and f.p.l.c. Superose 6 or Superose 12. The calibration proteins were (M_r values in parentheses), **1** thyroglobulin (669 000), **2** ferritin (440 000), **3** catalase (232 000), **4** aldolase (158 000), **5** albumin (68 000), **6** ovalbumin (45 000), **7** chymotrypsinogen A (25 000) and **8** cytochrome c (12 500).

(a) Proteins (□) were applied, either alone or in pairs, to a column of Sephacryl S300 (1.6 x 90 cm) in a volume of 2 ml. Fractions of 2 ml were collected and the proteins were located by measuring the A_{280} of the fractions. The buffer was 100 mM-potassium phosphate, pH 7.5/2 mM-DTT and the flow rate was 20 ml/h.

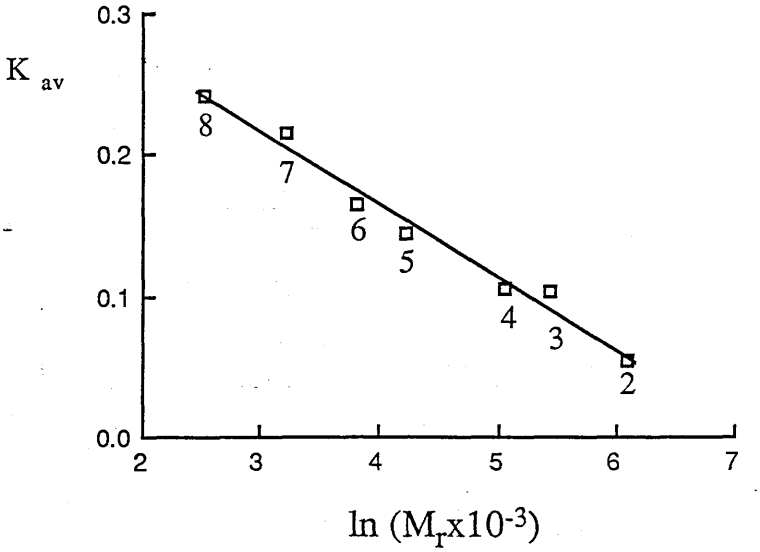
(b) Proteins (□) were applied to an f.p.l.c. Superose 6 column (1.0 x 30 cm) in a volume of 200 µl. The buffer was 100 mM-potassium phosphate, pH 7.5/2 mM-DTT and the flow rate was 0.3 ml/min. The elution volume of the proteins was found by continuously monitoring the A_{280} of the eluate (Methods 2.2.5).

(c) The Superose 12 column was run under exactly the same conditions as described in (b).

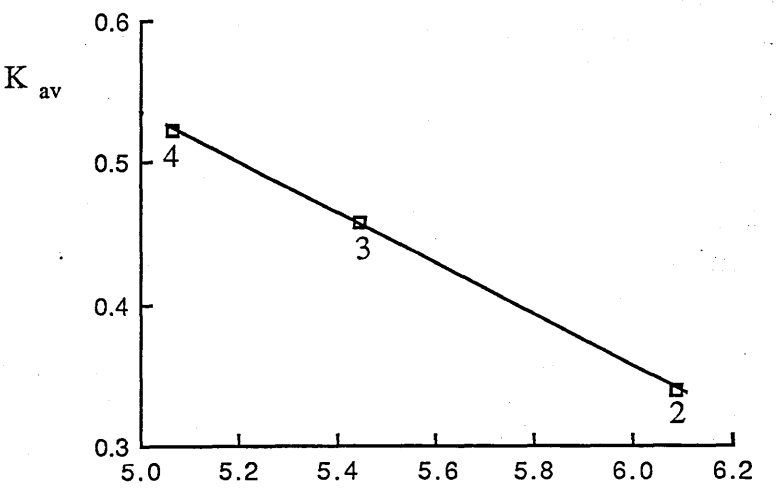
The total volume of the Sephacryl S300 column was assumed to be the elution volume of phenol red and the total volumes of the Superose columns were calculated from the dimensions of the columns. The void volume of the Sephacryl S300 column was taken as the elution volume of blue dextran and the void volumes of the Superose columns were assumed to be the elution volumes of thyroglobulin. The elution volumes of the proteins were used to calculate K_{av} values for each of the proteins which were plotted against ln (M_r). The equation of the line of best fit was calculated using the Cricket Graph V 1.1 program on an Apple Macintosh computer and from this the M_r values for the enzymes were calculated.

$K_{av} = (V_e-V_o)/(V_t-V_o)$, where V_e = elution volume of the protein,
V_o = void volume of the column and V_t = total volume of the column.

(a)


K av
0.3
0.2
0.1
0.0
8
7
6
5
4
3
2
2
3
4
5
6
7
ln (Mr x10-3)

(b)


K av
0.6
0.5
0.4
0.3
4
3
2
5.0
5.2
5.4
5.6
5.8
6.0
6.2
ln (Mr x10-3)

(c)

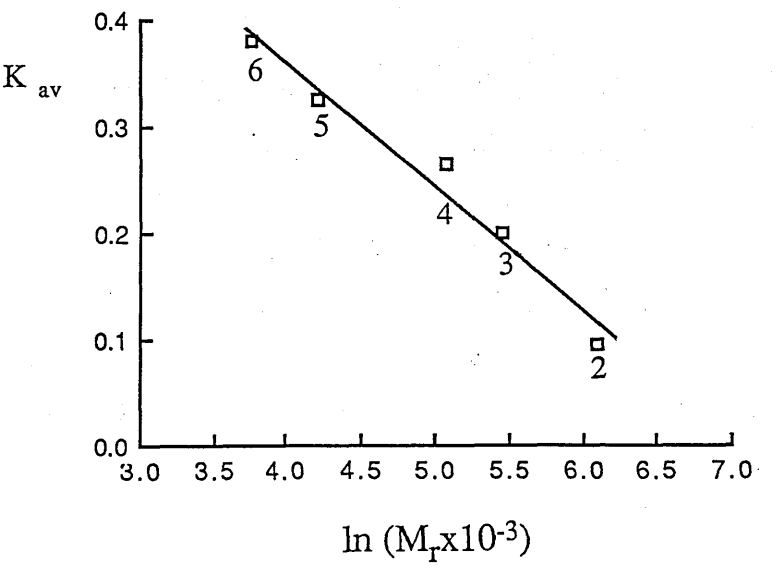

K av
0.4
0.3
0.2
0.1
0.0
6
5
4
3
2
3.0
3.5
4.0
4.5
5.0
5.5
6.0
6.5
7.0
ln (Mr x10-3)

Table 5.1 Summary of the apparent native M_r value determinations

The apparent native M_r values of the enzymes were determined by gel filtration as described in Figure 5.2. Most of the values are the means of duplicate determinations and the values given in parentheses are the individual values.

ND, not determined

Protein	M_r value		
	Superose 6	Superose 12	Sephacryl S300
BZDH I (Pure)	141 000 (both values identical)	ND	237 000 (234 000 239 000)
BZDH I (Extract)	ND	ND	172 000
TOL-BZDH (Pure)	ND	122 000 (115 000, 129 000)	ND
TOL-BZDH (Extract)	ND	ND	176 000 (174 000 178 000)
TOL-BADH (Pure)	ND	82 000 (81 000, 82 000)	114 000
TOL-BADH (Extract)	ND	ND	111 000 (108 000 114 000)

Figure 5.3 Cross-linking of enzyme subunits with dimethylsuberimidate

Enzyme subunits were cross-linked with dimethylsuberimidate as described in Methods 2.9.1, except that in Part d, Track C the enzyme was only treated once with dimethylsuberimidate and not twice like all of the other samples. The quaternary structures were examined using a phosphate buffered SDS-PAGE system and gels were silver stained for protein (Methods 2.7.4b).

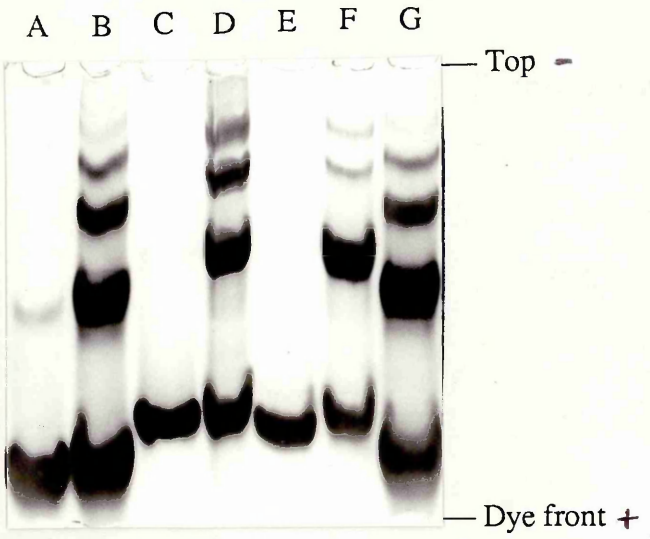
(a) Tracks: A, aldolase monomer (2.2 μg of protein); B, cross-linked aldolase (9.4 μg of protein); C, BZDH I monomer (2.2 μg of protein); D, cross-linked BZDH I (10 μg of protein); E, BZDH II monomer (8.2 μg of protein); F, cross-linked BZDH II (25 μg of protein); G, cross-linked aldolase (9.4 μg of protein).

(b) Tracks: A, cross-linked BADH (5, 25, 15 and 7.5 μg of protein from left to right); B, cross-linked aldolase (7.5, 10, 15 and 5 μg of protein from left to right).

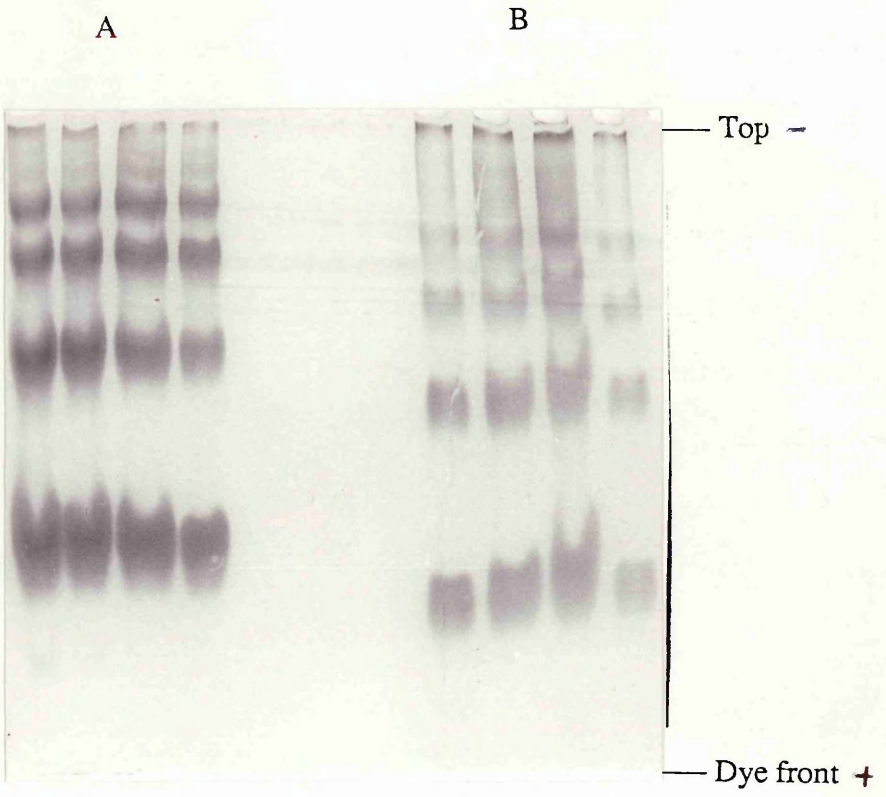
(c) Tracks: A, TOL-BZDH monomer (5 and 12.5 μg of protein from left to right); B, cross-linked TOL-BZDH (12.5 and 5 μg of protein from left to right). The bands corresponding to trimer and tetramer were each split into several minor bands. The centre of each group of minor bands is indicated and this was assumed to be the position of the band when calculating apparent M_r values (Figure 5.4; Table 5.2).

(d) Tracks: A, cross-linked TOL-BZDH (12.5 μg of protein); B, TOL-BADH monomer 1.75, 5 and 12.5 μg of protein from left to right); C, cross-linked TOL-BADH (2.5, 10 and 25 μg of protein from left to right); D, cross-linked TOL-BADH (2.5, 10 and 25 μg of protein from left to right).

(a)

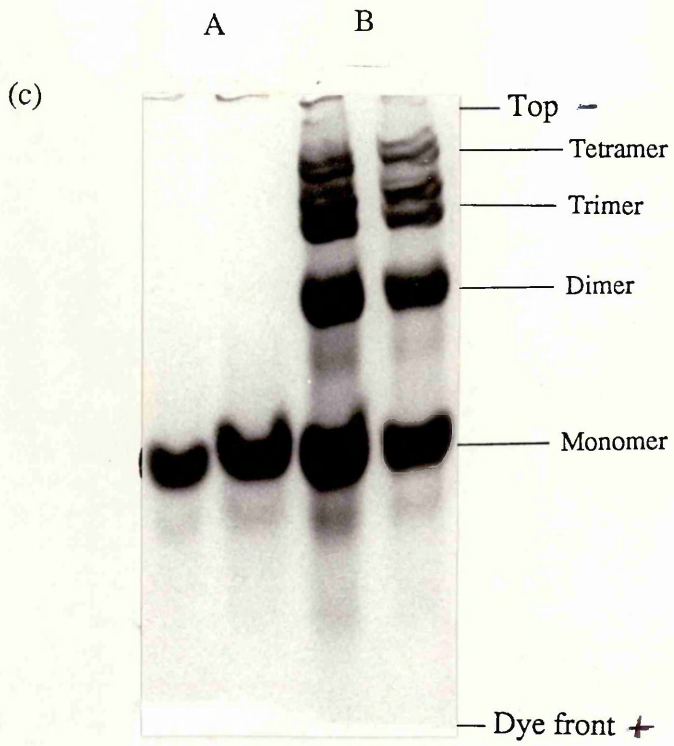


(b)



Continued:

Figure 5.3 Continued:



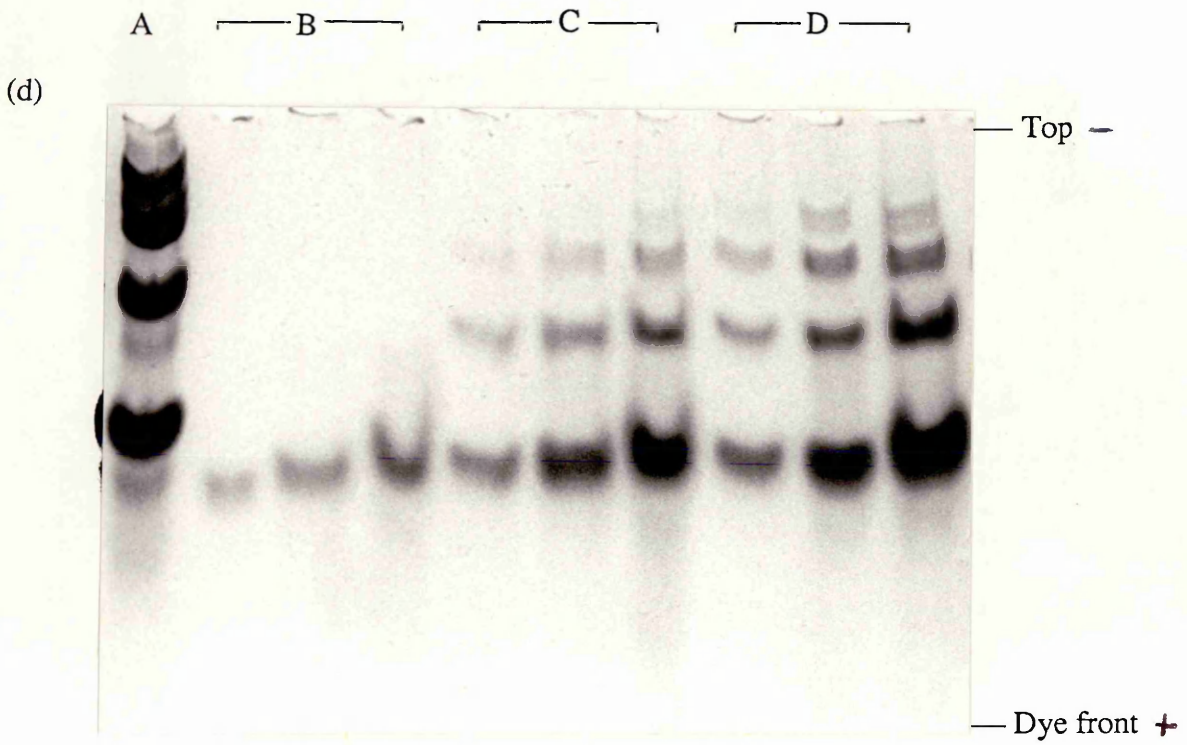


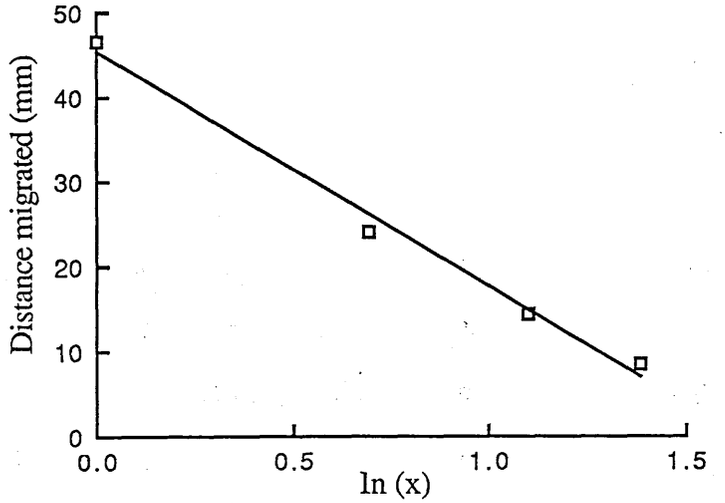
Figure 5.4 Subunit structure of BZDH I, BZDH II, TOL-BZDH, BADH and TOL-BADH

After cross-linking of the enzyme subunits by dimethylsuberimidate (Methods 2.9.1) the proteins were examined using a phosphate-buffered gel system (Methods 2.7.2) and some typical examples are shown in Figure 5.3.

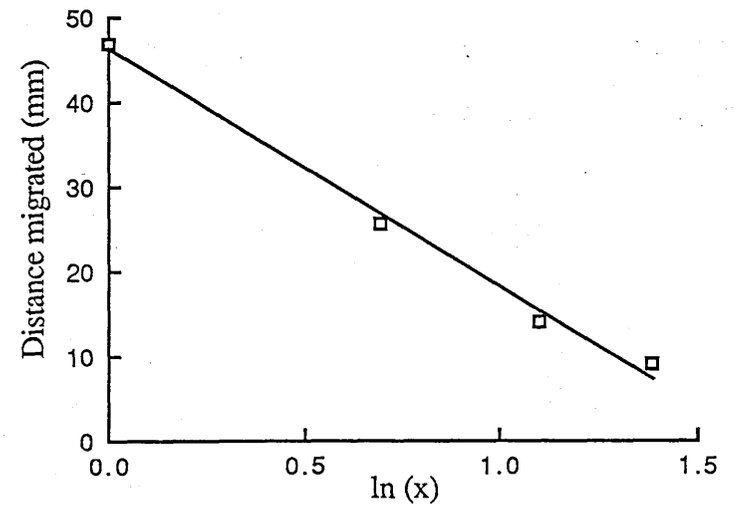
Plots were constructed of the distance migrated down the gel by each band (□), against ln (1.0), ln (2.0), ln (3.0) and ln (4.0). In each of the graphs these values are plotted on the axis labeled ln (x). If the four major bands observed in each of the enzyme preparations correspond to monomers, dimers, trimers and tetramers then the plots would be expected to be straight lines.

(a), BZDH I; (b), BZDH II; (c), TOL-BZDH; (d), BADH; (e), TOL-BADH; (f), aldolase.

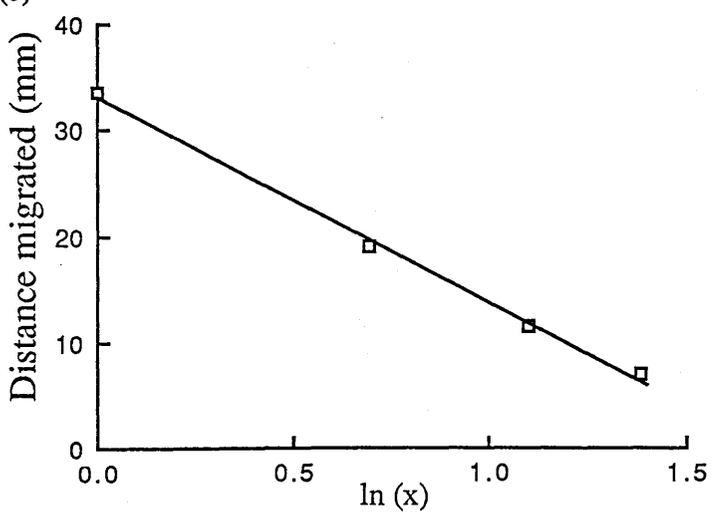
(a)



(b)

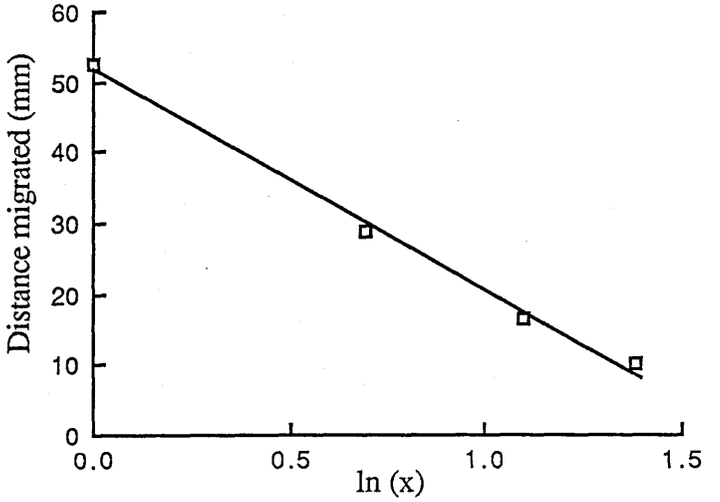


(c)

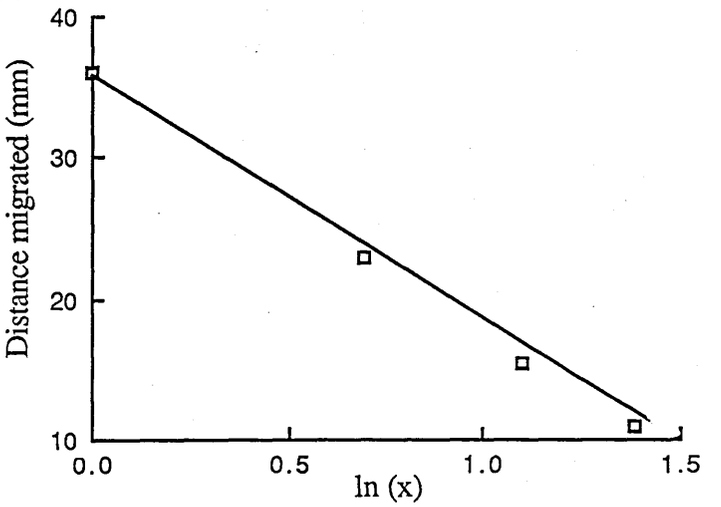


Continued:

(d)


Distance migrated (mm)
60
50
40
30
20
10
0
0.0
0.5
1.0
1.5
ln (x)

(e)


Distance migrated (mm)
40
30
20
10
0.0
0.5
1.0
1.5
ln (x)

(f)

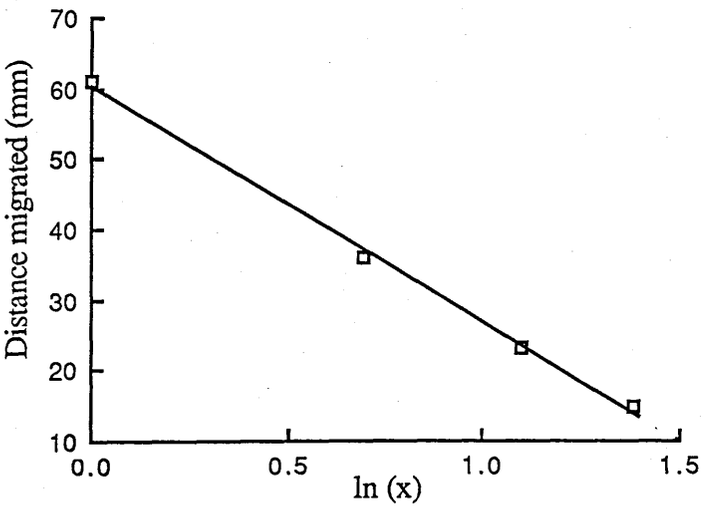

Distance migrated (mm)
70
60
50
40
30
20
10
0.0
0.5
1.0
1.5
ln (x)

Table 5.2 M_r values of each of the major bands in the cross-linked enzyme preparations

The cross-linking gels (Figure 5.3) were calibrated using the mobility of each of the major bands in the cross-linked aldolase preparation. The subunit M_r value of aldolase is 42 000 (Boehringer, 1987). The calibration curves were constructed by plotting the mobilities of each of the major bands in the aldolase preparation (Figure 5.4) against the natural logarithm of the predicted M_r values of aldolase monomers, dimers, trimers and tetramers. It was impractical to use R_F values instead of mobilities because the gels were electrophoresed for a considerable time after the tracker dye had reached the bottom of the gel (Methods 2.7.2). R_F values could have been calculated with respect to the mobility of the monomer but this was not done because the calibration curve would then only have consisted of three points.

	BZDH I	BZDH II	TOL-BZDH	BADH	TOL-BADH
Monomer	52 000	50 000	55 000	35 000	36 000
Dimer	111 000	105 000	105 000	75 000	74 000
Trimer	157 000	157 000	149 000	112 000	113 000
Tetramer	193 000	187 000	190 000	144 000	141 000

5.1.4 Amino terminal sequence analysis

The amino terminal sequences of purified BZDH I, BZDH II, TOL-BZDH, BADH and TOL-BADH (Methods 2.6) were determined (Table 5.3) after the proteins had been reduced and carboxymethylated (Methods 2.9.2).

Purified BZDH II and BADH each showed two major bands after non-denaturing PAGE (Sections 3.3.6 & 4.3.7; Figures 3.4 & 4.5). Both enzymes were pure as judged by several criteria (MacKintosh & Fewson, 1988a,b) and the most likely explanation for the double bands produced by BZDH II and BADH on non-denaturing gels is that there are two populations of molecules which each have different net charges. Other workers have ascribed similar effects to the deamidation of asparagine and glutamine residues (Julia *et al.*, 1988) or to differences in the oxidation states of thiol groups (Jornvall, 1973). In addition to the two major bands, BZDH II preparations also had a very minor third band that was seen after protein staining but not after activity staining (Section 4.3.7; Figure 4.5d). The relative abundance of protein and activity in each of the major bands was determined using a laser densitometer (Methods 2.7.5) to scan non-denaturing gels (Methods 2.7.3) that had been stained either for protein (Methods 2.7.4a) or for enzyme activity (Methods 2.7.4c). The upper band of BZDH II had 9% of the protein and 26% of the activity while the lower band had 91% of the protein and only 74% of the activity. The upper band of BADH had 35% of the protein and 56% of the activity while the lower band had 65% of the protein and only 44% of the activity. The amino terminal sequence of each of the major bands in each of the enzyme preparations was determined (Table 5.4) after the proteins had been electroeluted from non-denaturing gels (Methods 2.11.1). These sequences, together with the sequences of the purified proteins, were used in the compilation of the consensus sequences presented in Table 5.3. There were no discrepancies between the sequences of the proteins from either of the pairs of bands and the corresponding purified proteins at positions where amino acid residues were identified. Sequence data for BZDH II extended to residue 22 and residue 44 of the

upper and lower bands, respectively. The two bands from BADH were sequenced to residue 22 and residue 30. Unfortunately, the identities of the two samples became confused and there is no record of which was the upper or the lower band; however, since the sequences were identical this did not create a problem.

5.1.5 Amino acid compositions and the number of amino acid residues per enzyme subunit

The amino acid compositions of BZDH I, BZDH II, TOL-BZDH, BADH and TOL-BADH were determined (Methods 2.10). The amino acid composition of sperm whale apomyoglobin (a protein of known amino acid sequence) was also determined as a control (Table 5.5).

The amino acid composition data were used to calculate the average M_r value of an amino acid in each of the enzymes. This together with the subunit M_r values of the enzymes (Section 5.1.1 and MacKintosh & Fewson, 1988a) was used to estimate the number of amino acid residues in each subunit of each of the enzymes. The numbers of amino acid residues in each subunit of BZDH I, BZDH II, TOL-BZDH, BADH and TOL-BADH were 525, 525, 538, 381 and 412 respectively.

The amino acid compositions were compared in a pairwise fashion by the methods of Marchalonis & Weltman (1971) (Table 5.6) and Cornish-Bowden (1979) (Table 5.7). Comparisons were made among the aldehyde dehydrogenases and between the alcohol dehydrogenases. Comparisons were not made between the aldehyde and the alcohol dehydrogenases because the differences between the numbers of amino acid residues in each type of subunit are too great for the methods to give accurate results (Cornish-Bowden, 1979). Both methods of comparison involve the calculation of an index that is based on the sum of the differences squared between each type of amino acid in a pair of proteins. The method of Marchalonis & Weltman (1971) was established empirically after they had compared the amino acid compositions of a large number of proteins. The

method of Cornish-Bowden (1979) is based on statistical theory and unlike the other method incorporates a mechanism to account for any small difference in the number of amino acid residues in the pair of proteins under comparison.

Pairwise comparisons of the amino acid compositions of BZDH I, BZDH II and TOL-BZDH using the method of Marchalonis & Weltman indicated that all three enzymes are homologous to each other. In contrast, the Cornish-Bowden method indicated that none of the enzymes are homologous. A similar pattern was found when BADH and TOL-BADH were compared; the enzymes are homologous as judged by the criteria of Marchalonis & Weltman but unrelated using the criteria of Cornish-Bowden.

5.1.6 Thermal inactivation of enzyme activity

The rates of inactivation of BZDH I, BZDH II and TOL-BZDH were determined at 60 ^{0}C in incubation mixtures containing buffer alone or buffer plus 1 mM-benzaldehyde or 2 mM-NAD^+ (Figure 5.5). The times for each of the enzymes to reach 50% inhibition were estimated from the plots in Figure 5.5 and are summarised in Table 5.8. TOL-BZDH was the most stable of the three enzymes and was protected slightly by the presence of substrate or cofactor in the incubation mixtures. BZDH II, the least stable of the enzymes, was protected significantly by benzaldehyde and perhaps slightly by NAD^+. BZDH I was not protected by benzaldehyde and appeared to be destabilised by NAD^+ in the incubation mixture.

5.1.7 Absorption spectrum of BZDH I

The absorption spectrum of BZDH I showed no absorbance above 400 nm, indicating the absence of bound cofactor or prosthetic group (Figure 5.6). The only unusual feature of the spectrum was a small but distinct blip at 285 nm.

Table 5.3 *N*-Terminal amino acid sequences of purified BZDH I, BZDH II, TOL-BZDH, BADH and TOL-BADH

The *N*-terminal amino acid sequences of BZDH I, BZDH II and BADH were determined at Leads University and also at Aberdeen University (Methods 2.11). The *N*-terminal amino acid sequences of TOL-BZDH and TOL-BADH were determined only at Leeds University (Methods 2.11). The total numbers of sequencing runs carried out at Leeds and Aberdeen Universities were twelve and six, respectively. BZDH I, BZDH II, TOL-BZDH, BADH and TOL-BADH were sequenced four, five, two, five and two times each, respectively. This includes the samples of BZDH II and BADH that were sequenced after they had been eluted from non-denaturing gels (Table 5.4). Consensus sequences for each of the proteins are presented using the one letter amino acid code (Dayhoff, 1978) and the figure in parentheses after each residue indicates the total number of sequencing runs that confirmed this assignment. If the residue has been assigned on the basis of only a single sequencing run then it should be considered a tentative assignment. The data quantifying the recovery of each amino acid are taken from the single best sequencing run for each of the proteins; in all cases this was one of those carried out at Leeds University. The recovery of lysine residues (K) was consistently low because this residue was used to attach the protein to the sequencing matrix.

Cycle No.	BZDH I Residue	BZDH I Quantity (pmol)	BZDH II Residue	BZDH II Quantity (pmol)	TOL-BZDH Residue	TOL-BZDH Quantity (pmol)
1	P (2)	-	S (3)	-	M (2)	64.3
2	N (4)	36.1	I (4)	497	R (2)	464
3	I (4)	13.7	F (5)	623	E (2)	346
4	Q (4)	27.7	T (5)	311	T (2)	207
5	T (2)	18.4	K (5)	74.5	K (1)	29.8
6	K (2)	-	E (5)	326	E (2)	272
7	I (3)	20.3	L (5)	248	Q (2)	222
8	I (3)	20.3	W (4)	227	P (2)	169
9	E (3)	23.0	D (4)	341	I (1)	85.4
10	Q (4)	22.1	K (3)	18.0	W (2)	129
11	I (3)	-	K (2)	-	Y (2)	158
12	W (4)	7.7	L (5)	139	G (2)	137
13	K (3)	1.9	F (5)	225	K (2)	8.6
14	E (4)	13.7	N (4)	208	V (2)	77.8
15	H (2)	19.2	G (5)	108	F (2)	122
16	I (3)	-	S (5)	49.1	S (1)	46.5
17	F (2)	5.7	W (4)	111	S (1)	62.3
18	Q (4)	11.1	Q (5)	140	N (2)	75.8
19	G (2)	5.2	S (3)	30.7	W (2)	52.2
20	T (1)	3.6	A (5)	42.9	V (1)	52.3
21	V (2)	-	Q (5)	94.7	E (1)	55.9
22	V (2)	-	D (5)	133	A (2)	76.8
23	K (1)	-	T (3)	34.9	R (2)	89.0
24			Y (3)	116	G (2)	60.2
25			S (3)	18.9	G (2)	105
26			V (2)	40.4	V (2)	41.7
27			I (2)	50.1	A (2)	45.3
28			E (3)	65.5	N (2)	44.5
29			V (3)	39.7	V (2)	30.8
30			A (3)	44.8	V (2)	53.1
31			T (2)	34.2	D (2)	38.0
32			G (2)	-	P (2)	34.7
33			-	-	S (1)	17.8
34			V (1)	35.4	N (2)	15.7
35			L (2)	30.3	G (2)	16.0
36			G (1)	-	D (2)	12.4
37			E (1)	32.6	I (1)	-
38			I (1)	27.4	L (1)	-
39			G (1)	-	G (1)	11.5
40			Y (2)	55.9	I (2)	14.2
41			A (1)	29.8	T (1)	6.8
42			T (1)	15.1	G (1)	10.5
43			A (1)	30.1	V (1)	12.7
44			Q (1)	11.3	A (1)	21.2

Continued:

Table 5.3 Continued:

Cycle No.	BADH		TOL-BADH	
	Residue	Quantity (pmol)	Residue	Quantity (pmol)
1	S (2)	72.3	M (2)	68.3
2	E (5)	380	E (2)	495
3	L (5)	242	I (2)	507
4	K (5)	77.8	K (2)	12.5
5	D (5)	322	A (2)	473
6	I (5)	106	A (2)	569
7	I (5)	102	I (2)	389
8	A (5)	139	V (2)	470
9	A (4)	161	R (2)	619
10	V (5)	83.4	Q (2)	288
11	T (4)	133	K (2)	16.2
12	P (4)	64.9	N (2)	143
13	C (3)	165	G (2)	207
14	K (4)	13.2	P (2)	120
15	G (5)	113	F (2)	196
16	A (5)	83.6	L (2)	146
17	D (5)	151	L (2)	192
18	F (4)	112	E (2)	106
19	E (5)	119	-	-
20	L (4)	64.2	V (2)	145
21	Q (5)	97.7	A (2)	112
22	A (4)	78.6	L (2)	60.7
23	L (3)	68.5	N (2)	111
24	K (3)	8.3	E (2)	66.1
25	I (3)	38.3	P (2)	53.6
26	R (3)	130	A (2)	63.5
27	Q (4)	107	-	-
28	P (2)	23.8	D (2)	57.0
29	Q (3)	35.3	Q (2)	49.7
30	G (4)	68.4	V (2)	71.0
31	D (2)	76.2	L (2)	49.7
32	E (3)	64.9	V (2)	34.2
33	V (3)	22.9	R (2)	69.4
34	L (1)	30.0	L (2)	41.3
35	V (2)	14.5	V (1)	37.9
36	K (2)	3.0	A (2)	31.4
37	-	-	T (1)	33.5
38	-	-	G (1)	15.3
39	A (2)	25.2	L (1)	34.3
40	T (2)	13.7	P (1)	9.2
41	G (2)	18.1	G (1)	30.6
42	M (1)	11.9	T (1)	10.6

Continued:

Table 5.3 Continued:

Cycle No.	TOL-BADH	
	Residue	Quantity (pmol)
43	D (1)	40.0
44	L (1)	19.8
45	V (1)	14.6
46	-	-
47	R (1)	25.9
48	D (1)	20.2
49	Q (1)	10.0
50	G (1)	14.3
51	Y (1)	16.6
52	P (1)	9.7
53	V (1)	12.4

Table 5.4 *N*-Terminal amino acid sequences of BZDH II and BADH after electroelution from non-denaturing polyacrylamide gels

Purified BZDH II and BADH (Methods 2.6.3) were electrophoresed on non-denaturing polyacrylamide gels and each of the two major bands from each of the enzymes was electroeluted from the gels and sequenced at Leeds University (Methods 2.11). Unfortunately, the identities of the two major bands of BADH were confused and there is no record of which was the top or the bottom band on the gel. The sequences are presented using the one letter amino acid code (Dayhoff, 1978).

BZDH II:

Cycle No.	Bottom Band		Top Band	
	Residue	Quantity (pmol)	Residue	Quantity (pmol)
1	-	-	-	-
2	I	497	I	86.4
3	F	623	F	79.3
4	T	311	T	52.4
5	K	74.5	-	-
6	E	326	E	68.6
7	L	248	L	56.9
8	W	227	W	17.5
9	D	341	D	61.6
10	K	18.0	K	1.5
11	-	-	-	-
12	L	139	L	24.4
13	F	225	F	32.2
14	N	208	N	23.7
15	G	108	G	19.9
16	S	49.1	S	9.5
17	W	111	W	3.2
18	Q	140	Q	11.6
19	S	30.7	S	6.1
20	A	42.9	A	13.9
21	Q	94.7	Q	10.2
22	D	133	D	8.7
23	T	34.9	T	6.4
24	Y	116	Y	10.0
25	S	18.9		
26	V	40.4		
27	I	50.1		
28	E	65.5		
29	V	39.7		
30	A	44.8		
31	T	34.2		
32	-	-		
33	-	-		
34	V	35.4		
35	L	30.3		
36	G	-		
37	E	32.6		
38	I	27.4		
39	-	-		
40	Y	55.9		
41	A	29.8		
42	T	15.1		
43	A	30.1		
44	Q	11.3		

Continued:

Table 5.4 Continued:

BADH:

Cycle No.	Band 1		Band 2	
	Residue	Quantity (pmol)	Residue	Quantity (pmol)
1	-	-	-	-
2	E	31.1	E	54.1
3	L	71.9	L	50.9
4	K	7.4	K	3.9
5	D	50.0	D	54.6
6	I	106	I	65.7
7	I	50.8	I	20.3
8	A	31.5	A	11.7
9	A	34.5	-	-
10	V	36.0	V	9.7
11	T	35.7	T	4.1
12	P	17.4	-	-
13	-	-	-	-
14	K	1.5	-	-
15	G	21.4	G	13.4
16	A	15.4	A	6.4
17	D	37.4	D	20.8
18	F	36.8	F	14.8
19	E	14.9	E	10.0
20	L	14.6	L	7.2
21	Q	16.7	Q	8.7
22	A	13.4	A	4.0
23	L	13.3		
24	-	-		
25	I	10.2		
26	R	23.9		
27	Q	21.4		
28	-	-		
29	Q	17.1		
30	G	7.5		

Table 5.5 Amino acid compositions

The amino acid compositions of purified BZDH I, BZDH II, TOL-BZDH, BADH, TOL-BADH and apomyoglobin (as a control) were determined after the acid hydrolysis of the proteins (Methods 2.10).

The amino acid composition of apomyoglobin and each of the *Acinetobacter* enzymes was determined twice in completely independent experiments that were carried out at least several weeks apart. In one of the experiments the amino acid composition of the hydrolysed protein was determined in duplicate. The results of the duplicate determinations were first averaged together and then this composition was averaged with the independent determination to give the results presented here. The amino acid composition of single samples of each of the TOL plasmid-encoded enzymes was determined in duplicate and the mean values are presented. In general the duplicate and the independent amino acid compositions for each enzyme were in very good agreement and in some cases the compositions differed only by a fraction of a percent. Indications of the reproducibility of the amino acid composition determinations are given in Tables 5.6 and 5.7 where the independent determinations or the duplicate determinations are compared by methods based on the differences squared between the percentage composition of each amino acid.

* The amino acid composition of sperm whale apomyoglobin was calculated from the amino acid sequence (Edmundson, 1965).

Amino acid	% Composition						
	BZDH I	BZDH II	TOL-BZDH	BADH	TOL-BADH	Apo-myoglobin	Apo-myoglobin*
Asx	9.68	9.09	10.3	8.61	10.3	5.70	5.23
Thr	4.74	5.19	3.76	5.28	4.54	2.94	3.27
Ser	7.61	5.53	4.75	4.84	6.45	3.59	3.92
Glx	14.4	15.7	10.4	11.5	9.55	14.0	12.4
Pro	4.24	3.93	5.79	4.63	6.40	2.49	2.61
Gly	8.33	9.90	10.4	11.0	10.6	7.92	7.19
Ala	7.35	10.9	12.0	9.51	8.79	11.4	11.1
Cys	1.46	0.80	0.13	2.61	0.10	0	0
Val	7.17	8.16	9.23	8.00	9.60	5.11	5.23
Met	2.52	0.46	2.3	1.32	1.18	1.57	1.31
Ile	9.47	6.24	5.53	6.77	5.97	5.45	5.88
Leu	6.95	9.35	7.72	9.04	9.55	11.9	11.8
Tyr	1.53	1.91	1.25	2.08	1.62	2.10	1.96
Phe	4.21	4.33	3.55	4.72	4.54	4.05	3.92
His	2.46	1.50	2.09	2.26	1.96	7.70	7.84
Lys	5.78	4.69	4.96	5.86	5.11	12.0	12.4
Arg	2.14	2.32	5.79	1.93	3.77	2.66	2.61

Table 5.6 Comparison of the amino acid compositions by the method of Marchalonis & Weltman (1971)

The amino acid compositions of the enzymes were compared in a pairwise fashion by calculating the quantity SΔQ which is the sum of the difference squared between the percentage composition of each type of amino acid residue in the proteins. A good indication of sequence homology between two proteins is given by values of less than 100.

When the same enzyme appears at the top of a column and the start of a row, SΔQ (shown in bold) is a comparison between the two independent or duplicate amino acid composition determinations that were averaged to give the values presented in Table 5.5. These values are an indication of the reproducibility of the amino acid composition determinations. The value in parentheses at the bottom right hand corner of the table is the SΔQ for the comparison of the experimentally determined amino acid composition of apomyoglobin and that calculated from the amino acid sequence of the protein (Edmundson, 1965).

ND, not determined

SΔQ

Protein	BZDH I	BZDH II	TOL-BZDH	BADH	TOL-BADH	Apomyoglobin
BZDH I	**1.6**					
BZDH II	46	**3.2**				
TOL-BZDH	90	58	**0.5**			
BADH	ND	ND	ND	**11.4**		
TOL-BADH	ND	ND	ND	28	**0.8**	
Apomyoglobin	ND	ND	ND	ND	ND	**3.2** (4.1)

Table 5.7 Comparison of the amino acid compositions by the method of Cornish-Bowden (1979)

The amino acid compositions of the enzymes (Table 5.5) were compared in a pairwise fashion by calculating the quantity SΔn [equation (5) in the reference given above] which is based on the sum of the difference squared between the numbers of each type of amino acid residue in the proteins. SΔn has a component that accounts for any difference in the number of amino acid residues in each protein. SΔn is presented as a ratio with the number of amino acid residues per subunit for the smaller of the two proteins being compared. Values of less than 0.42 indicate sequence homology between the two proteins and satisfy a formal 95% significance test.

For the purpose of calculating SΔn the number of amino acid residues per subunit for both BADH and TOL-BADH was assumed to be 381 (see Section 5.5.1)

When the same enzyme appears at the top of a column and the start of a row, SΔn (shown in bold) is a comparison between the independent or duplicate amino acid composition determinations that were averaged to give the values presented in Table 5.5. These values are an indication of the reproducibility of the amino acid composition determinations. The value in parentheses at the bottom right hand corner of the table is the value for the comparison of the experimentally determined amino acid composition of apomyoglobin and that calculated from the amino acid sequence of the protein (Edmundson, 1965).

ND, not determined

SΔn/amino acid residues per subunit

Protein	BZDH I	BZDH II	TOL-BZDH	BADH	TOL-BADH	Apomyoglobin
BZDH I	**0.04**					
BZDH II	1.26	**0.09**				
TOL-BZDH	2.44	1.53	**0.01**			
BADH	ND	ND	ND	**0.23**		
TOL-BADH	ND	ND	ND	0.55	**0.02**	
Apomyoglobin	ND	ND	ND	ND	ND	**0.03** (0.03)

Figure 5.5 Thermal inactivation of BZDH I, BZDH II and TOL-BZDH

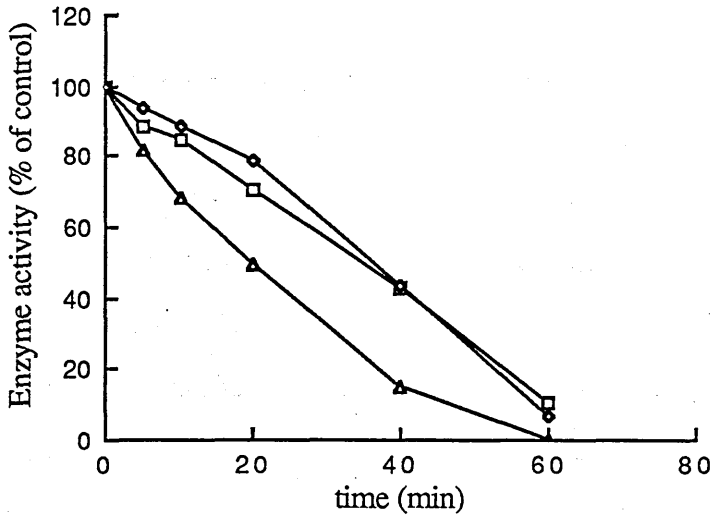
The enzymes were diluted in buffer (100 mM-potassium phosphate, pH 7.5/ 5 mM-DTT) and incubated in sealed Eppendorf tubes in a water bath at 60 ^{0}C. Samples were also incubated under the same conditions in buffer plus 2 mM-NAD^+ or 1 mM-benzaldehyde. At various times after the start of the incubation samples were removed, cooled on ice for 2 min and centrifuged (Methods 2.2.8) for 2 min. Supernates were then assayed for enzyme activity (Methods 2.5.1a & 2.5.2). Results are shown for a single representative experiment and several other experiments gave similar results.

A preliminary experiment showed that after heating solutions of 1 mM-benzaldehyde or 2 mM-NAD^+, in buffer at 60 ^{0}C for 90 min, 0.74 mM-benzaldehyde and 1.1 mM-NAD^+ remained, as judged by the ΔA_{340} after complete enzymic oxidation or reduction. These values are greater that 10 times the K'_m values of the enzymes (Section 5.3.3) and the effects of the compounds on the enzymes were therefore probably constant throughout the time course of the experiment. The incubation buffer contained 5 mM-DTT because a preliminary experiment indicated that samples of BZDH I, in buffer containing different amounts of DTT, were stable in the following order after heating for 1 h: 0 mM < 2 mM < 5 mM = 15 mM.

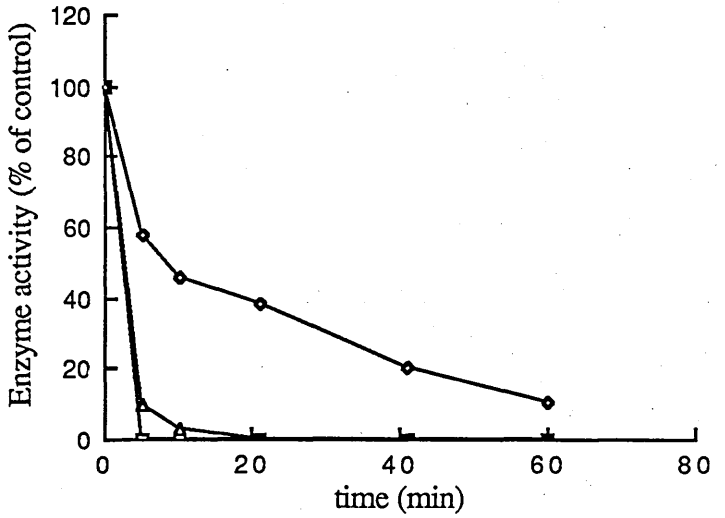
The enzyme concentrations in the incubation and the assay mixtures respectively were: BZDH I, 16.4 µg/ml and 1.2 µg/ml; BZDH II, 6.0 µg/ml and 0.45 µg/ml; TOL-BZDH, 8.6 µg/ml and 0.43 µg/ml.

Plots of % enzyme activity remaining against time: (a), BZDH I; (b), BZDH II; (c), TOL-BZDH.
Incubation buffers: buffer alone, □; buffer plus 2 mM-NAD^+, △; buffer plus 1 mM-benzaldehyde, ◇.

(a)


Enzyme activity (% of control)
120
100
80
60
40
20
0
0
20
40
60
80
time (min)

(b)


Enzyme activity (% of control)
120
100
80
60
40
20
0
0
20
40
60
80
time (min)

(c)

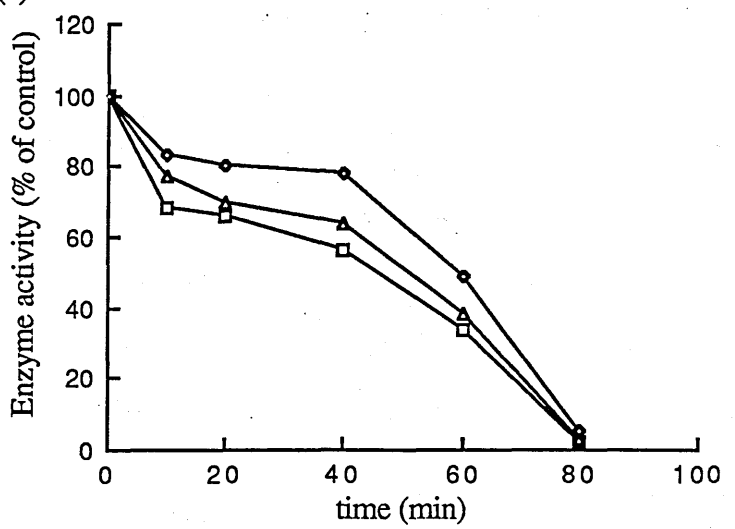

Enzyme activity (% of control)
120
100
80
60
40
20
0
0
20
40
60
80
100
time (min)

Table 5.8 Summary of the thermal inactivation experiments

The times to reach 50% inactivation for each of the aldehyde dehydrogenases in each of the different incubation mixtures were estimated from the plots in Figure 5.5.

Incubation mixture	Time (min) to reach 50% inactivation at 60 ^{0}C		
	BZDH I	BZDH II	TOL-BZDH
Buffer	34	<2	50
Buffer + NAD^+	20	2	62
Buffer + benzaldehyde	36	9	80

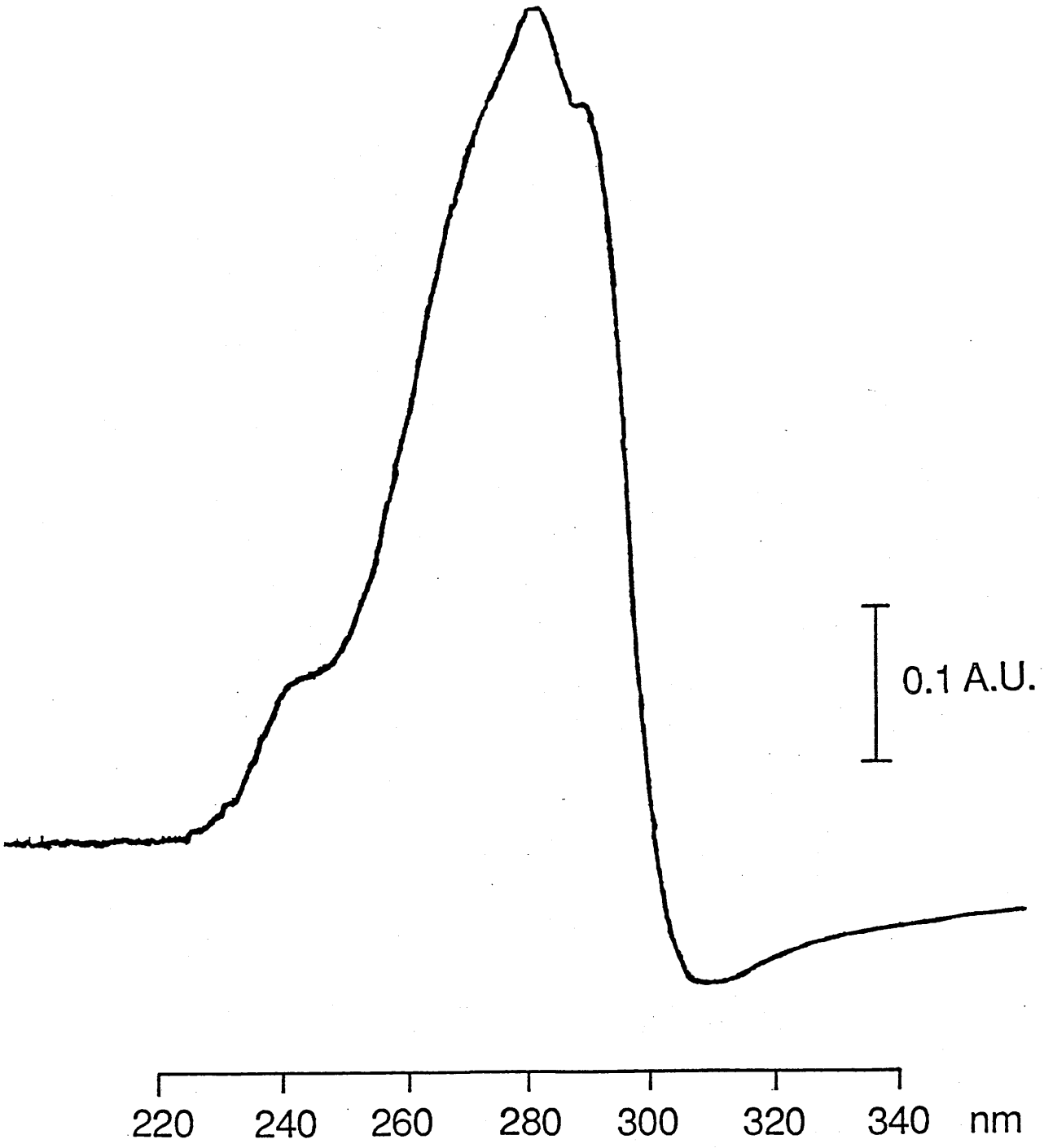


Figure 5.6 Absorption spectrum of BZDH I

The absorption spectrum of BZDH I in 100 mM-potassium phosphate buffer, pH 7.5/2 mM-DTT was recorded in a Pye Unicam SP8-100 spectrophotometer set to scan at 0.5 nm/sec. The light path was 1 cm and buffer without protein was placed in the reference beam. The protein concentration of the sample was 0.38 mg/ml.

5.2 Preliminary kinetic studies

5.2.1 Development of the assay procedures for BZDH I. TOL-BZDH and TOL-BADH

Preliminary experiments showed that the apparent K_m (K'_m) values of BZDH I for benzaldehyde and NAD^+ were approximately 1 μM and 100 μM respectively and that approximately 100 mM-K^+ was required for maximum activity. Subsequent experiments were therefore carried out with substrate concentrations at least ten times the values given above and with 100 mM-KCl.

Extract (Methods 2.4.4b) containing BZDH I was assayed using the standard procedure (Methods 2.5.1a) except that the following buffers (113 mM) replaced the 0.15 M-glycine/KOH assay buffer: sodium pyrophosphate/HCl, glycylglycine/KOH, glycine/KOH, Bicine/HCl and Ches/KOH buffers plus 100 mM-KCl with pH values between 8 and 10. The highest activity was in glycine/KOH buffer, pH 9.5 and this was adopted as the assay buffer for BZDH I.

The original TOL-BZDH assay contained glycine/NaOH buffer, pH 9.4 (Worsey & Williams, 1975). This was changed to glycine/KOH so that BZDH I, BZDH II and TOL-BZDH were all assayed in the same buffer. TOL-BZDH was substrate-inhibited by benzaldehyde concentrations above approximately 10 μM.

The original assay for TOL-BADH (Worsey & Williams, 1975) contained Tris/HCl buffer, pH 8.7 and the progress curve of the reaction became shallower as the assay proceeded. This was possibly due to the occurrence of the reverse reaction as benzaldehyde accumulated in the assay mixture. Linear progress curves were obtained by using the BADH assay (Methods 2.5.2), which contained hydrazine to trap the aldehyde product (MacKintosh & Fewson, 1988a).

5.2.2 pH profiles for BZDH I, TOL-BZDH and TOL-BADH activities

The pH profile of BZDH I activity in assay mixtures containing 0.15 M-glycine/KOH buffer is shown in Figure 5.7. The pH optimum for the reaction was very close to 9.5 with either 1 mM or 5 mM-NAD^+ in the reaction mixtures. The experiment was carried out at two different NAD^+ concentrations to confirm that the pH optimum of the reaction was not different when the NAD^+ concentration was less-than-saturating (1 mM).

The pH profiles for TOL-BZDH and TOL-BADH in glycine and Bicine/hydrazine buffers respectively are shown in Figure 5.8. The pH optima for TOL-BZDH and TOL-BADH were pH 9.3 and 9.4 respectively.

5.2.3 Dependence of BZDH I activity upon protein concentration

The relationship between the amount of BZDH I in an assay mixture and the rate of the reaction was determined. When the assay contained 100 µM-benzaldehyde the rate was proportional to the amount of enzyme added when the change in absorbance was 12×10^{-3} - 82×10^{-3}/min (Figure 5.9a). If the initial benzaldehyde concentration in the assay mixture was lowered to between 3 and 12 µM the rate was proportional to the amount of enzyme added when the change in absorbance was 1.5×10^{-3} - 9.5×10^{-3}/min (Figure 5.9b).

5.2.4 Reproducibility of the BZDH I assay

BZDH I was assayed in an LKB Ultrospec spectrophotometer using the High Range assay (Methods 2.5.1a) with 100 µM-benzaldehyde as substrate. When 190 ng of enzyme was assayed repeatedly the average rate of the reaction was 4.3 nmol/min, (SD = 0.146, n = 12) and when 475 ng of enzyme was repeatedly assayed the average rate was 12.3 nmol/min, (SD = 0.245 , n = 11).

BZDH I was also assayed in a Pye Unicam SP8-100 spectrophotometer using the Low Range assay (Methods 2.5.1b) with 7.5 μM-benzaldehyde as substrate. When samples containing 20 ng of enzyme were assayed repeatedly the average rate of the reaction was 0.57 nmol/min, (SD = 0.027, n = 14) and when 40 ng of enzyme was repeatedly assayed the average rate was 1.12 nmol/min, (SD = 0.07, n = 10).

5.2.5 Effects of salts on BZDH I activity

BZDH I was activated by K^+, Rb^+ and NH_4^+. The activation of BZDH I by KCl and NH_4Cl is shown in Figure 5.10. The activation curve for Rb^+ was almost identical to that for K^+. The data from this experiment were analysed by the Direct Linear method (Methods 2.5.4), and the K'_m for K^+ was determined to be 11.2 mM (68% confidence limits = 10.8 - 12.1 mM). The activation curves for K^+ and Rb^+ were each in the form of a rectangular hyperbola. The response to NH_4^+ was sigmoidal but I have no explanation for this observation and I know of no other documented cases of this type. The activation of BZDH I by various salts was also determined under the standard fluorimetric assay conditions (Methods 2.5.1c), except that 0.15 M-ethanolamine/HCl replaced the standard glycine assay buffer and 70 mM-salt was also included in the assay mixture. BZDH I was first changed into 50 mM-Tris/HCl, pH 7.5 by gel filtration (Methods 2.2.4), so that no K^+ was added to the assay mixtures in the enzyme stock solution. The activation of BZDH I by KCl, RbCl and NH_4Cl was 164, 135 and 129-fold respectively. BZDH I was not activated by Na^+, nor was there evidence from experiments using various buffers that either Tris or ethanolamine give any significant activation of BZDH I.

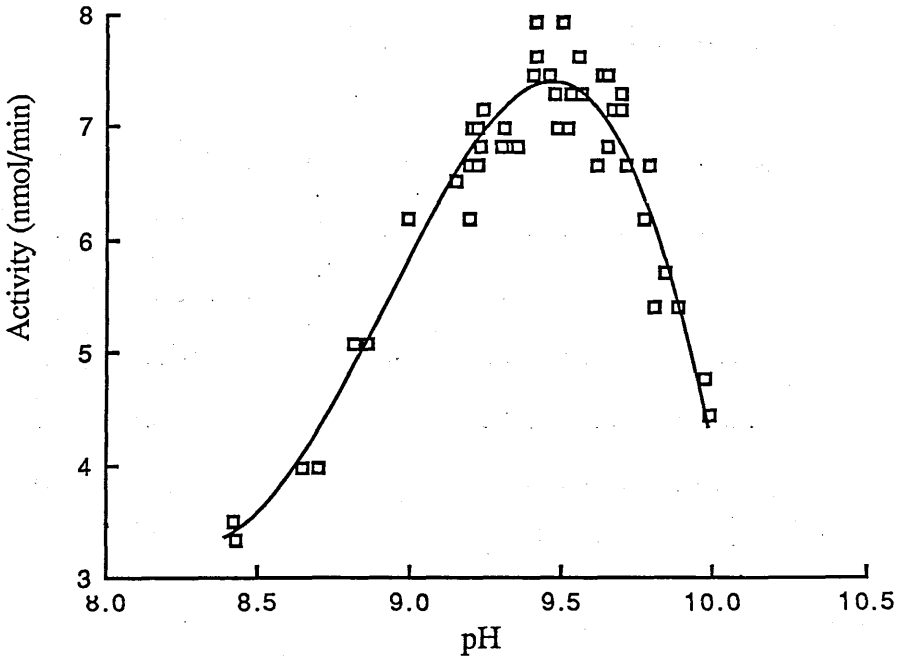
The stability of BZDH I (at 4 ^{0}C) was not lowered in the absence of K^+ and there was no detectable lag phase (>30 s) when K^+-depleted enzyme was assayed in the presence of K^+.

Figure 5.7 Effect of pH on the activity of BZDH I

BZDH I was assayed under standard assay conditions (Methods 2.5.1a) except that the assay mixtures contained 100 mM-KCl and either 1.0 or 5.0 mM-NAD^+. Assays were carried out at a series of pH values which were measured at room temperature after the assays had gone to completion. The concentration of BZDH I in the assay mixture was 0.38 μg/ml. KCl was included in the assay mixtures to ensure that the enzyme activity (which is K^+-activated; Section 5.2.5) was not effected by the different amounts of KOH that were used to adjust the assay buffer to different pH values.

(a), 1.0 mM-NAD^+; (b), 5.0 mM-NAD^+.

(a)


Activity (nmol/min)
8
7
6
5
4
3
8.0
8.5
9.0
9.5
10.0
10.5
pH

(b)

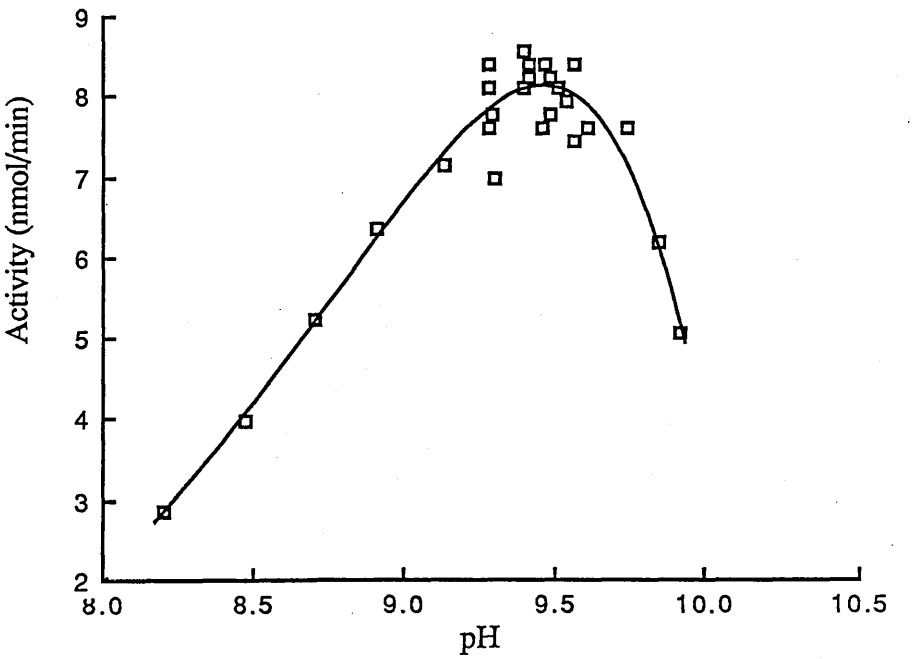
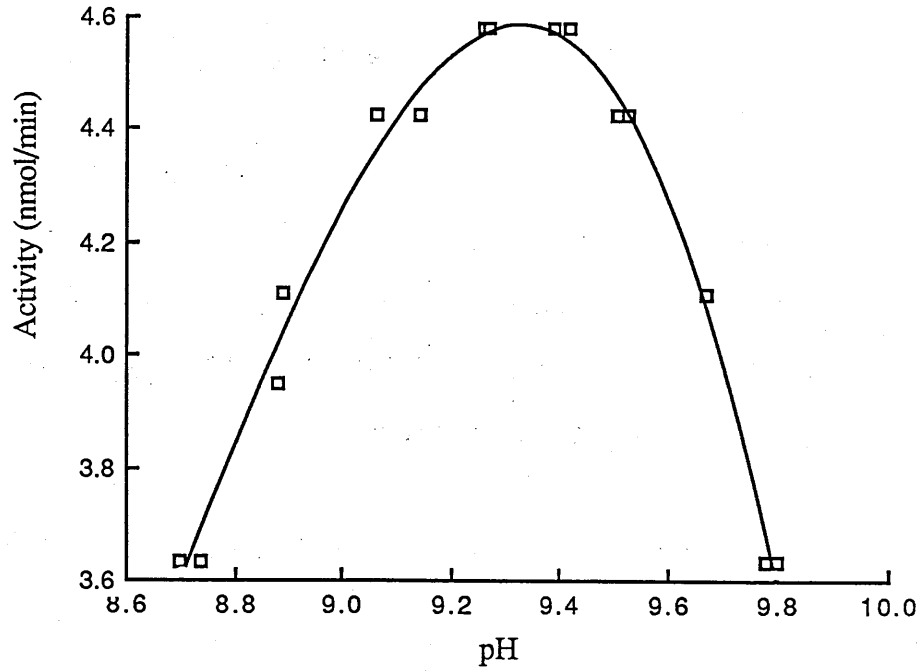

Activity (nmol/min)
9
8
7
6
5
4
3
2
8.0
8.5
9.0
9.5
10.0
10.5
pH

Figure 5.8 Effect of pH on the activity of TOL-BZDH and TOL-BADH

Enzymes were assayed in their respective assay mixtures (Methods 2.5.1a & 2.5.2) except that the assays were carried out at a series of different pH values. The pH values of the assay mixtures were measured at room temperature after the assays had gone to completion. The concentrations of TOL-BZDH and TOL-BADH in the assay mixtures were 0.13 and 0.05 μg/ml respectively.

(a), TOL-BZDH; (b), TOL-BADH.

(a)


Activity (nmol/min)
4.6
4.4
4.2
4.0
3.8
3.6
8.6
8.8
9.0
9.2
9.4
9.6
9.8
10.0
pH

(b)

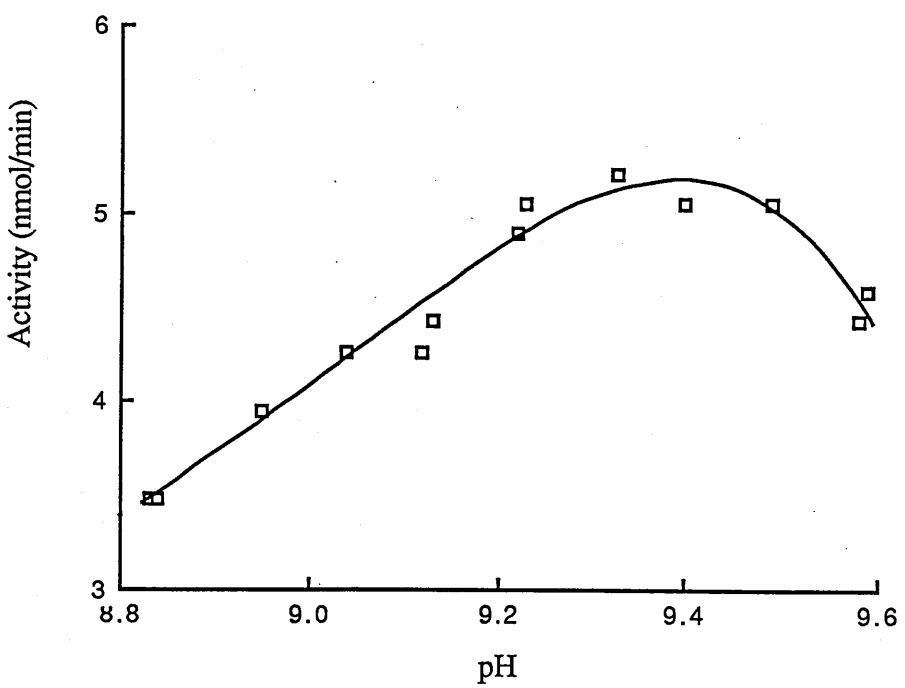

Activity (nmol/min)
6
5
4
3
8.8
9.0
9.2
9.4
9.6
pH

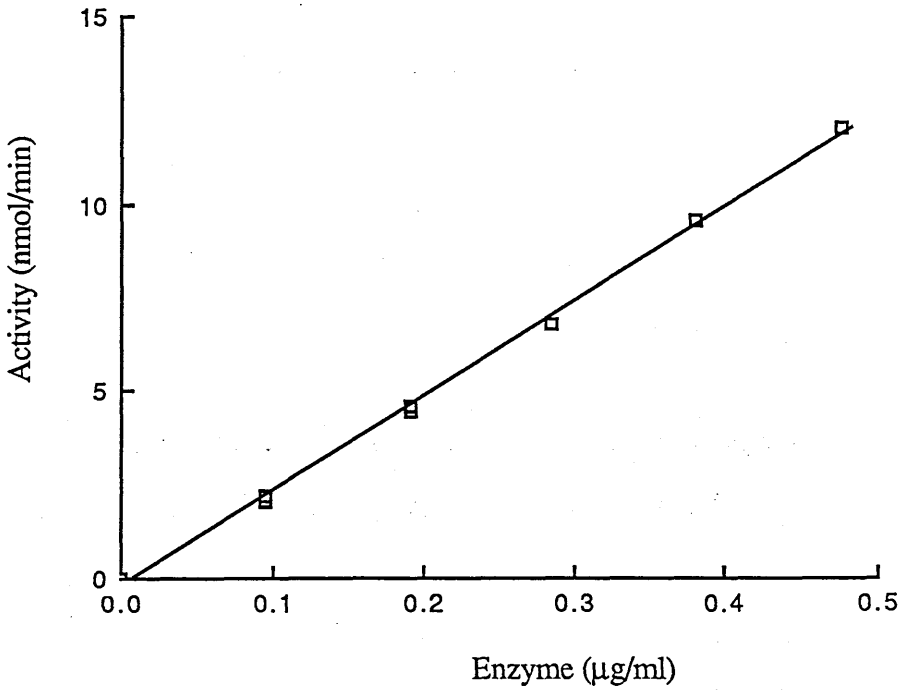
Figure 5.9 Dependence of the rate of the BZDH I reaction upon the protein concentration

Various amounts of BZDH I were assayed under the standard conditions (Methods 2.5).

(a) The High Range assay was carried out using an LKB Ultrospec spectrophotometer (Methods 2.5.1a). Assays were carried out in duplicate and when only one point is shown the values were identical.

(b) The Low Range assay was carried out using a Pye Unicam SP8-100 spectrophotometer (Methods 2.5.1b). Assays were carried out in quadruplicate and when only one point is shown the values were identical.

(a)


15
10
5
0
Activity (nmol/min)
0.0
0.1
0.2
0.3
0.4
0.5
Enzyme (μg/ml)

(b)

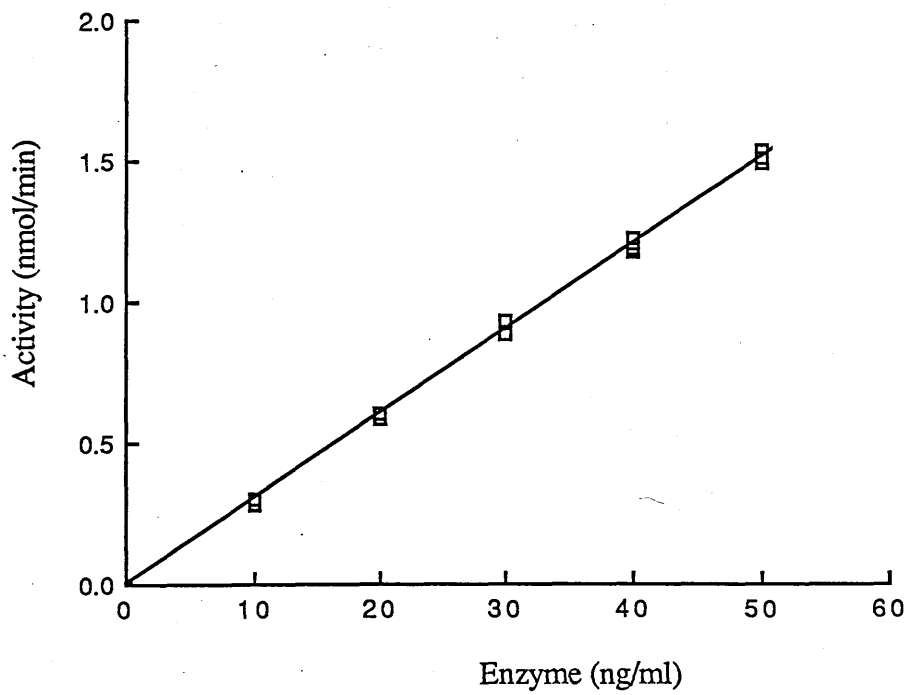
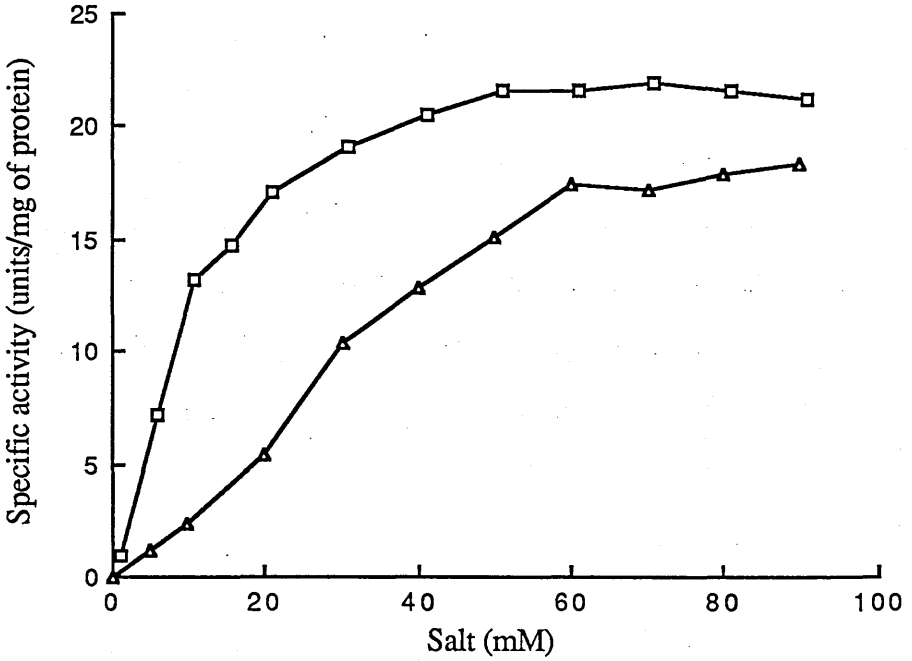

2.0
1.5
1.0
0.5
0.0
Activity (nmol/min)
0
10
20
30
40
50
60
Enzyme (ng/ml)

Figure 5.10 Effects of KCl and NH_4Cl on the activity of BZDH I

BZDH I was assayed under the standard assay conditions (Methods 2.5.1a) except that 0.15 M-ethanolamine/HCl buffer replaced the standard glycine assay buffer and various concentrations of salts were also included in the assay mixture. All the assay mixtures contained 0.48 μg BZDH I and 890 μM-K^+ from the enzyme stock solution. Assays were carried out in duplicate for the activation by KCl and the mean values were plotted. The values of the duplicate assays differed by no more than 5%.

KCl, □; NH_4Cl, △.


Specific activity (units/mg of protein)
25
20
15
10
5
0
0
20
40
60
80
100
Salt (mM)

5.2.6 Cofactor specificity of BZDH I, BZDH II, TOL-BZDH, BADH and TOL-BADH

The enzyme activities were each measured in the standard assay mixtures (Methods 2.5.1a & 2.5.2) except that 2 mM-$NADP^+$ replaced 2 mM-NAD^+. The mean activities with 2 mM-$NADP^+$, relative to the activities with 2 mM-NAD^+, were as follows (numbers of determinations in parentheses): BZDH I, 4.9% (4); BZDH II, 4.3% (4); TOL-BZDH, 27% (8); BADH, 7.8% (4) and TOL-BADH, 1.9% (6).

5.3 Steady-state kinetics

5.3.1 Substrate inhibition of BZDH I by benzaldehyde

BZDH I was progressively inhibited by benzaldehyde concentrations above approximately 4 μM (Figure 5.11).

5.3.2 Substrate specificity of BZDH I

BZDH I was assayed with a variety of compounds as substrates (Table 5.9). The enzyme did not use either of the aliphatic aldehydes tested and with the exception of 4-hydroxybenzaldehyde it had a higher tolerance for substitutions at the *meta* and *para* positions of the aromatic ring than at the *ortho* position.

5.3.3 Kinetic coefficients of selected substrates of BZDH I, BZDH II, TOL-BZDH, BADH and TOL-BADH

The apparent maximum velocity (V'_{max}) and the K'_m values were determined for BZDH I, BZDH II and TOL-BZDH with NAD^+, benzaldehyde and the three possible

monomethyl substituted benzaldehydes (Table 5.10). The kinetic constants were also determined for BADH and TOL-BADH with NAD^+, benzyl alcohol and the three possible monomethyl substituted benzyl alcohols (Table 5.10). The values of V'_{max} and K'_m determined for BZDH II and BADH, with NAD^+ and the unsubstituted substrates, were approximately the same as those previously determined by MacKintosh & Fewson (1987a,b).

The kinetic coefficients were used to calculate the apparent specificity constants (V'_{max}/K'_m) of the enzymes for the substrates (Table 5.11). V'_{max}/K'_m is analogous to the specificity constant, K_{cat}/K_m, except that instead of relating the activity to the number of active sites, the activity is related to the amount of protein. This approach was taken because the aim of this work was to compare the relative specificities of the enzymes for each of the substrates and any attempt to incorporate the number of active sites into the calculations can only increase the experimental error.

In general, the *ortho* substituted compounds were poor substrates for the enzymes while the *meta* and *para* substituted compounds were as good, or better, than the unsubstituted analogue.

The V'_{max} value of BZDH I with benzaldehyde (Table 5.10) was approximately four times the value of the specific activity of the purified enzyme (Table 3.1) measured using the High Range assay (Methods 2.5.1a). This difference was presumably caused by substrate inhibition of the enzyme (Section 5.3.1) by the high benzaldehyde concentration (100 μM) in the High Range assay mixture.

The experiments with TOL-BADH were carried out with a batch of enzyme that was purified before the final purification scheme (Methods 2.6.2) was established. The specific activity of this batch of enzyme (approximately 90 units/mg of protein), with NAD^+ and benzyl alcohol as substrates, was lower than that for the enzyme prepared by the final procedure (376 units/mg of protein; Table 4.1). This difference was probably caused by the longer time required for the preliminary purification scheme and the larger number of steps and manipulations carried out. The lower specific activity of the batch of

TOL-BADH used in these experiments presumably had no effect on the relative values of the specificity constants determined for the various substrates.

5.3.4 Esterase activities of BZDH I and BZDH II

Both BZDH I and BZDH II had esterase activity with 4-nitrophenol acetate as substrate. The rates at pH 8.5 were each approximately 2% of the dehydrogenase activities and the enzymes were both activated by low concentrations of NAD^+ (Figure 5.12).

5.4 Effects of thiol-blocking reagents on BZDH I

The effects of various thiol-blocking reagents on BZDH I are shown in Figure 5.13 and the data are summarised in Table 5.12. The enzyme was sensitive to inhibition by the reagents in the following order: 4-chloromercuribenzoate = *N*-ethylmaleimide > iodoacetamide > iodoacetate. BZDH I was protected from inhibition by iodoacetate and *N*-ethylmaleimide by benzaldehyde and NAD^+ (Figure 5.14 and 5.17).

BZDH I was gel filtered (Methods 2.2.4) into DTT-free buffer and made 10 or 50 μM in *N*-ethylmaleimide (Figure 5.15). The time to reach 50% inhibition ($t_{1/2}$) was estimated to be 2 minutes and < 2 minutes, respectively. These times are much shorter than those from the experiments described above (30 min and 4 min, respectively; Table 5.12) which were carried out in the presence of 32 μM-DTT added in the enzyme stock solution. This difference in $t_{1/2}$ values is consistent with a reaction between the thiol-blocking reagent and DTT that occurs at approximately the same rate as the reaction between the thiol-blocking reagent and the protein.

Figure 5.11 Substrate inhibition of BZDH I by benzaldehyde

BZDH I was assayed fluorimetrically (Methods 2.5.1c) at various concentrations of benzaldehyde with a close-to-saturating concentration of NAD^+ (2 mM). Results are shown for a single representative experiment and other experiments gave similar results.

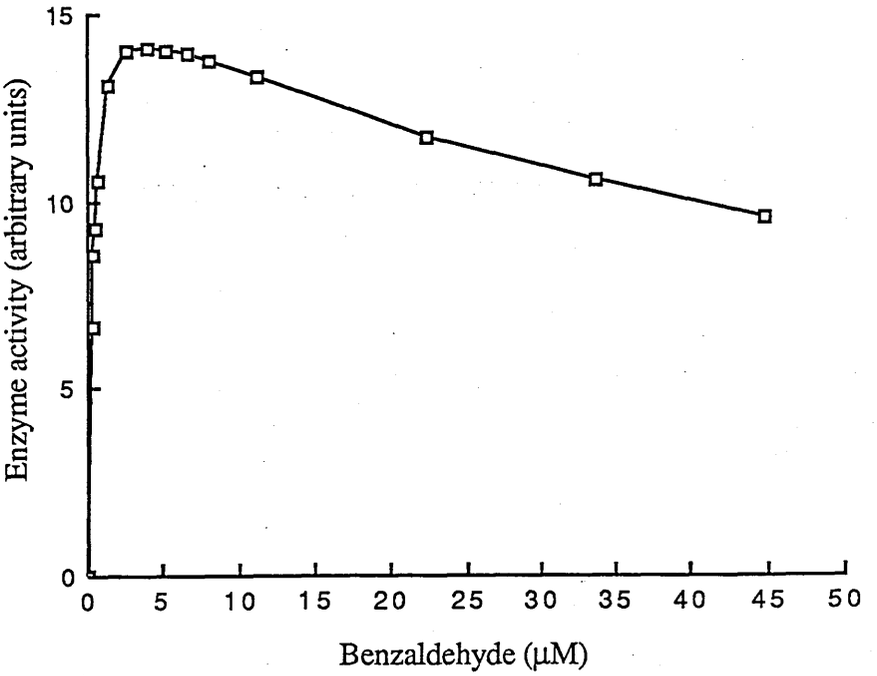

Enzyme activity (arbitrary units)
15
10
5
0
0
5
10
15
20
25
30
35
40
45
50
Benzaldehyde (μM)

Table 5.9 Substrate specificity of BZDH I

BZDH I was assayed under the standard assay conditions (Methods 2.5.1b) except that alternative compounds (100 μM) replaced benzaldehyde. Each 1 ml assay contained 70 ng of enzyme and the rate of the reaction with benzaldehyde was 1.48 nmol/min. The rates of the reactions with substrates 7, 8 and 12 have been adjusted using a conversion factor (MacKintosh & Fewson, 1988b) because either the substrates or the products absorb light at 340 nm.

Compound		% of activity with benzaldehyde
1	2-Fluorobenzaldehyde	41
2	3-Fluorobenzaldehyde	98
3	4-Fluorobenzaldehyde	210
4	2-Methylbenzaldehyde	13
5	3-Methylbenzaldehyde	119
6	4-Methylbenzaldehyde	240
7	2-Hydroxybenzaldehyde	39
8	3-Hydroxybenzaldehyde	222
9	4-Hydroxybenzaldehyde	0
10	3,4-Dihydroxybenzaldehyde	0
11	2-Methoxybenzaldehyde	0
12	3-Methoxybenzaldehyde	235
13	2-Bromobenzaldehyde	0
14	3-Methoxy-4-hydroxybenzaldehyde	0
15	3-Hydroxy-4-methoxybenzaldehyde	0
16	Hexan 1-al	0
17	Octan 1-al	0
18	Perillaldehyde	0

Table 5.10 Kinetic constants of selected substrates of BZDH I, BZDH II, TOL-BZDH, BADH and TOL-BADH

The K'_m and V'_{max} values for the substrates were determined by measuring the initial reaction rates at various non-saturating concentrations of the first substrate in the presence of a fixed concentration of the second substrate. The standard assay mixtures (Methods 2.5) were used for each of the enzymes except that the concentration of one of the substrates was varied. When the K'_m was determined for BZDH I with NAD^+ a saturating concentration of benzaldehyde was not used (because of substrate inhibition). The concentrations of benzaldehyde used in each of the $K'_m{}^{NAD^+}$ determinations is given in parentheses in the table. The kinetic parameters were calculated from initial velocity measurements by the Direct Linear method (Methods 2.5.4) and the 68% confidence limits are given in parentheses. The concentrations of the substrate stock solutions were calculated from the ΔA_{340} after the complete enzymic oxidation or reduction of one substrate in the presence of excess amounts of the second substrate.

The kinetic constants for the aldehyde dehydrogenases were determined using fluorimetric assays (Methods 2.5.1c) and the V'_{max} values have been converted from arbitrary units to μmol/min/mg of protein using a conversion factor (Methods 2.5.1c).

Enzyme	Substrate	K'_m (μM)		V'_{max} (units/ mg of protein)	
BZDH I	NAD^+ (3.7 μM-benzaldehyde)	96.0	(73.0 - 106)	110	(99.6 - 117)
	benzaldehyde	0.69	(0.53 - 0.80)	106	(95.4 - 120)
	2-methylbenzaldehyde	26.8	(24.8 - 28.4)	8.59	(8.40 - 8.70)
	3-methylbenzaldehyde	1.22	(1.11 - 1.32)	100	(93.1 - 106)
	4-methylbenzaldehyde	3.02	(2.67 - 3.35)	49.1	(47.5 - 50.6)
BZDH II	NAD^+ (11.2 μM-benzaldehyde)	121	(113 - 133)	60.8	(58.9 - 63.1)
	benzaldehyde	0.63	(0.56 - 0.73)	63.5	(60.1 - 63.1)
	2-methylbenzaldehyde	28.2	(22.8 - 33.0)	17.5	(15.8 - 18.3)
	3-methylbenzaldehyde	0.38	(0.33 - 0.41)	22.3	(21.6 - 22.9)
	4-methylbenzaldehyde	0.97	(0.82 - 1.13)	54.0	(51.3 - 58.2)
TOL-BZDH	NAD^+ (22.2 μM-benzaldehyde)	79.0	(74.3 - 84.4)	96.2	(93.2 - 98.9)
	benzaldehyde	0.79	(0.71 - 0.98)	104	(97.7 - 110)
	2-methylbenzaldehyde	88.3	(80.4 - 123)	3.92	(3.70 - 4.90)
	3-methylbenzaldehyde	0.60	(0.56 - 0.65)	46.8	(45.6 - 47.1)
	4-methylbenzaldehyde	1.06	(1.00 - 1.13)	94.7	(93.6 - 97.0)

Continued:

Table 5.10 Continued:

Enzyme	Substrate	K'_m (μM)		V'_{max} (units/ mg of protein)	
BADH	NAD^+	40.6	(38.6 - 40.8)	357	(372 - 379)
	benzyl alcohol	121	(115 - 127)	351	(337 - 356)
	2-methylbenzyl alcohol	992	(920 - 1080)	405	(387 - 418)
	3-methylbenzyl alcohol	146	(136 - 152)	498	(485 - 507)
	4-methylbenzyl alcohol	118	(110 - 126)	593	(564 - 620)
TOL-BADH	NAD^+	219	(209 - 229)	90.8	(87.7 - 92.8)
	benzyl alcohol	233	(209 - 259)	96.4	(92.2 - 106)
	2-methylbenzyl alcohol	605	(578 - 617)	55.4	(53.8 - 59.9)
	3-methylbenzyl alcohol	81.2	(76.8 - 85.7)	83.6	(82.4 - 85.2)
	4-methylbenzyl alcohol	106	(99.6 - 113)	90.5	(87.1 - 93.0)

Table 5.11 Apparent specificity constants of selected substrates of BZDH I, BZDH II, TOL-BZDH, BADH and TOL-BADH

The apparent specificity constants (V'_{max}/K'_m) for the substrates were calculated from the data in Table 5.10.

Enzyme	Substrate	V'_{max}/K'_m [l .min^{-1}.(mg of protein)$^{-1}$]
BZDH I	NAD^+	1.15
	benzaldehyde	154
	2-methylbenzaldehyde	0.32
	3-methylbenzaldehyde	82.0
	4-methylbenzaldehyde	16.3
BZDH II	NAD^+	0.50
	benzaldehyde	101
	2-methylbenzaldehyde	0.62
	3-methylbenzaldehyde	58.7
	4-methylbenzaldehyde	55.6
TOL-BZDH	NAD^+	1.22
	benzaldehyde	132
	2-methylbenzaldehyde	0.05
	3-methylbenzaldehyde	78.0
	4-methylbenzaldehyde	89.3
BADH	NAD^+	8.79
	benzyl alcohol	2.90
	2-methylbenzyl alcohol	0.41
	3-methylbenzyl alcohol	3.41
	4-methylbenzyl alcohol	5.02
TOL-BADH	NAD^+	0.42
	benzyl alcohol	0.41
	2-methylbenzyl alcohol	0.09
	3-methylbenzyl alcohol	1.03
	4-methylbenzyl alcohol	0.85

Figure 5.12 Esterase activities of BZDH I and BZDH II

The esterase activities of BZDH I and BZDH II were measured with 4-nitrophenol acetate as substrate (Methods 2.5.5) in the presence of various concentrations of NAD^+. Results are shown for a single representative experiment and other experiments gave similar results.

BZDH I, □; BZDH II, △.

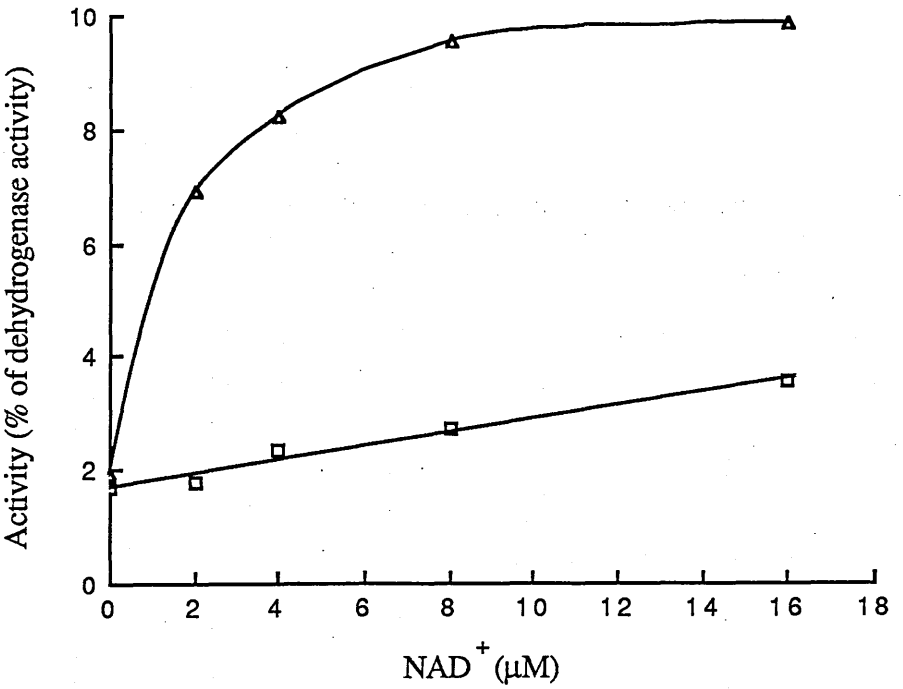

Activity (% of dehydrogenase activity)
10
8
6
4
2
0
0
2
4
6
8
10
12
14
16
18
NAD+ (μM)

Figure 5.13 Effects of thiol-blocking reagents on BZDH I

BZDH I was diluted 62-fold in 25 mM-potassium phosphate buffer, pH 7.0 containing various concentrations of thiol-blocking reagents. The mixtures were incubated on ice and at various times after the start of the incubation samples were removed and assayed (Methods 2.5.1a). A sample of enzyme was also incubated in buffer without thiol-blocking reagent as a control. The concentrations of enzyme in the incubation and the assay mixtures were 11.9 and 0.6 μg/ml respectively.

Results are shown for a single representative experiment and other experiments gave similar results.

Some of the inactivation curves in parts (c) and (d) reached a steady-state because the concentration of thiol-blocking reagent was very low and the incubations were carried out in the presence of a thiol-reagent (32 μM-DTT) added in the enzyme stock solution.

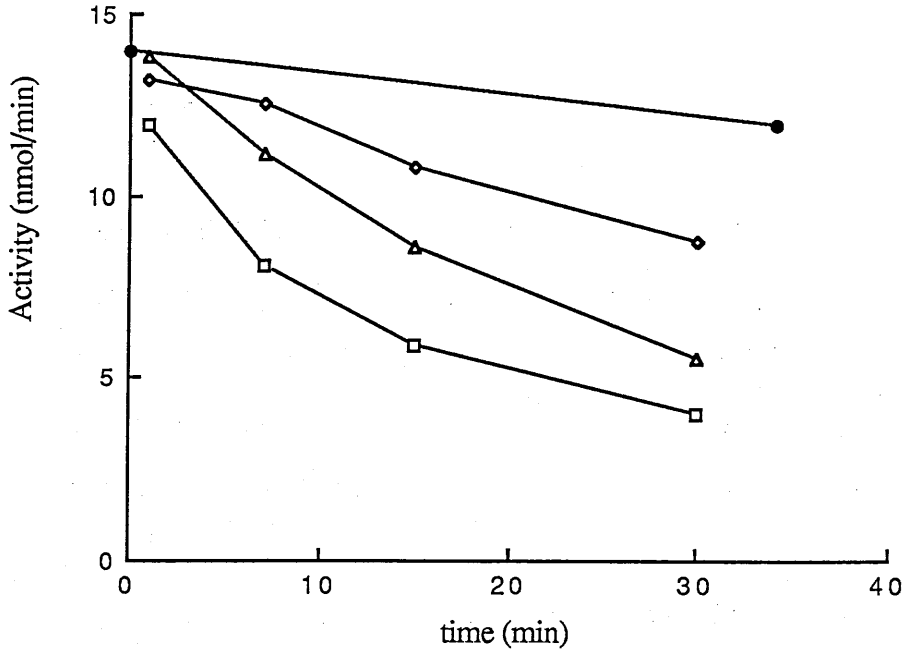
(a) Iodoacetate: control, ●; 10 mM, □; 5 mM, △; 1 mM, ◇.

(b) Iodoacetamide: control, ●; 10 mM, □; 1 mM, △; 0.5 mM, ◇.

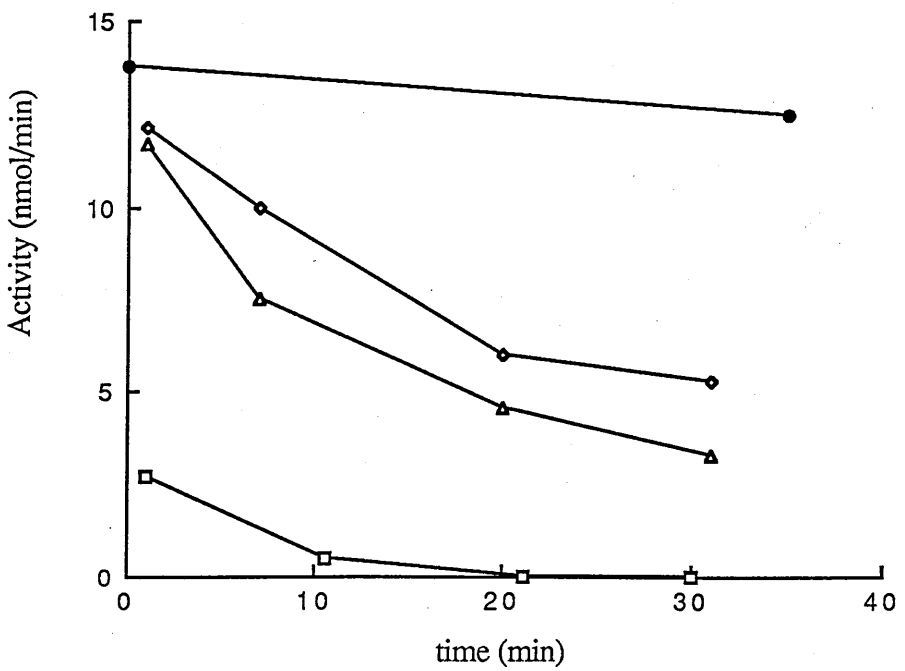
(c) 4-chloromercuribenzoate: control, ●; 100 μM, □; 50 μM, △; 30 μM, ◇.

(d) *N*-ethylmaleimide: control, ●; 100 μM, □; 50 μM, △; 10 μM, ◇.

(a)



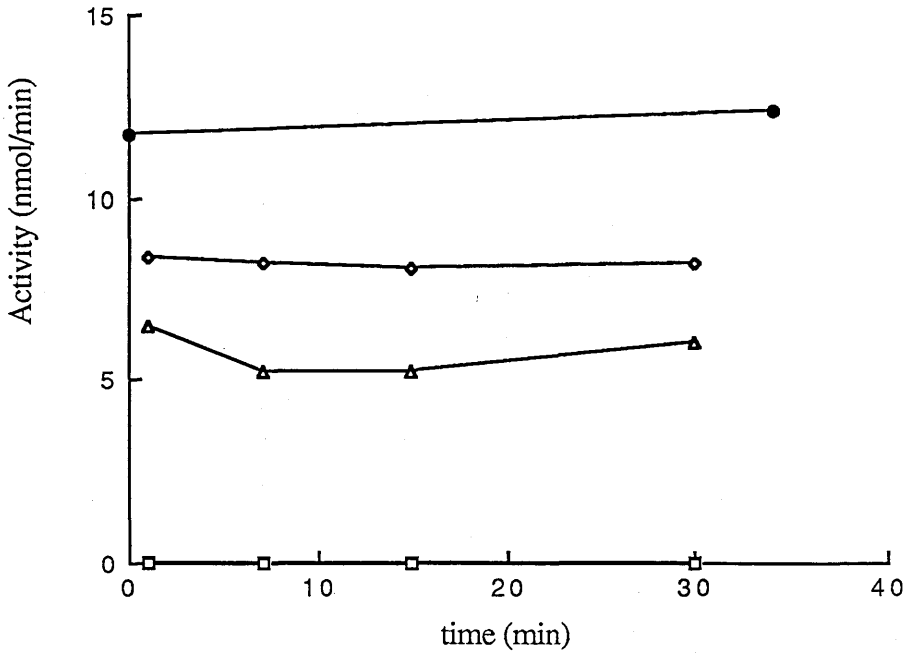
(b)



Continued:

Figure 5.13 Continued:

(c)



(d)

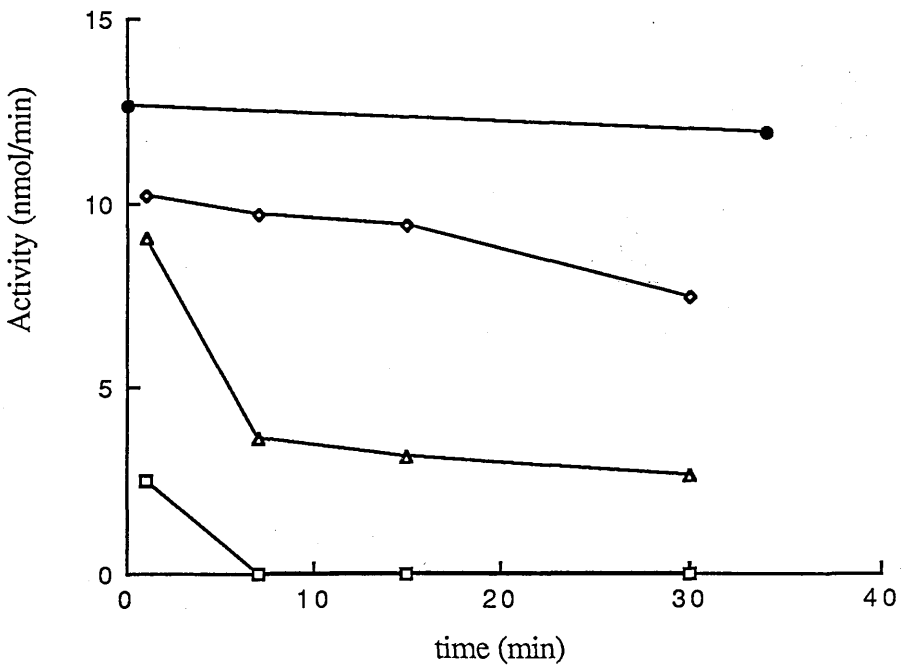


Table 5.12 Effects of thiol-blocking reagents on BZDH I

The time to reach 50% inhibition in various concentrations of thiol-blocking reagents was estimated from the plots in Figure 5.13.

Inhibitor	Inhibitor concentration (μM)	Time to reach 50% inhibition (min)
Iodoacetete	10 000	11
	5 000	23
	1 000	>30
Iodoacetamide	10 000	>1
	1 000	10
	500	17
4-Chloromercuribenzoate	100	<1
	50	4
	30	>30
N-Ethylmaleimide	100	<1
	50	4
	10	>30

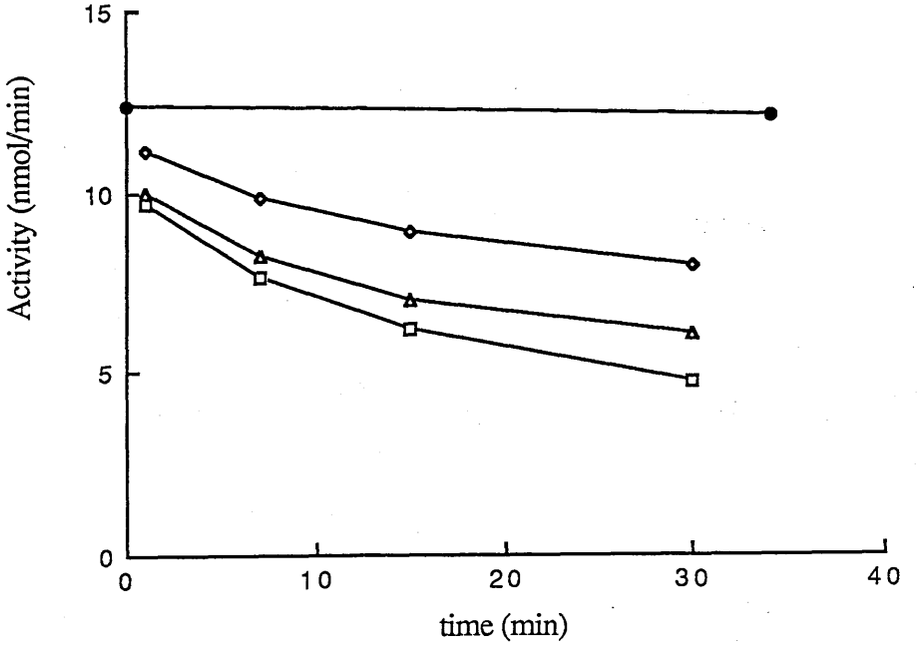
Figure 5.14 Protection of BZDH I from thiol-blocking reagents by substrate

The experiment was carried out exactly as described in Figure 5.13 except that benzaldehyde or NAD^+ was included in some of the incubation mixtures. Results are shown for a single representative experiment and other experiments gave similar results.

(a) Control, ●. 10 mM-iodoacetate, □ plus: 10 μM-benzaldehyde, △; 2 mM-NAD^+, ◇.

(b) Control, ●. 50 μM-*N*-ethylmaleimide, □ plus: 10 μM-benzaldehyde, △; 2 mM-NAD^+, ◇.

(a)


15
10
5
0
Activity (nmol/min)
0
10
20
30
40
time (min)

(b)

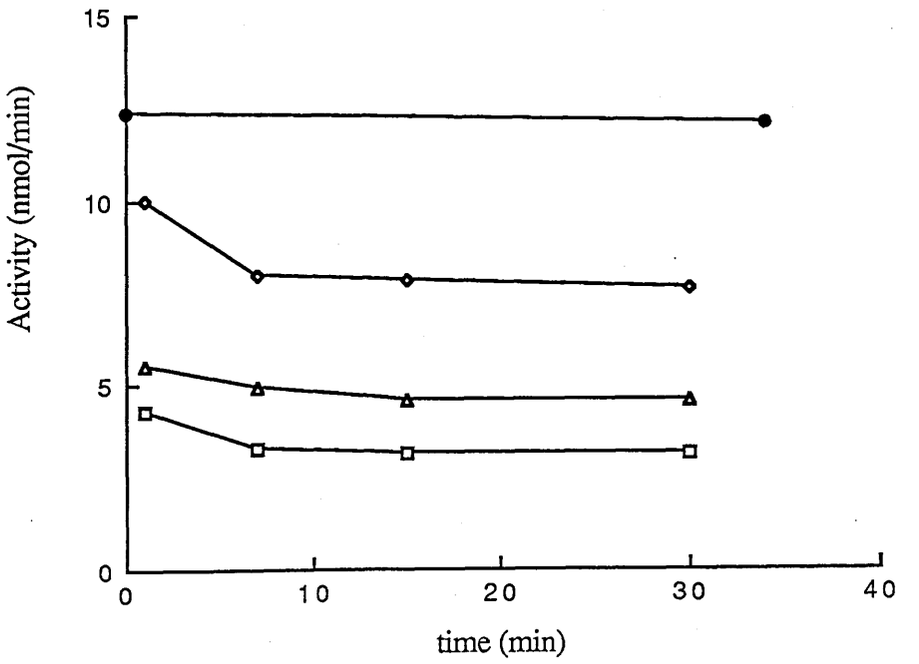

15
10
5
0
Activity (nmol/min)
0
10
20
30
40
time (min)

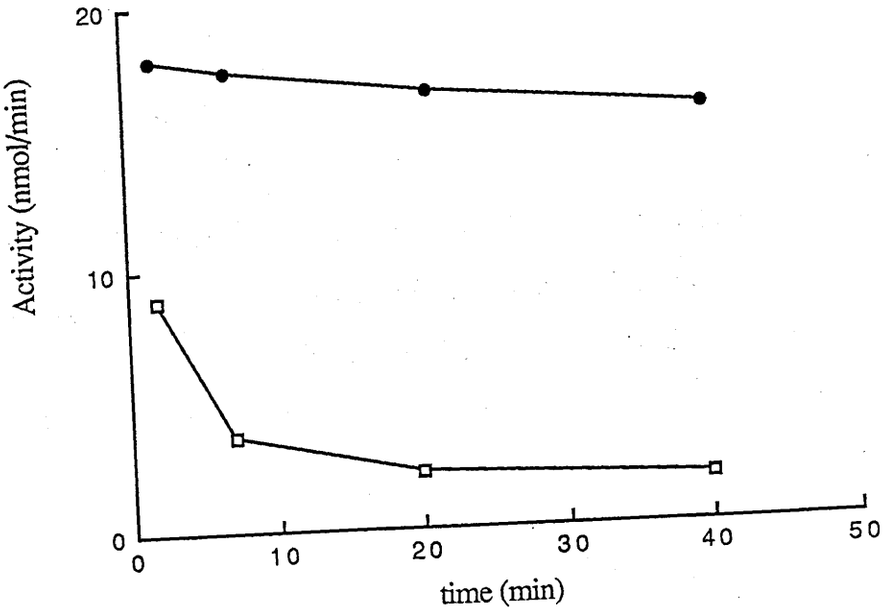
Figure 5.15 Effect of *N*-ethylmaleimide on BZDH I in the absence of DTT

The experiment was carried out exactly as described in Figure 5.13 and 5.14 except that BZDH I was first gel filtered (Methods 2.2.4) into DTT-free buffer (25 mM-potassium phosphate, pH 7.0) and the concentrations of BZDH I in the incubation and the assay mixtures were 11.2 and 0.56 μg/ml respectively. Except for the sample of enzyme incubated without thiol-blocking reagent (as a control), all of the assays in part (a) were carried out in triplicate and none of the values differed by more than 3%. Results are shown for a single representative experiment and other experiments gave similar results.

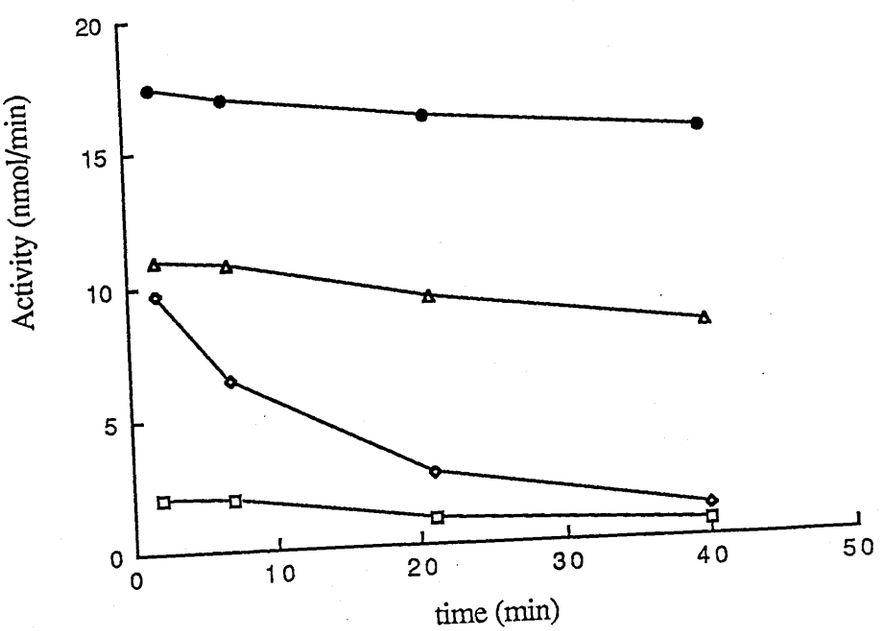
(a) Control, ●; 10 μM-*N*-ethylmaleimide, □.

(b) Control, ●. 50 μM-*N*-ethylmaleimide, □ plus: 1 mM-benzaldehyde, △
2 mM-NAD^+, ◇.

(a)


20
10
0
0
10
20
30
40
50
Activity (nmol/min)
time (min)

(b)


20
15
10
5
0
0
10
20
30
40
50
Activity (nmol/min)
time (min)

5.5 Discussion

5.5.1 M_r values, numbers of amino acid residues per subunit and subunit structures

BZDH I, BZDH II and TOL-BZDH run in almost identical positions on SDS-PAGE (Figure 5.16) and have subunit M_r values close to 56 000 (Section 5.1.1; MacKintosh & Fewson, 1988a). This is in good agreement with the molecular mass value for the benzaldehyde dehydrogenase subunit encoded by the archetypal TOL plasmid pWWO (Section 1.3) which has been determined to be 57 000 Da using *E. coli* maxicells (Harayama *et al.*, 1989).

The subunit M_r values of BZDH I, BZDH II and TOL-BZDH, together with their amino acid compositions (Section 5.1.5), were used to calculate the numbers of amino acid residues per enzyme subunit, which were all very close to 530 (Section 5.1.5).

BADH was reported to have a subunit M_r value of 39 700 (MacKintosh & Fewson, 1988a) and TOL-BADH has a subunit M_r value of 43 000 (Section 5.5.1), yet they appear to run in almost identical positions ahead of the ovalbumin standard protein (M_r = 43 000) on 12.5% SDS-polyacrylamide gels (Figure 5.16). They also have almost identical subunit M_r values on the phosphate-buffered SDS-polyacrylamide gels used in the cross-linking experiments (Section 5.1.2; Table 5.2). The discrepancy between the estimates of the M_r values is due to the fact that the value for BADH was determined using 10%-polyacrylamide gels (MacKintosh & Fewson, 1988a), whereas that for TOL-BADH was determined using 12.5%-polyacrylamide gels (Section 5.1.1). On a plot of mobility against ln (M_r) for the calibration proteins after electrophoresis on a 12.5%-polyacrylamide gel, ovalbumin does not lie on the line of best fit through all of the points (Figure 5.1). Ovalbumin appears to have a lower mobility than is predicted from the calibration curve and its subunit M_r value of 43 000. Thus, TOL-BADH clearly has a higher mobility than ovalbumin on 12.5%-polyacrylamide gels (Figure 5.16) yet the calibration curve (Figure 5.1), which takes account of the mobilities of all of the

calibration proteins, indicates that both proteins have identical M_r values. When plots of mobility against ln (M_r) are constructed for the calibration proteins on 12.5%-polyacrylamide gels, the first and third heaviest of the proteins lie at some distance from the line of best fit (Figure 5.1). When similar plots are constructed for the same proteins on 10% polyacrylamide gels (MacKintosh, 1988) the two lightest of the proteins are shifted from the line of best fit. Apparently it is the variable behaviour of the calibration proteins on different polyacrylamide gels that caused the discrepancy between the reported M_r values of BADH and TOL-BADH.

It is not immediately obvious from the data discussed above which of the two subunit M_r values for the alcohol dehydrogenases is the most accurate and additional facts must therefore to be considered. First, the subunit molecular mass of the benzyl alcohol dehydrogenase encoded by the archaetypal TOL plasmid pWWO (Section 1.3) has been determined be 40 000 Da using *E. coli* maxicells (Harayama *et al.*, 1989). Although there are considerable differences between the TOL plasmids pWWO and pWW53 there is considerable conservation of restriction sites within the coding regions and it is likely that the benzyl alcohol dehydrogenases encoded by each plasmid are highly homologous (Keil *et al.*, 1987a,b). Secondly, the amino terminal sequences of BADH and TOL-BADH (Section 5.1.4) are consistent with the enzymes belonging to the superfamily of zinc-dependent alcohol/polyol dehydrogenases (to be discussed in Section 5.5.6.2). The complete amino acid sequences of 20 members of the superfamily have been determined and the numbers of amino acid residues per enzyme subunit range from 347 to 391 (Figure 5.19 and references therein). The number of amino acid residues in the TOL-BADH subunit (412) calculated using a M_r value of 43 000 (Section 5.1.5) is well outside of this range, whereas that calculated using a M_r value of 39 700 (381) is within the range.

Overall the data are consistent with both BADH and TOL-BADH having subunits with approximately 381 amino acid residues and M_r values of 39 700.

BZDH II has an apparent native M_r value of 222 500 by gel filtration (MacKintosh &

Fewson, 1988a), which is consistent with a tetrameric arrangement of the enzyme subunit in the native state. BZDH I and TOL-BZDH have apparent native M_r values of 141 000 - 237 000 and 122 000 - 176 000 respectively, depending upon which type of gel-filtration column is used (Section 5.1.1). The native and subunit M_r values are consistent with a trimeric or a tetrameric arrangement of enzyme subunits of BZDH I and a dimeric or trimeric arrangement of TOL-BZDH.

BADH and TOL-BADH have subunit M_r values close to 40 000 (see above; Section 5.1.1; MacKintosh & Fewson, 1988a). BADH has an apparent native M_r value of 155 000 (MacKintosh & Fewson, 1988a) which is consistent with a tetrameric arrangement of enzyme subunits. However, TOL-BADH has an apparent native M_r value of between 82 000 and 114 000, depending upon the type of gel filtration column used (Section 5.1.1), and this is consistent with a dimeric or trimeric arrangement of enzyme subunits.

The determinations of the apparent native M_r values were inconsistent and failed to indicate clearly the subunit arrangement of the enzymes. Many dehydrogenase enzymes have been shown to form dimers or tetramers but none, to my knowledge, has ever been shown to form trimers. In order to clarify the results of the M_r value determinations all five enzymes were cross-linked using dimethylsuberimidate (Section 5.1.2). When samples were analysed using SDS-PAGE four major bands appeared in all five enzyme preparations, corresponding in M_r value to monomers, dimers, trimers and tetramers (Figure 5.3). It is therefore most likely that all five enzymes are tetrameric but that BZDH I, TOL-BZDH and TOL-BADH run anomalously through the gel-filtration columns, particularly the f.p.l.c. Superose columns.

Aldolase was included as a control on each of the gels used to analysed the cross-linked enzyme samples although it is not shown on all of the photographs in Figure 5.3. Except for TOL-BZDH, all of the cross-linked enzyme samples produced a pattern similar to aldolase on the gels. TOL-BZDH was different because the major bands corresponding in M_r to trimers and tetramers were each split into several minor bands.

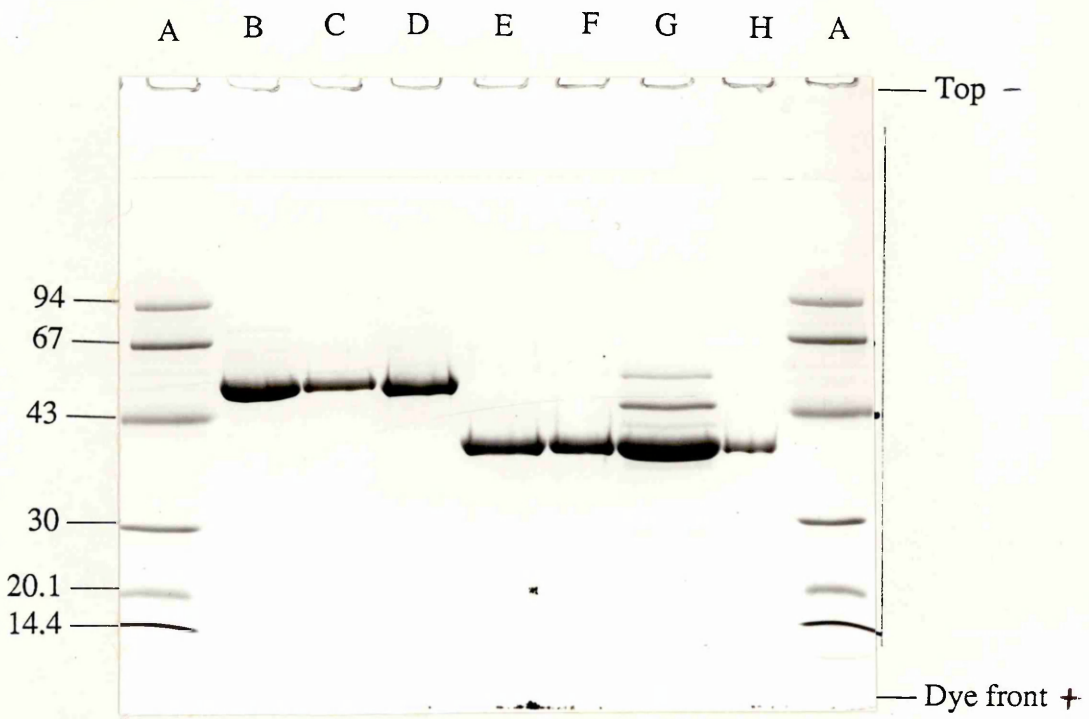


Figure 5.16 SDS-PAGE of all five enzymes

A denaturing 12.5% (w/v) polyacrylamide slab gel containing 0.1% SDS (Methods 2.7.1) was stained for protein (Methods 2.7.4a). Tracks: A, M_r markers (approx. 1 µg of each); B, BZDH I (5 µg of protein); C, BZDH II (2 µg of protein); D, TOL-BZDH (4 µg of protein); E, BADH (3 µg of protein); F, BADH (3 µg of protein); G, partially purified BADH (7 µg of protein); H, TOL-BADH (2 µg of protein).

This effect confused the interpretation of the gels to an extent that depended upon the amount of protein loaded in the individual tracks. This was because when a large amount of protein was loaded the trimer and tetramer bands appeared as a single broad band, and when a small amount of protein was loaded the minor bands appeared to stand alone. Overall, the impression gained from studying tracks loaded with various amounts of protein was of two major bands each consisting of a number of minor bands. Furthermore, the mobilities of the major bands were those expected of trimers and tetramers (Figure 5.4; Table 5.2).

5.5.2 Effects of pH and cations

The pH optima of the dehydrogenase activities of BZDH I, BZDH II and TOL-BZDH were almost identical and they were very close to pH 9.5 (Section 5.2.2; MacKintosh & Fewson, 1988a). These values are higher than those for aldehyde dehydrogenases from some other sources; for example the $NADP^+$-specific aldehyde dehydrogenase of *P. putida* (pH 8.5; Stachow *et al.*, 1967) and the yeast K^+-activated aldehyde dehydrogenase (pH 8.0; Bostian & Betts, 1978a).

The isoelectric point of BZDH I was pH 5.5 (Section 5.1.3), whereas the values for BZDH II and BADH were 4.6 and 5.0 respectively (MacKintosh & Fewson, 1988a).

BZDH I, in contrast to BZDH II (MacKintosh & Fewson, 1988a), was activated by monovalent cations with the maximum activity at approximately 70 mM-K^+ (Section 5.2.5). In common with the yeast aldehyde dehydrogenase (Dickinson & Haywood, 1987) and the $NADP^+$-specific aldehyde dehydrogenase from *P. putida* (Stachow *et al.*, 1967), which are also K^+-activated, Rb^+ and NH_4^+ were good substitutes for K^+ (Section 5.2.5). BZDH I and the yeast aldehyde dehydrogenase are both activated more than 100-fold by K^+, whereas the $NADP^+$-specific aldehyde dehydrogenase is activated only 3-fold. Furthermore, the extents of activation of BZDH I and the yeast enzyme were in the order, $K^+ > Rb^+ > NH_4^+$, whereas the order for the *P. putida* enzyme was,

$K^+ = NH_4^+ > Rb^+$. None of the enzymes were activated by Na^+ which has an atomic radius substantially different from K^+, Rb^+ and NH_4^+ (Section 5.2.5; Stachow *et al.*, 1967; Dickinson & Haywood, 1987). In contrast to the yeast (Dickinson & Haywood, 1987) and the *P. putida* enzymes (Stachow *et al.*, 1967) which are very unstable in the absence of K^+, the stability of BZDH I (at 4 ^{0}C) is not lowered in the absence of K^+, nor is there a detectable lag phase (>30 s) when K^+-depleted enzyme is assayed in the presence of K^+ (Section 5.2.5).

Dickinson & Haywood (1987) suggested that the role of the metal ion in the catalytic mechanism of the yeast aldehyde dehydrogenase is possibly to increase the rate of the acyl-enzyme hydrolysis step and that the properties of the lag phase when K^+-depleted enzyme is assayed in the presence of K^+ are consistent with a K^+-induced conformational change of the enzyme. Springham & Betts (1973) have also suggested that the K^+-dependent stability of the enzyme is due to a conformational change when K^+ binds. BZDH I and the yeast aldehyde dehydrogenase have similar ratios of activation by K^+, Rb^+ and NH_4^+, however BZDH I is stable when depleted of K^+ and there is no detectable lag phase when K^+-depleted enzyme is assayed in the presence of K^+ (Section 5.2.5). These differences between the properties of each of the enzymes are inconsistent with identical roles for the metal ion in each of the catalytic mechanisms.

Subsequent work in this laboratory by Mr. A. J. Scott and Professor C. A. Fewson has shown that the purified TOL-BZDH is also activated by K^+, Rb^+ and NH_4^+ but not by Na^+.

The pH optima for the oxidation of benzyl alcohol by BADH and TOL-BADH were both very close to pH 9.2 (Section 5.2.2; MacKintosh & Fewson, 1988a) and neither enzyme is activated by monovalent cations (MacKintosh & Fewson 1988a; C. A. Fewson, personal communication).

5.5.3 Cofactor specificity

BZDH I, BZDH II, BADH and TOL-BADH were relatively specific for NAD^+ and had only a limited ability to use $NADP^+$ (Section 5.2.6; MacKintosh & Fewson, 1988a). In contrast, TOL-BZDH was able to use both NAD^+ and $NADP^+$ as electron acceptors (Section 5.2.6). In this respect TOL-BZDH resembles the yeast K^+-activated aldehyde dehydrogenase (Bostian & Betts, 1978b). Worsey & Williams (1975) noted that in extracts of *P. putida* mt-2 the BZDH encoded by the TOL plasmid pWW0 (Section 1.3) was able to use $NADP^+$ at approximately 10% of the rate of NAD^+, when the concentration of the cofactors was 2.5 mM in the assay mixtures.

5.5.4 Substrate specificity

BZDH I was inhibited by low concentrations of benzaldehyde with maximum activity at approximately 4 - 5 μM (Section 5.3.1). The K'_m for benzaldehyde was 0.69 μM (Section 5.3.3). Several other aldehyde dehydrogenases have substrate K'_m values below 1 μM and are substrate-inhibited (Jacoby, 1963); these include BZDH II (MacKintosh & Fewson, 1988a), TOL-BZDH (Sections 5.2.1) and the $NADP^+$-specific aldehyde dehydrogenase from *P. putida* (Stachow *et al.*, 1967).

The ability of BZDH I to oxidise a range of compounds, including many substituted benzaldehydes, was tested (Section 5.3.2). The enzyme was more active when benzaldehyde was substituted at the *meta* and *para* positions of the aromatic ring than when it was substituted at the *ortho* position. The results were broadly similar to those for BZDH II (MacKintosh & Fewson, 1988b) except that BZDH I did not oxidise either of the aliphatic aldehydes tested. Many aldehyde dehydrogenases, including the yeast K^+-activated enzyme (Bostian & Betts, 1978b), accept both aromatic and aliphatic aldehydes as substrates.

In extracts, the BZDH encoded by the TOL plasmid pWW0 (Section 1.3) used the

meta and *para* monomethyl-substituted benzaldehydes at approximately 60% of the rate with benzaldehyde. Likewise, the BADH encoded by the same plasmid was able to use the *meta* and *para* monomethyl-substituted benzyl alcohols at approximately 60% and 70% respectively of the rate with benzyl alcohol (Worsey & Williams, 1975). The K'_m and V'_{max} values were determined for BZDH I, BZDH II and TOL-BZDH with NAD^+, benzaldehyde and the three possible monomethyl substituted benzaldehydes (Section 5.3.2). The V'_{max} values for BZDH I with each of the substrates were in the order benzaldehyde > 3-methylbenzaldehyde > 4-methylbenzaldehyde > 2-methylbenzaldehyde (Table 5.10). In contrast, the rate of the reactions when the substrates were tested at a fixed concentration (100 μM) were in the order, 4-methylbenzaldehyde > 3-methylbenzaldehyde > benzaldehyde > 2-methylbenzaldehyde (Table 5.9). This difference was probably caused by substrate inhibition of the enzyme by benzaldehyde (Section 5.3.1) and 3-methylbenzaldehyde. Values of K'_m and V'_{max} were also determined for BADH and TOL-BADH with NAD^+, benzyl alcohol and the three possible monomethyl substituted benzyl alcohols (Table 5.10).

The V'_{max} and K'_m values were used to calculate the apparent specificity constants (V'_{max}/K'_m) for each of the enzymes with each of the substrates (Table 5.11). A high value of the specificity constant (K_{kat}/K_m) is thought to indicate complementarity between the active site of the enzyme and the transition-state analogue of the substrate (Fersht, 1984). The specificity constant should therefore indicate how well an enzyme is adapted to utilise a particular substrate.

The relative values of the apparent specificity constants of BZDH I, BZDH II and TOL-BZDH, with benzaldehyde and the three monomethyl substituted benzaldehydes, were broadly similar (Figure 5.17). Benzaldehyde was the substrate with the highest specificity constant for all three enzymes. BZDH I was exceptional because it had a relatively low specificity constant for 4-methylbenzaldehyde and TOL-BZDH differed from the other two enzymes because the specificity constant for 2-methylbenzaldehyde was particularly low.

Figure 5.17 Relative values of the apparent specificity constants for unsubstituted and monomethyl-substituted substrates

The apparent specificity constants (Table 5.10) for each of the enzymes with each of the substrates are presented as a percentage of the apparent specificity constants for the unsubstituted substrates.

(a) Benzaldehyde, 

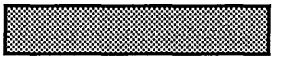
2-Methylbenzaldehyde, 

3-Methylbenzaldehyde, 

4-Methylbenzaldehyde, 

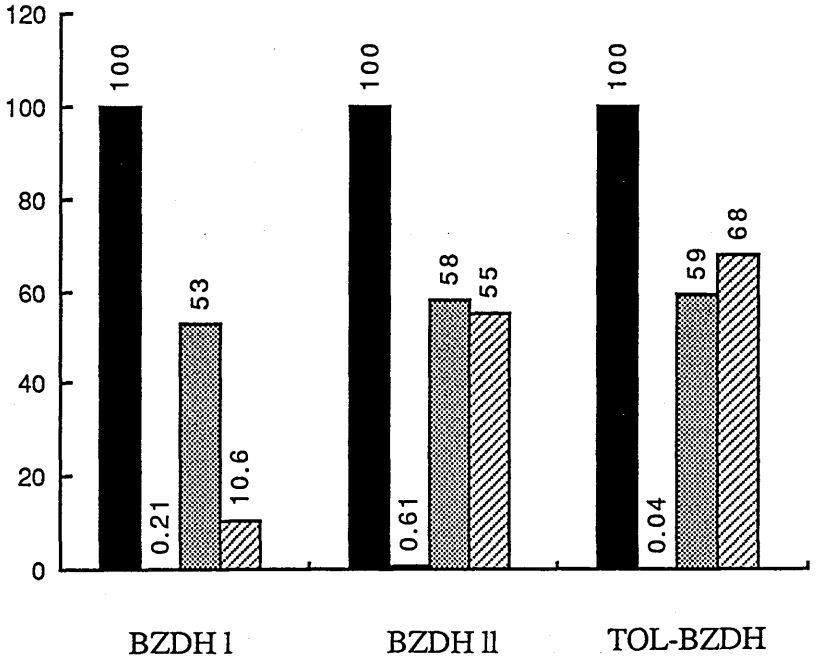
(b) Benzyl alcohol, 

2-Methylbenzyl alcohol, 

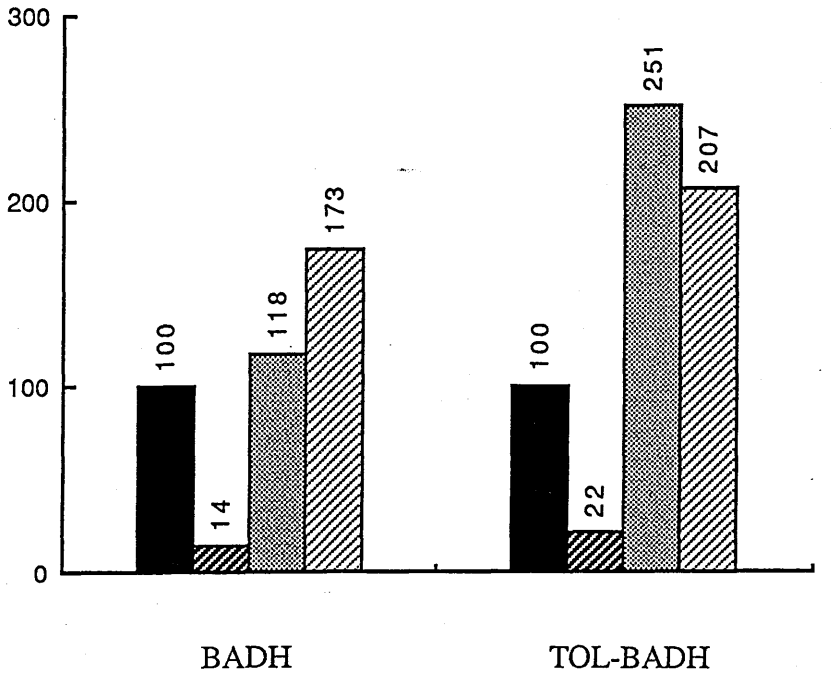
3-Methylbenzyl alcohol, 

4-Methylbenzyl alcohol, 

(a)


120
100
80
60
40
20
0
100
0.21
53
10.6
100
0.61
58
55
100
0.04
59
68
BZDH I
BZDH II
TOL-BZDH

(b)


300
200
100
0
100
14
118
173
100
22
251
207
BADH
TOL-BADH

The natural substrates of TOL-BZDH are 3- and 4-methylbenzaldehyde (Burlage *et al.*, 1989) and it is remarkable that the enzyme is not more highly adapted to their use (as judged by the apparent specificity constants) than BZDH I and BZDH II from *A. calcoaceticus*, which is unable to grow on either of these substrates (Kennedy & Fewson 1968a). Presumably this indicates that either the oxidation of substituted benzaldehydes is not the rate limiting step in the pathway encoded by the TOL plasmid or that the enzymes have been acquired too recently for adaptation to have taken place. Overall, the relative values of the apparent specificity constants for TOL-BZDH resemble the values for BZDH II more closely than they resemble those for BZDH I.

The relative values of the apparent specificity constants for BADH and TOL-BADH with benzyl alcohol and the three possible monomethyl substituted benzyl alcohols were different (Figure 5.17). The substrates that gave the highest specificity constants with BADH and TOL-BADH were 4-methylbenzyl alcohol and 3-methylbenzyl alcohol respectively.

5.5.5 Esterase activity and the effect of thiol-blocking reagents

Both BZDH I and BZDH II had esterase activity with 4-nitrophenol acetate as substrate (Section 5.3.4). The rates at pH 8.5 were each approximately 2% of the dehydrogenase activities, which is in good agreement with the rate previously determined under similar conditions for BZDH II (MacKintosh & Fewson, 1988b). It was impractical to measure esterase activity accurately at the pH optimum of the dehydrogenase activities (pH 9.5) because of the spontaneous hydrolysis of the substrate.

The esterase activities of BZDH I and BZDH II were both activated by low concentrations of NAD^+ (Section 5.3.4). BZDH II resembled the yeast K^+-activated aldehyde dehydrogenase (Dickinson & Haywood, 1987) in that maximum activation was achieved with approximately 10 μM-NAD^+, whereas the NAD^+-activation of BZDH I

was still increasing at 16 μM-NAD^+. Activation of the esterase function of these enzymes by NAD^+ is consistent with a single active site for both the esterase and the dehydrogenase activities. In contrast, the esterase activities of bovine lense and rabbit liver aldehyde dehydrogenases are both inhibited by preincubation with NAD^+ (Ting & Crabbe, 1983).

Many mammalian aldehyde dehydrogenases have been shown to have esterase activity with 4-nitrophenol acetate and there has been some debate as to whether a single active site thiol group is involved in both the esterase and the dehydrogenase activities (Tu & Weiner, 1988). In general, the kinetic data are consistent with a single active site, while the results of chemical modification studies are most compatible with distinct active sites for the esterase and the dehydrogenase reactions (Tu & Weiner, 1988). However, the most unequivocal evidence supports a single active site model; the acyl-enzyme intermediate, formed by reaction of the enzyme with an activated ester, can be reduced by NADH to produce free enzyme and the corresponding aldehyde (Loomes & Kitson, 1986).

Certain aldehyde dehydrogenases may have two esterase sites; the first site is involved in both dehydrogenase and esterase reactions, while the second site is involved only in ester hydrolysis. It is unclear whether or not the putative second active site has a specific function and it is possible that it consists of no more than a single reactive thiol group or the remnant of an ancestral active site (Tu & Weiner, 1988).

BZDH I was inhibited by a number of thiol-blocking reagents (Section 5.4). It was similar to BZDH II in being far more sensitive towards 4-chloromercuribenzoate and *N*-ethylmaleimide than towards iodoacetate or iodoacetamide (Section 5.4; MacKintosh & Fewson, 1988b). BZDH I also resembled BZDH II in that considerable protection from inhibition was afforded by the presence of benzaldehyde (Section 5.4; MacKintosh & Fewson, 1988b).

Jacoby (1963) proposed that the reaction mechanism of aldehyde dehydrogenases involved a thioester acyl-intermediate formed by a nucleophilic attack on the aldehyde

substrate. The protection of BZDH I and BZDH II by substrate from inhibition by thiol-blocking reagents and their NAD^+-activated esterase activities is consistent with this type of mechanism.

5.5.6 Amino terminal sequences

Of all the experimental methods used to assess possible evolutionary relationships among proteins (Section 1.5), amino acid sequence comparison is the most powerful because: **1**, it is generally applicable since amino acid sequences are now available for a large number of proteins and can also be rapidly determined; **2**, it is a sensitive method capable of discerning relationships among distantly related proteins; and **3**, the significance of the results of comparisons can be assessed using statistical techniques.

There are 20 possible character states for each position in an amino acid sequence, corresponding to the 20 common amino acids. If two random sequences, of unbiased amino acid composition, are compared they will therefore have, on average, identical residues at 5% of positions. The probability that two sequences are related can be assessed by comparing the number of identical positions with the normal distribution curve for the numbers of identities between pairs of random sequences of the same length. This type of significance test works well for closely related sequences; however, there are two other important factors that must be taken into account when comparing distantly related sequences. Firstly, the substitution of amino acids in evolving proteins is not random and there are great differences in the observed frequencies of the 180 possible amino acid point substitutions. This is because only substitutions that are permitted by natural selection are observed. Secondly, distantly related sequences are often observed to have insertions or deletions which confuse statistical tests because every position in one sequence does not have a counterpart in the second sequence.

Scoring matrices have been developed to take account of the different frequencies of substitution of different amino acid pairs. Some of these are based on the degeneracy of the genetic code (Needleman & Wunsch, 1970) or on apparent chemical similarities

among amino acids (Haber & Koshland, 1970), while still others are based on the observed frequencies of substitutions in various families of proteins (Haber & Koshland, 1970; McLachlan, 1971; Dayhoff *et al.*, 1978). At present the most commonly used scoring matrix for distantly related proteins appears to be the PAM 250 matrix which is based on the observed frequency of substitutions in 71 groups of closely related proteins (Dayhoff *et al.*, 1978). When two sequences have identical residues at the same position a score is assigned which depends upon the type of amino acid residue; for example, matching cysteines score +119 whereas matching leucines score only +59. When the residues at any particular position are different the score depends upon the observed frequency of substitution; for example leucine/isoleucine substitutions are frequent in related proteins and score +24, whereas cysteine/leucine substitutions are rare and score -60. The sum of the scores for each position in each pair of sequences is thus a more sensitive indication of an evolutionary relationship than is the number of identical amino acid residues.

All modern methods of sequence comparison allow the introduction of gaps to accommodate the insertion and deletion of genetic material that has occurred during the evolutionary process. It is necessary to penalise gaps with a negative score otherwise almost any two sequences would appear to be homologous after the insertion of a sufficient number of gaps. The gap penalty is usually chosen arbitrarily and this complicates any statistical analysis of the alignment among sequences.

In the following sections (Sections 5.5.6.1 & 5.5.6.2) the sequence homologies among BZDH I, BZDH II and TOL-BZDH, as well as those between BADH and TOL-BADH, are discussed. The sequences are short enough to be aligned by eye and recourse to computer aided techniques was not necessary. The sequence identity between the alcohol dehydrogenases is so extensive that a statistical analysis of the alignment is not necessary, however the aldehyde dehydrogenase sequence alignments require statistical analysis. This was carried out using the Monte Carlo analysis provided by the computer program RDF2 (Pearson & Lipman, 1988). This program finds the best

alignment between two sequences using the PAM 250 scoring matrix (see above) and the LFASTA algorithm (Pearson & Lipman, 1988). RDF2 then repeatedly shuffles the first sequence and realigns it with the second sequence. The program provides a score for the aligned sequences together with the highest score for the shuffled sequences and the mean score and standard deviation for the shuffled sequences. The significance of the alignment is judged by the number of standard deviations that the alignment score is above the mean of the shuffled scores. This would be a rigorous test of significance if the shuffled scores had a normal distribution. Unfortunately this is not always the case and in addition a qualitative judgement is required based upon the difference between the alignment score and the highest of the shuffled scores, together with the number of shuffles carried out.

5.5.6.1 Aldehyde dehydrogenase amino terminal sequences

The amino terminal sequences of BZDH I, BZDH II and TOL-BZDH (Table 5.3) were used in a computer search of the U.K. Protein Engineering Club's OWL (Version 6.0) data base (Akrigg *et al.*, 1988) which contains more than 19 000 amino acid and translated DNA sequences. The data base was searched using the SWEEP program (Version 3.02) and no biologically meaningful homologies were found. Furthermore, the data base contained 25 aldehyde and semialdehyde dehydrogenases none of which were picked out.

The amino terminal sequences of BZDH I, BZDH II and TOL-BZDH were aligned with each other so as to maximise the numbers of identical residues and minimise the number of gaps (Figure 5.18). In each of the three pairwise alignments, among all three enzymes, three gaps were introduced. When all three enzymes were aligned simultaneously (Figure 5.18d) there were a total of eight gaps. The inclusion of so many gaps in sequences as short as these complicates the assessment of whether or not the sequence homologies are consistent with evolutionary relationships among the proteins.

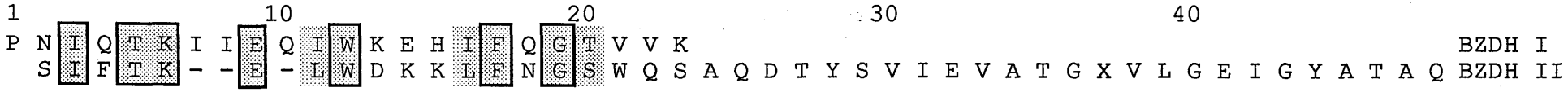
This is because the gap penalty (Section 5.5.6) is set to an arbitrary value that to some extent depends upon the prejudices of the experimenter. In order to eliminate the uncertain influence of so many gaps the residues between positions 1 and 11 of the alignment in Figure 5.18d were excluded from statistical analysis of the alignments. The remaining sequences, although considerably shorter, contain only one gap. The truncated sequences have the following levels of identity: BZDH I/BZDH II, 23%; BZDH I/TOL-BZDH, 31%; BZDH II/TOL-BZDH, 28%. The sequence alignments were analysed using the RDF2 program (Section 5.5.6) and the results are shown in Table 5.13. The alignment of the BZDH I sequence with BZDH II and TOL-BZDH had a score that was between two and three standard deviations above the mean of the shuffled alignments, however in both cases the highest shuffled score exceeded the alignment score. The alignment of the BZDH II sequence with TOL-BZDH had a score almost nine standard deviations above the mean of the shuffled scores and the highest of the shuffled scores was considerably lower than the alignment score. Overall, these results indicate that there is possibly a genuine evolutionary relationship between BZDH II and TOL-BZDH but that the homology between BZDH I and each of the other two enzymes (Figure 5.18) may be due to chance. However, the results should be considered in their biological context and there are a number of reasons why the possibility of evolutionary relationships among BZDH I and the other two enzymes should not be ruled out. Firstly, the amino terminal sequence of BZDH I is short and when it is truncated to eliminate the gaps introduced near the beginning of the sequence it is only 13 residues long. This length represents only 2.5% of the complete sequence (Section 5.1.5) and it seems unreasonable to base any conclusions on such a small region. Secondly, the BZDH I subunit is probably 525 residues in length (Section 5.1.5) and the available sequence, which extends for only 23 residues, is unlikely to include any of the active site residues which are likely to be more highly conserved than the amino terminal residues which possibly contribute very little to the three-dimensional structure of the subunit.

Figure 5.18 Aldehyde dehydrogenase sequence alignments

The amino terminal sequences of BZDH I, BZDH II and TOL-BZDH (Table 5.3) are aligned with each other. The sequences are presented in an alignment that is consistent throughout all of the comparisons and gives the maximum number of identities (boxed and shaded residues), the maximum number of conservative replacements (shaded residues) and the minimum number of gaps.

Conservative replacements are based on the PAM 250 scoring matrix (Dayhoff *et al.*, 1978) and are: I-L-V; D-E; K-R; T-S.

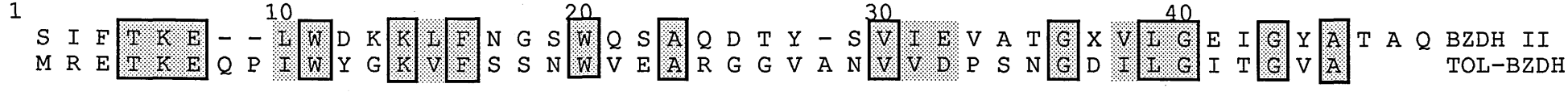
(a)


1 10 20 30 40
P N I Q T K I I E Q I W K E H I F Q G T V V K BZDH I
S I F T K - - E - L W D K K L F N G S W Q S A Q D T Y S V I E V A T G X V L G E I G Y A T A Q BZDH II

(b)


1 10 20 30 40
P N I Q T K I I E Q - I W K E H I F Q G T V V K BZDH I
M R E T K - - E Q P I W Y G K V F S S N W V E A R G G V A N V V D P S N G D I L G I T G V A TOL-BZDH

(c)


1 10 20 30 40
S I F T K E - - L W D K K L F N G S W Q S A Q D T Y - S V I E V A T G X V L G E I G Y A T A Q BZDH II
M R E T K E Q P I W Y G K V F S S N W V E A R G G V A N V V D P S N G D I L G I T G V A TOL-BZDH

(d)

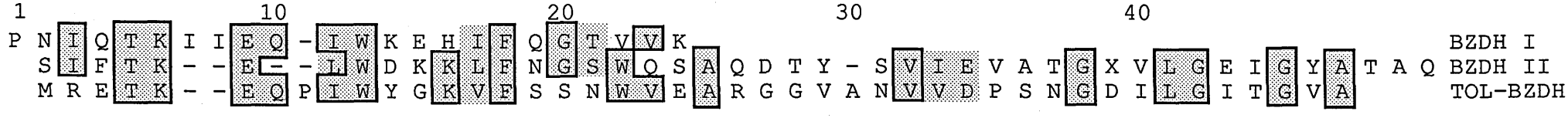

1 10 20 30 40
P N I Q T K I I E Q - I W K E H I F Q G T V V K BZDH I
S I F T K - - E - I W D K K L F N G S W Q S A Q D T Y - S V I E V A T G X V L G E I G Y A T A Q BZDH II
M R E T K - - E Q P I W Y G K V F S S N W V E A R G G V A N V V D P S N G D I L G I T G V A TOL-BZDH

Table 5.13 Statistical analysis of the aldehyde dehydrogenase sequences

The alignment of the amino terminal sequences of BZDH I, BZDH II and TOL-BZDH from Figure 5.18d were analysed using the RDF2 computer program (Pearson & Lipman, 1988). Residues between positions 1 and 11 were eliminated from the analysis because of the number of gaps that were required to maximise the number of identical residues. Each sequence was shuffled 1 000 times and the *ktup* factor was set to 1. Shuffles were carried out using the 'window' mode to maintain local amino acid composition and all other variables were at their default values.

S.D., standard deviation

Comparison	Alignment score	Number of S.D. above mean of shuffled scores	Highest shuffled score
BZDH I/BZDH II	37	3.4	40
BZDH I/TOL-BZDH	31	2.1	42
BZDH II/TOL-BZDH	74	8.9	51

5.5.6.2 Alcohol dehydrogenase amino terminal sequences

The amino terminal sequences of BADH and TOL-BADH (Table 5.3) were used in a search of the OWL sequence data base in an identical way to the aldehyde dehydrogenase sequences (Section 5.5.6.1). The search revealed that the amino terminal sequences of BADH and TOL-BADH share approximately 30% identity with chicken liver alcohol dehydrogenase and sheep liver sorbitol dehydrogenase. Reference to the relevant literature revealed that these enzymes were members of a superfamily of long-chain, zinc-dependent alcohol/polyol dehydrogenases that are found in mammals, higher plants, yeasts and bacteria (e.g. Jornvall *et al.*, 1987a). The amino terminal sequences of BADH and TOL-BADH were aligned with those of horse liver alcohol dehydrogenase, sheep liver sorbitol dehydrogenase and four other members of the superfamily from bacterial sources (Figure 5.19). All seven of the sequences share approximately 25% identity over the regions compared, although certain pairs of enzymes differ considerably from this value. For example, BADH and TOL-BADH have 36% identity, whereas TOL-BADH and the alcohol dehydrogenase from *Thermoanaerobium brockii* have only 8% identity.

At present the zinc-dependent alcohol/polyol dehydrogenase superfamily contains 16 eukaryotic alcohol dehydrogenases, three bacterial alcohol dehydrogenases (in addition to BADH and TOL-BADH) and sheep liver sorbitol dehydrogenase (Jornvall *et al.*, 1987a and references in Figure 5.19). Another zinc-dependent enzyme that may also be a member of the superfamily is alcohol dehydrogenase-1 from *Zymomonas mobilis*. The sequence of its first 31 amino terminal residues is known (Neale *et al.*, 1986) but apart from the first five residues (which are identical to those of the *B. stearothermophilus* enzyme; Figure 5.19) there is no significant homology with established members of the superfamily.

Members of the zinc-dependent alcohol dehydrogenase superfamily share amino acid sequence identities that vary between approximately 10% and 90%. The patterns of

relationships among the enzymes have been used to develop an evolutionary history for the mammalian members of the superfamily (e.g. Jornvall *et al.*, 1987b). The relationships among the human members of the family are illustrated schematically in Figure 5.20, together with their relationships to the other major groups within the superfamily. There have been at least three levels of gene duplication in the mammalian line. The first duplication event (Level 1 in Figure 5.20) led to the divergence of the tetrameric sorbitol dehydrogenase and the dimeric mammalian alcohol dehydrogenase from their common ancestor . The second duplication event (Level 2 in Figure 5.20) involved at least two individual duplications and produced the three classes of mammalian alcohol dehydrogenases that are commonly distinguished by their relative sensitivities to inhibition by pyrazole (Julia *et al.*, 1988). The third and final duplication event (Level 3 in Figure 5.20) also involved at least two individual duplication events and produced the α, β and γ isoenzyme subunits of the human Class 1 alcohol dehydrogenase. The first two levels of gene duplication are common among different mammalian species, however the third and most recent level of duplication occurred after the divergence of the mammals and is different in different species (Jornvall *et al.*, 1987b).

The three-dimensional structure of horse liver alcohol dehydrogenase (type E subunit) has been solved at 2.4 Å resolution (Eklund *et al.*, 1976) and the primary structures of all of the other eukaryotic members of the superfamily can be accommodated by the model (Jornvall *et al.*, 1987b). This is surprising because these structures include dimeric and tetrameric alcohol dehydrogenases and a sorbitol dehydrogenase which have amino acid sequence identities that range approximately between 25% and 90%. Local differences in these structures together with the three-dimensional model can be used to explain differences in the physical and catalytic properties of the enzymes, however because of the great divergence of the sequences additional three-dimensional structures are required to gain a complete understanding of the properties of the members of the superfamily (Jornvall *et al.*, 1987b).

Figure 5.19 BADH and TOL-BADH are members of a superfamily of alcohol dehydrogenases

The amino terminal sequences of BADH and TOL-BADH (Table 5.3) are compared with each other and with the corresponding sequences of other selected members of the same superfamily of alcohol dehydrogenases. The sequences are presented in an alignment which is consistent throughout all of the comparisons and gives the maximum number of identities (boxed and shaded residues) and the minimum number of gaps. The secondary structure of the horse liver ADH is indicated above the sequence and is taken from the complete three-dimensional structure which has been solved at 2.4 Å resolution (Eklund *et al.*, 1976).

The comparisons are presented as four different diagrams because the box patterns in a single diagram with all seven sequences are so complicated that the best alignment of the sequences is very difficult to find. The residues indicated by asterisks are conserved among 17 enzymes from mammals, higher plants and yeasts (Jornvall *et al.*, 1987a). Horse liver ADH (type E subunit) and sheep liver sorbitol dehydrogenase sequences are from Jornvall *et al.* (1987a); *Bacillus stearothermophilus* ADH sequence is from Bridgen *et al.* (1973) and Jeck *et al.* (1979); *Alcaligenes eutrophus* ADH sequence is from Jendrossek *et al.* (1988); *Thermoanaerobium brockii* ADH sequence is from Peretz & Burstein (1989).

ADH, alcohol dehydrogenase; HLADH, horse liver ADH (Type E subunit); SDH, sorbitol dehydrogenase; BSADH, *Bacillus stearothermophilus* ADH; AEADH, *Alcaligenes eutrophus* ADH, TBADH, *Thermoanaerobium brockii* ADH.

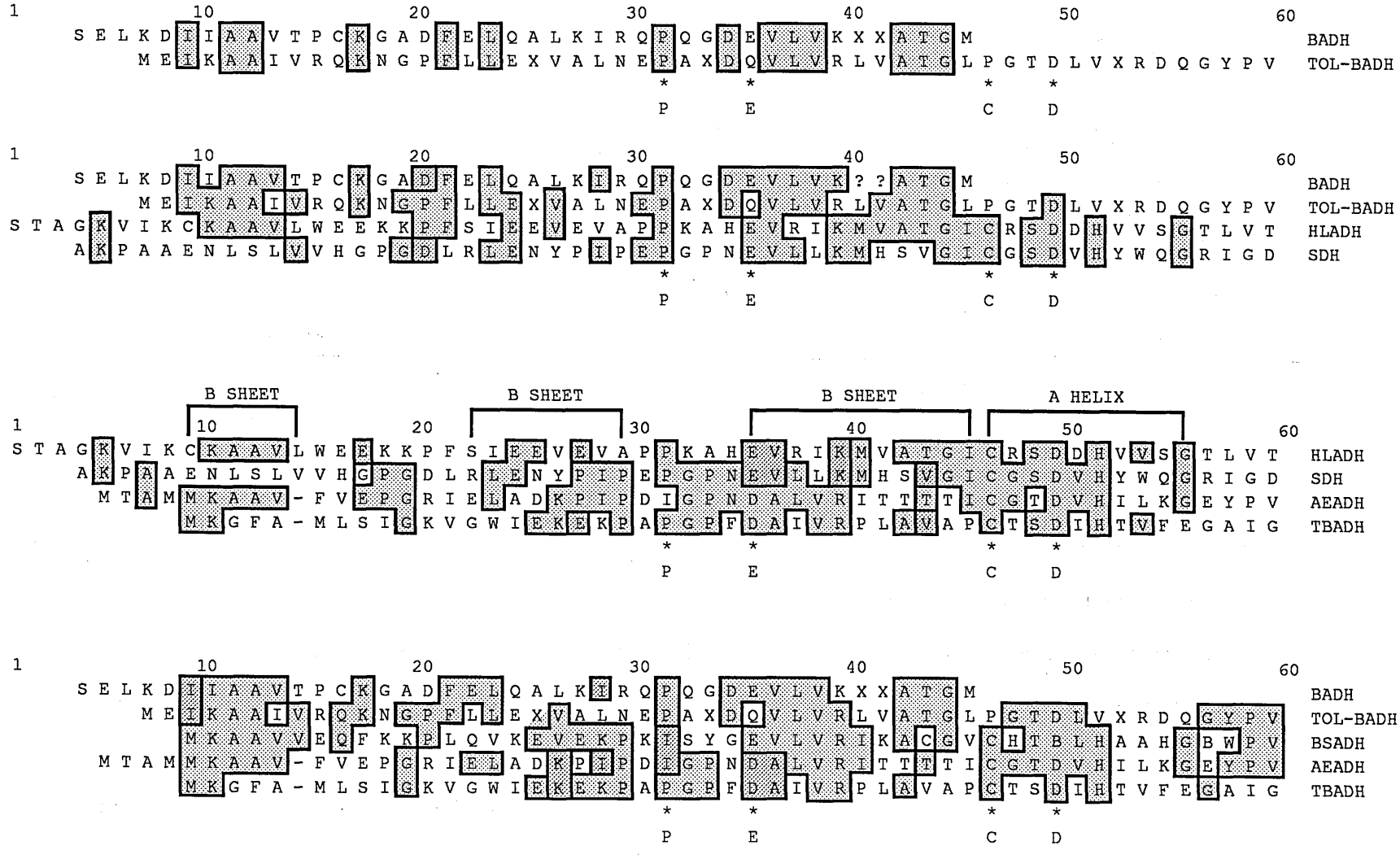
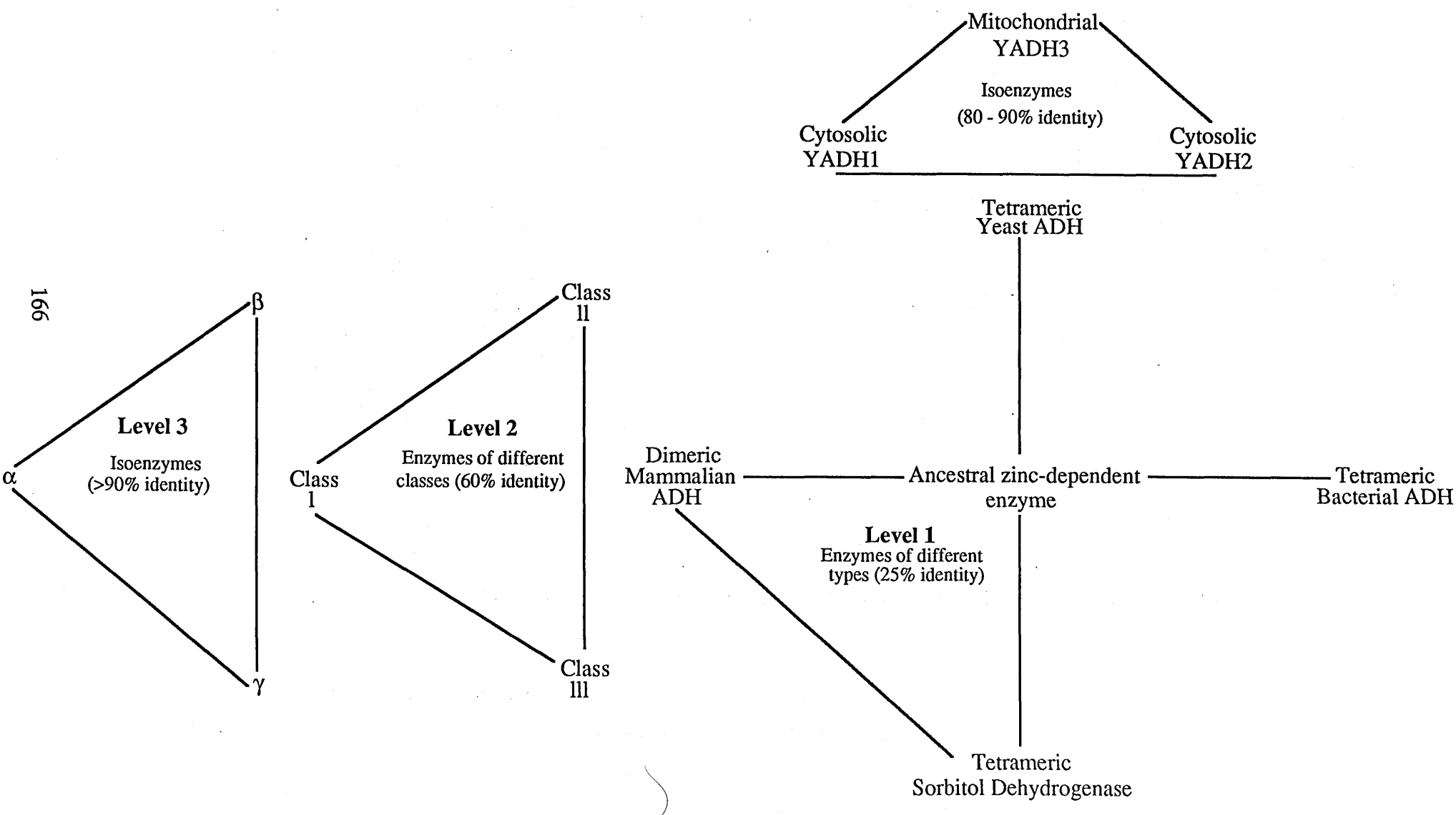

1 10 20 30 40 50 60
S E L K D I I A A V T P C K G A D F E L Q A L K I R Q P Q G D E V L V K X X A T G M BADH
M E I K A A I V R Q K N G P F L L E X V A L N E P A X D Q V L V R L V A T G L P G T D L V X R D Q G Y P V TOL-BADH
P E C D
S T A G K V I K C K A A V L W E E K K P F S I E E V E V A P P K A H E V R I K M V A T G I C R S D D H V V S G T L V T HLADH
A K P A A E N L S L V V H G P G D L R L E N Y P I P E P G P N E V L L K M H S V G I C G S D V H Y W Q G R I G D SDH
B SHEET
B SHEET
B SHEET
A HELIX
M T A M M K A A V - F V E P G R I E L A D K P I P D I G P N D A L V R I T T T T I C G T D V H I L K G E Y P V AEADH
M K G F A - M L S I G K V G W I E K E K P A P G P F D A I V R P L A V A P C T S D I H T V F E G A I G TBADH
M K A A V V E Q F K K P L Q V K E V E K P K I S Y G E V L V R I K A C G V C H T B L H A A H G B W P V BSADH

Figure 5.20 Evolutionary history of the human long-chain alcohol dehydrogenases

Triangles join duplicated forms existing in parallel in one organism; ADH, alcohol dehydrogenase. Based on Jornvall *et al.* (1987b).


Mitochondrial
YADH3
Isoenzymes
(80 - 90% identity)
Cytosolic
YADH1
Cytosolic
YADH2
Tetrameric
Yeast ADH
β
α
γ
Level 3
Isoenzymes
(>90% identity)
Class
ll
Class
l
Class
lll
Level 2
Enzymes of different
classes (60% identity)
Dimeric
Mammalian
ADH
Ancestral zinc-dependent
enzyme
Tetrameric
Bacterial ADH
Level 1
Enzymes of different
types (25% identity)
Tetrameric
Sorbitol Dehydrogenase

The horse liver alcohol dehydrogenase subunit consists of an amino terminal catalytic domain and a carboxy terminal coenzyme-binding domain (Eklund *et al.*, 1976). The catalytic domain binds two zinc atoms but only one of these is involved in catalysis. The catalytic zinc atom is located 20 Å from the surface in a cleft between the catalytic and the coenzyme-binding domains. It is ligated by two sulfur atoms from Cysteine46 and Cysteine174 and one nitrogen atom from Histidine67. A water molecule, which is involved in a system of hydrogen bonds with Serine48 and Histidine51, completes the tetrahedral co-ordination of the catalytic zinc atom. The three amino acid zinc ligands are conserved in almost all of the members of the superfamily for which there is unambiguous sequence data available (Figure 5.19 and references therein). The amino terminal sequence data for BADH only extends to the residue immediately before the first zinc ligand, however the extensive sequence identities with other members of the superfamily between residues 34 and 44 are consistent with the expectation of a cysteine residue at position 46 (Figure 5.19). The residue at position 46 in the TOL-BADH sequence is designated as proline (Figure 5.19). Only one of the two sequencing runs carried out with TOL-BADH extended into this region and proline at this position must be considered a tentative assignment because of the relatively low recovery (see Cycle 40 in Table 5.3). There are a considerable number of sequence identities between TOL-BADH and other members of the superfamily in the regions on either side of this position and it is very likely that the residue at this position is in fact cysteine and not proline.

Serine48 in the horse liver alcohol dehydrogenase is one of the residues involved in hydrogen bonding the water molecule that completes the tetrahedral co-ordination of the catalytic zinc atom (see above). This residue is either conserved or is replaced by threonine in all other members of the superfamily including TOL-BADH (Figure 5.19 and references therein). The second residue that hydrogen bonds the active site water molecule, Histidine51, is replaced by valine in TOL-BADH and by serine or tyrosine in various other members of the superfamily (Figure 5.19 and references therein).

In addition to the zinc ligand Cysteine46 and the two residues involved in hydrogen bonds with the active site water molecule the amino terminal sequence of horse liver

alcohol dehydrogenase has three other residues that are highly conserved in other members of the superfamily. Proline31, Glutamate35 and Aspartate49 are conserved within the 17 members of the superfamily from mammals, higher plants and yeasts (Figure 5.19; Jornvall *et al.*, 1987a). Proline31 and Glutamate35 are both conserved in BADH, however TOL-BADH has glutamine at position 35 (Figure 5.19). Aspartate49 is conserved in all members of the superfamily including TOL-BADH (Figure 5.19 and references therein). This residue binds to the catalytic zinc ligand Histidine67 and this function is presumably the basis for its conservation in all of the related structures examined to date (Eklund *et al.*, 1976; Figure 5.19 and references therein).

The three-dimensional structure of a triclinic ternary complex between horse liver alcohol dehydrogenase, NADH and the inhibitor dimethylsulfoxide has been solved at 2.9 Å resolution and compared to the structure previously determined for the apoenzyme (Eklund *et al.*, 1981). The coenzyme binding domain of the zinc-dependent alcohol dehydrogenases is located in the carboxy terminal half of the molecule, however the three-dimensional structure of the horse liver enzyme shows that Arginine47 and Histidine51 might be involved in coenzyme binding.

The coenzyme binding domains of the subunits interact to form a central compact core with the catalytic domain of each subunit on either side. The dimer has approximate dimensions of 40 Å x 55 Å x 100 Å and the catalytic domains are situated at the ends of the longest side. When the apoenzyme binds NADH the catalytic domains appear to rotate 7.5 ^{0} with respect to the the central core of the dimer. Although some of the α-carbons at the surfaces of the catalytic domains move by more than 10 Å the zinc ligands are relatively unaffected because they are close to the hinge region. In the model of the ternary complex the guanidinium group of Arginine47 is very close to the phosphate group of the adenosine monophosphate part of the coenzyme and the side chain of Histidine51 is within hydrogen bond distance of the O-3' of the nicotinamide ribose (Eklund *et al.*, 1981). The contribution made to coenzyme binding by Arginine47 is doubtful because it is not well conserved in the superfamily. Most of the members have histidine at this position and some have glycine (Figure 5.19 and references therein). Within the superfamily Histidine51 is well conserved (see above), although it is

occasionally replaced by the chemically dissimilar residues tyrosine or serine and in TOL-BADH by valine (Figure 5.19 and references therein). As previously mentioned Histidine51 is also thought to be involved in a system of hydrogen bonds with a water molecule that ligates the catalytic zinc atom (Eklund *et al.*, 1976) and its conservation (15 out of 20 enzymes) can not therefore be used as a fact in support of a direct role in coenzyme binding.

There are two other established groups of alcohol dehydrogenases in addition to the zinc-dependent superfamily. The first, and more thoroughly characterised, is a group of metal-independent enzymes that have approximately 250 amino acid residues per subunit (Jornvall *et al.*, 1984). These are known as the short-chain alcohol dehydrogenases by analogy with members of the zinc-dependent superfamily which each have approximately 350 amino acid residues per subunit (Section 5.5.1). There are three established members of the short-chain alcohol dehydrogenase group; *Drosophila* alcohol dehydrogenase, *Klebsiella aerogenes* ribitol dehydrogenase and *B. megaterium* glucose dehydrogenase (Jornvall *et al.*, 1984). The second, and more recently discovered, additional group of alcohol dehydrogenases contains four enzymes; alcohol dehydrogenase-2 from *Z. mobilis* (Conway *et al.*, 1987), alcohol dehydrogenase-4 from *Saccharomyces cerevisiae* (Williamson & Paquin, 1987), propanediol oxidoreductase from *Escherichia coli* (Conway & Ingram, 1989) and the $NADP^+$-dependent butanol dehydrogenase from *Clostridium acetobutylicum* (Youngleson *et al.*, 1989). These enzymes have approximately 380 amino acid residues per subunit and are thought to be iron-dependent, although there is one report that alcohol dehydrogenase-4 from *S. cerevisae* is zinc-dependent (Drewke & Ciriacy, 1988).

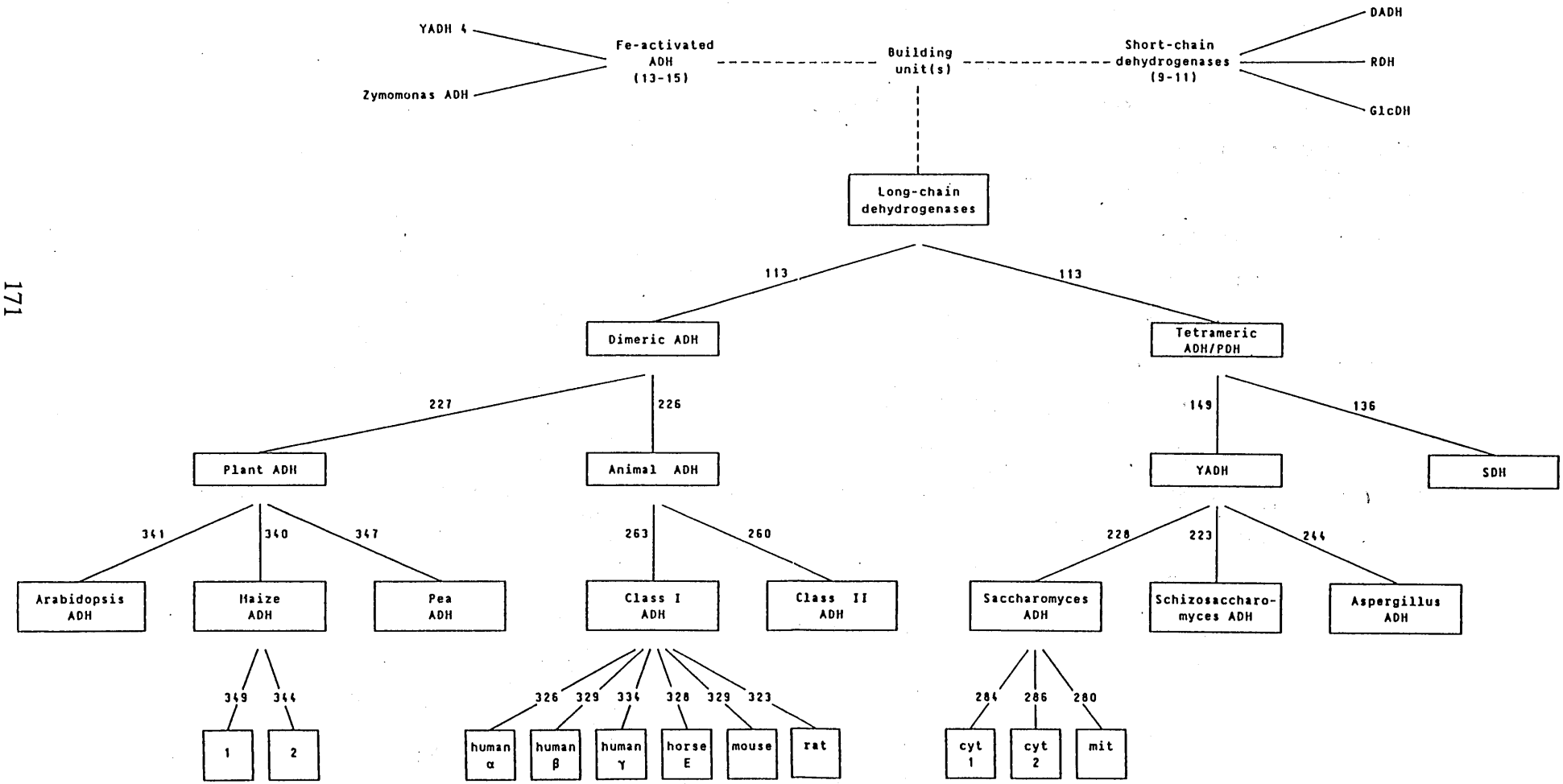
Overall, BADH and TOL-BADH appear to be members of the long-chain, zinc-dependent superfamily of alcohol dehydrogenases although there is no evidence from experiments using metal chelating agents that the putative active site zinc atom can be removed easily (MacKintosh & Fewson 1988b; C. A. Fewson personal communication). BADH and TOL-BADH appear to be unrelated to either the iron-dependent or the metal-independent groups of alcohol dehydrogenases. The

complete amino acid sequences of 19 members of the zinc-dependent superfamily are known (Figure 5.19 and references therein) and 17 of these, from mammals, higher plant and yeasts have been used to construct a family tree (Figure 5.21; Jornvall *et al.*, 1987a). The family tree, which has two major branches representing the dimeric and the tetrameric enzymes, was constructed by deducing the ancestral amino acid sequences using the modern sequences and the accepted phylogeny of the organisms concerned. Rough estimates of the rate of evolutionary change in the superfamily are of the order of about 10 accepted point mutations per 100 residues per 100 million years irrespective of which part of the family tree is used (Jornvall *et al.*, 1987a). Jornvall has suggested that this is consistent with a similar set of functional constraints acting on all of the enzymes throughout their evolutionary history, which is interesting because neither the exact metabolic role nor the principle substrates of the enzymes are established (Jornvall *et al.*, 1987a).

The functional roles of some of the bacterial members of the superfamily are established (references in Figure 5.19) but they are not represented in the present family tree (Figure 5.21). BADH and TOL-BADH have similar functional roles and principle substrates. Both enzymes have similar substrate specificities (Section 5.3.3) and are members of peripheral catabolic pathways that degrade a range of aromatic compounds (Sections 1.2 & 1.3). Judging by this limited amount of information it appears that the properties of BADH and TOL-BADH have been maintained despite substantial evolutionary change; using the calculations of Dayhoff *et al.* (1978), 160 amino acid substitutions per 100 residues would be required to account for the observed sequence differences between BADH and TOL-BADH (assuming that the amino terminal regions are typical of the whole of the molecules). If the metabolic roles and substrate specificities of BADH and TOL-BADH have been maintained throughout this amount of evolutionary change then (neglecting the possibility of convergent evolution or gene transfer) it is feasible that Jornvalls' suggestion (implied above) is correct and all of the eukaryotic members of the superfamily have similar metabolic roles.

Figure 5.21 The zinc-dependent alcohol dehydrogenase family tree

The long-chain zinc-containing forms are the ones presently outlined, short-chain forms without zinc and a third line with iron-activated forms, as well as possible interconnections (dashed) with ancestral building unit(s) are indicated (top) to complete the relationships but are not further detailed. Values for the subunit relationships of the zinc-containing enzymes show residue identities (in subunits of 347-391 residues). Except for the present day structures at the bottom level, these estimates involve comparisons with deduced consensus structures. At each level, residues of these deductions follow the most abundant residue of the alternatives at the level below (or, if non-conclusive from those alternatives, but invariable in other branches, the residues in the latter). The top values obtained are identical because of the method of consensus construction and give an impression of the long-chain ancestor as intermediate between the dimer and tetramer descendents. Only positions with identities in consensus descendents are presently counted, and remaining positions, not yet possible to judge, may modify the final relationships. Branch lengths are not drawn to scale. ADH, alcohol dehydrogenase; YADH, yeast-type alcohol dehydrogenase; DADH, *Drosophila* alcohol dehydrogenase; PDH, polyol dehydrogenase; SDH, sorbitol dehydrogenase; GlcDH, glucose dehydrogenase; RDH, ribitol dehydrogenase; α, β, γ, subunits of class 1 human ADH; E, subunit of horse Class 1 ADH; cyt, cytosolic; mit, mitochondrial; Saccharomyces, *Saccharomyces cerevisiae*; *Schizosaccharomyces pombe*; Aspergillus, *Aspergillus nidulans*; Zymomonas, *Zymomonas mobilis*. From Jornvall *et al.*, (1987a).


YADH 4
Zymomonas ADH
Fe-activated ADH (13-15)
Building unit(s)
Short-chain dehydrogenases (9-11)
DADH
RDH
GlcDH
Long-chain dehydrogenases
113
113
Dimeric ADH
Tetrameric ADH/PDH
227
226
149
136
Plant ADH
Animal ADH
YADH
SDH
341
340
347
263
260
228
223
244
Arabidopsis ADH
Maize ADH
Pea ADH
Class I ADH
Class II ADH
Saccharomyces ADH
Schizosaccharo-myces ADH
Aspergillus ADH
349
344
326
329
334
328
329
323
284
286
280
1
2
human α
human β
human γ
horse E
mouse
rat
cyt 1
cyt 2
mit

5.5.7 Amino acid compositions

Comparisons of amino acid compositions can be used to infer homologies amongst proteins. Methods for pairwise comparisons have been developed by Marchalonis & Weltman (1971) and Cornish-Bowden (1979). The Marchalonis & Weltman method is based on calculating the quantity SΔQ which for any pair of proteins is the sum of the differences squared between the percent composition for each amino acid. They divided over 100 proteins of known amino acid composition into related or unrelated families and performed over 5,000 pairwise comparisons. In almost no cases did proteins thought to be related give a SΔQ value greater than 50 and in only 2% of comparisons did proteins thought to be unrelated give a value of less than 100. When plots were constructed of percent sequence difference verses SΔQ for related families of proteins of known sequence linear regression analysis by least squares revealed a significant correlation. The method of Cornish-Bowden is similar to that of Marchalonis & Weltman because it is based on calculating the sum of the differences squared between the numbers of each type of amino acid residue in each of two proteins. However, the method was developed using statistical theory and is based on several assumptions. The most important assumption is that the differences between the amino acid sequences of two closely related proteins occur at sites where there is a weak requirement for any particular type of amino acid and that the substitutions at these sites are random. This assumption is clearly not true, as Cornish-Bowden (1979) has pointed out, however the assumption may reflect the true behaviour of proteins closely enough for the method to be useful. When 163 pairwise comparisons were made between proteins of known sequence only seven of the comparisons indicated homology. When the same comparisons were made using the amino acid sequences of the proteins an additional 12 relationships were detected. When used to compare homologous proteins the method therefore appears less likely to falsely indicate homology than to falsely indicate lack of homology. Cornish-Bowden (1979) constructed a phylogenetic tree of 37 related snake venom toxins based upon pairwise

comparisons of their amino acid compositions. The resulting tree was the same "in most important respects" to the tree constructed on the basis of the known amino acid sequences of the proteins.

The amino acid compositions of BZDH I, BZDH II, TOL-BZDH, BADH and TOL-BADH were determined and compared in a pairwise fashion by the methods of Marchalonis & Weltman and Cornish-Bowden (Section 5.1.5). The method of Marchalonis & Weltman indicated that the three aldehyde dehydrogenases are homologous to each other, as are the two alcohol dehydrogenases. In contrast, the Cornish-Bowden method indicated that none of the aldehyde dehydrogenases are homologous to each other. The Cornish-Bowden method also indicated that the the two alcohol dehydrogenases are not homologous to each other, however $S\Delta n$ was closer to its critical value than the precision of the duplicate amino acid composition determinations for BADH (Table 5.7).

The amino terminal sequences of the enzymes (Section 5.1.4 & 5.5.6) indicate that BADH and TOL-BADH are related, as are BZDH II, TOL-BZDH and possibly also BZDH I. Judging by the sequences, BADH and TOL-BADH are more closely related than are any of the three aldehyde dehydrogenases, and BZDH II and TOL-BZDH are more closely related to each other than to BZDH I. The results of the amino acid composition comparisons, using either of the two methods, are qualitatively consistent with these relationships. The comparative method of Cornish-Bowden (1979), which failed to indicate homology between any of the pairs of enzymes examined, appears to be a more conservative test of evolutionary relationships than is the method of Marchalonis & Weltman (1971).

This type of analysis is indirect and there is no way to establish whether apparent differences or apparent similarities are due to lack of sensitivity or to the chance convergence of the amino acid compositions. Furthermore, conclusions based on the comparison of amino acid compositions must be viewed with caution since pairs of proteins are known to satisfy the most stringent requirements of these methods while

lacking any discernable sequence homology (e.g. van Heyningen & Coulson, 1987).

5.5.8 Thermal inactivation of enzyme activity

The thermostability of protein structure is an interesting problem for two reasons. Firstly, an understanding of the forces that stabilise the native conformation of a protein illuminates the process of protein folding which is a fundamental biological process. Secondly, detailed knowledge of stabilising interactions facilitates attempts to engineer proteins with enhanced thermostability and this is potentially advantageous in the application of biotechnology.

Folded proteins are only marginally stable and the net difference between the free energies of the folded and unfolded forms is generally 5 - 20 Kcal/mol (Matthews, 1987). The small difference in free energy between the two forms, together with the large size and complexity of most polypeptides, makes detailed calculations of free energies very difficult. In general terms, folded proteins are stabilised by hydrophobic interactions among amino acids in the interior of the structure and destabilised by the difference between the entropy of the folded and the unfolded forms (Matthews, 1987).

In many protein structures there is a compromise between flexibility and rigidity. Allosteric regulation and the conformational changes that often occur on substrate binding require flexibility, while thermostability requires a rigid structure (Vihinen, 1987). Hydrogen-exchange experiments have shown that thermophilic enzymes are generally more rigid than mesophilic enzymes but at their respective temperature optima both types of enzyme have similar flexibilities (Vihinen, 1987).

In some cases (e.g. lysozyme) the unfolding of the three-dimensional structure at high temperatures can be reversed simply by cooling, although it is often necessary to control the redox potential and the rate of cooling so as to prevent the formation of non-native disulphide bridges or folding patterns (Ahern & Klibanov, 1988). There are a number of reasons why thermal inactivation may also be irreversible. The unfolded molecule may form aggregates because of interactions among hydrophobic regions that

are buried within the native structure (Ahern & Klibanov, 1988). Alternatively, aspartate-X peptide bonds may hydrolyse at high temperatures or essential thiol groups, that are protected in the native structure, may be exposed to oxidative damage upon unfolding (Bridgen *et al.*, 1973; Skerker & Clark, 1989).

The rates of thermal inactivation of BZDH I, BZDH II and TOL-BZDH were determined at 60 ^{0}C in the presence and absence of benzaldehyde and NAD^+ (Section 5.1.6). TOL-BZDH is slightly more stable than BZDH I and much more stable than BZDH II. Inclusion of either substrate in the incubation mixture had little effect on the stability of BZDH I or TOL-BZDH, whereas benzaldehyde greatly increased the stability of BZDH II. These results indicate that the three-dimensional structure of BZDH II is probably more flexible than those of BZDH I and TOL-BZDH. Furthermore, the stabilising effect of benzaldehyde on BZDH II is consistent with a conformational change on substrate binding that makes the structure more rigid and therefore less prone to unfolding.

The time course for the thermal inactivation of BZDH II was approximately exponential, however those for BZDH I and TOL-BZDH were almost arithmetic (Figure 5.5). Preliminary experiments established that this was not due to the destruction of DTT, benzaldehyde or NAD^+ during the time course of the experiment (Figure 5.5). With the small amount of data available on the thermal inactivation of these enzymes it is not possible at present to explain why the rates of inactivation did not have a classical exponential decay.

A more detailed investigation of the thermostability of these enzymes is be beyond the scope of this thesis (Section 1.6). The small amount of data presented is included largely because it confirms the original proposition of Kennedy & Fewson (1968b) that *A. calcoaceticus* N.C.I.B. 8250 has isofunctional benzaldehyde dehydrogenases that are differentially expressed under various growth conditions and have different thermostabilities. Their experiments with cell extracts showed that one benzaldehyde dehydrogenase was heat stable and K^+-activated while the other was heat labile and

K^+-independent. The work presented and referenced in this thesis confirms that the enzymes retain these properties after purification (Sections 5.1.6 & 5.2.5). Questions of more fundamental interest will arise if BZDH I and BZDH II do indeed prove to be homologous. The differences between the primary structures of the two enzymes might then help to explain why they differ so much in their thermostability.

CHAPTER 6

Immunological Relationships Amongst the Purified Benzaldehyde and Benzyl Alcohol Dehydrogenases from *Acinetobacter* and *Pseudomonas*

Some of the descriptions and interpretations presented in this chapter may appear to be rather long-winded. This is necessary because this work is important in supporting some of the conclusions drawn from the homologies among the amino terminal sequences of the aldehyde and alcohol dehydrogenases (Section 5.5.6), and also because the interdependence of the experiments must be fully explained. In order to make the text as concise as possible I have used the following terms:

1 inoculum - the sample of protein originally used to elicit the production of antibodies in the experimental animals
2 homologous antigen - this term is used with reference to a particular antiserum and means the protein against which the antiserum was raised
3 heterologous antigen - any antigen other than the homologous antigen even if both the homologous and heterologous antigens are homologous to each other with respect to their amino acid sequences

6.1 Preparation of antigens and antisera

In order to study the immunological relationships amongst BZDH I, BZDH II, TOL-BZDH, BADH and TOL-BADH, antisera were prepared from rabbit blood after the animals had been inoculated with highly purified samples of the enzymes (Methods 2.8.1).

6.1.1 Purity of the antigens

Immunological assays are among the most sensitive probes used in biological research and great care must be taken to ensure that the antigens used for inoculation are highly purified. The immunological response of an experimental animal may depend not only upon the amount of antigen injected but also upon its antigenicity. In theory, the

immunological response following a challenge could therefore be as strong towards a highly antigenic minor contaminant as towards the major component of the inoculum.

Minor contamination of antigen preparations does not represent a problem when the antisera are used to study cross-reactions between antigens from different sources. For example, if a cross-reaction was observed between BZDH II and TOL-BADH then, except for gross experimental errors, there would be no reason to suppose that this was due to cross-contamination of the original inocula because the enzymes were prepared from different organisms. However, if a cross-reaction was observed between two antigens purified from the same starting material, such as BZDH II and BADH which were both prepared from *A. calcoaceticus,* then there would exist the possibility that the inocula were cross-contaminated.

Before the enzymes were injected into the rabbits the aldehyde dehydrogenases were assayed for the presence of alcohol dehydrogenase activity and vice versa (Methods 2.5). The upper limits of contamination were approximately 1 part in 600 for alcohol dehydrogenase contaminating aldehyde dehydrogenase and approximately 1 part in 200 for aldehyde dehydrogenase contaminating alcohol dehydrogenase. In order to quantify the extent of any possible cross-contamination of the inocula by inactive enzyme different amounts of each sample were subjected to SDS-PAGE (Methods 2.7.1). The extent of any possible cross-contamination was then estimated by making careful comparisons between tracks of the gel lightly loaded with one enzyme and adjacent tracks heavily loaded with a second enzyme. Tracks heavily loaded with one enzyme were examined to see if there was a band corresponding in position to the band in the adjacent track lightly loaded with a second enzyme. The upper limits of cross-contamination of the preparations were estimated to be as follows: BADH contaminating BZDH I, 1 part in 7 000; BADH contaminating BZDH II, 1 part in 1 500; BZDH I contaminating BADH, 1 part in 150; BZDH II contaminating BADH, 1 part in 700; TOL-BADH contaminating TOL-BZDH, 1 part in 50 and TOL-BZDH contaminating TOL-BADH, 1 part in 50. The possible extent of cross-contamination between the *Acinetobacter* and the *Pseudomonas*

enzymes were not estimated for the reason given above.

BZDH I and BZDH II were not separated by SDS-PAGE (Figure 5.1.6) and the extent of any cross-contamination was estimated using non-denaturing PAGE with an activity stain (Methods 2.7.3 & 2.7.4c; Figure 3.4). The upper limits of cross-contamination were estimated to be: BZDH II contaminating BZDH I, 1 part in 55 and BZDH I contaminating BZDH II, 1 part in 350, as judged by activity.

6.1.2 Production of antisera

Antisera were prepared from New Zealand rabbits after the animals had been inoculated by subcutaneous injection of various amounts of the purified enzyme antigens (Methods 2.8.1; Table 6.1). Antisera were raised against the enzymes in their native states and also against the *Acinetobacter* enzymes after they had been denatured by the reduction and carboxymethylation of their cysteine residues (Methods 2.9.2). Before inoculation, samples of the denatured enzymes were subjected to amino acid analysis (Methods 2.10) to confirm the complete modification of their cysteine residues. All of the cysteine in each of the enzyme preparations was detected as carboxymethylcysteine.

6.2 Immunoprecipitation assays

The magnitude of the reactions amongst the antisera and their homologous and heterologous antigens were determined by immunoprecipitation assays with the antigens in their native states. Two types of immunoprecipitation assays were used. The first, which will simply be referred to as the immunoprecipitation assay, relied on the spontaneous formation of an insoluble antibody/antigen complex when optimum proportions of antibody and antigen were mixed. The second type of assay, which will be referred to as the *Staphylococcus aureus* immunoprecipitation assay, relied on the ability of *S. aureus* cells to bind (and therefore precipitate) antibody/antigen complex (e.g. Kessler, 1975). Neither of the immunoprecipitation assays was used to distinguish

between antibodies against structural and/or active site epitopes. However, in the immunoprecipitation assay the loss of enzyme activity before centrifugation indicated that some of the antibodies were directed against active site epitopes, or against structural epitopes that were necessary to maintain the conformation of the active site (Figure 6.1).

6.2.1 Cross-reactions detected by immunoprecipitation titrations

The reactions amongst the antisera and each of the five enzymes were quantified by immunoprecipitation assays (Methods 2.8.2). Briefly, a serial dilution of the antigen was prepared and an equal volume of each dilution was mixed with an equal volume of antiserum. The antiserum was either used undiluted or after dilution, depending upon the strength of the cross-reaction to be determined. After sufficient time for the formation of antibody/antigen complex the mixtures were centrifuged and the supernatants were assayed for enzyme activity. If an excess of antigen was added to the antiserum the supernatant retained some enzyme activity. If the amount of antigen was progressively reduced the point of equivalence was reached when all of the enzyme activity was removed from the supernatant. The titre of an antiserum (μg of protein/ml of antiserum) is defined as the amount of antigen added at the point of equivalence multiplied by the dilution of the antiserum in the mixture. An example of an immunoprecipitation titration is shown in Figure 6.1.

The titres of the antisera with respect to their homologous antigens varied considerably, ranging from 5 to 1440 μg/ml (Table 6.2). Only two heterologous cross-reactions were observed. These were between anti-BZDH I and BZDH II and between anti-TOL-BADH and BADH (Table 6.2). The reciprocal cross-reactions, between anti-BZDH II and BZDH I, and between anti-BADH and TOL-BADH, were not observed.

Table 6.1 Production of the antisera

The purified enzymes were each mixed with 1.0 ml of Freund's complete adjuvant. The mixtures were made up to 2.0 ml with water and briefly sonicated to form an emulsion. The mixtures were subcutaneously injected into white New Zealand rabbits at six sites. After six weeks the animals were boosted with further injections as described above, except that Freund's incomplete adjuvant was used. After a further two weeks the animals were bled.

CM-, before inoculation the antigen was denatured by the reduction and carboxymethylation of its cysteine residues (Methods 2.9.2).

Rabbit Code No.	Protein	1st Inoculation (μg of protein)	2nd Inoculation (μg of protein)
022	BZDH I	100	50
023	BZDH I	500	250
254	BZDH I	350	250
389	CM-BZDH I	900	450
314	BZDH II	500	250
251	BZDH II	500	250
390	CM-BZDH II	900	450
1079	TOL-BZDH	300	150
1080	TOL-BZDH	600	300
392	BADH	640	320
417	CM-BADH	900	450
1077	TOL-BADH	310	155
1078	TOL-BADH	620	310

Figure 6.1 An example of an immunoprecipitation titration experiment

Anti-BZDH I (Code No. 022; Table 6.1) was titrated with BZDH I. The antiserum was diluted 1 : 64 in the antiserum/antigen mixture. The titre of the antiserum (493 μg of antigen/ml of antiserum) is the concentration of the antigen at the point of equivalence (7.66 μg/ml) multiplied by the dilution of the antiserum (64).

Enzyme activity before centrifugation, □; enzyme activity of supernatant after centrifugation, △.

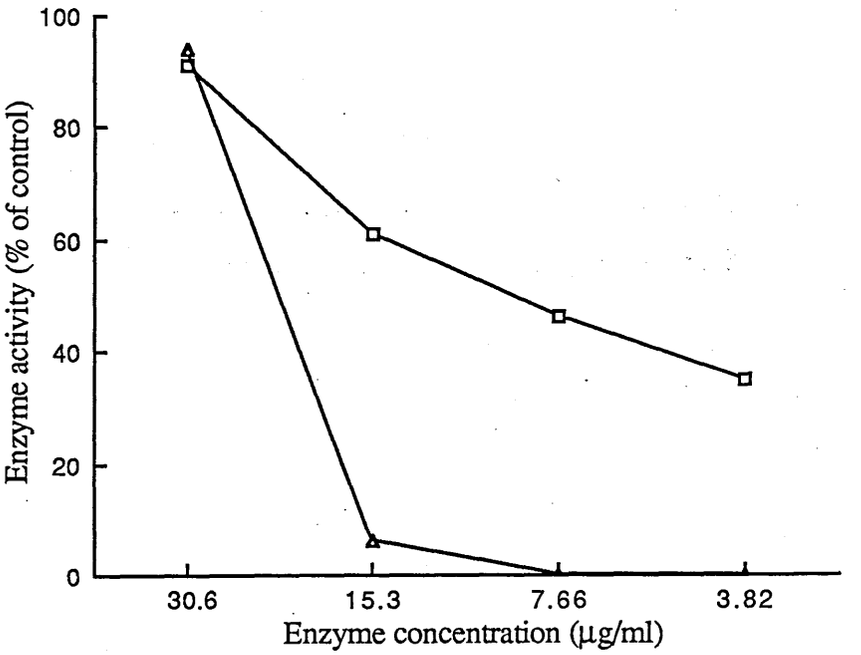

100
80
60
40
20
0
Enzyme activity (% of control)
30.6
15.3
7.66
3.82
Enzyme concentration (µg/ml)

Table 6.2 Immunoprecipitation titrations

All of the titrations (Methods 2.8.2) were carried out at least twice and the largest difference between duplicate determinations was 50%. The lowest values of duplicate determinations are recorded.

ND, not determined
*, antiserum raised against reduced and carboxymethylated antigen

		Antigen (μg bound/ml of antiserum)				
Antiserum	Code No.	BZDH I	BZDH II	TOL-BZDH	BADH	TOL-BADH
anti-BZDH I	022	493	75	0	0	0
	023	243	18	0	0	0
	254	1440	5	0	0	0
	389*	5	0	0	0	0
anti-BZDH II	314	0	30	0	0	0
	251	0	64	0	0	0
	390*	0	0	ND	0	ND
anti-TOL-BZDH	1077	0	0	230	0	0
	1078	0	0	230	0	0
anti-BADH	392	0	0	0	6	0
	417*	0	0	ND	0	ND
anti-TOL-BADH	1079	0	0	0	52	152
	1080	0	0	0	52	152

A single antiserum was raised against each of the *Acinetobacter* enzymes (BZDH I, BZDH II and BADH) after they had been denatured by the reduction and carboxymethylation of their cysteine residues (Table 6.1). The antiserum raised against denatured BZDH I was the only one of the three to react with the native homologous antigen. The titre of this antiserum was between 50 and 300-fold less than the titres of the three antisera raised against native BZDH I (Table 6.2).

In order to demonstrate that the cross-reaction between anti-BZDH I and BZDH II was not due to cross-contamination of the inoculum (see Section 6.1.1) a control experiment was devised. A sample of BZDH I was mixed with anti-BZDH I (code No. 022; Table 6.1). The amount of BZDH I used was equal to the titre of the antiserum (493 μg/ml; Table 6.2). The mixture was kept on ice for one hour and then microfuged (Methods 2.2.8) for five minutes. The supernatant lacked any detectable BZDH I activity thus indicating that the enzyme, and presumably the antibodies directed against it, had been removed from solution. The supernatant was then mixed with the smallest amount of BZDH II that was consistent with the ability easily to assay the enzyme activity. The mixture was again kept on ice for one hour and then four volumes of *S. aureus* cells (Methods 2.8.3a) were added. After a further 45 minutes on ice the mixture was microfuged (Methods 2.2.8) and the supernatant was assayed for BZDH II activity (Methods 2.5.1b). All of the BZDH II activity was found in the supernatant and none had been precipitated.

These experimental results can be interpreted in the following way. When anti-BZDH I was treated with BZDH I the antibodies and the enzyme were removed from the solution in the form of an insoluble antibody/antigen complex. The supernatant then lacked antibodies capable of binding BZDH II. Thus, it appears that anti-BZDH I contained a single population of antibodies some of which were directed against epitopes shared by BZDH I and BZDH II. If the original BZDH I inoculum had been cross-contaminated with BZDH II one would expect to detect a sub-population of antibodies, directed against BZDH II, which were not removed from solution by pre-treatment of the antiserum with BZDH I. It might be possible that a sub-population of antibodies,

directed only against BZDH II, was removed from the solution because it was somehow trapped within the antibody/BZDH I complex. This explanation of the experimental results was eliminated by repeating the experiment with a sample of anti-BZDH I (code No. 022; Table 6.1) which had been mixed with enough anti-BZDH II (code No. 314; Table 6.1) to double its titre towards BZDH II. The ability of the mixture to react with BZDH II was not abolished by pre-treatment with BZDH I, thus indicating that only antibodies directly involved in complex formation were removed from the solution.

6.2.2 *S. aureus* immunoprecipitation assays

The *S. aureus* immunoprecipitation assay (Methods 2.8.3) relied on the ability of *S. aureus* cells to bind immunoglobulin/antigen complexes (Kessler, 1975). The immunoglobulins are bound by protein A which is expressed on the cell surface. The assay does not depend on the formation of an insoluble antibody/antigen complex; therefore, in contrast to the immunoprecipitation assay (Section 6.2.1), the success of the assay does not depend on the relative proportion of antibody and antigen. Briefly, antiserum was mixed with enzyme solution and the mixture was incubated for one hour before *S. aureus* cells were added. After incubating for a further 45 minutes the mixture was centrifuged and the supernatant was assayed for enzyme activity.

No attempt was made to quantify the results of the *S. aureus* immunoprecipitation assays and they were used only to confirm the presence or absence of immunological reactions. All of the immunological comparisons made by immunoprecipitation assays (Table 6.2) were confirmed.

6.3 Quantitative immunoblotting

The reactions amongst the antisera and each of the five enzymes were quantified by immunoblotting. Immunoblotting was carried out by transferring the denatured antigens from SDS-polyacrylamide gels onto nitrocellulose membranes and also by direct

application of the antigens to the membranes.

6.3.1 Development of quantitative immunoblotting

Quantitative immunoblotting (Methods 2.8.4) was based on the methods described by Towbin *et al.* (1979) and Nimmo *et al.* (1986). Briefly, purified enzymes were subjected to SDS-PAGE and then electroblotted onto nitrocellulose membranes. Each membrane was first probed with the appropriate antiserum and then with [^{125}I] protein A. Radioactive bands (located by autoradiography) were cut out and counted.

Experiments were carried out to find conditions which would provide a linear relationship between the amount of purified protein loaded onto the gel and the number of counts per minute in the corresponding band on the nitrocellulose membrane. The following conditions were altered one at a time: gel running voltage, blotting voltage, concentration of antiserum, length of time for which the antiserum and the membrane were incubated together, the type of blotting membrane, type of blocking agent and the concentration of [^{125}I] protein A.

The experiments were reiterative in nature and no useful purpose would be served by describing them in detail. However, the following general conclusions are worth stating:

1 Some antisera raised in different animals or taken at different times after inoculation gave more quantitative results than others. The first batch of antiserum taken after inoculation worked best

2 Immobilon-P PVDF membranes (Millipore U.K. Ltd., Watford, Hertfordshire, U.K.) were almost as satisfactory as nitrocellulose membranes in terms of the total number of counts achieved and the linearity of the cross-reaction. However, the background counts on PVDF membranes were 10-fold lower than those with nitrocellulose membranes

3 The voltage and duration of blotting can affect the total counts achieved and the linearity of the cross-reaction. Blotting at 7.1 V/cm for 2 h gave lower total counts

but better linearity than blotting for 4 h at 14.2 V/cm

4 Tween 20 (0.5%) was a better blocking agent than bovine serum albumin (0.1%) in terms of the total number of counts achieved and the linearity of the cross-reaction

A protocol (Methods 2.8.4) was established that produced a linear relationship between the amount of protein loaded onto the gel (0.05 - 0.4 μg) and the number of counts in the corresponding band on the nitrocellulose membrane (Figure 6.2)

Immunoblotting of the enzymes from SDS-polyacrylamide gels confirmed that the reactions were located in a single band (Figure 6.2), corresponding in M_r value to the respective enzymes. For this reason there was no advantage in blotting the enzymes from SDS-polyacrylamide gels when application of the enzymes directly to the nitrocellulose membranes was more economical in terms of the materials used and the effort required. In many cases, when the enzymes were applied directly to the membranes (Methods 2.8.4c), the relationship between the amount of antigen applied and the counts per minute was linear between 0.05 to 0.8 μg of antigen (Figure 6.2). The linearity of the relationship between the amount of antigen and the counts per minute was not always as good as the example shown in Figure 6.2 but in all cases the relationship was positive and approximately linear up to 0.8 μg of antigen.

6.3.2 Cross-reactions detected by immunoblotting

Immunoblotting is a very sensitive technique and great care is needed to ensure that any weak cross-reactions are not artifacts caused by chance amino acid sequence homologies. A set of control proteins was selected; that is, proteins that were thought to be unlikely to have any very close evolutionary relationship with the aldehyde and alcohol dehydrogenases which were the subjects of this investigation (Section 1.6). The control proteins (Table 6.3) included alcohol and aldehyde dehydrogenases from different sources, as well as other proteins of diverse types. Alcohol and aldehyde dehydrogenases were included in order to show that any cross-reactions between the subject proteins were not due to general features of sequence or structure shared by all

dehydrogenase enzymes.

The proteins were applied directly to the nitrocellulose membranes (Methods 2.8.4) which were first probed with antiserum and then with [^{125}I] protein A (Methods 2.8.4). Each of the antisera was used to probe a different nitrocellulose membrane that had been loaded with a series of different amounts of homologous antigen (Figure 6.2) and also with 0.8 μg of each of the other proteins. A serial loading of the homologous antigen was included on every membrane to ensure that there was a positive and approximately linear relationship between the amount of antigen loaded and the number of counts in the corresponding band.

The extent of the cross-reactions among each of the antisera and 0.8 μg of each of the proteins is shown in Table 6.4. The antisera all reacted with their respective homologous antigens and a number of heterologous reactions were also recorded. There were cross-reactions among the three aldehyde dehydrogenases and also between the two alcohol dehydrogenases. All four antisera raised against BZDH I cross-reacted with BZDH II and TOL-BZDH less extensively than with the homologous antigen, but more extensively than with any of the control proteins. None of the three antisera raised against BZDH II cross-reacted with any of the control proteins, however they all cross-reacted with TOL-BZDH. Both of the antisera raised against TOL-BZDH cross-reacted with BZDH II to a much greater extent than with any of the control proteins. Each of the two antisera raised against each of the two alcohol dehydrogenases reacted with the heterologous alcohol dehydrogenase to a much greater extent than with any of the control proteins.

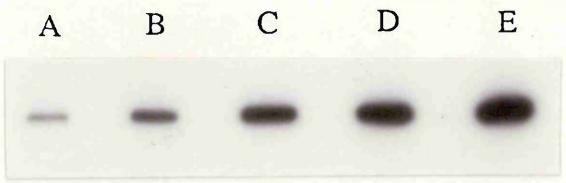


Figure 6.2 An example of a quantitative immunoblotting experiment

(a) The antigens were transferred to nitrocellulose membranes by electroblotting from SDS-polyacrylamide gels. After probing with antiserum and [^{125}I] protein A the membranes were autoradiographed.

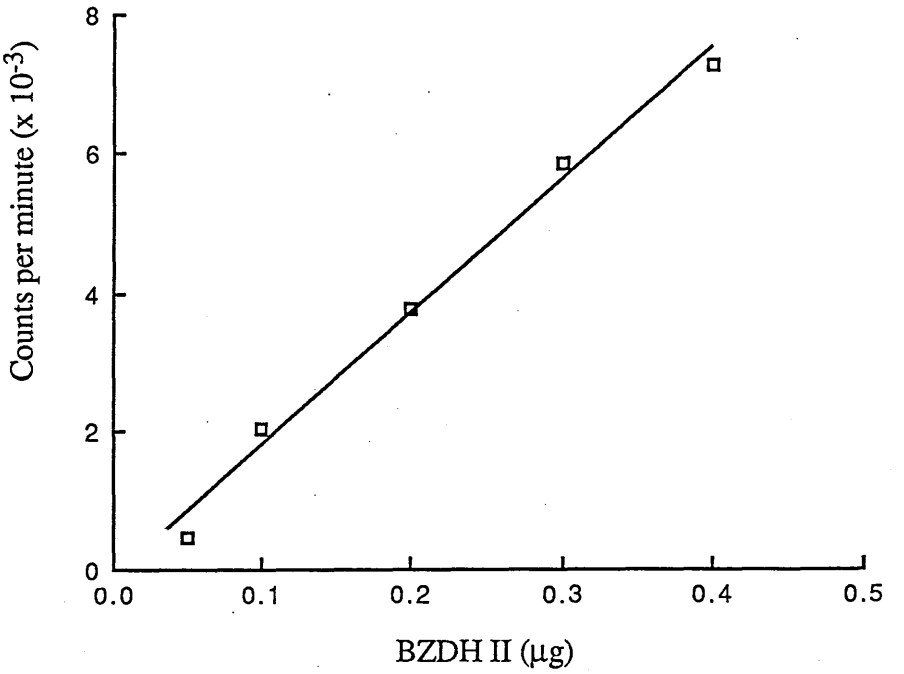
Tracks: A, 0.05 µg; B, 0.1 µg; C, 0.2 µg; D, 0.3 µg; E, 0.4 µg.

(b) Radioactive bands, located by autoradiography (see above) were cut out and counted.

(c) The antigen was applied directly to the nitrocellulose membrane using the Bio-Rad Dot Blot SF apparatus. After probing with antiserum and [^{125}I] protein A, appropriate sections of the membrane were cut out and counted.

In all cases the antiserum was anti-BZDH II (Code No. 314; Table 6.1).

(b)


Counts per minute (x 10^{-3})
8
6
4
2
0
0.0
0.1
0.2
0.3
0.4
0.5
BZDH II (µg)

(c)

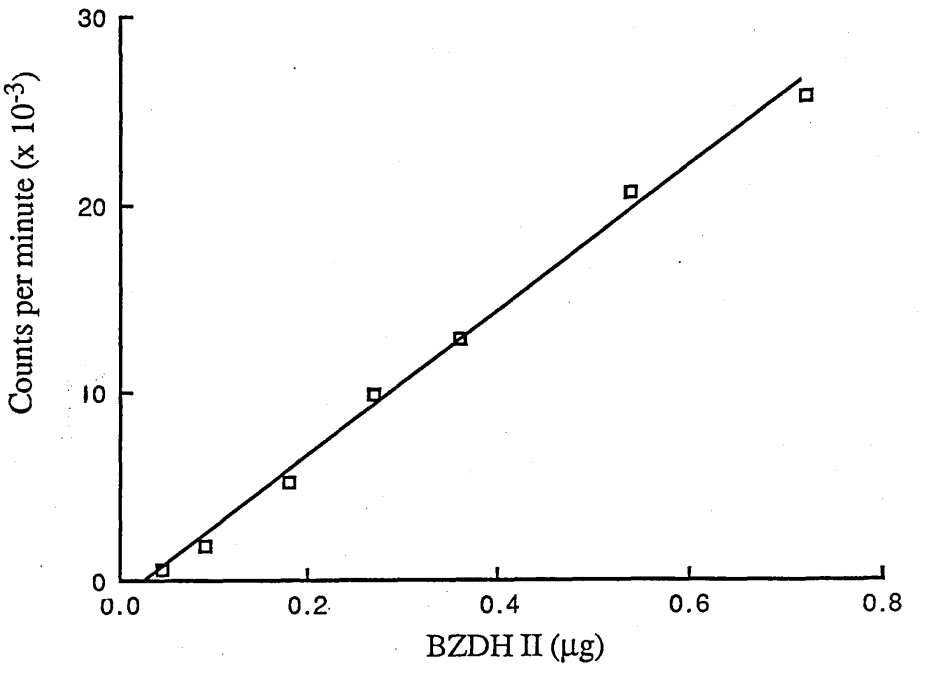

Counts per minute (x 10^{-3})
30
20
10
0
0.0
0.2
0.4
0.6
0.8
BZDH II (µg)

Table 6.3 Control proteins for immunoblotting

Various commercially available purified proteins were selected as control proteins for immunoblotting.

Protein	Source	Supplier
Serum albumin	Bovine	Pharmacia
Ovalbumin	Hen egg	"
Aldolase	Rabbit muscle	"
Ferritin	Horse spleen	'
Thyroglobulin	Bovine thyroid	"
Fumarase	Pig heart	Sigma
Catalase	Beef liver	"
Carbonic anhydrase	Bovine erythrocytes	"
Glycerol dehydrogenase	*Cellulomonas*	"
Alcohol dehydrogenase	*Thermoanaerobium brockii*	"
Alcohol dehydrogenase	Horse liver	Boehringer
Glyceraldehyde-3-phosphate dehydrogenase	Rabbit muscle	"
Aldehyde dehydrogenase	Yeast	Sigma
Formaldehyde dehydrogenase	*Pseudomonas putida*	"

Table 6.4 Quantitative immunoblotting

Antigen (0.8 μg) was applied to nitrocellulose membranes using the Bio-Rad Dot Blot SF apparatus. After probing with antiserum and [^{125}I] protein A appropriate sections of the membrane were cut out and counted. The results are from representative experiments and all the cross-reactions have been determined at least twice. Cross-reactions of less than 1% of that with the homologous antigen have been recorded as zero and control proteins (Table 6.3) that did not cross-react with any of the antisera are not shown on the table. The number of counts per minute associated with bands containing homologous antigen were quite constant in different experiments using the same antiserum, however they varied between 20 000 and 80 000 depending upon which antiserum was used.

Control proteins: C4, ferritin; C7, catalase; C10, alcohol dehydrogenase (*Thermoanaerobium*); C11 alcohol dehydrogenase (horse liver); C12, glyceraldehyde-3-phosphate dehydrogenase; C13 aldehyde dehydrogenase (yeast).

*, antiserum raised against reduced and carboxymethylated antigen

	% of cross-reaction with homologous antigen										
Antiserum	BZDH I	BZDH II	TOL-BZDH	BADH	Tol-BADH	C4	C7	C10	C11	C12	C13
BZDH I (022)	100	4	14	0	1	0	0	0	0	0	1
BZDH I (023)	100	20	14	1	0	0	0	0	0	0	0
BZDH I (254)	100	18	54	1	3	0	2	0	0	0	0
BZDH I (389*)	100	27	15	0	3	0	0	0	2	0	3
BZDH II (314)	0	100	7	0	0	0	0	0	0	0	0
BZDH II (251)	0	100	8	0	0	0	0	0	0	0	0
BZDH II (390*)	0	100	22	0	0	0	0	0	0	0	0
TOL-BZDH (1079)	1	23	100	0	0	3	0	6	0	0	2
TOL-BZDH (1080)	1	40	100	0	0	0	2	0	0	0	2
BADH (417*)	0	0	0	100	69	0	0	0	0	0	0
BADH (392)	0	0	0	100	94	0	0	0	3	0	0
TOL-BADH (1077)	0	4	0	38	100	0	0	0	0	0	0
TOL-BADH (1078)	0	0	0	24	100	0	0	0	0	0	1

6.4 Discussion

6.4.1 Design of control experiments for immunological assays

Immunological assays are very sensitive and may detect chance sequence similarities or similar structural features which show up as a background cross-reaction. When the conclusions drawn from a body of work rely on immunological relationships it is therefore particularly important to design stringent control experiments.

The only example of strong immunological cross-reactivity between proteins without sequence homology is the intensely sweet proteins monellin and thaumatin (Section 1.5; Ogata *et al.*, 1987). These proteins are a special case because they have a high specificity for the sweet receptor and therefore must share some structural features. Their common epitopes probably correspond to the three-dimensional structure of a sweet-receptor binding domain since they do not taste sweet when bound by antibodies nor when their structure is disrupted. This example illustrates how convergent evolution can potentially confuse the interpretation of the results of immunological experiments.

The most widely used type of control experiment is to test the antisera against a selection of control proteins that have been treated in the same way as the subject proteins (e.g. Zakin *et al.*, 1978; Chaffotte *et al.*, 1980; Pekkala-Flagan & Ruoslahti, 1982). This approach was used in the immunoblotting experiments (Section 6.3.2); the control proteins (Table 6.3) included several aldehyde and alcohol dehydrogenases from different sources in order to show that the cross-reactions among the subject proteins (Section 1.6) were not due to general features of sequence or structure shared by all enzymes of these types.

If experiments are carried out with antisera raised against denatured antigens it is important that two different methods are used for denaturation; one form is used during immunisation and the other during assay (e.g. Arnheim *et al.*, 1971; Mouhli *et al.*, 1980; Zakin *et al.*, 1980). This ensures that cross-reactions are not due to antibodies

directed against the denaturing agent. Antisera were raised against BZDH I, BZDH II and BADH after the enzymes had been denatured by the reduction and carboxymethylation of their cysteine residues (Section 6.1.2). The extent of the cross-reactions among the antisera and the enzymes were determined by quantitative immunoblotting after the enzymes had been denatured with SDS (Section 6.3.2).

The most stringent control experiment involves absorbing antibodies onto a column containing immobilised heterologous antigen; antibodies will be absorbed out of solution and can later be recovered (Arnon & Maron, 1971). If the recovered antibodies have a similar titre with the homologous and the heterologous antigen then it is certain that the cross-reaction was due to common epitopes and not to minor cross-contamination of the inoculum. This experiment was not attempted because it is technically very difficult, however a simpler experiment of a similar type was used to demonstrate that the cross-reaction between anti-BZDH I and native BZDH II was due to a single population of antibodies reacting with both of the enzymes and not to cross-contamination of the inoculum (Section 6.2.1).

6.4.2 Immunological cross-reactions give conservative indications of amino acid sequenc homologies

Immunological studies have often been used to demonstrate homology within families of isoenzymes (e.g. Arnon & Neurath, 1969; Arnheim *et al.*, 1971; Prager & Wilson, 1971a,b). In this type of study, polyclonal antisera, which contain antibodies directed against many different epitopes, are widely used in preference to monoclonal antibodies because related proteins share different numbers of epitopes depending upon the degree of divergence that has occurred. Monoclonal antibodies are not used because they will not detect a relationship between two proteins if they do not share the single epitope against which the antibody is directed. Although immunological cross-reactions only indicate homology amongst related proteins, there appears to be a usefully linear

relationship between the degree of cross-reaction and the extent of sequence divergence. Probably the best systematic study of the relationship between the extent of immunological cross-reactivity and amino acid sequence homology was carried out by Prager & Wilson (1971a,b) on a selection of purified lysozymes. They quantified cross-reactions by complement fixation and by precipitation assay and demonstrated that the values obtained could be interconverted. The results of the assays were used to calculate a defined term, 'immunological distance', which is a measure of the extent of the cross-reactivity between two proteins. A linear relationship was found between the immunological distance and the percent sequence difference when pairwise comparisons were made between members of the lysozyme family. The slope of the relationship was such that a 1% sequence change should cause a 2.6% difference in the cross-reactivity, measured in precipitation assays. This relationship predicts that immunological cross-reactions will not be detected beyond 40 % sequence difference.

Prager & Wilson (1971a,b) went on to survey the literature to find the extent of cross-reactivity between pairs of proteins of known sequence. They found zero cross-reaction to occur between 12 and 40% sequence difference in a set of 29 pairwise comparisons. When a plot of percent sequence difference against percent precipitation difference was constructed a linear relationship was again observed. The relationship was not as good as the one for the lysozyme family, however all of the points fell within an area representing 17% of the graph. They accounted for the large scatter by pointing out that the protocols used by other workers for immunisation and assay were not as stringent as their own. They had used antisera from four rabbits (mixed in inverse proportion to their homologous titres) taken after five boosts, over a six month period. They also expressed the extent of cross-reaction as an average of anti-X against Y and anti-Y against X.

The immunological comparisons reported in this chapter were carried out under less stringent conditions than the ones described by Prager & Wilson (1971a,b). The antisera were prepared after only two inoculations over a period of two months and they were used individually without mixing. The present study was never intended to define

precisely the extent of any sequence homology between the enzymes. Rather, it was intended to establish whether or not any of the proteins were homologous, which in turn would establish whether or not it would be fruitful to extend the work, at some future date, to include the cloning and sequencing of the genes that encode the enzymes (Section 1.6).

The extent of the cross-reactions among each of the 13 antisera (Table 6.1) and each of the five enzymes was studied by immunoprecipitation assay with the enzymes in their native states (Section 6.2). All three antisera raised against native BZDH I cross-reacted with BZDH II, and both of the antisera raised against native TOL-BADH cross-reacted with BADH. The putative homology between BZDH I and BZDH II, and between BADH and TOL-BADH, was not confirmed by reciprocal cross-reactions between BZDH I and antisera raised against BZDH II, nor by a cross-reaction between TOL-BADH and the single antiserum raised against native BADH. The reciprocal cross-reactions might be absent because the sequence identity between each of the pairs of enzymes is close to 60%, which Prager & Wilson (1971a,b) concluded to be the approximate limit of the sensitivity of the technique.

Yeh & Ornston (1982) have also described a pair of purified enzymes from the β-ketoadipate pathway in *P. putida* (Figure 1.3) that do not have reciprocity of immunological cross-reactions. They found that antisera raised against β-ketoadipate succinyl Co A transferase inhibit and precipitate carboxymuconolactone decarboxylase, but antisera raised against the decarboxylase do not react with the transferase.

Overall, cross-reactions between native proteins appear to indicate that they have greater than approximately 60% amino acid sequence identity. Sequence identity as low as 15%, between proteins of unbiased amino acid composition, is generally considered to be a statistically significant indication of an evolutionary relationship (e.g. Doolittle, 1981; Jornvall, 1980). Immunological cross-reactions therefore appear to be very conservative indicators of evolutionary relationships, provided that appropriate control experiments have been carried out.

6.4.3 Denatured antigens increases the range of sequence difference over which immunological relationships can be detected

The range of sequence difference over which immunological techniques can detect sequence homologies can be extended by using denatured antigens during immunisation and assay. The best known methods for denaturing antigens are SDS treatment (e.g. Lompre *et al.*, 1979; Schwartz *et al.*, 1980) and the modification of amino acid side chains (e.g. Arnon & Maron, 1971; Arnon, 1973). Hen egg-white lysozyme and human leukemia lysozyme have only 53 differences out of 129 amino acids, yet they show no cross-reactivity if antisera are raised against native proteins. Arnheim *et al.* (1971) demonstrated an immunological cross-reaction between these antigens after they had been denatured by the reduction and carboxymethylation of their cysteine residues. When the sequence of hen egg-white lysozyme is fitted into the three-dimensional structure known for human lysozyme it appears that 79% of the differences are on the surface of the molecule (Blake and Swan, 1971). In a small protein such as lysozyme only one in three amino acids is buried and when this is taken into account 74% of the internal residues and only 52% of the surface residues are conserved (Arnheim *et al.*, 1971). Presumably, substitution of the surface residues is more frequent because they have fewer conformational restraints than do residues in the centre of the molecule. In this example immunological tests were unable to detect the limited homology between the protein surfaces; however, denaturation of the antigens made the well-conserved buried regions available for study.

A similar procedure was used by Cohen and colleagues to probe the similarities amongst the isofunctional aspartokinase homoserine dehydrogenases (designated I and II), that catalyse the first and third steps in the methionine and threonine biosythetic pathways of *Escherichia coli* K12 (Figure 1.10; Zakin *et al.*, 1978; Sibilli *et al.*, 1982). In addition to catalysing the same reactions these enzymes have similar subunit M_r values and both have binding sites for aspartate, ATP, K^+, NADPH, homoserine and

aspartate semialdehyde. Despite these similarities they did cross-react unless they were denatured by modification of their cysteine residues. This work has been extended to show cross-reactions between aspartokinase I homoserine dehydrogenase I and aspartokinase III from the same organism (Mouhli *et al.*, 1980). Subsequent sequencing of the genes demonstrated sequence identity of approximately 40% amongst all three enzymes (Katinka *et al.*, 1980; Zakin *et al.*, 1983; Cassan *et al.*, 1986).

The immunological relationships amongst a selection of glyceraldehyde-3-phosphate dehydrogenases have been the subject of a broad investigation into the usefulness of studying denatured antigens (Zakin *et al.*, 1980). Antisera directed against native glyceraldehyde-3-phosphate dehydrogenase from *E. coli* K12 will recognise only glyceraldehyde-3-phosphate dehydrogenase from *Bacillus stearothermophilus* and to a lesser extent halibut. However, when antisera are raised against the denatured enzyme they will also recognise the denatured enzymes from man, ostrich, chicken, sturgeon, lobster and yeast. The complete amino acid sequences of five glyceraldehyde-3-phosphate dehydrogenases are known (only three of the ones above) and they have significant homology. When these sequences are fitted into the two known three-dimensional structures for glyceraldehyde-3-phosphate dehydrogenase the conserved residues are seen to be located mainly in buried areas where they are unavailable to immunological techniques using native antigens (Zakin *et al.*, 1980). As in the previous examples, the well conserved buries regions of the proteins were available for study only after denaturation of the native structure.

An example of a negative result using denatured antigens is found in the comparison of tryptophanase and the α and β_2 subunit of tryptophan synthetase which catalyse three consecutive reactions in the metabolism of tryptophan by *E. coli* K12 (Chaffotte *et al.*, 1980). All three proteins have a binding site for indole and the β_2 subunit of tryptophan synthetase shares in common with tryptophanase a binding site for tryptophan, pyridoxal 5-phosphate as cofactor and a similar catalytic mechanism. On the basis of these similarities it was suspected that the three enzymes arose by a process of retrograde evolution (see Section 1.4.2). Antisera were raised against native and denatured forms of

the β_2 subunit of tryptophan synthetase but they did not cross-react with the other enzymes (Chaffotte *et al.*, 1980).

BZDH I, BZDH II, TOL-BZDH, BADH and TOL-BADH were denatured with SDS and the extent of the cross-reactions with each of the 13 antisera (Table 6.1) was determined by quantitative immunoblotting (Section 6.3.2). Immunoprecipitation assays, using the native enzymes, had previously demonstrated cross-reactions between BZDH II and antisera raised against BZDH I, and also between BADH and antisera raised against TOL-BADH (Sections 6.2.1 & 6.4.2). These cross-reactions were confirmed using immunoblotting and additional ones were also observed. Antisera raised against BZDH II cross-reacted with TOL-BZDH and the reciprocal cross-reaction between BZDH II and antisera raised against TOL-BZDH was also observed. Antisera raised against BZDH I cross-reacted with TOL-BZDH, however there was no significant reciprocal cross-reaction between BZDH I and antisera raised against TOL-BZDH. Each of the antisera raised against BADH and TOL-BADH cross-reacted with the heterologous alcohol dehydrogenase.

Antisera were raised against BZDH I, BZDH II and BADH after the enzymes had been denatured by the reduction and carboxymethylation of their cysteine residues (Section 6.1.2). In immunoprecipitation assays these antisera had relatively low or undetectable titres for their native homologous antigens (Section 6.2.1), but overall they were little different from the antisera raised against the native enzymes in their cross-reactivity towards SDS-denatured enzymes (Section 6.3.2). This indicates that the protocol for the preparation of the inoculum (Methods 2.8.1) possibly denatures some significant proportion of the sample and elicits the production of antibodies against both the native and the denatured forms of the antigen. Alternatively, the protocol for probing the nitrocellulose membranes (Methods 2.8.3) may allow some proportion of the bound antigen to regain its native conformation.

6.4.4 Possible homologies among the aldehyde dehydrogenases and between the alcohol dehydrogenases

Immunological cross-reactions are probably detectable between denatured enzymes that have greater than 40% sequence identity (see above; Katinka *et al.*, 1980; Zakin *et al.*, 1983; Cassan *et al.*, 1986). The cross-reactions among the three aldehyde dehydrogenases in the immunoblotting experiments therefore indicate that they possibly share as much as 40% sequence identity (Section 6.3.2). Immunoprecipitation assays using the native proteins (Section 6.2.1) previously showed cross-reactions between BZDH I and BZDH II. This indicates that they possibly share as much as 60% sequence identity, which is the approximate limit of the sensitivity of immunological assays using native antigens (Section 6.4.2). Overall, the results of the immunological assays, using both native and denatured enzyme antigens, are consistent with all three aldehyde dehydrogenases sharing between 40% and 60% amino acid sequence identity.

Antisera raised against TOL-BADH cross-reacted with native BADH but the reciprocal cross-reaction between anti-BADH and TOL-BADH was not detected (Section 6.2.1). This is consistent with BADH and TOL-BADH having sequence identity close to 60% (Section 6.4.2). When experiments were carried out with denatured enzymes the reciprocal cross-reaction between anti-BADH and TOL-BADH was also observed (Section 6.3.2). Thus, like the aldehyde dehydrogenases, BADH and TOL-BADH might therefore be expected to have between 40% and 60% sequence identity.

The amino terminal sequence identity between BADH and TOL-BADH is 36% (Figure 5.19), and this lies just outside the range predicted on the basis of the immunological study. It is possible that the amino terminal sequences are atypical of the complete sequences, however this seems unlikely because many of these residues are involved in the structure of the active site (Section 5.5.6.2) and may therefore be more highly conserved than the complete amino acid sequence.

Cross-reactions were not detected between any of the antisera raised against BADH

or TOL-BADH and either of the alcohol dehydrogenase enzymes from horse liver or *Thermoanaerobium brockii* that were used as control proteins for immunoblotting (Section 6.3.2). At first sight this might seem to be surprising because BADH and TOL-BADH appear to belong to the same superfamily of zinc-dependent alcohol dehydrogenases as these control proteins (Section 5.5.6.2). Generally, immunological cross-reactions can be demonstrated between proteins if they have greater than 40% sequence identity (see above). If the sequence identities among BADH, TOL-BADH and the two control proteins are less than this, immunological techniques will not detect their relationships. BADH and TOL-BADH have 8% and 16% amino terminal sequence identity, respectively, with the *Thermoanaerobium* enzyme, and both have 26% identity with the horse liver enzyme (Figure 5.19). These levels of sequence identities are thus well outside the range demonstrable using immunological techniques, which explains why cross-reactions were not detected.

CHAPTER 7

Conclusions

7.1 Introduction

The properties of BZDH I described in this thesis complete the characterisation of all of the enzymes of the mandelate and benzyl alcohol pathways in *Acinetobacter calcoaceticus* N.C.I.B. 8250 (Section 1.6). BZDH I is induced co-ordinately with the other enzymes of the mandelate pathway, mandelate dehydrogenase and phenylglyoxylate decarboxylase (Section 1.2). Mandelate dehydrogenase, which is located on the inner cytoplasmic membrane, is $NAD(P)^+$-independent and its natural cofactor is probably a flavin (Fewson, 1988b). In contrast, phenylglyoxylate decarboxylase requires thiamin pyrophosphate for activity and appears to be freely soluble (Barrowman & Fewson, 1985). Both of the enzymes of the benzyl alcohol dehydrogenase pathway (BZDH II & BADH) are similar to BZDH I since they are NAD^+-dependent and also appear to be freely soluble (MacKintosh & Fewson, 1988a).

Two hypotheses have been proposed to account for the great variety of catabolic capabilities of microbial species: **1** retrograde evolution, in which each enzyme of a pathway is derived from the following enzyme by gene duplication and mutation (Section 1.4.2) and **2** evolution by gene recruitment, in which copies of genes for existing enzymes are recruited and integrated into evolving pathways by gene duplication and translocation (Section 1.4.3). There are many examples of evolution by gene recruitment but very few good examples of retrograde evolution. The detailed comparison of BZDH I, BZDH II and BADH appeared to be a good opportunity to test the hypothesis of retrograde evolution because the three enzymes share superficial characteristics that would be expected of enzymes that had evolved by retrograde evolution: all three enzymes share two common substrates (NAD^+ and benzaldehyde) and participate in peripheral metabolic pathways that may be more likely to undergo rapid evolution than the central pathways, which are constrained because of the necessity of conserving features concerned with regulation (Sections 1.4.2 & 1.6). The project was subsequently enlarged to include the purification and characterisation of TOL-BZDH and TOL-BADH,

because this would allow a more broadly-based comparison of five aldehyde and alcohol dehydrogenases, including a comparison of chromosomal and plasmid encoded enzymes.

7.2 Are the physical, chemical and kinetic properties of BZDH I, BZDH II, TOL-BZDH, BADH and TOL-BADH consistent with evolutionary relationships among any of the enzymes?

Some of the physical, chemical and kinetic properties of BZDH I, TOL-BZDH and TOL-BADH were determined (Chapter 5) largely to allow comparison of the enzymes with each other and with BZDH II and BADH which had already been studied extensively (MacKintosh & Fewson, 1988a,b). Although the aldehyde and the alcohol dehydrogenases share a number of similarities, the properties of the aldehyde dehydrogenases were in general more similar to each other than to the properties of the alcohol dehydrogenases.

If a classical taxonomic approach was adopted as a means of deciding whether or not there were any evolutionary relationships among the enzymes then each of the properties of each of the enzymes would be given equal weighting. Unfortunately this approach is inappropriate; partly because too few unique phenotypic characters have been measured but most importantly because the various characteristics of the enzymes are not equally related to the nucleotide sequences of the genes that encode the enzymes (Section 1.5). For example, appropriate substitution of a few amino acid residues on the surface of a protein may have a large effect on its isoelectric point and the similarity of the isoelectric points of BZDH I and BADH relative to that of BZDH II (Section 5.5.2) is therefore poor evidence of an evolutionary relationship between BZDH I and BADH. Furthermore, the very similar pH optima of the five aldehyde and alcohol dehydrogenases (pH 9.2 - 9.5; Section 5.5.2) is not evidence of evolutionary relationships among the five enzymes since the pH optima of the reactions are probably more strongly influenced by factors such as the thermodynamics and/or kinetics of the

reactions than by the nucleotide sequences of the genes that encode the enzymes.

The strongest evidence available at present of evolutionary relationships among any of the five enzymes are the homologies among the amino terminal sequences (Sections 1.5 & 5.5.6). BADH and TOL-BADH have 36% sequence identity and there can be little doubt that they are related to each other and also to the superfamily of long-chain, zinc-dependent alcohol/polyol dehydrogenases (Section 5.5.6). The interpretation of the homologies among the aldehyde dehydrogenase amino terminal sequences was more complicated than those between the alcohol dehydrogenases because a number of gaps were introduced into the sequences (Section 5.5.6). Overall, there appears to be a genuine evolutionary relationship between BZDH II and TOL-BZDH, however it is unclear whether or not the similarities among the sequences of these enzymes and the sequence of BZDH I is due to chance (Section 5.5.6) or to an evolutionary relationship.

There was no evidence from the comparison of the amino terminal sequences of any evolutionary relationships between any of the aldehyde dehydrogenases and any of the alcohol dehydrogenases (Section 5.5.6). However, the possibility that such a relationship exists can not be eliminated without knowledge of the complete amino acid sequences of the enzymes. The amino terminal sequences of BADH and TOL-BADH extended into the region where the active site zinc ligands are located (Section 5.5.6). In contrast, there is no evidence that the amino terminal sequences of any of the aldehyde dehydrogenases are involved in the structure of the active site and there is at least a formal possibility (though in my opinion unlikely) that the aldehyde and alcohol dehydrogenases are in fact related and that the differences among the subunit M_r values (Section 5.5.1) are due to the presence of extra residues on the amino terminal ends of the aldehyde dehydrogenase subunits.

7.3 Immunological cross-reactions among the aldehyde dehydrogenases and between the alcohol dehydrogenases

Homologies between the amino terminal sequences of BADH and TOL-BADH, and also between those of BZDH II and TOL-BZDH, are consistent with evolutionary relationships between each respective pair of enzymes (Section 7.2). These relationships were confirmed by the immunological cross-reactions that were observed between BADH and TOL-BADH and also between BZDH II and TOL-BZDH (Sections 6.2.1 & 6.3.2).

Additional cross-reactions were also observed that helped to clarify some of the more tentative conclusions drawn from the amino terminal sequence comparisons (Section 7.2). It was unclear from the sequence comparisons whether or not there is an evolutionary relationship between BZDH I and either of the other two aldehyde dehydrogenases (Section 7.2). All of the antisera raise against native BZDH I cross-reacted with native BZDH II (Section 6.2.1). When experiments were carried out with SDS-denatured antigens, cross-reactions were observed among all of the antisera raised against BZDH I and both of the other aldehyde dehydrogenases (Section 6.3.2). This is consistent with evolutionary relationships among all three aldehyde dehydrogenases (Section 6.4), although this should be considered a tentative conclusion since the reciprocal cross-reactions between BZDH I and any of the antisera raised against either BZDH II or TOL-BZDH were not observed (Sections 6.2.1 & 6.3.2). Finally, the absence of immunological cross-reactions among the aldehyde dehydrogenase and the alcohol dehydrogenase groups of enzymes (Sections 6.2.1 & 6.3.2) supports the conclusion drawn form the amino terminal sequence comparisons, that there is no evolutionary relationship between the two groups of enzymes.

7.4 Final conclusions and future experiments

The results of the amino terminal sequence comparisons (Section 7.2) and the immunological experiments (Section 7.3) are consistent with the alcohol dehydrogenases sharing a single ancestral gene with each other and also with the other 20 members of the superfamily of long-chain, zinc-dependent alcohol/polyol dehydrogenases that have so far been sequenced (Section 5.5.6). The interpretation of the results for the aldehyde dehydrogenases was less straight forward but overall they are consistent with BZDH II, TOL-BZDH and possibly also BZDH I sharing a single ancestral gene. One of the original aims of this projects was to test whether or not the alcohol dehydrogenases were derived from the aldehyde dehydrogenases by the process of retrograde evolution (Section 1.6). Neither the amino terminal sequence comparisons (Section 7.2) nor the immunological experiments (Section 7.3) provided any evidence of an evolutionary relationship between the aldehyde dehydrogenase and the alcohol dehydrogenase groups of enzymes. It therefore appears that the alcohol and aldehyde dehydrogenases have been brought together in the benzyl alcohol and the toluene pathways by gene recruitment rather than by retrograde evolution. If the alcohol dehydrogenases did in fact evolve from the aldehyde dehydrogenases by retrograde evolution then the relationships have been obscured by the effects of different selective pressures and genetic drift.

The most logical and worthwhile extension of the work described in this thesis is the cloning and sequencing of the genes for BZDH I, BZDH II, TOL-BZDH, BADH and TOL-BADH. This would firmly establish whether or not BZDH I and BZDH II evolved from a common ancestor, and if so it should illuminate the mechanism by which the putative duplicate genes were stabilised against elimination by unequal crossing over or by homologous recombination (Section 1.4.5). Furthermore, it would help to redress the imbalance between the large number (17) of complete sequences available for eukaryotic members of the zinc-dependent superfamily of alcohol dehydrogenases and the small number (3) of sequences available for prokaryotic members of the superfamily.

It may also be worthwhile to determine the complete three-dimensional structure of one of the aldehyde and one of the alcohol dehydrogenases. At present no aldehyde dehydrogenase three-dimensional structure is available and the different properties of BZDH I, BZDH II and TOL-BZDH may be sufficiently interesting to justify the considerable work and expense that would be necessary. For example, the three-dimensional structure of one of the enzymes would help to explain why BZDH I and TOL-BZDH are far more heat-stable than BZDH II (Section 5.5.8) and also why BZDH I and TOL-BZDH are K^+-activated while BZDH II is K^+-independent (Section 5.5.2). Members of the zinc-dependent superfamily of alcohol dehydrogenases have amino acid sequence identities that range approximately between 10% and 90% (Section 5.5.6.2). All of the sequences of the eukaryotic members of the superfamily can be accommodated by the three-dimensional structure of the horse liver enzyme but because of the great divergence of the sequences additional three-dimensional structures are required to gain a complete understanding of the properties of all of the members of the superfamily (Jornvall *et al.*, 1987b). The three-dimensional structure of either BADH or TOL-BADH would fulfill this requirement since the structure of a prokaryotic member of the superfamily would complement that of the horse liver enzyme.

References

Ahern, T. J. & Klibanov, A. M. (1988) Methods Biochem. Anal. **33**, 91-127

Akrigg, D., Bleasby, A. J., Dix, N. I. M., Findlay, J. B. C., North, A. C. T., Parry-Smith, D., Wootton, J. C., Blundell, T. L., Gardner, S. P., Hayes, F., Islam, S., Sternberg, M. J. E., Thornton, J. M., Tickle, I. J. & Murray-Rust, P. (1988) Nature (London) **355**, 745-746

Albertini, A. M., Hofer, M., Calos, M. P. & Miller, J. H. (1982) Cell (Cambridge, MA) **29**, 319-328

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

Amicon (1980), Dye Ligand Chromatography, Amicon Corp., Lexington, U.S.A.

Arnheim, N., Sobel, J. & Canfield, R. (1971) J. Mol. Biol. **61**, 237-250

Arnon, R. & Neurath, H. (1969) Proc. Natl. Acad. Sci. U.S.A. **64**, 1323-1328

Arnon., R. & Maron, E. (1971) J. Mol. Biol. **61**, 225-235

Arnon, R. (1973) in The Antigens (Sela, M., ed.), Vol. 1, pp. 88-159, Academic Press, New York

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

Bauchop, T. & Elsden, S. R. (1960) J. Gen. Microbiol. **23**, 457-469

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

Belfaiza, J., Parsot, C., Martel, A., de la Tour, C. B., Margarita, D., Cohen, G. N. & Saint-Girons, I. (1986) Proc. Natl. Acad. Sci. U.S.A. **83**, 867-871

Blake, C. C. F. & Swan, I. D. A. (1971) Nature (London) New Biol. **232**, 12-15

Boehringer (1987) Biochemica Service, Issue 6b.

Bostian, K. A. & Betts, G. F. (1978a) Biochem. J. **173**, 773-786

Bostian, K. A. & Betts, G. F. (1978b) Biochem. J. **173**, 787-798

Bridgen, J., Kolb, E. & Harris, J. I. (1973) FEBS Lett. **33**, 1-3

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

Cain, R. B. (1980) in Lignin Biodegradation: Microbiology, Chemistry & Potential Applications (Kirk, T. K., Higuchi, T. & Chang, H-M., eds.), Vol. 1, pp. 21-60, CRC Press, Boca Raton, Florida

Canovas, J. L. & Stanier, R. Y. (1967) Eur. J. Biochem. **1**, 289-300

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

Chaffotte, A. F., Zakin, M. M. & Goldberg, M. E. (1980) Biochem. Biophys. Res. Commun. **92**, 381-388

Chakrabarty, A. M. (1976) Annu. Rev. Genet. **10**, 7-30

Coggins, J. R. (1978) in Theory & Practice of Affinity Techniques (Sunderman, P. V. & Eckstein, F., eds.), pp. 89-112, Academic Press, New York

Conway, T., Sewell, G. W., Osman, Y. A. & Ingram, L. O. (1987) J. Bacteriol. **169**, 2591-2597

Conway, T. & Ingram, L. O. (1989) J. Bacteriol. **171**, 3754-3759

Cook, A. M., Beggs, J. D. & Fewson, C. A. (1975) J. Gen. Microbiol. **91**, 325-337

Cornish-Bowden, A. (1979) J. Theor. Biol. **76**, 369-386

Crawford, I. P. (1975) Bacteriol. Rev. **39**, 87-120

Crawford, R. L. (1981). Lignin Biodegradation & Transformation, John Wiley & Sons, New York

Dagley, S. (1975) in Essays in Biochemistry (Campbell, P. N. & Aldridge, W. N., eds.), pp. 81-138, Academic Press, New York

Dayhoff, M. O. (1978) in Atlas of Protein Sequence & Structure (Dayhoff, M. O., ed.), Vol. 5, Supplement 3, p. 26, National Biomedical Research Foundation, Washington DC

Dayhoff, M. O., Schwartz, R. M. & Orcutt, B. C. (1978) in Atlas of Protein Sequence & Structure (Dayhoff, M. O., ed.), Vol. 5, Supplement 3, pp. 345-352, National Biomedical Research Foundation, Washington DC

Dean, P. D. G. & Watson, D. H. (1979) J. Chromatogr. **165**, 301-319

Dickinson, F. M. & Haywood, G. W. (1987) Biochem. J. **247**, 377-384

Doolittle, R. F. (1981) Science **214**, 149-159

Drewke, C. & Ciriacy, M. (1988) Biochim. Biophys. Acta **950**, 54-60

Duchange, N., Zakin, M. M., Ferrara, P., Saint-Girons, I., Park, I., Tran, S.V., Py, M-C. & Cohen, G. N. (1983) J. Biol. Chem. **258**, 14868-14871

Edmundson, A. B. (1965) Nature (London) **205**, 883-887

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

Eklund, H., Nordstrom B., Zeppezauer, E., Soderlund, G., Ohlsson, I., Boiwe, T., Soderberg, B-O., Tapia, O., Branden, C-I. & Akeson, A. (1976) J. Mol. Biol. **102**, 27-59

Eklund, H., Samama, J-P., Wallen, L., Branden, C-I., Akeson, A. & Jones, T. A. (1981) J. Mol. Biol. **146**, 561-587

Fersht, A. (1984). Enzyme Structure & Mechanism, Freeman, San Francisco

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

Fewson, C. A. (1967b) J. Gen. Microbiol. **48**, 107-110

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

Fewson, C. A. (1988a) Trends Biotechnol. **6**, 148-153

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

Findlay, J. B. C., Pappin, D. J. C. & Keen, J. N. (1989) in Protein Sequencing - A Practical Approach (Findlay, J. B. C. & Geisow, M., eds.), pp. 69-84, IRL Press, Oxford

Flavin, M. (1975) in Metabolic Pathways (Greenberg, D. M., ed.), Vol. 7, pp. 457-503, Academic Press, New York

Fothergill-Gilmore, L. A. (1986) Trends Biochem. Sci. **1986**, 198-203

Haber, J. E. & Koshland, D. E. (1970) J. Mol. Biol. **50**, 617-639

Hall, B. G. (1984) in Micro-organisms as Model Systems for Studying Evolution (Mortlock, R. P., ed), pp. 165-185, Plenum Press, New York & London

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

Hartley, B. S. (1984) in Micro-organisms as Model Systems for Studying Evolution (Mortlock, R. P., ed), pp. 23-54, Plenum Press, New York & London

van Heyningen, S. & Coulson, A. F. W. (1987) Biochem. J. **244**, 805-806

Hills, C. A. & Fewson, C. A. (1983) J. Gen. Microbiol. **129**, 2009-2015

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

Holms, W. H. & Bennett, P. M. (1971) J. Gen. Microbiol. **65**, 57-68

Horiuchi, T., Horiuchi, S. & Novick, A. (1963) Genetics **48**, 157-169

Horowitz, N. H. (1945) Proc. Natl. Acad. Sci. U.S.A. **31**, 153-157

Horowitz, N. H. (1965) in Evolving Genes & Proteins (Bryson, V. & Vogel, H. J., eds.), pp. 15-23, Academic Press, New York

Jakoby, W. B. (1963) Enzymes, 2nd Edn. **7**, 203-221

Jain, R. K. & Sayler, G. S. (1987) Microbiol. Sci. **4**, 59-63

Jeck, R., Woenckhaus, C., Harris, J. I. & Runswick, M. J. (1979) Eur. J. Biochem. **93**, 57-64

Jeenes, D. J., Reineke, W., Knachmuss, H-J. & Williams, P. A. (1982) J. Bacteriol. **150**, 180-187

Jeffcoat, R., Dagley, S. (1973) Nature (London) New Biol. **241**, 186-187

Jendrossek, D., Steinbuchel, A. & Schlegel, H. G. (1988) J. Bacteriol. **170**, 5248-5256

Jensen, R. A. (1976) Ann. Rev. Microbiol. **30**, 409-425

Jornvall, H. (1973) Biochem. Biophys. Res. Commun. **53**, 1096-1101

Jornvall, H. (1980) in Dehydrogenases Requiring Nicatinamide Coenzymes (Jeffery, J., ed), pp. 126-148, Birkhauser Verlag, Basel, Boston & Stuttgart

Jornvall, H., von Bahr-Lindstrom, H., Jany, K-D., Ulmer, W. & Froschle, M. (1984) FEBS Lett. **165**, 190-196

Jornvall, H., Persson, B. & Jeffery, J. (1987a) Eur. J. Biochem. **167**, 195-201

Jornvall, H., Hoog, J-O., von Bahr-Lindstrom, H., Johansson, J., Kaiser, R. & Persson, B. (1987b) Biochem. Soc. Trans. **16**, 223-227

Julia, P., Pares, X. & Jornvall, H. (1988) Eur. J. Biochem. **172**, 73-83

Juni, E. (1978) Ann. Rev. Microbiol. **32**, 349-371

Katinka, M., Cossart, P., Sibilli, L., Saint-Girons, I., Chalvignac, M. A., Le Bras, G., Cohen, G. N. & Yaniv, M. (1980) Proc. Natl. Acad. Sci. U.S.A. **77**, 5730-5733

Keil, H., Keil, S., Pickup, R. W. & Williams, P. A. (1985) J. Bacteriol. **164**, 887-895

Keil, H., Saint, C. M. & Williams P. A. (1987a) J. Bacteriol. **169**, 764-770

Keil, H., Keil, S. & Williams, P. A. (1987b) J. Gen. Microbiol. **133**, 1149-1158

Kemper, J. (1984) in Micro-organisms as model systems for studying evolution (Mortlock, R. P., ed), pp. 255-284, Plenum Press, New York & London

Kennedy, S. I. T., Fewson, C. A. (1968a) J. Gen. Microbiol. **53**, 259-273

Kennedy, S. I. T., Fewson, C. A. (1968b) Biochem. J. **107**, 497-506

Kessler, S. W. (1975) J. Immunol. **115**, 1617-1624

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

Laursen, R. A. (1971) Eur. J. Biochem. **20**, 89-102

Lebioda, L. & Stec, B. (1988) Nature (London) **333**, 683-686

Levinson, G. & Gutman, G.A. (1987) Mol. Biol. Evol. **4**, 203-221

Lewin, B. (1987) Genes III, pp. 399-421, J. Wiley & Sons, New York

Lewis, E.B. (1951) Cold Spring Harbor Symp. Quant. Biol. **16**, 159-174

Livingstone, A. & Fewson, C. A. (1972) Biochem. J. **130**, 937-946

Livingstone, A., Fewson, C. A., Kennedy, S. I. T. & Zatman, L. J. (1972) Biochem. J. **130**, 927-935

Lompre, A-M., Bouveret, P., Leger, J. & Schwartz, K. (1979) J. Immunol. Methods **28**, 143-148

Loomes, K. M. & Kitson, T. M. (1986) Biochem. J. **238**, 617-619

Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. (1951) J. Biol. Chem. **193**, 265-275

Lumsden, J. & Coggins, J. R. (1977) Biochem. J. **161**, 599-607

Lumsden, J. & Coggins, J. R. (1978) Biochem. J. **169**, 441-444

MacKintosh, R. W. & Fewson, C. A. (1987) in Enzymology and Molecular Biology of Carbonyl Metabolism III: Aldehyde Dehydrogenase, Aldo-Keto Reductase, & Alcohol Dehydrogenase (Weiner, H. & Flynn, T.G. eds.), pp. 259-273, Alan R. Liss, New York

MacKintosh, R. W. (1988) Ph.D Thesis, University of Glasgow

MacKintosh, R. W. & Fewson, C. A. (1988a) Biochem. J. **250**, 743-751

MacKintosh, R. W. & Fewson, C. A. (1988b) Biochem. J. **255**, 653-661

McLachlan, A. D. (1971) J. Mol. Biol. **61**, 409-424

Marchalonis, J. J. & Weltman, J. K. (1971) Comp. Biochem. Physiol. **38B**, 609-625

Matthews, B. W. (1987) Biochemistry **26**, 6885-6888

Mortlock, R. P. (1984) in Micro-organisms as Model Systems for Studying Evolution (Mortlock, R. P., ed), pp. 1-21, Plenum Press, New York & London

Mouhli, H., Zakin, M. M., Richaud, C. & Cohen, G. N. (1980) Biochem. Int. **1**, 403-409

Murray, K., Duggleby, C. J., Sala-Trepat, J. M. & Williams, P. A. (1972) Eur. J. Biochem. **28**, 301-310

Neale, A. D., Scopes, R. K., Kelly, J. M. & Wettenhall, R. E. H. (1986) Eur. J. Biochem. **154**, 119-124

Needleman, S. B. & Wunsch, C. D. (1970) J. Mol. Biol. **48**, 443-453

Neidle, E. L., Hartnett, C., Bonitz, S. & Ornston, L. N. (1988) J. Bacteriol. **170**, 4874-4880

Nimmo, G. A., Nimmo, H. G., Hamilton, I. D., Fewson, C. A. & Wilkins, M. B. (1986) Biochem. J. **239**, 213-220

Ogata, C., Hatada, M., Tomlinson, G., Shin, W-C. & Kim, S-H. (1987) Nature (London) **328**, 739-742

Ohno S. (1970) Evolution by Gene Duplication, G. Allen & Unwin Ltd, London

Ornston, L. N. & Parke, D. (1977) Curr. Top. in Cell. Regul. **2**, 209-262

Ornston, L. N. & Yeh, W-K. (1982) in Biodegradation and Detoxification of Environmental Pollutants (Chakrabarty, A.M. ed), pp. 105-126, CRC Press, Boca Raton, Florida

Parsot, C. (1986) EMBO J. **5**, 3013-3019

Parsot, C., Saint-Girons, I. & Cohen, G.N. (1987) Microbiol. Sci. **4**, 258-262

Pearson, W. R. & Lipman, D. J. (1988) Proc. Natl. Acad. Sci. U.S.A. **85**, 2444-2448

Pekkala-Flagan, A. & Ruoslahti, E. (1982) J. Immunol. **128**, 1163-1167

Peretz, M. & Burstein, Y. (1989) Biochemistry **28**, 6549-6555

Phillips, D. C., Blow, D. M., Hartley, B. S. & Lowe, G. (1970) Phil. Trans. R. Soc. (London) **B257**, 63-146

Prager, E. M. & Wilson, A. C. (1971a) J. Biol. Chem. **246**, 5978-5989

Prager, E. M. & Wilson, A. C. (1971b) J. Biol. Chem. **246**, 7010-7017

Ramos, J. L., Mermod, N. & Timmis, K. N. (1987) Molec. Microbiol. **1**, 293-300

Ramos, J. L. & Timmis, K. N. (1987) Microbiol. Sci. **4**, 228-237

Riley, M. (1984) in Micro-organisms as Model Systems for Studying Evolution (Mortlock, R. P., ed), pp. 285-315, Plenum Press, New York & London

Rossmann, M. G. & Argos, P. (1977) J. Mol. Biol. **109**, 99-129

Sanderson, K. E. (1976) Annu. Rev. Microbiol. **30**, 327-349

Schwartz, K., Lompre, A-M., Bouveret, P., Wisnewsky, C. & Swynghedauw, B. (1980) Eur. J. Biochem. **104**, 341-346

Shapiro, A. L. & Maizel, J. V. (1969) Anal. Biochem. **29**, 505-514

Sibilli, L., Le Bras, G., Fazel, A., Bertrand, O., Dautry-Varsat, A., Cohen, G. N. & Zakin, M. M. (1982) Biochem. Int. **4**, 331-336

Skerker, P. S. & Clark, D. S. (1989) Biotechnol. & Bioeng. **33**, 62-71

Springham, M. G. & Betts, G. F. (1973) Biochim. Biophys. Acta **309**, 233-236

Stachow, C. S., Stevenson, I. L. & Day, D. (1967) J. Biol. Chem. **242**, 5294-5300

Stanier, R. Y. & Ornston, L. N. (1973) Adv. Microbiol. Physiol. **9**, 89-151

Sukhodolets, V. V. (1988) Microbiol. Sci. **5**, 202-206

Tautz, D., Trick, M. & Dover, G. A. (1986) Nature (London) **322**, 652-656

Ting, H-H. & Crabbe, M. J. C. (1983) Biochem. J. **215**, 361-368

Towbin, H., Staehelin, T. & Gordon, J. (1979) Proc. Natl. Acad. Sci. U.S.A. **76**, 4350-4354

Tu, G-C. & Weiner, H. (1988) J. Biol. Chem. **263**, 1218-1222

Vihinen, M. (1987) Protein Eng. **1**, 477-480

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

Williams, P. A. & Murray, K. (1974) J. Bacteriol. **120**, 416-423

Williams, P. A. & Worsey, M. J. (1976) J. Bacteriol. **125**, 818-828

Williamson, V. M. & Paquin, C. E. (1987) Mol. Gen. Genet. **209**, 374-381

Winstanley, C., Taylor, S. C. & Williams, P. A. (1987) Mol. Microbiol. **1**, 219-227

Worsey, M. J. & Williams, P. A. (1975) J. Bacteriol. **124**, 7-13

Wray, W., Boulikas, T., Wray, V. P. & Hancock, R. (1981) Anal. Biochem. **118**, 197-203

Yeh, W-K. & Ornston, L. N. (1982) J. Bacteriol. **149**, 374-377

You, K-s., Arnold, L. J., Allison, W. S. & Kaplan, N. O. (1978) Trends Biochem. Sci. **1978**, 265-268.

Youngleson, J. S., Jones, W. A., Jones, D. T. & Woods, D. R. (1989) Gene **78**, 355-364

Zakin, M. M., Garel, J-R., Dautry-Varsat, A., Cohen, G. N. & Boulot, G. (1978) Biochemistry **17**, 4318-4323

Zakin, M. M., Hirth, C., Garel, J-R. & Cohen, G. N. (1980) Mol. Immunol. **17**, 1373-1379

Zakin, M. M., Duchange, N., Ferrara, P. & Cohen, G. N. (1983) J. Biol. Chem. **258**, 3028-3031

Publications

Chalmers, R. M. & Fewson, C. A., (1988). Quantitative immunoblotting in the study of bacterial evolution. Biochem. Soc. Trans. **16**, 153-154.

Chalmers, R. M. & Fewson, C. A., (1988). The evolution of metabolic pathways: an immunological approach to the evolution of aromatic alcohol and aldehyde dehydrogenases. In Enzymology & Molecular Biology of Carbonyl Metabolism 2 (Weiner, H. & Flynn, T. G., eds.), pp. 193-207, Alan R. Liss, New York

Chalmers, R. M. & Fewson, C. A. (1989). Purification and characterisation of benzaldehyde dehydrogenase I from *Acinetobacter calcoaceticus*. Biochem. J. **263**, 913-919

Chalmers, R. M. & Fewson, C. A. (1990). Purification of the benzyl alcohol dehydrogenase and benzaldehyde dehydrogenase encoded by the TOL plasmid pWW53 of *Pseudomonas putida* MT53 and their preliminary comparison with benzyl alcohol dehydrogenase and benzaldehyde dehydrogenase I & II from *Acinetobacter calcoaceticus*. J. Gen. Microbiol. **136**, in the press