PHD

Studies of dihydrolipoamide dehydrogenase from the halophilic archaeon Haloferax volcanii

Jolley, Keith A.

Award date: 1996

Awarding institution: University of Bath

Link to publication
Studies of Dihydrolipoamide Dehydrogenase from
the Halophilic Archaeon *Haloferax volcanii*

submitted by Keith A. Jolley

for the degree of PhD

of the University of Bath

1996

COPYRIGHT

Attention is drawn to the fact that copyright of this thesis rests with its author.

This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with its author and that no quotation from the thesis and no information derived from it may be published without the prior written consent of the author.

This thesis may be made available for consultation within the University Library and may be photocopied or lent to other libraries for the purpose of consultation.
Abstract

The aim of this project was to create a high-level expression system for halophilic dihydrolipoamide dehydrogenase (DHLipDH), and use this system to express site-directed mutants to investigate structural features important in conferring halophilicity.

Attempts were made to express the enzyme in *Escherichia coli*. Although large quantities of a glutathione-S-transferase-DHLipDH fusion product were obtained, it was found difficult to purify and cleave the DHLipDH fragment from the rest of the fusion. Homologous expression was finally achieved using a shuttle vector, capable of replicating in both *Haloferax volcanii* and *E. coli*. The recombinant enzyme was purified to near homogeneity by a combination of hydroxylapatite and Cu²⁺-chelate chromatography. Crystallisation trials using the purified product have identified conditions around which further optimisation should yield crystals suitable for X-ray diffraction.

Using a molecular model, a putative K⁺-binding site was identified at the dimer interface of the enzyme, consisting of four coordinated glutamate residues. This site is apparently absent from the enzyme from non-halophilic sources, so was deemed a suitable target of site-directed mutagenesis. The site was perturbed and mutant proteins expressed in a strain of *H. volcanii* lacking DHLipDH activity, developed in collaboration with Dr. M. Dyall-Smith (University of Melbourne, Australia). Characterisation of these mutants have identified significant changes in the halophilic stability of the enzyme, and a role for the site in stabilising the protein in elevated salt conditions is discussed.

Additionally, the region upstream of the DHLipDH gene in the genome has been sequenced and coding regions for components homologous to those from pyruvate dehydrogenase multienzyme complexes (PDHC) identified. These complexes have never previously been found in the Archaea. The E2 component appears to have undergone an internal duplication event that would perturb formation of an active complex, possibly explaining the lack of activity of a PDHC in this organism.
Table of contents

Abstract .............................................................................................................. ii
Table of contents ............................................................................................. iii
List of figures .................................................................................................... vii
Abbreviations ................................................................................................... x
Acknowledgements ........................................................................................ xii
1. Introduction ................................................................................................. 1
  1.1 General phenotypes of the Archaea ........................................................... 1
      1.1.1 Physiology .............................................................................................. 1
  1.2 Archaeal phylogeny ..................................................................................... 2
      1.2.1 Creation of phylogenetic trees from molecular data ......................... 4
      1.2.2 Rooting the universal phylogenetic tree .............................................. 5
      1.2.3 Opposition to the universal tree .......................................................... 7
      1.2.4 The consensus view ............................................................................ 8
  1.3 The halophilic Archaea ............................................................................... 9
      1.3.1 Ecology, taxonomy and phenotypes .................................................... 9
      1.3.2 Haloferax volcanii ............................................................................... 11
  1.4 Enzymes in high salt environments ......................................................... 11
      1.4.1 Structural features important for halotolerance .............................. 12
      1.4.2 Alternative mechanisms for existence in high salt ....................... 13
  1.5 Dihydrolipoamide dehydrogenase (DHLipDH) ........................................ 14
      1.5.1 2-Oxo acid dehydrogenase multienzyme complexes ...................... 15
      1.5.2 The Archaeal alternative ................................................................. 17
      1.5.3 Other possible roles of DHLipDH ...................................................... 18
      1.5.4 Available data .................................................................................... 19
      1.5.5 DHLipDH from H. volcanii ............................................................... 19
      1.5.6 Molecular model of halophilic DHLipDH .......................................... 20
      1.5.7 Use of halophilic DHLipDH as a model enzyme .............................. 23
  1.6 The biotechnological potential of halophiles and their proteins ............ 23
      1.7 Project aims ........................................................................................... 25
2. Materials and Methods ............................................................................. 27
  2.1 Materials ..................................................................................................... 27
  2.2 Strains and culture conditions ................................................................. 28
  2.3 Biochemical Techniques .......................................................................... 30
      2.3.1 Preparation of cell extracts ............................................................... 30
      2.3.2 DHLipDH enzyme assay ................................................................. 30
      2.3.3 Preparation of dihydrolipoamide ...................................................... 31
      2.3.4 Citrate synthase assay ...................................................................... 31
      2.3.5 Acetylation of coenzyme A ............................................................. 32
      2.3.6 SDS polyacrylamide gel electrophoresis (SDS-PAGE) ................... 32
      2.3.7 Chromatographic techniques ............................................................ 33
      2.3.7.1 Hydroxylapatite chromatography ............................................... 33
      2.3.7.2 FPLC Superdex 200 gel filtration .................................................... 33
      2.3.7.3 Glutathione agarose affinity chromatography ............................ 33
      2.3.7.4 Copper chelate affinity chromatography ....................................... 34
      2.3.8 TCA precipitation of proteins ........................................................... 35
2.3.9 Protein estimation .................................................................................. 35
  2.3.9.1 Bradford Assay ............................................................................... 35
  2.3.9.2 BCA Assay ...................................................................................... 35
  2.3.10 Crystallisation trials .............................................................................. 36
2.4 Molecular Biological Methods ....................................................................... 36
  2.4.1 Quantitation of DNA ........................................................................... 36
  2.4.2 Ethanol precipitation of DNA ................................................................. 36
  2.4.3 DNA preparation methods ...................................................................... 37
    2.4.3.1 Preparation of \textit{H. volcanii} genomic DNA .................................. 37
    2.4.3.2 Large scale (Maxi) preparations of \textit{E. coli} plasmid DNA ............. 37
    2.4.3.3 Small scale (Mini) preparations of plasmid DNA - Alkaline lysis method ........................................................................... 38
    2.4.3.4 Preparations of \textit{H. volcanii} plasmid DNA ................................ 39
  2.4.4 Agarose gel electrophoresis ................................................................. 39
  2.4.5 Recovery and purification of DNA from agarose gels ............................ 40
    2.4.5.1 Sephadex recovery ........................................................................ 40
    2.4.5.2 Geneclean\textsuperscript{\textregistered} method ........................................ 40
  2.4.6 Restriction digestion of DNA ................................................................. 40
  2.4.7 Ligations ................................................................................................ 41
  2.4.8 Phosphorylation of oligonucleotides ......................................................... 41
  2.4.9 Removal of single-stranded overhangs ..................................................... 42
  2.4.10 Site-directed mutagenesis .................................................................... 42
  2.4.11 Polymerase chain reaction (PCR) amplification of DNA ....................... 42
  2.4.12 DNA sequencing ................................................................................... 43
  2.4.13 Transformation of \textit{E. coli} .................................................................. 44
    2.4.13.1 CaCl\textsubscript{2} method ................................................................. 44
    2.4.13.2 RbCl / CaCl\textsubscript{2} method ......................................................... 45
    2.4.13.3 TSS method .................................................................................. 45
  2.4.14 Transformation of \textit{H. volcanii} .......................................................... 46
  2.4.15 Purification of polyethylene glycol (PEG) 600 ....................................... 47

3. Expression in \textit{E. coli} ................................................................................. 48
  3.1 Introduction ................................................................................................ 48
  3.2 Amplification of gene from pNAT82 by PCR .............................................. 48
  3.3 Subcloning into the pMEX8 expression vector ........................................... 54
    3.3.1 Expression in JRG1342 (DHLipDH - minus) ...................................... 56
    3.3.2 Growth of JRG1342 in elevated salt concentrations ............................ 58
    3.3.3 Inclusion of FAD in extracts to facilitate correct folding ....................... 60
    3.3.4 Growth of cells at 30\textdegree{}C ............................................................ 60
  3.4 Expression using the pGEX-KG fusion protein vector ................................ 60
    3.4.1 Subcloning of DHLipDH gene into pGEX-KG ..................................... 62
    3.4.2 Expression of a GST-DHLipDH fusion protein .................................... 64
    3.4.3 Thrombin cleavage of GST-DHLipDH ................................................. 66
    3.4.4 Heat stability of the fusion proteins ...................................................... 70
  3.5 Subcloning into the pKK223-3 expression vector ...................................... 72
  3.6 Use of compatible solutes .......................................................................... 73
    3.6.1 Elevated salt and inclusion of betaine in growth media .......................... 73
    3.6.2 Osmotic stress conditioning .................................................................. 74
  3.7 Discussion .................................................................................................. 75
    3.7.1 Use of PCR in gene cloning ................................................................. 75
    3.7.2 The suitability of \textit{E. coli} for expression of halophilic proteins .......... 76
    3.7.3 GST-DHLipDH fusion proteins ............................................................. 78
4. Over-expression in H. volcanii .............................................................. 80
  4.1 Introduction .............................................................................................. 80
  4.2 Subcloning DHLipDH gene into the pMDS20 shuttle vector ............... 80
  4.3 Introduction of a rRNA promoter .............................................................. 84
  4.4 Change of antibiotic resistance marker ................................................... 87
  4.5 Concluding remarks ................................................................................. 93

5. Purification and characterisation of recombinant halophilic DHLipDH ........................................................................................................ 94
  5.1 Introduction .............................................................................................. 94
  5.2 Development of the purification protocol ................................................ 94
    5.2.1 Heat treatment...................................................................................... 94
    5.2.2 Hydroxylapatite chromatography.......................................................... 96
    5.2.3 Mono-Q ion exchange.......................................................................... 96
    5.2.4 Gel filtration.......................................................................................... 98
    5.2.5 Cu²⁺ chelate chromatography............................................................... 98
    5.2.6 The final protocol................................................................................100
  5.3 Characterisation of the purified enzyme ................................................102
    5.3.1 Kinetic properties of the enzyme.........................................................102
    5.3.2 Salt optima..........................................................................................103
    5.3.3 pH optimum........................................................................................106
    5.3.4 UV-visible absorption spectrum..........................................................107
  5.4 Crystallisation trials ................................................................................107
  5.5 Concluding remarks ...............................................................................109

6. Creation of a DHLipDH - minus strain of H. volcanii ...................... 111
  6.1 Introduction ................................................................................  111
  6.2 Insertional inactivation of the DHLipDH gene ....................................... 111
  6.3 Confirmation of DHLipDH gene disruption ............................................ 113
  6.4 Enzyme assays of the DHLipDH - minus strain .................................... 115
  6.5 Concluding remarks ...............................................................................116

7. Site-directed mutagenesis of halophilic DHLipDH ............................. 118
  7.1 Introduction .............................................................................................118
  7.2 A putative K⁺-binding site and design of mutants .................................. 118
  7.3 Mutagenesis ............................................................................................122
    7.3.1 The Altered Sites® II in vitro mutagenesis system............................122
    7.3.2 Cloning of halophilic DHLipDH into pALTER-1.................................124
    7.3.3 Mutagenesis reactions........................................................................124
  7.4 Homologous expression of mutants......................................................128
  7.5 Purification of the mutant enzymes .......................................................134
  7.6 Characterisation of mutants ...................................................................136
    7.6.1 Kinetic parameters ............................................................................136
    7.6.2 Salt optima.........................................................................................136
    7.6.3 pH optima..........................................................................................138
    7.6.4 Thermal stability. .................................................................................139
      7.6.4.1 Stability in the presence of salt....................................................139
List of figures

Figure 1.1: Pathways of glucose catabolism in halophilic and thermoacidophilic Archaea........................................................................................................3
Figure 1.2: Rooted universal phylogenetic tree derived from 16S and 18S rRNA sequence comparisons using distance matrix and maximum parsimony algorithms ..................................................................................................... 6
Figure 1.3: Lake’s rooted tree, constructed using the evolutionary parsimony algorithm ...................................................................................................... 9
Figure 1.4: NAD⁺-dependent oxidation of dihydrolipoamide catalysed by dihydrolipoamide dehydrogenase .................................................................... 15
Figure 1.5: Mechanism of DHLipDH ................................................................ 15
Figure 1.6: The reaction mechanism of the 2-oxo acid dehydrogenase multi- enzyme complexes ..................................................................................... 16
Figure 1.7: The catalytic cycle of a 2-oxoacid : ferredoxin oxidoreductase, found in Archaea and anaerobic Bacteria .............................................................. 17
Figure 1.8: Molecular model of H. volcanii DHLipDH ........................................ 22

Figure 3.1: Partial restriction map of pNAT82 .................................................... 49
Figure 3.2: Restriction digestion of pNAT82 ...................................................... 50
Figure 3.3: Binding positions of PCR primers KAJ-2 and KAJ-3 on pNAT82 .... 51
Figure 3.4: PCR amplification of DHLipDH gene from pNAT82 ........................ 53
Figure 3.5: The pMEX8 expression vector ...................................................... 55
Figure 3.6: Restriction digestion of DHLipDH cloned into pMEX8 ...................... 55
Figure 3.7: Halotolerance of E. coli DHLipDH .................................................. 56
Figure 3.8: Growth profile of E. coli JRG1342 .................................................. 57
Figure 3.9: SDS-PAGE of extracts of transformed JRG1342 cells grown in a range of NaCl concentrations ................................................................. 59
Figure 3.10: The pGEX-KG fusion expression vector ....................................... 61
Figure 3.11: Binding positions of PCR primers KAJ-4 and KAJ-5 on pNAT82 ... 62
Figure 3.12: PCR amplification of DHLipDH from pNAT82 using primers KAJ-2 and KAJ-4 ................................................................. 63
Figure 3.13: Expression of GST-DHLipDH fusion protein (with linker) .......... 65
Figure 3.14: Thrombin cleavage of GST-DHLipDH fusion protein in cell-free extract ........................................................................................................ 67
Figure 3.15: Thrombin-cleavage time course ................................................... 68
Figure 3.16: Time-course showing DHLipDH activity in extracts of cells transformed with pGEX-KG constructs following reaction with thrombin for varying periods .............................................................................. 69
Figure 3.17: DHLipDH activity in extracts of cells transformed with pGEX-KG constructs after reaction with thrombin for varying periods and thermal renaturation .................................................................................... 70
Figure 3.18: Heat stability of the pGEX-DHLipDH fusion proteins .................... 71
Figure 3.19: The pKK223-3 expression vector ................................................. 72
Figure 3.20: Growth profile of E. coli under salt-induced osmotic stress, in the presence or absence of 100 µM betaine ..................................................... 74
Figure 3.21: Growth profile of E. coli under salt-induced osmotic stress, in the presence or absence of 100 µM betaine, following osmotic stress conditioning .............................................................................................. 75
Figure 4.1: The pMDS20 shuttle vector .............................................................. 81
Figure 4.2: Binding positions of PCR primers KAJ-1 and KAJ-2 on pNAT82 ...... 82
Figure 4.3: PCR amplification of DHLpDH gene from pNAT82 ......................... 83
Figure 4.4: Construction of the pMDS24 shuttle / expression vector ............... 86
Figure 4.5: Binding positions of primers KAJ-6 and KAJ-7 to the novobiocin resistance determinant of pMDS20 ............................................................. 87
Figure 4.6: PCR amplification of the rovobiocin resistance determinant from pMDS20 ..................................................................................................... 89
Figure 4.7: Optimised PCR amplification of the novobiocin resistance determinant from pMDS20 ................................................................................. 90
Figure 4.8: pBAP5003 shuttle vector for over-expression of DHLipDH in H. volcanii ....................................................................................................... 91
Figure 4.9: SDS-PAGE of whole cell extracts from H. volcanii BAS5005 cells transformed with pBAP5003 ............................................................. 92

Figure 5.1: Plot showing relative specific activity and the proportion of protein remaining following incubation of cell-free extracts at 95°C for varying lengths of time ............................................................................................................. 95
Figure 5.2: SDS-PAGE showing level of purification following heat-step and hydroxylapatite chromatography ................................................................................................................................. 97
Figure 5.3: SDS-PAGE of adjacent fractions eluted from gel filtration .......... 99
Figure 5.4: SDS-PAGE of adjacent fractions eluted from the Cu²⁺-chelate column ........................................................................................................... 99
Figure 5.5: SDS-PAGE showing each stage of the purification procedure ...... 101
Figure 5.6: Hanes-Woolf plot showing S/Vₘₐₓ⁡ₕₖₚₚₚ against dihydrolipoamide concentration for recombinant halophilic DHLipDH ................................................ 102
Figure 5.7: Hanes-Woolf plot showing S/Vₘₐₓ⁡ₕₖₚₚ against NAD⁺ concentration for recombinant halophilic DHLipDH ................................................... 103
Figure 5.8: Salt activity profile of recombinant halophilic DHLipDH (NaCl) .. 104
Figure 5.9: Salt activity profile of wild-type non-recombinant halophilic DHLipDH (NaCl) .......................................................... 105
Figure 5.10: Salt activity profile of recombinant halophilic DHLipDH (KCl) ...... 105
Figure 5.11: Normalised pH profile of recombinant DHLipDH ................. 106
Figure 5.12: UV-visible absorption spectra of oxidised recombinant DHLipDH and FAD ......................................................................................... 107
Figure 5.13: Crystals of halophilic DHLipDH .................................................... 110

Figure 6.1: Insertional inactivation of the DHLipDH gene ............................... 112
Figure 6.2: PCR analysis of BAS5005 (DHLipDH - minus strain of H. volcanii) 114
Figure 6.3: DHLipDH and citrate synthase activities of four mevinolin-resistant transformants .......................................................................................... 115

Figure 7.1: The putative K⁺ binding site, consisting of four coordinated glutamate residues ................................................................. 119
Figure 7.2: Molecular model of H. volcanii DHLipDH, based on the structure from Pseudomonas fluorescens ............................................................ 120
Figure 7.3: Alignment of available DHLipDH sequences, showing the putative K⁺ binding site residues, E423 and E426 .................. 121
Figure 7.4: The Altered Sites® II in vitro mutagenesis procedure ................. 123
Figure 7.5: Cloning of DHLipDH into pALTER-1 ........................................... 125
## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATP</td>
<td>adenosine triphosphate</td>
</tr>
<tr>
<td>bp</td>
<td>base pairs</td>
</tr>
<tr>
<td>BCDHC</td>
<td>branched-chain 2-oxo acid dehydrogenase multienzyme complex</td>
</tr>
<tr>
<td>BSA</td>
<td>bovine serum albumin</td>
</tr>
<tr>
<td>CoA</td>
<td>coenzyme A</td>
</tr>
<tr>
<td>CS</td>
<td>citrate synthase (EC 4.1.3.7)</td>
</tr>
<tr>
<td>DHLipDH</td>
<td>dihydrolipoamide dehydrogenase (EC 1.8.1.4)</td>
</tr>
<tr>
<td>DMSO</td>
<td>dimethyl sulphoxide</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>DTNB</td>
<td>5,5'-dithio bis-2-nitrobenzoate</td>
</tr>
<tr>
<td>EDTA</td>
<td>(disodium) ethylenediamine tetraacetate</td>
</tr>
<tr>
<td>FAD</td>
<td>flavin adenine dinucleotide</td>
</tr>
<tr>
<td>Fd</td>
<td>ferredoxin</td>
</tr>
<tr>
<td>GST</td>
<td>glutathione-S-transferase (EC 2.5.1.18)</td>
</tr>
<tr>
<td>HMG-CoA</td>
<td>3-hydroxy-3-methylglutaryl coenzyme A</td>
</tr>
<tr>
<td>IPTG</td>
<td>isopropyl-β-D-thiogalactoside</td>
</tr>
<tr>
<td>kb</td>
<td>kilobase pairs</td>
</tr>
<tr>
<td>kD</td>
<td>kilodaltons</td>
</tr>
<tr>
<td>LB</td>
<td>Luria-Bertani</td>
</tr>
<tr>
<td>MGM</td>
<td>modified growth medium</td>
</tr>
<tr>
<td>MOPS</td>
<td>3-(N-morpholino) propane-sulphonic acid</td>
</tr>
<tr>
<td>Mr</td>
<td>molecular mass</td>
</tr>
<tr>
<td>NAD(H)</td>
<td>nicotinamide adenine dinucleotide (reduced)</td>
</tr>
<tr>
<td>OGDH</td>
<td>2-oxoglutarate dehydrogenase multienzyme complex</td>
</tr>
<tr>
<td>PAGE</td>
<td>polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>PDHC</td>
<td>pyruvate dehydrogenase multienzyme complex</td>
</tr>
<tr>
<td>PHB</td>
<td>poly-β-hydroxybutyrate</td>
</tr>
<tr>
<td>PECSS</td>
<td>Perkin Elmer computer spectroscopy software</td>
</tr>
<tr>
<td>PEG</td>
<td>polyethylene glycol</td>
</tr>
<tr>
<td>PMSF</td>
<td>phenylmethane sulphonyl fluoride</td>
</tr>
<tr>
<td>RNase A</td>
<td>ribonuclease A (EC 3.1.27.5)</td>
</tr>
<tr>
<td>rRNA</td>
<td>ribosomal RNA</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>-------------</td>
<td>-------------</td>
</tr>
<tr>
<td>SDS</td>
<td>sodium dodecyl sulphate</td>
</tr>
<tr>
<td>TAE</td>
<td>tris-acetate EDTA</td>
</tr>
<tr>
<td>TCA</td>
<td>trichloroacetic acid</td>
</tr>
<tr>
<td>TE(8)</td>
<td>10 mM Tris-HCl, 1 mM EDTA (pH 8)</td>
</tr>
<tr>
<td>TEMED</td>
<td>N, N, N', N' tetramethylethylene diamine</td>
</tr>
<tr>
<td>ThDP</td>
<td>thiamin diphosphate</td>
</tr>
<tr>
<td>Tris</td>
<td>tris-(hydroxymethyl)-methylamine</td>
</tr>
<tr>
<td>TSS</td>
<td>transformation and storage solution</td>
</tr>
<tr>
<td>X-Gal</td>
<td>5-bromo-4-chloro-3-indolyl-β-D-galactoside</td>
</tr>
</tbody>
</table>
Acknowledgements

I would like to thank Drs. Michael Danson and David Hough for their guidance and supervision of this project. I am also very grateful to Dr. Michael Dyall-Smith (University of Melbourne, Australia) who was instrumental in the development of the shuttle vector and DHLipDH - minus strain of *H. volcanii*. His experience in the field of halobacterial genetics was invaluable and his enthusiasm led to many helpful discussions. I would also like to thank Dr. Rupert Russell for creation of the molecular model and advice on the design of site-directed mutants.

Special mention must go to all the members of the lab for their help and for making my time at Bath enjoyable.

Financial support for this work was provided by the BBSRC.
1. Introduction

1.1 General phenotypes of the Archaea

The Archaea (previously known as archaebacteria) encompass three main phenotypes, namely thermophilic, halophilic and methanogenic (reviewed by Woese & Olsen, 1986; Fewson, 1986; Woese, 1987; Danson et al., 1992). The thermophiles are sulphur-dependent organisms found in thermophilic environments (55->110°C), such as submarine vents and hot springs, and many can reduce or oxidise sulphur to produce energy. Some are also acidophiles, growing at acidities as low as pH 0. The methanogens are obligate anaerobes, reducing carbon dioxide to methane (Jones et al., 1987). They are found in a wide variety of habitats, such as in cattle rumen, sewage-treatment plants and anaerobic sediments. The halophiles require high concentrations (up to 5 M) of sodium chloride for growth and survival. Some are also highly alkaliphilic. Their internal environment has a high concentration of potassium chloride, isotonic with the exterior. Recently, a psychrophilic archaean has been isolated and this may represent a fourth main phenotype (Preston et al., 1996).

The Archaea display a diverse range of morphological forms: irregular disks, cocci or straight or curved rods. They vary in size from 0.2 μm to 10 μm in diameter.

1.1.1 Physiology

Archaea possess some distinctive biochemical pathways involved with central metabolism (Danson, 1993). Some halophilic Archaea have been shown to catabolise glucose and galactose via a modified Entner-Doudoroff pathway (Tomlinson et al., 1974). Glucose is oxidised to gluconate by an NAD⁺-
dependent dehydrogenase and is then converted to 2-keto-3-deoxygluconate. This is phosphorylated with ATP to form 2-keto-3-deoxy-6-phosphogluconate, which then undergoes aldol cleavage to produce pyruvate and glyceraldehyde 3-phosphate. Thermoacidophilic Archaea catabolise glucose through a further modification of this pathway, in which no phosphorylated intermediates are used to produce the first molecule of pyruvate (Figure 1.1).

Archaea have some other biochemical features that phenotypically distinguish them from other organisms. They lack muramic acid-containing peptidoglycan in their cell envelopes (Kushner et al., 1964; Brown & Cho, 1970) and their membrane lipids are composed of ether-linked isoprenyl phosphoglycerides (Tornabene & Langworthy, 1979). Archaea generally lack sensitivity to many antibiotics concerned with cell wall and protein synthesis (Woese et al., 1978; Bonelo et al., 1984).

1.2 Archaeal phylogeny

Phylogeny is the evolutionary history of a taxonomic group, such as a species or a group of related organisms. Historically, taxonomic classifications were based on the fossil record, visible traits and, more recently, the results of biochemical investigations of the organisms. Such classification schemes were more data retrieval systems, conveying information such as structure, physiology and behaviour, than tools for making inferences about the evolutionary relationships between organisms.
Figure 1.1: Pathways of glucose catabolism in halophilic and thermoacidophilic Archaea. The modified Entner-Doudoroff pathways of halophiles (—►) and the non-phosphorylated Entner-Doudoroff pathway of Sulfolobus solfataricus and Thermoplasma acidophilum (—►—►) are shown alongside the classical Entner-Doudoroff pathway of eubacteria (—►). Conversion of glyceraldehyde into pyruvate via glycerate has been demonstrated only in T. acidophilum (Danson, 1993).

The use of non-comparable properties to create phylogenetic relationships is fraught with difficulties. Comparison of genetic sequences has, for this reason, been found to be a more reliable and useful method for such analyses. Ribosomal RNA (rRNA) sequences have proved to be a suitable subject for such
studies. They are found in all living organisms, are readily isolated and evolve only slowly with time, enabling the relatedness of distant species to be determined. The first rigorous analysis of the 16S and 18S rRNA of a wide range of organisms (Woese & Fox, 1977; Woese et al., 1978) clustered the organisms into three primary kingdoms or "urkingdoms" with the proposed names; the eubacteria, the eukaryotes and the archaeabacteria. This replaced the dichotomy between prokaryotes and eukaryotes that was previously accepted. It appeared that the archaebacteria were no more related to typical bacteria than the bacteria were related to the eukaryotes. It was later proposed that a new taxon be introduced above the level of kingdom called a "domain", with all life comprising three domains: the Bacteria, the Archaea and the Eucarya, each containing at least two kingdoms (Woese et al., 1990). The term "archaebacteria" was abandoned due to its suggestion of a specific relationship between the Archaea and the Bacteria.

The archaeal domain comprises two kingdoms. The first of these, the *Euryarchaeota*, is a phenotypically-diverse group containing the methanogens, the extreme halophiles, sulphate reducers and some thermophiles. The other archaeal kingdom, the *Crenarchaeota*, is a more homogenous grouping of thermophilic phenotypes, comprising acidophilic and sulphur-dependent thermophiles.

1.2.1 Creation of phylogenetic trees from molecular data

Phylogenetic trees can be produced from sequence data of molecular species, such as rRNA, by a number of methods (for a review see Beanland and Howe, 1992). Any of these procedures can produce incorrect trees because they all involve assumptions and approximations. The general procedure of all the methods involves creating a sequence alignment to generate data for
phylogenetic inference. Following the phylogenetic inference steps, such as parsimony, maximum likelihood or evolutionary parsimony (method of invariants) the tree is tested to see how strongly it is favoured over others.

During the sequence alignment step it must be confirmed that the sequences are homologous, that is whether they have a degree of similarity greater than would be expected by chance or convergence. Many alignments are not suitable for direct use by inference programs as gaps can lead to grouping artefacts.

Parsimony is one of the most popular inference methods due to its relative simplicity and ease of interpretation. The theory is that simple explanations should be favoured over complex ones. In other words, a parsimonious tree is the tree that requires the least evolutionary changes to explain the data.

Maximum likelihood methods use sophisticated statistical models, taking into account factors such as unequal rates of nucleotide substitution in different lineages, site-specific rate variability and the distribution of variable sites. These methods are gaining favour due to their overtly statistical nature and the fact that many of them are less restrictive in the assumptions they follow than other techniques.

1.2.2 Rooting the universal phylogenetic tree

A phylogenetic tree derived from a comparison of a single molecular species from a range of organisms is generally unrooted, because the universal ancestor cannot be determined. A sequence-based tree relating a restricted group of organisms can be rooted by establishing at which point it is closer to an outgroup, but for a tree relating all organisms there can be no such outgroup (Doolittle & Brown, 1994). In order to root the phylogenetic tree derived from 16S rRNA
(Woese et al., 1990; Wheelis et al., 1992) the Dayhoff strategy was used - a comparison of the amino acid sequences of pairs of paralogous proteins, assumed to have arisen by duplication events, whose common ancestor predates the most recent universal ancestor. The proteins used were the α and β subunits of ATPases and the translation elongation factors EF-1α (Tu) and EF-2 (G), each of which shows extensive sequence homology across members of the three domains (Gogarten et al., 1989; Iwabe et al., 1989). By inferring a composite phylogenetic tree from such comparisons the root can be positioned where the two genes are thought to have diverged. The root on the rRNA tree has been positioned at the same point (Figure 1.2)

![Phylogenetic Tree]

**Figure 1.2:** Rooted universal phylogenetic tree derived from 16S and 18S rRNA sequence comparisons using distance matrix and maximum parsimony algorithms (reproduced from Wheelis et al., 1992). Rooting of the tree was achieved by using data from duplicated genes encoding the α and β subunits of ATPases and elongation factors EF-1α and EF-2 (Iwabe et al., 1989)

The validity of rooting the tree in such a manner is not universally acclaimed, however. Using a similar procedure with glutamate dehydrogenase genes as
was used by Iwabe et al. (1989), it was found to be impossible to root the tree conclusively (Benachenhou-Lahfa, Forterre & Labedan, 1993; Forterre et al., 1993). With this discrepancy in mind, and the fact that the currently-accepted position of the root is based on just two pairs of sequences, doubts must remain as to the accuracy of the placing.

The major conclusions from Woese's tree are that the Archaea are a monophyletic evolutionary group that lie on the Eukaryl lineage. The shortest branches of the tree are to the hyperthermophilic Archaea, suggesting that they are the most primitive.

1.2.3 Opposition to the universal tree

Not every member of the evolutionary community was prepared to embrace Woese's universal phylogenetic tree. Objections ranged from both philosophical debate about the purpose of phylogenetic categorisation to questioning the statistical validity of the methods employed.

Mayr (1990) argues that basing a classification solely on differences in rRNA sequences ignores the apparently enormous evolutionary step from the prokaryotes to the eukaryotes. The general organisation of Archaea and Bacteria, with regard to lack of a nucleus, chromosomes and organelles, is far removed from that of the Eucarya. Mayr would prefer a system based on two domains: the eukaryotes and the prokaryotes.

Margulis and Guerrero question the assumption that molecules of a single member of a species yield information typical of that species (Margulis & Guerrero, 1991; Margulis, 1996). Greater concern is shown when the chosen sequence varies between different tissues of an organism, or at different stages in its life cycle. Their modification of Whittaker's original "five kingdoms" scheme is based on molecular biological results as well as more traditional taxonomic
methods. In their scheme, they have two superkingdoms - the Procaryotae and the Eucaryotae. The prokaryotic domain contains just one kingdom, the Monera, which includes all the eubacteria and Archaea. The eukaryotes are divided into three higher kingdoms; Animalia, Fungi and Plantae. The last kingdom is the Protoctista, containing all living things with a nucleus that do not display animal, plant or fungal qualities.

Lake (1988, 1989) has argued that the methods employed in the construction of Woese's phylogenetic tree are prone to errors when rates of nucleotide substitution are greatly unequal in juxtaposed branches, and when mutations are frequent. He suggested use of the evolutionary parsimony algorithm as being the least sensitive to unequal rate effects. Using this algorithm he obtained a tree with a different branching pattern (Figure 1.3), suggesting two superkingdoms, with the methanogens and halophiles grouping with the eubacteria and the sulphur-dependent thermoacidophiles grouping with the eukaryotes. In this proposal, therefore, Archaea are polyphyletic.

This view has been rejected by Olsen and Woese (1993) on the grounds of the alignment used and the fact that methods other than evolutionary parsimony give different results. In addition, the same method using 23S rRNA also gives a different grouping to Lake (Gouy & Li, 1989).

1.2.4 The consensus view

It is now commonly accepted that life on Earth can be divided into three domains, although the position of the root of the tree is not without controversy. Since the original suggestion of the archaeal tree, more supporting evidence has become available. The resolution of the tree has been enhanced by the complete sequencing of many 16S rRNAs and trees created using other molecular species, such as 23S rRNA (Leffers et al., 1987), support the three domain view.
1.3 The halophilic Archaea

1.3.1 Ecology, taxonomy and phenotypes

The Halobacteriaceae comprise six genera: Haloarcula, Halobacterium, Halococcus, Haloferax, Natronobacterium and Natronococcus (for a review see Rodriguez-Valera, 1988). They are found in hypersaline environments such as salt lakes and salterns (where sea water is evaporated to produce salt) and even in highly-alkaline soda lakes (natronobacteria). In such habitats they are easily identifiable by their red colour, mainly caused by their carotenoid pigments. The halobacteria can only grow in salt solutions above 15% (w/v) and most require 20
- 25% (w/v) for good growth (sea water is ~3.5% salt). Such hypersaline environments are inhibitory to growth and survival of most other forms of life.

The halobacteria are all either strict aerobes or facultative anaerobes. They are chemoorganotrophic and grow on complex media, while some species of *Halobacteria* can utilise solar energy through their production of bacteriorhodopsin. Generation times are considerably longer than most bacteria, with cells dividing every 3-4 hours in the case of *Haloferax volcanii* to 8-12 hours for species of *Halobacterium*. The osmotic pressure of the external environment is balanced by isotonic accumulation of K⁺ ions (Kushner, 1988). This is maintained against a large concentration gradient.

In general, the halophilic Archaea do not have a rigid cell wall (with the exception of the halococci). The cell boundary consists of a single layer of glycoprotein, called the S-layer. This S-layer is held together by divalent cations and can be completely stripped from the cells by treatment with EDTA. This property is utilised in the preparation of competent cells for transformation. Many of the halobacteria have a rod-shaped morphology, although some take on square or triangular forms. *Haloferax* species commonly have thin cup-shaped cells. Due to the lack of a cell wall most halobacteria are Gram-negative, while the halococci are Gram-positive.

Halobacteria are insensitive to antibiotics usually used in bacterial selection procedures. They are, however, sensitive to mevinolin, an inhibitor of 3-hydroxy-3-methylglutaryl CoA (HMG-CoA) reductase - a key enzyme in the biosynthesis of isoprenoids and cholesterol. They also have sensitivity to novobiocin, which inhibits DNA gyrase by blocking binding of ATP. Plasmids containing resistant genes to each of these antibiotics are available and their development has enabled transformant selection.
1.3.2 Haloferax volcanii

_Haloferax volcanii_ was first isolated from the Dead Sea in 1975 (Mullakhanbhai & Larsen, 1975). It requires at least 1 M NaCl for growth, with cells lysing upon dilution of the culture medium with water. Optimum growth conditions are NaCl concentrations between 1.7 and 2.5 M at 30 to 40°C. These salt conditions categorise _H. volcanii_ as a moderately-extreme halophile. The organism has a high requirement for Mg\(^{2+}\), with an optimum concentration of 0.2 M, higher than many other halophiles.

The use of _H. volcanii_ for molecular biological studies has the advantage over that of many halophiles in that its genome is relatively stable. Many halophiles have a large number of mobile insertion sequences that lead to genetic instability, making it necessary to avoid repetitive sub-culturing.

Transformation techniques have been developed for _H. volcanii_ (Charlebois _et al._, 1987) and a number of shuttle vectors are now available that will replicate in both _E. coli_ and _H. volcanii_ (Lam & Doolittle, 1989; Holmes _et al._, 1991; Holmes _et al._, 1994). These developments have established the organism as perhaps the species of choice for genetic-level studies of halophilic Archaea.

1.4 Enzymes in high salt environments

The effects of high concentrations of certain salts on protein stability are related to modifications of the solvent properties of the aqueous solution by the salt (Von Hippel & Schleich, 1969). The internal environment of the extremely halophilic Archaea is close to a saturating concentration of KCl, requiring the proteins of these organisms to be themselves halophilic. They are usually only stable in solvents of very high salt concentration (Eisenberg & Wachtel, 1987). Under
these conditions, most proteins from non-halophilic organisms are likely to precipitate or unfold, depending on the salt.

The folding of a protein is determined by the interaction of the amino acid side-chains with water (Dill, 1990; Zaks, 1992). In order to minimise the entropically unfavourable solvation of hydrophobic residues, these are forced into the hydrophobic interior of the molecule. In contrast, charged and polar amino acids occupy the surface area, resulting in a tightly-bound hydration layer. The surface-energy of the protein-water interface, caused by the hydrogen-bonding nature of the solvent, results in the molecule folding to minimise its surface area. Addition of solutes, such as salts, to the bulk solvent disturb the hydrogen-bonding nature of the water and hence the surface energy, resulting in a change of conformation of the protein.

Increases in surface tension resulting from increasing salt concentration induce proteins to further minimise their surface area and they may aggregate, resulting in salt-precipitation. Since the surfaces of proteins are charged, however, they are able to attract hydrated salt ions, counter-balancing the exclusion of the salt. Effect of salt on protein conformation, therefore, is a balance of these two effects (Timasheff, 1992).

1.4.1 Structural features important for halotolerance

Halophilic proteins generally have a high proportion of acidic amino acids, with an excess of acidic over basic residues of about 17%. In addition, they also have a lower hydrophobic amino acid content (Reistad, 1970; Lanyi, 1974).

A large proportion of the work investigating stability of halophilic proteins has been conducted on malate dehydrogenase from \textit{Halobacterium marismortui} (Mevarech, Eisenberg & Neumann, 1977; Mevarech & Neumann, 1977; Pundak & Eisenberg, 1981; Pundak \textit{et al.}, 1981; Reich \textit{et al.}, 1982; Zaccai, Wachtel &
Eisenberg, 1986; Zaccai, Bunick & Eisenberg, 1986; Zaccai et al., 1989; Hecht et al., 1989; Bonneté et al., 1993) but other work suggests that the main findings are relevant to other halophilic proteins (Guinet et al., 1988). The X-ray crystallographic structure of *H. marismortui* malate dehydrogenase has been resolved to 3.2Å (Dym et al., 1995). The structure is seen to be stabilised by arginine, glutamate and aspartate salt bridges, both interior and exterior to the monomer subunits and there are increased negative charges over the surface when compared to non-halophilic homologues.

Proteins are not able to compete with a highly concentrated salt solution for water molecules through normal hydration interactions of their amino acid residues. Halophilic proteins, however, seem to have evolved a folded structure that can coordinate hydrated salt ions on the surface, at a higher local concentration than in the bulk solvent. Generally, they are rich in the acidic residues glutamate and aspartate. A recent model of *H. volcanii* dihydrofolate reductase (Böhm & Jaenicke, 1994) has an asymmetrical charge distribution over the surface of the protein, with positive charges around the active site and negative charges on the opposite side of the molecule. The negative charges form clusters that are shielded at high salt concentrations but are repulsed at lower concentrations. This repulsion leads to the destabilisation of the protein in the absence of salt.

### 1.4.2 Alternative mechanisms for existence in high salt

In contrast to the halophilic Archaea, halotolerant Bacteria and methanogenic Archaea accumulate high concentrations of organic osmolytes to balance the osmotic pressure of the external environment (Severin et al., 1992; Lai et al., 1991). These compounds, now commonly known as "compatible solutes" (Brown, 1976) must fulfil their function while remaining physiologically inert.
A wide range of compatible solutes are involved in osmotic adaptation and include sugars, polyols, amino acids and derivatives of them, tetrahydropyrimidines, betaines and ectoines (a class of cyclic amino acids). In times of need, some compatible solutes can be synthesised de novo, but in preference, they are generally accumulated from the culture medium.

Some bacteria that are not usually halotolerant can be induced to limited tolerance by osmotic stress conditioning (Ishida et al., 1994). Some strains of *E. coli*, when exposed to a moderate osmotic stress (~0.5 M NaCl for instance) for a short time, in the order of 30 minutes, are subsequently able to grow in a medium containing up to 1.2 M NaCl. Inhibition of this acquired osmotolerance can be achieved by treatment of the cells with chloramphenicol or rifampicin, suggesting that synthesis of new protein or RNA may be important.

1.5 Dihydrolipoamide dehydrogenase (DHLipDH)

Dihydrolipoamide dehydrogenase (EC 1.8.1.4) catalyses the NAD⁺-dependent oxidation of dihydrolipoamide (Williams, 1976) through a catalytic mechanism involving the alternate oxidation and reduction of an intrachain disulphide bond and a base on the enzyme (Figure 1.4). It has a tightly, but non-covalently, bound FAD cofactor that participates in the transfer of electrons from dihydrolipoamide to NAD⁺.

The enzyme exists as a dimer, with the active site comprising residues from both subunits. Stopped-flow spectroscopy has revealed that the enzyme follows a ping-pong mechanism, which can be divided into two half-reactions, i) reduction of the oxidised enzyme to the reduced form, and ii) the transfer of electrons from the reduced form to the nicotinamide cofactor (Figure 1.5) (Massey et al., 1960; Massey & Veeger, 1961).
Figure 1.4: NAD⁺-dependent oxidation of dihydrolipoamide catalysed by dihydrolipoamide dehydrogenase (Danson, 1988).

\[
\begin{align*}
E + \text{lip(SH)}_2\text{NH}_2 & \xrightleftharpoons[k_2]{k_1} [E - \text{lip(SH)}_2\text{NH}_2] \\
& \xrightarrow{k_3} \text{EH}_2 - \text{lipS}_2\text{NH}_2 \\
& \xrightarrow{k_4} \text{EH}_2 + \text{lipS}_2\text{NH}_2 \\
\text{EH}_2 + \text{NAD}^+ & \xrightleftharpoons[k_6]{k_5} \text{[EH}_2 - \text{NAD} + E - \text{NADH}]} \\
& \xrightarrow{k_7} E + \text{NADH} + \text{H}^+ \\
\end{align*}
\]

Figure 1.5: Mechanism of DHLipDH. Key: E = oxidised enzyme, EH₂ = 2-electron-reduced enzyme, lip(SH)₂NH₂ = dihydrolipoamide, lipS₂NH₂ = lipoamide.

1.5.1 2-Oxo acid dehydrogenase multienzyme complexes

In all eukaryotes and most eubacteria DHLipDH is an integral component of the pyruvate, 2-oxoglutarate, and branched-chain 2-oxoacid dehydrogenase multienzyme complexes (Perham, 1991; Mattevi, de Kok & Perham, 1992) (Figure 1.6), as well as the glycine cleavage system (Kikuchi & Hiraga, 1982).
These multienzyme complexes consist of three components, a 2-oxoacid dehydrogenase (E1), a dihydrolipoamide transacetylase (E2) and a dihydrolipoamide dehydrogenase (E3). The E2 components form a structural core to which the E1 and E3 components are bound non-covalently. The pyruvate dehydrogenase (PDHC) and 2-oxoglutarate dehydrogenase (OGDHC) multienzyme complexes of Gram-negative bacteria and the OGDHC and branched-chain 2-oxo acid dehydrogenase (BCDHC) complexes of mammals contain 24 copies of the E2 component arranged with octahedral symmetry. In the PDHC of eukaryotes and Gram-positive bacteria, however, the core comprises 60 E2 chains arranged with icosahedral symmetry. These complexes have a high relative molecular mass of \((5-10) \times 10^6\). The number of E1 and E3 chains in each complex varies dependent on the symmetry and source, with for example, an E1:E2:E3 ratio of 1:1:0.5 in the PDHC of \(E. coli\).

**Figure 1.6:** The reaction mechanism of the 2-oxo acid dehydrogenase multienzyme complexes (modified from Perham, 1991). ThDP, thiamin diphosphate; Lip, lipoic acid.
The substrate is effectively channelled through the reactions by attachment to a lipoyl group, which is in turn attached to the E2 component. The lipoyl group is able to rotate freely within the complex, carrying the substrate in turn to each of the three separate reaction sites.

The E1 and E2 components of the complexes display specificity for the substrate of the overall reaction, whereas the E3 component is normally identical in the PDHC and 2-oxoglutarate dehydrogenase complexes, and is interchangeable between the two.

1.5.2 The Archaeal alternative

The Archaea, however, convert pyruvate and 2-oxoglutarate to acyl-CoA thioesters via less complex oxidoreductases (Kerscher & Oesterhelt, 1982) that involve ferredoxin as the electron acceptor, rather than NAD⁺, and contain no lipoic acid (Kerscher & Oesterhelt, 1981a,b) (Figure 1.7).

---

Figure 1.7: The catalytic cycle of a 2-oxoacid : ferredoxin oxidoreductase, found in Archaea and anaerobic Bacteria (modified from Kerscher & Oesterhelt, 1982). ThDP, thiamin diphosphate, Fd, ferredoxin.
It was, therefore, unexpected to find that halophilic, and methanogenic Archaea, as well as *Thermoplasma* sp., i.e. the Euryarchaeota, possess both a DHLipDH and lipoic acid (reviewed in Danson, 1993). As with the eukaryotic and eubacterial enzyme, DHLipDH from halophiles is specific for both NAD⁺ and dihydrolipoamide (Danson *et al.*, 1984), but the lack of 2-oxoacid dehydrogenase complexes in the Archaea suggests that the enzyme may fulfil an additional role not previously elucidated.

### 1.5.3 Other possible roles of DHLipDH

There are a number of cases where DHLipDH apparently exists as an uncomplexed enzyme rather than as a component of a multienzyme complex.

A second DHLipDH from *E. coli* has been implicated in galactose and maltose transport (Richarme, 1985; Richarme & Heine, 1986). The enzyme is located at the plasma membrane where the oxidoreduction of dihydrolipoic acid and NAD⁺ stimulates the transport. This DHLipDH was shown to be distinct from the normal gene product involved in the multienzyme complexes, and failed to restore PDHC and OGDHC activities in a mutant strain lacking the DHLipDH from these complexes.

A DHLipDH from the trypanosomatidae, the causative agents of African sleeping sickness in humans and nagana in cattle, is also associated with the plasma membrane (Danson *et al.*, 1987; Jackman *et al.*, 1990; Else *et al.*, 1993). The enzyme is present throughout the life-cycle of *Trypanosoma brucei*, including the bloodstream form where there is no functional citric acid cycle. In contrast to the case in *E. coli*, there is only one DHLipDH gene and no differences have been discovered between the complexed and uncomplexed forms of the enzyme.
1.5.4 Available data

A large number of primary sequences of DHLipDH from a wide range of organisms has been determined (Table 1-1).

<table>
<thead>
<tr>
<th>Organism</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acholeplasma laidlawii</td>
<td>Wallbrandt et al., 1992</td>
</tr>
<tr>
<td>Azotobacter vinelandii</td>
<td>Westphal &amp; De Kok, 1988</td>
</tr>
<tr>
<td>Bacillus stearothermophilus</td>
<td>Hawkins et al., 1990</td>
</tr>
<tr>
<td></td>
<td>Borges et al., 1990</td>
</tr>
<tr>
<td>Bacillus subtilis</td>
<td>Hemila et al., 1990</td>
</tr>
<tr>
<td>Escherichia coli</td>
<td>Stephens et al., 1983</td>
</tr>
<tr>
<td>Garden pea</td>
<td>Bourguignon et al., 1992</td>
</tr>
<tr>
<td></td>
<td>Turner et al., 1992</td>
</tr>
<tr>
<td>Haloferax volcanii</td>
<td>Vettakkorumakankav &amp; Stevenson, 1992</td>
</tr>
<tr>
<td>Human</td>
<td>Otulakowski &amp; Robinson, 1987</td>
</tr>
<tr>
<td></td>
<td>Pons et al., 1988</td>
</tr>
<tr>
<td>Pig</td>
<td>Otulakowski &amp; Robinson, 1987</td>
</tr>
<tr>
<td>Pseudomonas fluorescens</td>
<td>Benen et al., 1989</td>
</tr>
<tr>
<td>Pseudomonas putida 1</td>
<td>Burns et al., 1988</td>
</tr>
<tr>
<td></td>
<td>Burns et al., 1989</td>
</tr>
<tr>
<td>Pseudomonas putida 2</td>
<td>Palmer, Hatter &amp; Sokatch, 1991</td>
</tr>
<tr>
<td>Pseudomonas putida 3</td>
<td>Palmer, Madhusudhan, Hatter &amp; Sokatch, 1991</td>
</tr>
<tr>
<td>Saccharomyces cerevisiae</td>
<td>Browning et al., 1988</td>
</tr>
<tr>
<td></td>
<td>Ross et al., 1988</td>
</tr>
<tr>
<td>Trypanosoma brucei brucei</td>
<td>Else et al., 1993</td>
</tr>
</tbody>
</table>

Table 1-1: Available DHLipDH gene sequences

In addition, the 3D crystal structures of the enzyme from *Pseudomonas putida* (Mattevi, Obmolova, Sokatch, Betzel & Hol, 1992), *Pseudomonas fluorescens* (Mattevi et al., 1993) and *Azotobacter vinelandii* (Schierbeek et al., 1989; Mattevi et al., 1991) have been resolved.

1.5.5 DHLipDH from *H. volcanii*

The gene encoding DHLipDH from *Haloferax volcanii* has been cloned and sequenced (Vettakkorumakankav & Stevenson, 1992). The open reading frame codes for 475 residues and the sequence shows a high degree of homology with DHLipDH from a number of non-halophilic sources, especially around the active-
site thiols. The gene has a high percentage of G and C nucleotides (69%) although such a G+C content is not uncommon in halophilic genes.

The pH profile of the enzyme shows activity over a pH range of 6-10.5. It has an optimum pH of 9, significantly above the optimal growth pH of 7 for the organism (Vettakkorumakankav et al., 1992). It has been shown by NMR spectroscopy that the internal pH of the halophiles *Halobacterium halobium* and *H. marismortui* closely matches that of the external environment (Quirk & Campbell, 1990), which would suggest that the enzyme exists natively at a pH of about 7. At this pH the enzyme is approximately 50% active, based on the activity profile.

The participation of a redox disulphide bond in the active site of the enzyme was evidenced by studies using p-amino-phenyl arsenoxyde, a trivalent organoarsenical that selectively forms covalent bonds with two spatially close thiols (Vettakkorumakankav et al., 1992). Rapid inactivation of the enzyme was observed upon exposure to the arsenical, indicating the presence of the redox disulphide in a position accessible to the substrate.

### 1.5.6 Molecular model of halophilic DHLipDH

The three-dimensional structure of DHLipDH from *H. volcanii* has been modelled on the homologous structure from *Pseudomonas fluorescens*, a psychrotrophic bacterium (Dr. R.J.M. Russell, University of Bath, unpublished) (Figure 1.8). The two molecules share a high degree of homology (43% identity, 65% similarity of residues).

Characteristic of other halophilic proteins, the surface of halophilic DHLipDH has a high degree of negative charge relative to non-halophilic homologues. Each monomer has 31 basic and 86 acidic residues, resulting in a net charge of -55 if all side chains were present in their conjugate forms. In
contrast, the P. fluorescens structure has a charge of -7 per monomer. In the absence of salt it is predicted that the molecular surface of the protein is predominantly negatively-charged, whereas at a salt concentration of 1 M NaCl the charge profile is comparable to non-halophilic proteins in their native environment. This is due to a screening of charges rather than to a change of conformation.

The molecular model of dihydrofolate reductase from H. volcanii (Böhm & Jaenicke, 1994) has a number of negatively-charged clusters on the surface, and a positively-charged area near the active site that is thought to facilitate orientation of the substrate. The model of DHLipDH, however, suggests that the entire surface of the protein forms a macrocluster. Around the active site the prevalence of acidic residues is approximately half, and that of basic residues approximately double, that over the surface taken as a whole. Again, this is probably due to the need for orientating the substrates.

Protein stability can be increased by the presence of disulphide bonds. Although the structure of DHLipDH from H. volcanii has an extra two cysteine residues compared to P. fluorescens, they do not appear to be involved in disulphide bond formation.

The enzyme's active site appears to be structurally identical to that of P. fluorescens, and is conserved throughout the available structures. The sequence between the cysteines of the redox disulphide bridge (CLNYGC) is changed by one residue from that conserved between other organisms (CLNVGC). The overall shape of the molecule bears close resemblance, with the main significant differences being in the loop regions. The positions of the α-helices and β-sheets are relatively conserved.
Figure 1.8: Molecular model of *H. volcanii* DHLipDH, homology-modelled on the *P. fluorescens* DHLipDH structure.
1.5.7 Use of halophilic DHLipDH as a model enzyme

The halophilic DHLipDH exists as a homodimer with a total relative molecular mass (M_r) of 119,000, determined by gel filtration (Danson et al., 1984). It displays considerable thermal stability, showing no detectable loss of activity after a 15 min incubation at 95°C, whereas other halophilic enzymes are inactivated at much lower temperatures (Danson et al., 1984). As with many other halophilic enzymes, DHLipDH is shown to be inactive when assayed in the absence of salt. Re-addition of 2 M NaCl to the inactive protein, however, results in complete restoration of activity almost immediately, even after storage for 24 h at 4°C in the absence of salt (Danson et al., 1986). This stability makes the enzyme a good candidate for studies into the structural properties of halophilic enzymes.

The discovery of DHLipDH in the Archaea means that it is present in all three evolutionary domains, suggesting that it may be an “ancient” enzyme that evolved before the divergence from the common progenote. This provides the opportunity to conduct phylogenetic analysis, as well as comparative enzymology.

1.6 The biotechnological potential of halophiles and their proteins

The Archaea are a diverse group of organisms with certain physical characteristics very different from most eubacterial cells. It is this sheer diversity that provides the huge potential of archaeal systems to biotechnological applications.

The proton-pumping, photosensitive membranes of certain halobacteria have enormous potential for use as optical switching elements in bio-electronics (Hong, 1986), from uses such as optical recording materials to components of opto-molecular computers (Moses, 1991). Bacteriorhodopsin has been isolated
in a stable and active form, that is readily immobilised on solid substrates, or embedded in polymers. The photoreceptor has been extensively studied in molecular detail and its responses found to be extremely reproducible (Hong, 1986).

The considerable halophilicity, and in many cases thermostability, of enzymes from halobacteria has great potential for industrial applications. Reactions in organic solvents can be favoured by the use of halophilic enzymes (Klibanov, 1989) and such conditions may have the advantage of altered substrate and enantiomeric specificities, due to their increased rigidity, as well as improved yields. Halobacterial enzymes are much more efficient at maintaining their hydration layer in low-water conditions than non-halophilic proteins. Such enzymes would be expected to be suitable for use in processes that require the presence of organic solvents and where water activity is low, such as where high substrate or product concentrations are employed (Hough & Danson, 1989).

Some halobacteria, such as *Haloferax* (Antón et al., 1988) and *Haloarcula*, produce extracellular polysaccharides. These compounds have been shown to have good rheological properties (high viscosity at low concentration), and are resistant to salt, temperature and pH extremes. Applications exist for such compounds, especially in the field of oil recovery.

The high molecular weight polymer, poly-β-hydroxybutyrate (PHB) can be used for the production of commodity plastics that are both biodegradable and biocompatible. At present, this polymer is harvested from *Alcaligenes eutrophus* by ICI to produce the plastic Biopol™. *Haloferax mediterranei* has been shown to accumulate high levels of PHB (Lillo & Rodriguez-Valera, 1990), with the advantage that starch can be used as a carbon source rather than glucose, which is required for *A. eutrophus*. Since starch is considerably cheaper than
glucose the cost of the process could be reduced. The fragility of halobacteria when exposed to low salt concentrations may also enable a simplified recovery process.

Fundamental research into the structural properties of halophilic enzymes is important for the long-term goal of engineering halophilicity into existing enzymes with useful catalytic properties. Although the favoured method of commercial enterprises for discovering new activities is to screen new microorganisms, there may be instances where rational engineering of existing non-halophilic enzymes is more appropriate. Even if this is not shown to be the case, one could argue that the fundamental understanding that such research endows is invaluable.

1.7 Project aims

Our research group at Bath has the aim of investigating the structural features that enable archaeal enzymes to be functionally active in extremes of temperature and salinity. To achieve this goal, the three-dimensional structures of equivalent enzymes from organisms living across the range of conditions, from mesophilic to the most extreme, are compared. Coupled with this is the testing of hypotheses and deductions with mutational studies.

The aim of this project has been to express dihydrolipoamide dehydrogenase from the halophilic archaeon *H. volcanii* at high levels, either in a non-halophilic host such as *E. coli* or in a mutant strain of *H. volcanii* that lacks a functional DHLipDH. Using the developed expression system, sufficient protein was to be purified to enable crystallisation trials to be undertaken, with the long-term aim of resolving a three-dimensional crystal structure for the enzyme. Prior to this development, a molecular model of the structure would be used to analyse and design site-directed mutants, which could be then expressed in the system.
developed, to test particular features that may play a role in conferring halophilicity.

In addition to this, investigations were to be undertaken to probe the function of the enzyme in the Archaea. Comparisons of the growth characteristics of wild-type *H. volcanii* and a strain with the gene inactivated were to be made to try and assess the effect of the deficiency. To complement this, the region upstream of the DHLipDH gene was to be sequenced to enable database searching for homologous sequences that may provide further clues as to function.
2. Materials and Methods

2.1 Materials

Cell culture: Yeast extract, tryptone and bacto-agar were supplied by Difco Laboratories, MI, USA. Bacteriological peptone was from Oxoid, UK. Ampicillin (sodium salt), tetracycline and novobiocin were obtained from Sigma, UK. Mevinolin was a gift of Merck Research Laboratories (USA).

Molecular biology: Restriction enzymes were supplied by Pharmacia, Sweden; New England Biolabs, MA, USA or Gibco BRL Life Technologies Ltd., UK. Taq DNA polymerase, T4 polynucleotide kinase and T4 DNA ligase were from Boehringer Mannheim (Germany). Vent™ DNA polymerase and mung bean nuclease were purchased from New England Biolabs. 1 kb ladder DNA markers were obtained from Gibco. Phenol : chloroform : isooamyl alcohol (25:24:1) and RNAse A were supplied by Sigma. Magic™ Miniprep and Altered Sites™ II mutagenesis kits were obtained from Promega Corporation, WI, USA. QiAprep™ spin columns were supplied by Qiagen, UK. Sequenase™ version 2.0 kits were supplied by United States Biochemical Corp., OH, USA. Geneclean® II kits were from Bio 101, CA, USA. IPTG and X-Gal were purchased from Calbiochem, UK. [α-^35S] dATP Redivue was from Amersham International, UK. X-Ray film was purchased from Fuji Photo Film Co., Japan.

Plasmids: pKK223-3 and pGEX-KG were obtained from Pharmacia and pMEX8 was from Medac (Germany).

Chromatography: Hydroxylapatite Bio Gel HT was supplied by Bio-Rad Laboratories Ltd., UK. Glutathione agarose was from Sigma.

Enzyme assays: DL-lipoamide and oxaloacetate were from Sigma and coenzyme A was from Calbiochem, UK. NAD^+ was from Boehringer Mannheim.
Electrophoresis: Acrylamide Sequagel™ and Protogel™ were purchased from National Diagnostics (USA). TEMED, agarose (standard and low melting point) and ethidium bromide solution came from Sigma. Protein low molecular weight standards were obtained from Bio-Rad Laboratories Ltd.

Crystallisation: Crystal Screen™ reagents were supplied by Hampton Research, CA, USA.

Oligonucleotides were ordered from Pharmacia, Perkin Elmer or Severn Biotech, UK. BCA reagent A was from Pierce, IL, USA. All other reagents (SLR, Analar and Aristar grades) were purchased from BDH Ltd., UK, Sigma, UK and Fisons, UK.

DNA and amino acid sequence analysis were carried out using the GCG Sequence Analysis Software Package Version 7.0. Molecular visualisation was predominantly conducted using Rasmol Version 2.5 (PC) written by R. Sayle (Glaxo Group Research).

2.2 Strains and culture conditions

H. volcanii WFD11, a derivative of H. volcanii DS2 cured of the endogenous plasmid pHV2 (Charlebois et al., 1987) (obtained from Dr. M. Dyall-Smith, University of Melbourne, Australia) and strains derived from it were grown in 18% (w/v) salt water modified growth medium (MGM), consisting of 14.4% (w/v) NaCl, 1.8% (w/v) MgCl₂.6H₂O, 2.1% (w/v) MgSO₄.7H₂O, 0.42% (w/v) KCl, 0.5% (w/v) peptone, 0.1% (w/v) yeast extract, pH 7.2, at 37°C with shaking at 200 rpm. Minimal medium was used for growth experiments and consisted of the same salts as modified growth medium with 0.5% (v/v) glycerol, 0.05% (w/v) succinate, 10 mM NH₄Cl, 1 mM K₂HPO₄ buffer, 20 mM Tris-HCl (pH 7.5), and 0.1% (v/v) trace elements solution. Trace elements solution consisted of 0.36 mg/ml MnCl₂.4H₂O, 0.44 mg/ml ZnSO₄.7H₂O, 3.3 mg/ml FeSO₄.7H₂O and 0.05% (w/v)
CuSO₄·5H₂O. Where required the medium was solidified by addition of 1.5% (w/v) agar. Frozen stocks were maintained at -20°C and -70°C in the same medium containing 30% (v/v) glycerol.

Antibiotic resistance plasmids were maintained or selected for in solid and liquid media by addition of 0.3 µg novobiocin/ml or 4 µg mevinolin/ml.

_E. coli_ strains TG1 (supE, hsdΔ5, thi, Δ(lac-proAB) F'[traD36 proAB lacF' lacZΔM15], DH5α (supE44, ΔlacU169, (φ80 lacZΔM15), hsdR17, recA1, endA1, gyrA96, thi-1, relA1); JM105 (thi, rpsL, endA, sbcB15, hspR4, Δ(lac-proAB), [F' traD36 proAB lacF' lacZΔM15]); JM109 (endA1, recA1, gyrA96, thi, hsdR17 (r-, mR+), relA1, supE44, λ-, Δ(lac-proAB), [F' traD36, proAB, lacF' lacZΔM15]); BMH 71-18 mutS (thi, supE, Δ(lac-proAB) [mutS::Tn10][F' proAB, lacF' lacZΔM15]) and ES1301 mutS (lacZ53, mutS201::Tn5, thyA36, rha-5, metB1, deoC, IN(mD-rmE)) and strains derived from them, were grown in LB broth (1% (w/v) tryptone, 0.5% (w/v) yeast extract and 1% (w/v) NaCl) at 37°C with shaking at 200 rpm. F' strains were maintained on minimal M-9 plates (0.6% (w/v) Na₂HPO₄, 0.3% (w/v) KH₂PO₄, 0.05% (w/v) NaCl, 0.1% (w/v) NH₄Cl, 0.2% (w/v) glucose, 2 mM MgSO₄, 100 µM CaCl₂, pH 7.4) supplemented with 1 mM thiamine-HCl. JRG1342 (Δ(aroP-lpD), DE74, metB1, met-105, azi, pox, pps-1, tsx-87?, ton?, relA1, rpsL, recA1) (Guest et al., 1985) was grown in LB broth supplemented with 1% (w/v) glucose. Where required the medium was solidified by addition of 1.5% (w/v) agar. Frozen stocks were maintained at -20°C and -70°C in the same medium containing 30% (v/v) glycerol.

Antibiotic resistance plasmids were maintained or selected for in solid and liquid media by addition of 100 µg ampicillin/ml or 12.5 µg tetracycline/ml (solid media) or 10 µg tetracycline/ml (liquid culture).
2.3 Biochemical Techniques

2.3.1 Preparation of cell extracts

*E. coli* cells were harvested by centrifugation at 10,800 x g in a Sorvall GSA rotor for 25 min at 4°C. The cells were then resuspended in 10 mM Tris-HCl (pH 8.0), 1 mM EDTA containing 1 mM PMSF at a density of 0.2 g/ml and lysed by sonication on ice in an MSE 150 W Ultrasonic Disintegrator Mk2 using a 9.5 mm end diameter titanium probe. Three 30 s bursts were used with a 90 s cooling time between each. Cell debris was removed by centrifugation at 15,600 x g for 30 min at 4°C in a Sorvall SS34 rotor.

For small volumes, cells were harvested at full speed in a microcentrifuge for 2 min in 1.5 ml Eppendorf tubes, and resuspended in a reduced volume. The samples were sonicated using a 3 mm end diameter exponential probe. Three 10 s blasts were used with approximately 1 min cooling time between each. Cell debris was removed by further centrifugation at full speed for 10 min.

*H. volcanii* extracts were prepared in a similar manner with the exception that cells were resuspended in 2 M KCl, 10 mM Tris-HCl (pH 8.0), 1 mM EDTA prior to sonication.

2.3.2 DHLipDH enzyme assay

Catalytic activity was assayed spectrophotometrically at 30 or 37°C using a Perkin Elmer Lambda Bio or Lambda 11 spectrophotometer. The reaction reduces NAD⁺ to NADH, which has an absorption maximum at 340 nm and a molar extinction coefficient of 6,200 l mol⁻¹ cm⁻¹. Initial rates were calculated using the Perkin Elmer Computer Spectroscopy Software (PECSS) version 4.31 supplied with the instrument.
Unless otherwise stated, assays were carried out in 2 M KCl, 10 mM Tris-HCl (pH 8.0), 1 mM EDTA, 1 mM NAD$^+$ and 0.4 mM dihydrolipoamide, in a final volume of 1 ml. 1 unit of enzyme activity is defined as the amount required to produce 1 µmol of product per min.

2.3.3 Preparation of dihydrolipoamide

DL-dihydrolipoamide was prepared by the reduction of DL-lipoamide with NaBH$_4$ (Reed et al., 1958). 800 mg DL-lipoamide was dissolved in 20 ml methanol : water (4:1) over ice. 800 mg NaBH$_4$ was added in 4 ml cold distilled water, and the mixture stirred over ice for 2 h. The solution was acidified with 1 M HCl until effervescence stopped. The product was then extracted three times with 50 ml chloroform, transferred to a round-bottomed flask and rotary evaporated to dryness. The resulting crystals were dissolved in 40 ml toluene with gentle warming. 16 ml hexane was added and the precipitate collected by vacuum filtration, followed by drying in a desiccator.

Purity was determined spectrophotometrically by titration with 5,5'-dithio bis-2-nitrobenzoate (DTNB). 0.2 mM DTNB was reacted with 0.05 mM dihydrolipoamide (0.1 mM thiol groups) in a volume of 1 ml TE buffer. For 100% purity the expected absorbance at 412 nm was 1.36 (E$_m$=13,600 l mol$^{-1}$ cm$^{-1}$). Preparations were shown to be over 98% pure.

2.3.4 Citrate synthase assay

Activity was assayed spectrophotometrically at 30°C in a Perkin Elmer Lambda Bio or Lambda 11 spectrophotometer by the method of Srere et al. (1963). Free coenzyme A reacts with DTNB, releasing thionitrobenzoate. This has an absorption maximum at 412 nm and a molar extinction coefficient of 13,600
I mol\(^{-1}\) cm\(^{-1}\). Initial rates were calculated using the PECSS software supplied with the machine.

Assays were carried out in 2 M KCl, 20 mM Tris-HCl (pH 8.0), 2 mM EDTA, 0.2 mM oxaloacetate, 0.4 mM acetyl CoA and 0.1 mM DTNB, in a final volume of 1 ml.

2.3.5 Acetylation of coenzyme A

Coenzyme A was acetylated by dissolving 10 mg in 1 ml Milli-Q water, cooling on ice and adding 0.2 ml 1 M KHCO\(_3\), followed by 0.1 ml 1 M acetic anhydride freshly-diluted with Milli-Q water. After incubation on ice for 10 min, acetylation was checked by addition of 20 \(\mu\)l to 1 ml of assay buffer containing 100 \(\mu\)M DTNB. A yellow colour indicated incomplete acetylation and more acetic anhydride was added as well as more KHCO\(_3\) to maintain the pH. The final concentration of acetyl-CoA was \(~\)7.5 mM.

2.3.6 SDS polyacrylamide gel electrophoresis (SDS-PAGE)

Samples were boiled for 5 min in loading buffer (63 mM Tris, pH 6.8, 2\% (w/v) SDS, 10\% (w/v) sucrose, 5\% (v/v) \(\beta\)-mercaptoethanol, 0.02\% (w/v) bromophenol blue). They were then loaded onto a 10\% acrylamide gel. The resolving gel consisted of 4 ml Protogel\textsuperscript{TM} (30\% (w/v) acrylamide, 0.8\% (w/v) N N'-methylene bisacrylamide stock solution), 3 ml buffer (1.5M Tris-HCl, pH 8.9, 0.4\% (w/v) SDS), 5 ml dH\(_2\)O, 50 \(\mu\)l 10\% (w/v) ammonium persulphate and 12.5\(\mu\)l TEMED. The stacking gel consisted of 0.9 ml Protogel\textsuperscript{TM}, 2.4 ml buffer (487 mM Tris-HCl, pH 6.8, 0.4\% (w/v) SDS), 3.6 ml dH\(_2\)O, 50\(\mu\)l 10\% (w/v) ammonium persulphate and 10\(\mu\)l TEMED.

Gels were run in 52 mM Tris-HCl, 0.1\% (w/v) SDS, 0.4\% (w/v) glycine on an Atto Corp. Mini-Atto system at 10 mA until samples entered the resolving gel.
and then at 20 mA. The samples were stained for 30 min in Coomassie stain (0.25% (w/v) Coomassie Blue R, 45.4% (v/v) methanol, 9.2% (v/v) glacial acetic acid) and then destained for a few days in 7.5% (v/v) glacial acetic acid, 5% (v/v) methanol.

2.3.7 Chromatographic techniques

2.3.7.1 Hydroxylapatite chromatography

A 2.2 cm x 20 cm column of hydroxylapatite Bio-Gel HT was equilibrated with 50 mM potassium phosphate buffer, pH 7.0, containing 1 M NaCl. The sample was applied and the column washed with equilibration buffer until all unbound solutes were removed. The product was eluted by the application of a phosphate gradient (50 mM to 500 mM, pH 7.0), in the presence of 1 M NaCl, over a volume of 400 ml at a flow rate of 30 ml/h.

2.3.7.2 FPLC Superdex 200 gel filtration

The column was equilibrated with 50 mM potassium phosphate buffer, pH 7.0, containing 1 M NaCl. The sample was loaded using a 2 ml loop and eluted at a flow rate of 60 ml/h. 1 ml fractions were collected after the void volume (30 ml).

2.3.7.3 Glutathione agarose affinity chromatography

The gel was prepared by swelling lyophilised glutathione agarose in water for 2 h. This was packed into 2 ml disposable columns, avoiding the introduction of air bubbles, and washed with 10 column volumes of water added in several aliquots. The column was then equilibrated with several column volumes of PBST buffer (150 mM NaCl, 16 mM Na₂HPO₄, 4 mM NaH₂PO₄, 1% Triton X-100, pH 7.3).

The sample was loaded in 3 ml PBST buffer supplemented with 2 mM EDTA, 0.1% (v/v) β-mercaptoethanol, 0.2 mM PMSF and 5 mM benzamidine.
The column was washed with 8 ml PBST and the sample eluted with 2.5 ml cleavage buffer (50 mM Tris-HCl (pH 8.0), 150 mM NaCl, 2.5 mM CaCl₂, 0.1% (v/v) β-mercaptoethanol) containing 10 mM glutathione.

The column was regenerated by washing with 5 volumes 0.1 M borate buffer, pH 8-9, containing 0.5 M NaCl, followed by 5 column volumes of water, 5 column volumes of 0.1 M acetate buffer, pH 4, containing 0.5 M NaCl, followed by another 5 volumes of water.

Alternatively, a batch procedure was used whereby glutathione agarose was added directly to the sample and incubated for 30 min at 4°C. Washing and elution was accomplished by centrifugation at 1000 rpm for 5 min in a microcentrifuge, followed by removal of the supernatant. The same buffers were used as for the columns.

2.3.7.4 Copper chelate affinity chromatography

Copper chelate affinity chromatography was carried out using a 5 ml HiTrap chelating column (Pharmacia). The column was primed with 2.5 ml 0.1 M CuSO₄ using a syringe, washed with 15 ml Milli-Q water and equilibrated with 35 ml 50 mM potassium phosphate buffer, pH 7.0, containing 1 M NaCl. The sample was loaded and the column washed with 25 ml equilibration buffer. Proteins were eluted with an imidazole gradient (0 - 100 mM) in 50 mM potassium phosphate buffer, pH 7.0, containing 1 M NaCl, over 16 ml at a flow rate of 2 ml/min. 1 ml fractions were collected.

The column was regenerated by stripping the metal ions with 25 ml equilibration buffer supplemented with 50 mM EDTA, washed with 25 ml Milli-Q water and stored in 20% ethanol.
2.3.8 *TCA precipitation of proteins*

Proteins were precipitated in 10% (w/v) trichloroacetic acid (TCA) for 30 min on ice. The precipitates were spun down at full speed in a microcentrifuge for 10 min, washed three times with ethanol : diethyl ether (1:1), and dried under vacuum.

2.3.9 *Protein estimation*

2.3.9.1 *Bradford Assay*

100 µl of appropriately-diluted sample was mixed with 1 ml Bradford Reagent (0.1 mg/ml Coomassie Brilliant Blue G-250, 10% (v/v) phosphoric acid, 5% (v/v) ethanol) in a cuvette. This was incubated at room temperature for approximately 10 min. The absorbance at 590 nm was determined and the protein concentration interpolated from known standards. The standards used were bovine serum albumin over a range of concentrations from 0 to 200 µg/ml (Bradford, 1976).

2.3.9.2 *BCA Assay*

A 50 µl sample was mixed with 1 ml working reagent in a cuvette. The working reagent consisted of 50 parts BCA Reagent A (1% (w/v) bicinchoninic acid, 2% (w/v) Na₂HCO₃·H₂O, 0.16% (w/v) Na₂ tartrate, 0.4% (w/v) NaOH, 0.95% (w/v) NaHCO₃, pH 11.25) to 1 part Reagent B (4% (w/v) CuSO₄). The cuvette was incubated at 37°C for 30 min and then the absorbance at 562 nm was determined. Protein concentration was interpolated from known standards. The standards used were bovine serum albumin over a range of concentrations from 0 to 1200 µg/ml (Smith *et al.*, 1985).
2.3.10 Crystallisation trials

Protein crystallisation trials were performed using the hanging drop vapour diffusion method (McPherson, 1990). Drops contained 2 μl concentrated protein solution (10 mg/ml) mixed with 2 μl of the precipitant solution. These were placed on silanised glass cover slips and suspended over 700 μl of the precipitant solution in 24-well tissue culture plates. A seal was formed between the cover slip and the well plate by a layer of vacuum grease. The trials were incubated at room temperature over a period of weeks.

2.4 Molecular Biological Methods

2.4.1 Quantitation of DNA

DNA concentration and purity were determined spectrophotometrically by the measurements of absorbance at 260 and 280 nm. An absorbance of 1 at 260 nm (1 cm path length) corresponds to approximately 50 μg/ml (double-stranded) DNA. An $A_{260} / A_{280}$ ratio of greater than 1.8 was taken to indicate purity. (Sambrook et al., 1989).

2.4.2 Ethanol precipitation of DNA

Ethanol precipitation was used to concentrate a DNA sample or perform an exchange of solvent. Three volumes of absolute ethanol (stored at -20°C) and one tenth volume of 3 M sodium acetate, pH 5.2 (0.3 M final concentration) was added to the sample. This was incubated at -20°C for 30 min or longer and the DNA recovered by centrifugation at top speed in a microcentrifuge for 10 min. The DNA pellet was washed by adding a small volume (~400 μl in a 1.5 ml Eppendorf tube) of 70% (v/v) ethanol, incubating on ice for 2 min and spinning for
2 min at top speed. The ethanol was removed, the pellet dried under vacuum and then resuspended in the desired volume of buffer.

2.4.3 DNA preparation methods

2.4.3.1 Preparation of H. volcanii genomic DNA

200 μl of stationary phase culture was harvested and resuspended quickly in the same volume of distilled water to lyse the cells. An equal volume of buffer-saturated phenol (pH 8.0) was added, incubated at 65°C for 10 min and the aqueous layer removed. The DNA was ethanol precipitated out of solution, washed and resuspended in 200 μl distilled water. DNA was further purified using Geneclean®.

2.4.3.2 Large scale (Maxi) preparations of E. coli plasmid DNA

100 ml of cells from an overnight culture were harvested by centrifugation at 1,300 x g in a bench top centrifuge for 10 min. The cells were resuspended in 4 ml 50 mM glucose, 25 mM Tris-HCl (pH 8.0), 10 mM EDTA and lysed with addition of 8 ml 0.2M NaOH, 1% SDS. After standing on ice for 10 min, 6 ml 3M KAc (pH 4.8) was added quickly and the mixture was left on ice for 15-30 min. The precipitate was removed by centrifugation. Nucleic acids were precipitated from the supernatant by addition of 18 ml isopropanol followed by incubation on ice for 30 min. The precipitate was removed by centrifugation and resuspended in 1 ml 10 mM Tris-HCl (pH 8.0), 1 mM EDTA. 1 ml of 6 M LiCl was added and the mixture was left on ice for 15 min. The precipitate was removed by centrifugation and DNA was isolated from the supernatant by ethanol precipitation. The DNA was resuspended in 100 μl 10 mM Tris-HCl (pH 8.0), 1 mM EDTA and incubated with 0.1 mg/ml RNase A at 37°C for 30 min. The
RNAse was inactivated by adding 0.5% (w/v) SDS and heating to 75°C for 10 min. An equal volume of 6 M LiCl was added and the mixture was left for 15 min at room temperature. The precipitate was removed and the DNA extracted from the supernatant by ethanol precipitation. Following centrifugation for 10 min in a microcentrifuge, the pellet was dissolved in 400 μl Milli-Q water. The DNA solution was extracted once with phenol : chloroform : isoamyl alcohol (25:24:1) and twice with chloroform : isoamyl alcohol (24:1). This was followed by ethanol precipitation and resuspension in 100 μl Milli-Q water.

**2.4.3.3 Small scale (Mini) preparations of plasmid DNA - Alkaline lysis method**

1.5 - 3 ml of an overnight culture was spun down at full speed in a microcentrifuge for 2 min. The medium was removed by aspiration, leaving the bacterial pellet as dry as possible. The cell pellet was then resuspended in 100 μl ice-cold miniprep lysis buffer (50 mM glucose, 10 mM EDTA, 25 mM Tris-HCl, pH 8.0), and incubated for 5 min at room temperature. 200 μl of a freshly prepared solution containing 0.2 M NaOH, 1% (w/v) SDS was added and mixed by inversion. This was incubated on ice for 5 min. 150 μl ice-cold 3 M potassium acetate solution, pH 4.8, was added, mixed by inversion and incubated on ice for 5 min. The mixture was then spun down at full speed in a microcentrifuge for 5 min and the supernatant transferred to a fresh tube. 1 volume of phenol : chloroform : isoamyl alcohol (25:24:1) was added, mixed by vortexing for 1 min and centrifuged at full speed for 2 min. The upper, aqueous phase was transferred to a fresh tube and 1 volume of chloroform : isoamyl alcohol (24:1) added and mixed as before. The aqueous phase was again transferred to a fresh tube and ethanol precipitated, as previously described. The DNA pellet
was resuspended in 50 μl Milli-Q water and 0.5 μl 10 mg/ml RNase added and incubated for 5 min at room temperature.

Alternatively, Promega's Magic™ Miniprep or Qiagen's QIAprep™ kits were used in accordance with the manufacturer's instructions.

2.4.3.4 Preparations of *H. volcanii* plasmid DNA

Plasmids isolated from *H. volcanii* were prepared using the same procedures as ones from *E. coli* with the exception that the cells were initially resuspended in 1 M NaCl, 50 mM glucose, 25 mM Tris-HCl (pH 8.0), 10 mM EDTA.

2.4.4 Agarose gel electrophoresis

DNA fragments were routinely analysed and separated by horizontal agarose gel electrophoresis, using 1% (w/v) gels. These were made by dissolving agarose in the desired volume of TAE buffer (40 mM Tris-acetate, 1 mM EDTA) by heating in a microwave oven. Ethidium bromide was added to a concentration of 0.5 μg/ml and the hot agarose poured into a perspex gel mould, with a comb in place to form wells, and allowed to set either at room temperature or at 4°C. Once set, the comb was removed and the gel placed in an electrophoresis tank and covered in TAE buffer. Sample loading buffer (6x) was 40% (w/v) sucrose, 0.25% (w/v) bromophenol blue, 0.25% (w/v) xylene cyanol. Samples were electrophoresed at a constant voltage of 50 - 90 V. In cases where the DNA was to be excised and purified from the gel, ultra-pure low-melting point agarose was used. The size-separated DNA was visualised under UV transillumination.
2.4.5 Recovery and purification of DNA from agarose gels

2.4.5.1 Sephadex recovery

A Sephadex G-50 spin-column was made by plugging the bottom of a 0.6 ml Eppendorf tube with a small amount of sterile silane-treated glass wool. A volume of Sephadex G-50 dry beads (equilibrated with TE buffer, autoclaved and dried) was placed on top of the glass wool. The tube was spun at 3,300 x g in a swing-out rotor to pack the resin. The DNA was electrophoresed in 1% (w/v) low-melting-point agarose in TAE buffer as described. The desired band was excised from the gel under long wave UV light and the gel slice placed on top of the Sephadex beads. The tube was closed and the gel melted in a 60°C water bath for 5-10 min. As the gel slice melted, it entered the Sephadex resin. A volume of TE8 buffer was added and incubated for a further 5 min at 60°C. The tube was then placed on ice for 5 min. Small holes were made in the top and bottom of the tube with a 21-gauge needle. The tube was placed in a 1.5 ml Eppendorf tube and spun at 3,000 x g in a refrigerated microcentrifuge at 4°C. The flow-through contained approximately 80% of the total DNA. (Method modified from Mukhopadhyay & Roth, 1991).

2.4.5.2 Geneclean® method

The Geneclean® method uses a silica-based matrix called Glassmilk® that has a high affinity for DNA. The DNA was purified from agarose as detailed in the manufacturer's instructions.

2.4.6 Restriction digestion of DNA

DNA was digested with the appropriate restriction endonuclease as specified in the manufacturer's instructions. The DNA was incubated with the enzyme in
buffer, supplied as a 10x concentrate with the enzyme, at 37°C (or 25°C for Smal). For digestion with two endonucleases, the enzymes could be used simultaneously if the buffer requirements were similar. If this was not the case, however, the enzymes were used sequentially, digesting first with the enzyme requiring the lowest buffer concentration. The buffer concentration was then adjusted to suit the second reaction and the enzyme added. Alternatively, the DNA was ethanol precipitated after the first incubation and resuspended in the second reaction buffer. Incubation times varied from 1 h to overnight.

2.4.7 Ligation

Ligation of cohesive ends was carried out using 1 unit of T4 DNA ligase (supplied at a concentration of 1 unit/μl) in the buffer supplied with the enzyme. A ratio of insert to vector of 2 - 5 to 1 was used in a total volume of 30 μl. The reaction was incubated overnight at 15°C or at room temperature for 3 h. 15 μl of the reaction mixture was used to transform competent *E. coli* cells, with the remainder kept at 4°C for use if the first transformation failed.

2.4.8 Phosphorylation of oligonucleotides

100 pmol of each oligonucleotide was 5' phosphorylated using 5 units of T4 polynucleotide kinase (supplied at a concentration of 10 units/μl) in the buffer supplied with the enzyme, supplemented with 1 mM ATP. The final volume was 25 μl. The reaction was incubated at 37°C for 30 min and the kinase then inactivated by incubating at 70°C for 10 min. Phosphorylated oligonucleotides were stored at -20°C.
2.4.9 Removal of single-stranded overhangs

Single-stranded DNA overhangs, produced by cutting with restriction enzymes, were removed to produce blunt ends by incubation with mung bean nuclease. The DNA was suspended in reaction buffer (supplied with the enzyme) or in NEBuffer 1, 2 or 4 supplemented with 1 mM ZnSO₄ at a concentration of 0.1 μg/μl. 1 unit of enzyme was added for each μg of DNA and the reaction was incubated at 30°C for 30 min. After this time, the enzyme was inactivated by addition of SDS to a concentration of 0.01% (w/v) and the DNA recovered by ethanol precipitation.

2.4.10 Site-directed mutagenesis

Site-directed mutants were produced using Promega's Altered Sites® II in vitro mutagenesis system. The protocols supplied with the kit were followed with a few modifications. Firstly, the length of the DNA denaturation stage was increased from 5 min at room temperature to 20 min at 37°C, and secondly, BMH 71-18 mutS cells were substituted for the ES1301 mutS strain supplied with the kit due to problems encountered in making the latter competent.

2.4.11 Polymerase chain reaction (PCR) amplification of DNA

Reactions contained approximately 100 ng target DNA, 0.1 μM of each primer, and 50 μM each of dATP, dGTP, dCTP and dTTP in a 100 μl final volume. For reactions with Vent® DNA polymerase the reaction also contained 20 mM Tris-HCl (pH 8.8), 10 mM KCl, 10 mM (NH₄)₂SO₄, 2 mM MgSO₄ and 0.1% (v/v) Triton X-100. For reactions with Taq DNA polymerase the reaction contained 10 mM Tris-HCl (pH 8.3), 50 mM KCl and 1.5 mM MgCl₂. The reaction mixture was overlaid with 50 μl mineral oil to prevent evaporation. Reactions were incubated
at 96°C for 5 min before addition of 1-2.5 units of Vent® DNA polymerase (exo+) or Taq DNA polymerase. The amplification program varied, but consisted of a melting step (usually 96°C for 75 s), an annealing step and an extension step of 72°C (length determined by size of expected product, usually 1 min per kilobase). A final 10 min extension was run after the last cycle was complete. A 10 µl sample of the reaction mixture was visualised on a 1% (w/v) agarose gel.

2.4.12 DNA sequencing

DNA for sequencing was prepared using the maxiprep method previously described. Sequencing was performed using the dideoxy method of Sanger et al. (1977) using a Sequenase™ version 2.0 kit. Approximately 5 µg double-stranded DNA was alkali-denatured by incubating in freshly prepared 200 mM NaOH, 0.1 mM EDTA in a volume of 25 µl, at 37°C for 20 min. The DNA was ethanol precipitated, with two washes in 70% (v/v) ethanol, and resuspended in 6 µl Milli-Q water.

10 pmol of primer was added to the denatured DNA and annealed in reaction buffer (supplied with the Sequenase™ kit) in a total volume of 10 µl. Annealing was achieved by heating to 65°C for 2 min and cooling to 30°C over a period of 30 min in a water bath. The labelling reaction was performed as detailed in the manufacturer’s instructions, using 5 - 10 µCi [α-35S] dATP (supplied at a concentration of 1000 µCi/mmol). The extension and termination reactions were followed, again as specified in the instructions.

The sequencing reactions were run on 0.4 mm thick 6% polyacrylamide-urea gels prepared using Sequagel™ reagents. A Hybaid sequencing apparatus (20 x 45 cm) was used to prepare the gel, with one plate coated with silane. The Sequagel™ components were mixed and polymerisation started by addition of
ammonium persulphate and TEMED, as described in the manufacturer's instructions. The gel was poured and allowed to polymerise overnight, with a comb in position.

Gels were pre-run at 38 mA for 15 - 30 min in TBE (90 mM Tris-borate, 2mM EDTA) buffer. The samples were heated to 80°C for 5 min and loaded immediately. The gel was run at 38 mA. After running, the gel was transferred to Whatman 3MM paper, dried at 80°C under vacuum and exposed to X-ray film for 1-7 days. The film was developed in an Amersham Hyperprocessor.

2.4.13 Transformation of E. coli

2.4.13.1 CaCl₂ method

100 ml of LB medium was inoculated with 1 ml from an overnight culture and incubated at 37°C until the $A_{600}$ reached about 0.4 (after ~3 h). The cells were then aseptically transferred to sterile, chilled 50 ml Falcon tubes and kept on ice for 10 min. The cells were recovered by spinning at 3,300 x g for 10 min in a benchtop centrifuge and each pellet was then resuspended in 25 ml ice-cold CaCl₂ solution (50 mM CaCl₂, 10 mM Tris, pH 8) and stored on ice for 15 min. The cells were spun as before and the fluid decanted. Each pellet was resuspended in 3.5 ml ice-cold CaCl₂ solution. 200 μl aliquots of cells were transferred to 1.5 ml Eppendorf tubes and incubated overnight at 4°C. The following day, the transforming DNA was added (up to 50 ng in a volume of 10 μl or less). The contents were mixed gently and the tubes stored on ice for 30 min. The tubes were then heat-shocked in a 42°C water bath for exactly 90 s and then immediately put on ice for 2 min. 800 μl of SOC medium (2% (w/v) bacto-tryptone, 0.5% (w/v) bacto-yeast extract, 0.05% (w/v) NaCl, 20 mM glucose, 2.5 mM KCl, 10 mM MgCl₂, pH 7.0) was added to each tube and the cultures
incubated at 37°C for 1 h to allow the bacteria to recover and express ampicillin resistance. 200 μl of transformed cells were plated out on LB agar, containing 100 μg ampicillin/ml. The plates were incubated overnight at 37°C and transformed colonies selected. (Modified from Sambrook et al., 1989)

2.4.13.2 RbCl / CaCl₂ method

20 ml of LB medium was inoculated with 200 μl of an overnight culture and grown at 37°C until the A₆₀₀ reached 0.3 - 0.5 (approximately 2-3 h). The cells were then harvested by centrifugation at 3,300 x g for 10 min and resuspended in 1 ml ice-cold solution A (10 mM MOPS, 10 mM RbCl, pH 7.0). The volume was brought up to 10 ml with solution A and the cells spun as before. The supernatant was decanted and the cells resuspended in 1 ml ice-cold solution B (100 mM MOPS, 50 mM CaCl₂, 10 mM RbCl, pH 6.5). The volume was brought up to 10 ml with solution B and the cells incubated on ice for 30 min. The cells were spun again and gently resuspended in 2 ml solution B. For long term storage, glycerol was added to a final volume of 10% (v/v), the cells snap-frozen on dry ice in 200 μl aliquots and stored at -70°C.

3 μl DMSO was added to each 200 μl aliquot of competent cells and DNA added (up to 200 ng in 15 μl). The cells were incubated on ice for 30 min and then heat-shocked at 42°C for 90 s. The cells were put back on ice for 2 min and 1 ml LB medium added to each tube. Cells were allowed to recover for 1 h before being plated out on selectable media.

2.4.13.3 TSS method

100 ml LB media was inoculated with 1 ml overnight culture and incubated at 37°C until the A₆₀₀ reached about 0.4 (after ~3 h). Cells were then transferred to sterile 50 ml Falcon tubes, harvested at 3,300 x g in a benchtop centrifuge and
the media decanted off. The cells were resuspended in one tenth their original volume (5 ml per tube) ice-cold TSS (Transformation and storage solution: LB medium with 10% (w/v) PEG 8000, 5% (v/v) DMSO, 50 mM MgCl₂, pH 6.5). 100 µl aliquots of cells were pipetted into cold 1.5 ml Eppendorf tubes and transforming DNA (up to 200 ng in 10 µl) added. The tubes were mixed gently and incubated at 4°C for 30 min. After this time, 900 µl LB + 20 mM glucose was added and the cells incubated at 37°C for 1 h to allow the cells to express antibiotic resistance. Transformants were then plated out on selectable media. (Chung, Niemela & Miller, 1989).

2.4.14 Transformation of H. volcanii

Transformations were performed using the PEG method described by Charlesbois et al. (1987). Freshly inoculated cultures were grown until late log (A₆₀₀ of 0.8-1.0) and then harvested by centrifugation at 3,300 x g for 15 min, at room temperature, in a Sorvall SS34 rotor. The cells were then washed in 1/10 volume of buffered spheroplasting solution (1 M NaCl, 27 mM KCl, 50 mM Tris-HCl (pH 8.2), 15% (w/v) sucrose) and resuspended in buffered spheroplasting solution containing 15% (w/v) glycerol. Spheroplasts formed on addition of 45 mM EDTA (pH 8.0). Spheroplasts were added to tubes containing 1-2 µg DNA and an equal volume of 60% (v/v) purified PEG₆₀₀ added. The cells were incubated for 20-30 min at room temperature after which they were allowed to recover in 18% (w/v) salt water modified growth medium for a few hours. Selection of transformants was by plating onto 18% (w/v) salt water MGM agar containing 0.3 µg novobiocin/ml or 4 µg mevinolin/ml and incubating at 37°C for 5-10 days.
2.4.15 Purification of polyethylene glycol (PEG) 600

PEG₆₀₀ was purified by the method of Klebe et al. (1983), as outlined by Cline et al. (1989). 100 g of PEG₆₀₀ was weighed into a 1 or 2 l vacuum flask. 200 ml of benzene was added to dissolve the PEG and mixed well. 200 ml of iso-octane was added, mixed well and the mixture put on ice for at least 1 h. The liquid separated into two phases. On ice, the lower phase became like a gel, and the top phase (solvent) was poured off and discarded. The flask was taken off the ice and 100 ml benzene added to dissolve the PEG layer. 5 volumes of diethyl ether were added and mixed overnight using a magnetic stirrer at 4°C to precipitate the PEG completely. The precipitate was filtered using a chilled Buchner funnel on ice and washed with 500 ml chilled diethyl ether. The PEG was scraped off and the residual ether removed in a vacuum desiccator for several hours. The final fluid was stored, tightly capped, at 4°C.
3. Expression in *E. coli*

3.1 Introduction

In order to obtain a high level of expression of halophilic DHLipDH it was thought desirable to express the recombinant protein in *E. coli*. Using an *E. coli* system has a number of potential advantages over expression in the native organism. Firstly, the rate of growth of *E. coli* is significantly faster than that of *H. volcanii* even when the latter organism is grown optimally. Cultures of *E. coli* are typically incubated overnight whereas *H. volcanii* require a week to ten days to reach an equivalent cell density. Secondly, genetic manipulation techniques are at an advanced stage for *E. coli* compared to the halophilic Archaea, which have only recently seen the development of some of the methods that would be required, such as transformation. Thirdly, differences in the properties of an expressed halophilic protein and those from *E. coli* could be exploited in a purification procedure. Addition of a high-salt solution to an *E. coli* extract would be expected to precipitate many of the *E. coli* proteins, leaving the expressed halophilic enzyme in a near-native environment. In addition, the relative thermostability of halophilic DHLipDH could be utilised.

3.2 Amplification of gene from pNAT82 by PCR

The *H. volcanii* DHLipDH gene had previously been cloned into the pBluescript KS+ plasmid as part of a 4.3 kb genomic *Mbol* restriction fragment (Vettakkorumakankav & Stevenson, 1992). The resulting construct was designated pNAT82 (Figure 3.1).
Figure 3.1: Partial restriction map of pNAT82 (Vettakkorumakankav & Stevenson, 1992). 1.6 kb of this clone has been sequenced and accounts for the DHLipDH-coding region.

The DHLipDH-coding region accounts for approximately 1.6 kb of this fragment, shown in the figure. About 40 bp upstream from the transcription initiation site is located a putative Box A archaeal promoter region.

E. coli TG1 cells were transformed with pNAT82 and a miniprep performed to produce template DNA for the PCR reaction. A series of restriction digests were carried out, which confirmed the identity of the plasmid (Figure 3.2). Digestion of pNAT82 with PstI and XbaI cut out the genomic insert, giving bands of 4.3 kb and 3.0 kb, corresponding to insert and vector respectively. Smal digestion produced bands of 2.8 kb and 4.5 kb, in agreement with the expected pattern. Sall digestion linearised the plasmid, giving a band of 7.3 kb. Digestion with Sall produced 1.0 kb and 3.6 kb fragments, as expected from the partial restriction map, as well as other bands indicative of Sall sites within the 2.7 kb of the genomic fragment that is so far unsequenced. Taken together, these results confirm the identity of the plasmid.
Figure 3.2: Restriction digestion of pNAT82. Key: (1) PstI + XbaI, (2) SalI, (3) SmaI, (4) Scal, (P) PstIλ markers, (H) HindIIIλ markers
In order to ligate the gene into an *E. coli* expression vector, new restriction sites needed to be introduced at either end of the coding region. Two PCR primers were designed to bind to the regions shown in Figure 3.3. Primer KAJ-2 contained a *Hind*III site at the 3' end of the gene, after both the stop codon and a possible stem-loop structure. Primer KAJ-3 contained an *Eco*RI site and was designed to bind next to the start codon, after the putative Box A region.

![Diagram showing binding positions of PCR primers KAJ-2 and KAJ-3 on pNAT82. Mismatches are represented by asterisks.](image)

Figure 3.3: Binding positions of PCR primers KAJ-2 and KAJ-3 on pNAT82. Mismatches are represented by asterisks.
One would have expected little trouble in amplifying from a plasmid using nearly consensus primers; however, this was not the case. Initially, a melting step of 5 min at 94°C was used, followed by 30 cycles of melting at 94°C, annealing at temperatures ranging from 60°C down to 37°C, and extension at 72°C. Concentrations of template, primers, nucleotides and Mg²⁺ were varied but none yielded any product. An attempt was then made using an initial melting step of 96°C for 5 min, followed by introduction of Taq DNA polymerase, 3 cycles melting at 95°C and then melting at 94°C for the remaining cycles. This cautious approach to raising the melting temperature was in deference to the limited thermostability of the enzyme. This time, a product of about 800 bp was obtained, while the expected product should have been about 1.5 kb.

It seemed that melting temperature was the critical factor, so a temperature of 96°C was used for all cycles and a product of 1.5 kb was obtained (Figure 3.4).

The PCR program finally used was as follows: 96°C for 5 min followed by addition of enzyme, then 30 cycles of 96°C for 75 s, 60°C for 90 s, 72°C for 2 min. A final extension step of 72°C for 10 min was included. Table 3-1 summarises the reaction conditions tried.
Figure 3.4: PCR amplification of the DHLipDH gene from pNAT82. Key: (M) 
PstI, (1) Primers KAJ-2 + KAJ-3, (2) KAJ-3 + universal M13-40 primer. The 
reaction using the universal primer was performed to test whether the 
initial failure was due to poor binding of KAJ-2.
<table>
<thead>
<tr>
<th>Template conc. (ng/ml)</th>
<th>10-1000</th>
<th>10-1000</th>
<th>1000-5000</th>
<th>200</th>
<th>1000</th>
<th>1000</th>
<th>1000</th>
<th>1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primers</td>
<td>2+3</td>
<td>2+3</td>
<td>2+3</td>
<td>2+3</td>
<td>2+3</td>
<td>3+M</td>
<td>3+M</td>
<td>2+3</td>
</tr>
<tr>
<td>Primer conc. (µM)</td>
<td>1</td>
<td>1</td>
<td>0.5</td>
<td>0.1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Initial 5 min</td>
<td>94</td>
<td>94</td>
<td>94</td>
<td>94</td>
<td>96</td>
<td>96</td>
<td>96</td>
<td>96</td>
</tr>
<tr>
<td>melting temp (°C)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Melting temp (°C)</td>
<td>94</td>
<td>94</td>
<td>94</td>
<td>94</td>
<td>95(3)</td>
<td>94(3)</td>
<td>94</td>
<td>96</td>
</tr>
<tr>
<td>Annealing temp (°C)</td>
<td>55</td>
<td>40</td>
<td>65</td>
<td>37</td>
<td>37</td>
<td>37</td>
<td>60</td>
<td>60</td>
</tr>
<tr>
<td>Extension temp (°C)</td>
<td>72</td>
<td>72</td>
<td>72</td>
<td>72</td>
<td>72</td>
<td>72</td>
<td>72</td>
<td>72</td>
</tr>
<tr>
<td>Product size (kb)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>~0.8</td>
<td>~1.6</td>
<td>~1.6</td>
<td>1.5</td>
</tr>
</tbody>
</table>

Table 3-1: Summary of PCR reactions to amplify DHLipDH from pNAT82. Template concentrations are approximate, based on expected yield from a miniprep. Primer M = M13-40 universal primer.

3.3 Subcloning into the pMEX8 expression vector

The PCR-amplified gene was cut with EcoRI and HindIII, gel-purified using Sephadex G-50 recovery, and ligated into similarly-cut and gel-purified pMEX8 (Figure 3.5). This vector contains a tac promoter for high-level inducible expression, a multiple cloning site and a ribosomal termination sequence. Transformants are selectable by growth on ampicillin-containing media.

The gene was successfully ligated into the vector, as determined by restriction digestion (Figure 3.6), and *E. coli* TG1 cells were transformed using the CaCl₂ method. Digestion with EcoRI and Sall individually linearised the construct, the latter cutting within the insert, producing a fragment of 5.1 kb. The insert could be cut out using EcoRI and HindIII, producing fragments of 3.6 kb and 1.5 kb.
Figure 3.5: The pMEX8 expression vector (Medac)

Figure 3.6: Restriction digestion of DHLipDH cloned into pMEX8. Key: (M) PstI markers, (1) EcoRI linearises within the vector, (2) EcoRI + HindIII cuts out the insert, (3) SalI linearises within the insert.
To check expression of halophilic DHLipDH, 200 ml of transformed cells were grown overnight in media containing 100 μg ampicillin/ml and 1 mM IPTG. They were harvested and sonicated in TE8 containing 2 M KCl, as previously described. After removal of the cell debris by centrifugation, the supernatant was assayed for DHLipDH activity. No activity was detected above that of *E. coli* transformed with pMEX8 alone, used as a control. SDS-PAGE analysis of both soluble and insoluble fractions showed no noticeable difference between the extracts.

### 3.3.1 Expression in JRG1342 (DHLipDH - minus)

Expression in TG1 cells is problematic due to the presence of the host cell DHLipDH, which displays a significant degree of halotolerance (Figure 3.7). This leads to a high background activity that may hide expression if it is present but at a low level.

![Figure 3.7: Halotolerance of *E. coli* DHLipDH. The enzyme activity in whole-cell extracts was assayed at various salt concentrations.](image)
To circumvent this problem, *E. coli* JRG1342 cells were obtained (Prof. J.R. Guest, University of Sheffield). These cells are an ace-lpd deletion strain of *E. coli* K12 that lack a functional DHLipDH and grow very slowly relative to wild type. In order to determine an optimum time for transformation, a growth profile was determined (Figure 3.8).

![Figure 3.8: Growth profile of *E. coli* JRG1342. The $A_{550}$ of a shaking culture at 37°C was measured every hour.](image)

From inoculation, the cells reached an $A_{550}$ of 0.4 after approximately 8 hours. This was the cell density used for transformation.

JRG1342 cells were successfully transformed with the construct, using the TSS method, and grown overnight in LB medium containing 100 μg ampicillin/ml. The following morning, IPTG was added to a final concentration of 1 mM and the cells incubated for a further 4 h at 37°C. The cells were harvested and sonicated, as before, but no enzyme activity was detected. No activity was
observed in JRG1342 cells transformed with pMEX8 alone (lacking an insert). SDS-PAGE analysis showed no difference between extracts.

3.3.2 Growth of JRG1342 in elevated salt concentrations

As the enzyme is folded under relatively high salt conditions in *H. volcanii*, it was thought that the internal environment of *E. coli* might be unsuitable for the protein to take up its correct conformation, and it might be susceptible to proteolytic digestion (or it would have been seen on SDS-PAGE gels). By growing *E. coli* in a medium containing a higher than normal level of salt, it was thought possible that the internal concentration might be increased and the enzyme would be more likely to fold correctly.

Transformed JRG1342 cells were grown in LB medium containing 0.25 - 1.0 M NaCl + 50 mM KCl. No growth occurred in 1 M salt, but slow growth was found in 0.75 M (A550 of 0.3 after 36 h). The cells were treated as before, but still no activity was detectable. SDS-PAGE, however, showed a large band of an apparent molecular weight of 59 kD, the predicted size for DHLipDH, in the soluble fraction of cells grown in 0.75 M NaCl (Figure 3.9).

This band was assumed to be halophilic DHLipDH even though no activity was detected. This experiment was initially designed to test whether the cells would grow in such conditions, rather than to check expression, so no untransformed cell controls were included. Later, when controls grown in the same elevated salt concentrations were tested, the same band was found in them and so the earlier assumption of expression seemed falsely optimistic. The band observed was most likely an *E. coli* protein that happened to be soluble in the high salt conditions of the assay.
Figure 3.9: SDS-PAGE of extracts of transformed JRG1342 cells grown in a range of NaCl concentrations. After cell lysis and centrifugation, electrophoresis was performed on both the supernatant (soluble fraction) and the pelleted debris (insoluble fraction). Key: (M) Low molecular weight markers, (1) 0.25 M soluble fraction, (2) 0.5 M soluble, (3) 0.75 M soluble, (4) 0.25 M insoluble fraction, (5) 0.5 M insoluble, (6) 0.75 M insoluble.
As the gene had been amplified using Taq DNA polymerase, which is known to introduce mutations, on average, every $10^4$ bases (Tindall & Kunkel, 1988), the possibility was considered that the construct may code for a mutated, and therefore possibly inactive, protein. For this reason, the gene was reamplified using Vent® DNA polymerase (exo+) which has a proof-reading ability and therefore higher fidelity than Taq DNA polymerase. The PCR conditions used were the same as before. The gene was successfully amplified and ligated into pMEX8, again determined by restriction digestion and also sequencing across the 5' terminal join, but still no expression could be detected in cells transformed with the construct.

### 3.3.3 Inclusion of FAD in extracts to facilitate correct folding

DHLipDH uses FAD as a cofactor, and therefore cell extracts were prepared, and the enzyme assayed, in the presence of 0.1 mM FAD, as a lack of this cofactor may prevent the enzyme from taking up its correct conformation. No difference, however, was seen between extracts containing the cofactor and those not.

### 3.3.4 Growth of cells at 30°C

Cells were also grown at 30°C, as a lower temperature can influence solubility of an expressed protein. The cells were grown for 5 days and IPTG added to a final concentration of 0.5 mM 18 h prior to assay. Again, no activity was detected, but as previously there appeared to be no expression rather than expression of an insoluble product this was not too surprising.

### 3.4 Expression using the pGEX-KG fusion protein vector

As an alternative expression system, the pGEX-KG fusion protein system was tried (Figure 3.10). This vector produces a fusion of the protein under study with
glutathione S-transferase (GST) (Smith & Johnson, 1988; Guan & Dixon, 1991). The fusion protein can be purified from a cell extract by affinity binding to glutathione agarose and elution with glutathione, after which the two proteins of the fusion can be separated by cleavage with thrombin. The gene to be expressed can be cloned in to the vector either contiguously with the coding region for GST, or utilising a glycine linker region between the two proteins. Use of the linker facilitates more effective cleavage but has the disadvantage of modifying the N-terminus of the expressed protein.

![Diagram of pGEX-KG fusion expression vector](image)

**Figure 3.10:** The pGEX-KG fusion expression vector (Pharmacia). The gene can be cloned in next to the coding region for GST, either utilising the linker (EcoRI site) or not (BamHI site).
3.4.1 Subcloning of DHLipDH gene into pGEX-KG

Primers were designed for amplifying DHLipDH with restriction sites for cloning into the vector, both with and without the linker. KAJ-2, used previously, was the reverse primer, binding to the 3' end of the gene and introducing a HindIII site. KAJ-4 and KAJ-5 are forward primers that introduce a BamHI and EcoRI site respectively.

KAJ-4: 5' ggtgttgatccatgtcgtcggagacatgc 3' 32-mer
KAJ-5: 5' ggtggaattctaatgtcgtcggagacatgc 3' 32-mer

The DHLipDH gene was amplified from pNAT82 using Vent® DNA polymerase (exo⁺) as described in the Materials and Methods (Figure 3.12). Mg²⁺ concentrations in the range 2 mM - 6 mM were used to optimise the reaction, and the program used was as follows: 96°C for 5 min, followed by 30 cycles of 96°C for 75 s, 60°C for 60 s and 72°C for 90 s.
Figure 3.12: PCR amplification of DHLipDH from pNAT82 using primers KAJ-2 and KAJ-4. Key: (M) PstI markers, (1) 2 mM Mg$^{2+}$, (2) 4 mM Mg$^{2+}$, (3) 6 mM Mg$^{2+}$. Amplification with KAJ-2 and KAJ-5 produced products of the same size.
The products from the reactions containing 2 mM Mg\(^{2+}\) were cut with the appropriate restriction enzymes, gel-purified and ligated into similarly cut and purified pGEX-KG vector to create pBAP5000 (with linker) and pBAP5001 (without linker).

### 3.4.2 Expression of a GST-DHLipDH fusion protein

*E. coli* DH5α cells transformed with either pBAP5000 or pBAP5001, when induced with 0.2 mM IPTG for three hours, produced a soluble fusion product of approximately 86 kD, determined by SDS-PAGE. Control cells transformed with the pGEX-KG vector without insert produced a soluble product of approximately 26 kD, corresponding to the GST part of the fusion. Addition of glutathione agarose removed some of the fusion protein from the solution but a large fraction remained soluble, as shown by SDS-PAGE analysis (Figure 3.13). Upon centrifugation, removal of the supernatant and elution of the product from the glutathione agarose with 10 mM free glutathione, a faint band (relative to the starting material) was visible at 86 kD, which is the size of the fusion protein. Addition of 4 μg thrombin to cleave the fusion left only very faint bands at approximately 58 kD (not visible in figure) and 26 kD.

No DHLipDH enzyme activity was detected when assayed in 2 M KCl. Incubation in 3 M KCl for 3-4 hours at 4°C in an attempt to induce a native conformation did not result in activity. No difference was seen between extracts of cells transformed with pBAP5000 and pBAP5001.

As the level of binding of the fusion protein to glutathione agarose was low, columns were used in place of the batch procedure. However, no increased binding was observed.
Figure 3.13: Expression of GST-DHLipDH fusion protein (with linker). Key: (M) low molecular weight markers, (1-2) control and fusion soluble fractions, (3-4) supernatants after glutathione agarose treatment, (5-6) elution from agarose, (7-8) thrombin cleaved, (9-10) after glutathione agarose to remove free GST. The double bands in lane 7 are due to partial removal of the linker region from GST.
3.4.3 **Thrombin cleavage of GST-DHLipDH**

An attempt was also made to thrombin-cleave the fusion in the whole cell extract. This was done in the hope that any active protein released could be purified in ways other than affinity binding, possibly utilising the relative thermostability of halophilic DHLipDH. The cell-free extracts were incubated overnight with thrombin at room temperature. SDS-PAGE analysis of the extracts showed that the fusion was cleaved, releasing a 26 kD fragment (GST) but the DHLipDH was further cleaved into approximately 32 kD and 34 kD fragments (Figure 3.14).

A faint band at about 58 kD was present, which is the expected size for free DHLipDH, but only a relatively small increase in activity above control levels (approximately 2.5 times) was detected when assayed in 2 M KCl.

Since the DHLipDH cleaved from the fusion was undergoing further breakdown, a time-course for thrombin cleavage was undertaken. 4 µg thrombin was added to 2 ml extracts and these were incubated at room temperature for between 0 and 5 hours. 200 µl samples were taken at appropriate time points and the reactions were stopped by addition of 0.2 mM PMSF and 5 mM benzamidine. Following overnight incubation in 4 M KCl at 4°C, the samples were assayed for enzyme activity and an SDS-PAGE gel run (Figure 3.15).

No significant increase in enzyme activity was detected in samples during the time-course, although activity in extracts from cells expressing the fusion proteins were approximately double those from the controls (Figure 3.16). From the SDS-PAGE gels, bands at 56 kD, thought to be DHLipDH, appear after 30 minutes, and breakdown products of this appear at apparent molecular weights of approximately 34 kD and 32 kD after about 2 hours.
Figure 3.14: Thrombin cleavage of GST-DHLipDH fusion protein in cell-free extract. Key: (M) low molecular weight markers, (1-3) control, pGEX-DHLipDH (with linker), pGEX-DHLipDH (without linker) before cleavage, (4-6) post-cleavage (same order).
Figure 3.15: Thrombin-cleavage time course. Key: (M) low molecular weight markers, (1-3) pGEX control, pGEX-DHLipDH (with linker), pGEX-DHLipDH (without linker) at t=0, (4-6) t=30 min, (7-9) t=1 h, (10-12) t=2 h, (13-15) t=3 h, (16-18) t=4 h, (19-21) t=5 h
Figure 3.16: Time-course showing DHLipDH activity in extracts of cells transformed with pGEX-KG constructs following reaction with thrombin for varying periods.

The band at 56 kD was thought to be DHLipDH in a largely inactive conformation. Attempts were made at thermal renaturation of the protein, utilising the relative thermostability of the enzyme. Samples from the thrombin-cleavage time-course were heated to 65°C for 15 minutes in the presence of 3 M KCl and reassayed (Figure 3.17). Overall activities remain similar, with no significant change in activity observed between samples exposed to thrombin for various times and those not.
3.4.4 Heat stability of the fusion proteins

As one half of the fusion proteins consisted of the thermotolerant halophilic DHLipDH, it was thought possible that the fusions as a whole would exhibit such a property. If the fusion proteins were thermotolerant, a heat step could be used to purify them partially from the cell extracts, even though the problems with cleavage and affinity binding made it unlikely that the system would eventually be used.

100 µl of transformed cell extracts were heated to 65°C, 75°C and 85°C for 15 min. The samples were centrifuged at 13,000 rpm in a microcentrifuge for 10 min and the supernatants analysed by SDS-PAGE (Figure 3.18).

The faint bands at 86 kD following 65°C incubation, compared with those from the unheated samples, suggest that the fusions show very limited thermostability. The same bands are barely present in the extracts heated to 75°C and 85°C.
Figure 3.18: Heat stability of the pGEX-DHLipDH fusion proteins. Key: (M) low molecular weight markers, (1-2) pGEX-DHLipDH (with linker), pGEX-DHLipDH (without linker) not heated, (3-4) 65°C, (5-6) 75°C, (7-8) 85°C
3.5 Subcloning into the pKK223-3 expression vector

Due to the ultimate lack of success with the GST fusion protein expression, a new alternative was tried. pKK223-3 (Brosius & Holy, 1984) is similar to pMEX8 in that it has a tac promoter for inducible expression and selection is by ampicillin resistance (Figure 3.19). The reason this vector was tried, when expression in pMEX8 had failed, was because pKK223-3 was available in the lab and had been used successfully to express the citrate synthase from the thermophilic archaeon *Pyrococcus furiosus* (Muir *et al.*, 1995). The DHLipDH gene was PCR amplified from pNAT82 using oligos KAJ-2 and KAJ-3, the same as for ligation into pMEX8, introducing *Hind*III and *EcoR*I sites at the 3' and 5' end respectively.

![Figure 3.19: The pKK223-3 expression vector (Pharmacia)](image)

The product was successfully ligated into the *EcoR*I and *Hind*III sites of pKK223-3, as determined by restriction digestion, to create pBAP5002. *E. coli* JM105
cells were transformed with the plasmid but there was no apparent expression, as determined by SDS-PAGE of both the soluble and insoluble fractions of the cell extract and by enzyme assay.

Variation of the growth temperature (30°C or 37°C) and of induction times (from inoculation to 3 hours before harvest) and concentrations of IPTG failed to yield any expressed protein.

3.6 Use of compatible solutes

As there was a possibility that the internal environment of *E. coli* was unsuitable for expression of halophilic DHLipDH, attempts were made to increase the osmolarity of the cells' cytoplasm. Previously the growth of JRG1342 in elevated salt concentrations was described, but only slow growth was achieved in 0.75 M NaCl. In an attempt to improve on this, the use of compatible solutes and osmotic stress conditioning of cells was investigated. Successes have been reported of induced halotolerance of *E. coli* by these methods (Jebbar *et al.*, 1992; Peddie *et al.*, 1994).

3.6.1 Elevated salt and inclusion of betaine in growth media

*E. coli* JM105 cells were grown overnight in LB medium at 37°C. 1 ml of this culture was used to inoculate six 100 ml cultures containing 0.17 M, 0.5 M or 1.0 M NaCl, in both the presence and absence of 100 μM betaine. The cultures additionally contained 50 mM KCl, as this may have been taken up by the cells in order to reduce the osmotic pressure. These were incubated as before and samples taken every hour to determine growth (Figure 3.20).
Figure 3.20: Growth profile of *E. coli* under salt-induced osmotic stress, in the presence or absence of 100 μM betaine.

The growth profiles showed no significant enhancement of growth in saline media in the presence of betaine.

### 3.6.2 Osmotic stress conditioning

Reports have suggested that osmotic stress conditioning of cells can lead to heightened halotolerance (Ishida *et al.*, 1994). *E. coli* JM105 cells were pre-incubated in LB medium containing 0.5 M NaCl for 30 min at 37°C. 1 ml of this culture was used to inoculate flasks containing 100 ml LB + 50 mM KCl containing 0.17 M, 0.5 M or 1.0 M NaCl, with or without 100 μM betaine, and incubated as before. The A₆₀₀ of the cultures was measured every 60 min. Again, no significant increase in halotolerance was observed between cultures containing betaine and those not. The rate of growth, however, was greater in the cultures containing 0.5 M and 1.0 M NaCl than in those without the pre-incubation, although the overall growth in 1.0 M NaCl was negligible (Figure 3.21).
Figure 3.21: Growth profile of E. coli under salt-induced osmotic stress, in the presence or absence of 100 μM betaine, following osmotic stress conditioning.

3.7 Discussion

3.7.1 Use of PCR in gene cloning

PCR is a quick and effective method for producing large amounts of DNA for cloning, with the advantage that restriction sites can be easily introduced by including mismatches in the primer sequences. One should be aware, however, of the potential problems resulting from polymerase fidelity.

DNA polymerase fidelity has been extensively studied (reviewed by Bloch, 1991). Taq DNA polymerase has been shown to have a base substitution error rate of approximately $10^{-4}$ errors/bp/duplication, with frameshift errors occurring at a frequency of $2 \times 10^{-5}$ errors/bp/duplication (Tindall & Kunkel, 1988). These values vary depending on a range of factors, including template, nucleotide and Mg$^{2+}$ concentration, reaction buffer and annealing temperature, as well as the assessment method (Bloch, 1991). In recent years, new thermostable DNA
polymerases with improved fidelity have been introduced. Vent® DNA polymerase, for example, has 3' - 5' proof-reading exonuclease activity resulting in levels of fidelity at least three times those of Taq DNA polymerase (Keohavong et al., 1993).

When amplifying the DHLipDH gene from pNAT82, products from more than one reaction were used in subcloning and both Taq and Vent® DNA polymerases were tried.

The problems initially encountered with amplifying the DHLipDH gene were unexpected since the primers were matched to the known sequence and the template was a plasmid. The high melting temperature of 96°C that was found to be required can be explained by the high G+C content of the gene (69%). When using a 96°C hot-start and 3 cycles of melting at 95°C, followed by cycles melting at 94°C, an 800 bp band was seen on the gel. It is not clear what this product was, but as it seems that the template was not fully melting, it is likely to be an artefact due to incomplete replication of the template. Melting at 96°C over all cycles resolved this problem and the gene was successfully amplified.

3.7.2 The suitability of E. coli for expression of halophilic proteins

The lack of success in expressing halophilic DHLipDH in E. coli as an unmodified protein suggests the possibility that the internal environment is incompatible with either its translation or its stability. If the protein was being translated in E. coli but not folding correctly, two possibilities present themselves: either the protein forms an insoluble aggregation, or it is rapidly degraded by intracellular proteases. Since there was no evidence of any expressed protein in the insoluble component of the cell extract when run on SDS-PAGE, one should consider the possibility that the protein was rapidly degraded after translation. If
this was indeed the case, attempts to modify the internal environment, such as raising the osmolarity of the cytoplasm, might have lead to the protein taking up its native conformation upon translation.

It was for this reason that attempts were made to grow \textit{E. coli} in elevated salt concentrations. As was shown, only very slow growth was achieved in 0.75 M NaCl and there was no significant growth in 1 M. Growth in concentrations of salt greater than this would be desirable, but even if it was possible, the sacrifice of growth rate would reduce the advantages of expressing in \textit{E. coli}, as fast growth was seen as one of its major benefits. The use of betaine as a compatible solute to balance the external osmotic pressure failed to enhance growth rates significantly. It is possible that induced osmotolerance is strain specific and use of other compatible solutes may have given improved results. It was felt though that this was an area requiring too much optimisation with only a small chance of ultimate success. For this reason, no further attempts at inducing osmotolerance were made.

One can not exclude the possibility that the protein was not translated in the first place. Great care was exercised, however, when designing the experiments for cloning into pMEX8 and pKK223-3, that the start codon was optimally placed relative to the ribosomal binding site. Both these vectors have been successfully used in our lab for the expression of foreign proteins, so the reasons for the failure in this case are unknown. The possibility exists that a mutation introduced during PCR amplification created a stop codon shortly after the initiation site. I feel this is unlikely to be the cause of the problem though, as the reaction was repeated using both Taq and Vent® DNA polymerase, and unless the mutation occurred early in the sequence a truncated product would be expected to be visible by SDS-PAGE. Short of sequencing the entire gene one
can not completely rule out the possibility. This option was not adopted due to the prohibitive length of time that it would likely have taken.

It must be acknowledged that *E. coli* has been successfully used to express halophilic proteins in the past. Malate dehydrogenase from *Haloarcula marismortui* was expressed in large amounts in a soluble but inactive form. Activation of the enzyme was achieved by increasing the salt concentration of the extract to 3 M NaCl (Cendrin *et al*., 1993). Dihydrofolate reductase from *H. volcanii* was also expressed, this time in an insoluble form. Activation was achieved by dissolving the aggregates in 6 M guanidine hydrochloride followed by dilution into salt solutions (Blecher *et al*., 1993). Both these proteins were expressed in pET11 vectors (Studier *et al*., 1990), in which cloned genes are transcribed from the T7 promoter by T7 polymerase, whose gene resides in the chromosome of the *E. coli* strain BL21 (DE3).

HMG-CoA reductase from *H. volcanii* was expressed using pT7-7, a plasmid also containing the T7 promoter, in *E. coli* BL21 (DE3) (Bischoff & Rodwell, 1996). Activity was detected after extracts were exposed to 3 M KCl.

The disadvantage of these plasmids is that expression can only be done in strains containing the T7 polymerase gene. In order to express in other strains, such as the DHLipDH - minus mutant JRG1342, the cell needs to be infected with a bacteriophage containing the polymerase gene. This may well have been detrimental to the growth of an already slow-growing mutant strain. It was for this reason that a similar plasmid was not tried for the expression of halophilic DHLipDH.

### 3.7.3 GST-DHLipDH fusion proteins

Expression was obtained by use of the pGEX-KG vector to create a GST-DHLipDH fusion protein. By being part of a fusion, the expressed DHLipDH was
probably stabilised and folded into a soluble form in the low salt environment. Unfortunately several problems were encountered. Firstly, there was the lack of binding of the fusion protein to glutathione agarose. Binding of GST fusion proteins diminishes with increasing molecular mass, and 86 kD is at the upper limit of what would be expected to bind (Frangioni & Neel, 1993). The second problem was the internal cleavage of the DHLipDH part of the fusion with thrombin. This was not predicted, although the possibility had not been discounted, due to the fact that the recognition site for thrombin is somewhat vague, namely Arg-Gly peptide bonds in specific sequences only. Since the thrombin-cleavage site between the components of the fusion was specifically designed and engineered as such, it was expected that cleavage would occur there preferentially.

3.7.4 Concluding remarks

The overall lack of success in expressing halophilic DHLipDH in E. coli was disappointing but it was felt by this stage that it was time to abandon further attempts and concentrate on expression in the native organism. Work towards this end had been conducted in parallel with the E. coli expression studies and sufficient progress had been made to suggest that this path would ultimately prove successful.
4. Over-expression in *H. volcanii*

4.1 Introduction

Expression of proteins in the native organism has potential advantages over expression in *E. coli*. While the latter may grow considerably quicker, the significant differences in the osmotic environments of the cytoplasms of the respective species can lead to problems with protein folding and stability.

It was for this reason that work was initiated to express halophilic DHLipDH back in *H. volcanii*. A small number of shuttle vectors have been developed over the last few years that are able to replicate in both *E. coli* and *H. volcanii* (Lam & Doolittle, 1989; Holmes et al., 1991; Holmes et al., 1994). This ability is due to the inclusion of origins of replication for both species. Similarly, the vectors are selectable in each host by display of two types of antibiotic resistance.

4.2 Subcloning DHLipDH gene into the pMDS20 shuttle vector

pMDS20 is a shuttle vector containing the *E. coli* ColE1 plasmid *ori* region and the ampicillin-resistance-conferring β-lactamase gene. It also contains the *Haloferax* pHK2 replicon region and novobiocin-resistance-encoding *gyrB* gene (Holmes et al., 1994) (Figure 4.1). The vector is hence able to be maintained and selected for in both hosts. pMDS20 has a copy number of 7-8 copies per chromosome, unchanged from that of pHK2. It contains a number of unique sites for cloning.
Figure 4.1: The pMDS20 shuttle vector

As with subcloning into the E. coli expression vectors, new restriction sites needed to be introduced at either end of the coding region of the DHLipDH gene in order to ligate it into pMDS20. This was again achieved by PCR. A PCR primer, KAJ-1, was designed to bind to a region upstream and including the 5' end of the putative Box A promoter in pNAT82. This primer had mismatches to introduce a BamHI site. Primer KAJ-2, which was used in the amplifications for cloning into the E. coli expression vectors, introduced a HindIII site at the 3' end of the coding region (Figure 4.2).

KAJ-1 : 5' agggcgtcggatccgaaccggtgaagg 3' 29-mer
KAJ-2 : 5' tctcttaagcttggaactcgatgcgacctg 3' 31-mer
Figure 4.2: Binding positions of PCR primers KAJ-1 and KAJ-2 on pNAT82. Mismatches are represented by asterisks.

The same PCR conditions were used as those optimised in Chapter 3, namely the following program: 96°C for 5 min followed by addition of 2.5 units Taq DNA polymerase, then 30 cycles of 96°C for 75 s, 60°C for 90 s, 72°C for 2 min. A final extension step of 72°C for 10 min was included.
Figure 4.3: PCR amplification of DHLipDH gene from pNAT82. Key: (M) PstI, (1) Primers KAJ-1 + KAJ-2
The PCR product was gel-purified using Sephadex G-50 recovery, cut with BamHI and HindIII and gel-purified again. This was ligated into similarly-cut and gel-purified pMDS20. The resulting construct, pMDS21, was transformed into *E. coli* in order to prepare enough DNA for transformation into *H. volcanii*. A miniprep was done by the alkaline lysis method and the purified plasmid used to transform *H. volcanii* WFD11. [NB: The actual ligation of the purified PCR product and transformation were done at the University of Melbourne, Australia by Dr. M. Dyall-Smith and Dr. M.J. Danson].

Cells transformed with pMDS21 were harvested and assayed for DHLipDH activity. No increase in activity above levels found in control cells transformed with pMDS20 (no insert) was detected. The cell-free extracts, as well as solubilised whole cells, were analysed by SDS-PAGE. No differences were observed between pMDS21-transformed cells and the controls. Additionally, a mutant strain of *H. volcanii*, BAS5005, lacking a functional DHLipDH gene, whose construction is described in Chapter 6, was transformed with pMDS21. No activity was detected.

The lack of expression suggests that the putative Box A region may not represent an active promoter. In order to overexpress the protein, introduction of a strong promoter to the gene was considered desirable.

### 4.3 Introduction of a rRNA promoter

The promoter that was chosen to overexpress the DHLipDH gene was the strong P2 promoter from the rRNA operon of *Halobacterium cutirubrum* (Brown *et al.*, 1989).

The promoter was synthesised as follows. Two overlapping oligos (WGW-1 and 2) were designed to bind to the consensus region of the promoter, as well as to all of the region downstream to 2 bases past the indicated
transcription starting point. The forward oligo had half an EcoRV site, and the reverse contained a BamHI site. The oligos were annealed, extended with T4 DNA polymerase and the resultant reaction run on a 10% polyacrylamide gel. The band was excised and purified, cut with BamHI and initially cloned into the EcoRV and BamHI sites of pUK21 (Vieira & Messing, 1991). The resultant construct was designated pMDS65. The promoter was cut out of pMDS65 with EcoRV and BamHI and annealed to the BamHI end of the DHLipDH PCR product amplified with KAJ-1 and KAJ-2. The resultant construct was ligated into pWL102 (Lam & Doolittle, 1989), at the KpnI and XbaI sites, to create pMDS24 (Figure 4.4).

WGW-1: 5' atcgatgcccttaagtacaacagggtacttcgg 3' 33-mer
WGW-2: 5' gcggatccgttcgcattccaccgaagtaccctgtt 3' 35-mer

[NB: The pMDS24 overexpression plasmid was created by W. Woods and Dr. M. Dyall-Smith (University of Melbourne, Australia) as part of an active collaboration (Jolley et al., 1996)].

pMDS24 was introduced in to H. volcanii WFD11, following passage through E. coli to produce enough DNA for transformation. Mevinolin-resistant transformants were selected and cell extracts produced, as described in the Materials and Methods.

The transformants had a DHLipDH activity of 0.38 ± 0.03 units/mg (95% confidence level) at 30°C, compared to a wild type activity of 0.03 units/mg. This corresponds to an average 13-fold over-expression of enzyme activity.
Figure 4.4: Construction of the pMDS24 shuttle / expression vector
4.4 Change of antibiotic resistance marker

In order to over-express the enzyme in the DHLipDH - minus strain of *H. volcanii*, described in Chapter 6, the antibiotic resistance marker needed to be changed. This was because the mutant strain was constructed by inserting the mevinolin resistance gene into the chromosomal DHLipDH gene by homologous recombination. Consequently, the mutant strain was mevinolin resistant and pMDS24 was not selectable.

Primers were designed to PCR-amplify the novobiocin resistance marker from pMDS20, introducing *SphI* and *KpnI* sites at the ends so it could be exchanged with the mevinolin resistance marker of pMDS24.

Figure 4.5: Binding positions of primers KAJ-6 and KAJ-7 to the novobiocin resistance determinant of pMDS20. Mismatches are represented by asterisks.
KAJ-6: 5' cgcaggtacccagtccgcttaagtactc 3' 30-mer
KAJ-7: 5' ggttcgaggttgcatgcccgaagcgcggtc 3' 30-mer

A PCR reaction was set up containing 10 ng pMDS20 template DNA and 50 pmol primers KAJ-6 and KAJ-7. The buffer was as detailed in the Materials and Methods for amplification with Vent® DNA polymerase. Mg²⁺ concentration was varied over the range of 2 mM to 6 mM. After a hot-start of 96°C for 5 min, Vent® DNA polymerase was added and the following program run over 30 cycles: 96°C for 75 s, 60°C for 60 s and 72°C for 3 min. A final extension step of 10 min at 72°C was included. See Figure 4.6.

The presence of extra bands in the reaction meant that optimisation of the conditions was required. Using the same concentration of reactants, the annealing temperature was raised to 72°C, in effect producing a two-step PCR reaction. 30 cycles of the following program were used: 96°C for 75 s, 72°C for 4 min. The same range of Mg²⁺ concentrations was used. See Figure 4.7.
Figure 4.6: PCR amplification of the novobiocin resistance determinant from pMDS20. Key: (M) PstI markers, (1) 2 mM Mg$^{2+}$, (2) 4 mM Mg$^{2+}$, (3) 6 mM Mg$^{2+}$. 
Figure 4.7: Optimised PCR amplification of the novobiocin resistance determinant from pMDS20. Key: (M) PstI markers, (1) 2 mM Mg$^{2+}$, (2) 4 mM Mg$^{2+}$, (3) 6 mM Mg$^{2+}$. 
The product from lane 3 (6 mM Mg$^{2+}$) was gel-purified, cut with KpnI and SphI and exchanged with the mevinolin resistance marker of pMDS24 to create pBAP5003 (Figure 4.8).

![Diagram of pBAP5003 shuttle vector](image)

Figure 4.8: pBAP5003 shuttle vector for over-expression of DHLipDH in *H. volcanii*

When pBAP5003 was transformed into the DHLipDH - minus strain of *H. volcanii*, BAS5005, the development of which is detailed in Chapter 6, DHLipDH activity was 0.84 ± 0.14 units/mg (95% confidence level) at 30°C. This corresponds to an average 28-fold over-expression of enzyme activity, compared to wild type.

SDS-PAGE analysis of novobiocin-resistant transformed cells shows a clear band at an apparent molecular weight of 66 kD which is absent in the BAS5005 cell extract (Figure 4.9). One of the extracts tested had no activity, and in this the 66 kD band was shown to be absent on SDS-PAGE.
Figure 4.9: SDS-PAGE of whole cell extracts from *H. volcanii* BAS5005 cells transformed with pBAP5003. Key: (M) low molecular weight markers, (1) BAS5005 extract, (2-6) extracts from novobiocin-resistant colonies. No DHLipDH activity was detected in the extract from lane 4.
4.5 Concluding remarks

The lack of success in expressing active halophilic DHLipDH in *E. coli* necessitated the investigation of other systems to attain sufficient quantities of the enzyme for further studies. The only reported cases of heterologous expression of halophilic enzymes, malate dehydrogenase from *Haloarcula marismortui* (Cendrin *et al.*, 1993), dihydrofolate reductase from *H. volcanii* (Blecher *et al.*, 1993) and Hmg-CoA reductase from *H. volcanii* (Bischoff & Rodwell, 1996) produced inactive enzymes which required refolding or salt-activation treatments. Homologous expression of halophilic enzymes, however, allows the translated product to take up its natural conformation in a native environment and ensures that correct post-translational modifications will be carried out.

The level of expression achieved for halophilic DHLipDH is sufficient for convenient purification and structural studies. Additionally, use of the developed vector systems, using the halophilic rRNA promoter, should provide a suitable means of expression for any cloned halophilic gene.
5. Purification and characterisation of recombinant halophilic DHLipDH

5.1 Introduction

DHLipDH has previously been purified from several sources, including the halophiles *Halobacterium halobium* (Danson et al., 1986; Sundquist & Fahey, 1988) and *H. volcanii* (Vettakkorumakankav et al., 1992). The procedures utilised have included heat treatment, hydroxylapatite chromatography, immobilised-metal-ion chromatography, ion exchange and gel filtration. Many of these techniques can utilise buffers containing high salt concentrations, generally preferable for the purification of halophilic enzymes, although not essential in this case as halophilic DHLipDH is not irreversibly denatured in the absence of salt (Danson et al., 1986).

Overexpression of the enzyme, described in Chapter 4, has enabled much greater levels of enzyme to be purified from a relatively small volume of cells, as well as reducing the number of steps required in the purification scheme.

5.2 Development of the purification protocol

5.2.1 Heat treatment

Due to the extreme thermostability of DHLipDH from all sources studied, a heat step was included as the first stage of the purification procedure. Danson et al. (1984) showed that the enzyme from a number of halophilic sources maintained full activity following incubation at 95°C for 15 minutes in the presence of 4 M NaCl. In contrast, many other halophilic enzymes were completely inactivated at 60°C.
Cell extracts from *H. volcanii* transformed with the overexpressing plasmid pBAP5003 (see Chapter 4) were subjected to temperatures ranging from 50°C to 95°C for 10 min. No appreciable decrease in activity was detected following any of these treatments (84% activity was remaining following incubation at 95°C for 10 min). The greatest increase in specific activity was achieved by incubation at 95°C.

To determine the optimum incubation period, cell extracts were heated to 95°C for times up to 10 min. They were then cooled on ice, centrifuged at full speed in a microfuge for 10 min to remove any precipitated protein and assayed for DHLipDH activity. As shown in Figure 5.1, protein was precipitated over the first 4-5 min with a concomitant increase in the specific activity of DHLipDH. Specific activity then remained at a plateau for approximately 4 min before starting to decrease.

![Figure 5.1](image)

**Figure 5.1:** Plot showing relative specific activity of DHLipDH and the proportion of protein remaining following incubation of cell-free extracts at 95°C for varying lengths of time. Values are expressed relative to those at time = 0 min. [Data provided by E. Rapaport (University of Bath)].
Based on these data, a heat step of 95°C for 6 min was chosen, the specific activity increasing approximately 3-fold from this step.

5.2.2 Hydroxylapatite chromatography

Hydroxylapatite chromatography has been used previously in DHLipDH purification protocols (Danson et al., 1986). A 2.2 cm x 20 cm column of hydroxylapatite Bio-Gel HT was prepared as described in the Materials and Methods. Heat-treated sample was applied to the column in equilibration buffer (50 mM phosphate buffer, pH 7.0, 1 M NaCl) and was shown to bind as evidenced by the lack of activity in the wash (100 ml equilibration buffer at 30 ml/h).

The enzyme was eluted early in a phosphate gradient (50 - 500 mM phosphate buffer, pH 7.0, 1 M NaCl). A significant increase (approximately 7-fold) in specific activity was achieved from this step, as can be visualised by SDS-PAGE (Figure 5.2).

5.2.3 Mono-Q ion exchange

Active fractions from the hydroxylapatite step were pooled and dialysed against 20 mM Tris-HCl, pH 7.5 at 4°C to remove the NaCl. As previously mentioned, halophilic DHLipDH while largely inactive in the absence of salt will regain full activity following its reintroduction. The sample was loaded onto a Mono-Q anion exchange column on a Pharmacia FPLC system. Protein was eluted with a linear gradient of 0-2 M NaCl. It was found that most protein eluted in a single major peak with no resultant increase in specific activity. Consequently this step was not used in the final protocol.

[Data provided by E. Rapaport, University of Bath].
Figure 5.2: SDS-PAGE showing level of purification following heat-step and hydroxylapatite chromatography. Key: (M) low-range molecular weight markers, (1) cell-free extract, (2) heat-step, (3) hydroxylapatite.
5.2.4 Gel filtration

Active fractions pooled from the hydroxylapatite step were concentrated to a volume of 1 ml using a Centricon 30 concentrator and loaded onto an FPLC Superdex 200 gel filtration column. The sample was eluted in 50 mM phosphate buffer, pH 7.0, containing 1 M NaCl at 1 ml/min and 1 ml fractions were collected. The specific activity of DHLipDH was not significantly increased from this step and it was consequently not used in the final purification procedure. SDS-PAGE of active fractions clearly shows the presence of a major contaminating band at about 48 kD (Figure 5.3).

5.2.5 Cu²⁺ chelate chromatography

A Pharmacia Hi-Trap chelating column was used as detailed in the Materials and Methods. The active fractions from the hydroxylapatite step were pooled and applied to the Cu²⁺-primed column in 50 mM phosphate buffer (pH 7.0) containing 1 M NaCl. The sample was eluted using a 0-100 mM imidazole gradient. This step resulted in an approximate 2-fold increase in purity, but SDS-PAGE revealed the presence of a small number of contaminating proteins, albeit in low quantities (Figure 5.4).

Following stripping and re-priming with Cu²⁺, the column was re-equilibrated with 50 mM phosphate buffer (pH 7.0) containing 1.23 M (NH₄)₂SO₄. The active fractions eluted from the first run were pooled and diluted 4-fold in equilibration buffer to lower the concentration of imidazole which would otherwise inhibit binding. The sample was loaded and eluted as before, but this time in buffer containing 1.23 M (NH₄)₂SO₄ rather than NaCl. A different elution profile was obtained and a further 15% increase in purity was achieved with only a small loss of total activity.
Figure 5.3: SDS-PAGE of adjacent fractions eluted from gel filtration.

Figure 5.4: SDS-PAGE of adjacent fractions eluted from the Cu²⁺-chelate column. The sample was run in buffer containing 1 M NaCl.
5.2.6 The final protocol

It was originally decided to utilise a heat-step as the first stage of the purification protocol. One of the mutants described in Chapter 7, however, was found to be susceptible to irreversible thermal denaturation and consequently the step was dropped. This was done also for the purification of wild-type enzyme in order that there would be no differences in the methods of purification and hence direct comparisons of properties could be made.

The final purification procedure consisted of hydroxylapatite chromatography, followed by 2 runs of Cu\(^{2+}\)-chelate chromatography in buffers containing NaCl and (NH\(_4\))\(_2\)SO\(_4\) respectively. The product obtained following these steps was deemed to be over 99% pure by SDS-PAGE (Figure 5.5). The levels of purification and yields for each step are outlined in Table 5-1 below.

<table>
<thead>
<tr>
<th>Stage</th>
<th>Protein (mg/ml)</th>
<th>Units</th>
<th>Yield (%)</th>
<th>Protein (mg)</th>
<th>Specific activity (units/mg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell-extract</td>
<td>21.5</td>
<td>488</td>
<td>100</td>
<td>344</td>
<td>1.42</td>
</tr>
<tr>
<td>Hydroxylapatite</td>
<td>0.76</td>
<td>435</td>
<td>89.3</td>
<td>38.0</td>
<td>11.5</td>
</tr>
<tr>
<td>Cu(^{2+})-chelate (NaCl)</td>
<td>5.25</td>
<td>248</td>
<td>50.8</td>
<td>10.5</td>
<td>23.6</td>
</tr>
<tr>
<td>Cu(^{2+})-chelate ((NH(_4))(_2)SO(_4))</td>
<td>1.86</td>
<td>203</td>
<td>41.6</td>
<td>7.44</td>
<td>27.3</td>
</tr>
</tbody>
</table>

Table 5-1: Purification table for recombinant wild-type DHLipDH. Samples were assayed in 10 mM Tris (pH 8), 1 mM EDTA, 2 M KCl at 30°C. 1 unit = 1 \(\mu\)mol NADH/min.
Figure 5.5: SDS-PAGE showing each stage of the purification procedure. Key: (M) low-range molecular weight markers, (1) cell-free extract, (2) hydroxylapatite, (3) Cu$^{2+}$-chelate (NaCl), (4) Cu$^{2+}$-chelate [(NH$_4$)$_2$SO$_4$].
5.3 Characterisation of the purified enzyme

5.3.1 Kinetic properties of the enzyme

The recombinant enzyme was shown to follow Michaelis-Menten kinetics, with a $K_m$ (DHLip) of $27 \pm 15 \ \mu M$, $K_m$ (NAD$^+$) of $156 \pm 43 \ \mu M$ and a $V_{max}$ of $19.1 \pm 1.9$ units/mg (Figures 5.6, 5.7). Samples were assayed in 50 mM potassium phosphate buffer (pH 7) containing 1 M KCl at 30°C. The data were analysed by the direct linear plot of Eisenthal and Cornish-Bowden (1974). These data are similar to those previously published for the native enzyme (Vettakkorumakanakav et al., 1992); $K_m$ (NAD$^+$) of $80 \pm 5 \ \mu M$ and a $V_{max}$ of $11.3 \pm 0.1$ units/mg measured at 25°C in 50 mM potassium phosphate (pH 7.2) containing 1 M NaCl.

![Hanes-Woolf plot](image)

Figure 5.6: Hanes-Woolf plot showing $S/V_{max}^{app}$ against dihydrolipoamide concentrations for recombinant halophilic DHLipDH. The points were calculated from direct linear plots using five concentrations of DHLip at five concentrations of NAD$^+$. 

102
Figure 5.7: Hanes-Woolf plot showing $S/V_{\text{max}}^{\text{app}}$ against $\text{NAD}^+$ concentrations for recombinant halophilic DHLipDH. The points were calculated from direct linear plots using five concentrations of DHLip at five concentrations of $\text{NAD}^+$.

5.3.2 Salt optima

The enzyme was assayed over a range of salt concentrations in 50 mM phosphate buffer (pH 7) at 30°C. The activities in both NaCl (Figure 5.8) and KCl (Figure 5.10) were determined. NaCl was used in order to be able to make direct comparisons between the recombinant enzyme with previously published data for the native enzyme. KCl was used because activity in this salt is more relevant to the physiological conditions.
Figure 5.8: Salt activity profile of recombinant halophilic DHLipDH (NaCl)

The optimum concentration of NaCl at pH 7 was shown to be approximately 1.5 M for the recombinant enzyme. Previously published data for the wild-type enzyme has an optimum concentration of 1 M, under similar conditions, with a significantly greater reduction of activity in the higher concentrations of salt (Vettakkorumakan et al., 1992). In order to determine whether this was a genuine difference between the native and the recombinant enzyme, the NaCl activity profile was repeated using wild-type whole-cell extract (Figure 5.9). The profile obtained was similar to that of the recombinant enzyme.
Figure 5.9: Salt activity profile of wild-type non-recombinant halophilic DHLipDH (NaCl)

Figure 5.10: Salt activity profile of recombinant halophilic DHLipDH (KCI)
The optimum concentration of KCl at pH 7 was found to be approximately 1.5 M, the same as for NaCl. Above this concentration, activity remains near maximum with only a slight loss as KCl levels increase.

5.3.3 pH optimum

The enzyme was assayed over the range of pH 4-10 (Figure 5.11). The buffers used were 50 mM acetate (pH 4-6), 50 mM phosphate (pH 6-8) and 50 mM Tris (pH 8-10), each containing 1 M KCl. The enzyme exhibited slightly different specific activities in each buffer at equivalent pH, with the activity in phosphate 1.39 times that of the activity in acetate; the activity in Tris was 1.02 times the activity in phosphate. Consequently, the profile was normalised to activity in phosphate buffer, chosen because this was the buffer used for pH 7 which is the physiological condition. For this reason, phosphate buffer was used for the other characterisation assays, allowing direct comparison of results across the assays. Optimal activity was found between pH 8.5 and 9, consistent with data for the non-recombinant enzyme (Vettakkorumakankav et al., 1992).

Figure 5.11: Normalised pH profile of recombinant DHLipDH
5.3.4 UV-visible absorption spectrum

A UV-visible absorption spectrum of purified recombinant DHLipDH confirmed the presence of a flavin cofactor (Williams, 1976) (Figure 5.12).

![UV-visible absorption spectrum](image)

Figure 5.12: UV-visible absorption spectra of oxidised recombinant DHLipDH and FAD. The peaks at approximately 375, 460 and 485 nm are characteristic of a flavin cofactor.

5.4 Crystallisation trials

Crystallisation trials of the purified recombinant enzyme were carried out using the hanging drop vapour diffusion procedure (McPherson, 1990) outlined in the Materials and Methods section. Crystal Screen™ (Hampton Research) reagents were used as the precipitating solutions (Table 5-2). Trials were conducted with the enzyme in 50 mM potassium phosphate buffer (pH 7) and either 1 M KCl or 10% (v/v) glycerol. The trials were left at room temperature for a few weeks.
<table>
<thead>
<tr>
<th>Tube</th>
<th>Salt</th>
<th>Buffer</th>
<th>Precipitant</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.02M CaCl&lt;sub&gt;2&lt;/sub&gt;</td>
<td>0.1M Na acetate pH 4.6</td>
<td>30% 2-methyl-2,4-pentanediol</td>
</tr>
<tr>
<td>2</td>
<td>None</td>
<td>None</td>
<td>0.4M K, Na tartrate</td>
</tr>
<tr>
<td>3</td>
<td>None</td>
<td>None</td>
<td>0.4M Ammonium dihydrogen phosphate</td>
</tr>
<tr>
<td>4</td>
<td>None</td>
<td>0.1M Tris-HCl pH 8.5</td>
<td>2.0M Ammonium sulphate</td>
</tr>
<tr>
<td>5</td>
<td>0.2M tri-sodium citrate</td>
<td>0.1M Na Hepes pH 7.5</td>
<td>30% 2-methyl-2,4-pentanediol</td>
</tr>
<tr>
<td>6</td>
<td>0.2M MgCl&lt;sub&gt;2&lt;/sub&gt;</td>
<td>0.1M Tris-HCl pH 8.5</td>
<td>30% PEG 4000</td>
</tr>
<tr>
<td>7</td>
<td>None</td>
<td>0.1M Na cacodylate pH 6.5</td>
<td>1.4M Na acetate</td>
</tr>
<tr>
<td>8</td>
<td>0.2M tri-sodium citrate</td>
<td>0.1M Na cacodylate pH 6.5</td>
<td>30% 2-propanol</td>
</tr>
<tr>
<td>9</td>
<td>0.2M Ammonium acetate</td>
<td>0.1M Na citrate pH 5.6</td>
<td>30% PEG 4000</td>
</tr>
<tr>
<td>10</td>
<td>0.2M Ammonium acetate</td>
<td>0.1M Na acetate pH 4.6</td>
<td>30% PEG 4000</td>
</tr>
<tr>
<td>11</td>
<td>None</td>
<td>0.1M Na citrate pH 5.6</td>
<td>1.0M Ammonium dihydrogen phosphate</td>
</tr>
<tr>
<td>12</td>
<td>0.2M MgCl&lt;sub&gt;2&lt;/sub&gt;</td>
<td>0.1M Na Hepes pH 7.5</td>
<td>30% 2-propanol</td>
</tr>
<tr>
<td>13</td>
<td>0.2M tri-sodium citrate</td>
<td>0.1M Tris-HCl pH 8.5</td>
<td>30% PEG 4000</td>
</tr>
<tr>
<td>14</td>
<td>0.2M CaCl&lt;sub&gt;2&lt;/sub&gt;</td>
<td>0.1M Na Hepes pH 7.5</td>
<td>28% PEG 4000</td>
</tr>
<tr>
<td>15</td>
<td>0.2M Ammonium sulphate</td>
<td>0.1M Na cacodylate pH 6.5</td>
<td>30% PEG 8000</td>
</tr>
<tr>
<td>16</td>
<td>None</td>
<td>0.1M Na Hepes pH 7.5</td>
<td>1.5M Lithium sulphate</td>
</tr>
<tr>
<td>17</td>
<td>0.2M Lithium sulphate</td>
<td>0.1M Tris-HCl, pH 8.5</td>
<td>30% PEG 4000</td>
</tr>
<tr>
<td>18</td>
<td>0.2M Mg acetate</td>
<td>0.1M Na cacodylate pH 6.5</td>
<td>20% PEG 8000</td>
</tr>
<tr>
<td>19</td>
<td>0.2M Ammonium acetate</td>
<td>0.1M Tris-HCl, pH 8.5</td>
<td>30% 2-propanol</td>
</tr>
<tr>
<td>20</td>
<td>0.2M Ammonium sulphate</td>
<td>0.1M Na acetate pH 4.6</td>
<td>25% PEG 4000</td>
</tr>
<tr>
<td>21</td>
<td>0.2M Mg acetate</td>
<td>0.1M Na cacodylate pH 6.5</td>
<td>30% 2-methyl-2,4-pentanediol</td>
</tr>
<tr>
<td>22</td>
<td>0.2M Na acetate</td>
<td>0.1M Tris-HCl pH 8.5</td>
<td>30% PEG 4000</td>
</tr>
<tr>
<td>23</td>
<td>0.2M MgCl&lt;sub&gt;2&lt;/sub&gt;</td>
<td>0.1M Na Hepes pH 7.5</td>
<td>30% PEG 4000</td>
</tr>
<tr>
<td>24</td>
<td>0.2M CaCl&lt;sub&gt;2&lt;/sub&gt;</td>
<td>0.1M Na acetate pH 4.6</td>
<td>20% 2-propanol</td>
</tr>
<tr>
<td>25</td>
<td>None</td>
<td>0.1M Imidazole pH 6.5</td>
<td>1.0M Na acetate</td>
</tr>
<tr>
<td>26</td>
<td>0.2M Ammonium acetate</td>
<td>0.1M Na citrate</td>
<td>30% 2-methyl-2,4-pentanediol</td>
</tr>
<tr>
<td>27</td>
<td>0.2M tri-sodium citrate</td>
<td>0.1M Na Hepes pH 7.5</td>
<td>20% 2-propanol</td>
</tr>
<tr>
<td>28</td>
<td>0.2M Na acetate</td>
<td>0.1M Na cacodylate pH 6.5</td>
<td>30% PEG 8000</td>
</tr>
<tr>
<td>29</td>
<td>None</td>
<td>0.1M Na Hepes pH 7.5</td>
<td>0.8M K, Na tartrate</td>
</tr>
<tr>
<td>30</td>
<td>0.2M Ammonium sulphate</td>
<td>None</td>
<td>30% PEG 8000</td>
</tr>
<tr>
<td>31</td>
<td>0.2M Ammonium sulphate</td>
<td>None</td>
<td>30% PEG 4000</td>
</tr>
<tr>
<td>32</td>
<td>None</td>
<td>None</td>
<td>2.0M Ammonium sulphate</td>
</tr>
<tr>
<td>33</td>
<td>None</td>
<td>None</td>
<td>4.0M Na formate</td>
</tr>
<tr>
<td>34</td>
<td>None</td>
<td>0.1M Na acetate pH 4.6</td>
<td>2.0M Na formate</td>
</tr>
<tr>
<td>35</td>
<td>None</td>
<td>0.1M Na Hepes pH 7.5</td>
<td>0.8M Na phosphates, 0.8M K phosphate</td>
</tr>
<tr>
<td>36</td>
<td>None</td>
<td>0.1M Tris-HCl pH 8.5</td>
<td>8% PEG 8000</td>
</tr>
<tr>
<td>37</td>
<td>None</td>
<td>0.1M Na acetate pH 4.6</td>
<td>8% PEG 4000</td>
</tr>
<tr>
<td>38</td>
<td>None</td>
<td>0.1M Na Hepes pH 7.5</td>
<td>1.4M Na citrate</td>
</tr>
<tr>
<td>39</td>
<td>None</td>
<td>0.1M Na Hepes pH 7.5</td>
<td>2% PEG 400 &amp; 2.0M Ammonium sulphate</td>
</tr>
<tr>
<td>40</td>
<td>None</td>
<td>0.1M Na citrate pH 5.6</td>
<td>20% 2-propanol &amp; 20% PEG 4000</td>
</tr>
<tr>
<td>41</td>
<td>None</td>
<td>0.1M Na Hepes pH 7.5</td>
<td>10% 2-propanol &amp; 20% PEG 4000</td>
</tr>
<tr>
<td>42</td>
<td>0.05M K phosphate monobasic</td>
<td>None</td>
<td>20% PEG 8000</td>
</tr>
<tr>
<td>43</td>
<td>None</td>
<td>None</td>
<td>20% PEG 1500</td>
</tr>
<tr>
<td>44</td>
<td>None</td>
<td>None</td>
<td>0.2M Mg formate</td>
</tr>
<tr>
<td>45</td>
<td>0.2M Zn acetate</td>
<td>0.1M Na cacodylate pH 6.5</td>
<td>18% PEG 8000</td>
</tr>
<tr>
<td>46</td>
<td>0.2M Ca acetate</td>
<td>0.1M Na cacodylate pH 6.5</td>
<td>18% PEG 8000</td>
</tr>
<tr>
<td>47</td>
<td>None</td>
<td>0.1M Na acetate pH 4.6</td>
<td>2.0M Ammonium sulphate</td>
</tr>
<tr>
<td>48</td>
<td>None</td>
<td>0.1M Tris-HCl pH 8.5</td>
<td>2.0M Ammonium dihydrogen phosphate</td>
</tr>
<tr>
<td>49</td>
<td>1.0M Lithium sulphate</td>
<td>None</td>
<td>2% PEG 8000</td>
</tr>
<tr>
<td>50</td>
<td>0.5M Lithium sulphate</td>
<td>None</td>
<td>15% PEG 8000</td>
</tr>
</tbody>
</table>

Table 5-2: The Crystal Screen™ reagents used in the initial trials (Hampton Research)

In the absence of salt many of the trials produced a precipitate, but in a number of these small irregular crystals could be seen. One of the best examples of
these was in trial 45. A larger irregular crystal, along with a smaller more regular crystal, was seen in trial 23, also in the absence of salt. In the presence of salt, trial 16 produced hundreds of micro-crystals, and trial 23 produced a number of larger, more regular, crystals. These can be seen in Figure 5.13. Using the conditions in this trial as a starting point, a future project will hopefully yield crystals suitable for X-ray diffraction. Time constraints prevented further optimisation.

5.5 Concluding remarks

The purification scheme detailed in this chapter provides a suitable means of isolating overexpressed DHLipDH from halophilic sources. The absence of a heat-step in the protocol means that heat-labile mutants can be purified without modification of the procedure. The characterisation of the recombinant DHLipDH demonstrates that its properties are indistinguishable from those of the non-recombinant enzyme.

I gratefully acknowledge the significant contribution made by Elizabeth Rapaport, whose undergraduate final year project was involved with the purification of halophilic DHLipDH (Rapaport, 1995). I am also thankful for the help received from Dr. Rupert Russell (University of Bath) in the setting up of the crystallisation trials.
Figure 5.13: Crystals of halophilic DHLipDH
6. Creation of a DHLipDH - minus strain of \textit{H. volcanii}

6.1 Introduction

In order to express mutant DHLipDH within \textit{H. volcanii} without host background enzyme activity, a strain needed to be developed with the gene inactivated. Creation of the strain was done by Dr. M. Dyall-Smith (University of Melbourne, Australia) as part of a collaboration and sent to Bath for assay and confirmation of gene disruption (Jolley \textit{et al.} 1996).

6.2 Insertional inactivation of the DHLipDH gene

To create a strain with an inactive DHLipDH gene, a plasmid was constructed which carried a mevinolin-resistance marker (Mev\textsuperscript{R}) inserted into the unique \textit{Xhol} site within the coding region (nt 593-598 of the published sequence). This was introduced into \textit{H. volcanii} cells and allowed to recombine at the DHLipDH locus. The plasmid was constructed as detailed in Figure 6.1 and its legend. Mevinolin-resistant transformants were selected. Recombination at the HmgCoA reductase locus was also possible and would lead to mevinolin-resistant colonies, but it would destroy an \textit{MluI} site in the wild-type HmgCoA reductase gene. Presence of an intact wild-type HmgCoA reductase locus was confirmed by digesting chromosomal DNA with \textit{MluI} and probing with a Mev\textsuperscript{R} gene. Transformants lacking DHLipDH activity were designated strain BAS5005.
Figure 6.1: Insertional inactivation of the DHLipDH gene. The 4.3 kb insert from plasmid pNAT82 was excised with HindIII and Xbal and ligated to HindIII-Xbal-cut pUC18 (A). Plasmid pMDS51 (B), containing a mevinolin resistance marker, was cut with Xhol and ligated to the Xhol-cut pUC18 clone, thus inserting pMDS51 into the open reading frame of the DHLipDH gene. The E. coli plasmid component of the inserted pMDS51 was removed by Spel digestion, thus leaving only the mevinolin resistance marker (C). The resulting plasmid was linearised at the HindIII site and introduced into H. volcanii cells (D).
6.3 Confirmation of DHLipDH gene disruption

Correct construction of the mutant strain was confirmed by PCR. PCR primers were designed to bind approximately 200 bp either side of the \textit{XhoI} insertion site within the DHLipDH gene.

KAJ-8 (nt 381-399): 5' tcgagaagctctgtaaggc 3' (19-mer)
KAJ-9 (nt 786-768): 5' atgtcgatgccgagctctt 3' (19-mer)

Genomic DNA from both wild-type and the mutant strain was prepared as detailed in the Materials and Methods and the following PCR program run with primers KAJ-8 and KAJ-9: 96°C for 5 min followed by addition of Vent\textsuperscript{®} DNA polymerase, then 30 cycles of 96°C for 75 s, 37°C for 90 s, 72°C for 5 min. A final extension step of 72°C for 10 min was included.

With wild-type gDNA, a band of 400 bp was obtained consistent with an intact DHLipDH gene. With gDNA from the mutant strain, however, a band of approximately 2.3 kb was obtained consistent with disruption of the DHLipDH gene with the mevinolin resistance marker (Figure 6.2). The 2.3 kb band was of a lower intensity than the 400 bp band due to less efficient amplification of longer sequences.
Figure 6.2: PCR analysis of BAS5005 (DHLipDH - minus strain of \textit{H. volcanii}). Key: (M) 1 kb ladder, (1) wild-type gDNA, (2) BAS5005 gDNA, (3) wild-type gDNA
6.4 Enzyme assays of the DHLipDH - minus strain

The mevinolin-resistant cells were assayed for DHLipDH activity. As a control, the citrate synthase activities were also determined to show that a reduction in the level of DHLipDH activity was not accompanied by a large overall reduction in protein synthesis. Out of four transformants tested, two had completely lost DHLipDH activity (Figure 6.3).

![Graph showing DHLipDH and citrate synthase activities](image)

**Figure 6.3: DHLipDH and citrate synthase activities of four mevinolin-resistant transformants.**

All the transformants tested had levels of citrate synthase activity approximately 80% of those of wild-type. Transformants #3 and #6 had no detectable DHLipDH activity, while #8 had approximately 28% wild-type activity. Interestingly, transformant #12 had a DHLipDH activity of 157% of wild-type levels. The levels of expression seen in transformants #8 and #12 may be the result of modifications in the spacing of control elements due to recombination events upstream of the DHLipDH coding region. Transformant #3 was designated strain BAS5005.
An alternative, although highly unlikely, explanation for the loss of DHLipDH activity in transformants #3 and #6 is the activation of an inhibitor of DHLipDH or the switching on of a dehydrogenase that rapidly uses up any available NAD⁺. If such a scenario was occurring, one would expect a significant reduction in the DHLipDH activity in the cell-free extract of transformant #12 when mixed with the extract of #3 or #6. This was not observed, effectively ruling out such a conclusion.

The DHLipDH - minus strain (BAS5005) was transformed to DHLipDH - plus by introduction of plasmid pBAP5003 (see Chapter 4), which contains the DHLipDH gene under control of a rRNA promoter. Transformation of the strain with pMDS21, which contains the DHLipDH gene with no additional promoter, yielded no enzyme activity. This confirms that the DHLipDH gene is not expressed from the putative Box A promoter, unless additional enhancer elements are required.

6.5 Concluding remarks

Creation of a H. volcanii strain lacking a functional DHLipDH gene enables the homologous expression of mutant DHLipDH proteins, without the masking presence of wild-type activity. A potential problem is that as the DHLipDH gene has been interrupted, but not deleted, the possibility exists for an incoming plasmid containing a DHLipDH gene to recombine at the chromosomal sequence. In practice this was found not to be a problem.

Ideally, it would be preferable to delete the chromosomal DHLipDH gene entirely. The difficulty with achieving this is that the gene is right at one end of the original pNAT82 clone, meaning that there is very little flanking DNA to allow recombination back with the chromosomal locus.
The presence of the Mev<sup>R</sup> marker within the DHLipDH gene has the advantage of maintaining strain selection. This does mean, of course, that to select plasmids within the strain another antibiotic marker must be used. The novobiocin resistance gene (Nov<sup>R</sup>), used in the creation of the overexpressing plasmid pBAP5003, is a suitable candidate.
7. Site-directed mutagenesis of halophilic DHLipDH

7.1 Introduction

With the creation of a plasmid for homologous overexpression and a strain of *H. volcanii* lacking DHLipDH activity, the tools are in place for the production of mutants of the enzyme.

The molecular model of halophilic DHLipDH, described in the Introduction (Dr. R.J.M. Russell, University of Bath, unpublished) was used to identify features that may play a role in conferring halophilicity to the protein. As with other halophilic proteins a large negative surface charge, that would be shielded at elevated salt concentrations, was the predominant characteristic. Since mutagenesis of a significant proportion of these surface residues would be a monumental task, other features that are at variance with the non-halophilic structures were sought.

7.2 A putative K⁺-binding site and design of mutants

Using the program GRID (Molecular Discovery Ltd.), a search was made of the structure for K⁺-binding sites. GRID determines energetically-favourable binding sites for a small probe on a target molecule of known three-dimensional structure (Goodford, 1985). The strongest potential site was found at the dimer interface, and had a predicted binding energy of -88 kcal/mol, compared to -57 kcal/mol in the *P. fluorescens* structure. The site consists of four coordinated glutamate residues, two from each monomer (Figure 7.1). The nearest catalytic cysteine is 17Å away (measured between the α-carbons). Figure 7.2 shows a representation of the molecule, with the position of the putative K⁺-binding site and the catalytic residues.
Figure 7.1: The putative K⁺ binding site, consisting of four coordinated glutamate residues.
Figure 7.2: Molecular model of *H. volcanii* DHLipDH, based on the structure from *Pseudomonas fluorescens*. The catalytic cysteines and histidines are coloured yellow and pale blue respectively. The binding sites of the FAD cofactor adenine are coloured magenta and the putative $K^+$ binding site is coloured red.
One of the glutamate residues, E426, is conserved throughout most of the sequences available, and in the two cases where it is not it is replaced by the similarly-charged aspartate. The other glutamate, E423 however, is only present in the *H. volcanii* structure (Figure 7.3).

<table>
<thead>
<tr>
<th>Organism</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Haloferax volcanii</td>
<td>SGFVLGAQIV GPEASELIAE LAFA1EMGAT</td>
</tr>
<tr>
<td>Pseudomonas fluorescens</td>
<td>TDRVLGVHVI GPEASELVOQ GAIGMEFGTS</td>
</tr>
<tr>
<td>Azotobacter vinelandii</td>
<td>TDRVLGVHVI GPEASELVOQ GAIGMEFGTS</td>
</tr>
<tr>
<td>Bacillus stearothermophilus</td>
<td>DGVIIQAQII GPNASDMIAE LGLA1EAGMT</td>
</tr>
<tr>
<td>Bacillus subtilis</td>
<td>DGLVIGAQIIA GASASDMISE LSLA1EGGM</td>
</tr>
<tr>
<td>Escherichia coli</td>
<td>SHRVIGGAIV GTNGEGHAA AG1LAE1GA</td>
</tr>
<tr>
<td>Garden Pea</td>
<td>TDKILGVHIM APNAGELIHE AAIALQYDAS</td>
</tr>
<tr>
<td>Human</td>
<td>TDRVLGAHIL GPGAGEMVNE AALALEYGAS</td>
</tr>
<tr>
<td>Pig</td>
<td>TDRVLGAHII GPGAGEMINE AALALEYGAS</td>
</tr>
<tr>
<td>Pseudomonas putida 1</td>
<td>NHLILGWQAV GAVASELSTA FQSELEMGAC</td>
</tr>
<tr>
<td>Pseudomonas putida 2</td>
<td>TDRVLGVHVI GPEASELVOQ GAIGMEFGTS</td>
</tr>
<tr>
<td>Pseudomonas putida 3</td>
<td>TDEVLGVHLV GPSVSEMIGE FCVAMEFSA</td>
</tr>
<tr>
<td>Saccharomyces cerevisiae</td>
<td>TERILGAHII GPNAGEMIAE AGLALEYGAS</td>
</tr>
<tr>
<td>Trypanosoma brucei brucei</td>
<td>TDRILGVHIV CSAAGELTAG ALLAMEYGAS</td>
</tr>
</tbody>
</table>

Figure 7.3: Alignment of available DHLipDH sequences, showing the putative K⁺ binding site residues, E423 and E426. Residues appearing in half or more of the sequences are shaded.

It was decided to mutate residue E423 in order to investigate the role of the site in conferring halophilicity and stability to the enzyme. Four mutations were chosen:

- **E423S** - serine is the residue present in the *P. fluorescens* structure on which the model was based.
- **E423Q** - glutamine is similarly-sized to glutamate but without a negative charge.
- **E423D** - aspartate has the same charge as glutamate but is smaller.
- **E423A** - alanine is a small aliphatic residue.
GRID analysis of the energy minimised mutant structures predicted the following binding energies at the K⁺-binding site: E423S: -80 kcal/mol, E423Q: -83 kcal/mol, E423D: -82 kcal/mol, E423A: -80 kcal/mol, compared with -88 kcal/mol for the wild-type.

7.3 Mutagenesis

7.3.1 *The Altered Sites*® *II in vitro mutagenesis system*

The Altered Sites® II mutagenesis system (Promega) uses antibiotic selection as a means to obtain a high frequency of mutants (Figure 7.4). The gene to be mutated is cloned into the pALTER-1 vector, which contains tetracycline and ampicillin resistance genes. The ampicillin resistance gene, though, is inactive due to a frame-shift mutation. An ampicillin resistance repair oligonucleotide, which is annealed to the single-stranded, or denatured double-stranded, template at the same time as a mutagenic oligonucleotide restores ampicillin resistance and enables selection. Following synthesis of the second strand with T4 DNA polymerase and T4 DNA ligase, the plasmid is transformed into *E. coli* cells. The mutagenic oligonucleotide is present in greater quantities than the repair oligonucleotide in the reaction, so providing both anneal with similar affinity, a large proportion of the transformants resistant to ampicillin should contain mutant genes.

Initially, the mutagenesis reaction is transformed into a repair-minus strain of *E. coli* (ES1301 *mutS*) to avoid selection against the mutation. Strain transfer into JM109 cells subsequently ensures proper segregation of mutant and wild-type plasmids.
1. Clone insert into pALTER-1 vector
2. Isolate dsDNA
3. Alkaline denature and anneal mutagenic and ampicillin repair oligonucleotides.
4. Synthesise mutant strand with T4 DNA polymerase and T4 DNA ligase

Figure 7.4: The Altered Sites® II in vitro mutagenesis procedure (Promega)
7.3.2 Cloning of halophilic DHLipDH into pALTER-1

A maxiprep of pALTER-1 was prepared from JM109 cells transformed with the plasmid. This was restricted with BamHI and KpnI. The DHLipDH insert was excised from the overexpressing plasmid pMDS24 with the same enzymes. Following gel purification of both fragments, the DHLipDH insert was ligated into the vector and the resultant construct transformed into JM109 cells. Recombinant colonies were selected by plating on LB plates containing 12.5 μg tetracycline/ml, 0.5 mM IPTG and 40 μg X-Gal/ml. White colonies were picked and plasmids prepared from them. These were linearised with BamHI to check for the presence of the insert (Figure 7.5). The pALTER-1 vector with the DHLipDH insert was designated pBAP5004.

7.3.3 Mutagenesis reactions

Four mutagenic oligonucleotides were designed to change residue E423. In the list below, the mutant codon is in bold.

KAJ-12 (E423Q): 5' tcggaggcctgaggccgac 3' 20-mer
KAJ-13 (E423S): 5' gagttcggaggccgaggggccgacgatt 3' 28-mer
KAJ-14 (E423D): 5' ttcggaggcgtcggggccgac 3' 21-mer
KAJ-15 (E423A): 5' tcggaggccgcgggccgac 3' 20-mer

The mutagenesis reactions were carried out as detailed in the kit manufacturer's instructions, but two modifications were required before the procedure worked.
Figure 7.5: Cloning of DHLipDH into pALTER-1. Key: (M) 1 kb ladder, (1) pALTER-1 uncut, (2) pALTER-1 cut with BamHI, (3) pBAP5004 uncut, (4) pBAP5004 cut with BamHI
The first step of the procedure requires alkali-denaturation of the double-stranded target DNA. The denaturation conditions were changed from 5 min at room temperature to 20 min at 37°C. Prior to changing this, no ampicillin resistant colonies were obtained.

The second problem with the procedure was the level of competence of the _E. coli_ ES1301 _mutS_ strain. Use of the RbCl / CaCl₂ (recommended by Promega), CaCl₂, or TSS transformation methods, outlined in the Materials and Methods, failed to produce competent cells. Another repair-minus strain of _E. coli_, BMH 71-18 _mutS_, was tried and was successfully made competent using the TSS method.

Following the mutagenesis reaction, transformation into BMH 71-18 _mutS_, and strain transfer to JM109 cells, plasmid DNA was prepared and manually sequenced. Only one of the reactions resulted in a mutant gene, KAJ-13 (E423S), while the rest resulted in wild-type. After a number of unsuccessful attempts at producing the remaining mutants, new oligonucleotides were designed based on KAJ-13. The oligonucleotides also contained a silent mutation that would introduce an _ApaI_ restriction site. This enabled mutants to be screened by restriction, rather than by sequencing.

<table>
<thead>
<tr>
<th>Oligonucleotide</th>
<th>sequence</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>KAJ-17 (E423Q)</td>
<td>5' gagttcggagggctgggggcccacgatttgcg 3'</td>
<td>32-mer</td>
</tr>
<tr>
<td>KAJ-18 (E423D)</td>
<td>5' gagttcggagggctgggggcccacgatttgcg 3'</td>
<td>32-mer</td>
</tr>
<tr>
<td>KAJ-19 (E423A)</td>
<td>5' gagttcggagggcgggggcccacgatttgcg 3'</td>
<td>32-mer</td>
</tr>
</tbody>
</table>

Using these new oligonucleotides, the remaining mutants were obtained. These were confirmed by restriction screening and manual sequencing (Figure 7.6).
Figure 7.6: Sequencing reactions of the wild-type and mutant genes. The mutant codon is in bold.
7.4 Homologous expression of mutants

All the mutant genes were cut out of the mutagenesis vector with BamHI and Kpnl. They were then gel-purified. The wild-type overexpressing plasmid, pBAP5003, was similarly cut, purified and the mutant genes ligated in, using a 2 hour, room temperature incubation. The new constructs were transformed into competent E. coli JM109 cells in order to produce enough DNA for transformation into H. volcanii. The constructs were designated pBAP5005 (E423S), pBAP5006 (E423Q), pBAP5007 (E423D) and pBAP5008 (E423A).

Minipreps from the transformed JM109 cells were restriction-digested to check the size and the presence of the insert. Plasmids were cut with Kpnl, HindIII and EcoRI. In all cases, the Kpnl and EcoRI digestions produced bands of the same size as with pBAP5003, consistent with the successful ligation of the insert into the vector. There was, however, a discrepancy with the HindIII digestion. pBAP5003, the original expression plasmid, produced three bands on a gel when cut with HindIII, the mutant plasmids all produced a doublet (Figure 7.7). From looking at the pBAP5003 plasmid map (Figure 4.8), it appeared that the HindIII site upstream of the promoter was missing. The explanation for this came from studying Lam's original paper (Lam & Doolittle, 1989) which details construction of pWL102, the precursor to pMDS24 and hence to pBAP5003. A BamHI site was present just upstream of the HindIII site near the promoter. This site had been overlooked and omitted from the restriction map. The effect of this was that cutting out the wild-type DHLipDH gene from pBAP5003 with BamHI and Kpnl, resulted in removal of the promoter. Not surprisingly, when transformed into H. volcanii BAS50005 (DHLipDH - minus), none of the mutant constructs resulted in expression of active enzyme.

The revised restriction map of pBAP5003 is shown in Figure 7.8.
Figure 7.7: Restriction digests of pBAP5005 and pBAP5003. Key: (M) 1kb ladder, (1) uncut pBAP5005, (2) uncut pBAP5003, (3-4) Kpnl cut pBAP5005 and pBAP5003, (5-6) HindIII cut pBAP5005, producing a doublet (visible under transillumination but appearing as a single band in the photograph), and pBAP5003, (7-8) EcoRI cut pBAP5005 and pBAP5003. The other mutant constructs produced the same restriction pattern as pBAP5005.
Figure 7.8: Revised restriction map of pBAP5003, showing a previously overlooked BamHI site upstream of the promoter. Note also the presence of a unique NcoI site next to the BamHI site.

In order to get expression of the mutants two possibilities presented themselves. Either the BamHI site upstream of the promoter in pBAP5003 needed to be removed, or the mutated region of the DHLipDH gene could be exchanged with the wild-type part using a unique site within the gene, upstream of the mutated site, and KpnI. No such site was present, however, leaving removal of the BamHI site the only option. Removal of this site also has the advantage that, in the future, other halophilic genes could easily be cloned into the vector for expression.

There is a unique NcoI site overlapping the BamHI site. This was highly advantageous as it facilitated a theoretically-simple means of removing the BamHI site, as outlined in Figure 7.9.
Figure 7.9: Removal of BamHI site from pBAP5003. The BamHI site overlaps with a unique NcoI site. The vector is cut with NcoI and blunt-ended with mung bean nuclease, removing the terminal nucleotide from the BamHI recognition site. The blunt ends are then re-ligated, destroying both the NcoI and BamHI sites. The vector is incubated with NcoI prior to transformation to linearise any plasmid still retaining the site.

In theory, this procedure should have been relatively straightforward. In practice, however, optimisation was required. In the first attempt, mung bean nuclease was used at a concentration of 1 unit/μg DNA and incubated for 30 min at 30°C, as recommended by the supplier. After transformation, plasmids from 12 colonies were checked for the presence of the NcoI site by restriction screening; none had lost it.
The nuclease incubation time was increased to 1 hour, and increased concentrations of Ncol and incubation time were used to ensure complete digestion. This time, out of twelve colonies checked, nine still had the site. Out of the three that had lost the site, however, two of the plasmids were significantly smaller than expected; up to 3 kb of the plasmid having been digested by the nuclease (Figure 7.10). The other preparation, though, was the correct size. The presence of the intact promoter was inferred by cutting with HindIII and observing three bands when run on a gel. This showed that the HindIII site between the promoter and the lost Ncol site was still there, and hence so would be the promoter.

The plasmid cured of the BamHI site was designated pBAP5009. This plasmid was cut with BamHI and KpnI and the mutant genes cloned in. The mutant constructs were then transformed into E. coli JM109 cells and plasmid DNA purified. Apal screening and manual sequencing confirmed the mutations. The new constructs were designated pBAP5010 (E423S), pBAP5011 (E423Q), pBAP5012 (E423D) and pBAP5013 (E423A).

H. volcanii BAS5005 (DHLipDH - minus) was transformed with the mutant constructs. Transformants were selected by growth in top agar containing 4 µg mevinolin/ml (selective for strain) and 0.3 µg novobiocin/ml (selective for plasmid).

Preliminary assays on cell-free extracts of the transformants indicated expression of active DHLipDH.
Figure 7.10: Restriction screening of plasmids following attempted removal of *Bam*HI and *Nco* sites. Plasmids were cut with *Kpnl* and *Nco*, which cuts out the DHLipDH insert in plasmids containing the *Nco* site. This results in a 1.7 kb band as seen in most lanes. The plasmids in lanes 5, 6, and 9 have lost the *Nco* and *Bam*HI sites, but only the plasmid in lane 6 is the right size.
7.5 Purification of the mutant enzymes

The mutant enzymes were purified from 1 litre of culture in the same manner as the wild-type enzyme, as detailed in Chapter 5. Preliminary investigation of the thermostability of the mutants showed that E423D was significantly less stable than wild-type, necessitating removal of the heat-step in the original purification scheme. Figure 7.11 shows SDS-PAGE gels of each stage of the purification.

Table 7-1 shows the yield and specific activities obtained from each step.

<table>
<thead>
<tr>
<th></th>
<th>E423S</th>
<th>E423Q</th>
<th>E423D</th>
<th>E423A</th>
<th>Wild-type</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Protein (mg/ml)</td>
<td>Protein (mg/ml)</td>
<td>Protein (mg/ml)</td>
<td>Protein (mg/ml)</td>
<td>Protein (mg/ml)</td>
</tr>
<tr>
<td>cell-free extract</td>
<td>21.96</td>
<td>24.43</td>
<td>24.64</td>
<td>22.52</td>
<td>21.47</td>
</tr>
<tr>
<td>Hydroxylapatite</td>
<td>3.39</td>
<td>3.90</td>
<td>2.34</td>
<td>1.17</td>
<td>0.76</td>
</tr>
<tr>
<td>Cu$^{2+}$ (NaCl)</td>
<td>5.20</td>
<td>0.59</td>
<td>0.59</td>
<td>4.57</td>
<td>5.26</td>
</tr>
<tr>
<td>Cu$^{2+}$ [(NH₄)₂SO₄]</td>
<td>2.52</td>
<td>3.02</td>
<td>1.94</td>
<td>1.52</td>
<td>1.86</td>
</tr>
</tbody>
</table>

Table 7-1: Purification table for mutant and recombinant wild-type halophilic DHLipDH. Samples were assayed in 10 mM Tris (pH 8), 1 mM EDTA, 2 M KCl at 30°C.
Figure 7.11: SDS-PAGE gels showing the purification of the mutants. Key:
(M) low-range molecular weight markers, (1) cell-free extract, (2) hydroxylapatite, (3) Cu²⁺-chelate (NaCl), (4) Cu²⁺-chelate [(NH₄)₂SO₄]
7.6 Characterisation of mutants

7.6.1 Kinetic parameters

Kinetic parameters were determined in 50 mM potassium phosphate buffer, pH 7.0, containing 1 M KCl. Assays were at 30°C. All measurements were done in triplicate and the direct linear plot of Eisenthal and Cornish-Bowden (1974) was used to determine $K_m$ and $V_{max}$ values, shown in Table 7-2 below.

<table>
<thead>
<tr>
<th></th>
<th>$K_m$ (DHLip)</th>
<th>$K_m$ (NAD$^+$)</th>
<th>$V_{max}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild type</td>
<td>26.9 ± 14.6 μM</td>
<td>156 ± 43 μM</td>
<td>19.1 ± 1.9 U/mg</td>
</tr>
<tr>
<td>E423S</td>
<td>23.0 ± 3.3 μM</td>
<td>211 ± 54 μM</td>
<td>5.8 ± 0.4 U/mg</td>
</tr>
<tr>
<td>E423Q</td>
<td>22.8 ± 5.4 μM</td>
<td>203 ± 73 μM</td>
<td>6.4 ± 0.8 U/mg</td>
</tr>
<tr>
<td>E423D</td>
<td>26.6 ± 5.4 μM</td>
<td>199 ± 28 μM</td>
<td>16.2 ± 0.8 U/mg</td>
</tr>
<tr>
<td>E423A</td>
<td>10.7 ± 4.8 μM</td>
<td>215 ± 47 μM</td>
<td>6.2 ± 0.4 U/mg</td>
</tr>
</tbody>
</table>

Table 7-2: Kinetic parameters of recombinant wild-type and mutant halophilic DHLipDH

The mutants, with the exception of E423D, were significantly less active than wild-type, whereas $K_m$ differences are not significant in most cases. One exception is the E423A mutant, which has an increased affinity for dihydrolipoamide.

7.6.2 Salt optima

Salt profiles of the mutants were determined in 50 mM potassium phosphate buffer, pH 7.0. Assays in both NaCl (Figure 7.12) and KCl (Figure 7.13) were performed. Measurements were made in triplicate.
Figure 7.12: Salt activity profiles for wild-type and mutant halophilic DHLipDH (NaCl). Error bars show the 95% confidence limits.

Figure 7.13: Salt activity profiles for wild-type and mutant halophilic DHLipDH (KCl). Error bars show the 95% confidence limits.
The activity profiles were similar in NaCl and KCl. Interestingly, the profiles show little difference in the activity of mutants E423S/Q/A upon increase of salt concentration. This is a significant loss of the activating effect of salt. The E423D mutant has practically no activity in the absence of salt, but is activated to levels approaching those of wild-type on its addition.

The optimum salt concentration (either salt) for E423D was found to be 1.5 M, the same as the wild-type enzyme.

7.6.3 pH optima

The pH profiles for the mutants were characterised over a pH range of 4 to 10 (Figure 7.14). Three buffer systems were used to attain this range; acetate (pH 4-6), phosphate (pH 6-8) and Tris (pH 8-10), all at 50 mM and containing 1 M KCl. Profiles were normalised to activity in phosphate buffer.

![Figure 7.14: Normalised pH profiles of mutant and recombinant wild-type DHLipDH. Error bars show the 95% confidence limits.](image-url)
The mutants all have a pH optimum of 8.5, unchanged from that of wild-type, which would suggest no difference in the electrostatic potential around the catalytically-important residues.

7.6.4 Thermal stability

Thermal stability studies were done both in the presence and absence of salt. Tubes containing 5 μg of protein in a volume of 100 μl were heated in thin-walled PCR reaction tubes in a heating block, for times up to 1 hour. Heat transfer was found to be equivalent to using glass tubes in a PEG bath. The temperature in a control tube was monitored constantly, and when it reached within 1°C of the required temperature the t=0 tube was removed and the timing started. The time required to reach temperature was under 2 minutes. Following removal from the heating block, the tubes were cooled rapidly on ice.

The tubes were briefly centrifuged before assaying at 30°C in triplicate. Activity values were expressed as a percentage of the t=0 value. Logarithmic plots of the inactivations were used to determine the first-order rate constants of inactivation, which were used to display the data on an Arrhenius plot.

7.6.4.1 Stability in the presence of salt

Heat inactivations were done in 50 mM potassium phosphate buffer, pH 7.0, containing 2 M KCl (Figure 7.15).

The thermal stability properties of the enzyme are comparable in mutants E423S/Q/A to those of wild-type, although if anything they appear to be slightly more stable. Mutant E423D, however, is inactivated at temperatures around 20°C lower than wild-type. This is clearly seen in the Arrhenius plot (Figure 7.16).
Figure 7.15: Logarithmic plots of heat-inactivations of mutant and recombinant wild-type halophilic DHLipDH in the presence of 2 M KCl
Figure 7.16: Arrhenius plot for the thermal inactivation of wild-type and mutant halophilic DHLipDH in the presence of 2 M KCl. For clarity, linear regressions are shown only for wild-type and E423D.

7.6.4.2 Stability in the absence of salt

Heat inactivations were also carried out in 50 mM potassium phosphate buffer, pH 7.0, containing no salt (Figure 7.17).

As was the case in the presence of salt, the thermal stability properties of mutants E423S/Q/A are comparable to those of wild-type, while mutant E423D is inactivated at temperatures just over 20°C lower. Again, it appears that the mutants E423S/Q/A may be slightly more thermostable than the wild-type enzyme. This can be see in the Arrhenius plot in Figure 7.18.

The mutants and the wild-type are less thermostable in the absence of salt than in its presence, typically denaturing at temperatures around 16-18°C lower.

---

141
Figure 7.17: Logarithmic plots of heat-inactivations of mutant and recombinant wild-type halophilic DHLipDH in the absence of salt.
Figure 7.18: Arrhenius plot for the thermal inactivation of wild-type and mutant halophilic DHLipDH in the absence of salt. For clarity, linear regressions are shown only for wild-type and E423D.

7.7 Discussion

7.7.1 Mutagenesis reactions

The initial problems encountered with the mutagenesis reactions highlight the potential problems of oligonucleotide design for G+C rich sequences. The original primers that failed to bind specifically, 20- or 21-mers, were clearly too short even though they only introduced one mismatch. The one oligonucleotide that worked first time, KAJ-13, was a 28-mer, a longer sequence being used as it introduced two mismatches. In order to increase specificity, one should either utilise regions containing a higher than average proportion of A and T, a lesson utilised in designing the primers for upstream sequencing described in the next chapter, or increase the overall length.

The second set of oligonucleotides designed for mutagenesis incorporated a silent Apal restriction site. This enabled rapid screening of
potential mutants by restriction analysis, and saved the need for time-consuming and expensive sequencing. Potential mutants detected by Apal screening were, of course, confirmed by sequencing. Inclusion of such a site was not done originally, as it meant creating another mismatch which would affect binding affinity. Also, the kit manufacturer claimed a consistent mutant success rate of over 90%. If this had been the case in practice, restriction screening would have been unnecessary.

The difficulties encountered in making the ES1301 mutS strain competent have since been recognised by Promega. They now sell ready-competent cells for use with this procedure.

7.7.2 Mutant enzyme properties

Energy minimisation of the mutant structures predicted no gross change in structure, but one should be cautious since the minimisation was carried out on an already-minimised structure, i.e. the residue was changed in the structure file of the wild-type enzyme in its native conformation and was then re-minimised. As the mutant residues were not larger than the wild-type residue and did not introduce new charges, energy minimisation would not be expected to show any large structural changes. The possibility still exists of a conformational change that could affect the active site.

All the mutants are less active than wild-type under optimal conditions, although E423D is only slightly so. The kinetic data show no significant change in the $K_m$ of NAD$^+$, suggesting that NAD$^+$ binding is unaffected. A similar case is found for the $K_m$ of dihydrolipoamide, except in E423A where the affinity is increased slightly. E423A is the least conservative of the mutations and may affect the conformation of the dihydrolipoamide binding region around the catalytic cysteines and histidine. The pH optima of all the mutants are
unchanged from that of wild-type, suggesting that there is no difference in the electrostatic potential around the catalytically-important residues.

At physiological salt concentrations, the activity in mutants E423S/Q/A is approximately 25% that of wild-type. At lower concentrations of salt, however, the wild-type enzyme rapidly loses activity, whereas mutants E423S/Q/A do not, to the point that in the total absence of salt the activities are similar. In the absence of salt, the putative K⁺-binding site would be unoccupied, resulting in the wild-type enzyme having four unshielded opposing glutamate residues at the dimer interface which would have a destabilising effect on the structure. Under the same conditions, E423S/Q/A would only have two unshielded glutamate residues (E426 from each chain) and so the destabilising effect should be lower than in the wild-type enzyme.

E423D is the most conservative mutant, maintaining the negative charge, and the binding site would be expected to maintain a high affinity for K⁺. Perhaps the most significant property of this mutant is its lack of activity in the absence of salt. Interestingly, although the residue is smaller than the glutamate present in the wild-type, the carboxy- oxygens of the two aspartate residues are nearer to each other at their closest point than those of the wild-type glutamates (3.60Å and 4.15Å respectively). The corresponding α-carbons are also closer (8.35Å compared to 8.59Å in the wild-type), although it must be remembered that these measurements are based on a model structure. This may mean that in the absence of salt, the backbone is displaced more in this mutant than in the wild-type, affecting the active site conformation and hence activity. As the salt concentration increases, however, occupancy of the K⁺ binding site could result in stabilisation that results in the backbone near the active site taking up a conformation similar to wild-type. This would explain the activation with
increasing salt concentration, seen in both this mutant and the wild-type. In the wild-type enzyme, there is a hydrogen bond between the backbone oxygen of E423 and the nitrogen of E426, on the same chain. This hydrogen bond is predicted to be missing in mutant E423D in the absence of salt, further evidence of reduced stability around this site, although again caution must be exercised with this hypothesis due to the calculations being derived from a model.

The thermal stability data show that mutants E423S/Q/A are as similarly resistant to heat denaturation as wild-type. It is E423D that denatures at temperatures around 20°C lower, again showing that this mutant is less stable both in the presence and absence of salt.

What could be the role of a K⁺-binding site at the dimer interface? One possibility is that it helps maintain the integrity of the core, especially the spacing of groups around the active site, as the salt concentration increases. It is known that increasing salt concentrations tend to lead to a compacting of protein structure, and such a site may strengthen the core under these conditions.

I would like to thank Dr. Rupert Russell for running the GRID analyses and energy minimisations.
8. The physiological role of halophilic DHLipDH

8.1 Introduction

Although DHLipDH has been found in a number of archaeal species in the apparent absence of 2-oxo acid dehydrogenase multienzyme complexes (Danson et al., 1984; Smith et al., 1987), its physiological role remains undetermined. In order to investigate its possible function in these organisms, two approaches can be followed. Firstly, sequencing upstream of the DHLipDH coding region could determine the identity of nearby genes that might be associated with the enzyme. Secondly, experiments comparing the growth properties of wild-type and BAS5005 (DHLipDH - minus) H. volcanii on different substrates may provide clues to the metabolic defect of the mutant strain. Both these approaches have been undertaken.

8.2 Sequencing of upstream region of pNAT82

pNAT82 is the pBluescript vector containing a 4.3 kb genomic fragment including the DHLipDH coding region at the 3' end (Vettakkorumakankav & Stevenson, 1992). The remaining 2.7 kb upstream of the DHLipDH gene of this insert has not previously been sequenced, but the 76 bases of the published sequence upstream of the DHLipDH start codon appear to be homologous to an E2 component of a 2-oxo acid dehydrogenase complex (Prof. K.J. Stevenson, University of Calgary, Canada, unpublished observation).

A short sequence of approximately 140 bases was obtained from the 5' end, using a universal sequencing primer (Dr. M. Dyall-Smith, University of Melbourne, Australia). Using this sequence, a primer, KAJ-21, was designed to
sequence the forward strand. Another primer, KAJ-20, was designed using the published DHLipDH sequence to start sequencing the reverse strand.

pNAT82 DNA was prepared using a Qiagen Tip-20 column. Individual reactions were sent for automated sequencing on an ABI Prism™ 377 DNA sequencer. Sequence obtained was used to design new primers until both strands of the entire 2.7 kb region had been completed. In all, 14 primers were required and these are listed below.

KAJ-20 : 5' gagcagttccttcacgcggttc 3' 22-mer
KAJ-21 : 5' gagcgagatgttcaaccacgcatac 3' 25-mer
KAJ-24 : 5' atctctgctggcgagtcttgagcatcc 3' 25-mer
KAJ-25 : 5' agattcagttctccggcttcat 3' 22-mer
KAJ-26 : 5' ggttcactcgcgcgcgtcgaccgt 3' 23-mer
KAJ-27 : 5' cgtggacatcgaaccatct 3' 20-mer
KAJ-28 : 5' cgaacgtgccttcggttc 3' 20-mer
KAJ-29 : 5' cgaactcgttaacggcgcacgtc 3' 22-mer
KAJ-30 : 5' agcacctgtttgtctcgtac 3' 22-mer
KAJ-31 : 5' gtcacgaaggcgacgtcg 3' 20-mer
KAJ-32 : 5' ttctcgacggcctccagcgtc 3' 21-mer
KAJ-33 : 5' cgcgacgggaggcgttcgtc 3' 20-mer
KAJ-34 : 5' ccatgtggctcagctttggt 3' 20-mer
KAJ-35 : 5' ctacatgcgccttgatga 3' 20-mer

Sequences were aligned and edited using the GCG package. The full sequence of the pNAT82 insert is shown in Figure 8.1, including the published DHLipDH sequence, with the binding positions of the primers shown in bold.
1  GACGACGAGG GCGTCAGGCGG CATCGAAGAA GACGTGAAAG GGCGGGTCGC
51  GGACGCCATGC AGGCCGCGCC AGTCCGACCC GCAGGCAGAC AGGACGGAGA
101  TGTTCACCA CGCATA GACCAGCACCC CGGAGATTCA AGCAGCTTAC
151  GAGGAGTTCG AGGCGTCCTCC CGAGAAATTC CGCAGAAGAG GATTCCTCAG
201  AGAATGAGCA GTCAAGACTC ACCACACGGG CAGCCGCTAC GGCAGGCTCT
251  CTACACCGAG ATGAACCTCG AGCAGCGAGGT GCTCGTACTG GGCGAAGACG
301  TGGGCAAGAA CGGCGCGGTC TTTCCGGGCTA CGGAGGGGCT CGGGCATCGT
351  TGGGCCAGAG ACCGCGTGAT CAGACGCACG CTGGCCAATG AGGCGATCGT
401  CGGCAGCGCC ATCGGATGCG CGGCGATGGG ACTGAAGCCG GTCCCGAGA
451  TTCAGTTCTC GGGCTTCTAG TACCCCGGAT TCGAGCAAGAT CTGGAGCCAC
501  ATGGAAGCTG ATCCCAAGCC AGCAGCAGCG CGGCCGCTGC GTGCCGCAGG
551  GGTCTGACGC CTGACACGACG CGGTCCGTCG CCCGCCGCGCG GTGGCTCGCC
601  AGCTCAACGA GAGTCTCTAC GGCACGCGAG CCGCCGCTGA GGTGCTACAT
651  CCCCCTGCGC CGGACTGACG GAAGGAGGCT CTGACTCTCG CGATCGCGGA
701  CCGAGCGCTG GGTATCTTCA TGGAGCGGAA GCTCATCTAC GGGCGCTGTC
751  GCAGGAGATC CCGCGAGAAG GACTACACGG TCCCGACGCG GGTGCTCGCC
801  GTGGCCCGCG AGGGCAGCGG GGTGCTCGCG TCCCGGCGCC GTGGCTCGCC
851  GGGCCGAGGC GTGGAGTTGC CGCAGGGCGG TGCGAGAACG GGTATCGAGC
901  CAGGAGGCTG CTGACAGCCG AGGCGCGCCG CACGCAAGGT GGTGCTCGCC
951  GTGGAAGGTC TCCAGGACGC GCGGCAGCGG GTGGCTCGCC ACGAGGCCC
1001  GAAGAACGCG GGCCTCGGAC CCCAACTCAG CGGCCGGTGG CAGAGGACC
1051  CGCTGCTGAT TCAAGGACGC CCGCCGCGCG GTGGCTGCGT ACGAGGCCC
1101  CGGCTCACAC GTGACAGCAG GAGGACACGT TACCTCTCCG CGGACGGCC
1151  CGTGCGAGAA GGTATTCTGC AGCCGGTCTG CTTCCTGGAC ATCAGCCGCTC
1201  AGGAATTCAA ACTCCCCAGC GTGGTGGAAG CCGGGCTGGG AGGGCAAACG
1251  GTAACGTGCC AGCTCGCGCC GGGCGACGAG GTAACCGAAG ACCAGGTGC
1301  CGCGGAGGTG GAGGCCGACA AGGCGTCGCG CGACGTCGCCC TCGCCGCTCG
2751 GAATAATGGT CGTCGGAGAC ATCGCAACCG GAACCGAACT GCTCGTCATC
2801 GCCGGCGGAC CGGGCGGCTA GTGCGCCGCC ATCCGCGCCG CACAGAACGG
2851 CATCGACACG ACGCTGGTTC AGAAGGACGC CTACGGGGGC ACCTGCCTCA
2901 ACTACGGCTG TATCCCATCG AAGGCGCTCA TCACGGGGGC GAACCTCGCC
2951 CACGAGGCGG GCAACGCCGA GGAGATGGGC ATCCACGCCG ACCCCGTCGT
3001 GGACATGTCG CAACTGCGGC ACTGGAAGAG CGGCGTCGTC GACCAACTCA
3051 CCGGCGGCGT CGAGAAGCTC TGTTAAGGCGA ACGGCGTCGG TGACGTCGAG
3101 GGAACCGCCC GTCTAAAGGA CGAGAAGCGC GTCCGGCGGT CCCCCGTGGG
3151 CGAGGCGGAC CGGCGGCGC GTGCGTCGTC AAGGCGCTCA ACATCGGCCG
3201 CCAGGCGGCT CGGCGGCTCT CGTTATCGGT ATGGCGCGGC GTCCGGCGGT
3251 GTCTGGGTAGT GCTGTCGCCA CCTCGAGGTC TGTCCCGTCT GCTGGCGGTG
3301 GGACGGGCGG GGAGATGCCG ACTGCAAGAG CGGCGTCGTC GACCAACTCA
3351 CCGGCGGCGT CGAGAAGCTC TGTTAAGGCGA ACGGCGTCGG TGACGTCGAG
3401 GGAACCGCCC GTCTAAAGGA CGAGAAGCGC GTCCGGCGGT CCCCCGTGGG
3451 CGAGGCGGAC CGGCGGCGC GTGCGTCGTC AAGGCGCTCA ACATCGGCCG
3501 CCAGGCGGCT CGGCGGCTCT CGTTATCGGT ATGGCGCGGC GTCCGGCGGT
3551 GTCTGGGTAGT GCTGTCGCCA CCTCGAGGTC TGTCCCGTCT GCTGGCGGTG
3601 GGACGGGCGG GGAGATGCCG ACTGCAAGAG CGGCGTCGTC GACCAACTCA
3651 TCTCTCTCGG TGACGGCACC CGGCCCGACG ACGTGGAGCA CATCTAGGCC
3701 GTGCGGCGGC TGGTGCGAGA CAGCAGGATG CTCGGCCACG TCGACTCAGA
3751 GGAGGCGATC GTGCGGCCGC AGCACTGCGC CGGGCGGTTC GGGCGCGCTC
3801 ACACGCAGTG GACATCGGA ACGCCGGGCT CGAGAGGCAG CACCGCGCTG
3851 GTGCCGATCG CGAGGCGGCA CGGCCGCCGT GGGCGGTTTC GGGCGGCGTG
3901 CCGACGCGGT CGGCGGGCGT GTGCGGCGAG AGAGGCGGCT CTGCTCTCGT
3951 CAGAGGTGG CACCCTCGGG GTGCGCGCAG CGAGGACGCT CGGCGGGCGT
4001 GGGCCGAAAT CGTGCGGCGC CAGGCGCTTC GACATCGGCG CGAACCGCGC
4051 GTGCCGCTAC GAGATGGCAG CGACGCTCGA AGACGTGGCC TCGACCATCG
4101 ACACCCGCCA GACGCTCGGC GAAGCGGTCA TGAGACCGCG CGAGAAGCGC
4151 CTCGGACAGG CGATTACAGC CCTGACGTCG TGAGAGCGGT CGGAAAACGC
Figure 8.1: The complete nucleotide sequence of pNAT82

Figure 8.2, below, shows the direction and extent of the sequencing reactions.

**Figure 8.2: The sequencing strategy.** The direction and length of the arrows indicate the direction and extent of the individual sequencing reactions. The dotted arrow represents the published sequence of the DHLipDH gene.

8.3 Sequence analysis

8.3.1 Positions of open reading frames

Translation of the upstream sequence showed the presence of three open reading frames. These were found to be homologous to the C-terminus end of E1α, and complete E1β and E2 components of 2-oxo acid dehydrogenase multienzyme complexes. Figure 8.3 shows the position of these open reading frames.

**Figure 8.3: Positions of the three open reading frames upstream of the DHLipDH (E3) gene.**
The E1α and E1β coding regions overlap by two nucleotides, with the ATG start codon of E1β sharing bases with the TGA stop codon of E1α. E1β and E2 are in frame with each other, separated by three bases. E2 and E3 share a single nucleotide, with the TAA stop codon of E2 overlapping with the ATG start codon of E3. The close spacing of these genes suggests that they are regulated as an operon.

8.3.2 Homology scores

8.3.2.1 The E2 gene

The E2 gene consists of 1566 bases which translate to 521 amino acids. It has a G+C content of 70.3%. The closest homology match is to the dihydrolipoamide acetyltransferase component (E2) of the pyruvate dehydrogenase complex of *Bacillus stearothermophilus*, with 37.8% identity in a 405 amino acid overlap. A dendrogram of the ten best matches is shown in Figure 8.4.

![Dendrogram](image.png)

Figure 8.4: Dendrogram showing relatedness of the ten best matches to the *H. volcanii* E2 sequence. Produced using PILEUP (GCG).
The ten best matches are aligned in Figure 8.5 below. Shaded residues are ones identical to those in the \textit{H. volcanii} protein.

\begin{verbatim}
H. volcanii E2
B. steaerothermophilus PDHC E2
B. subtilis PDHC E2
A. laidlawii PDHC E2
P. putida BCDC E2
A. vinelandii PDHC E2
bovine BCDC E2
human BCDC E2
E. coli PDHC E2
H. influenza OGDHC E2
B. subtilis BCDC E2

H. volcanii E2
B. steaerothermophilus PDHC E2
B. subtilis PDHC E2
A. laidlawii PDHC E2
P. putida BCDC E2
A. vinelandii PDHC E2
bovine BCDC E2
human BCDC E2
E. coli PDHC E2
H. influenza OGDHC E2
B. subtilis BCDC E2

H. volcanii E2
B. steaerothermophilus PDHC E2
B. subtilis PDHC E2
A. laidlawii PDHC E2
P. putida BCDC E2
A. vinelandii PDHC E2
bovine BCDC E2
human BCDC E2
E. coli PDHC E2
H. influenza OGDHC E2
B. subtilis BCDC E2

H. volcanii E2
B. steaerothermophilus PDHC E2
B. subtilis PDHC E2
A. laidlawii PDHC E2
P. putida BCDC E2
A. vinelandii PDHC E2
bovine BCDC E2
human BCDC E2
E. coli PDHC E2
H. influenza OGDHC E2
B. subtilis BCDC E2

H. volcanii E2
B. steaerothermophilus PDHC E2
B. subtilis PDHC E2
A. laidlawii PDHC E2
P. putida BCDC E2
A. vinelandii PDHC E2
bovine BCDC E2
human BCDC E2
E. coli PDHC E2
H. influenza OGDHC E2
B. subtilis BCDC E2

H. volcanii E2
B. steaerothermophilus PDHC E2
B. subtilis PDHC E2
A. laidlawii PDHC E2
P. putida BCDC E2
A. vinelandii PDHC E2
bovine BCDC E2
human BCDC E2
E. coli PDHC E2
H. influenza OGDHC E2
B. subtilis BCDC E2

H. volcanii E2
B. steaerothermophilus PDHC E2
B. subtilis PDHC E2
A. laidlawii PDHC E2
P. putida BCDC E2
A. vinelandii PDHC E2
bovine BCDC E2
human BCDC E2
E. coli PDHC E2
H. influenza OGDHC E2
B. subtilis BCDC E2

H. volcanii E2
B. steaerothermophilus PDHC E2
B. subtilis PDHC E2
A. laidlawii PDHC E2
P. putida BCDC E2
A. vinelandii PDHC E2
bovine BCDC E2
human BCDC E2
E. coli PDHC E2
H. influenza OGDHC E2
B. subtilis BCDC E2

H. volcanii E2
B. steaerothermophilus PDHC E2
B. subtilis PDHC E2
A. laidlawii PDHC E2
P. putida BCDC E2
A. vinelandii PDHC E2
bovine BCDC E2
human BCDC E2
E. coli PDHC E2
H. influenza OGDHC E2
B. subtilis BCDC E2

H. volcanii E2
B. steaerothermophilus PDHC E2
B. subtilis PDHC E2
A. laidlawii PDHC E2
P. putida BCDC E2
A. vinelandii PDHC E2
bovine BCDC E2
human BCDC E2
E. coli PDHC E2
H. influenza OGDHC E2
B. subtilis BCDC E2

H. volcanii E2
B. steaerothermophilus PDHC E2
B. subtilis PDHC E2
A. laidlawii PDHC E2
P. putida BCDC E2
A. vinelandii PDHC E2
bovine BCDC E2
human BCDC E2
E. coli PDHC E2
H. influenza OGDHC E2
B. subtilis BCDC E2

H. volcanii E2
B. steaerothermophilus PDHC E2
B. subtilis PDHC E2
A. laidlawii PDHC E2
P. putida BCDC E2
A. vinelandii PDHC E2
bovine BCDC E2
human BCDC E2
E. coli PDHC E2
H. influenza OGDHC E2
B. subtilis BCDC E2
\end{verbatim}
The *H. volcanii* E2 gene groups with the E2 of pyruvate dehydrogenase multienzyme complexes of Gram-positive bacteria. Perhaps the most interesting observation from the sequence alignments is that the *H. volcanii* gene is predicted to translate an extra hundred amino acids within the middle of the sequence (amino acid residues 155-256), not present in any of the other sequences.

There are at least five possible explanations for these extra residues. Firstly, it is possible that they have a role forming a large loop within the protein structure, but such a large extra structure seems unlikely. Secondly, they may be...
due to the presence of an intron, but these have never been found within archaeal structural genes. Thirdly, they may represent a fragment of protein that is spliced out of the final product, i.e. an intein (Colston & Davis, 1994). The inteins so far discovered have conserved junctions; these are not present in the *H. volcanii* E2 sequence making the intein explanation unlikely. By far the most likely explanations though, are that the extra sequence is due to either a transposition or a duplication event.

Evidence that either a transposon has interrupted the gene or that a duplication event has occurred is the presence of a ten nucleotide direct repeat at either end of the extra sequence. Such a repeat is often observed upon introduction of a transposable element. There are, however, no terminal inverted repeats. These are normally seen with transposons but are not absolutely essential as there are at least two cases in the literature of halobacterial transposons that lack them (Schnabel *et al.*, 1984; Pfeifer *et al.*, 1984).

To test whether a transposition or a duplication event is the most likely explanation, the extra sequence was translated in all six reading frames and compared with the SwissProt database. The greatest homology was to a hypothetical protein of a transposable element from *Streptomyces fradiae*, with an identity of 34% over 106 amino acids. This was in a different reading frame from the E2 gene. The extra sequence is G+C rich and will consequently give translations with high arginine and proline (R+P) content in most non-coding open reading frames. These will align with R+P rich proteins in the database, of which the transposon protein is one. This casts doubt on the significance of the alignment.
More convincing evidence exists for a duplication event, however. The extra translated sequence aligns partially with an upstream region of E2 (Figure 8.6).

<table>
<thead>
<tr>
<th>Insert</th>
<th>LAAPATRALAKEEGVDIDAV</th>
</tr>
</thead>
<tbody>
<tr>
<td>H. volcanii E2</td>
<td>FAPPSVRLARELGVDLDAV 139</td>
</tr>
</tbody>
</table>

Figure 8.6: Partial alignment of the E2 insert sequence against an upstream region of the translated sequence.

The possible duplication does not seem to correspond to a domain, since it interrupts the peripheral subunit-binding region, as evidenced by the identified position of this region on the aligned sequences. This region is approximately 35 residues in length, whereas the extra sequence is over a hundred amino acids. This would suggest that the duplication is not a functional domain rearrangement.

With the extra sequence removed, the homology scores were recalculated and are shown in Table 8-1.

<table>
<thead>
<tr>
<th>Organism/protein</th>
<th>% Identity (amino acid overlap)</th>
</tr>
</thead>
<tbody>
<tr>
<td>B. stearothermophilus PDHC E2</td>
<td>47.8% (425 aa)</td>
</tr>
<tr>
<td>B. subtilis PDHC E2</td>
<td>43.0% (386 aa)</td>
</tr>
<tr>
<td>A. laidlawii PDHC E2</td>
<td>41.5% (429aa)</td>
</tr>
<tr>
<td>E. coli PDHC E2</td>
<td>38.8% (428 aa)</td>
</tr>
<tr>
<td>B. subtilis BCDHC E2</td>
<td>37.8% (421 aa)</td>
</tr>
<tr>
<td>A. vinelandii PDHC E2</td>
<td>36.7% (417 aa)</td>
</tr>
<tr>
<td>bovine BCDHC E2</td>
<td>35.5% (425 aa)</td>
</tr>
<tr>
<td>P. putida BCDHC E2</td>
<td>34.4% (425 aa)</td>
</tr>
<tr>
<td>human BCDHC E2</td>
<td>34.4% (425 aa)</td>
</tr>
<tr>
<td>E. coli OGDHC E2</td>
<td>34.1% (419 aa)</td>
</tr>
</tbody>
</table>

Table 8-1: Homology scores for the H. volcanii E2 protein.

8.3.2.2 The E1β gene

The E1β gene consists of 984 bases, which translate to 327 amino acids. It has a G+C content of 66.2%. The closest homology match is to the pyruvate
dehydrogenase E1β component of *B. stearothermophilus*, with 56.3% identity in a 323 amino acid overlap. A dendrogram of the ten best matches is shown in Figure 8.7.

![Dendrogram showing relatedness of the ten best matches to the H. volcanii E1β sequence. Produced using PILEUP (GCG).](image)

The ten best matches are aligned in Figure 8.8 below. Shaded residues are ones identical to those in the *H. volcanii* protein.

**N. volcanii E1β**
B. stearothermophilus PDHC E1β
B. subtilis PDHC E1β
A. laidlawii PDHC E1β
B. subtilis BCDHC E1β
P. putida BCDHC E1β
M. genitalium PDHC E1β
human BCDHC E1β
A. eutrophus acetoin ox/red β
bovine BCDHC E1β
rat BCDHC E1β

MAVAAAAGW LLRLRAAGAE GHWRRLPGAG LARGFLHPAA TVEDAAQRRQ 50
MAVAAAFAGW LLRLRAAGAD GPWRRLCGAG LSRGFLQSAS AYGAAAQRRQ 50
MAVAAAAGW LLRLRAAGAD GPWRRLCGAG LSRGFLQSAS AYGAAAQRRQ 50
MAVAAAAGW LLRLRAAGAD GPWRRLCGAG LSRGFLQSAS AYGAAAQRRQ 50
MAVAAAAGW LLRLRAAGAD GPWRRLCGAG LSRGFLQSAS AYGAAAQRRQ 50
MAVAAAAGW LLRLRAAGAD GPWRRLCGAG LSRGFLQSAS AYGAAAQRRQ 50
MAVAAAAGW LLRLRAAGAD GPWRRLCGAG LSRGFLQSAS AYGAAAQRRQ 50
MAVAAAAGW LLRLRAAGAD GPWRRLCGAG LSRGFLQSAS AYGAAAQRRQ 50
MAVAAAAGW LLRLRAAGAD GPWRRLCGAG LSRGFLQSAS AYGAAAQRRQ 50

159
H. volcanii EIP

- B. steathermophilus PDNC EIP
- M. genitalium PDHC Eip
- A. eutrophus Acetoin ox/red P
- A. laidlawii PDHC Eip
- H. volcanii Eip

- B. subtilis PDHC Eip
- Rat BCDHC Eip
- P. putida BCDHC Eip
- B. subtilis BCDHC Eip
- P. putida BCDHC Eip

- A. eutrophus Acetoin ox/red β
- B. subtilis PDHC Eip
- Rat BCDHC Eip

- B. steathermophilus PDNC EIP
- M. genitalium PDHC Eip
- A. laidlawii PDHC Eip
- B. subtilis BCDHC Eip
- P. putida BCDHC Eip
- H. volcanii Eip

- B. subtilis PDHC Eip
- Rat BCDHC Eip

- B. steathermophilus PDNC EIP
- M. genitalium PDHC Eip
- A. laidlawii PDHC Eip
- B. subtilis BCDHC Eip
- P. putida BCDHC Eip
- H. volcanii Eip

- B. subtilis PDHC Eip
- Rat BCDHC Eip

- B. steathermophilus PDNC EIP
- M. genitalium PDHC Eip
- A. laidlawii PDHC Eip
- B. subtilis BCDHC Eip
- P. putida BCDHC Eip
- H. volcanii Eip

- B. subtilis PDHC Eip
- Rat BCDHC Eip

- B. steathermophilus PDNC EIP
- M. genitalium PDHC Eip
- A. laidlawii PDHC Eip
- B. subtilis BCDHC Eip
- P. putida BCDHC Eip
- H. volcanii Eip

- B. subtilis PDHC Eip
- Rat BCDHC Eip

- B. steathermophilus PDNC EIP
- M. genitalium PDHC Eip
- A. laidlawii PDHC Eip
- B. subtilis BCDHC Eip
- P. putida BCDHC Eip
- H. volcanii Eip

- B. subtilis PDHC Eip
- Rat BCDHC Eip

- B. steathermophilus PDNC EIP
- M. genitalium PDHC Eip
- A. laidlawii PDHC Eip
- B. subtilis BCDHC Eip
- P. putida BCDHC Eip
- H. volcanii Eip

- B. subtilis PDHC Eip
- Rat BCDHC Eip

- B. steathermophilus PDNC EIP
- M. genitalium PDHC Eip
- A. laidlawii PDHC Eip
- B. subtilis BCDHC Eip
- P. putida BCDHC Eip
- H. volcanii Eip

- B. subtilis PDHC Eip
- Rat BCDHC Eip

- B. steathermophilus PDNC EIP
- M. genitalium PDHC Eip
- A. laidlawii PDHC Eip
- B. subtilis BCDHC Eip
- P. putida BCDHC Eip
- H. volcanii Eip

- B. subtilis PDHC Eip
- Rat BCDHC Eip

- B. steathermophilus PDNC EIP
- M. genitalium PDHC Eip
- A. laidlawii PDHC Eip
- B. subtilis BCDHC Eip
- P. putida BCDHC Eip
- H. volcanii Eip

- B. subtilis PDHC Eip
- Rat BCDHC Eip

- B. steathermophilus PDNC EIP
- M. genitalium PDHC Eip
- A. laidlawii PDHC Eip
- B. subtilis BCDHC Eip
- P. putida BCDHC Eip
- H. volcanii Eip

- B. subtilis PDHC Eip
- Rat BCDHC Eip

- B. steathermophilus PDNC EIP
- M. genitalium PDHC Eip
- A. laidlawii PDHC Eip
- B. subtilis BCDHC Eip
- P. putida BCDHC Eip
- H. volcanii Eip

- B. subtilis PDHC Eip
- Rat BCDHC Eip

- B. steathermophilus PDNC EIP
- M. genitalium PDHC Eip
- A. laidlawii PDHC Eip
- B. subtilis BCDHC Eip
- P. putida BCDHC Eip
- H. volcanii Eip

- B. subtilis PDHC Eip
- Rat BCDHC Eip

- B. steathermophilus PDNC EIP
- M. genitalium PDHC Eip
- A. laidlawii PDHC Eip
- B. subtilis BCDHC Eip
- P. putida BCDHC Eip
- H. volcanii Eip

- B. subtilis PDHC Eip
- Rat BCDHC Eip
Figure 8.8: Alignment of *H. volcanii* E1β with the ten closest matching sequences.

The homology scores are shown in Table 8-2 below.

<table>
<thead>
<tr>
<th>Organism/protein</th>
<th>% Identity (amino acid overlap)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>B. stearothermophilus</em> PDHC E1β</td>
<td>56.3% (323 aa)</td>
</tr>
<tr>
<td><em>B. subtilis</em> PDHC E1β</td>
<td>54.2% (323 aa)</td>
</tr>
<tr>
<td><em>A. laidlawii</em> PDHC E1β</td>
<td>50.8% (317 aa)</td>
</tr>
<tr>
<td><em>P. putida</em> BCDHC E1β</td>
<td>48.0% (342 aa)</td>
</tr>
<tr>
<td><em>B. subtilis</em> BCDHC E1β</td>
<td>46.9% (324 aa)</td>
</tr>
<tr>
<td><em>M. genitalium</em> PDHC E1β</td>
<td>46.0% (326 aa)</td>
</tr>
<tr>
<td><em>A. eutrophus</em> Acetoin ox/red β subunit</td>
<td>44.2% (335 aa)</td>
</tr>
<tr>
<td>human BCDHC E1β</td>
<td>42.2% (327 aa)</td>
</tr>
<tr>
<td>bovine BCDHC E1β</td>
<td>41.8% (328 aa)</td>
</tr>
<tr>
<td>rat BCDHC E1β</td>
<td>37.6% (327 aa)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Organism/protein</th>
<th>% Identity (amino acid overlap)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>H. volcanii</em> E1β</td>
<td>56.3% (323 aa)</td>
</tr>
<tr>
<td><em>B. stearothermophilus</em> PDHC E1β</td>
<td>54.2% (323 aa)</td>
</tr>
<tr>
<td><em>B. subtilis</em> PDHC E1β</td>
<td>50.8% (317 aa)</td>
</tr>
<tr>
<td><em>A. laidlawii</em> PDHC E1β</td>
<td>48.0% (342 aa)</td>
</tr>
<tr>
<td><em>P. putida</em> BCDHC E1β</td>
<td>46.9% (324 aa)</td>
</tr>
<tr>
<td><em>M. genitalium</em> PDHC E1β</td>
<td>46.0% (326 aa)</td>
</tr>
<tr>
<td><em>A. eutrophus</em> Acetoin ox/red β subunit</td>
<td>44.2% (335 aa)</td>
</tr>
<tr>
<td>human BCDHC E1β</td>
<td>42.2% (327 aa)</td>
</tr>
<tr>
<td>bovine BCDHC E1β</td>
<td>41.8% (328 aa)</td>
</tr>
<tr>
<td>rat BCDHC E1β</td>
<td>37.6% (327 aa)</td>
</tr>
</tbody>
</table>

Table 8-2: Homology scores for the *H. volcanii* E1β protein.

### 8.3.2.3 The E1α gene

Only 207 nucleotides of the 3' terminus of an E1α was sequenced from pNAT82. These translate to 68 amino acid residues. Homologies to any known sequences were low but the greatest was to the PDHC E1α component from *A. laidlawii* (36.5% identity over 52 amino acids).
8.3.3 Evidence of regulatory regions

A search was made of the entire pNAT82 sequence for possible promoters and other regulatory regions. A putative Box A element was identified by Vettakkorumakankav & Stevenson (1992) in their paper describing the cloning and sequencing of the DHLipDH gene. This sequence, TGAAGGAA, appears 35 nucleotides upstream of the DHLipDH start codon (nt 2713-2720 in the full sequence), but is not consistent with the halobacterial Box A consensus of (T/C)(T/A)(A/T)A (Palmer & Daniels, 1995). The sequence was not found to be sufficient for expression as the region was included in the DHLipDH insert in our first homologous expression experiment, which did not result in protein production. Vettakkorumakankav & Stevenson (1992) claim, however, that Northern blot analysis revealed that the DHLipDH gene was transcribed from its own promoters.

The only region that identified with the consensus Box A sequence in the entire pNAT82 clone was nt 4212-4216. This is at the 3' end of the DHLipDH coding region and is thought to be part of the stem-loop termination sequence.

Box B sequences contain the transcription start site. They have only weak consensus and consist of two nucleotides, (T/C)(G/A) (Hain et al., 1992). It was pointless searching for such a motif as it would appear too frequently throughout the sequence.

The motif AAGTTA is considered important in the binding of RNA polymerase (Dennis, 1985). The sequence was not found within pNAT82.

If the genes are transcribed as an operon, one would expect to find the control regions upstream of the E1α gene, which has yet to be sequenced.
8.3.4 Structural motifs of the E2 gene

E2 polypeptide chains consist of N-terminal lipoyl domains followed by a peripheral subunit-binding domain and the C-terminal acyltransferase (inner-core) domain (Figure 8.9). The number of lipoyl domains depends on the species of origin. Domains are connected by extended, flexible linker regions (Perham, 1996).

Figure 8.9: The segmented structure of the E2 polypeptide chain. The N-terminal lipoyl domains (approximately 80 residues) are followed by the peripheral subunit-binding domain (approximately 35 residues) and the C-terminal acyltransferase domain. The domains are connected by flexible but extended linker regions (approximately 25-30 residues). The number of lipoyl domains varies with the origin of the E2 chain. Taken from Perham (1996).
8.3.4.1 The lipooyl domain

Lipooyl domains consist of approximately 80 residues arranged in two β-sheets forming a flattened β-barrel, with a core of hydrophobic residues (Green et al., 1995). The lipooyl domain of the E2 from *H. volcanii*, extending from residues 1 to 84, is evident by the conserved DKA motif around the lipooyl lysine, K44 (Dardel et al., 1993), as well as high homology throughout the region. There is also a large proportion of hydrophobic residues in this area. The structure is similar to the *B. stearothermophilus* PDHC E2 in having a single lipooyl domain.

8.3.4.2 The peripheral subunit-binding domain

Peripheral subunit-binding domains of E2 chains consist of approximately 35 residues and form two parallel α-helices, separated by a short strand, a helix-like turn and an irregular loop, with a hydrophobic core (Robien et al., 1992; Kalia et al., 1993). From the sequence alignments, the domain in *H. volcanii* consists of residues 119-260, interrupted by the putative duplication sequence. It has been suggested that an aspartate and a threonine residue in the *B. stearothermophilus* PDHC E2 form a hydrogen bond crucial to the domain stability (Kalia et al., 1993). The aspartate is conserved throughout all known E2 sequences, including the *H. volcanii* E2 (D152), whereas the threonine is replaced by a serine in many structures. In *H. volcanii* E2 it appears as S142. Since the two residues are similar in structure, it is predicted that the stabilising hydrogen bond is present in the *H. volcanii* structure.

Many of the hydrophobic residues in this region are also conserved, whereas others are exchanged for different hydrophobic groups.

Most of the residues involved in binding the E3 dimer are conserved between the *H. volcanii* and *B. stearothermophilus* structures (detailed in Mandé
et al., 1996). S124, R126, R130, R147 are all conserved, whereas R127 is exchanged with a similarly-charged lysine residue in the *B. stearothermophilus* structure. P123 is also conserved and acts as a helix promoter.

G141 is conserved throughout most E2 sequences aligned and has been implicated in the binding of E1 (Kalia *et al.*, 1993). E150, however, is not conserved in the *H. volcanii* structure. In most other sequences it is either a lysine or an arginine and has been implicated in electrostatic interactions with E1.

### 8.3.4.3 The catalytic domain

The catalytic domain, from residues 264-521 approximately, is highly conserved, especially around the active site histidine, H493.

### 8.4 The remaining E1α sequence

In order to obtain the remaining E1α sequence not included in the pNAT82 clone, a *H. volcanii* cosmid library was probed. The library is located at the University of Alicante in Spain and was created by Consuelo Ferrer by the methodology of Charlebois *et al.* (1989). The probe was created by PCR amplifying pNAT82 using primers KAJ-26 and KAJ-27 with Vent® DNA polymerase. The PCR program used was as follows: 96°C for 4 minutes, followed by 30 cycles of 96°C for 75 s, 50°C for 90 s and 72°C for 2 minutes. The product of this reaction was approximately 1 kb in length. This was sent to Spain and was found to hybridise to a single cosmid (126) in the library, which was sent back.

The cosmid was purified using a Qiagen Tip 20 column. Attempts were made to sequence the E1α coding region directly from the cosmid using primer KAJ-36.

KAJ-36: 5' tgcgtggtgaacatctcgctcgg 3' 24-mer
No success at sequencing was achieved, due to a high noise to signal ratio. This is a common problem with attempts to sequence large DNA molecules, the cosmid being 40 kb in length.

In order to obtain the remaining sequence, it is envisaged that the cosmid DNA will be size-restricted and subcloned into a more convenient sequencing vector. Time constraints prevented this approach being followed in this project.

8.5 Growth experiments

Experiments comparing growth properties of wild-type and BAS5005 (DHLipDH - minus) *H. volcanii* were conducted in an attempt to identify the metabolic defect of the mutant. Cells were grown on minimal medium agar containing a range of compounds as sole carbon sources at concentrations of 0.2% to 0.5% (w/v). This work was conducted mainly by Dr. M.J. Danson (University of Bath) and Dr. M. Dyall-Smith (University of Melbourne, Australia). Table 8-3 summarises the results of these experiments. No differences in growth were observed between wild-type and BAS5005.

Analysis of the first sequencing reactions, prior to completion of pNAT82, suggested homology to an acetoin dehydrogenase complex. To test if such a complex was present in *H. volcanii*, wild-type, BAS5005 cells and BAS5005 transformed with the DHLipDH overexpressing plasmid pBAP5003 were grown on minimal media plates with acetoin as the sole carbon source. Similar low levels of growth were observed for all strains on 0.5% acetoin, suggesting that DHLipDH plays no essential role in acetoin metabolism.
Table 8-3: Growth properties of wild-type and BAS5005 (DHLipDH - minus) strains of \textit{H. volcanii}. Key: -, no growth; w, weak growth; +, ++, or +++., positive growth (Jolley et al., 1996).

<table>
<thead>
<tr>
<th>Carbon source</th>
<th>Wild-type growth</th>
<th>BAS5005 growth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acetate</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Citrate</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Fructose</td>
<td>+++</td>
<td>+++</td>
</tr>
<tr>
<td>Galactose</td>
<td>w</td>
<td>w</td>
</tr>
<tr>
<td>Glucose</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td>Glutamate</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td>Glycerol</td>
<td>+++</td>
<td>+++</td>
</tr>
<tr>
<td>Glycine</td>
<td>w</td>
<td>w</td>
</tr>
<tr>
<td>Isoleucine</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Leucine</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Pyruvate</td>
<td>+++</td>
<td>+++</td>
</tr>
<tr>
<td>Sorbitol</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Succinate</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td>Sucrose</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td>Valine</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

8.6 Discussion

The identification of two E1 genes in \textit{H. volcanii} places its putative 2-oxo acid dehydrogenase complex nearer to the eukaryote and Gram-positive bacterial PDH complexes. These organisms have an \(\alpha_2\beta_2\) E1 heterotetramer in the PDHC, whereas Gram-negative bacteria usually have a homodimeric E1 protein. The mycoplasmas are an exception, although these group phylogenetically with Gram-positive bacteria. Both the \textit{H. volcanii} E1\(\beta\) and E2 genes group with homologous components of Gram-positive PDH complexes. The E3 gene similarly groups with the DHLipDH gene from Gram-positive bacteria (data not shown).

Although no activities for any 2-oxo acid dehydrogenase multienzyme complex have been reported in the Archaea, the sequencing of E2 and E1 components in \textit{H. volcanii} provides evidence of their one-time presence. As the components group with those from PDH complexes rather than from BCDHC or
OGDHC, it is possible that an ancestral PDHC existed. The presence of a putative duplication sequence within the \textit{H. volcanii} E2 gene provides an explanation for its apparent lack of functional enzymic activity.

The similar degrees of identity of the E1\textbeta, E2 and E3 genes of \textit{H. volcanii} to the homologous genes of the \textit{B. stearothermophilus} PDHC might suggest that any inactivation of the halophilic E2 gene occurred relatively recently in evolution. The identity scores are 49.6\% for E3, 47.8\% for E2 (not including the insert sequence) and 56.3\% for the E1\textbeta. Further evidence for a recent inactivation is the high degree of conservation in the binding and catalytic regions of the putative E2 gene product. If this is the case, the question of whether DHLipDH is a remnant of a now inactive complex or whether it occupies another role still needs to be asked. If it was solely a remnant, there would be no evolutionary pressure selecting for its continued activity and yet the gene codes for an active DHLipDH and it is clearly expressed within the cell. Its role, however, remains unclear, especially in the light that no metabolic defect was identified in the DHLipDH - minus mutant strain. If the DHLipDH gene is part of an operon, as it appears to be, the other genes within the operon must also be expressed.

The recently-completed sequencing of the entire genome of \textit{Methanococcus jannaschii} (Bult et al., 1996) shows the presence of a DHLipDH gene in the apparent absence of genes for E1 and E2 components of a 2-oxo acid dehydrogenase complex. This would suggest that the possible explanation for a lack of PDHC activity in \textit{H. volcanii} is not applicable to all other Archaea and it seems likely that the DHLipDH in \textit{H. volcanii} occupies a role other than as a remnant of a PDH complex. DHLipDH is also found in the thermophilic archaeon \textit{Thermoplasma acidophilum} in the apparent absence of a PDHC (Smith et al., 1987).
Dihydrolipoamide is known to possess antioxidant properties (Suzuki et al., 1994) and it has been implicated in the repair of membrane thiol groups in mitochondria (Zimmer et al., 1991). The role of DHLipDH in *H. volcanii* and other Archaea may, therefore, be in repair mechanisms. Such a role would not be evident from simple growth experiments.
9. General discussion

9.1 Expression studies

Heterologous expression of foreign proteins in bacteria such as *E. coli* is a compromise. The advantages of such a system often outweigh potential disadvantages. These advantages include high cell densities, rapid growth and detailed knowledge of the surrogate genetic system. Transformation and other manipulative techniques are routine, abrogating the requirements of development and optimisation of such protocols in the natural host. There is a considerable choice of commercial expression vectors available, offering a range of inducible promoters that provide high levels of control over expression. This inducibility allows cells to reach high cell densities rapidly before their translation machinery is subverted to the production of the foreign protein.

Clearly there are difficulties in using a eubacterial system for expressing proteins whose native environment is highly saline. Although success has been achieved in the past, this project has demonstrated that such expression is not necessarily facile. The advantages of using a homologous expression system, therefore, are obvious. Any protein translated in its native environment would be expected to adopt its correct conformation and to undergo any necessary post-translational modifications. The disadvantages, though, are slow growth, a lack of inducible promoter systems and an incomplete understanding of the intracellular biochemistry of these organisms. Another consideration comes when considering a scale-up of expression using large-scale fermenters. The saline growth medium required of halophilic organisms is highly corrosive to stainless steel, often used in such systems, which would probably need to be replaced with more resistant materials such as glass. For research purposes,
however, on a laboratory scale, homologous expression of halophilic proteins represents a useful system and creation of the pBAP5009 shuttle vector should enable the expression of any cloned halophilic gene.

9.2 The nature of halophilic proteins

It is generally accepted that the predominant characteristic of halophilic proteins is a high degree of negatively-charged residues on the surface. These allow the protein to maintain a hydration shell by competing with the bulk solvent for hydrated salt ions. A concomitant reduction in the surface hydrophobicity counteracts the tendency of proteins to precipitate under these saline conditions. DHLipDH from *H. volcanii* conforms to these observations.

The identification of a potential *K*⁺-binding site at the dimer interface, however, suggests that the surface charge is not the sole structural feature important in conferring halophilicity. Halophilic proteins need to maintain the spatial alignment of key residues in the catalytic and binding regions of the core, while under increasingly saline conditions the molecule is under pressure to contract from the increased surface tension at the protein-solvent interface. Occupancy of the *K*⁺-binding site may well increase the stability of the core, especially as the salt concentration increases. As more halophilic structures become available, it will be interesting to see if such features are common. The tetrameric structure of malate dehydrogenase from *H. marismortui* (Dym et al., 1995), consisting of two dimers, does not however appear to have such a site between the monomers making up each dimer.

The hypotheses of structural features involved in conferring halophilicity, based on precise spatial arrangements of atoms, must be treated with a degree of caution when those arrangements are based on a molecular model. Clearly, an X-ray crystal structure is desirable, and the crystallisation trials have shown
that halophilic DHLipDH will form crystals and have identified starting conditions around which future trials can be optimised. It is still valid to make predictions from model structures as long as one remains aware of the potential positional uncertainties.

9.3 Phylogenetic implications

The discovery of genes coding for putative E1 and E2 components of a 2-oxo acid dehydrogenase complex in the Archaea is the first indication of their possible existence within these organisms. The presence of a large duplicated region within the E2 gene that would be expected to perturb its binding of the E3 component provides an explanation for the lack of detectable activity. Disruption of binding between E3 and the putative E2 gene product complements the discovery that in gel filtration of whole-cell extracts of halophilic Archaea, all DHLipDH activity is present in a single peak representing free enzyme (Danson et al., 1984). In these experiments, the possibility exists that in the process of making cell extracts, large multienzyme complex structures may have dissociated. Since no activity at all was detected in fractions larger than the DHLipDH dimer, however, it seems likely that the enzyme exists solely in an unassociated form in these organisms. The function of an unassociated DHLipDH may be in oxidative repair mechanisms (Suzuki et al., 1994; Zimmer et al., 1991) or in sugar transport (Richarme, 1989; Richarme & Heine, 1986), although the growth experiments seem to rule out the latter.

If a 2-oxo acid dehydrogenase multienzyme complex existed at one time in the halophilic Archaea, it would appear that the case is not universal across the archaeal domain. Since the putative E2 and E1 genes and the functional DHLipDH gene of *H. volcanii* have very similar percentage identities to the homologous components of the *B. stearothermophilus* PDHC, any evolutionary
event that resulted in a loss of function for the complex would be expected to
have occurred relatively recently. This is consistent with the analysis of the
sequencing of the entire genome of Methanococcus janaschii (Bult et al., 1996).
This organism, while having a DHLipDH gene, lacks genes for either the E1 or
E2 components, unless their coding regions are so dissimilar from their
homologues that they are not recognisable.

One experiment that may yield conclusive insight to the former role of the
halophilic 2-oxo acid dehydrogenase complex would be to remove the duplicated
region of the E2 gene. It is possible that removal of this extraneous DNA would
allow the expression of a fully active complex. Whether this would work depends
partly on how long ago the duplication event occurred which would effect the
degree of mutation within the gene. Once the remainder of the E1α gene
sequence is obtained, it may be possible to express all the components of the
multienzyme complex. It cannot be absolutely ruled out that the genes do code
for an active complex whose expression is tightly regulated and is hence not
detectable under normal assay conditions. Sequencing of regions upstream of
the E1α gene may provide clues to the regulation of the operon and the
identification of other nearby genes may provide further evidence of its evolved
role.
References


Smith, P.K., Krohn, R.I., Hermanson, G.T., Mallia, A.K., Gartner, F.H.,
Provenzano, M.D., Fujimoto, E.K., Goeke, N.M., Olsen, B.J. & Klenk, D.C.
76-85

Srere, P.A., Brazil, H. & Gonen, L. (1963) The citrate condensing enzyme of

sequence of the lipoamide dehydrogenase gene of *Escherichia coli* K12. *Eur. J.
Biochem.*, 135, 519-527

T7 RNA polymerase to direct expression of cloned genes. *Methods in Enzymology*, 185, 60-89

glutamylcystine reductase and dihydrolipoamide dehydrogenase from
*Halobacterium halobium*: Purification by immobilized-metal-ion affinity
chromatography and properties of the enzymes. *J. Bacteriol.*, 170, 3459-3467

Suzuki, Y.J., Tsuchiya, M. & Parker, L. (1994) Determination of structure-

Biol.*, 2, 35-39

Tindall, K.R. & Kunkel, T.A. (1988) Fidelity of DNA synthesis by the *Thermus
aquaticus* DNA polymerase. *Biochemistry*, 27, 6008-6013

carbohydrates by extremely halophilic bacteria: glucose metabolism via a


Arch. Biochem. Biophys., 288, 609-613