LEABHARLANN CHOLÁISTE NA TRÍONÓIDE, BAILE ÁTHA CLIATH | TRINITY COLLEGE LIBRARY DUBLIN
Ollscoil Átha Cliath | The University of Dublin

Terms and Conditions of Use of Digitised Theses from Trinity College Library Dublin

Copyright statement

All material supplied by Trinity College Library is protected by copyright (under the Copyright and Related Rights Act, 2000 as amended) and other relevant Intellectual Property Rights. By accessing and using a Digitised Thesis from Trinity College Library you acknowledge that all Intellectual Property Rights in any Works supplied are the sole and exclusive property of the copyright and/or other IPR holder. Specific copyright holders may not be explicitly identified. Use of materials from other sources within a thesis should not be construed as a claim over them.

A non-exclusive, non-transferable licence is hereby granted to those using or reproducing, in whole or in part, the material for valid purposes, providing the copyright owners are acknowledged using the normal conventions. Where specific permission to use material is required, this is identified and such permission must be sought from the copyright holder or agency cited.

Liability statement

By using a Digitised Thesis, I accept that Trinity College Dublin bears no legal responsibility for the accuracy, legality or comprehensiveness of materials contained within the thesis, and that Trinity College Dublin accepts no liability for indirect, consequential, or incidental, damages or losses arising from use of the thesis for whatever reason. Information located in a thesis may be subject to specific use constraints, details of which may not be explicitly described. It is the responsibility of potential and actual users to be aware of such constraints and to abide by them. By making use of material from a digitised thesis, you accept these copyright and disclaimer provisions. Where it is brought to the attention of Trinity College Library that there may be a breach of copyright or other restraint, it is the policy to withdraw or take down access to a thesis while the issue is being resolved.

Access Agreement

By using a Digitised Thesis from Trinity College Library you are bound by the following Terms & Conditions. Please read them carefully.

I have read and I understand the following statement: All material supplied via a Digitised Thesis from Trinity College Library is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of a thesis is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form providing the copyright owners are acknowledged using the normal conventions. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone. This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

THE ALCOHOL DEHYDROGENASES OF *Bacillus stearothermophilus*

A dissertation submitted to University of Dublin, Trinity College,

for the degree of Doctor of Philosophy

by

Glynis A. Robinson, BA (Open)

Department of Biochemistry,

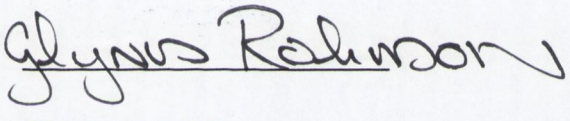
Trinity College, Dublin.

2000

TRINITY COLLEGE
24 JAN 2001
LIBRARY DUBLIN
Thesis
6158

DECLARATION

I declare that this thesis is entirely my own work except where duly acknowledged. This thesis has not previously been submitted for a degree at this or at any other university. I agree that this thesis may be lent or copied at the discretion of the Librarian, Trinity College, Dublin.



Glynis A. Robinson

ACKNOWLEDGEMENTS

At the very first, I wish to thank my supervisor, Dr. Chistopher J. Bailey, for his guidance, good advice, patience and for putting up with me. A special thank you to Dr. Mark Carrington of the Department of Biochemistry, University of Cambridge, for his assistance with the construction of the expression system, and for the use of his facilities. I thank Dr. David Doff and Mr. Sean Walsh for assistance with atomic absorption spectroscopy and Dr. David Doff for the use of his facilities. I thank Dr. Barbara Dowds for the pADH1b clone, Ms Ruth Kelleher for the mp18*adh* clone and Mr M.J. Runswick, Medical Research Council, Hills Road, Cambridge, for his generous gift of purified ADH 1503. I thank Dr. Ken Wolfe who performed the automated DNA-sequencing analysis and Dr. Mark Lalor who performed the oligonucleotide synthesis. I am eternally grateful to Maria Redpath, Ciaran Cronin, Mary O'Reilly and Orla Ennis for revealing to me the mysteries of molecular biology. I thank Louis Duffy for his willing and friendly help in taking and printing photographs. I thank Sinéad Boyce, Bernie Butler, Ed Franklin and Gavin McManus for proof reading the final text, and a big thank you to Natalie Voorheis for sorting out the references.

I thank Professor Keith Tipton for his advice, and for the use of the facilities of the Biochemistry Department during his time as Head of Department. I thank the present Head of Department, Professor Clive Williams for the use of the Department's facilities and, especially, for the assistance which he has given me during the writing up of this thesis. I thank Dr. Timothy Mantle for his assistance during his time as Director of The Centre For Pharmaceutical Biotechnology and I acknowledge the financial assistance of BioResearch Ireland.

I thank Maria Redpath for her friendship, her companionship during our days in Lab. 6, and for her help during my time in Cambridge. For his interest, help, friendship, and inexhaustible store of references, I thank Thomas Dowling. Many members of The Biochemistry, Genetics and Microbiology Departments and The Hamilton Library have helped me in many ways and I would particularly like to acknowledge Tim Mantle, Joe Carrol, Frank Winder, Jane Irwin, Larry Vaughan, Mary Kelly, Mary Meehan, Sandra Watkins and the late Helen Doherty.

For her support and encouragement, over many years, I thank Professor Jane Nelson of the Open University.

I thank Paul for everything.

SUMMARY

A 1.5-kb fragment of *Bacillus stearothermophilus* DSM 2334 DNA was subcloned into pUC18, and the nucleotide sequence was determined. The fragment contained a complete open reading frame that encoded a 339-residue amino-acid sequence. The deduced amino-acid sequence was typical of a bacterial, NAD^+-dependent, alcohol dehydrogenase (ADH). Seven residues that constitute the ligands to two zinc ions in horse liver ADH were strictly conserved. The N-terminal amino-acid sequence (forty residues) of purified *B. stearothermophilus* DSM 2334 alcohol dehydrogenase (ADH 2334) was determined. The amino-acid sequence-data confirmed that the nucleotide sequence of the complete open reading frame was that of the gene (*adh*2334) that encodes ADH 2334.

Sequence motifs typical of *Bacillus* transcription and translation signals were identified upstream from *adh*2334. The *adh*2334 stop codon overlapped the start codon of a second open reading frame. The second open reading frame encoded a seventy-three amino-acid sequence that had 46% identity with the N-terminal amino-acid sequence of human cytosolic aldehyde dehydrogenase. A putative ribosome-binding site for the gene encoding aldehyde dehydrogenase was identified within the coding sequence of *adh*2334. It is proposed that, in *B. stearothermophilus* DSM 2334, the genes encoding alcohol dehydrogenase and aldehyde dehydrogenase constitute part of an operon, and that the function of the operon is to co-ordinately express the genes encoding those enzymes involved in the conversion of ethanol to acetyl-CoA.

Peptides of ADH from *B. stearothermophilus* NCA 1503 (ADH 1503) were purified. N-terminal amino-acid sequencing of the peptides, and of the intact protein, provided sufficient data to confirm the amino-acid sequence deduced from the nucleotide sequence. An alignment of the amino-acid sequences of ADH 2334 and ADH 1503 revealed that ninety-five amino acids were not conserved between the two enzymes.

The amino acids involved in binding alcohol and NAD^+ to ADH 2334 and to ADH 1503 were determined, by alignment of the amino-acid sequences with those of horse liver ADH and fermentative ADH of yeast. The area of lowest sequence identity between the *B. stearothermophilus* ADHs was within a region identified as part of the NAD^+-binding domain. The residues deduced to be in contact with the nicotinamide ring, the nicotinamide ribose and diphosphate were conserved between the *B. stearothermophilus* sequences, but, of the nine residues that were assigned a role in adenine-binding, six were not conserved. ADH 2334 and ADH 1503 were different at three of the fifteen positions which were assigned a role in binding to

alcohol. The primary structure of the alcohol-binding pocket of both enzymes was found to be more like that of yeast ADH than that of horse liver ADH. The relevance of the differences between the primary structure of the *B. stearothermophilus* enzymes to the differences between substrate specificties and between the kinetic mechanisms of the two enzymes are discussed.

The *B. stearothermophilus* amino-acid sequences were examined for features purported to convey thermostability to enzymes. The substitutions Ile to Val and Lys to Glu were identified as being the most frequent substitutions of the type Xaa to Val and Xaa to Glu, when substitutions that convert typical mesophile ADH to *B. stearothermophilus* ADH were surveyed. Two proline residues were identified as being unique to *B. stearothermophilus* ADHs.

The polymerase chain reaction was used to introduce restriction sites at each end of *adh*2334. A 1 kb-*adh* PCR product, which contained the complete coding sequence of ADH 2334, was cloned into the expression vector pET21d. The recombinant *adh* was expressed in *Escherichia coli* at the level of at least 30% of the soluble protein produced by the host cell. The recombinant protein was purified and characterised. That the recombinant protein was ADH 2334 was confirmed by polyacrylamide-gel electrophoresis, amino-acid sequence analysis and enzyme assay. The specific activity of the purified recombinant protein was lower than that reported for the purified wild-type enzyme, and the recombinant enzyme was notably unstable. Atomic absorption spectroscopy demonstrated that the recombinant protein contained 1 Zn^{2+} per mol of subunit ADH and was, therefore, deficient in Zn^{2+}. *In vitro* addition of Zn^{2+} to the purified enzyme did not increase the specific activity.

Supplementation of the culture medium with Zn^{2+} increased ADH activity by up to four-fold. The recombinant enzyme, purified from culture medium supplemented with Zn^{2+}, was stable, contained about 2 Zn^{2+} per mol of subunit ADH, and had a specific activity of 2.3 U mg^{-1}, slightly higher than the value of 2.1 U mg^{-1} recorded for the wild-type enzyme. It is proposed that high-level expression of recombinant, zinc-dependent proteins can overwhelm the ability of *E. coli* to provide sufficient zinc co-factor, unless culture conditions are suitably adjusted.

Recombinant ADH, both the zinc-deficient enzyme and the enzyme purified from culture medium supplemented with Zn^{2+}, was found to be cold-inactivated at 0 °C. The cold-inactivation was reversible over 10 min at 37 °C.

Abbreviations

Abbreviations used in the text and not found in this list are those defined as 'accepted' in The Biochemical Journal Instructions To Authors (1998). Non standard abbreviations used in the text are defined below.

ADH	Alcohol dehydrogenase
adh	Gene encoding alcohol dehydrogenase
ADH 1503	NAD^+-dependent alcohol dehydrogenase of *Bacillus stearothermophilus* strain NCA 1503
*adh*1503	gene encoding ADH 1503
ADH 2334	NAD^+-dependent alcohol dehydrogenase of *B. stearothermophilus* strain DSM 2334
*adh*2334	gene encoding ADH 2334
ADH 12403	NAD^+-dependent alcohol dehydrogenase of *B. stearothermophilus* strain NCIMB12403
*adh*12403	gene encoding ADH 12403
ADHE	96,000-M_r subunit, multifuntional protein of *Escherichia coli* which includes an ADH
ALDH	Aldehyde dehydrogenase
aldh	Gene encoding aldehyde dehydrogenase
*aldh*1503	Gene encoding aldehyde dehydrogenase in *B. stearothermophilus* strain NCA 1503
ALDH 2334	Aldehyde dehydrogenase of *B. stearothermophilus* strain DSM 2334
*aldh*2334	Gene encoding aldehyde dehydrogenase in *B. stearothermophilus* strain DSM 2334
ATCC	American Type Culture Collection
A_x	Absorbance at x nm
BLAST	Basic Local Alignment Search Tool
CAPS	3-[cyclohexylamino]-1-propanesulfonic acid
dATP	2'-Deoxyadenosine-5'-triphosphate
dCTP	2'-Deoxycytidine-5'-triphosphate
dGTP	2'-Deoxyguanosine-5'-triphosphate
DSM	Deutsche Sammlung von Mikroorganismen
dTTP	2'-Deoxythymidine-5'-triphosphate
E. coli BL21	*E. coli* strain BL21(DE3)
E. coli pLysS	*E. coli* strain BL21(DE3)pLysS

GenBank	DNA database at the National centre for Biotechnology Information , Bethesda, Maryland, USA.
gst	gene encoding glutathione *S*-transferase
His-Tag	C-terminal fusion peptide encoded by plasmid pET21a
HLADH	ADH from horse liver
h.p.l.c.	High-performance liquid chromatography
IPTG	Isopropylthio-β-D-galactoside
k_{cat}	Catalytic rate constant
K_m	Michaelis-Menten constant
lac repressor	lactose operon repressor protein
Mes	4-morpholine-ethanesulfonic acid
$NAD(P)^+$	NAD^+ or $NADP^+$
Na_2 EDTA	Disodium ethylenediaminetetra-acetate
NCA	National Canners Association
NCIMB	National Collection of Industrial and Marine Bacteria, Aberdeen, Scotland
PIR	Protein Identification Resource databank maintained at the National Biomedical Research Foundation, Georgetown, Washington DC
PQQ	pyrrolo-quinoline quinone
PTH	Phenylthiohydantoin
rADH	Recombinant ADH 2334
[^{35}S]-dATP	Deoxyadenosine 5'- (α-[^{35}S] thio) triphosphate, triethylammonium salt
S.D.	Standard Deviation
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide-gel electrophoresis
sample buffer	2% (w/v) SDS, 0.025 M Tris-HCl, pH 6.8, 2.5% (w/v) glycerol, 0.003% (w/v) bromophenol blue, 1% (v/v) 2-mercaptoethanol
Swiss-Prot	Swiss Protein Bank at the University of Geneva, Switzerland.
TBE	89 mM Tris, 89 mM boric acid, 2.5 mM Na_2EDTA, pH 8.3
TBS	10 mM Tris-HCl, 0.9% (w/v) NaCl, pH 7.4
TE	10 mM Tris-HCl, 1 mM Na_2EDTA, pH 8.0
TES-Sucrose	30 mM Tris-HCl, 5 mM Na_2EDTA, 50 mM NaCl, 20% (w/v) sucrose, pH 8
Tricine-SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide-gel electrophoresis with Tricine in the cathode buffer
Tris-HCl	A Tris solution of stated molarity adjusted to the stated pH with HCl

v_o	Initial reaction velocity
V_{max}	Maximum reaction velocity
V8 protease	Endoproteinase Glu-C from *Staphylococcus aureus* strain V8
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactoside
YADH1	Cytoplasmic, fermentative isoenzyme of *Saccharomyces cerevisiae* ADH1
YADH2	Cytoplasmic oxidative isoenzyme of *S. cerevisiae* ADH1
YADH3	Mitochondrial isoenzyme of *S. cerevisiae* ADH1
YADH4	Type 3 (formally iron-dependent) ADH of *S. cerevisiae*
yeast	*Saccharomyces cerevisiae*
ZADH1	*Zymomonas mobilis* zinc-dependent ADH
ZADH2	*Z. mobilis* iron-dependent ADH

Dedicated to Henry and Emma Robinson

TABLE OF CONTENTS

		Page
Declaration		I
Acknowledgements		II
Summary		III
Abbreviations		V
CHAPTER 1: GENERAL INTRODUCTION		1-23
1.1	INTRODUCTION	1
1.2	THE PHYSIOLOGICAL ROLE OF ADH	2
1.3	CLASSIFICATION OF NAD(P)$^+$-DEPENDENT ADHS	4
1.3.1	Short-Chain, Metal-Independent Alcohol Dehydrogenase	5
1.3.2	Medium-Chain, Zinc-Dependent ADHs	5
1.3.2.1	Multiple Classes and Isoenzymes of Medium-Chain, Zinc-Dependent ADH	6
1.3.3	Long-Chain ADH	7
1.3.4	*Z. mobilis* ADH and Related Enzymes: Type 3 ADH	8
1.4	PRIMARY AND HIGHER ORDER STRUCTURE OF MEDIUM-CHAIN, ZINC-DEPENDENT HLADH	9
1.4.1	Amino-Acid Sequence Alignments to HLADH	9
1.4.2	HLADH as the Model for the Tertiary Structure	11
1.4.2.1	Zinc Ions in the Tertiary Structure	12
1.4.2.2	The NAD$^+$-Binding Site	13
1.4.2.3	The Substrate-Binding Pocket	14
1.5	CATALYTIC MECHANISM OF HLADH AND YADH	14
1.5.1	The Ordered Mechanism of Alcohol Oxidation	14
1.5.2	Proton Release and Hydride Transfer	17
1.6	*B. stearothermophilus* ADH	20
1.6.1	ADH from *B. stearothermophilus* NCA 1503	20
1.6.2	ADH from *B. stearothermophilus* DSM 2334 Comparison with ADH 1503	20
1.6.3	Substrate Specificites, Kinetic Properties and Regulation of ADH 1503 and ADH 2334	20

1.7 OBJECTIVES 23

CHAPTER 2: AMINO-ACID SEQUENCING OF ADH 1503 AND ADH 2334 24-33

2.1 MATERIALS AND METHODS 24
2.1.1 Sources of Purified ADH 1503 and Purified ADH 2334 24
2.1.2 Generation of Peptides from ADH 1503 24
2.1.2.1 Denaturation, Reduction and Alkylation 24
2.1.2.2 Cleavage with CNBr 25
2.1.2.3 Cleavage with V8 Protease 25
2.1.3 Reverse-Phase High-Performance Liquid Chromatography 25
2.1.3.1 Separation of Peptides 26
2.1.4 Tricine-Sodium Dodecyl Sulfate-Polyacrylamide-Gel Electrophoresis 27
2.1.5 Electroblotting of Peptides onto Immobilon PVDF 28
2.1.6 Automated Amino-Acid Sequencing 28

2.2 RESULTS 30
2.2.1 Recovery of *S*-Pyridylethylated ADH 1503 30
2.2.2 Cleavage of ADH 1503 with V8 Protease, Reverse-Phase Chromatography of the Resulting Peptides 30
2.2.3 Cleavage of ADH 1503 with CNBr, Reverse-Phase Chromatography of the Resulting Peptides 31
2.2.4 Amino-Acid Sequencing of Peptides of ADH 1503 31
2.2.5 N-Terminal Amino-Acid Sequencing of ADH 1503 and ADH 2334 32

2.3 DISCUSSION 33
2.3.1 N-Terminal Amino-Acid Sequences of ADH 1503 and ADH 2334 33
2.3.2 Amino-Acid Sequences of Peptides of ADH 1503 33

CHAPTER 3: SUBCLONING AND NUCLEOTIDE SEQUENCING OF THE GENE ENCODING ADH 2334 34-68

3.1 INTRODUCTION 34
3.1.1 Cloning of the Gene Encoding ADH 2334 34

3.2	MATERIALS AND METHODS	35
3.2.1	Aseptic Techniques	35
3.2.2	Cloning Host, Plasmids and Bacteriophage	35
3.2.3	Composition of Bacteriological Culture Media	35
3.2.4	Storage of *E. coli* TG1	36
3.2.5	Growth of *E. coli* TG1 in Liquid Culture	36
3.2.6	Storage of Bacteriophage M13	37
3.2.7	Preparation of Infective Stocks of Bacteriophage M13	37
3.2.8	Preparation of Competent *E. coli* TG1	37
3.2.9	Transformation of *E. coli* TG1 with Plasmid DNA	37
3.2.10	α-Complementation Test	38
3.2.11	ADH 2334 Antiserum	38
3.2.12	Colony Immunoblotting	38
3.2.13	Purification of Plasmid DNA	39
3.2.13.1	Preparation of Equilibrated Liquid Phenol	39
3.2.13.2	Small-Scale Purification of Plasmid DNA from *E. coli*	39
3.2.13.3	Large-Scale Purification of Plasmid DNA from *E. coli* TG1	40
3.2.13.4	Recovery of Plasmid DNA from Agarose Gels	41
3.2.14	Purification of Single-Stranded Bacteriophage DNA	42
3.2.15	Estimation of DNA Concentration	42
3.2.16	Agarose Gel-Electrophoresis	43
3.2.17	Digestion of Purified Plasmid DNA with Restriction Enzymes	43
3.2.17.1	Analytical Restriction Digests	43
3.2.17.2	Preparative Restriction Digests	44
3.2.18	Ligation Reactions	44
3.2.19	Manual DNA-Sequencing	44
3.2.19.1	Sequencing Primers	45
3.2.19.2	Denaturation and Recovery of Double-Stranded DNA Templates	45
3.2.19.3	Sequencing Reactions	45
3.2.19.4	Electrophoresis of Sequencing Reaction Mixtures	46
3.2.20	Automated DNA-Sequencing	47
3.2.21	Analyses of Nucleotide and Amino-Acid Sequences	47
3.3	RESULTS	48
3.3.1	Characterisation of Plasmid pADH1b	48
3.3.2	Subcloning *adh*2334 into pUC18	48

3.3.3	Nucleotide Sequencing of the *B. stearothermophilus* Insert in pADH2 and in mp18*adh*	49
3.4	DISCUSSION	51
3.4.1	DNA Sequencing in pADH1b	51
3.4.2	Analysis of the Nucleotide Sequence of the *B. stearothermophilus* insert in pADH2	51
3.4.2.1	Confirmation that pADH2 Contained the Complete *adh*2334 Coding Sequence	51
3.4.2.2	Confirmation that the Noncoding Upstream Nucleotide Sequence in pADH2 was Intact	52
3.4.2.3	Ribosome-binding Sites	53
3.4.2.4	Transcription Promoter Elements	53
3.4.2.5	Transcription Termination Sites	54
3.4.2.6	Translation of the Second Open Reading Frame: Identification of a Downstream Aldehyde Dehydrogenase	55
3.4.3	Genes Encoding Alcohol Dehydrogenase and Aldehyde Dehydrogenase Constitute Part of an Operon	55
3.4.4	Comparison of the Amino-Acid Sequences of *B. stearothermophilus* ADHs	59
3.4.5	Alignment of *B. stearothermophilus* ADHs with HLADH and YADH1	59
3.4.5.1	ADH 2334 is an NAD^+-Dependent, Medium-Chain, Zinc-Dependent ADH	60
3.4.5.2	Amino-Acid Residues Associated with Alcohol Substrate Binding in *B. stearothermophilus* ADHs	60
3.4.5.3	Amino-Acid Residues Associated with Coenzyme -Binding in *B. stearothermophilus* ADHs	63
3.4.6	Identities Between *B. stearothermophilus* ADHs and other Medium-Chain, Zinc-Dependent ADHs	65
3.4.7	Primary Structure of ADH 2334 in the Context of Thermostability	65
3.4.7.1	Amino-Acid Composition	65
3.4.7.2	Specific Substitutions Involving Proline, Glutamic Acid and Lysine	67

CHAPTER 4: THE CONSTRUCTION OF AN EXPRESSION SYSTEM FOR *adh*2334 69-95

4.1	INTRODUCTION	69
4.1.1	Expression of Recombinant Proteins in *E. coli*	69
4.1.1.1	The pET Expression System	69
4.1.1.2	The pTrc Expression System	70
4.1.2	Polymerase Chain Reaction	71
4.1.3	Selection of a Suitable Expression System for *adh*2334	71
4.2	MATERIALS AND METHODS	74
4.2.1	Cloning Hosts, Expression Hosts and Plasmids	74
4.2.2	Working Stocks of Cells, Culture Media and Culture Conditions	74
4.2.3	Storage of *E. coli* as Cryocultures	74
4.2.4	Synthetic Oligonucleotides	75
4.2.5	Polyacrylamide-Gel Electrophoresis of Oligonucleotides	75
4.2.6	Phosphorylating and Annealing Synthetic Oligonucleotides	76
4.2.7	Purification of Plasmid DNA	77
4.2.8	Electrophoresis of Plasmid DNA	77
4.2.9	Recovery of DNA in Centricon Concentrators	78
4.2.10	Polymerase Chain Reaction	78
4.2.11	Restriction of DNA	79
4.2.12	Dephosphorylation of Restricted Plasmid DNA	79
4.2.13	Ligation Reactions	80
4.2.14	Preparation of Competent *E. coli* Cells	80
4.2.14.1	Preparation of Competent *E. coli* Cells by the Method of Morrison	81
4.2.14.2	Preparation of Competent *E. coli* Cells by the Method of Hanahan	81
4.2.15	Transformation of *E. coli* cells	82
4.2.16	Screening *E. coli* Transformants for IPTG-Dependent Expression	82
4.2.17.	Polyacrylamide-Gel Electrophoresis	82
4.2.18	N-Terminal Amino-Acid Sequence Analysis of Recombinant Protein	84
4.3	RESULTS	85
4.3.1	Restriction Analyses of Plasmids pADH2 and pET21a and	

	Isolation of Relevant Restriction Fragments	85
4.3.2	Monitoring Construction of the Double-Stranded *Nde*I-*Hpa*I Adapter	85
4.3.3	Ligation of the *Nde*I-*Hpa*I Adapter to the 1-kb*Hpa*I-*Sac*I *adh* Fragment, Cloning of the Construct into *Nde*I-*Sac*I pET21a	85
4.3.4	Ligation of the 1-kb *Hpa*I-*Sac*I *adh* Fragment to Linear *Nde*I-*Sac*I pET21a, Formation of a 6.4-kb *adh*-pET21a Linear Construct	86
4.3.5	Ligation of the *Nde*I-*Hpa*I Adapter to the 6.4-kb *adh*-pET21a Linear Construct	87
4.3.6	Insertion of *Nde*I-*Hpa*I Adapter into pADH2, Construction of pADH2a	87
4.3.7	Cloning of the 1-kb *Nde*I-*Sac*I Fragment of pADH2a into pET21a	88
4.3.8	Subcloning of an *adh* PCR Product into pET21d and into pTrc99A	88
4.3.9	IPTG-Dependent Expression in pETADH and in pTrcADH Clones	89
4.3.10	Confirmation that the Product of Expression was Recombinant ADH 2334	90
4.4	DISCUSSION	91
4.4.1	Construction of an *adh* Expression System in pET21a	91
4.4.2	Construction of an Expression System Containing an *adh* PCR Product	93
CHAPTER 5: PURIFICATION AND CHARACTERISATION OF RECOMBINANT ADH 2334		96-133
5.1	INTRODUCTION	96
5.2	MATERIALS AND METHODS	98
5.2.1	Source of Recombinant ADH	98
5.2.2	Culture Medium, Culture Conditions and Estimation of Cell Number	99
5.2.3	Potassium Phosphate Buffers	99
5.2.4	Protease Inhibitors, Dithiothreitol and DNase I	99
5.2.5	*E. coli* Lysis-Buffers	99

5.2.6	Saturated Ammonium Sulfate Solution	100
5.2.7	Lysis of *E. coli* Cells	100
5.2.8	Ammonium Sulfate Fractionation	100
5.2.9	Purification of Recombinant ADH	101
5.2.10	Measurement of Protein Concentration	102
5.2.11	Polyacrylamide-Gel Electrophoresis	102
5.2.12	Determination of ADH Enzyme Activity by Spectrophotometric Assay	102
5.2.13	Determination of ADH Enzyme Activity by Fluorimetric Assay	103
5.2.14	Calibration Curves for Measurement of NADH Concentration	104
5.2.15	Calculation of Specific Activity	105
5.2.16	Determination of Kinetic Parameters	105
5.2.17	Measurement of Zinc Ion Concentration	105
5.3	RESULTS	107
5.3.1	Optimum Conditions for IPTG-Dependent Expression	107
5.3.2	Ammonium Sulfate Fractionation of a Cell Lysate and Specific Activity of Recombinant ADH in the Fractions	108
5.3.3	Addition of Zinc Ions to Culture Medium Resulted in Increased Specific Activity of Recombinant ADH	109
5.3.4	Depletion of Zinc Ions from Culture Medium During Expression	111
5.3.5	Ammmonium Sulfate Fractionation of a Cell Lysate Prepared from Cells Grown in the Presence of Zinc Ions	112
5.3.6	Purification of Recombinant ADH	113
5.3.7	The Zinc Content of Purifed Recombinant ADH	114
5.3.8	Determination of the Kinetic Parameters of Recombinant ADH Using a Fluorimetric Assay	115
5.3.9	Addition of Zinc Ions to Purified -Zn Recombinant ADH	116
5.3.10	The Affect of Temperature on Recombinant ADH	116
5.4	DISCUSSION	118
5.4.1	Expression from pETADH in the Absence and Presence of Supplementary Zinc Ions, Specific Activity of the Enzymes Produced	118
5.4.2	Heterogeneity in the Recombinant Protein Produced in the Absence of Supplementary Zinc Ions	119

5.4.3	The Zinc Content of Recombinant ADH Produced in the Absence and Presence of Supplementary Zinc Ions	121
5.4.4	Extracellular and Intracellular Concentrations of Zinc Ions During Expression from pETADH	122
5.4.5	Addition of Zinc Ions to Purified -Zn Recombinant ADH	124
5.4.6	Zinc Limitation May Be a Feature of High-Level Expression of Recombinant Zinc-Dependent Enzymes	126
5.4.7	Kinetic Parameters of Recombinant ADH Produced in the Absence and Presence of Supplementary Zinc Ions	129
5.4.8	Reversible Cold-Inactivation of rADH	130

FURTHER WORK 134

REFERENCES 135-153

CHAPTER 1
GENERAL INTRODUCTION

1.1 INTRODUCTION

Bacteria of the genus *Bacillus* are endospore-forming, Gram-positive rods, capable of growth under aerobic conditions. The taxonomy of the genus has been reviewed by Gordon (1981). The species *Bacillus stearothermophilus* was first described by Donk (1920) and is an obligate thermophile (will not grow at 37 °C) associated with compost, manure and the spoiling of canned food. Optimum growth temperature is 55 °C - 60 °C and maximum growth temperature is about 65 °C. The taxonomy and the ecology of the species are reviewed by Wolf & Sharp (1981). *B. stearothermophilus* is used as a source of large amounts of easily purified stable enzymes for basic research (Atkinson *et al.*, 1979; Schläpfer & Zuber, 1992; Zhu *et al.*, 1995; Witzmann & Bisswanger, 1998; Roitel *et al.*, 1999; Howard *et al.*, 2000) and for the biotechnology industry (Ljungdahl, 1979; Ng & Kenealy, 1986; Vieille *et al.*, 1996).

Two strains of *B. stearothermophilus* are relevant to this thesis. The two strains will be referred to as *B. stearothermophilus* NCA 1503 and *B. stearothermophilus* DSM 2334. *B. stearothermophilus* NCA 1503 was originally isolated from canned peas by Cameron & Esty (1926) at the National Canners Association (NCA), Washington DC. The organism was characterised by Gordon & Smith (1949) and is a facultative anaerobe. Other reference numbers for this strain are: ATCC 7954 in the American Type Culture Collection, Rockville, USA; N.C.I.B. 8924 in the National Collection of Industrial Bacteria, Aberdeen, Scotland; DSM 2027 in the Deutsche Sammlung von Mikroorganismen, Gottingen, Germany. *B. stearothermophilus* NCA 1503 is a popular source of the enzymes of central metabolism including: pyruvate dehydrogenase (Henderson *et al.*, 1979); lactate dehydrogenase (Clarke *et al.*, 1985; Barstow *et al.*, 1986; Cameron *et al.*, 1994) and 3-phosphoglycerate kinase (Davies *et al.*, 1992).

B. stearothermophilus var. *non diastaticus* (*B. stearothermophilus* DSM 2334) was isolated (from a manure pile) and characterised by Epstein & Grossowicz (1969). *B. stearothermophilus* DSM 2334 is a prototroph with simple nutritional requirements and differs from *B. stearothermophilus* NCA 1503 in that it is an obligate aerobe and is unable to hydrolyse starch. An inability to hydrolyse starch is uncommon among *B. stearothermophilus* strains (Wolf & Sharp, 1981). *B. stearothermophilus* DSM 2334, unlike the NCA 1503 strain, does not have an extensive literature. Apart from studies of ADH (Section 1.6), the literature concerning the biochemistry of *B. stearothermophilus* DSM 2334 is limited to a study of glutamate dehydrogenase (Epstein & Grossowicz, 1975), a study of respiration linked proton translocation (Chicken *et al.*, 1981), a study on the uptake and metabolism of acetate (Mallinder & Moir, 1991) and the cloning and nucleotide

sequencing of the genes encoding glycerol dehydrogenase (Mallinder *et al.*, 1992) and fumarase (Reaney *et al.*, 1993).

The enzyme alcohol dehydrogenase (ADH) catalyses the interconversion of alcohols, aldehydes and ketones. Substrates include primary and secondary aliphatic alcohols, aromatic alcohols, and the corresponding aldehydes and ketones. The known electron acceptors for ADH are nicotinamide adenine dinucleotide (NAD^+), nicotinamide adenine dinucleotide phosphate ($NADP^+$) and pyrrolo-quinoline quinone (PQQ). Known metal ion cofactors for ADH are Zn^{2+}, Fe^{2+} and Ca^{2+}.

NAD^+-dependent ADH (alcohol : NAD oxidoreductase; EC 1.1.1.1) is widely distributed and occurs in archaebacteria, eubacteria and eukaryotes. The interconversion of ethanol and acetaldehyde, catalysed by NAD^+-dependent ADH, is shown in Equation 1:

$$\mathbf{CH_3CH_2OH + NAD^+ \rightleftharpoons CH_3CHO + NADH + H^+} \qquad \textbf{(1)}$$

The forward reaction of Equation 1 occurs in mammalian liver and ADH catalysed oxidation of ethanol is the first step in the metabolism of beverage alcohol in humans. The reverse reaction, the last step in anaerobic glycolysis in *Saccharomyces cerevisiae* (yeast), is a key reaction in the commercial production of beverage alcohol. Consequently, mammalian liver ADH and yeast ADH (both are zinc-dependent) are among the most intensely studied enzymes (reviews by Sund & Theorell, 1963; Brändén *et al.*, 1975; Klinman, 1981; Eklund & Brändén, 1983; Pettersson, 1987).

With $NADP^+$-dependent ADH (alcohol : NADP oxidoreductase; EC 1.1.1.2), the catalysed reaction is overwhelmingly in favour of alcohol formation, and the mammalian enzyme, but not the prokaryotic enzyme, is usually described as aldehyde reductase (review by Flynn, 1982). Aldehyde reductase from kidney (Bosron & Prairie, 1972), brain (Tabakoff & Erwin, 1970) and liver (Udovikova & Wojtczak, 1998; O'Connor *et al.*, 1999) are among the mammalian aldehyde reductases that have been studied. Prokaryotic metal-dependent $NADP^+$-dependent ADHs and the metal-dependent NAD^+-dependent ADHs are often considered together as $NAD(P)^+$-dependent ADH.

1.2 THE PHYSIOLOGICAL ROLE OF ADH

ADHs are involved in an astonishingly wide range of metabolism. However, few ADHs are equally reactive towards, for example, primary and

secondary alcohols. Many organisms contain multiple ADHs (discussed in Section 1.3.2.1) and their physiological roles can sometimes prove difficult to disentangle.

In anaerobic bacteria and fungi, the reduction of aldehyde by NAD(P)H-dependent ADH results in the production of alcohol and oxidised coenzyme. The importance of the reaction to the organism lies in the fact that the regeneration of oxidised coenzyme is essential for other oxidative and energy-yielding metabolic processes to continue (Clark, 1989). Most organisms produce ethanol only, but 1-butanol and 2-propanol are produced by *Clostridium* (Ismaiel *et al.*, 1993). The oxidation of alcohols by $NAD(P)^+$-dependent ADH provides a mechanism whereby microorganisms can use alcohols as carbon and energy sources.

The study of mammalian ADH has focussed on the role of the enzyme in ethanol metabolism, particularly as it pertains to human alcohol abuse (Day *et al.*, 1991; Moreno & Parés, 1991; Han *et al.*, 1998). However, Krebs & Perkins (1970) demonstrated that ethanol is present in the gastro-intestinal tract of rats not exposed to alcohol in their diet. The ethanol was absorbed into the blood system and quantitatively removed by the liver. Furthermore, the concentration of ethanol in the blood (portal vein) was dramatically reduced when antibiotics were included in the diet, indicating that the source of the ethanol was microbial action in the gut. Krebs & Perkins (1970) present a fascinating list of the circumstances under which mammals and birds, in their normal living conditions, can be exposed to large amounts of alcohol. Most of these circumstances are related to the consumption of nectar and decaying fruit. It is generally accepted (Benach *et al.*, 1998) that the function of *Drosophila* ADH is detoxification and metabolic assimilation of alcohols that *Drosophila* encounter in their food source.

Mammalian liver ADHs have a very broad substrate specificity, and maximum activity is not observed with ethanol (Wagner *et al.*, 1983). The role of ADH in retinol (vitamin A) metabolism for the production of the transcriptional regulator retinoic acid has attracted considerable interest (Boleda *et al.*, 1993; Kedishvili *et al.*, 1997; Napoli, 1999). Mammalian ADH also functions in the metabolism of hydroxysteroids (McEvily *et al.*, 1988; Kedishvili *et al.*, 1997), in the metabolism of aldehydes produced by lipid peroxidation (Sellin *et al.*, 1991; Boleda *et al.*, 1993) and in the metabolism of compounds derived from dopamine and norepinephrine degradation (Mårdh & Vallee, 1986; Mårdh *et al.*, 1986).

PQQ-containing alcohol dehydrogenase (EC 1.1.99.8; Inoue *et al.*, 1989; Ameyama *et al.*, 1991; Anthony, 1996; Goodwin & Anthony, 1998; Matsushita *et al.*, 1999) occurs in the periplasm of Gram-negative bacteria capable of growth on ethanol, methanol or methane as the sole carbon source. In these organisms, ADH plays a key role in the production of both ATP and precursors for carbon assimilation (Groen *et al.*, 1984; Schrover *et al.*, 1993; Takemura *et al.*, 1993). Electrons

resulting from the oxidation of alcohol to aldehyde are passed *via* PQQ to cytochromes or ubiquinone and hence to electron transport chains. Re-oxidation of PQQ is accompanied by the release of protons from PQQ into the periplasim.

Methylophilus methylotrophus, which grows on methanol, has been considered for the commercial production of single cell protein, a potential animal-feed protein supplement (Anthony, 1982). Electrons from the PQQ-dependent oxidation of methanol in *M. methylotrophus* enter the electron transport chain at cytochrome *c* (Cox *et al.*, 1992), and therefore 1 mol of ATP is produced for every mol of methanol oxidised. Electrons from the NAD^+-dependent oxidation of methanol, which enter the electron transport chain at an earlier point, yield three mol of ATP for every mol of methanol oxidised. There should therefore be an energetic advantage, manifest in terms of cell yield, to the production of single cell protein from *M. methylotrophus* which was genetically modified to express an NAD^+-dependent methanol dehydrogenase.

1.3 CLASSIFICATION OF NAD(P)$^+$-DEPENDENT ADHS

Differences between the primary structures and subunit sizes of horse liver ADH [HLADH; Jörnvall, 1970(a)] and *Drosophila melanogaster* ADH (Thatcher, 1980) formed the basis for the original division of ADHs, by Jörnvall *et al.* (1981), into two evolutionary lines. The two classes were originally described as long and short ADHs, according to the relative sizes of the HLADH (long) and *D. melanogaster* ADH (short) subunit. The defining features of a long ADH were a subunit M_r of approximately 40,000, a requirement for Zn^{2+}, and sequence homology to HLADH in the neighbourhood of Cys-46 (a ligand to Zn^{2+}). The defining features of a short ADH were a subunit M_r of approximately 25,000, no requirement for metal ions, and no obvious amino-acid sequence homology to HLADH. Sorbitol dehydrogenase was found to have sequence homology to HLADH and a requirement for Zn^{2+} and was classified as a long ADH. Ribitol dehydrogenase was found to have sequence homology to *D. melanogaster* ADH and no requirement for Zn^{2+} and was classified as a short ADH.

The classification based on subunit size and amino-acid sequence homology was revised, following the discovery of an *Escherichia coli* gene encoding an ADH that displayed no sequence homology to either the *Drosophila* or the horse liver enzyme (Goodlove *et al.*, 1989). The *E. coli* ADH was part of a 96,000-M_r subunit, multifunctional protein (ADHE). On the basis of primary structure and subunit size, ADHs are classified (Persson *et al.*, 1991) as short-chain (*Drosophila* ADH-like), medium-chain (horse liver ADH-like) and long-chain (*E. coli* ADH-like).

The isolation of an iron-dependent ADH (subunit M_r~40,000) from *Zymomonas mobilis* (Neale *et al.*, 1986), the discovery that ADHE was iron-dependent (Kessler *et al.*, 1991) and the fact that the C-terminal region of long-chain ADHs show approximately 30% homology to *Z. mobilis* iron-dependent ADH (de Vries *et al.*, 1992; Rosenthal *et al.*, 1997) have given rise to an alternative classification system that is based on sequence homology to horse liver ADH, *D. melanogaster* ADH or *Z. mobilis* iron-dependent ADH, the three classes being labeled Type (or Group) 1, 2 and 3 respectively. There are, therefore, three classification schemes currently in use in the literature. The relationships between the schemes are shown in Table 1.3.1.

1.3.1 Short-Chain, Metal-Independent Alcohol Dehydrogenase

Short-chain alcohol dehydrogenases are distributed among *Drosophila* species, prokaryotes and mammals and are part of an extended superfamily of short-chain alcohol / polyol / sugar dehydrogenases [Jörnvall *et al.*, 1984(a); Persson *et al.*, 1991; Table 1.3.2]. *Drosophila* ADH catalyses the oxidation of secondary alcohols (Sofer & Ursprung, 1968; Winberg *et al.*, 1986) and the dismutation of aldehyde (Henehan *et al.*, 1995; Winberg & McKinley-McKee, 1998; Equation 2).

$$2CH_3CHO + H_2O \rightleftharpoons CH_3COOH + CH_3CH_2OH \qquad (2)$$

Detailed analyses of the defining characteristics of the short-chain alcohol / polyol / sugar dehydrogenases superfamily are given by Persson *et al.* (1991) and by Reid & Fewson (1994). There is no evidence for a metal requirement in any of the enzymes that have been characterised.

The three-dimensional structure has been solved for the *D. lebanonensis* enzyme (Gordon *et al.*, 1992; Benach *et al.*, 1998) and also for several other members of the extended superfamily (discussed in Benach *et al.*, 1998). Within the superfamily there are NAD^+-dependent enzymes, $NADP^+$-dependent enzymes and enzymes that can use both NAD^+ and $NADP^+$ (Table 1.3.2).

1.3.2 Medium-Chain, Zinc-Dependent ADHs

Medium-chain, zinc-dependent ADHs belong to a superfamily of medium-chain dehydrogenases / reductases defined by Persson *et al.* (1994) and Edwards *et al.* (1996). A representative selection of members of the superfamily and relevant references are presented in Table 1.3.3. ζ-crystallin is an eye lens structural protein with quinone oxidoreductase activity (Rao *et al.*, 1992) and high sequence identity to *E. coli* quinone oxidoreductase (Edwards *et al.*, 1996).

			Classification System		
Model Enzyme	**Metal cofactor**	**Approximate subunit M_r**	**Persson *et al.* (1991)**	**de Vries *et al.* (1992) Rosenthal *et al.* (1997) Antoine *et al.* (1999)**	**Reid & Fewson (1994)**
Horse liver ADH	Zn^{2+}	40,000	Medium-chain	Type 1	Group 1
D. melanogaster ADH	none	25,000	Short-chain	Type 2	Group 2
Z. mobilis iron-dependent ADH	Fe^{2+}	40,000	not included	Type 3	Group 3
E. coli ADHE	Fe^{2+}	96,000	Long-chain	Type 3	Group 3

Table 1.3.1 Classification Systems Used for Alcohol Dehydrogenases

Enzyme	Coenzyme	Reference(s)
D. melanogaster ADH	NAD^+	Thatcher (1980)
D. lebanonensis ADH	NAD^+	Villaroya *et al.* (1989)
Klebsiella aerogenes ribitol dehydrogenase EC 1.1.1.56	NAD^+	Dothie *et al.* (1985) Loviny *et al.* (1985)
E. coli enoyl reductase EC 1.3.1.9 EC 1.3.1.10	NAD^+ and $NADP^+$	Bergler *et al.* (1996 Baldock *et al.* (1998)
Pseudomonas testosteroni 3β-hydroxysteroid dehydrogenase EC 1.1.1.51	NAD^+ and $NADP^+$	Maser *et al.* (1992)
Rat liver D-β-hydroxybutyrate dehydrogenase EC 1.1.1.30	NAD^+	Churchill *et al.* (1992)
B. megaterium glucose dehydrogenase EC 1.1.1.47	NAD^+ and $NADP^+$	Jany *et al.* (1984)
Human placental 15-hydroxyprostaglandin dehydrogenase EC 1.1.1.141	NAD^+	Krook *et al.* (1990)

Table 1.3.2 **Examples of Short-Chain Alcohol / Polyol / Sugar Dehydrogenases**

Protein	Substrate	Subunit : Zn^{2+} stoichiometry	Coenzyme	Reference(s)
Horse liver ADH (HLADH) EC 1.1.1.1)	Ethanol	2 : 1	NAD^+	Pettersson (1987)
E. coli quinone oxidoreductase (EC 1.6.99.2)	Quinone	No Zn^{2+}	NADPH and NADH	Thorn *et al.* (1995)
Thermoplasma acidophilum glucose dehydrogenase (EC 1.1.1.47)	Glucose	Contains the ligands for two Zn^{2+}	NAD^+ and $NADP^+$	Bright *et al.* (1993) John *et al.* (1994)
ζ–crystallin from guinea pig eye lens	Quinone	No Zn^{2+}	NADPH	Borrás *et al.* (1989) Rodokanaki *et al.* (1989) Rao *et al.* (1992) Persson *et al.* (1994)
B. subtilis sorbitol dehydrogenase (EC 1.1.1.14)	Secondary alcohols only	1 : 1	NAD^+	Ng *et al.* (1992)
E. coli L-threonine dehydrogenase (EC 1.1.1.103)	Threonine	Uncertain, at least one Zn^{2+} per subunit	NAD^+	Aronson *et al.* (1989) Johnson *et al.* (1998)

Table 1.3.3 **The Superfamily of Medium-Chain Dehydrogenases / Reductases**

Comparison of the three-dimensional crystal structures of *E. coli* quinone oxidoreductase, *Thermoplasma acidophilum* glucose dehydrogenase and HLADH (Edwards *et al.,* 1996) confirms that members of the group possess overall structural homology in both the nucleotide-binding domain and the catalytic domain, despite low sequence identity, dissimilar cellular functions, dissimilar substrate specificities and dissimilar requirements for Zn^{2+} (Table 1.3.3). The dehydrogenases require at least one Zn^{2+} per subunit, the reductases do not necessarily require Zn^{2+}.

In vertebrates, medium-chain, zinc-dependent ADH is cytosolic, dimeric (subunit M_r ~ 40,000) and NAD^+-dependent (Duester *et al.*, 1999). Tetrameric ADH occurs in fungi and in bacteria (Reid & Fewson, 1994). However, thermostable NAD^+-dependent ADH from the archaebacterium *Sulfolobus solfataricus* is dimeric (Ammendola *et al.*, 1992). The majority of zinc-dependent ADHs contain two Zn^{2+} per subunit; one of these (the catalytic-zinc ion) is essential for catalytic activity (Brändén *et al.*, 1975, Sakoda & Imanaka, 1992), the other (the second-zinc ion) has been assigned a structural role [Drum *et al.*, 1969(a); Jeloková *et al.*, 1994].

1.3.2.1 Multiple Classes and Isoenzymes of Medium-Chain, Zinc-Dependent ADH

The labyrinthine subject of the nomenclature of the classes of medium-chain, zinc-dependent ADH, the isoenzymes within classes, and the genes encoding them, has recently been rationalised by Duester *et al.* (1999). Species-specific nomenclatures, Roman numerals and Greek symbols are replaced by a system where the enzyme alcohol dehydrogenase is referred to by the symbol ADH, Arabic numbers refer to class, and Arabic capital letters refer to the isoenzyme. The classical, ethanol-active, NAD^+-dependent, pyrazole-sensitive ADHs of horse liver and yeast are classed as ADH1 (previously class I). Glutathione formylating NAD^+-dependent formaldehyde dehydrogenase (EC 1.2.1.1) is classed as ADH3 (previously class III). ADH3, with the reduced form of glutathione (GSH), functions in the oxidative elimination of formaldehyde (Equation 3) and has no catalytic activity in the presence of ethanol (Koivusalo *et al.*, 1989; Holmquist & Vallee, 1991).

$$\mathbf{HCHO + GSH + NAD^+ \rightleftharpoons \textit{S}\text{-formylglutathione + NADH + H^+}} \qquad \mathbf{(3)}$$

A common origin has been established for mammalian ADH1 and ADH3 (Cederlund *et al.*, 1991) with ADH3 closer to the original zinc-dependent ancestor and ADH1 emerging from ADH3 after duplication (Kaiser *et al.*, 1989; Danielsson & Jörnvall, 1992; Kaiser *et al.*, 1993). Ethanol dehydrogenase activity is considered to have had multiple origins and to have evolved repeatedly (Danielsson & Jörnvall 1992).

Seven distinct classes of vertebrate medium-chain, zinc-dependent ADH have been identified (Duester *et al.*, 1999). Distinctions between the classes are made on the basis of electrophoretic mobility in starch gels, pI values, substrate specificity, kinetic properties and inhibition by pyrazole (Wagner *et al.*, 1983; Bosron & Li, 1987; Julià *et al.*, 1987; Moreno & Parés, 1991; Kedishvili *et al.*, 1995). Classes within the same organism share less than 70% amino-acid sequence identity (Danielsson & Jörnvall, 1992). Within a class, multiple isoenzymes, sharing greater than 80% sequence identity, can occur and any one tissue may contain several isoenzymes of the same class (Danielsson & Jörnvall, 1992; Estonius *et al.*, 1996; Svensson *et al.*, 1998). No single species that encodes all seven classes is known. Five distinct classes of human ADH have been described. The distribution of the five human classes across tissues vary, and the expression of some classes is linked to foetal development (Estonius *et al.*, 1996).

Three isoenzymes of HLADH have been identified [Jörnvall, 1970(b)]. The homo dimers EE and SS, and the hetero dimer ES, occur due to the expression of two separate genes. The S subunit is responsible for catalytic activity of HLADH towards 3 β-hydroxysteroids. The E subunit and the S subunit catalyse the oxidation of ethanol (Park & Plapp, 1992). The amino-acid sequence of the S subunit differs from the E subunit at ten positions, nine substitutions and one deletion (Park and Plapp, 1991). Horse liver also contains glutathione-dependent ADH3 (Kaiser *et al.*, 1989).

Three isoenzymes of yeast ADH1 are known (YADH1, YADH2 and YADH3). The three isoenzymes have different substrate specificities and are differentially expressed, at least partly in relation to the requirement for fermentation or oxidation of alcohols. YADH1 is the classic cytoplasmic fermentative enzyme expressed during anaerobic growth in the presence of glucose. YADH2, also cytoplasmic, is expressed during aerobic growth, is catabolite repressed by glucose and works preferentially in the direction of ethanol oxidation (Wills, 1976; Wills & Jörnvall, 1979). The YADH1 and YADH2 subunit amino-acid sequences differ at 23 positions (Russell *et al.*, 1983). YADH3 is located in the mitochondrion (Ciriacy 1975) and has 79% amino-acid sequence identity with YADH1 (Young & Pilgrim, 1985). The physiological function of this isoenzyme is unknown. Glutathione-dependent ADH3 also occurs in yeast (Wehner *et al.*, 1993).

1.3.3 Long-Chain ADH

Long-chain ADH is a multifunctional protein, containing the amino-acid sequence of an acetaldehyde-CoA dehydrogenase (EC 1.2.1.10) at the N-terminal end and the amino-acid sequence of a Type 3 ADH (Section 1.3.4) at the C-terminal end. The protein is the product of a single open reading frame which encodes both enzyme

activities (Kessler *et al.*, 1992; Bruchhaus & Tannich, 1994) and functions during the anaerobic metabolism of glucose where pyruvate is converted, *via* pyruvate formate lyase (EC 2.3.1.54), to acetyl-CoA plus formate (Knappe & Sawers, 1990). Acetaldehyde-CoA dehydrogenase converts acetyl-CoA to acetaldehyde, ADH then converts acetaldehyde to ethanol (Clark, 1989). Long-chain ADH has been isolated from eubacteria and fermentative protozoa (Table 1.3.4). The abbreviations used in Table 1.3.4 are those assigned to the enzymes in the original references. *E. coli* ADHE is a polymeric (up to 40 subunits) rod shaped protein. As well as exhibiting acetaldehyde-CoA dehydrogenase and ADH activity, each multifunctional subunit (891 amino acids) functions as pyruvate formate lyase deactivase (Kessler *et al.*, 1992). All three functions are iron-dependent (Kessler *et al.*, 1991). The *Entamoeba histolytica* and *Clostridium acetobutylicum* enzymes are each 872 amino acids per subunit and there is no evidence of pyruvate formate lyase deactivase activity in these enzymes (Bruchhaus & Tannich, 1994).

1.3.4 *Z. mobilis* ADH and Related Enzymes: Type 3 ADH

Z. mobilis is an obligately ethanologenic bacterium that contains a conventional medium-chain, zinc-dependent ADH (Keshav *et al.*, 1990; ZADH1). A second, iron-dependent ADH (ZADH2) was purified from *Z. mobilis* by Neale *et al.* (1986). ZADH2 has a subunit M_r~40,000 but displays no obvious amino-acid sequence homology to either HLADH or the *D. melanogaster* enzyme (Conway *et al.*, 1987). Homology to the ZADH2 amino-acid sequence is the criterion used to classify ADHs as Type 3 ADH.

All Type 3 ADHs identified to date are from microbial sources, typically obligate anaerobes. Some examples are presented in Table 1.3.5 where the abbreviations used are those assigned to the enzymes in the original references. The subunit M_r ranges from 41,000 - 48,000, both NADH- and NADPH-dependent enzymes occur and the metal contents of the purified enzymes vary. Glycerol dehydrogenase of *B. stearothermophilus* DSM 2334 (Mallinder *et al.*, 1992) shows sequence homology (25% identity) to the *C. acetobutylicum* ADH sequenced by Youngleson *et al.* (1989). The study of the Type 3 ADHs of sulfur reducing, anaerobic archaebacteria from the genera *Thermococcus* and *Pyrococcus* is currently an active area of research (Table 1.3.5). The *T. hydrothermalis* enzyme has a very wide substrate range including glycerol and monoterpenes and has been reported to have a high activity towards methanol. However, methanol dehydrogenase activity is absent from the other *Thermococcus* enzymes (Antoine *et al.*, 1999).

ZADH2, when purified in the presence of Fe^{2+}, contains one mole of Fe^{2+} and only trace amounts of Zn^{2+} per subunit (Neale *et al.*, 1986). The enzyme is inactivated by oxygen and by Zn^{2+} (Tamarit *et al.*, 1997). The metal ion content of

Enzyme	Coenzymes	Metal Cofactor	Reference
E. coli ADHE	NADH	Fe^{2+}	Goodlove *et al.* (1989) Kessler *et al.* (1991) Kessler *et al.* (1992)
Clostridium acetobutylicum (strain ATCC 824) AAD	NADH	Not reported	Nair *et al.* (1994)
Giardia lamblia ADHE	Not reported	Not reported	Rosenthal *et al.* (1997)
Entamoeba histolytica EhADH2	NADH	Not reported	Bruchhaus & Tannich (1994)

Table 1.3.4 **Examples of Long-Chain, Acetaldehyde Dehydrogenase / Alcohol Dehydrogenases**

Enzyme	Metal ion content	Reference to protein sequence
Z. mobilis ADH2	Fe^{2+} Neale *et al.* (1986)	Conway *et al.* (1987)
Yeast ADH4	Not reported	Williamson & Paquin (1987)
E. coli 1,2-propanediol oxidoreductase	Fe^{2+} and Zn^{2+} Obradors *et al.* (1998)	Conway & Ingram (1989)
C. acetobutylicum ADH1	Not reported	Youngleson *et al.* (1989)
C. acetobutylicum butanol dehydrogenases (2 isoenzymes)	Not reported	Walter *et al.* (1992)
B. methanolicus methanol dehydrogenase	Zn^{2+} and Mg^{2+} Vonck *et al.* (1991)	Vonck *et al.* (1991) de Vries *et al.* (1992)
Thermococcus litoralis	Fe^{2+} Ma *et al.* (1994)	Ma *et al.* (1994) Partial sequence only
Thermococcus strain ES-1 ADH	Fe^{2+} Ma *et al.* (1995)	Ma *et al.* (1995) Partial sequence only
T. zilligii (previously Thermococcus strain AN1)	Not reported	Li & Stevenson (1997)
T. hydrothermalis ADH	Fe^{2+}	Antoine *et al.* (1999)
Pyrococcus furiosus ADH	Fe^{2+} and Zn^{2+} Ma & Adams (1999)	Ma & Adams (1999) Partial sequence only

Table 1.3.5 **Type 3 Alcohol Dehydrogenases**

purified *E. coli* 1,2-propanediol oxidoreductase has been determined, by Obradors *et al.* (1998), to be ~ 0.5 atom of Fe^{2+} and 0.5 atom of Zn^{2+} per subunit. However, Sridhara *et al.* (1969) demonstrated that the enzyme was inactivated by Zn^{2+}. Drewke & Ciriacy (1988) reported that yeast ADH4 was activated by the addition of zinc ions *in vitro*, and not by iron. The metal content of the enzyme was not determined and there appear to be no subsequent reports about this enzyme.

An alignment of the enzymes for which complete sequences have been determined is discussed by Li & Stevenson (1997). A fifteen amino acid motif, containing three highly conserved His residues (Walter *et al.*, 1992; Cabiscol *et al.*, 1994; Li & Stevenson, 1997) has been confirmed, by site-directed mutagenesis, to be the iron-binding site in *E. coli* 1,2-propanediol oxidoreductase (Obradors *et al.*, 1998). There is no evidence for the presence in Type 3 ADHs of a second metal binding site, analogous to either of the zinc-binding sites in medium-chain ADHs. Tararit *et al.* (1997) have shown that Zn^{2+} replaces the Fe^{2+} in ZADH2 *in vitro* and one must conclude that the presence of Zn^{2+} in preparations of the enzymes listed in Table 1.3.5 is due to Zn^{2+} binding, *in vitro*, to the iron-binding site.

1.4 PRIMARY AND HIGHER ORDER STRUCTURE OF MEDIUM-CHAIN, ZINC-DEPENDENT HLADH

1.4.1 Amino-Acid Sequence Alignments to HLADH

The HLADH E subunit was the first ADH for which the complete protein primary structure was determined [Jörnvall, 1970(a)]. Alignments of medium-chain, zinc-dependent ADHs are usually made with reference to the HLADH E amino-acid sequence and have been influenced by the alignment of HLADH and ADH from an uncharacterised commercially available yeast suggested by Jörnvall *et al.* (1978). An ement of HLADH and the amino-acid sequence deduced from the nucleotide sequence of the gene encoding YADH1 (Bennetzen & Hall, 1982) is presented in Figure 1.4.1(a) and is based on the alignment of Jörnvall *et al.* (1978). When discussing alignments, and according to convention, numbering refers to the HLADH E subunit sequence.

YADH1 and HLADH are distantly related, displaying only 25% residue identity (Jörnvall *et al.*, 1978). The gap in the yeast sequence, equivalent to residues 119 - 139 of HLADH, was introduced into the original sequence alignment by Jörnvall *et al.* (1978) to preserve the alignment of functionally important residues. In HLADH residues 119 - 139 form a surface loop-structure (Jörnvall *et al.*, 1978). A scheme in which two evolutionary lines, indicated by the presence or absence of the segment, diverge from a common medium-chain ADH ancestor has been proposed by Kaiser *et al.* (1993). The presence of the segment has been associated with a dimeric

Figure 1.4.1(a) **The Primary structure of HLADH and of YADH1**

The amino-acid sequence of the E subunit of HLADH is that determined by Jörnvall [1970(a)]. The amino-acid sequence of YADH1 is that deduced from the nucleotide sequence determined by Bennetzen & Hall (1982). The amino-acid sequences were aligned manually, according to the criteria of Jörnvall *et al.* (1978). The twenty-two residues identifed by Jörnvall *et al.* (1987) as being highly conserved between ADHs are indicated by an asterisk under the YADH1 sequence. The nine residues identified by Sun & Plapp (1992) as being strictly conserved between medium-chain ADHs are indicated in green. Residues that form ligands to the catalytic-zinc ion are coloured red, ligands to the second-zinc ion are coloured blue. The second-sphere ligands to the catalytic-zinc ion are coloured magenta. Residues coloured cyan are Asp-223, responsible for discrimination against $NADP^+$ binding and the Ser and His residues involved in the proton-release mechanism.

Figure 1.4.1(b) **The Tertiary Structure of the HLADH E Subunit**

The conformation is that adopted by the E subunit complexed with Zn^{2+} (black spheres), NAD^+ and dimethyl sulfoxide (the closed conformation; 2.9 Å resolution; Eklund *et al.*, 1981). The file containing the crystal structure of HLADH was obtained from the Protein Data Bank at Brookhaven National Laboratory, Long Island New York and was Brookhaven code 6ADH. The NAD^+ and dimethyl sulfoxide molecules have been removed from the pictorial representation of the conformation. Colour conventions for individual amino acids are as in Figure 1.4.1(a).

```
HLADH       1 STAGKVIKCKAAVLWEEKKPFSIEEVEVAPPKAHEVRIKMVATGICRSDDHVVSGTL-V-  58
YADH1       1 ---SIPETQKGVIFYESHGKLEHKDIPVPKPKANELLINVKYSGVCHTDLHAWHGDWPLP  57
                                            *   *           *  *

HLADH      59 TPLPVIAGHEAAGIVESIGEGVTTVRPGDKVIPLFTP-QCGKCRVCKHPEGNFCLKNDLS 117
YADH1      58 VKLPLVGGHEGAGVVVGMGENVKGWKIGDYAGIKWLNGSCMACEYCELGNESNCPHADLS 117
                 *   ***  *     *  *     **                  *

HLADH     118 MPRGTMQDGTSRFTCRGKPIHHFLGTSTFSQYTVVDEISVAKIDAASPLEKVCLIGCGFS 177
YADH1     118 G----------------------YTHDGSFQQYATADAVQAAHIPQGTDLAEVAPILCAGI 156

HLADH     178 TGYGSAVKVAKVTQGSTCAVFGL-GGVGLSVIMGCKAAGAARIIGVDINKDKFAKAKEVG 236
YADH1     157 TVY-KALKSANLMAGHWVAISGAAGGLGSLAVQYAKAMGY-RVLGIDGGEGKEELFRSIG 214
                        *      *  *  *                  *             *

HLADH     237 ATECVNPQDYKKPIQEVLTEMSNGGVDFSFEVIGRLDTMVTALSCCQEAYGVSVIVGVPP 296
YADH1     215 GEVFIDFTKEKDIVGAVLK-ATDGGAHGVINVSVSEAAIEASTRYVR-ANGTTVLVGMPA 272
                                  *

HLADH     297 DSQNLSMNPMLLLSGRTWKGAIFGGFKSKDSVPKLVADFMAKKFALDPLITHVLPFEKIN 356
YADH1     273 GAKCCSDVFNQVVK----SISIVGSYVGNRADTREALDFFAR--GLVKSPIKVVGLSTLP 326
                                  *

HLADH     357 EGFDLLRSGE-SIRTILTF---                                        374
YADH1     327 EIYEKMEKGQIVGRYVVDTSK                                        347
```

Figure 1.4.1(a) The Primary Structure of HLADH and of YADH1

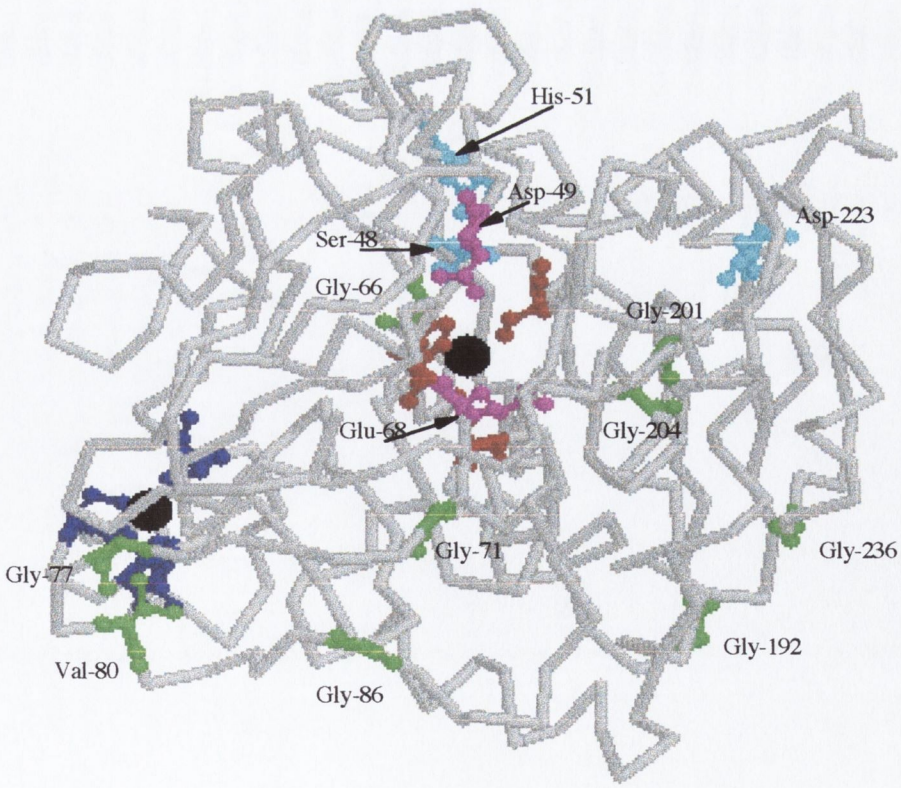


Figure 1.4.1(b) The Tertiary Structure of the HLADH E Subunit

quaternary structure, the absence of the segment with the tetrameric quaternary structure typical of microbial ADHs (Kaiser *et al.* (1993). However, Reid & Fewson (1994) have noted that the 119 - 139 segment from HLADH is also present in the primary structure of *E. coli* threonine dehydrogenase and *Alcaligenes eutrophus* fermentative ADH; both of these enzymes are tetrameric. The segment is absent in the primary structure of the dimeric ADH purified from *S. solfataricus* by Ammendola *et al.* (1992).

Jörnvall *et al.* (1987) aligned the amino-acid sequences of sixteen eukaryotic, NAD^+-dependent, medium-chain, zinc-dependent ADHs and sheep liver sorbitol dehydrogenase [Jörnvall *et al.*, 1984(b)]. Twenty-two strictly conserved residues, of which eleven were Gly residues, were identified and these are shown in Figure 1.4.1(a) where the twenty-two residues are indicated by an asterisk under the YADH1 sequence. The three residues indicated in red in Figure 1.4.1(a) (Cys-46, His-67, Cys-174) and the four Cys residues indicated in blue (Cys-97, Cys-100, Cys-103, Cys-111) are the ligands to the catalytic-zinc ion and the second-zinc ion, respectively (Eklund *et al.*, 1974). The Zn^{2+} ligands are conserved in the Jörnvall *et al.* (1987) alignment when sorbitol dehydrogenase is removed.

Sun & Plapp (1992) aligned the sequences of forty-seven members of the medium-chain dehydrogenase / reductase superfamily, including forty-two zinc-dependent ADHs (both NAD^+- and $NADP^+$-dependent), from twenty-six different species (three bacterial sequences). The number of strictly conserved residues were reduced to nine [eight Gly and Val-80; green in Figure 1.4.1(a)]. The low number of conserved residues is not surprising considering the functional and structural diversity of the different ADHs included in the alignment. The number of conserved residues increases to twelve if guinea pig ζ-crystallin is excluded.

The minimal requirements for an NAD^+-dependent, medium-chain, zinc-dependent ADH containing two zinc ions per subunit, were defined by Sun & Plapp (1992) and with reference to Figure 1.4.1(a) these requirements were: eight Gly and one Val residues (green), required to form the basic folded structure; three residues that are necessary to bind the catalytic-zinc (red); four Cys ligands to the second-zinc ion (blue); Asp-49 and Glu-68 (magenta), second-sphere ligands to the catalytic-zinc ion; Asp-223 (cyan), responsible for discrimination against $NADP^+$ binding; Ser or Thr at the position equivalent to Ser-48 (cyan), involved in the proton-release mechanism and His or Tyr at the position equivalent to His-51 (cyan), also involved in the proton-release mechanism. The position of these defining residues in the three-dimensional structure of the E subunit of HLADH are shown in Figure 1.4.1(b).

-1.4.2 HLADH as the Model for the Tertiary Structure

The tertiary structure of HLADH and the interactions between HLADH and coenzyme (the binary complex) and between HLADH, coenzyme and substrate (the ternary complex) have been extensively studied by X-ray crystallography: (Brändén *et al.*, 1973; Eklund *et al.*, 1974; Eklund *et al.*, 1976; Eklund *et al.*, 1981; Eklund *et al.*, 1982; Ramaswamy *et al.*, 1994(a), 1999). The other medium-chain, zinc-dependent ADHs for which crystal structures have been determined are summarised in Table 1.4.1. YADH1 and ADH of *S. solfataricus* have been crystallised, by Ramaswamy *et al.* [1994(b)] and Pearl *et al.* (1993), respectively, but only preliminary data have been published to date.

Each subunit consists of an N-terminal domain (residues 1 - 175), a C-terminal domain (residues 319 - 374) and a coenzyme-binding domain (residues 176 - 318, Eklund *et al.*, 1974). The coenzyme-binding domain is indicated in green in Figure 1.4.2. An active-site cleft (substrate-binding pocket) accommodates alcohol, the nicotinamide moiety of the coenzyme and the catalytic-zinc ion.

The N-terminal domain comprises mainly β-sheet structures, provides the amino-acid ligands to the two zinc ions and provides most of the groups that control substrate specificity and catalytic activity. The coenzyme-binding domain is formed from six parallel strands of pleated-sheet flanked by four helices (Eklund *et al.*, 1974), binds one molecule of NAD^+ and is similar in three-dimensional (but not primary) structure to the corresponding NAD^+-binding domains of other dehydrogenases (Eklund *et al.*, 1974; Ohlsson *et al.*, 1974; Rossmann *et al.*, 1975; Edwards *et al.*, 1996).

Enzyme	Reference(s)
Human ADH4	Xie *et al.* (1997)
Human ADH1B	Hurley *et al.* (1991) Davis *et al.* (1996)
Cod liver ADH	Ramaswamy *et al.* (1996)

Table 1.4.1 **References for Crystal Structures of Medium-Chain, Zinc-Dependent ADHs**

Figure 1.4.2 **The Coenzyme-Binding Domain of HLADH**

The primary structure [Figure 1.4.2(a)] and the tertiary structure [Figure 1.4.2(b)] of HLADH are those described in the legend to Figure 1.4.1. In the primary structure and in the tertiary structure, the coenzyme-binding domain is indicated in green. In the N-terminal domain residues that form ligands to the catalytic-zinc ion are coloured red, ligands to the second-zinc ion are coloured blue and the second-sphere ligands to the catalytic-zinc ion are coloured magenta. In the C-terminal domain an Arg residue that forms a salt-bridge with Glu-68 is coloured yellow.

```
  1 STAGKVIKCKAAVLWEEKKPFSIEEVEVAPPKAHEVRIKMVATGICRSDDHVVSGTL-V-  58
                                    *   *          *  *

 59 TPLPVIAGHEAAGIVESIGEGVTTVRPGDKVIPLFTP-QCGKCRVCKHPEGNFCLKNDLS 117
      *   ***  *     *  *     **                  *

118 MPRGTMQDGTSRFTCRGKPIHHFLGTSTFSQYTVVDEISVAKIDAASPLEKVCLIGCGFS 177


178 TGYGSAVKVAKVTQGSTCAVFGL-GGVGLSVIMGCKAAGAARIIGVDINKDKFAKAKEVG 236
                  *      *  *  *                  *           *

237 ATECVNPQDYKKPIQEVLTEMSNGGVDFSFEVIGRLDTMVTALSCCQEAYGVSVIVGVPP 296
                           *

297 DSQNLSMNPMLLLSGRTWKGAIFGGFKSKDSVPKLVADFMAKKFALDPLITHVLPFEKIN 356
                           *

357 EGFDLLRSGE-SIMTILTF---                                       374
```

Figure 1.4.2(a)

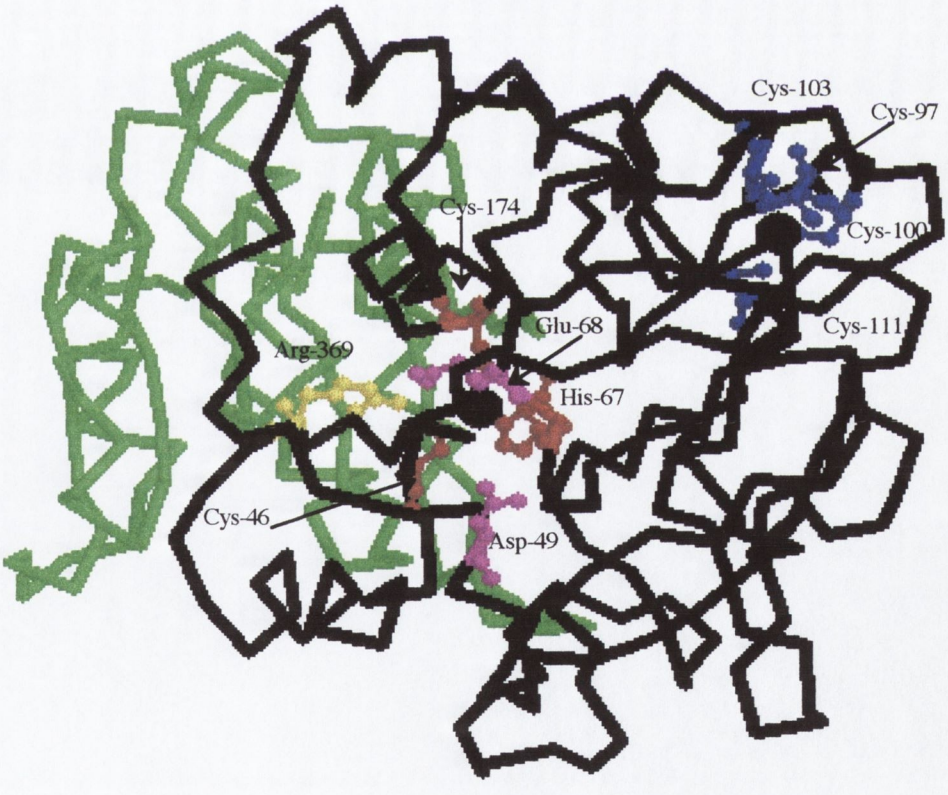
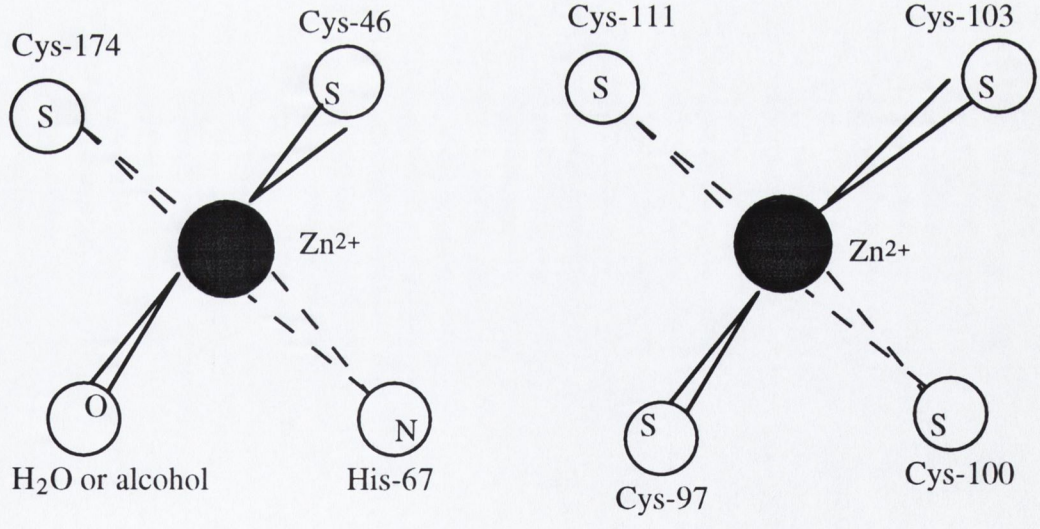


Figure 1.4.2(b)

1.4.2.1 Zinc Ions in the Tertiary Structure

Both zinc ions have the coordination number four and a distorted tetrahedral geometry (Figure 1.4.3). The catalytic-zinc ion is situated at the bottom of the substrate-binding pocket. The ligands are far apart in the primary sequence, and it is therefore assumed (Eklund *et al.*, 1976) that the N-terminal domain must be folded in the final conformation before the catalytic-zinc ion binds. The fourth coordination-position points into the substrate-binding pocket and is occupied by either H_2O or alcohol. When alcohol is liganded to Zn^{2+} it is positioned within hydrogen-bonding distance of the hydroxyl group of Ser-48 (Eklund *et al.*, 1982). The experimental evidence (Pettersson, 1987) indicates that alcohol is bound as the alcoholate anion. Asp-49 and Glu-68 (magenta in Figure 1.4.2) are in close proximity to, but not directly ligated to, the catalytic-zinc ion and are referred to as second-sphere ligands (Sun & Plapp, 1992). Replacing these residues with positively charged amino acids reduced the catalytic efficiency of YADH1 (Ganzhorn & Plapp, 1988). The carboxyl group of Asp-49 hydrogen-bonds to N-3 of the imidazole group of the catalytic-zinc ligand His-67. Glu-68 is located behind the metal ion, opposite the substrate-binding pocket (Figure 1.4.2). The guanidino group of Arg-369 (yellow in Figure 1.4.2) forms a salt-bridge with Glu-68 (Eklund *et al.*, 1976).

The ligands to the second-zinc ion are situated close together in the primary structure and form part of a lobe (residues 95 - 113) that participates in subunit interactions (Eklund *et al.*, 1976).



Ligands to the Catalytic-Zinc Ion **Ligands to the Second-Zinc Ion**

Figure 1.4.3 **Ligands to the Two Zinc Ions in HLADH**

1.4.2.2 The NAD^+-Binding Site

The three-dimensional structure of the E subunit of HLADH with NAD^+ bound, and the positions of residues involved in NAD^+ binding are shown in Figure 1.4.4.

Binding of NAD^+ is accompanied by a conformational change from the open, apoenzyme form, to a closed, holoenzyme conformation (Eklund *et al.*, 1981; Colonna-Cesari *et al.*, 1986). Hydrophobic stabilisation is thought to provide the major driving-force for the transition (Pettersson, 1987) and the conformational change is dependent on rearrangement of a loop (residues 293 - 298) within the coenzyme-binding domain (Ramaswamy *et al.*, 1999). However, the precise details are not fully understood (Hemmingsen *et al.*, 1996; Ramaswamy *et al.*, 1999). The transition from open to closed conformation narrows the substrate-binding pocket, brings the catalytic-zinc ion closer to the coenzyme-binding region and results in complete dehydration of the active-site region. The nicotinamide-binding site is buried in a slot at the bottom of the substrate-binding pocket. One side of the slot is formed by hydrophobic residues from the coenzyme-binding domain, while the other side is highly polar and is bounded by the catalytic-zinc ion and its ligands. There is stereospecific binding of the nicotinamide ring with its A-side directed towards the catalytic-zinc ion. HLADH is therefore a Class-A dehydrogenase, transferring the hydride ion to the *pro-R* position [Levy & Vennesland, 1957; Dickinson & Dalziel 1967(a)]. Thr-178 (cyan; Figure 1.4.4) is in van der Wals contact to C4 of the nicotinamide ring. There is hydrogen-bonding between Ser-48 (cyan) and the nicotinamide ribose, and hydrogen-bonding between His-51 (cyan) and the nicotinamide ribose.

The adenosine part of the coenzyme is accommodated in a shallow cleft at the surface of the coenzyme-binding domain with the adenine ring in a hydrophobic pocket. Two charged residues, Arg-271 and Asp-273 (yellow; Figure 1.4.4) form an ionic interaction at the surface of the pocket. Arg-271 and Asp-273 are not conserved in YADH1, the equivalent residues being Ser and Ala, respectively. The adenine N6 atom points out of the pocket towards the solution. The 3'-oxygen of the adenosine ribose is hydrogen-bonded to Lys-228 (magenta; Eklund *et al.*, 1981). Both of the adenosine ribose hydroxyl groups form hydrogen-bonds to Asp-223 (cyan). Hydrogen-bonding between adenosine ribose and Asp-223 is responsible for discrimination against $NADP^+$ binding (Eklund *et al.*, 1974; Fan *et al.*, 1991). Asp-223 is replaced by Gly in microbial medium-chain, $NADP^+$-dependent ADHs (Reid & Fewson, 1994).

The diphosphate group is bound close to the rim of the substrate-binding pocket and makes a bend over the edge of the pleated-sheet structure of the coenzyme-binding domain. Residues 194 - 223 form a compact βαβ unit typical of the ADP-

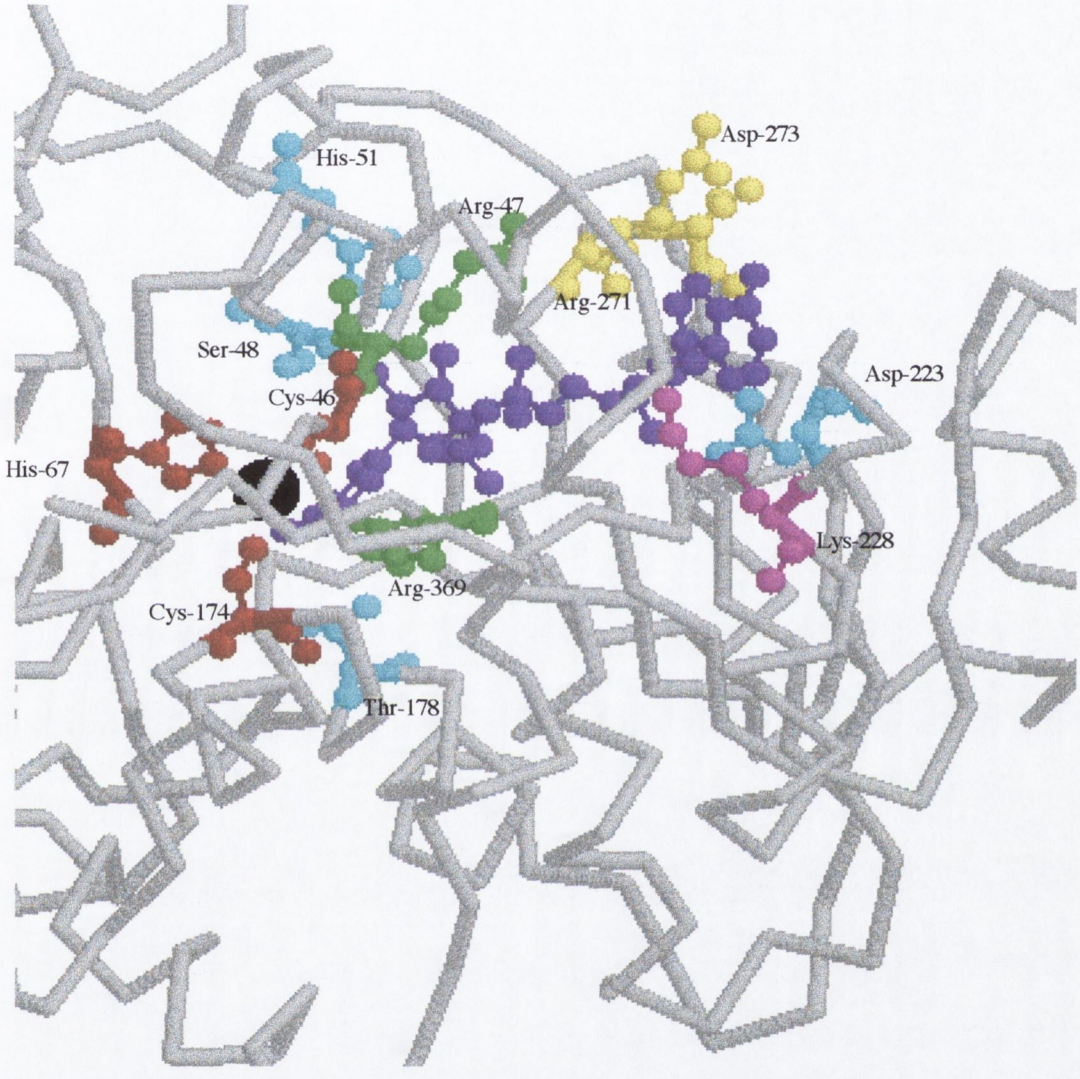


Figure 1.4.4 **Amino Acids Involved in Coenzyme Binding to HLADH**

The structure presented is a selected portion of the closed conformation of the HLADH E subunit that was shown in Figure 1.4.1. The NAD^+ molecule is coloured purple.

binding site of dinucleotide-binding proteins (Wierenga *et al.*, 1986). Residues 201 - 213, within the βαβ unit, form the diphosphate-binding helix. The negative charge on diphosphate interacts with the positive end of the α-helical dipole (Hol *et al.*, 1978; Wierenga *et al.*, 1985). There are ionic interactions between the nicotinamide phosphate and Arg-369 and between the adenosine phosphate and Arg-47 (Eklund *et al.*, 1981). There are also hydrogen-bonds between the phosphate oxygen atoms and the main-chain nitrogen atoms.

1.4.2.3 The Substrate-Binding Pocket

The inner part of the substrate-binding pocket is formed by the catalytic-zinc ion at the bottom, the ligands to the catalytic-zinc ion and the nicotinamide ring of the coenzyme. The main part of the substrate-binding pocket is lined with hydrophobic side chains. Fifteen amino acids have been identified by X-ray crystallography (Eklund *et al.*, 1974; 1982) as being associated with alcohol binding in HLADH and these are listed in Table 1.4.2. All fifteen residues do not necessarily come from the subunit that supplied the ligands to a particular catalytic-zinc ion (Table 1.4.2). Substrate-binding residues in YADH1 (Table 1.4.2) were identified by Jörnvall *et al.* (1978) by alignment of the horse liver and yeast sequences. YADH1 has a limited substrate specificity, compared to HLADH (Section 1.5.1), and that has been rationalised on the basis of the presence of bulky residues lining the wall of the YADH1 substrate-binding pocket (Jörnvall *et al.*, 1978; Eklund *et al.*, 1987; Table 1.4.2). Protein engineering studies (Table 1.4.2) have confirmed the role played by some of the substitutions in limiting the substrate specificity of YADH1 compared to HLADH. A similar approach has been taken to explain differences in substrate specificities of human liver isoenzymes (Light *et al.*, 1992; Xie *et al.*, 1997).

1.5 CATALYTIC MECHANISM OF HLADH AND YADH

1.5.1 The Ordered Mechanism of Alcohol Oxidation

HLADH has a very broad substrate specificity and catalyses the oxidation of a wide range of primary alcohols and secondary alcohols [Dalziel & Dickinson, 1966; Dickinson & Dalziel, 1967(b)], benzyl alcohol (Shearer *et al.*, 1993), cyclohexanol (Light *et al*, 1992), vitamin A and ethanediol (Sund & Theorell, 1963). Among the primary alcohols the value of the Michaelis-Menten constant (K_m) for alcohol decreases with increasing alkyl-chain length (Dalziel and Dickinson, 1966; Light *et al.*, 1992). Primary alcohols and aldehydes are better substrates (lower value of K_m) than secondary aldehydes and ketones (Dalziel & Dickinson, 1966). YADH1 does not catalyse the oxidation of cyclohexanol or secondary alcohols with alkyl groups longer than CH_3 (Sund and Theorell, 1963; Dickinson and Dalziel, 1967).

HLADH	Comments	YADH1	Reference(s) to protein engineering studies in YADH1
Ser-48	Bottom of pocket, side-chain position variable	Thr-45	Murali & Creaser (1986) Creaser *et al.* (1990) Green *et al.* (1993)
Leu-57	Top of pocket	Trp-54	Green *et al.* (1993) Weinhold & Benner (1995)
Phe-93	Bottom of pocket, space restricting	Trp-92	Murali & Creaser (1986) Creaser *et al.* (1990) Green *et al.* (1993)
Phe-110		Asn-110	
Leu-116 Ser-117		Leu-116 Ser-117	
Phe-140 Leu-141		Tyr-119 Thr-120	
Thr-178	Positions nicotinamide with respect to alcohol	Thr-157	
Val-294	In contact with nicotinamide ribose and Leu-57 (Eklund *et al.*, 1981)	Met-270	Ganzhorn *et al.* (1987)
Pro-296	On a loop that underoges conformational change upon coenzyme binding (Eklund *et al.*, 1981)	Ala-272	
Met-306	From the second subunit (Eklund & Brändén, 1983)	Asn-282	
Leu-309	From the second subunit (Eklund & Brändén, 1983)	Val-285	
Ser-310		Lys-286	
Ile-318		Ile-290	

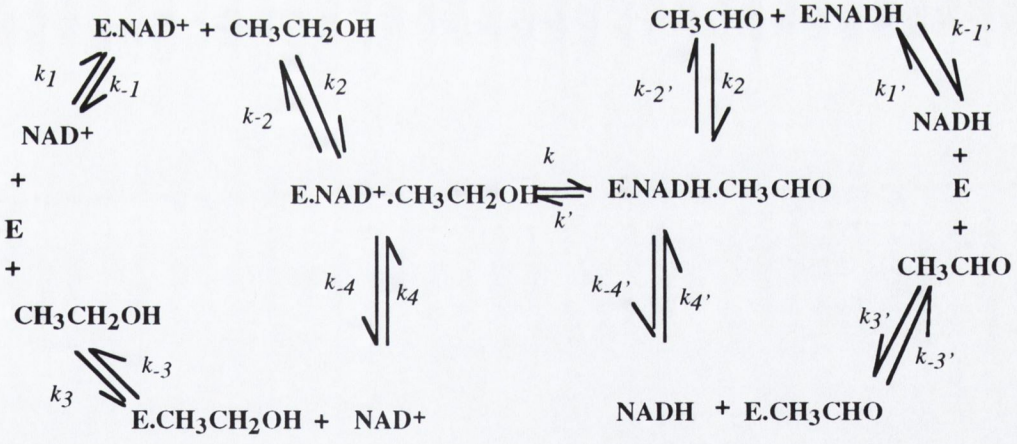
Table 1.4.2 **Residues Associated with Alcohol Binding in HLADH and YADH1**

HLADH, but not YADH1, will catalyse the oxidation of methanol (Sheehan *et al.*, 1988).

Studies on the kinetics and the mechanism of catalysis of HLADH (Theorell & Chance, 1951; Dalziel, 1962; Silverstein & Boyer, 1964; Wratten & Cleland, 1965; Dalziel & Dickinson, 1966) and studies on YADH1 [Dickinson & Dalziel, 1967(b); Dickinson & Monger, 1973] made significant contributions to the understanding of two-substrate reactions and fall within the category of 'classic kinetic studies'. These studies are reviewed by Dalziel (1975), Klinman (1981) and Pettersson (1987). The two subunits in the EE isoenzyme are accepted as equivalent as far as the kinetic mechanism is concerned (Brändén *et al.*, 1975; Pettersson, 1987).

Oxidation of alcohol results in the reduction of coenzyme and concurrent release of a proton, from the alcohol substrate, into solution. The simple, compulsory-order mechanism, which was proposed for HLADH by Theorell and Chance (1951), is shown in Equation 4. The binding of NAD^+ to enzyme is the first step and the dissociation of the E.NADH complex is rate-limiting. The mechanism proposed by Dalziel and Dickinson (1966) for HLADH, and which includes the formation of ternary complexes, is shown in Scheme 1.

$$\mathbf{E + NAD^+ \rightleftharpoons E.NAD^+ + CH_3CH_2OH \rightleftharpoons NADH + E.CH_3CHO \rightleftharpoons CH_3CHO + E} \qquad \mathbf{(4)}$$



Scheme 1 Mechanism for HLADH Catalysed Oxidation of Alcohols Proposed by Dalziel & Dickinson (1966)

Over substrate concentration ranges where Michaelis-Menten kinetics are obeyed, conversion of the ternary complex $E.NAD^+.CH_3CH_2OH$ to $E.NADH.CH_3CHO$ (k) and dissociation of CH_3CHO from $E.NADH.CH_3CHO$ (k_{-2}) are fast, therefore, the ternary complex concentration in the steady-state is negligible, and $E.NAD^+.CH_3CH_2OH$ and $E.NADH.CH_3CHO$ are not significant. Oxidation of primary alcohols can therefore be said to conform to the compulsory-order Theorell

and Chance mechanism (Equation 4). NAD^+ binds first (k_1), and the release of NADH from the E.NADH binary complex (k_{-1}) is the rate-determining step (Dalziel, 1975). Trends in the values of the catalytic rate constant (k_{cat}) and K_m, as substrate is varied, are used to establish the reaction mechanism of ADH. The diagnostic test for the compulsory order of substrate addition, with NAD^+ leading, is invariance in the values of $K_m^{NAD^+}/k_{cat}$ as alcohol substrate is varied (Dalziel & Dickinson, 1966). Furthermore, as the rate-limiting step does not involve an alcohol, or the product of an alcohol, k_{cat} values are also invariant as substrate alcohol is varied.

Substrate inhibition, which occurs at high concentration of primary alcohols, is attributed to the formation of an abortive complex, $E.NADH.CH_3CH_2OH$. The complex is not dead end; NADH does dissociate, but at a lower rate than from the enzyme-NADH-binary complex, (Dalziel & Dickinson, 1966).

Secondary alcohols do not conform to the ordered Theorell and Chance mechanism. According to the Dalziel & Dickinson mechanism for secondary alcohols, slow (rate-limiting) intramolecular rearrangement within the ternary complex $E.NAD^+ROH$ results in a significant concentration of $E.NAD^+.ROH$ and, therefore, significant dissociation of NAD^+ (k_{-4}) from the ternary complex.

The reverse reaction, aldehyde reduction, is random as far as NADH and aldehyde binding are concerned (Andersson *et al.*, 1984; Pettersson, 1987), however, release of products is ordered, with desorption of coenzyme being the last step. The reduction of aromatic aldehydes involves rate-determining dissociation of the enzyme-alcohol complex (Wratten & Cleland, 1965) whereas the reduction of acetaldehyde is limited by the chemical step of hydride transfer.

The affinity of the enzyme for NAD^+ increases when the pH is raised from 6 to 10, whereas the affinity for NADH remains the same up to pH 9 and then decreases as the pH is raised further (Brändén *et al.*, 1975). The binding of NADH is reliant on the ionization state of a group with a pKa of 9.2 in the apoenzyme and the same group has been proposed to be involved in the association of NAD^+ with the free enzyme (DeTraglia *et al.*, 1977; Andersson *et al.*, 1981). The binding of NAD^+ to form the holoenzyme perturbs the pKa of the ionizable group from 9.2 to 7.6.

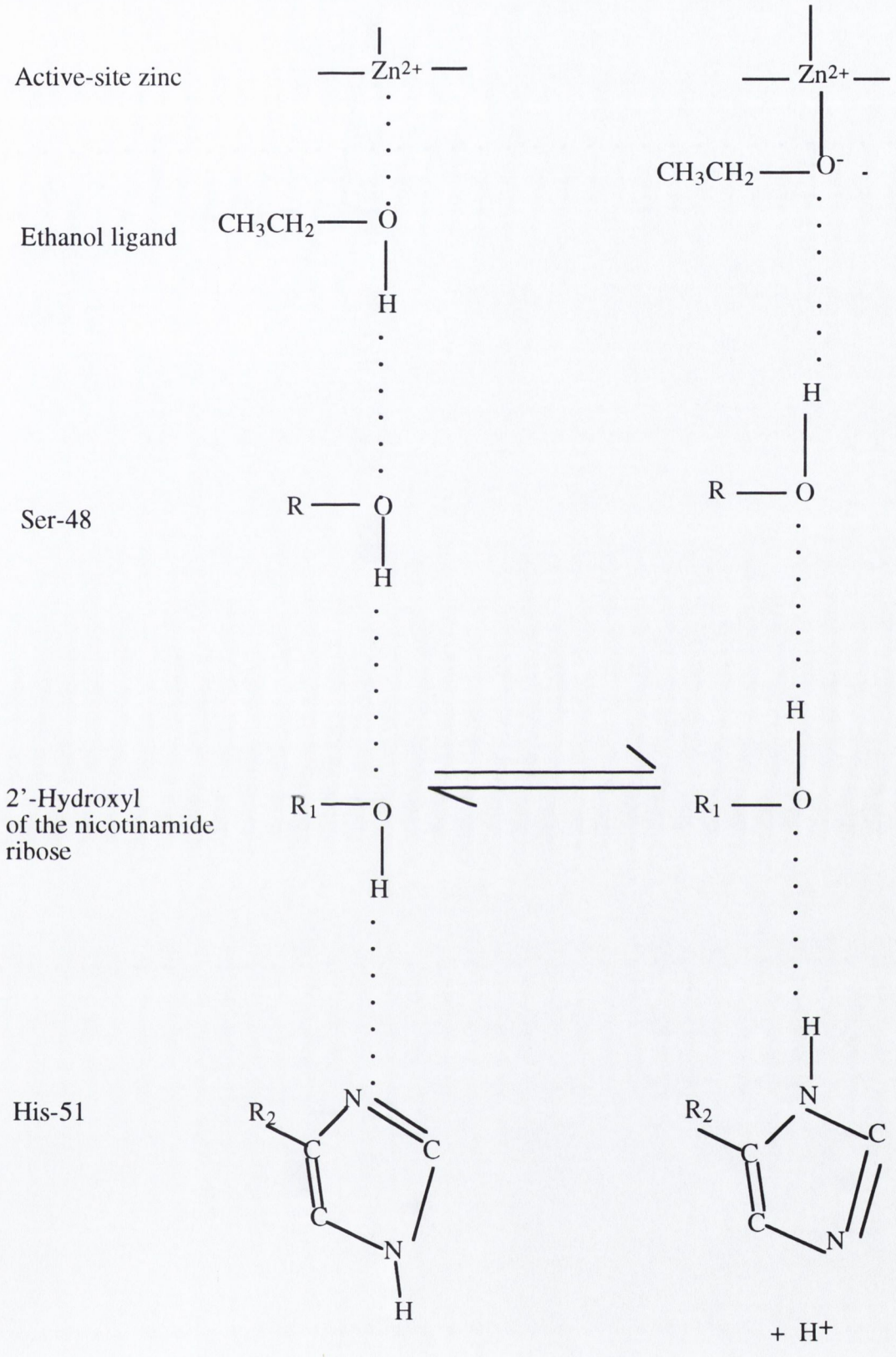
For YADH1, with ethanol as substrate, the mechanism approximates to the compulsory-order Theorell and Chance mechanism proposed for HLADH and primary alcohols. As with HLADH, dissociation of NADH from the binary complex E.NADH is rate-limiting. A compulsory-order mechanism is not obeyed with the longer chain, primary alcohols (Dickinson & Monger, 1973). The catalytic interconversion of the ternary complexes $E.NAD^+.RCH_2OH$ and E.NADH.RCHO is thought to be rate-limiting for oxidation of primary alcohols other than ethanol. Substrate inhibition by primary alcohols is not observed, indicating that the abortive

complex E.NADH.RCH_2OH is not formed. The K_m for alcohol increases with increasing alkyl chain length.

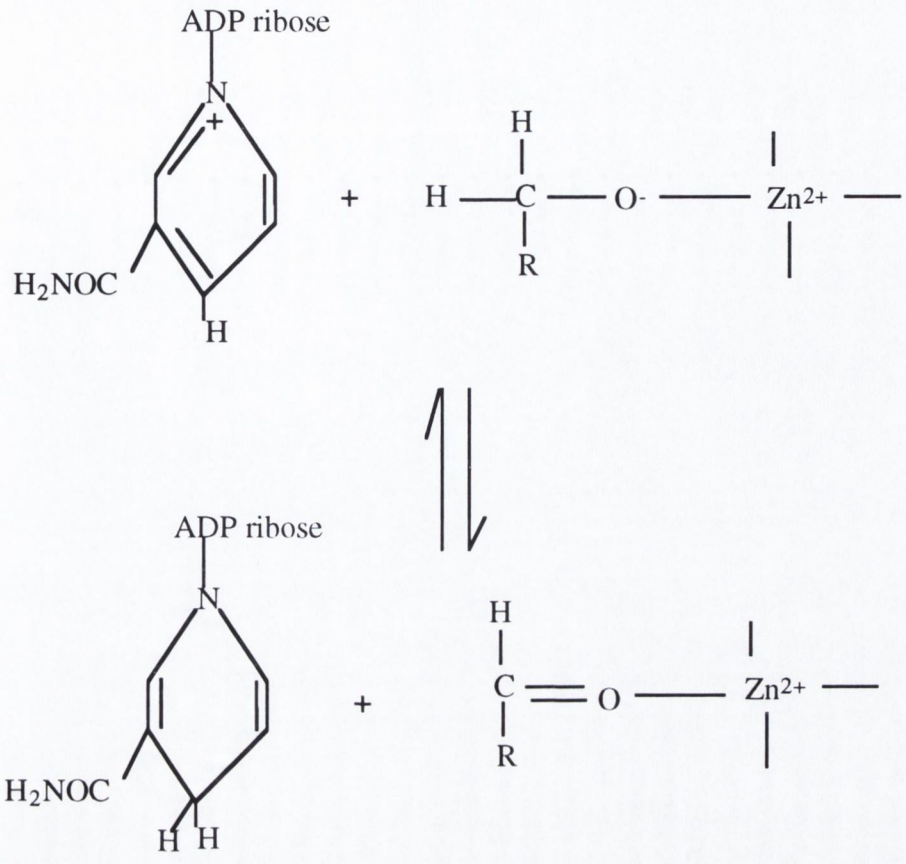
1.5.2 Proton Release and Hydride Transfer

The currently accepted mechanism for proton release (Pettersson, 1987) is the proton relay system originally proposed by Eklund *et al.* (1982). Proton release to solution during alcohol oxidation occurs from the ternary complex and precedes hydride transfer to NAD^+ [Kvassman & Pettersson, 1980(a); 1980(b)]. The binding of alcohol to Zn^{2+} in the presence of NAD^+ results in immediate deprotonation of the substrate. The substrate is bound in the active-site in an anhydrous environment and the proton is shuttled to the surface of the protein *via* a system of hydrogen-bonds (Andersson *et al.*, 1981; Eklund *et al.*, 1982; Hennecke & Plapp, 1983). His-51 acts as the proton acceptor (Ehrig *et al.*, 1991; Sakoda & Imanaka, 1992). The hydrogen-bond system involves Ser-48, the 2'-hydroxyl of nicotinamide ribose and His-51 (Scheme 2) and is supported by the crystallographic data defining the relative positions of the groups shown in Scheme 2 [Eklund *et al.*, 1974; Eklund *et al.*, 1982; Ramaswamy *et al.*, 1994(a)], the inhibitory affect of histidine modifying reagents on catalytic activity (Hennecke & Plapp, 1983), the pH dependence of alcohol binding and ternary complex formation (Kvassman *et al.*, 1981), and protein engineering studies (Sakoda & Imanaka, 1992).

The current view of hydride transfer (Pettersson, 1987) is that it occurs directly from the zinc-bound alcoholate ion (the transition-state intermediate) to NAD^+ (Scheme 3). Alcohol displaces the H_2O ligand to the catalytic-zinc ion. The zinc ion polarises the alcoholate ion so that direct hydride transfer and subsequent rearrangement to an aldehyde can occur. The *pro-R* hydrogen of ethanol is transferred to the *pro-R* (or A position) of nicotinamide C4 (Eklund *et al.*, 1982). In aldehyde reduction, the zinc ion polarises the aldehyde substrate in such a way that carbon atom one of the aldehyde increases in electrophilicity, making the transfer of a proton from the coenzyme favourable.



Scheme 2 **The Proton-Relay System of HLADH**



Scheme 3 **Hydride Transfer in HLADH**

1.6 *B. stearothermophilus* ADH

1.6.1 ADH from *B. stearothermophilus* NCA 1503 ADH

An NAD^+-dependent ethanol dehydrogenase (ADH 1503) was purified from *B. stearothermophilus* strain NCA 1503 by Kolb & Harris (1971). ADH 1503 had a pI of 5.5 and was tetrameric with an apparent subunit M_r of approximately 35,000. The N-terminal amino-acid sequence (forty-five residues) of ADH 1503 was determined by Bridgen *et al.* (1973) and was the first amino-acid sequence to be obtained for a prokaryotic ADH. The sequence data indicated that ADH 1503 was a medium-chain, zinc-dependent ADH. The amino-acid composition of the purified enzyme was determined by Runswick & Harris (1978).

1.6.2 ADH from *B. stearothermophilus* DSM 2334, Comparison with ADH 1503

B. stearothermophilus DSM 2334 was included in a range of bacteria (40 strains, 24 species) that were screened for NAD^+-dependent methanol dehydrogenase activity by D. W. Tonge and J. Foundling (unpublished). The only organism that exhibited a significant methanol dehydrogenase activity was *B. stearothermophilus* DSM 2334. The specific activity of the *B. stearothermophilus* ADH was 0.2 μmol min^{-1} mg^{-1} of crude cell extract with methanol as substrate and 2.0 μmol min^{-1} mg^{-1} of crude cell extract with ethanol as substrate.

The *B. stearothermophilus* DSM 2334 ADH (ADH 2334) was purified and characterised by Sheehan *et al.* (1988). The purified enzyme had an apparent subunit M_r of approximately 35,000. Amino-acid analyses of the purified enzyme revealed that the amino-acid composition was very similar to that of ADH 1503 (38 +/- 15 amino-acid changes). A six-amino-acid N-terminal sequence determined for ADH 2334 was identical to that of ADH 1503. Despite the apparent high primary structural identity between ADH 1503 and ADH 2334, polyclonal antibodies raised against ADH 2334 did not cross-react with ADH 1503, and under non-denaturing conditions the electrophoretic mobilities of the two purified ADHs were quite different, ADH 2334 having the greater anionic mobility. Furthermore, ADH 1503 displayed no catalytic activity in the presence of methanol. An enzyme-linked-immunoabsorbent-assay was used to show that ADH 1503 had less than 6% of its antigenic activity in common with ADH 2334.

1.6.3 Substrate Specificities, Kinetic Properties and Regulation of ADH 1503 and ADH 2334

Systematic studies (Sheehan *et al.*, 1988) of the kinetics of ADH 2334 and ADH 1503 showed that the two *B. stearothermophilus* ADHs had different pH

optima, substrate specificities and kinetic mechanisms. A summary of the kinetic parameters for the catalytic activity of ADH 2334 and ADH 1503 with primary alcohols, and the corresponding values for HLADH and YADH1, is given in Table 1.6.1. With methanol as substrate, the K_m for ADH 2334 and alcohol was 20 mM and was very similar to the value obtained for HLADH and methanol (30 mM). ADH 1503 and YADH1 are similar in that methanol dehydrogenase activity could not be demonstrated unequivocally for either enzyme. Both *B. stearothermophilus* ADHs catalysed the oxidation of ethanol, propan-1-ol and butan-1-ol, but with very different K_m and k_{cat} values. The values of K_m and k_{cat} obtained for ADH 2334 and ethanol were 82 μM and 1.3 s^{-1} receptively, which are both lower than those observed for ADH 1503 and ethanol (K_m value of 450 μM and k_{cat} value of 6.0 s^{-1}). Both *B. stearothermophilus* ADHs were similar to HLADH in that the value of K_m decreased with increasing substrate alkyl-chain length. In terms of the actual values of K_m and k_{cat}, ADH 2334 was more similar to HLADH than to ADH 1503.

The reaction mechanism determined for ADH 2334 was identical to that determined for HLADH and primary alcohols by Dalziel and Dickinson (1966). The values of $K_m^{NAD^+}/k_{cat}$ for ADH 2334 with ethanol, propan-1-ol or butan-1-ol as substrate are indistinguishable (Table 1.6.1), hence there is a compulsory order of substrate addition, with NAD^+ leading (Section 1.5.1). Furthermore, the invariance of the k_{cat} values for the three substrates requires that the rate-limiting step does not involve an alcohol, or the product of an alcohol, and must therefore be located at NADH release (Section 1.5.1). Substrate inhibition, which according to the mechanism proposed by Dalziel and Dickinson (1966) for HLADH can only occur if the binary enzyme-NADH complex dominates the steady-state population of enzyme species, was observed for ADH 2334. The values of $K_m^{NAD^+}/k_{cat}$ for ADH 1503 with ethanol, propan-1-ol or butan-1-ol as substrate are also indistinguishable, but the value of k_{cat} declines with increasing alkyl-chain length, indicating that NADH release is not rate-limiting. Substrate inhibition was not observed for ADH 1503.

The pH-dependence of ADH 2334 and ADH 1503 with ethanol as substrate was investigated by Sheehan *et al.* (1988) and revealed a further difference between the two enzymes. The pH dependence of ADH 1503 was similar to that of YADH1 and HLADH (Sund and Theorell, 1963) in that the pH optimum was 8. The pH optimum for ADH 2334 was 5.5 - 6.0.

The gene that encodes ADH 2334 (*adh*2334) was cloned by Dowds *et al.* (1988) and will be described in Chapter 3 (Section 3.1.1). The response of both strains of *B. stearothermophilus* to oxygen stress as well as the regulation of ADH 2334 production were also investigated by Dowds *et al.* (1988). *B. stearothermophilus* strain DSM 2334 was found to be an obligate aerobe and the expression of ADH 2334 was catabolite-repressed by glucose. The regulation of

	Methanol		Ethanol			Propan-1-ol			Butan-1-ol		
	k_{cat} (s^{-1})	K_m (mM)	k_{cat} (s^{-1})	K_m (μM)	$K_m^{NAD^+}/k_{cat}$ (μM s)	k_{cat} (s^{-1})	K_m (μM)	$K_m^{NAD^+}/k_{cat}$ (μM s)	k_{cat} (s^{-1})	K_m (μM)	$K_m^{NAD^+}/k_{cat}$ (μM s)
Enzyme											
ADH 2334	0.06	20	1.3	82	30	1.3	18	29	1.1	12	36
ADH 1503	<0.016		6.0	450	14	4.6	370	16	2.4	160	20
HLADH	0.048	30	2.7	180	1.1	3.2	61	1.2	2.9	11	1.1
YADH1	<0.03		450	2170	0.24	66	2920	2.3	25	3200	10

Table 1.6.1 **Kinetic Parameters for the ADH Catalysed Oxidation of Primary Alcohols**

The kinetic parameters for ADH 2334 and ADH 1503 are those determined by Sheehan *et al.* (1988). The kinetic parameters for HLADH were determined by Dalziel & Dickinson (1966) and those of YADH1 by Dickinson & Monger (1973), except for data with methanol as substrate, which were determined by Sheehan *et al.* (1988).

expression of ADH 2334 is therefore the same as that of the oxidative yeast isoenzyme YADH2 (Section 1.3.2.1). Consequently, it may be assumed that ADH 2334 is an enzyme adjusted to operate in the direction of alcohol oxidation, in the absence of a better carbon source. *B. stearothermophilus* 1503 is a facultative anaerobe but there are no data to indicate how production of ADH 1503 is regulated. It is tempting, however, to compare the occurrence and function of ADH 1503 and ADH 2334 to the occurrence and function of the high K_m fermentative YADH1 and the low K_m aerobic YADH2.

1.7 OBJECTIVES

ADH 1503 from *B. stearothermophilus* NCA 1503, and ADH 2334 from *B. stearothermophilus* DSM 2334, appeared to be closely related. The amino-acid compositions of the two enzymes were very similar, indicating that the amino-acid sequences were similar. However, the two enzymes were different in terms of kinetic mechanism, pH optimum and antigenic response. ADH 2334, but not ADH 1503, catalysed the NAD^+-dependent oxidation of methanol. The two ADHs presented an opportunity to examine how differences in substrate specificity and kinetic mechanism are arranged in the primary structure of closely-related enzymes.

The available amino-acid sequence-data were limited, a six residue N-terminal amino-acid sequence for ADH 2334 and a forty-five residue N-terminal amino-acid sequence for ADH 1503. The determination of the complete protein primary structure of the two enzymes was therefore the first objective. Large amounts of ADH 1503, purified by Runswick & Harris (1978) were available and it was proposed to determine the primary structure of ADH 1503 by amino-acid sequencing. *adh*2334 had been cloned by Dowds *et al.* (1988), and that made it possible to determine the primary structure of ADH 2334 by nucleotide sequencing.

The nucleotide sequence of the gene encoding ADH 1503 (Sakoda & Imanaka, 1992; *adh*1503) was published after the commencement of this work. The amino-acid sequence data of ADH 1503 that are presented (Chapter 2) were determined independent of, and prior to, the publication of the work of Sakoda & Imanaka (1992) and are confirmatory.

The second objective was to determine if the differences in kinetic mechanism and substrate specificity between the two enzymes could be linked to differences in the primary structures.

To facilitate study of the properties of ADH 2334 a supply of the enzyme was necessary. The wild-type enzyme had proved very difficult to purify (Sheehan *et al.*, 1988) and therefore, the third objective was the construction of an expression system for the production of recombinant ADH 2334.

CHAPTER 2
AMINO-ACID SEQUENCING OF ADH 1503 AND ADH 2334

2.1 MATERIALS AND METHODS

2.1.1 Sources of Purified ADH 1503 and Purified ADH 2334

The source of ADH 1503 was that purified by Runswick & Harris (1978). ADH 2334 was that purified by Sheehan *et al.* (1988). Both enzymes were received as salt-free freeze-dried powders and were stored as such at -20 °C.

2.1.2 Generation of Peptides from ADH 1503

Solutions used in the denaturation, reduction, alkylation, precipitation and cleavage of ADH 1503 were prepared in analytical grade water that had a specification of $A_{254}<0.002$ and was obtained from LAB-SCAN Analytical Sciences. All solutions were filtered by passage through disposable syringe filters (0.2 μm, Anotec). Where possible, procedures were performed in disposable 1.5 ml microcentrifuge tubes. Glass tubes were rinsed extensively in distilled water and dried in a dust free environment before use.

2.1.2.1 Denaturation, Reduction and Alkylation

ADH 1503 was denatured with guanidine hydrochloride or urea, reduced with dithiothreitol and the cysteine groups were alkylated with 4-vinylpyridine using procedures based on those described by Wilson & Yuan (1989). Solutions containing guanidine hydrochloride or urea were prepared immediately before use. The 4-vinylpyridine (Sigma) was distilled (7 mM Hg, 62 °C), dispensed into 1 ml aliquots and stored under nitrogen in sealed glass tubes at -20 °C.

ADH 1503 (1 mg) was dissolved in a solution of 6 M guanidine hydrochloride (or 6 M urea), 15 mM dithiothreitol, 1 mM Na_2 EDTA, 0.25 M Tris-HCl, pH 8.5 (150 μl) and incubated at room temperature for 2 h. Following the addition of 4-vinylpyridine (20 μl) the incubation was continued for a further 2 h in the dark. When guanidine hydrochloride was used as the denaturant *S*-pyridylethylated protein was recovered by reverse-phase high-performance liquid chromatography (reverse-phase h.p.l.c.; Section 2.1.3.3). When urea was used as the denaturant *S*-pyridylethylated protein was recovered by precipitation in the presence of trichloroacetic acid.

S-Pyridylethylated ADH 1503 was precipitated from solution by incubating on ice for 15 min in the presence of 12.5% (w/v) trichloroacetic acid (addition of an equal volume of ice-cold 25% (w/v) trichloroacetic acid). The precipitated protein was collected by centrifugation (10,000 ***g***, 10 min, 4 °C) and washed repeatedly (without resuspension) with 100 μl of ice-cold acetone until the pH of the acetone rose to pH 6.0. Finally, the protein pellet was dried (at room temperature and atmospheric pressure) and stored at -20 °C.

2.1.2.2 Cleavage with CNBr

S-Pyridylethylated ADH 1503 was cleaved with CNBr using the procedure described by Fontana & Gross (1986). Precipitated protein (Section 2.1.2.1; 1 mg) was dissolved in 100 µl of a freshly prepared solution of CNBr [20 mg ml^{-1} in 70% (v/v) formic acid] and left for 24 h at room temperature in the dark. Cleavage was terminated by the addition of 10 volumes of water and the cleavage mixture was either analysed immediately by reverse-phase h.p.l.c. or was dispensed into aliquots, equivalent to approximately 200 µg of starting protein, freeze-dried and stored at -20 °C.

2.1.2.3 Cleavage with V8 Protease

V8 protease (endoprotease Glu-C from *Staphylococcus aureus* strain V8), obtained from Sigma as a freeze-dried powder, was dissolved in H_2O (1 mg ml^{-1}) and stored, as small aliquots, at -20 °C.

Aliquots of desalted *S*-pyridylethylated ADH 1503 (Section 2.1.3.3; ~2 µg $µl^{-1}$) were adjusted to 4 M urea (addition of 10 M urea), 50 mM Tris-HCl (addition of 2 M Tris-HCl, pH 7.8) and 5% (w/w) V8 protease (addition of 1 mg ml^{-1} V8 protease) and then incubated at 37 °C for 1 - 24 h. Cleavage was terminated by freezing (-20 °C).

2.1.3 Reverse-Phase High-Performance Liquid Chromatography

The instrumentation used for reverse-phase h.p.l.c. consisted of a h.p.l.c pump (LKB model 2150), a solvent conditioner (LKB model 2156), a gradient mixer (LKB model 2040 - 203), a h.p.l.c controller and a spectrophotometer (LKB Ultraspec 11) that was equipped with a 8 µl flow cell and connected to a chart-recorder. Samples were injected onto the column using a Rheodyne injector fitted with a 200 µl sample loop.

The stationary phase was octylsilica (Aquapore RP-300 octylsilica, 300 Å pore-size, 7 µm particle diameter) obtained as pre-packed columns, 30 x 2.1 mm or 220 x 2.1 mm, from Brownlee Labs. A guard column, Brownlee Aquapore RP-8 octylsilica (300 Å pore-size, 7 µm particle diameter, 15 x 3.2 mm) was used with the 220 x 2.1 mm column.

The mobile phase was acetonitrile in trifluoroacetic acid. Acetonitrile was Far UV grade, obtained from LAB-SCAN Analytical Sciences. Water used as the diluent when preparing solvents and samples was analytical grade (Section 2.1.2). Gradients of increasing concentration of acetonitrile were obtained by mixing 77.77% (v/v) acetonitrile [diluted into 0.1% (v/v) trifluoroacetic acid] with 0.1% (v/v) trifluoroacetic acid.

Reverse-phase chromatography was performed at room temperature and a solvent flow rate of 0.2 ml min^{-1}. Columns were equilibrated with 0.1% (v/v)

trifluoroacetic acid before application of protein or peptides. Volumes in excess of 0.2 ml were loaded by multiple injections (trace-enrichment). The eluent from the column was monitored at 208 nm and collected manually after correcting for the dead volume of the tubing between the detector and the collection port. Column performance was monitored by noting the increase in pressure across the column that occurred following repeated use of a column and by counting the number of theoretical plates per peak (N) when RNase, myoglobin and lysozyme (20 μg of each) were separated on the column of interest (Johns, 1989). N is a measure of the zone spread of a peak throughout the chromatographic system and can be calculated from Equation 5 where t is the retention time (min) of the peak measured from the beginning of the gradient to the mid-point of the peak and where w is the peak width at half the peak height (min).

$$N = 5.54\ (t\ /\ w)^2 \qquad (5)$$

The smaller the zone spread the larger N and the more efficient the column. Columns were discarded when there was an obvious decrease in the value of N, or an increase in pressure.

2.1.3.1 Separation of Peptides

Peptides generated by CNBr cleavage were separated on 30 x 2.1 mm reverse-phase columns using linear gradients [0 - 77.77% (v/v) in 30 min] of acetonitrile. Peptides generated by V8 protease cleavage were separated on 220 x 2.1 mm columns using linear gradients [0 - 77.77% (v/v) in 60 min] of acetonitrile. Prior to chromatography, freeze-dried CNBr peptide mixtures (Section 2.1.2.2; 100 - 500 μg) were dissolved in a minimal volume of 4 M guanidine hydrochloride, 0.1% (v/v) trifluoroacetic acid and particulate material was removed by centrifugation (10,000 ***g***, 1 min, room temperature). V8 peptide mixtures (Section 2.1.2.3; 100 - 500 μg) were adjusted to 0.1% (v/v) trifluoroacetic acid [addition of 10% (v/v) trifluoroacetic acid] and to < 1 M urea [addition of 0.1% (v/v) trifluoroacetic acid] and centrifuged as above before chromatography. Recoveries (%) of peptides were estimated by comparing the total area under the curve of the elution profile with the area under the curve obtained when a known amount of *S*-pyridyethylated ADH 1503 was desalted by reverse-phase h.p.l.c.

Column eluent was collected in fractions and either stored at -20 °C or re-chromatographed using a variety of flow rates and acetonitrile gradients. When samples were re-chromatographed they were diluted two-fold with 0.1% (v/v) trifluoroacetic acid before application to the column. When the volume of the fraction to be re-chromatographed exceeded 0.2 ml, it was concentrated to approximately 0.2 ml by lyophilisation, before dilution into 0.1% (v/v) trifluoroacetic acid.

2.1.4 Tricine-Sodium Dodecyl Sulfate-Polyacrylamide-Gel Electrophoresis

Peptide mixtures were analysed using the Tricine-sodium dodecyl sulfate-polyacrylamide-gel electrophoresis (Tricine-SDS-PAGE) system developed by Schägger & von Jagow (1987). Electrophoresis was performed in an Atto vertical dual mini-slab-gel apparatus with gel mould dimensions of 90 x 80 x 1 mm and with 0.2 M Tris-HCl, pH 8.9 as the anode buffer and 0.1 M Tris-HCl, 0.1 M Tricine, 0.1% (w/v) SDS, pH 8.25 as the cathode buffer. A CNBr cleavage mixture of horse-heart myoglobin (LKB), containing peptides with M_r of 2512, 6214, 8159, 14,404 and 16,949, was used as peptide apparent M_r markers.

Gels consisted of three regions, a resolving gel (90 x 40 x 1 mm) overlaid with a "spacer" gel (90 x 10 x 1 mm) which in turn was overlaid with a stacking gel (90 x 18 x 1 mm). Resolving gels were 15.5% (w/v) acrylamide, 1% (w/v) bisacrylamide, 6 M urea (Aristar grade, BDH), 0.1 % (w/v) SDS, 1 M Tris-HCl, pH 8.45. Spacer gels were 9.8% (w/v) acrylamide, 0.3% (w/v) bisacrylamide, 0.1% (w/v) SDS, 1M Tris-HCl, pH 8.45. Stacking gels were 3.8% (w/v) acrylamide, 0.12% (w/v) bisacrylamide, 0.075% (w/v) SDS, 0.75 M Tris-HCl, pH 8.45. Gel solutions were polymerised, in the presence of 0.075% (w/v) ammonium persulfate (Electran grade, BDH) and 0.1% (v/v) *N,N,N',N'*-tetramethylethylenediamine (Electran grade, BDH), under a layer of water-saturated butanol. With the exception that stock solutions were filtered before use (8.0 µm Millipore filters), gels were prepared exactly as described by Schägger & von Jagow (1987). When peptides were electroblotted onto Immobilon PVDF, urea in the resolving gel was replaced by glycerol [13.3% (w/v) final concentration].

Peptides in solution in acetonitrile-trifluoroacetic acid (fractions from h.p.l.c) were prepared for electrophoresis as follows: fractions were dispensed into 1.5 ml microcentrifuge tubes, 7.5 volumes of a filtered solution of 0.4% (w/v) SDS, 5 mM Tris-HCl, pH 6.8 were added to each tube (maximum total volume per tube 300 µl), the solutions were frozen in liquid nitrogen and then freeze-dried to complete dryness in a Savant SpeedVac. The freeze-dried material was solubilised in 6 M urea, 0.01% SERVA Blue G (SERVA Feinbiochemica), to yield a final concentration of 4% (w/v) SDS and 50 mM Tris and loaded onto the gel.

Electrophoresis was performed at room temperature and at 25 V (constant voltage) until the SERVA Blue G tracker dye reached the interface between the "spacer" and resolving gels. The voltage was then increased to 50 V and electrophoresis continued (at room temperature) until the tracker dye had completely "run off" the bottom of the resolving gel. Following electrophoresis gels were either electroblotted immediately or were fixed [shaking for 30 min in 50% (v/v) methanol, 10% (v/v) acetic acid] and then stained [shaking for 1 h in 0.025% (w/v) SERVA

Blue G, 10% (v/v) acetic acid]. Background staining was removed by shaking the gels in multiple changes of 10% (v/v) acetic acid.

2.1.5 Electroblotting of Peptides onto Immobilon PVDF

Following Tricine-SDS-PAGE (Section 2.1.4), peptides were transferred onto Immobilon PVDF (Xu & Shively, 1988) by the standard "wet blotting" Western blotting technique (Harlow & Lane, 1988). The transfers were performed in a Bio-Rad TransBlot "wet blotting" apparatus at room temperature and with 10 mM 3-[cyclohexylamino]-1-propanesulfonic acid (CAPS), 10% (v/v) methanol, pH 11, as the transfer buffer. Transfer buffer was prepared from a stock solution of 100 mM CAPS, adjusted to pH 11 with 4 M NaOH. Immobilon PVDF was obtained from Millipore.

Immobilon was rinsed with methanol, distilled water and, finally, transfer buffer, immediately before use. Following electrophoresis, the gel was soaked in transfer buffer (20 min, room temperature) and then placed in a Western blotting "sandwich" in contact with Immobilon. The transfer of peptides to Immobilon was achieved by applying a constant current of 0.3 A for 30 min. Peptides were detected by staining the Immobilon with 0.1% (w/v) Coomassie blue in 50% (v/v) methanol, 10% (v/v) acetic acid for 30 min. Background staining was removed by soaking the Immobilon in 50% (v/v) methanol, 10% (v/v) acetic acid for 5 min. The blot was left to dry at room temperature. The Coomassie blue stained peptide "bands" of interest were excised from the dry blot and stored at -20 °C in sealed plastic bags.

2.1.6 Automated Amino-Acid Sequencing

Amino-acid sequence analyses of proteins and peptides were performed using the automated Edman procedure (Edman & Begg, 1967; Hewick *et al.*, 1981) in an Applied Biosystems 477A pulsed-liquid protein sequencer (described by Geisow & Aitken, 1989). Phenylthiohydantoin (PTH) derivatives of amino acids were analysed by reverse-phase h.p.l.c. in an Applied Biosystems 120A PTH analyser connected to a dual-pen chart-recorder. Chromatography and identification of PTH amino acids by their relative retention times were according to the criteria of Hunkapiller (1985). The protein sequencer and PTH analyser were operated using Applied Biosystems version 1.61 software.

Trifluoroacetic acid, used to solubilise freeze-dried proteins prior to sequencing, was obtained from Sigma (product T-6S08) and was diluted to 50% (v/v) with water (analytical grade; Section 2.1.2) before use. All other sequencing reagents and all accessories were obtained from Applied Biosystems. Trifluoroacetic acid-treated glass fibre discs were coated with 3 mg of the polycationic carrier BioBrene (Tarr *et al.*, 1978; Klapper *et al.*, 1978) and precycled once before use.

Volumes (<30 µl) equivalent to 100 - 500 pmol were applied directly to the precycled glass fibre discs. Peptides that had been excised from Immobilon (Section 2.1.5) were placed on top of precycled glass fibre discs. Peptides in solution in acetonitrile (fractions from reverse-phase h.p.l.c) were pipetted directly on to the precycled glass fibre discs (maximum volume loaded was 0.2 ml).

The sequence data were analysed both manually, by comparing the retention times of PTH amino acid peaks on the recorder traces with those of standards, and using Applied Biosystems version 1.61 data analysis software. The computer generated assignments were used as an aid to manual determination and as confirmation of the manual assignment. When there was a discrepancy between the manual assignment and the computer assignment of an amino acid, either the manual assignment was taken to be correct or no assignment was made. Cysteine was identified as the PTH derivative of *S*-pyridylethylcysteine (Friedman *et al.*, 1970) using *S*-pyridylethylated lysozyme (Cys at residue 6), as a standard. *S*-pyridylethylated lysozyme was prepared by the method of Amons (1987).

2.2 RESULTS

2.2.1 Recovery of *S*-Pyridylethylated ADH 1503

When ADH 1503 was denatured in guanidine hydrochloride, reduced and alkylated, the *S*-pyridylethylated protein was desalted on a 30 x 2.1 mm reverse-phase column. Salts were removed by washing the column with 0.1% (v/v) trifluoroacetic acid (6 min) and protein was then eluted from the column by applying a linear gradient of acetonitrile [0 - 77.77% (v/v) in 4 min] followed by a wash with 77.7% (v/v) acetonitrile for a further 4 min. In a typical desalting 200 µg of protein were eluted in a final volume of 1 ml and collected as a single fraction (Figure 2.2.1).

When the acetonitrile was removed by freeze-drying to complete dryness, the *S*-pyridylethylated protein was completely insoluble in aqueous solutions of low ionic strength and was unsuitable for cleavage with proteases. Desalted *S*-pyridylethylated protein that had been concentrated to approximately 3 mg ml^{-1} by partial freeze-drying precipitated when the solution was adjusted to 50 mM NH_4HCO_3 pH 7.8 or 50 mM Tris-HCl pH 7.8. The problem was solved by partially freeze-drying the *S*-pyridylethylated protein to approximately 2 mg ml^{-1} after reverse-phase chromatography, adjusting the solution to 4 M urea by addition of 10 M urea and then adjusting the buffer concentration to 50 mM Tris-HCl pH 7.8 by addition of the appropriate volume of 2 M Tris-HCl pH 7.8. *S*-pyridylethylated protein that was maintained in solution in the aforementioned manner was suitable for cleavage by V8 protease.

In an alternative denaturation procedure, urea was used as the denaturant and the *S*-pyridylethylated protein was recovered by precipitation from solution in the presence of trichloroacetic acid. *S*-pyridylethylated protein that had been precipitated from trichloroacetic acid solution proved to be equally insoluble in dilute aqueous solution with that which had been freeze-dried but was soluble in 70% (v/v) formic acid and was, therefore, suitable for cleavage by CNBr.

2.2.2 Cleavage of ADH 1503 with V8 Protease, Reverse-Phase Chromatography of the Resulting Peptides

When *S*-pyridylethylated ADH 1503 was incubated with V8 protease, in the presence of urea (4 M) and 50 mM Tris-HCl pH 7.8, a precipitate was observed in the cleavage mixture 1 - 2 h after the addition of V8 protease. When this particulate material was removed by centrifugation and the cleared supernatants were analysed by reverse-phase h.p.l.c. no protein, or at most 10% (w/w) of the starting material, was recovered. An estimated 100% (w/w) of the starting material was recovered as a complex mixture of peptides when the particulate material was solubilised by addition of 10% (v/v) trifluoroacetic acid (Figure 2.2.2). Cleavage of ADH 1503 was

complete after 4 h, as judged by comparison of the reverse-phase h.p.l.c. elution profiles of the peptides resulting from incubation with V8 protease for 4 h and 24 h.

Peptides that co-eluted as a single fraction, for example Fractions B and C in Figure 2.2.2, were subjected to further chromatography under a variety of conditions of solvent flow rate and acetonitrile gradient. Fraction B resolved into three peaks when re-chromatographed on a 30 x 2.1 mm Aquapore RP-300 octylsilica column with an acetonitrile gradient of 15.55 - 38.88% (v/v) in 36 min. In all cases rechromatography was hampered by the decrease in yield, as much as 50% (w/w), encountered each time a fraction was subjected to further chromatography.

2.2.3 Cleavage of ADH 1503 with CNBr, Reverse-Phase Chromatography of the Resulting Peptides.

Peptides generated by cleavage of *S*-pyridylethylated ADH 1503 with CNBr remained in solution over the entire time-scale of the cleavage and when the formic acid concentration was reduced to approximately 7% (v/v) prior to reverse-phase h.p.l.c. However, freeze-dried peptide mixtures were completely insoluble in 0.1% (v/v) trifluoroacetic acid and were dissolved in 4 M guanidine hydrochloride, 0.1% (v/v) trifluoroacetic acid prior to chromatography.

A typical elution profile obtained when a CNBr cleavage mixture was analysed by reverse-phase h.p.l.c. is presented in Figure 2.2.3. Analyses of aliquots of the reverse-phase h.p.l.c. fractions by Tricine-SDS-PAGE indicated that Fraction E was homogeneous and contained a peptide with an apparent M_r of 1000 (1000-M_r CNBr peptide). Fraction F was also homogeneous and contained a 20,000-M_r CNBr peptide. Fraction G contained a mixture of uncleaved protein and a 14,000-M_r CNBr peptide.

2.2.4 Amino-Acid Sequencing of Peptides of ADH 1503

The V8 protease peptides Fractions A and D (Figure 2.2.2), the 1000-M_r CNBr peptide and the 20,000-M_r CNBr peptide were sequenced directly from the reverse-phase h.p.l.c. fractions. The 14,000-M_r CNBr peptide and uncleaved protein, Fraction G, were resolved by Tricine-SDS-PAGE, the 14,000-M_r peptide was electroblotted onto Immobilon and sequenced from the blot. The amino-acid sequences obtained from the peptides are presented in Table 2.2.1. The data obtained for those peptides that yielded a single sequence were unambiguous over the amino-acid residues presented. The first sequencing cycle of the V8 protease peptide Fraction D did not yield a unique N-terminal amino acid. However, subsequent sequencing cycles did yield a major sequence. The computer-generated raw-data tabulation from the sequencing of Fraction D is presented in Figure 2.2.4.

2.2.5 N-Terminal Amino-Acid Sequencing of ADH 1503 and ADH 2334

Freeze-dried proteins (ADH 1503 and ADH 2334; Section 2.1.1) were dissolved in 50% (v/v) trifluoroacetic acid (~mg ml^{-1}) immediately prior to sequencing. Molar amounts of both proteins were estimated assuming a subunit M_r of 35,000 as determined by Sheehan *et al.* (1988). A thirty-seven-residue N-terminal sequence (including one unknown) was determined for ADH 1503 and a forty-residue N-terminal sequence (including two unknowns) was determined for ADH 2334. The two N-terminal amino-acid sequences are aligned in Table 2.2.1. The computer-generated raw data from the sequence analysis of ADH 1503 are presented in Figure 2.2.5. The N-terminal sequence of ADH 2334 was determined using the data presented in Figure 2.2.6 and the results of a second, confirmative sequence analysis (not shown). The data for both proteins are unambiguous and clearly show that both proteins had been purified to homogeneity.

Figure 2.2.1 **Desalting *S*-Pyridylethylated ADH 1503 by Reverse-Phase h.p.l.c.**

ADH 1503 was denatured with guanidine hydrochloride, and *S*-pyridylethylated (Section 2.1.2.2). Reverse-phase h.p.l.c. was as described in Section 2.1.3. The *S*-pyridylethylation reaction mixture (170 μg in 100 μl) was applied to the column at elution time 0 min, the column was washed with 0.1% (v/v) trifluoroacetic acid for 6 min, and then with a linear gradient of increasing concentration of acetonitrile [0 - 77.77% (v/v) in 4 min; - - -], followed by 77.77% (v/v) acetonitrile for a further 4 min. Column eluent was monitored at 208 nm (——). Salt eluted as a broad front between 0 - 6 min. *S*-Pyridylethylated ADH 1503, which eluted as a sharp peak at 12 - 17 min, was collected manually as a single fraction (~0.8 ml).

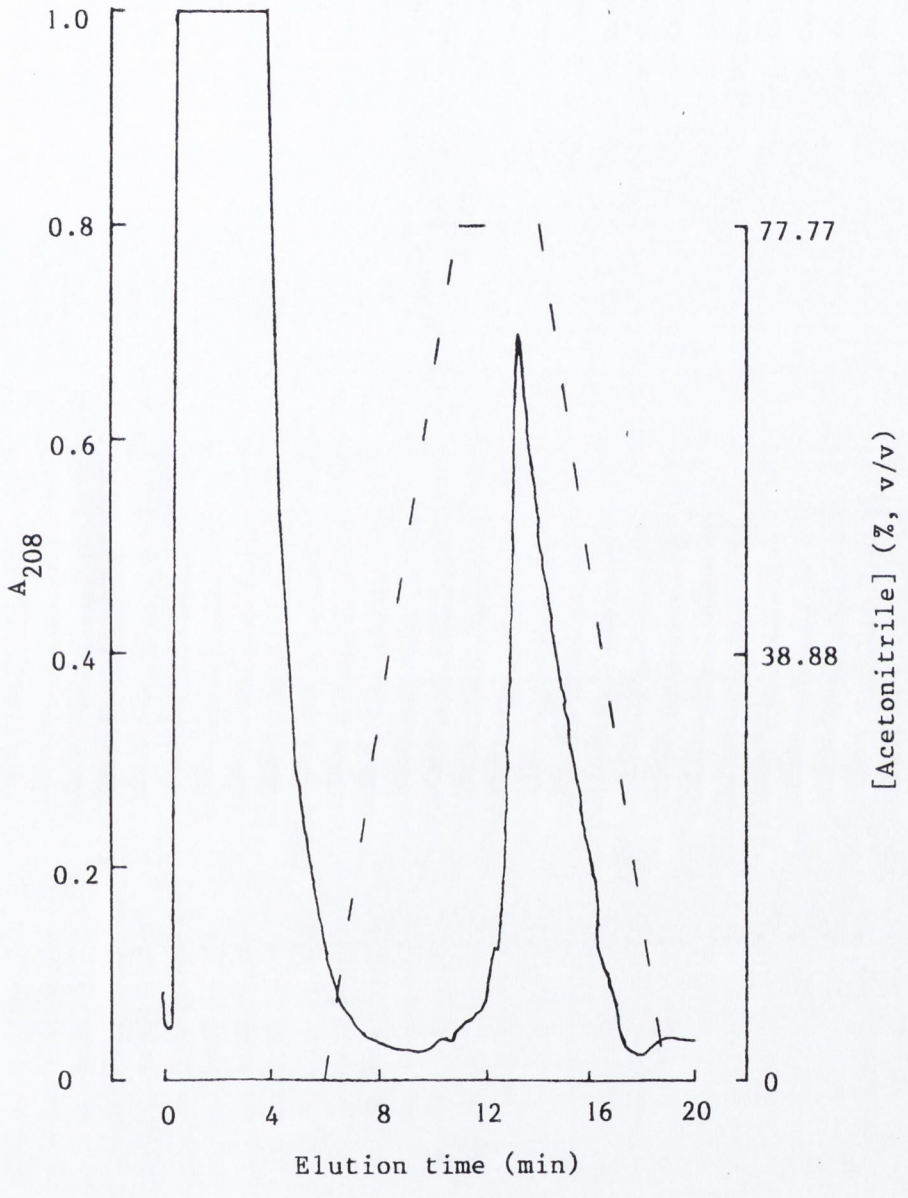

1.0
0.8
0.6
0.4
0.2
0
A_{208}
77.77
38.88
0
[Acetonitrile] (%, v/v)
0
4
8
12
16
20
Elution time (min)

Figure 2.2.1

Figure 2.2.2 **Separation of V8 Peptides of *S*-Pyridylethylated ADH 1503 by Reverse-Phase h.p.l.c.**

S-Pyridylethylated ADH 1503 was cleaved with V8 protease for 6 h (Section 2.1.2.5). An aliquot of the reaction mixture, equivalent to 500 μg of protein, was adjusted to 0.1% (v/v) trifluoroacetic acid and to 0.73 M urea before chromatography (Section 2.1.3.4). Reverse-phase h.p.l.c. was as described in Section 2.1.3. The peptide mixture was applied to the column at elution time 0 min. The column was washed with 0.1% (v/v) trifluoroacetic acid for 5 min, and then with a linear gradient of increasing concentration of acetonitrile [0 - 77.77% (v/v) for 60 min; - - -]. Column eluent was monitored at 208 nm (——), and collected as fractions that corresponded to individual peaks in the elution profile. Fractions discussed in the text (Sections 2.2.2 and 2.2.4) are indicated by the letters A, B, C and D.

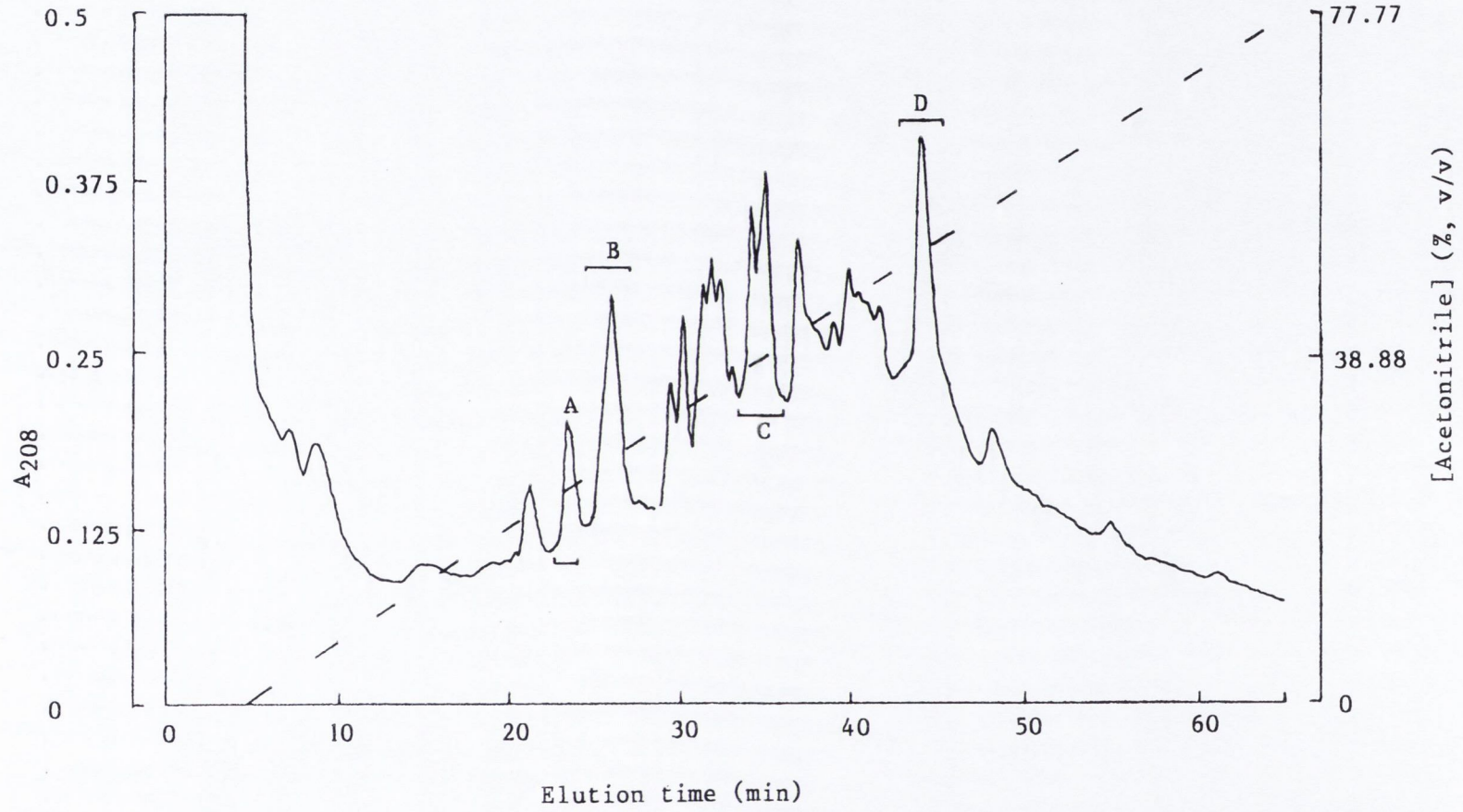

0.5
0.375
0.25
0.125
0
A208
77.77
38.88
0
[Acetonitrile] (%, v/v)
A
B
C
D
0
10
20
30
40
50
60
Elution time (min)

Figure 2.2.2

Figure 2.2.3 **Separation of CNBr Peptides of *S*-Pyridylethylated ADH 1503 Reverse-Phase by h.p.l.c.**

S-Pyridylethylated ADH 1503 was cleaved with CNBr and stored as a freeze-dried powder (Section 2.1.2.4). Reverse-phase h.p.l.c. was as described in Section 2.1.3. Freeze-dried cleavage mixture (~200 μg) was solubilised in 4 M guanidine hydrochloride, 0.1% (v/v) trifluoroacetic acid, and applied to the column at elution time 0 min. The column was washed with 0.1% (v/v) trifluoroacetic acid for 2 min, and then with a linear gradient of increasing concentration of acetonitrile [0 - 77.77% (v/v) in 30 min; - - -]. Column eluent was monitored at 208 nm (——), and collected as fractions that corresponded to individual peaks in the elution profile. Fractions discussed in the text (Section 2.2.3 and Section 2.2.4) are indicated by the letters E, F and G.

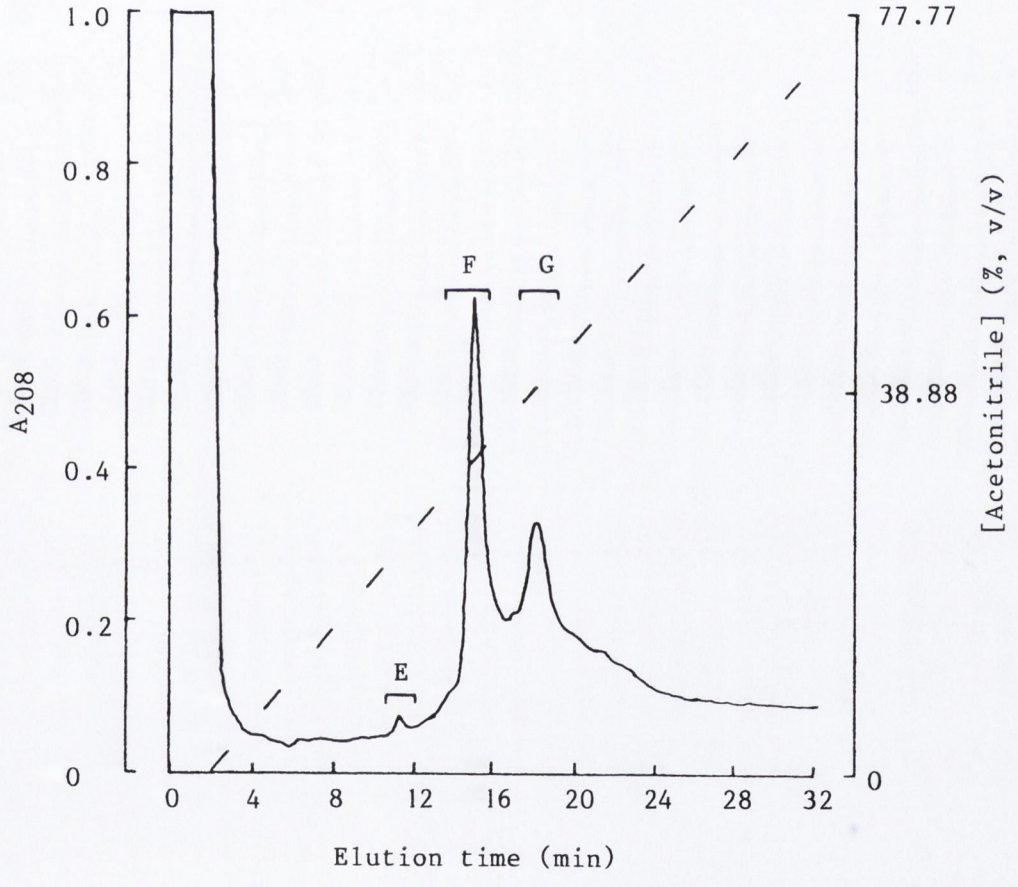

1.0
0.8
0.6
0.4
0.2
0
A208
77.77
38.88
0
[Acetonitrile] (%, v/v)
F
G
E
0
4
8
12
16
20
24
28
32
Elution time (min)

Figure 2.2.3

Protein or ADH 1503 peptide	Amino-acid sequence(s)
ADH 2334 N-terminus	1 MKAAVVNEFKKALEIKEVERPKLEEGEVLVKIEA-GV-HT 40
ADH 1503 N-terminus	1 MKAAVVEQFKKPLQVKEVEKPKISYGEVLVRIKA-GV 37
V8 protease Fraction A (Figure 2.2.2)	(1) MKAAVV (2) QFKKPLQ
20,000-M_r CNBr Fraction F (Figure 2.2.3)	(1) MKAAVV (2) KAAVVE
V8 protease Fraction D (Figure 2.2.2)	-AYKSIRRGGACVLVGLP-E
1000-M_r CNBr peptide, Fraction E (Figure 2.2.3)	LKGQINGRVVLKV
14,000-M_r CNBr peptide isolated from Fraction G (Figure 2.2.3) by Tricine-SDS-PAGE and recovered by electroblotting onto Immobilon	GLNVVAVDLG

Table 2.2.1 **The Amino-Acid Sequence Data Obtained for ADH 2334 and ADH 1503**

The notation - indicates that an assignment could not be made on the basis of the available data.

Figure 2.2.4 **Amino-acid Sequence Analysis of *S*-Pyridylethylated ADH 1503 V8 Protease Fraction D**

The data are that obtained when Fraction D (Figure 2.2.2) was amino-acid sequenced without further purification.

1. A facsimile of the computer-generated raw-data tabulation. The amino-acid sequence can be determined from the tabulation by: first, identifying the amino acid present in each sequencing cycle at a molar-yield significantly above that of all other amino acids in that cycle; second, confirming that the molar-yield of the selected amino acid is above the background values for that particular amino acid in other cycles. Molar amounts that are significantly above background are highlighted in each cycle. The computer analysis was not calibrated to detect *S*-pyridylethylated cysteine.

2. The amino-acid sequence, compiled by combining the data obtained from the recorder traces (not shown) with those obtained from the computer-generated raw-data tabulation. Cysteine in cycle 12 was identified in the recorder traces by comparison of the retention time of the significant peak with that of the predetermined retention time of *S*-pyridylethylated cysteine (Section 2.1.6). The notation - indicates that it was not possible to make an assignment on the basis of the available data.

1.

– Applied Biosystems Tabulation Report –

SAMPLE : RUN125 GLYNIS RV8P17

Sample Amount : 1000 pmol
Date/Time of Run : 13 February 1991 1:09pm
Cycles Run : 1 STD, 30 AA

Raw Data Tabulation ... (in pmols)

Cycle	A ALA	R ARG	N ASN	D ASP	C CYS	E GLU	Q GLN	G GLY	H HIS	I ILE	L LEU	K LYS	M MET	F PHE	P PRO	S SER	T THR	W TRP	Y TYR	V VAL
1	131.3	4.9	1.8	128.8	0.0	6.0	33.1	24.9	3.3	11.0	27.4	4.9	22.1	5.2	6.5	165.8	39.0	59.9	4.8	88.9
2	390.1	0.0	3.1	80.3	0.0	17.6	19.2	46.1	2.1	79.0	25.3	5.7	4.4	29.3	5.0	31.9	13.5	35.3	3.5	35.6
3	82.5	10.3	2.9	51.9	0.0	11.2	14.6	59.8	1.0	24.3	34.8	54.9	1.1	10.7	79.5	24.7	13.7	29.8	197.0	83.2
4	71.0	4.0	28.1	29.6	0.0	14.3	17.3	60.9	3.5	101.0	0.0	368.5	2.6	11.7	75.1	30.2	23.7	25.0	30.1	41.3
5	80.5	3.6	22.7	29.4	0.0	16.8	6.1	60.7	2.1	56.5	44.3	53.2	2.1	78.4	51.1	292.1	22.8	24.8	12.9	52.3
6	52.3	9.9	17.7	16.2	0.0	9.2	8.8	50.0	1.8	319.7	57.5	0.0	2.3	37.9	21.7	35.7	53.4	29.5	7.9	52.3
7	119.2	62.0	20.0	23.6	0.0	23.4	27.1	53.0	3.5	74.8	51.1	16.3	4.0	18.3	18.5	39.5	41.8	24.3	15.6	38.7
8	86.3	79.0	17.4	14.3	0.0	25.0	41.5	111.8	4.1	39.4	47.8	20.9	6.7	12.8	19.0	39.3	42.0	26.8	21.4	62.0
9	60.5	26.0	12.2	19.6	0.0	16.7	26.2	293.0	4.3	38.2	49.1	39.9	3.0	21.9	15.7	23.5	51.5	27.6	15.7	120.9
10	47.2	20.7	7.5	15.4	0.0	20.0	18.5	236.4	2.6	26.0	50.0	28.1	5.1	22.5	13.1	18.9	67.1	19.7	8.3	88.6
11	283.1	15.3	13.4	25.4	0.0	27.1	24.8	112.7	3.6	32.8	47.2	38.0	3.9	24.1	24.3	22.8	113.2	30.8	13.6	103.8
12	98.0	9.0	9.3	11.5	0.0	20.5	16.2	55.9	5.3	31.6	51.4	19.4	3.1	17.3	19.4	11.5	61.7	23.6	22.0	68.5
13	71.5	22.9	10.4	24.8	0.0	29.4	16.7	65.4	3.8	29.7	59.1	55.4	3.6	14.8	25.7	22.3	53.2	26.2	20.6	277.1
14	91.7	18.0	12.8	24.4	0.0	28.5	23.0	65.1	1.0	29.9	250.5	44.1	7.3	13.8	35.6	25.6	59.8	29.4	17.2	92.6
15	69.2	18.9	13.2	21.9	0.0	22.8	19.6	64.5	1.8	37.8	113.6	40.6	5.3	12.2	29.2	31.2	47.9	30.9	13.5	235.3
16	61.5	15.7	13.7	22.7	0.0	21.2	18.6	193.4	3.5	42.6	67.0	70.9	4.6	12.2	28.6	26.8	39.4	29.7	12.9	87.0
17	63.9	11.7	10.7	19.1	0.0	18.9	13.9	94.6	0.2	40.1	201.7	44.9	3.4	13.6	23.4	26.8	32.3	25.1	12.0	80.2
18	65.9	18.0	12.4	19.3	0.0	22.6	17.6	78.7	4.3	35.2	78.2	37.2	4.3	13.2	143.4	34.0	69.7	29.3	16.4	65.6
19	39.7	10.3	4.5	6.2	0.0	10.1	9.2	58.9	2.3	25.5	40.7	12.1	2.1	8.9	64.7	18.9	22.6	24.0	6.9	37.0
20	76.7	16.6	11.4	12.7	0.0	59.6	16.5	84.0	1.0	32.1	50.5	20.4	3.4	11.7	57.3	26.6	35.2	20.8	16.7	51.5
21	11.8	2.7	4.9	5.9	0.0	32.4	9.8	46.7	2.5	23.8	40.9	18.0	2.8	9.3	20.2	12.0	13.7	26.9	8.5	43.1
22	63.3	13.0	10.5	22.0	0.0	41.4	14.9	73.9	5.4	180.2	51.9	20.3	4.3	11.7	40.9	19.2	24.3	25.7	13.5	56.1
23	38.6	0.4	5.6	8.4	0.0	12.0	8.7	50.2	3.0	48.3	36.2	8.8	3.3	11.9	40.5	8.6	18.2	26.9	7.4	44.9
24	46.0	18.0	9.5	15.7	0.0	21.7	11.9	68.2	1.3	79.3	39.1	11.3	2.5	13.3	45.2	22.5	27.1	18.0	11.0	50.3
25	50.9	2.2	9.2	11.2	0.0	15.5	12.9	63.7	4.3	66.6	49.6	12.0	2.9	12.2	51.5	18.7	26.1	23.5	12.0	48.3
26	49.5	11.7	9.0	13.0	0.0	16.2	12.7	67.5	4.6	72.7	45.6	14.6	2.8	11.9	44.9	18.9	22.0	23.8	10.8	52.1
27	55.7	15.3	7.8	9.0	0.0	10.2	11.7	57.6	1.2	57.4	40.8	12.2	2.1	30.4	26.3	17.0	21.9	21.8	8.4	43.3
28	64.8	18.4	10.6	25.4	0.0	16.6	13.6	70.1	1.3	57.9	42.4	15.6	2.8	35.3	30.8	25.2	23.0	21.2	13.1	46.8
29	43.3	10.3	7.3	14.9	0.0	10.6	11.3	51.2	2.5	35.6	33.5	11.9	2.6	22.0	17.6	15.1	39.6	17.1	10.7	36.3
30	45.1	3.1	6.8	10.6	0.0	9.2	11.1	64.1	3.5	32.9	36.5	11.2	2.3	16.1	16.4	8.6	34.4	17.8	11.6	55.5

2. -AYKSIRRGGACVLVGLP-E

Figure 2.2.4

Figure 2.2.5 **N-Terminal Amino-Acid Sequence-Analysis of ADH 1503**

Freeze-dried desalted ADH 1503 (Section 2.1.1) was dissolved in 50% (v/v) trifluoroacetic acid, to yield a protein concentration of 1 mg ml^{-1} (28 nmol ml^{-1}). Automated amino-acid sequence-analysis (Section 2.1.6) was performed on a volume equivalent to 500 pmol of protein.

1. A facsimile of the computer-generated raw-data tabulation in which cycle numbers correspond to amino acid positions, commencing at the N-terminal amino acid of the protein. The amino-acid sequence can be determined from the tabulation as described in the legend to Figure 2.2.4.

2. The amino-acid sequence, compiled by combining the data obtained from the recorder traces (not shown) with those obtained from the computer-generated raw-data tabulation. The notation - indicates that it was not possible to make an assignment on the basis of the available data.

1.

– Applied Biosystems Tabulation Report –

SAMPLE : RUN152 GLYNIS ADH1503

Sample Amount : 500 pmol
Date/Time of Run : 12 November 1991 1:13pm
Cycles Run : 1 STD, 37 AA

Raw Data Tabulation ... (in pmols)

Cycle	A ALA	R ARG	N ASN	D ASP	C CYS	E GLU	Q GLN	G GLY	H HIS	I ILE	L LEU	K LYS	M MET	F PHE	P PRO	S SER	T THR	W TRP	Y TYR	V VAL
1	92.4	0.6	1.7	0.8	0.0	4.1	3.4	17.4	1.6	5.7	6.5	3.0	503.8	3.4	3.6	30.7	33.7	17.7	1.9	9.3
2	16.6	1.2	3.0	2.7	0.0	5.3	4.0	15.4	0.1	59.8	14.0	198.4	29.2	6.7	4.0	11.8	6.7	19.6	2.2	45.5
3	368.8	1.5	3.3	2.2	0.0	5.9	3.8	17.1	0.7	27.7	21.0	38.6	4.8	7.2	5.9	10.7	11.7	13.4	4.7	63.1
4	353.3	1.2	11.5	3.6	0.0	6.6	4.6	18.2	0.1	37.9	30.4	8.4	4.6	8.5	6.7	10.8	13.9	11.9	5.3	75.0
5	85.5	1.6	5.7	3.0	0.0	10.1	5.1	43.7	0.3	31.6	38.6	13.3	10.4	16.2	7.8	13.4	14.6	12.2	5.9	100.0
6	37.2	0.8	3.8	3.7	0.0	8.1	5.2	25.8	1.3	63.0	41.1	9.1	5.5	11.9	8.5	11.0	29.9	12.7	5.1	392.9
7	51.7	1.9	18.1	12.3	0.0	99.6	11.8	39.7	0.3	41.5	53.4	18.8	5.9	12.2	11.8	17.7	24.1	15.0	8.2	101.9
8	42.8	2.5	6.1	7.7	0.0	45.2	139.9	42.8	1.3	36.7	50.8	15.7	5.8	12.2	11.8	15.1	17.9	13.7	7.1	66.9
9	54.0	2.5	7.7	9.4	0.0	23.2	33.5	38.9	0.5	48.5	66.9	21.2	5.7	304.3	14.9	16.3	36.2	14.4	9.4	68.4
10	53.6	3.0	9.4	10.8	0.0	17.7	17.4	53.0	0.8	39.3	62.6	151.7	5.6	60.2	14.0	15.8	28.4	18.6	8.6	70.7
11	87.9	7.3	10.3	11.7	0.0	19.0	21.3	54.6	0.7	45.4	82.2	191.8	7.3	26.9	18.6	15.6	27.6	20.7	11.1	83.0
12	79.9	4.9	8.7	10.9	0.0	18.1	17.3	42.1	1.4	71.1	71.0	37.7	6.2	22.7	109.8	19.7	25.8	19.9	10.8	74.7
13	84.6	6.7	11.4	15.5	0.0	25.7	21.3	69.3	1.6	58.0	250.8	29.7	6.7	26.2	72.1	24.2	28.6	21.3	12.7	82.8
14	87.3	7.2	11.7	14.0	0.0	34.3	97.0	57.7	1.2	57.0	159.5	25.8	6.3	27.1	40.6	23.5	37.8	16.6	12.5	85.9
15	93.9	7.0	21.0	15.8	0.0	29.7	63.3	60.4	1.3	56.9	115.4	33.9	7.2	25.7	32.0	23.5	33.5	16.3	13.2	243.0
16	96.0	5.8	13.3	18.2	0.0	23.8	36.5	62.1	1.5	56.3	98.8	100.9	6.9	31.9	27.5	26.4	30.6	20.6	13.7	186.5
17	95.5	7.8	13.2	18.4	0.0	49.1	29.0	69.9	1.0	54.4	89.2	64.9	7.3	48.0	26.8	27.3	32.2	21.8	16.5	123.6
18	94.9	8.2	13.6	17.3	0.0	35.1	25.0	65.9	1.4	60.7	94.6	48.5	7.9	31.8	24.8	27.2	32.3	16.7	15.5	225.3
19	96.8	6.5	10.4	13.9	0.0	35.9	18.5	54.5	0.7	56.5	82.0	31.0	8.8	27.6	21.4	21.0	31.3	14.7	13.8	145.7
20	106.4	7.5	12.8	17.3	0.0	31.2	20.6	61.1	1.9	59.0	92.2	80.9	8.8	28.0	25.2	26.9	30.9	20.6	16.1	119.7
21	95.8	7.4	12.2	17.1	0.0	23.7	20.9	65.2	1.4	61.7	102.2	56.1	8.6	28.6	66.1	24.8	30.7	16.9	16.8	113.6
22	101.8	9.8	13.3	20.8	0.0	23.2	22.2	70.5	2.2	63.7	92.1	94.6	8.7	28.5	65.6	28.8	36.0	19.1	19.1	109.8
23	88.9	10.2	21.0	24.7	0.0	22.8	22.4	66.6	2.8	122.6	79.9	56.8	8.1	32.0	49.9	26.2	33.2	32.0	19.4	94.8
24	105.2	8.0	15.6	17.9	0.0	21.0	23.1	81.3	1.6	125.4	90.5	49.7	9.5	30.0	47.3	61.8	36.1	16.3	19.8	109.9
25	103.1	10.7	14.3	20.6	0.0	20.8	24.6	76.9	2.8	98.0	88.3	38.3	8.6	27.6	36.3	54.4	34.7	15.0	44.1	102.9
26	101.2	10.6	14.5	19.4	0.0	20.0	24.5	112.9	1.4	91.2	86.7	37.2	9.7	27.4	36.5	41.4	32.8	14.8	40.6	103.9
27	95.5	9.4	13.9	18.1	0.0	30.2	23.2	102.7	2.7	76.4	85.4	31.8	9.2	28.9	31.2	33.7	35.5	12.8	29.8	99.5
28	95.9	7.9	13.5	16.3	0.0	26.9	23.1	91.7	2.2	70.1	87.2	30.6	9.7	27.6	30.2	26.4	33.8	13.1	23.7	148.0
29	93.6	6.5	11.1	12.0	0.0	19.0	19.4	82.5	0.0	67.1	121.7	26.1	9.6	26.3	26.5	25.4	29.3	12.3	18.7	146.6
30	104.4	9.3	14.1	17.4	0.0	21.7	22.6	83.5	1.4	69.9	128.0	28.7	9.1	27.3	32.5	26.3	33.6	14.3	19.6	167.9
31	99.6	14.7	14.7	16.8	0.0	20.2	23.7	81.3	1.9	69.3	104.7	29.5	8.9	26.8	31.3	26.6	33.5	13.7	19.8	156.1
32	99.0	12.3	16.8	16.2	0.0	18.6	23.0	80.1	1.8	104.4	95.7	26.1	9.2	26.5	31.9	24.9	33.5	14.5	19.5	139.7
33	101.5	14.4	16.7	20.0	0.0	19.9	24.3	80.7	1.4	103.1	91.1	45.7	10.5	26.9	32.8	28.4	35.5	18.0	20.9	127.7
34	138.8	11.6	17.4	20.4	0.0	20.7	26.2	86.4	1.9	94.6	96.2	45.6	11.1	27.8	35.9	29.8	40.7	15.6	22.0	126.9
35	138.6	10.1	17.7	21.1	0.0	20.4	26.0	85.6	3.1	85.5	94.4	42.8	10.7	28.5	36.3	29.3	42.2	14.4	22.7	126.3
36	106.7	8.0	14.9	16.7	0.0	17.2	22.1	92.5	0.4	70.9	79.1	30.0	9.6	24.8	29.4	22.2	33.2	14.3	19.2	105.5
37	114.1	10.4	16.9	18.3	0.0	19.3	24.9	104.3	1.8	73.1	81.9	32.6	10.3	26.3	34.7	25.3	37.2	13.8	21.1	138.6

2. MKAAVVEQFKKPLQVKEVEKPKISYGEVLVRIKA-GV

Figure 2.2.5

Figure 2.2.6 **N-Terminal Amino-Acid Sequence-Analysis of ADH 2334**

Freeze-dried desalted ADH 2334 (Section 2.1.1) was dissolved in 50% (v/v) trifluoroacetic acid to yield a final concentration of 2.8 mg ml^{-1} (78.4 nmol ml^{-1}). Automated amino-acid sequence-analysis (Section 2.1.6) was performed on a volume equivalent to 100 pmol of protein.

1. A facsimile of the computer-generated raw-data tabulation, from one of two sequence determinations.
2. The amino-acid sequence, compiled by combining all of the data obtained from two seperate sequence determinations. The notation - indicates that it was not possible to make an assignment on the basis of the available data. All other details were as described in the legend to Figure 2.2.5.

1.

- Applied Biosystems Tabulation Report -

SAMPLE : 03. C.BAILEY. ADH.

Sample Amount : 100 pmol
Date/Time of Run : 9 June 1988 10:23pm
Cycles Run : 1 STD, 39 AA

Raw Data Tabulation ... (in pmols)

Cycle	A ALA	R ARG	N ASN	D ASP	C CYS	E GLU	Q GLN	G GLY	H HIS	I ILE	L LEU	K LYS	M MET	F PHE	P PRO	S SER	T THR	W TRP	Y TYR	V VAL
1	7.8	13.9	0.3	14.6	0.0	1.6	0.3	2.6	0.1	0.9	1.1	0.9	56.4	0.4	0.9	11.8	36.5	5.8	0.6	1.0
2	4.1	4.2	0.4	7.6	0.0	2.0	0.3	3.4	0.2	1.0	0.7	17.7	1.4	0.5	1.3	4.3	3.9	8.1	0.6	1.3
3	57.1	22.2	0.4	6.8	0.0	3.3	1.9	3.9	0.1	1.9	3.9	5.1	0.4	0.8	1.5	3.3	2.0	10.5	1.2	2.5
4	44.6	61.1	1.0	6.9	0.0	2.1	0.9	4.0	0.1	2.0	3.0	2.5	0.2	0.6	1.5	4.0	1.7	18.0	1.2	3.1
5	8.3	52.8	0.9	6.6	0.0	3.0	0.9	5.4	0.2	2.8	7.8	4.6	0.6	1.4	2.2	3.2	1.5	11.9	1.8	55.0
6	4.5	40.3	0.6	4.7	0.0	2.3	1.1	5.0	0.1	2.7	4.7	3.0	0.6	3.5	1.8	4.5	1.3	17.4	1.4	38.2
7	6.5	79.2	33.8	9.5	0.0	4.2	3.0	7.3	0.3	4.8	6.6	5.4	0.9	1.1	2.9	4.1	2.6	13.6	1.8	7.9
8	7.3	108.3	3.7	6.9	0.0	30.2	2.0	7.8	0.3	5.1	8.1	5.8	1.0	1.2	3.6	5.7	2.3	16.6	2.0	7.8
9	6.8	91.7	2.0	4.7	0.0	4.7	1.9	7.6	0.3	3.4	7.5	5.5	1.2	32.6	4.1	5.1	2.9	8.2	1.6	7.2
10	9.1	63.9	2.8	5.4	0.0	4.3	2.1	9.3	0.4	4.9	8.0	37.9	1.5	3.3	4.0	5.3	2.6	7.7	1.8	7.3
11	9.9	126.4	2.7	6.7	0.0	5.8	2.7	10.7	0.6	5.1	9.2	51.4	1.2	1.8	5.1	7.1	3.4	10.1	2.5	9.8
12	45.0	119.4	2.4	7.5	0.0	5.4	4.0	10.7	0.5	4.9	9.3	8.9	1.2	2.0	5.3	6.4	3.2	16.4	3.3	9.3
13	8.6	61.1	1.3	4.0	0.0	2.9	1.7	7.9	0.2	4.2	28.4	3.9	1.3	1.7	3.1	4.6	1.7	9.3	1.5	6.4
14	13.5	163.9	3.1	7.5	0.0	28.3	4.2	12.9	0.6	5.6	13.0	7.4	1.3	2.4	6.9	6.9	3.3	10.6	2.7	9.4
15	11.3	144.4	2.8	7.6	0.0	7.4	3.2	12.9	0.6	32.7	10.3	6.8	1.4	2.6	6.0	6.3	3.9	12.4	2.6	9.3
16	11.3	93.1	2.3	5.7	0.0	4.8	2.4	11.2	0.3	7.6	9.2	30.2	1.4	2.4	5.8	5.2	3.2	12.6	2.2	8.2
17	10.9	177.8	3.7	7.1	0.0	21.0	2.9	12.3	0.6	6.0	9.2	8.7	1.4	2.3	6.2	7.0	4.2	13.3	2.6	8.5
18	10.9	186.9	2.4	6.1	0.0	6.2	2.5	12.1	0.4	6.1	9.1	6.9	1.4	2.3	5.7	6.2	2.9	15.5	2.4	28.5
19	11.9	165.3	2.8	7.2	0.0	18.9	2.9	13.1	0.6	7.2	10.8	7.1	1.4	2.4	6.4	7.0	3.9	12.6	3.0	11.4
20	12.8	1061.1	3.1	8.3	0.0	8.0	3.1	13.5	0.6	7.5	10.7	8.2	1.6	2.5	7.4	7.2	4.7	11.2	3.9	10.6
21	9.7	159.7	2.0	5.2	0.0	3.8	2.2	10.9	0.3	6.4	8.3	5.8	1.2	2.1	16.5	5.7	3.0	12.9	2.4	8.4
22	11.0	159.7	2.5	5.7	0.0	4.9	2.6	12.0	0.4	7.6	8.7	19.7	1.4	2.3	8.0	6.2	3.4	12.1	2.8	9.0
23	12.1	155.6	2.8	6.4	0.0	5.2	2.9	13.2	0.4	6.9	23.8	9.9	1.4	2.4	6.8	6.4	3.5	11.6	3.2	9.6
24	9.4	120.8	2.4	5.4	0.0	18.1	2.4	11.8	0.4	6.0	11.0	6.2	1.1	2.0	5.9	6.2	2.9	14.7	2.6	8.1
25	10.4	169.4	3.2	6.8	0.0	12.9	2.6	12.0	0.6	6.9	9.7	6.5	1.2	2.1	5.4	5.5	3.3	14.5	3.2	8.5
26	8.6	98.6	1.8	4.5	0.0	4.1	2.0	15.5	0.2	5.3	8.7	4.2	1.2	2.1	3.6	4.7	2.3	15.6	2.1	7.5
27	8.9	102.8	1.9	5.1	0.0	7.8	2.4	13.2	0.4	5.5	8.5	4.8	1.1	2.0	4.4	4.9	2.6	12.4	2.4	7.7
28	9.6	101.4	2.1	5.5	0.0	5.4	2.4	12.6	0.4	5.9	8.3	4.8	1.1	2.1	4.6	5.0	2.7	11.4	2.5	14.2
29	7.1	48.6	1.4	3.6	0.0	2.6	1.7	9.3	0.4	4.9	12.3	3.1	1.0	1.6	3.2	4.9	1.9	10.4	1.8	7.8
30	11.4	165.3	3.0	6.7	0.0	6.5	3.3	14.3	0.7	6.7	12.0	6.6	1.3	2.2	6.4	5.4	3.3	8.9	3.3	17.3
31	10.4	131.9	2.5	6.0	0.0	4.6	2.7	12.4	0.5	6.6	9.8	11.7	1.2	2.0	5.3	5.3	2.8	13.6	2.7	10.9
32	10.3	120.8	2.5	5.4	0.0	4.7	2.7	12.3	0.5	12.1	8.9	7.5	1.1	2.2	5.3	5.7	3.2	10.9	2.7	9.4
33	9.0	73.6	1.7	3.8	0.0	5.3	2.1	10.9	0.4	7.8	8.4	4.8	1.1	1.9	4.8	4.4	2.3	9.2	2.3	8.0
34	18.5	155.6	2.8	6.6	0.0	6.1	3.1	13.8	0.6	0.0	8.2	6.0	1.3	2.1	5.1	5.7	3.3	10.9	3.2	9.4
35	12.8	155.6	2.5	5.9	0.0	3.7	2.6	12.2	0.6	8.0	9.4	7.4	1.3	2.1	5.1	4.7	3.0	10.0	3.2	9.4
36	15.7	248.6	3.6	9.5	0.0	5.9	4.2	21.4	0.9	8.2	11.0	10.4	1.6	2.6	7.4	6.7	4.0	14.3	4.2	10.8
37	10.5	118.1	2.1	4.9	0.8	3.2	2.0	12.7	0.4	6.0	7.6	5.6	1.0	1.9	4.0	4.6	2.7	8.5	2.4	10.4
38	11.3	120.8	2.3	5.7	0.0	2.1	2.4	13.3	0.6	6.8	8.8	6.4	1.6	2.2	4.7	5.2	2.5	7.5	2.7	10.1

2. MKAAVVNEFKKALEIKEVERPKLEEGEVLVKIEA-GV-HT

Figure 2.2.6

2.3 DISCUSSION

2.3.1 N-Terminal Amino-Acid Sequences of ADH 1503 and ADH 2334

The sequence determined for ADH 1503 (Table 2.2.1) is in agreement with the N-terminal amino-acid sequence originally determined by Bridgen *et al.* (1973) and presented in Figure 1 of the published result. The present analysis was performed to clarify confusion arising from an error in the text of Bridgen *et al.* (1973), and an error in the ADH 1503 N-terminal sequence entered in the Protein Identification Resource databank (PIR, Accession number A10207). The sample of ADH 1503 that was sequenced in the present work was not chemically modified to facilitate the detection of cysteine and the unknown at residue 35 corresponds to Cys-35 in the sequence determined from the carboxymethylated protein by Bridgen *et al.* (1973).

A comparison of the amino-acid composition of ADH 1503 and ADH 2334, carried out by Sheehan *et al.* (1988), indicated that the contents of only four amino-acids (His, Thr, Ala and Phe) were significantly different between the two enzymes. A high sequence identity was therefore expected between the two enzymes. The % identity over the first thirty-four residues (Table 2.2.1) is 67.6%.

2.3.2 Amino-Acid Sequences of Peptides of ADH 1503

V8 protease Fraction A was found to be, by comparison with the N-terminal sequence of ADH 1503, a mixture of the two N-terminal peptides generated by cleavage at Glu-7 (Table 2.2.1). Similarly, the 20,000-M_r CNBr peptide (Fraction F) was a mixture of the N-terminal peptide and the peptide generated by cleavage at Met-1. V8 protease Fraction D, the 1000-M_r CNBr peptide (Fraction E) and the 14,000-M_r CNBr peptide (from Fraction G) were internal peptides.

In Figure 2.3.1 the data from Table 2.2.1 are presented in the context of the amino-acid sequence deduced from the nucleotide sequence of *adh*1503 (Sakoda & Imanka, 1992). There is complete agreement between the deduced amino-acid sequence and the amino-acid sequences determined from the protein. Notably and fortuitously, the sequence of the C-terminal peptide (Fraction E) removed the possibility of any substantial error in the ADH 1503 amino-acid sequence deduced from the nucleotide sequence.

```
ADH 1503     1 MKAAVVEQFKKPLQVKEVEKPKISYGEVLVRIKACGVCHTDLHAAHGDWPVKPKLPLIPG  60
ADH 2334     1 MKAAVVNEFKKALEIKEVERPKLEEGEVLVKIEA-GV-HT                      40
               ******  *** *  **** **   ***** * * ** **

ADH 1503    61 HEGVGVIEEVGPGVTHLKVGDRVGIPWLYSACGHCDYCLSGQETLCERQQNAGYSVDGGY 120


ADH 1503   121 AEYCRAAADYVVKIPDNLSFEEAAPIFCAGVTTYKALKVTGAKPGEWVAIYGIGGLGHVA 180

                    Fraction G
ADH 1503   181 VQYAKAMGLNVVAVDLGDEKLELAKQLGADLVVNPKHDDAAQWIKEKVGGVHATVVTAVS 240

                  Fraction D
ADH 1503   241 KAAFESAYKSIRRGGACVLVGLPPEEIPIPIFDTVLNGVKIIGSIVGTRKDLQEALQFAA 300

                                  Fraction E
ADH 1503   301 EGKVKTIVEVQPLENINDVFDRMLKGQINGRVVLKVD                        337
```

Figure 2.3.1 **The Amino-Acid Sequence of ADH 1503**

The complete amino-acid sequence of ADH 1503 is that deduced from the gene sequence of Sakoda & Imanaka (1992). The N-terminal amino-acid sequence of ADH 1503, and the amino-acid sequences of internal peptides of ADH 1503, determined in this work, are underlined. Labels assigned to the peptides correspond to those used in Table 2.3.1. The N-terminal amino-acid sequence of ADH 2334 is aligned with the N-terminal sequence of ADH 1503 and identical residues are indicated by an asterisk.

CHAPTER 3
SUBCLONING AND NUCLEOTIDE SEQUENCING OF THE GENE ENCODING ADH 2334

3.1 INTRODUCTION

3.1.1 Cloning of the Gene Encoding ADH 2334

The gene encoding ADH 2334 was cloned, by Dowds *et al.* (1988), into the shuttle plasmid pRK59 (Figure 3.1.1). Plasmid pRK59 contains the marker for ampicillin resistance and the pUC18 origin of replication and can, therefore, be selected for, and maintained, in *E. coli*. The marker for chloroamphenicol resistance and the pBAA1 origin of replication allow selection and maintenance of the plasmid in *B. subtilis*. A recombinant plasmid, which carried a 4.5-kb *B. stearothermophilus* DNA insert and directed the expression of ADH 2334, was identified in *E. coli* TG1 (pADH1; Figure 3.1.2). Plasmid pADH1 was maintained, but not expressed, in *B. subtilis* SO113.

Subcloning of fragments of the 4.5-kb insert of pADH1 into pRK59 localised *adh*2334 to the 2.7-kb fragment, which extends from the *Eco*RI site in the polylinker region of the vector to the *B. stearothermophilus Bam*HI site (Figure 3.1.2). The plasmid containing this insert was named pADH1a. Inversion of the insert in pADH1a did not alter its ability to specify ADH antigenic activity, demonstrating that the gene was expressed from its own promoter rather than from a vector promoter. Subsequent exonuclease BAL 31 deletions of the 2.7-kb fragment from the *Bam*HI end revealed that a fragment approximately 1.5 kb long, extending from the *Eco*RI polylinker site to just beyond the *B. stearothermophilus* internal *Sac*I site, contained the entire gene. Following the addition of a *Bam*HI linker, the 1.5-kb fragment was subcloned into pRK59, to create the plasmid pADH1b (Figure 3.1.3).

The three plasmids, pADH1, pADH1a and pADH1b contain the *B. stearothermophilus* DNA insert downstream from the pUC18 DNA sequencing reverse priming site (Figure 3.1.3) and all three were available. The plasmid containing the smallest *B. stearothermophilus* DNA insert (pADH1b) had been shown to contain the complete *adh*2334 coding sequence and was, therefore, the starting point for the determination of the nucleotide sequence of *adh*2334.

The nucleotide sequence of *adh*2334 presented in the results section of this chapter has been published as Robinson *et al.* (1994). The GenBank accession number is Z25544.

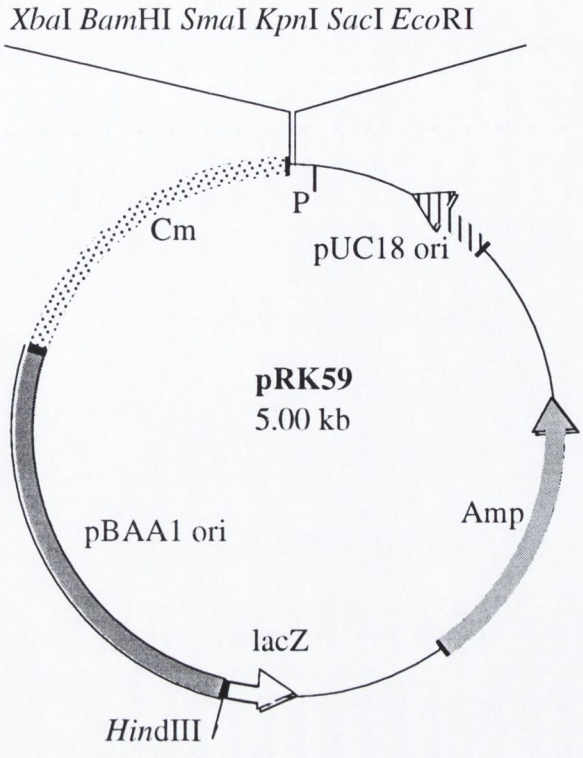


Figure 3.1.1 **Structure of the Plasmid pRK59**

The plasmid was constructed by inserting a 1 kb *Hpa*II fragment containing the chloroamphenicol resistance gene of pBD64 (Gryczan *et al.*, 1980) into the *Acc*I site of pUC18 (Yanisch-Perron *et al.*, 1985), and a 2.2 kb *Hin*dIII fragment from pBAA1 (Devine *et al.*, 1989) into the *Hin*dIII site of pUC18. Structural instability of this plasmid led to a deletion of 0.9 kb of the pBAA1 fragment, including one of the *Hin*dIII sites, however, the origin of DNA replication was retained. The notation pBAA1 ori indicates the origin of replication of pBAA1, Amp indicates the ampicillin resistance gene, pUC18 ori indicates the origin of replication of pUC18, lacZ indicates the fragment of the *lacZ* gene that encodes the N-terminus of β-galactosidase, P indicates the *lacZ* promoter, and Cm indicates the chloroamphenicol resistance gene. Arrows indicate the direction of transcription. Redrawn from Dowds *et al.* (1988).

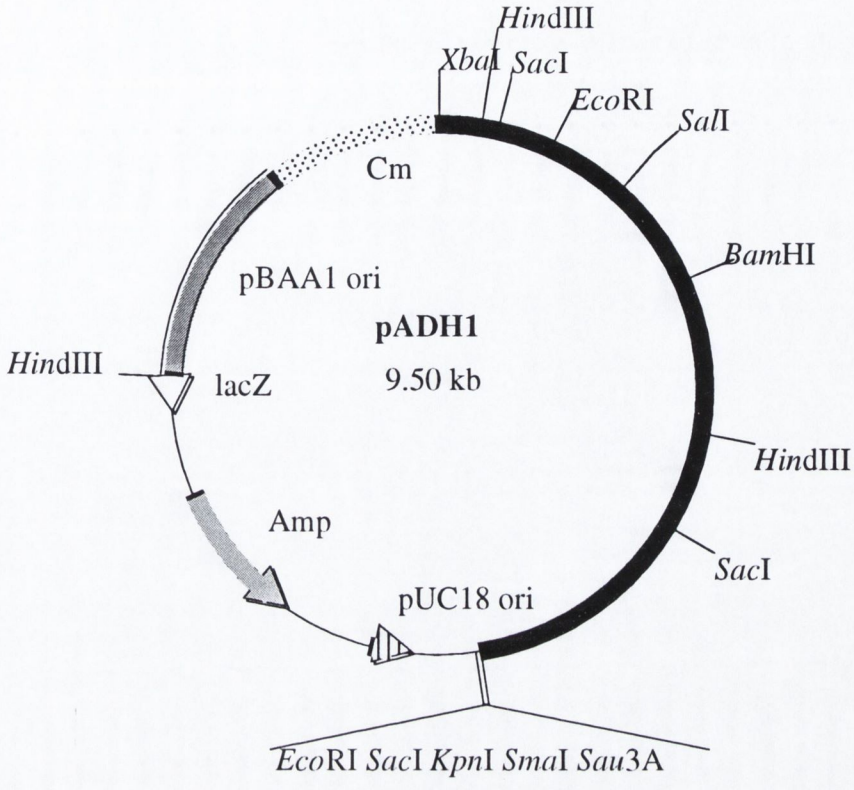


Figure 3.1.2 **Restriction Map of Plasmid pADH1**

The plasmid contains a 4.5 kb *Sau*3A fragment of *B. stearothermophilus* DNA (black shading) inserted into the *Bam*HI site of pRK59 (Figure 3.1.1). The notations used are defined in the legend to Figure 3.1.1.

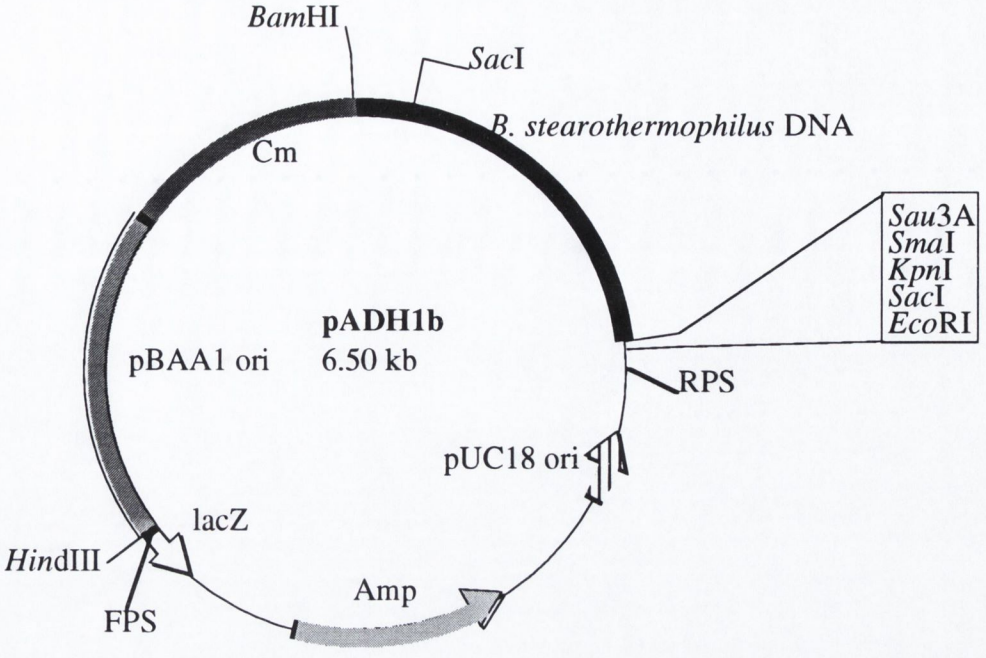


Figure 3.1.3 **Structure of the Plasmid pADH1b**

The 1.5 kb fragment (*B. stearothermophilus* DNA), which encodes ADH 2334, was derived from pADH1 (Figure 3.1.2) by exonuclease BAL 31 deletions, from the *Bam*HI end, of the 2.7 kb *Eco*RI-*Bam*HI fragment of the *B. stearothermophilus* DNA insert. Following addition of a *Bam*HI linker, the 1.5 kb fragment was inserted into pRK59 (Figure 3.1.1). The expected positions of the pUC18 forward (FPS) and reverse (RPS) priming sites are indicated. Primer-directed DNA synthesis was expected in the clockwise direction from FPS and in the anti-clockwise direction from RPS. The *B. stearothermophilus* DNA extends from (and includes) the *Sau*3A site to (but not including) the *Bam*HI site. All other details are as in the legend to Figure 3.1.3.

3.2 MATERIALS AND METHODS

3.2.1 Aseptic Techniques

Routine microbiological aseptic techniques were used when handling bacteria and DNA.

Bacteriological culture media, H_2O, culture flasks, culture tubes, centrifuge tubes and micropipette tips were sterilised by autoclaving at 120 °C and 103.5 kPa for 20 min. All solutions used in the purification and manipulation of DNA were sterilised as described above, unless sterilisation was deemed to be unnecessary (organic solvents and 10 M NaOH), or when autoclaving was deemed to be inappropriate (proteins, thiamine, antibiotics, and some solutions of sucrose and glucose). Solutions of sucrose and glucose at concentrations in excess of 15% (w/v) and solutions of enzymes were sterilised by filtration through 0.2 μm pore-size sterile syringe filters (Sigma).

Restriction enzymes, ligase, buffers supplied with these enzymes, and all solutions supplied with the DNA sequencing kit were assumed to be sterile.

3.2.2 Cloning Host, Plasmids and Bacteriophage

The cloning host, *E. coli* TG1, K12 Δ*(lac-pro)* , *sup*E, *thi* , *hsdD5/* F' *traD36* *proA*$^+$*B*$^+$, *lacI*^{q}, *lacZ*Δ*M15* (Carter *et al.*, 1985) was obtained from Amersham. The vectors, plasmids and bacteriophage used in subcloning and nucleotide sequencing are listed in Table 3.2.1. Plasmids pRK59, pADH1a and pADH1b, described in Dowds *et al.* (1988), were obtained from Barbara Dowds as solutions of purified plasmid DNA that had been stored at -20 °C. Bacteriophage M13mp18 containing *adh*2334 (mp18*adh*) was obtained from R. Kelleher as bacteriophage plaques on agar. The single-stranded form of mp18*adh* contained the sense strand of the 1.5-kb *Eco*RI-*Bam*HI fragment of *B. stearothermophilus* DNA.

3.2.3 Composition of Bacteriological Culture Media

Bacto-tryptone extract, bacto-yeast extract and bacto-agar were obtained from Difco.

LB Medium (Luria-Bertani Medium) LB medium was prepared by dissolving 10 g of bacto-tryptone extract, 5 g of bacto-yeast extract and 10 g of NaCl in 950 ml of H_2O. The pH was adjusted to 7 with 10 M NaOH before adjusting the final volume to 1 l.

2TY Medium 2TY medium was 16 g of bacto-tryptone, 10 g of bacto-yeast extract and 5 g of NaCl per 1 l of H_2O.

LB Agar Bacto-agar (15 g) was added to 1 l of LB medium prior to autoclaving.

Plasmids and bacteriophage	Reference
pRK59	Dowds *et al.* (1988) Devine *et al.* (1989) Section 3.1, Figure 3.1.1
pADH1a	Dowds *et al.* (1988) Section 3.1.1
pADH1b	Dowds *et al.* (1988) Section 3.1, Figure 3.1.3
pUC18	Yanisch-Perron *et al.* (1985)
M13mp18	Norrander *et al.* (1983)
mp18*adh*	Ms Ruth Kelleher [Senior Sophistor Project, Biochemistry Department, Trinity College, Dublin (1993)]

Table 3.2.1 **Vectors used in Nucleotide Sequencing of *adh2334***

M9 Salts M9 salts contained 6 g of Na_2HPO_4, 3 g of KH_2PO_4, 1 g of NH_4Cl and 0.5 g of NaCl per 1 l of H_2O.

M9 Minimal-Salts Agar The following reagents were sterilised separately and then mixed: agar, 1.5 g, suspended in 100 ml of M9 salts; 1 M $MgSO_4$ (0.1 ml); 50 mM $CaCl_2$ (0.2 ml); 1 M thiamine-HCl (0.1 ml) and 20% (w/v) glucose (1 ml).

Ampicillin A stock solution (200 mg ml^{-1}) was stored at -20 °C. The final concentration in media was 0.2 mg ml^{-1}.

Wrights beef infusion broth Fat-free minced ox-heart (500 g), peptone (10 g) and NaCl (5 g) were stirred into H_2O (final volume of 1l). After standing at 4 °C for 16 h the mixture was heated to 60 °C, stirred for 20 min and, finally, boiled for 20 min. Prior to autoclaving, the medium was filtered and the pH was adjusted to 7.6 - 7.8. After autoclaving, the medium was filtered again, the pH was adjusted to 7.5 and the medium was autoclaved again.

Dorset egg medium Whole chicken eggs (2 volumes) were mixed with Wrights beef infusion broth (1 volume), filtered through gauze and aliquoted into 3 ml bottles. The medium was sterilised, in a sloping position, by heating at 75° C for 30 min on 3 successive days (inspissated).

3.2.4 Storage of *E. coli* TG1

Laboratory stocks of *E. coli* TG1 and *E. coli* TG1 that had been transformed with plasmid DNA were stored, for up to one year, as streak cultures on Dorset egg slopes, at room temperature in the dark. Working stocks of transformed cells were prepared by streaking a single colony taken from a Dorset egg slope onto a LB + ampicillin agar plate, incubating the plate at 37 °C for 16 h and then storing the resulting colonies (for up to three weeks) at 4 °C. Working stocks of *E. coli* TG1 were prepared by streaking a single colony taken from a Dorset egg slope onto a LB agar plate, incubating the plate as described above and then streaking a colony taken from the LB agar onto M9 minimal-salts agar. Cultures on M9 minimal-salts agar were grown at 37 °C for 24 h and were then stored (for up to three weeks) at 4 °C.

3.2.5 Growth of *E. coli* TG1 in Liquid Culture

All liquid cultures were grown at 37 °C in an orbital shaker (250 revolutions min^{-1}). Small-scale cultures typically contained 2 ml of media inoculated with a single *E. coli* colony. Large-scale cultures contained 100 - 500 ml of media inoculated (100-fold dilution) with a small-scale culture that had been grown to stationary-phase (16 h). Large-scale cultures were grown for approximately 2 h (log-phase culture) or for 16 h (stationary-phase culture). When transformed, *E. coli* were grown in liquid medium containing ampicillin (0.2 mg ml^{-1}).

3.2.6 Storage of Bacteriophage M13

Bacteriophage M13 plaques were removed from agar plates by stabbing through the underlying agar with a sterile Pasteur pipette. The agar plug containing the plaque was expelled into 1 ml of 2TY medium. The suspension of agar was left to stand at room temperature for 2 h, to allow bacteriophage particles to diffuse out of the agar. The suspension was then stored at -20 °C.

3.2.7 Preparation of Infective Stocks of Bacteriophage M13

2TY broth (1.5 ml) was inoculated with 37.5 µl of a stationary-phase culture of *E. coli,* and with 0.1 ml of a bacteriophage suspension (Section 3.2.6). The infected culture was grown for 5 h at 37 °C (300 revolutions min^{-1}, with the tubes held in a slanting position). The liquid culture were stored at 4 °C until required.

3.2.8 Preparation of Competent *E. coli* TG1

E. coli cells were made competent by the method of Mandel & Higa (1970). *E. coli* (100-ml culture) were grown until the A_{600} value reached 0.5 - 0.6. The bacteria were then harvested by centrifugation (2000 ***g***, 10 min, 4 °C), resuspended in 10 ml of ice-cold $CaCl_2$ (50 mM) and were placed on ice for 20 min. Competent cells were recovered by centrifugation (2000 ***g***, 10 min, 4 °C), resuspended in 1 ml of ice-cold $CaCl_2$ (50 mM) and were placed on ice for at least 1 h before being transformed.

3.2.9 Transformation of *E. coli* TG1 with Plasmid DNA

Competent *E. coli* cells were transformed with plasmid DNA using the method of Cohen *et al.* (1972). Purified plasmid DNA (typically 5 ng, with the exception of controls), or a ligation reaction mixture, was added to a 1.5 ml microcentrifuge tube containing 0.1 ml of competent cells and 0.1 ml of $CaCl_2$ (50 mM). Negative transformation controls contained 0.1 ml of competent cells and 0.1 ml of $CaCl_2$ (50 mM). Positive transformation controls contained 20 ng of pRK59 or 32 ng of pUC18, as appropriate, 0.1 ml of competent cells and 0.1 ml of $CaCl_2$ (50 mM). Transformation mixtures were incubated on ice for 1 h, heat-shocked at 42 °C for 2 min and incubated on ice again for 5 min. The contents of each tube were then added to tubes containing LB medium (2 ml) and were incubated at 37 °C for 2 h in an orbital shaker (250 revolutions min^{-1}). Dilutions of the cultures were made into fresh LB medium (10-fold, 100-fold, 1000-fold). Diluted and undiluted cultures (0.1 ml) were then spread on LB + ampicillin agar, and the agar plates were incubated for 16 h at 37 °C.

3.2.10 α-Complementation Test

E. coli cells that carried the recombinant plasmid pADH2 were identified by the failure of colonies to α–complement, failure to express the *lacZ* gene (Ullmann *et al.*, 1967). Complementation was assayed by monitoring the colony's ability to express a blue colour when grown in the presence of the chromogenic substrate 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-gal) and the inducer isopropylthio-β-D-galactoside (IPTG) (Horwitz *et al.*, 1964).

LB + ampicillin agar plates were prepared for the α-complementation test by spreading a mixture containing 40 µl of X-gal [2% (w/v) in dimethylformamide] and 4 µl of IPTG (200 mg ml^{-1} in H_2O) on the surface of each plate. The plates were then dried at 37 °C, inoculated with cultures of transformed *E. coli* cells (Section 3.2.9), and incubated for 16 h at 37 °C.

3.2.11 ADH 2334 Antiserum

Rabbit polyclonal antiserum, raised against ADH 2334 and stored as 20 ml aliquots of serum at -20 °C (Sheehan *et al.*, 1988) was obtained from Marian Sheehan. Prior to being used in this study, possible anti-*E. coli* antibodies were removed from the antiserum by the procedure described below.

Frozen antiserum was thawed, mixed with an equal volume of *E. coli* TG1 whole-cell lysate (see below) and the mixture was placed at 4 °C for 16 h. The resulting immunoprecipitate was removed by centrifugation (2500 ***g***, 15 min, 4 °C). The anti-*E. coli* -free antiserum was dispensed into aliquots and stored at -20 °C.

E. coli cells were harvested from a stationary-phase culture by centrifugation (2000 ***g***, 10 min, 4 °C). The cell pellet was resuspended in 4 ml of ice-cold Tris-HCl (50 mM), sucrose (0.25 M), Na_2EDTA (1mM), pH 8. Lysozyme was added (0.14 ml of a 12 mg ml^{-1} solution in H_2O), the suspension was placed on ice for 5 min and then sonicated (5 x 30 s at 150 W). Cell debris was removed by centrifugation (2500 ***g***, 15 min, 4 °C), the cleared lysate was then dispensed into 5 ml aliquots and stored at -20 °C.

3.2.12 Colony Immunoblotting

The colony immunoblotting technique described by Helfman *et al.* (1983) was adapted to detect *E. coli* colonies that expressed *adh*2334. Selected colonies were stab inoculated onto LB agar (the blotting plate) and onto LB + ampicillin agar (the master plate), and grown for 16 h at 37 °C. The blotting plate included 4 colonies of nontransformed *E. coli,* as a control for nonspecific antibody-binding. The colonies that grew on the blotting plate were transferred to a nitrocellulose disc (0.4 µm pore-size, 82 mm diameter; Schleicher & Schuell) by overlaying the agar with nitrocellulose (4 °C, 30 min). The colonies that adhered to the nitrocellulose

were killed by exposure to chloroform vapour (15 min) and were lysed by immersion in 50 mM Tris-HCl, 150 mM NaCl, 5 mM $MgCl_2$, 3% (w/v) bovine serum albumin, 1 μg ml^{-1} DNase, 40 μg ml^{-1} lysozyme, 2 mM phenylmethane sulphonyl fluoride, pH 7.5 (10 ml, 16 h, 4 °C). Nonspecific antibody-binding sites were then blocked by shaking the nitrocellulose gently (30 min, room temperature) in a solution [4% (w/v)] of commercial milk powder in 10 mM Tris-HCl, 0.9% (w/v) NaCl, pH 7.4 (TBS). The nitrocellulose was then shaken gently (2 h, room temperature) in 5 ml of primary antibody (Section 3.2.11) that had been diluted 100-fold into TBS containing 4% (w/v) commercial milk powder. The blot was washed (10 min each wash) first with 10 ml of TBS, then with 0.05% (w/v) polyoxyethylene sorbitan monolaurate in TBS and finally with TBS. The blot was then shaken gently (90 min, room temperature) in 5 ml of goat anti-rabbit IgG conjugated to horseradish peroxidase (Bio-Rad). Before use, goat anti-rabbit IgG was diluted 1000-fold into TBS containing 4% (w/v) commercial milk powder.

Colonies that were positive for ADH 2344 stained brown when the nitrocellulose was immersed (2 to 5 min, room temperature) in a freshly prepared solution containing 1 mM diaminobenzidine, 0.019% (v/v) H_2O_2, 25 mM citric acid and 50 mM Na_2HPO_4.

3.2.13 Purification of Plasmid DNA

3.2.13.1 Preparation of Equilibrated Liquid Phenol

Solid phenol (100 g) was liquefied at 68 °C, 8-hydroxyquinoline (0.1 g) was added and the solution was cooled to 40 °C. Tris-HCl (0.5 M, pH 8.0, 100 ml) was added, the mixture was stirred vigorously and then left to settle. The aqueous (upper) layer was removed, replaced with 0.1 M Tris-HCl, pH 8.0, and the equilibration step repeated until the pH of the phenol reached at least 7.0. The equilibrated liquid phenol was stored under 0.1 M Tris-HCl, pH 8.0, at 4 °C in the dark. Equilibrated liquid phenol was used in the preparation of phenol-chloroform-isoamyl alcohol solutions, which were 25 volumes of phenol mixed with 24 volumes of chloroform and 1 volume of isoamyl alcohol.

3.2.13.2 Small-Scale Purification of Plasmid DNA from *E. coli*

Plasmid DNA was purified from small-scale cultures of transformed *E. coli* by the "boiled-lysis" method of Holmes & Quigley (1981).

Transformed *E. coli* were grown for 16 - 18 h in 2 ml of LB + ampicillin medium. Bacteria were harvested by centrifugation (10,000 ***g***, 20 min, room temperature). The pellet of bacteria was resuspended in 60 μl of Tris-sucrose-EDTA buffer (50 mM Tris-HCl, 2 mM Na_2EDTA, 25% (w/v) sucrose, pH 8.0). The suspension was incubated on ice for 5 min, following the addition of 20 μl of lysozyme solution (10 mg ml^{-1} in 10 mM Tris-HCl, pH 8.0). The lysozyme-treated

bacteria were then lysed by the addition of 0.55 ml of Triton-sucrose lysing mix (50 mM Tris-HCl, 50 mM Na_2EDTA, 5% (w/v) sucrose, 5% (w/v) Triton X-100, pH 8.0). The lysate was vortex-mixed, placed in a boiling water-bath for 1 min and then centrifuged (10,000 ***g***, 20 min, room temperature). The resulting pellet (containing chromosomal DNA and bacterial membranes) was removed from the centrifuge tube with a sterile cocktail stick and discarded. The supernatant was incubated at 50 °C for 20 min, following the addition of RNase A (10 µl of a 10 mg ml^{-1} solution).

Protein was removed from the RNase-treated supernatant by extraction into phenol-chloroform-isoamyl alcohol. An equal volume of phenol-chloroform-isoamyl alcohol was added to the supernatant, the mixture was vortex-mixed and centrifuged (10,000 ***g***, 10 min, room temperature). Following centrifugation, the aqueous layer was transferred to a fresh tube and extracted again with phenol-chloroform-isoamyl alcohol. The organic (lower) layer was mixed with an equal volume of 10 mM Tris-HCl, 1 mM Na_2EDTA, pH 8.0 (TE), and centrifuged as before. The two aqueous layers were combined and extracted, once, with an equal volume of chloroform-isoamyl alcohol (24:1, v/v). Plasmid DNA was recovered from the final aqueous layer by precipitation in the presence of 0.2 M NaCl and 2 volumes of ethanol (-70 °C, 2 h). The precipitated DNA was collected by centrifugation (10,000 ***g***, 30 min, room temperature), washed twice (without resuspension) by adding 0.2 ml of 70% (v/v) ethanol and recentrifuging as before, dried under vacuum, dissolved in a minimum volume of TE and stored at -20 °C.

3.2.13.3 Large-Scale Purification of Plasmid DNA from *E. coli* TG1

Plasmid DNA was purified from 500-ml stationary-phase cultures of transformed *E. coli* cells by a combination of detergent lysis (adapted from the original method of Godson & Sinsheimer (1967), precipitation with polyethylene glycol and equilibrium centrifugation in a CsCl-ethidium bromide gradient (Sambrook *et al.*, 1989).

Cells were harvested (2000 ***g***, 4 °C, 15 min), washed by resuspension in 200 ml of TE, recentrifuged as before and resuspended in 10 ml of TES-sucrose (30 mM Tris-HCl, 5 mM Na_2EDTA, 50 mM NaCl, 20% (w/v) sucrose, pH 8.0). The cell suspension was incubated at 37 °C for 10 min, following the addition of 1ml of lysozyme (10 mg ml^{-1} in 10 mM Tris-HCl, pH 8.0). The suspension was incubated for an additional 10 min at 37 °C, following the addition of 5 ml of Na_2EDTA (250 mM, pH 8.0). The bacteria were lysed by the addition of an equal volume of Triton-lysing mix [50 mM Tris-HCl, 60 mM Na_2EDTA, 2% (w/v) Triton X-100, pH 8.0] and the resulting lysate was clarified by centrifugation (43,000 ***g***, 45 min, 4 °C).

The cleared lysate was digested with RNase A (20 µl of a 10 mg ml^{-1} solution, 37 °C, 20 min) and the plasmid DNA was precipitated by the addition of 1/9 volumes of NaCl (5 M) and sufficient solid polyethylene glycol (PEG 8000) to bring

the PEG 8000 concentration to 10% (w/v). The mixture was vortex-mixed vigorously (to solubilise the PEG 8000), incubated on ice for 90 min (to precipitate the plasmid DNA), warmed to room temperature and, finally, the precipitate of (impure) plasmid DNA was collected by centrifugation (7500 ***g***, 10 min, 15 °C) and dissolved in a minimal volume (~5 ml) of TE. Solid CsCl was added, to give a final concentration of 1 g ml^{-1}, and the mixture was warmed to 30 °C to solubilise the CsCl. An aqueous solution of ethidium bromide (10 mg ml^{-1}) was added to give a final concentration of ethidium bromide of 0.8 mg ml^{-1}. The solution was centrifuged (2000 ***g***, 10 min, 15 °C) to remove the layer of proteinaceous debris that formed on the surface of the solution and the infranatant was transferred to heat-sealable ultracentrifuge tubes (13 x 51 mm; Beckman). Following ultracentrifugation (125,000 ***g***, 16 h, 18 °C; Beckman VT165 vertical rotor), the plasmid DNA was withdrawn from the tubes using a 2 ml syringe fitted with an 18 gauge needle.

Ethidium bromide was removed from plasmid DNA by extraction into water-saturated butan-1-ol. Plasmid DNA was dispensed in 0.3 ml aliquots into 1.5 ml microcentrifuge tubes. Water-saturated butan-1-ol was added, the tubes were vortex-mixed (10 s) and centrifuged (10,000 ***g***, 3 min). The pink butan-1-ol phase was discarded and the extraction was repeated until no pink colour was observed (typically 3 extractions).

Pure plasmid DNA was recovered from the CsCl solution by precipitation with ethanol. To every 1 ml of CsCl solution, 2 ml of H_2O and 6 ml of ethanol were added. The tubes were left at room temperature for 4 h. Precipitated DNA was collected by centrifugation (10,000 ***g***, 30 min, room temperature), washed twice (without resuspension) by the addition of 10 ml of 70% (v/v) ethanol and recentrifuging as before, dried under vacuum, dissolved in a minimum volume of TE and stored at -20 °C.

3.2.13.4 Recovery of Plasmid DNA from Agarose Gels

Fluorescent bands, corresponding to DNA fragments, were excised from agarose gels, placed on Petri dishes and cut into thin slices. The agarose slices, corresponding to a single band, were transferred to a microcentrifuge tube, crushed with a sealed micropipette tip, suspended in an equal volume of equilibrated liquid phenol and vortex-mixed. The resulting slurry was frozen (-70 °C, 30 min) and then centrifuged (10,000 ***g*** 10 min, 4 °C). After centrifugation, the agarose was visible as a discrete white layer at the interface between the organic layer and the aqueous layer. The aqueous layer, containing the plasmid DNA, was removed into a clean microcentrifuge tube and the phenol-extraction was repeated. An equal volume of TE buffer was added to the tube containing the original organic layer plus agarose interface, and this mixture was vortex-mixed and centrifuged as before. The two aqueous layers were then combined, an equal volume of chloroform-isoamyl alcohol

(24:1, v/v) was added and the mixture was centrifuged as before. The DNA was recovered from the final aqueous layer by precipitation with ethanol (as in Section 3.2.13.2, but in the presence of 0.5 M NaCl). The final, dry, pellet of DNA was dissolved in a minimum volume of TE and stored at -20 °C.

3.2.14 Purification of Single-Stranded Bacteriophage DNA

The procedure used to purify single-stranded bacteriophage DNA was based on the protocol provided with the Amersham RPN 1523 oligonucleotide-directed *in vitro* mutagenesis kit (version 2).

2TY medium (50 ml) was inoculated with 0.5 ml of a stationary-phase culture of *E. coli* and 1 ml of an infective stock of bacteriophage M13 particles (Section 3.2.7). The culture was grown for 5 h at 37 °C (300 revolutions min^{-1}). Cells were harvested (5000 ***g***, 30 min, 4 °C) and the supernatant, containing single-stranded bacteriophage DNA, was transferred to clean centrifuge tubes. PEG 8000 [0.2 volumes of a 20% (w/v) solution in 2.5 M NaCl; PEG-NaCl] was added to each tube and the tubes were incubated for 1 h at 4 °C. Precipitated DNA was collected by centrifugation (5000 ***g***, 20 min, 4 °C), the supernatant was removed and, to collect all remaining traces of PEG-NaCl from the sides of the tubes, the tubes were recentrifuged (5000 ***g***, 5 min, 4 °C). To remove any remaining cells, precipitated DNA was resuspended in 0.25 ml of TE and centrifuged (10,000 ***g***, 5 min, room temperature). DNA was reprecipitated from the cell-free supernatant by addition of 0.1 ml of PEG-NaCl, incubation at room temperature for 15 min, and centrifugation (10,000 ***g***, room temperature, 5 min). The final DNA precipitate was resuspended in 0.25 ml of TE and extracted with phenol-chloroform-isoamyl alcohol (as in Section 3.2.13.2).

Single-stranded bacteriophage DNA was recovered from the final aqueous layer by precipitation with ethanol (as in Section 3.2.13.2, but in the presence of 0.3 M sodium acetate). The final DNA precipitate was resuspended in a minimal volume of TE and stored at -20 °C (Section 3.2.13.2).

3.2.15 Estimation of DNA Concentration

The absorbance values of appropriate dilutions of plasmid DNA solutions were measured at 260 nm. The concentrations of DNA in the original solutions were estimated using Equation 6 (Sambrook *et al.*, 1989).

$$\mathbf{A_{260} \times dilution\ factor \times 50 = \mu g\ of\ DNA\ ml^{-1}} \qquad \mathbf{(6)}$$

Alternatively, a qualitative estimate of the DNA concentration was obtained by comparing the fluorescence of a range of dilutions of the DNA sample of interest and a plasmid DNA sample of known concentration after agarose gel-electrophoresis (Section 3.2.16). The concentration of single-stranded bacteriophage DNA solutions was estimated by the electrophoresis method using a solution of M13mp18 single-stranded DNA (United States Biochemical; 0.2 $\mu g\ l^{-1}$) as the standard. Molarity was calculated using 324.5 as the average M_r of a deoxynucleotide base.

3.2.16 Agarose Gel-Electrophoresis

Plasmid DNA was analysed by electrophoresis through 1% (w/v) agarose gels, with 89 mM Tris, 89 mM boric acid, 2.5 mM Na_2EDTA, pH 8.3 (TBE), as the buffer system. An Atto, horizontal submerged gel-electrophoresis apparatus was used. Gel dimensions were 10 x 8 x 0.5 cm. Ultra-pure agarose was obtained from Gibco BRL. An *Eco*RI-*Hin*dIII digest of phage λ DNA (Boehringer) was used as DNA size markers. The fragment sizes were 21,226, 5148, 4973, 4268, 3530, 2027, 1904, 1584, 1375, 947, 831, 564 and 125 bp. Before electrophoresis, and to prevent hybridisation of the 3530-bp fragment to the 21,226-bp fragment, an aliquot of the size markers was incubated at 65 °C for 5 min.

Plasmid DNA samples were mixed with 0.25 volume of 0.25% (w/v) bromophenol blue, 30% (w/v) glycerol in TBE (agarose gel loading-buffer) before loading onto the gel. Electrophoresis was performed at room temperature and at 50 V until the bromophenol blue in the loading-buffer had migrated to within 2 cm of the anode end of the gel.

Following electrophoresis, gels were soaked in an aqueous solution of ethidium bromide (5 $\mu g\ ml^{-1}$, 15 min), rinsed with H_2O and viewed on a transilluminator.

3.2.17 Digestion of Purified Plasmid DNA with Restriction Enzymes

*Eco*RI and *Bam*HI were obtained from Boehringer and were stored, as supplied, as 10 $U\ \mu l^{-1}$ stocks. Boehringer Buffer B was 100 mM Tris-HCl, 50 mM $MgCl_2$, 1 M NaCl, 10 mM 2-mercaptoethanol, pH 8.0. Working solutions of restriction enzymes were 1 $U\ \mu l^{-1}$ in a ten-fold dilution of Boehringer Buffer B and were freshly prepared for each digest.

3.2.17.1 Analytical Restriction Digests

Plasmid DNA was digested with a single enzyme (*Eco*RI or *Bam*HI) or with both enzymes simultaneously, in a final volume of 20 µl. Aliquots of DNA,

corresponding to 1 - 2 µg, were mixed with 2 µl of undiluted Boehringer Buffer B, the appropriate volume of H_2O and, finally, 2 U of each enzyme or 2 U of a single enzyme, per µg of DNA. Following incubation for 2 h at 37 °C, 5 µl of agarose gel loading-buffer were added and the digests (25 µl) were immediately analysed by agarose gel-electrophoresis (Section 3.2.16).

3.2.17.2 Preparative Restriction Digests

Plasmid DNA was digested, with *Eco*RI and *Bam*HI simultaneously, in a final volume of 40 µl. Aliquots of DNA , equivalent to ~ 6 µg, were mixed with 4 µl of Boehringer Buffer B, the appropriate volume of H_2O and 3 U of each enzyme per µg of DNA. Following incubation for 2 h at 37 °C, 8 µl of agarose gel loading-buffer (Section 3.2.16) were added and the digests were analysed by agarose gel-electrophoresis (Section 3.2.16).

3.2.18 Ligation Reactions

T4 DNA ligase was purchased from Promega, dispensed into small aliquots and stored at -20 °C. Promega ligation buffer consisted of 300 mM Tris-HCl, 100 mM $MgCl_2$, 100 mM dithiothreitol, 10 mM ATP, pH 7.8. A Weiss unit of T4 DNA ligase is defined as the amount of enzyme that catalyzes the exchange of 1 nmol of ^{32}P from diphosphate into [γ, β–^{32}P] ATP in 20 min at 37 °C (Weiss *et al.*, 1968).

Plasmids were restricted on a preparative scale (Section 3.2.17.2) and the restriction products required for ligation reactions were purified by agarose gel-electrophoresis (Section 3.2.16). Aliquots of vector (0.345 pmol) and insert (1 pmol) were mixed with 4.5 µl of Promega ligation buffer, the appropriate volume of H_2O (for a final volume of 45 µl) and 1 Weiss unit of T4 DNA ligase per 0.1 µg of vector. The ligation mixes were incubated for 18 h at 18 °C and then transformed (Section 3.2.9) into *E. coli.*

3.2.19 Manual DNA-Sequencing

The dideoxy chain-termination method (Sanger *et al.*, 1977; Biggin *et al.*, 1983) was used to sequence double-stranded plasmid DNA (Murphy & Ward, 1989) and single-stranded phage M13 DNA. Single-stranded DNA-binding protein was purchased from Pharmacia and deoxyadenosine 5'-(α-[^{35}S]thio)triphosphate, triethylammonium salt, ~ 37 TBq $mmol^{-1}$ ([^{35}S]-dATP) was from Amersham. Sequencing reagents (Table 3.2.2) and polymerase were obtained as a kit from United States Biochemical (Sequenase version 2.0 DNA sequencing kit). The polymerase was Sequenase version 2.0, a variant of bacteriophage T7 polymerase (Tabor & Richardson, 1989).

Reagent	Composition
Sequenase reaction buffer	0.2 M Tris-HCl, 0.1 M $MgCl_2$, 0.25 M NaCl, pH 7.5
Sequenase dilution buffer	10 mM Tris-HCl, 5 mM dithiothreitol, 0.5 mg ml^{-1} bovine serum albumin
Labeling mix (5x stock)	7.5 μM 2'-deoxyguanosine-5'-triphosphate (dGTP), 7.5 μM 2'-deoxycytidine-5'-triphosphate (dCTP), 7.5 μM 2'-deoxythymidine-5'-triphosphate (dTTP)
Dithiothreitol	0.1 M dithiothreitol
ddG Termination mix	80 μM dGTP, 80 μM dCTP, 80 μM dTTP, 80 μM 2'-deoxyadenosine-5'-triphosphate (dATP), 8 μM dideoxyguanosine-5'-triphosphate (ddGTP), 50 mM NaCl
ddA Termination mix	As for ddG termination mix, with dideoxyadenosine-5'-triphosphate (ddATP) replacing ddGTP
ddT Termination mix	As for ddG termination mix, with dideoxythymidine-5'-triphosphate (ddTTP) replacing ddGTP
ddC Termination mix	As for ddG termination mix, with dideoxycytidine-5'-triphosphate (ddCTP) replacing ddGTP
Mn buffer	0.15 M sodium isocitrate, 0.1 M $MnCl_2$
Stop solution	95% (w/v) formamide, 20 mM Na_2EDTA, 0.05% (w/v) bromophenol blue, 0.05% (w/v) xylene cyanol

Table 3.2.2 **Composition of Sequencing Reagents**
The reagents were supplied in a Sequenase version 2.0 DNA sequencing kit, obtained from United States Biochemical.

3.2.19.1 Sequencing Primers

The M13/pUC forward sequencing primer was obtained from United States Biochemical. The M13/pUC reverse sequencing primer was obtained from Boehringer. All other sequencing primers (Section 3.3.3; Figure 3.3.7) were synthesised (nmolar scale) at the Department of Genetics, Trinity College, Dublin, using an Applied Biosystems model 391 PCR MATE DNA synthetiser and the phosphite-triester method (Itakura *et al.*, 1984).

Oligonucleotides were obtained from the DNA synthetiser, attached to the solid support, but with the dimethoxy trityl protecting groups removed. Oligonucleotides were cleaved from the support by 3 successive washes with 1 ml of ammonia solution [Merck, 25% (w/v), 15 min per wash]. Washes were combined, incubated for 16 h at 55 °C and the ammonia was removed under vacuum. The typical yield was 160 nmol.

The recovered oligonucleotides were dissolved in H_2O (40 µl), centrifuged (10,000 ***g***, 5 min) to remove particulate material, and adjusted to 0.5 M ammonium acetate (addition of 2 M ammonium acetate). Oligonucleotides were precipitated from solution by the addition of 1.4 ml of ethanol and incubated at -70 °C for 30 min. The precipitate was collected by centrifugation (10,000 ***g***, 15 min), washed twice (without resuspension) with 0.2 ml of 70% (v/v) ethanol and dried under vacuum. The dried pellet was dissolved in 0.1 ml of TE, centrifuged (10,000 ***g***, 5 min), to remove particulate material, and stored as small aliquots at -20 °C.

The oligonucleotide concentration was estimated by determining the A_{260} of appropriate dilutions of synthetic oligonucleotide solutions and applying Equation 6, with a multiplication factor of 20.

3.2.19.2 Denaturation and Recovery of Double-Stranded DNA Templates

Plasmid DNA (~16 µg in 20 µl of TE) was mixed with 5 µl of 1M NaOH, 1mM Na_2EDTA, and incubated at 37 °C for 15 min. The denatured DNA template was immediately recovered by spin-dialysis.

Spin-dialysis columns of Sepharose-CL6B, pre-equilibrated in 10 mM Tris-HCl, 0.1 mM Na_2EDTA, pH 8.0, were prepared in 0.5 ml microcentrifuge tubes, and were used, exactly as described by Murphy & Ward (1989). A single column was used to desalt 16 µg (25 µl) of denatured plasmid DNA. The DNA was recovered in a volume of approximately 25 µl, and stored (up to one month) at -20 °C in 7 µl aliquots.

3.2.19.3 Sequencing Reactions

For each set of four termination reactions, a single annealing (and subsequent labeling) reaction was used.

Annealing template and primer Denatured spin-dialysed plasmid DNA (5 μg in 7 μl) or single-stranded bacteriophage M13 DNA (2 - 4 μg at 1 μg ml^{-1}) was mixed with 2 μl of Sequenase reaction buffer and sequencing primer (1.5 pmol of primer per μg of plasmid DNA, or 0.75 pmol of primer per μg of bacteriophage DNA). The mixture was incubated at 37 °C for 15 min (plasmid DNA) or 5 min at 65 °C, followed by slow cooling (30 min) to 35 °C (bacteriophage DNA).

Labeling reaction Labeling mix was diluted five-fold with sterile H_2O before use. Sequenase was diluted seven-fold with Sequenase dilution buffer before use. Freshly annealed template and primer (10 μl) were mixed with dithiothreitol (1 μl), diluted labeling mix (2 μl), [^{35}S]-dATP (1 - 2 μl) and diluted Sequenase (2 μl). When sequence data were required adjacent to the priming site, Mn buffer (1 μl) was included in the labeling reaction. When single-stranded DNA-binding protein was included, 1 μl (1.5 μg) was added to annealed template and primer prior to the addition of dithiothreitol. Labeling was performed at room temperature for 5 min.

Termination Labeling reaction mix (3.5 μl) was dispensed into four pre-warmed tubes, each tube containing one of the four termination mixes. The tubes were mixed and incubated at 37 °C for 5 min. The termination reactions were stopped by adding 4 μl of stop solution and were either electrophoresed immediately (5 μl of sequence reaction per track), or stored at -20 °C (for up to one week).

3.2.19.4 Electrophoresis of Sequencing Reaction Mixtures

The electrophoresis apparatus was a Bio-Rad Sequi-Gen nucleic acid sequencing cell. The gel mould dimensions were 17 cm x 40 cm x 0.4 mm.

Sequencing gels were composed of 5.7% (w/v) acrylamide, 2% (w/v) bisacrylamide, 42% (w/v) urea (Aristar grade, BDH), 89 mM Tris, 89 mM boric acid, 2.5 mM Na_2EDTA, pH 8.3. The electrophoresis buffer system was TBE (Section 3.2.16). The acrylamide-bisacrylamide stock solution was 38% (w/v) acrylamide, 2% (w/v) *N*, *N'*-methylenebisacrylamide. Before use, the stock solution (100 ml) was deionised by adding Amberlite MB-1 resin (5 g) and stirring (30 min). The resin was removed by filtration through a glass fibre filter (Whatman GF/C). The acrylamide stock solution was stored for up to one month at 4 °C.

The gels were polymerised in the presence of 0.075% (w/v) ammonium persulfate (Electran grade, BDH) and 0.1% (v/v) *N*,*N*,*N'*,*N'*-tetramethylethylenediamine (Electran grade, BDH). Gels were cast according to the instructions provided with the sequencing apparatus (Bio-Rad instruction manual) and were stored for 12 h before use. The gels were heated to 55 °C (2500 V, 41 mA) before loading the samples.

Sequencing reactions were heated at 75 °C for 2 min immediately prior to electrophoresis. Electrophoresis was performed at 50 °C (~30 mA, 1800 V) for 1.25 - 6 h. Following electrophoresis, gels were soaked in 1% (v/v) methanol, 10% (v/v)

acetic acid, transferred to 3MM Whatman filter paper and dried on a Savant gel drier. Dried gels were left in contact with Kodak X-omat RP X-ray film for 2 - 6 days. The film was developed according to the manufacturer's instructions.

3.2.20 Automated DNA-Sequencing

Automated DNA-sequencing was performed by the dye-labeled primer method (M13 universal forward primer) using an Applied Biosystems model 373A DNA sequencer (Hunkapiller *et al.*, 1991).

3.2.21 Analyses of Nucleotide and Amino-Acid Sequences

Nucleotide and amino-acid sequence analyses and database interrogations were performed using a Unix operating system on the "Acer" workstation at the Irish National Centre for BioInformatics. The Genetics Computer Group sequence analysis software package (GCG, version 8, September 1994; Devereux *et al.*, 1984) was used to calculate pairwise amino-acid sequence identities (programme Bestfit, gap weight of 3, gap length of 0.1), to perform multiple amino-acid sequence alignments (programme Pileup, gap weight of 3, gap length of 0.1), to obtain amino-acid compositions of proteins (programme Composition) and to interrogate the Swiss Protein Bank at the University of Geneva, Switzerland (Swiss-Prot, Release 34, 1996). Amino-acid and nucleotide sequence homology searches were performed using the Basic Local Alignment Search Tool programme (BLAST; Altschul *et al.*, 1990). The programme Query was used to search the GenBank DNA database (National Centre for Biotechnology Information, Bethesda, Maryland). The programme Querypir was used to search the Protein Identification Resource amino-acid sequence database (PIR, Release 49, 1996), National Biomedical Research Foundation, Georgetown University Medical Center, Washington, DC, Sidman *et al.*,1988).

The free energy of potential DNA stem loop structures was calculated by the method described by Cantor & Schimmel (1980).

3.3 RESULTS

3.3.1 Characterisation of Plasmid pADH1b

When competent *E. coli* TG1 cells were transformed with the plasmid pADH1b, and grown on LB + ampicillin agar, 617 colonies ng^{-1} of plated plasmid DNA were observed. Twenty-six colonies were selected at random, transferred to nitrocellulose, and screened with anti-ADH 2334 antiserum. Twenty-five colonies cross-reacted with the antiserum, one colony was selected at random and the plasmid DNA was purified from a small-scale culture of that colony. The purified plasmid DNA was restricted with *Eco*RI and *Bam*HI, and the expected restriction fragments, 1.5 kb long (corresponding to *adh*2334) and 5 kb long (corresponding to the pRK59 vector), were observed (Figure 3.3.1; Lane 5). Plasmid pADH1a was transformed into *E. coli* TG1 and the plasmid DNA was purified from a small-scale culture of one of the resulting transformants. An *Eco*RI plus *Bam*HI restriction analysis of the pADH1a DNA is shown in Figure 3.3.1 (Lane 3). The 2.7 kb-fragment in Lane 3 corresponds to the *B. stearothermophilus* DNA insert in pADH1a.

Plasmid pADH1b DNA was purified from a large-scale culture of the selected colony by equilibrium centrifugation in a CsCl-ethidium bromide gradient. Manual double-stranded DNA sequencing was performed on the purified pADH1b DNA, using the M13/pUC forward sequencing primer and the M13/pUC reverse sequencing primer. The expected positions of the forward and reverse priming sites, with respect to the *B. stearothermophilus* DNA insert, are indicated in Figure 3.3.2(a). The 158 nucleotide sequence obtained, using the forward primer, is presented in Figure 3.3.2(b) (underlined nucleotides). The underlined sequence is complementary to nucleotides 2215 - 2058 of the *Hin*dIII-fragment of the plasmid pBAA1 that had been used in the construction of the vector pRK59 (Devine *et al.*, 1989). Sequencing with the reverse primer was attempted using two different commercial preparations of reverse primer, two different preparations of DNA and multiple ratios of reverse primer to DNA. No sequence information was obtained as a result of any of these experiments.

3.3.2 Subcloning *adh*2334 into pUC18

The *Eco*RI-*Bam*HI *B. stearothermophilus* insert was excised from pADH1b and was ligated into pUC18 restricted with *Eco*RI and *Bam*HI (Figure 3.3.3). The ligation reaction mixture was transformed into *E. coli* TG1, and the transformants were grown on LB agar containing ampicillin, X-gal and IPTG. Seven white colonies were obtained.

All seven white colonies cross-reacted with anti-ADH 2334 antiserum (Figure 3.3.4; Rows 2 and 3). Plasmid DNA was purified from a small-scale culture

of each of the seven colonies, and restricted with *Eco*RI alone, and with *Eco*RI and *Bam*HI simultaneously. The seven plasmids were all approximately 4.2 kb in length and all carried a 1.5-kb *Eco*RI-*Bam*HI insert (Figure 3.3.5).

One of the ADH-positive colonies was selected, arbitrarily, and was the source of the DNA used in the remainder of this work. The new plasmid was named pADH2. A plasmid map of pADH2 is presented in Figure 3.3.6.

3.3.3 Nucleotide Sequencing of the *B. stearothermophilus* Insert in pADH2 and in mp18*adh*

Sequencing data was obtained when plasmid pADH2 was primed with the M13/pUC forward sequencing primer and when the plasmid was primed with the M13/pUC reverse sequencing primer. The nucleotide sequence of each strand of the *B. stearothermophilus* DNA insert in pADH2 was determined using manual double-stranded DNA sequencing and a series of synthetic oligonucleotide primers. The oligonucleotide sequences of the primers, positions of the priming sites, sequence data generated using each primer and the extent of the overlaps between sequence data are indicated in Figure 3.3.7. A typical example of an autoradiogram is presented in Figure 3.3.8.

The data were, in general, good quality and contained only two examples of compression (consecutive bands not resolved). However, there were many stops (the appearance of bands at the same position in all four lanes). Addition of single-stranded DNA-binding protein to labeling reactions, or doubling the amount of sequenase added to labeling reactions, had no affect on the appearance of stops. When all of the information from the manual sequencing of plasmid pADH2 was combined, a total of twenty-eight bases, twenty on the sense strand and eight on the antisense strand, remained as unknowns. The twenty-eight unknown bases are indicated by N in Figure 3.3.7. With the possible exception of nucleotide 1366 (Figure 3.3.7), unknowns on the sense strand did not overlap with unknowns on the antisense strand.

The single-stranded form of bacteriophage mp18*adh* contained the sense strand of *adh*2334 (Section 3.2.2) and was, therefore, an alternative source of sequence data for the antisense strand. The single-stranded DNA sequencing technique is reputedly less vulnerable to the appearance of stops. Manual single-stranded DNA sequencing of mp18*adh*, with primer P4, did indeed result in the resolution of the stops obtained with P4 and plasmid sequencing. Furthermore, it was possible, with P8 and single-stranded sequencing, to extend the antisense strand data to overlap with the entire extent of the 5' end of the sense sequence.

As a final check, plasmid pADH2 was submitted for automated DNA sequence analysis, from the forward priming site. A 432 nucleotide sequence

(antisense strand) was obtained. The data resolved the unknowns situated between the forward priming site and the P4 priming site (Figure 3.3.7), and confirmed the sequence that had been obtained for this region using the manual method. The final sequence, determined for the antisense strand by combining all of the available information from plasmid and single-stranded DNA sequencing, is presented in Figure 3.3.9. The sequence of the sense strand is presented in the same format in Figure 3.3.10. The single compression remaining on the antisense strand in the region of P6 (Figure 3.3.9) was unambiguously resolved by the sequence data from the sense strand (Figure 3.3.11). The stops and the single compression on the sense strand were unambiguously resolved by the sequence data of the antisense strand.

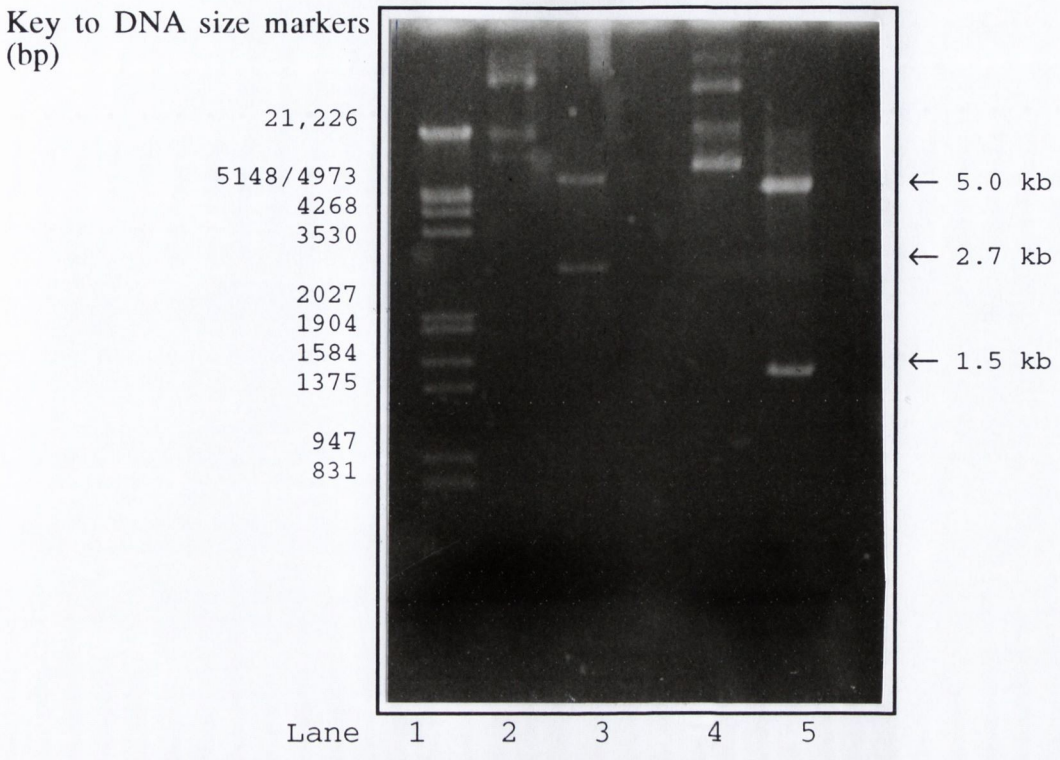
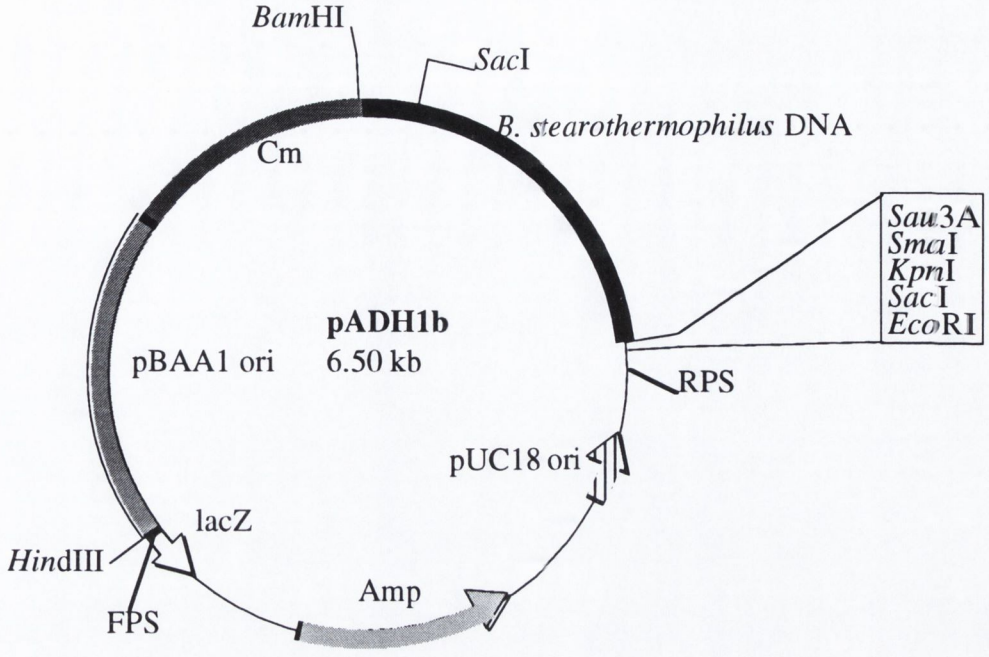


Figure 3.3.1 **Restriction Analysis of pADH1b**

Plasmid DNA was purified (Section 3.2.13.2) from a 2 ml culture of a pADH1b transformant, which had stained positive for ADH antigen. Lane 1 contained DNA size markers: phage λ DNA (0.5 μg), restricted with *Eco*RI and *Hin*d III. Lane 2 contained an aliquot of pADH1a DNA. Lane 3 contained pADH1a, restricted with *Eco*RI and *Bam*HI. Lane 4 contained an aliquot of pADH1b and Lane 5 contained pADH1b, restricted with *Eco*RI and *Bam*HI. Arrows indicate DNA fragments that are 5, 2.7 and 1.5 kb long.



(a)

```
                FP
-61 ...cgccagggttttcccagtcacgacgttgtaaaacgacggccagtgcc -15
    HindIII
-14 AAGCTTTGTTTTCGGAGCAATGACAACAGTATATTTAAACTATCCTAAGA  36

 37 GCAAACCAGTTTATGTAACTATTCAACACAGACAATGTAAAGATTCCCTA  86

 87 GGTTATCATAGACAATTAATCATAAAAAGCTACAAAGATAAAAAACGAAC  13

137 AAAACATATTTCAACACAATAC                             158
```

(b)

Figure 3.3.2 **Nucleotide Sequencing of pADH1b**

Manual DNA-sequencing of plasmid DNA was performed using the M13/pUC forward sequencing primer and the M13/pUC reverse sequencing primer (Section 3.2.19).

(a) Plasmid map of pADH1b showing the expected positions of the forward (FPS) and reverse (RPS) priming sites. Primer-directed DNA synthesis was expected in the clockwise direction from FPS and in the anti-clockwise direction from RPS. Other details as in the legend to Figure 3.1.3.

(b) Sequence data (upper-case letters, underlined, 1 - 158) obtained using the M13/pUC forward sequencing primer. The sequence of the forward primer (FP) is shown in bold lower-case letters. Lower-case letters indicate pUC18 DNA, upper-case letters indicate pBAA1 DNA. The pBAA1 *Hin*dIII site is nucleotides -14 to -9 inclusive.

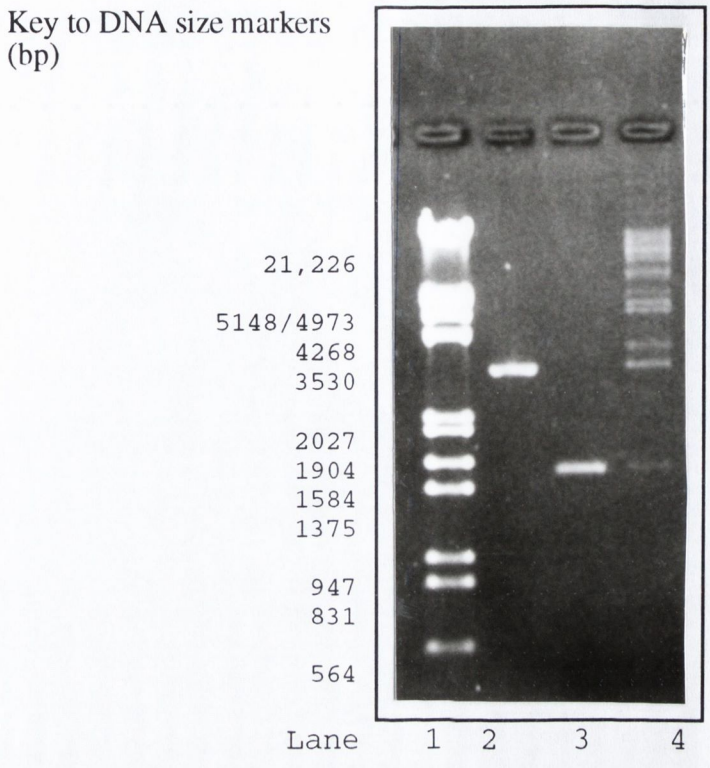


Figure 3.3.3 **Ligation of the 1.5-kb *Eco*RI-*Bam*HI Fragment of pADH1b into pUC18**

Plasmids pUC18 and pADH1b were restricted with *Eco*RI and *Bam*HI (Section 3.2.17.2). Restriction digest mixtures were electrophoresed through 1% (w/v) agarose, the appropriate restriction fragments were excised from the agarose and purified (Section 3.2.13.4). The purity and yield of vector and insert were determined by electrophoresis (this Figure). Lane 1 contained DNA size markers: phage λ DNA (1.25 μg) restricted with *Eco*RI and *Hin*dIII. Lane 2 contained the 2.665-kb *Eco*RI-*Bam*HI fragment of pUC18, which had been purified from agarose, 5.7% of the final yield from 5 μg of uncut pUC18. Lane 3 contained the 1.5-kb *Eco*RI-*Bam* HI fragment of pADH1b, which had been purified from agarose, 11% of the final yield from 6 μg of uncut pADH1b. Lane 4 contained an aliquot of the ligation mixture, equivalent to 0.1 μg of purified cut vector and 0.2 μg of purified insert.

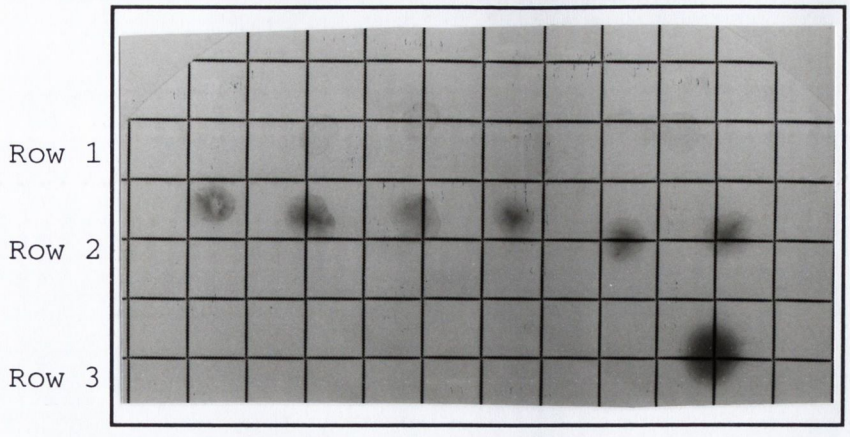


Figure 3.3.4 **Immunoblot of *E. coli* Colonies Transformed with the Products of the Ligation of the 1.5-kb *Eco*RI-*Bam*HI Fragment of pADH1b into pUC18**

Colony immunoblotting was as described in Section 3.2.12. Row 1 contained four separate nontransformed *E. coli* colonies that had been grown on M9 minimal-salts agar. Rows 2 and 3 contained the seven white colonies, obtained when *E. coli* TG1 cells were transformed with the products of the ligation of the 1.5-kb *Eco*RI-*Bam*HI fragment of pADH1b into pUC18. Colonies that were positive for ADH antigen show dark staining.

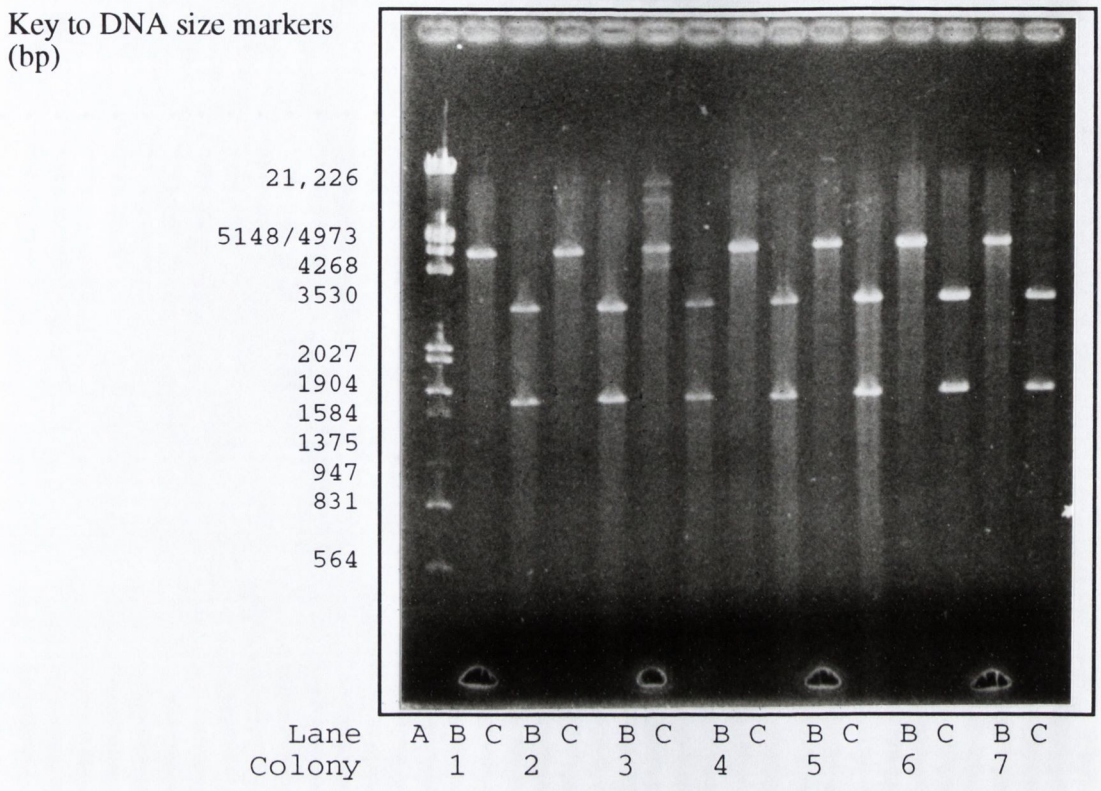


Figure 3.3.5 **Restriction Analyses of the Products of the Ligation of the 1.5-kb *Eco*RI-*Bam*HI Fragment of pADH1b into pUC18**

Plasmid DNA was purified from 2 ml cultures of the seven ADH antigen positive transformants shown in Figure 3.3.4. The seven preparations of DNA were restricted with *Eco*RI (Lanes B), or with *Eco*RI and *Bam*HI (Lanes C). Lane A contained DNA size markers: phage λ DNA (1.25 μg) restricted with *Eco*RI and with *Hin*dIII. Numbers 1 - 7 refer to the colonies from which the DNA was purified.

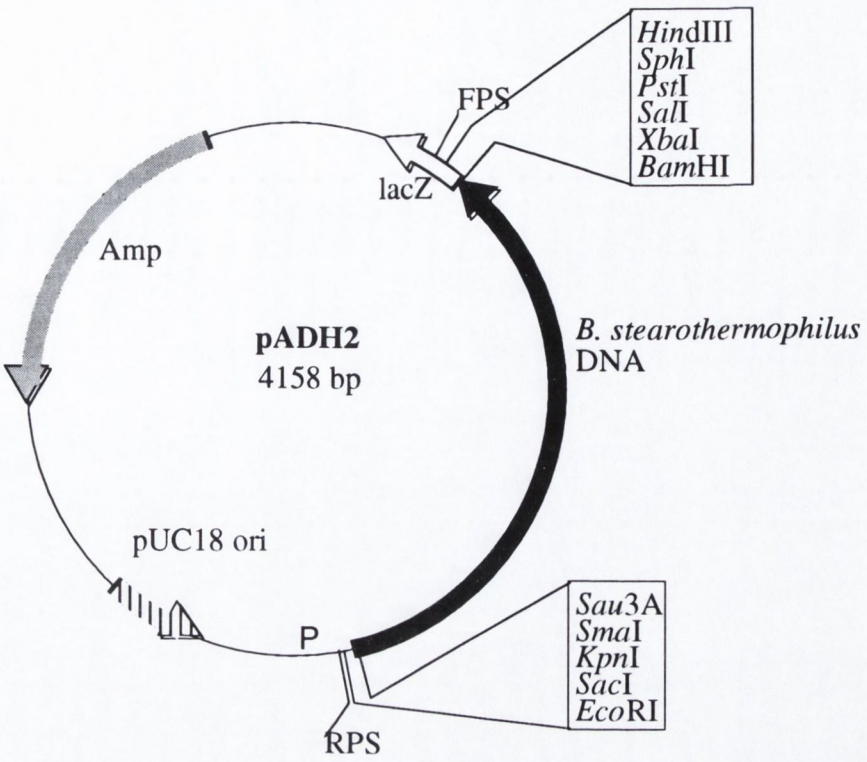


Figure 3.3.6 Structure of the Plasmid pADH2

The plasmid was constructed by inserting the 1.5 kb *Eco*RI-*Bam*HI fragment *(B. stearothermophilus* DNA) of pADH1b (Figure 3.3.2) into pUC18. The notations FPS and RPS indicate the forward and reverse sequencing priming sites respectively. Primer-directed DNA synthesis was in the clockwise direction from FPS and in the anti-clockwise direction from RPS. Arrows indicate the direction of transcription. All other notations as defined in the legend to Figure 3.1.1 (Section 3.1.1).

Figure 3.3.7 **Primer-Directed Nucleotide Sequencing of the *B. stearothermophilus* DNA Insert in pADH2**

The nucleotide sequences presented in upper-case letters were obtained by manual double-stranded DNA sequencing of pADH2 (Section 3.2.19). The sense strand is indicated in red, with unidentified nucleotide bases (N) in black, and reads in the 5' to 3' direction from 1 to 1498. *B. stearothermophilus* DNA extends from, and includes, the *Sau*3A site (GATC) to, and including, nucleotide 1485. The antisense strand is indicated in black, with unidentified nucleotide bases (N) in red. The sequences of the internal sequencing primers P1, P3, P5, P7 and P9 (sense strand), P2, P4, P6 and P8 (antisense strand) and the *Sau*3A site are underlined. The sequence data obtained have been extrapolated (lower-case letters) to show the sequence of the M13/pUC reverse primer (sense strand, RP, underlined), and the sequence of the M13/pUC forward primer (antisense strand, FP, underlined). The overlining (in red) on the sense strand indicates the extent of the sequence information obtained using each primer. For example, the information obtained using RP commences at nucleotide 1 and terminates at nucleotide 218, the information obtained using P1 commences at nucleotide 168 and terminates at nucleotide 347. The extent of the sequence information obtained, using each antisense strand primer, is indicated by underlining (in black) of the antisense strand.

RP
agcggataacaatttcacacaggaaacagctatgaccatgattacgaattcga
Sau3A
1 GCTCGGTACCCGGGGATCATTCATGCTTGAGCAAAAACTGAACACTATATTTTTAGCAGG 60
TTTTTGACTTGTGATATAAAAATCGTCC
P1
61 TTTCCAACTGGAACAACCTGCTTTTTCTTTTCATGCAACAATTCTGCCAATATTTTGTTC 120
AAAGGTTGACCTTGTTGGACGAAAAAGAAAAGTACGTTGTTAAGACGGTTATAAAACAAG
121 AAGTCTGCCAAAAATTCGGCAAAATTCTATTTTGTATTTTTCTAATATTCCGCAAATTAA 180
TTCAGACGGTTTTTAAGCCGTTTTAAGATAAAACATAAAAAGATTATAAGGCGTTTAATT
181 GCGCTTTCAATTTTGGTACGGTTCTTGCATGATATAAAGTCGAAAAAAATGAAAGGGGTT 240
CGCGAAAGTTAAAACCATGCCAAGAACGTACTATATTTCAGCTTTTTTTACTTTCCCCAA
241 GATTTGAGTGAAAGCAGCAGTAGTTAACGAATTTAAAAAANCACTTGAAATTAAAGAGGT 300
CTAAACTCACTTTCGTCGTCATCAATTGCTTAAATTTTTTCGTGAACTTTAATTTCTCCA
P3
301 GGAAAGACCCAAACTAGAGGAAGGTGAAGTTTTAGTTAAGATCGAAGCATGCGGTGTGTG 360
CCTTTCTGGGTTTGATCTCCTTCCACTTCAAAATCAATTCTAGCTTCGTACGCCACACAC
361 CCATACGGATTTACATGCAGCCCATGGTGACTGGCCGATTAAACCGAAATTGCCGTTAAT 420
GGTATGCCTAAATGTACGTCGGGTACCACTGACCGGCTAATTTGGCTTTAACGGCAATTA

Figure 3.3.7

Figure 3.3.7 (continued)

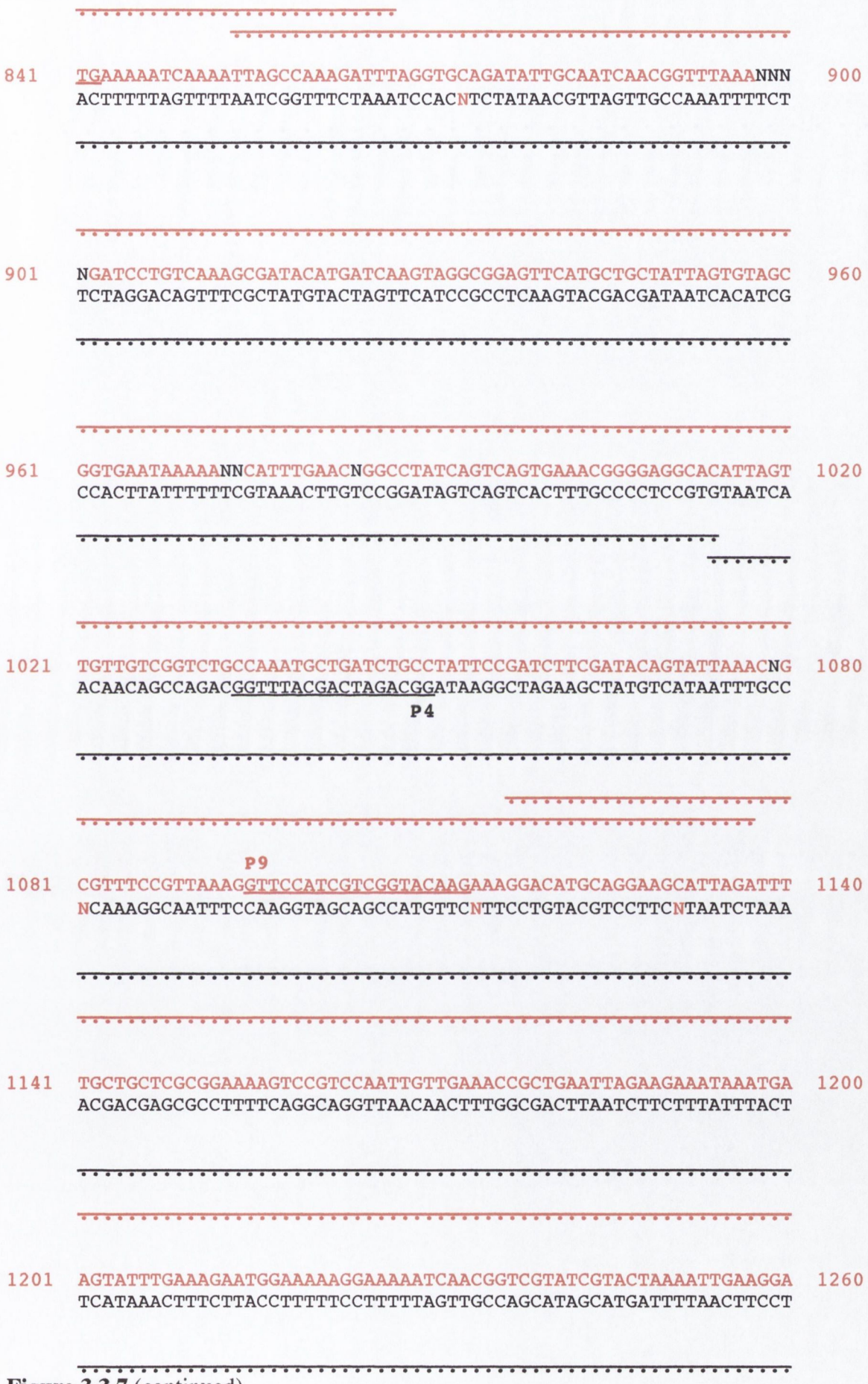


Figure 3.3.7 (continued)

Figure 3.3.7 (continued)

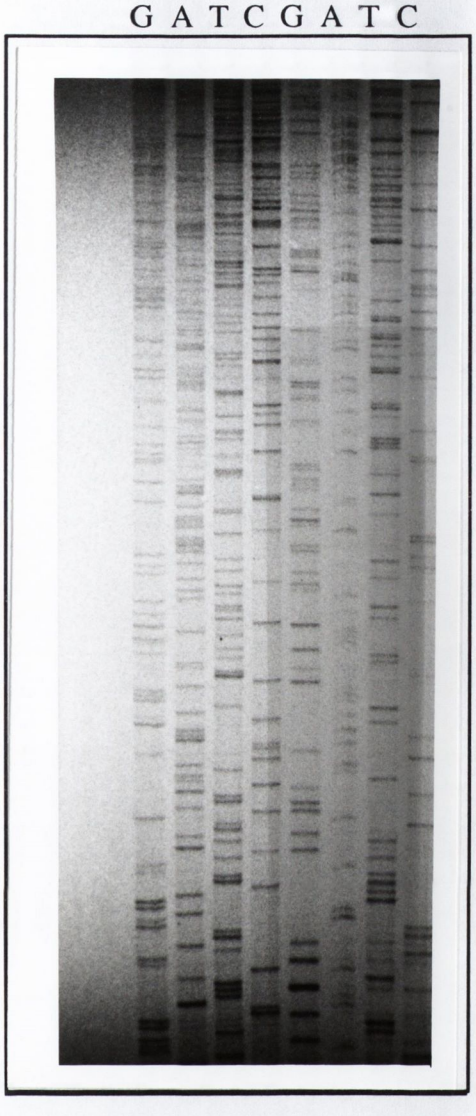


Figure 3.3.8 **An Example of an Autoradiogram from Double-Stranded DNA Sequencing of pADH2**

The data shown are for the sense strand with primer P5, and were obtained after 5 h of electrophoresis (left-hand side of autoradiogram), and after 3 h of electrophoresis (right-hand side of the autoradiogram). Manual DNA-sequencing was as described in Section 3.2.19. The gel was loaded GATC from left to right and the nucleotide sequences are read from the bottom of the gel.

Figure 3.3.9 **Nucleotide Sequence of the *B. stearothermophilus* DNA Insert in pADH2 and in mp18*adh* (Antisense strand)**

The nucleotide sequence of the antisense strand (upper-case letters) is that obtained when the results of automated sequencing of pADH2 DNA (nucleotides 38 - 469 inclusive), and manual sequencing of single-stranded mp18*adh* DNA (nucleotides 652 - 670 inclusive, 901 - 920 inclusive and 1265 - 1531 inclusive), were superimposed on the sequence data obtained for the antisense strand by manual sequencing of plasmid pADH2 (Figure 3.3.7). Two unidentified nucleotide bases remained and those are indicated by the letter N (in red). The sequences of primers P2, P4, P6, and P8 are underlined, and the sequence data have been extrapolated back (lower-case letters) to show the sequence of the M13/pUC forward primer (FP, underlined). *B. stearothermophilus* DNA extends from, and includes, nucleotide 47 to, and including, the *Sau*3A site.

FP

gttttcccagtcacgacgttgtaaaacgacg

*Bam*HI

```
   1 GCCAGTGCCAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCGGTCAAACGCTTCTC   60

  61 TGGCTGCCTTCACAGCGCGGTCGATATCTTCAGCATCTCCTTCATAAACTGTTGCCAATC  120

                                                              P2
 121 TTTCTCCCGTTGCAGGGTTAGGAGTGTCAAATGTTTTTCCTGAAGCGCTTTCAACAAAGC  180

 181 TGCCGTTTATATACATTTTTTTCTTGCCGGATAAAAACTTTTCAACTTTTTCTTTTAATT  240

 241 TAGGGGCTGCAATTGAGCTCATCATCAATCCTCCTTCAATTTTAGTACGATACGACCGTT  300

 301 GATTTTTCCTTTTTCCATTCTTTCAAATACTTCATTTATTTCTTCTAATTCAGCGGTTTC  360

 361 AACAATTGGACGGACTTTTCCGCGAGCAGCAAAATCTAATGCTTCCTGCATGTCCTTTCT  420

 421 TGTACCGACGATGGAACCTTTAACGGAAACGCCGTTTAATACTGTATCGAAGATCGGAAT  480

      P4
 481 AGGCAGATCAGCATTTGGCAGACCGACAACAACTAATGTGCCTCCCCGTTTCACTGACTG  540

 541 ATAGGCCTGTTCAAATGCTTTTTTATTCACCGCTACACTAATAGCAGCATGAACTCCGCC  600

 601 TACTTGATCATGTATCGCTTTGACAGGATCTTCTTTTAAACCGTTGATTGCAATATCTGC  660

 661 ACCTAAATCTTTGGCTAATTTTGATTTTTCATCACTGATATCAACAGCTACCACATTAAG  720

 721 TCCCATCGCTTTGGCGTATTGTAATGCAATATGACCAAGTCCTCCAATACCGTAAATGGC  780

                                                         P6
 781 CACCCATTCACCTGGCCGAGCACCTGATACTTTCAACGCTTTATATGTCGTTACGCCTGC  840

 841 ACATAAAATTGGCGCTACTTCAACTGGATCTAAGTTATCAGGAATTTTCGCCACATAATC  900

 901 TGCGGGANNTTTACAGTATTCCGCATATCCGCCATCAACGGAATATCCGCCATTCAATTG  960

 961 ATGCGGGCAAAGTGTTTCTTGACCAGTTAAACAATATTCACATTCACCACATGCAGAATA 1020

1021 TAACCATGGGATGCCGACGCGATCTCCCACTTTTATTGATTTAACTCCCTTTGCGACTTC 1080

                P8
1081 TACAACAATTCCGACTCCTTCATGCCCCGGAATTAACGGCAATTTCGGTTTAATCGGCCA 1140

1141 GTCACCATGGGCTGCATGTAAATCCGTATGGCACACACCGCATGCTTCGATCTTAACTAA 1200

1201 AACTTCACCTTCCTCTAGTTTGGGTCTTTCCACCTCTTTAATTTCAAGTGCTTTTTTAAA 1260

1261 TTCGTTAACTACTGCTGCTTTCACTCAAATCAACCCCTTTCATTTTTTTCGACTTTATAT 1320

1321 CATGCAAGAACCGTACCAAAATTGAAAGCGCTTAATTTGCGGAATATTAGAAAAATACAA 1380

1381 AATAGAATTTTGCCGAATTTTTGGCAGACTTGAACAAAATATTGGCAGAATTGTTGCATG 1440

1441 AAAAGAAAAAGCAGGTTGTTCCAGTTGGAAACCTGCTAAAAATATAGTGTTCAGTTTTTG 1500

          *Sau*3A
1501 CTCAAGCATGAATGATCCCCGGGTACCGAGCTC                              1532
```

Figure 3.3.9

Figure 3.3.10 **Nucleotide Sequence of the *B. stearothermophilus* DNA Insert in pADH2 (Sense strand)**

The data (nucleotides 1 - 1498 inclusive, upper-case letters) were obtained by manual double-stranded DNA sequencing of pADH2 (Section 3.2.19). Unidentified nucleotide bases are indicated by the letter N (in red). The sequences of primers P1, P3, P5, P7 and P9 are underlined, and restriction sites relevant to establishing the integrity of the cloned fragment are indicated. The sequence data have been extrapolated back (lower-case letters) to show the sequence of the M13/pUC reverse primer (RP, underlined), and to complete the sequence across the pADH1b cloning site. *B. stearothermophilus* DNA extends from, and includes, the *Sau*3A site (GATC) to, and including, nucleotide 1485. The pADH1b polylinker site extends from, and includes, the *Eco*RI site (presence confirmed by restriction analysis; Figure 3.3.5) to, and including, the *Sma*I site.

RP *Eco*RI

agcggataacaatttcacacaggaaacagctatgaccatgattacgaattcga

*Sma*I *Sau*3A

1 GCTCGGTACCCGGGGATCATTCATGCTTGAGCAAAAACTGAACACTATATTTTTAGCAGG 60

P1

61 TTTCCAACTGGAACAACCTGCTTTTTCTTTTCATGCAACAATTCTGCCAATATTTTGTTC 120

121 AAGTCTGCCAAAAATTCGGCAAAATTCTATTTTGTATTTTTCTAATATTCCGCAAATTAA 180

181 GCGCTTTCAATTTTGGTACGGTTCTTGCATGATATAAAGTCGAAAAAAATGAAAGGGGTT 240

241 GATTTGAGTGAAAGCAGCAGTAGTTAACGAATTTAAAAAANCACTTGAAATTAAAGAGGT 300

P3

301 GGAAAGACCCAAACTAGAGGAAGGTGAAGTTTTAGTTAAGATCGAAGCATGCGGTGTGTG 360

361 CCATACGGATTTACATGCAGCCCATGGTGACTGGCCGATTAAACCGAAATTGCCGTTAAT 420

421 TCCGGGGCATGAAGGAGTCGGAATTGTTGTAGAAGTCGCAAAGGGAGTTAAATCAATAAA 480

481 AGTGGGAGATCGCNTNNNCATCCCATGGTTATATTCTGCATGTGGTGAATGTGAATATTG 540

P5

541 TTTAACTGGTCAAGAAACACTTTGCCCGCATCAATTGAATGGCGGATATTCCGTTGATGG 600

601 CGGATATGCGGAATACTGTAAAGCTCCCGCAGATTATGTGGCGAAAATTCCTGATAACTT 660

661 AGATCCAGTTGAAGTAGCGCCAATTTTATGTGCAGGCGTAACGACATATAAAGCGTTGAA 720

721 AGTATCAGGTGCTCGGCCAGGTGAATGGGTGGCCATTTACGGTATTGGAGGACTTGGTCA 780

P7

781 TATTGCATTACAATACGCCAAAGCGATGGGACTTAATGTGGTAGCTGTTGATATCAGTGA 840

841 TGAAAAATCAAAATTAGCCAAAGATTTAGGTGCAGATATTGCAATCAACGGTTTAAANNN 900

901 NGATCCTGTCAAAGCGATACATGATCAAGTAGGCGGAGTTCATGCTGCTATTAGTGTAGC 960

961 GGTGAATAAAAANNCATTTGAACNGGCCTATCAGTCAGTGAAACGGGGAGGCACATTAGT 1020

1021 TGTTGTCGGTCTGCCAAATGCTGATCTGCCTATTCCGATCTTCGATACAGTATTAAACNG 1080

P9

1081 CGTTTCCGTTAAAGGTTCCATCGTCGGTACAAGAAAGGACATGCAGGAAGCATTAGATTT 1140

1141 TGCTGCTCGCGGAAAAGTCCGTCCAATTGTTGAAACCGCTGAATTAGAAGAAATAAATGA 1200

1201 AGTATTTGAAAGAATGGAAAAAGGAAAAATCAACGGTCGTATCGTACTAAAATTGAAGGA 1260

1261 GGATTGATGATGAGCTCAATTGCAGCCCCTAAATTAAAAGAAAAAGTTGAAAAGTTTTTA 1320

1321 TCCGGCAAGAAAAAAATGTATATAAACNGCAGCTTTGTTGAAAGCNCTTCAGGAAAAACA 1380

1381 TTTGACACTCCTAACCCTGCAACGGGAGAAAGATTGGCAACAGTTTATGAAGGAGATNCT 1440

*Bam*HI

1441 GAAGATATCGACCNCNCTGTGAANNCAGCCAGAGAAGCGTTTGACCGGATCCTCTAGA 1498

Figure 3.3.10

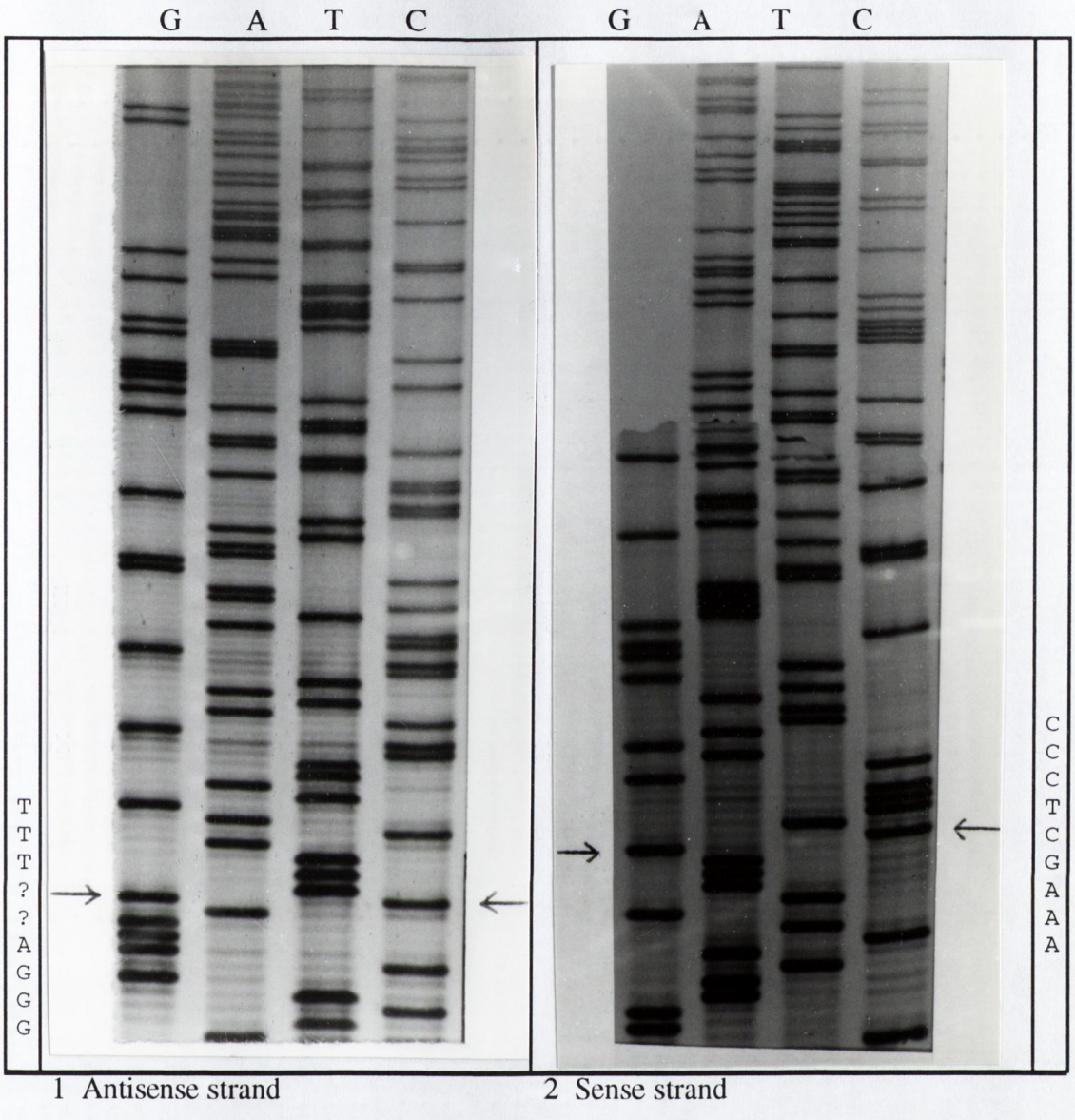


Figure 3.3.11 **Resolution of the G/C Compression on the Antisense Strand**

The data for the antisense strand (1) were obtained using the primer P6 in the presence of Mn^{2+}. The data for the sense strand were obtained using the primer P5 in the presence of Mn^{2+}. Sequencing gels were loaded GATC from left to right and the nucleotide sequence is read from the bottom of the gel. Overlapping regions of the antisense and sense strands are shown. On the antisense strand (1) the sequence begins A, nucleotide 898 (Figure 3.3.9). On the sense strand (2) the sequence begins C, nucleotide 609 (Figure 3.3.10). A compression of G and C (arrowed) was present on the antisense strand at nucleotides 908/909. The compression was unambiguously resolved into G followed by C (nucleotides 623 - 624) in the sequence data from the sense strand.

3.4 DISCUSSION

3.4.1 DNA Sequencing in pADH1b

Plasmid pADH1b contained *adh*2334 on a 1.5-kb *Eco*RI-*Bam*HI restriction fragment (Dowds *et al.*, 1988; Section 3.1.1). *E. coli* TG1 cells that were transformed with pADH1b produced the ADH 2334 antigen, and the plasmid DNA purified from the transformants carried an *Eco*RI-*Bam*HI restriction fragment of the correct size (Section 3.3.1). However, *adh*2334 could not be sequenced in pADH1b.

Sequence data were obtained from pADH1b when priming was initiated from the forward priming site into the vector backbone (Section 3.3.1). Therefore, the absence of data, when sequencing was initiated from the reverse priming site into the *B. stearothermophilus* insert, was not due to some systematic error in the methodology of the sequencing reactions. The sequence data obtained from pADH2 established that the pADH1b multiple cloning site, and the *B. stearothermophilus* DNA downstream from the multiple cloning site, was intact (see Section 3.4.2.2 below). Therefore, the problem encountered when attempting to sequence *adh*2334 in pADH1b was due to some property of the DNA upstream from the *Eco*RI cloning site. It seems most likely that the reverse priming site was deleted, either during the construction of pRK59 or during the cloning of *adh*2334. If this latter explanation were correct, the size of the deleted fragment would have had to have been small enough not to have affected the observed size of the vector backbone following restriction (Figure 3.3.1).

3.4.2 Analysis of the Nucleotide Sequence of the *B. stearothermophilus* insert in pADH2

3.4.2.1 Confirmation that pADH2 Contained the Complete *adh*2334 Coding Sequence

Analysis of the complete sequence of the sense strand of the *B. stearothermophilus* insert in pADH2 revealed two consecutive open reading frames (Figure 3.4.1). The first (major open reading frame, underlined) is initiated from GTG (red) at 234 - 236, encodes a 339 residue protein sequence (M_r 36,205) and is followed by a stop codon (TGA, red, overlined). The second, incomplete, open reading frame is initiated from either, the ATG at 1253 - 1255 (red) or, from the ATG at 1256 - 1258, and encodes a sequence of seventy-three (or seventy-two) amino acids. Putative transcription terminator sequences (126 - 140 and 1338 - 1351, green), -35 and -10 promoter elements (179 - 184 and 199 - 204 respectively, magenta), and two ribosome-binding sites (217 - 222 and 1242 - 1248, blue) were also identified.

Figure 3.4.1 **Translation of the *B. stearothermophilus* Nucleotide Sequence**

1 - 4 inclusive correspond to residues 15 - 18 inclusive in Figure 3.3.10 (Section 3.3.3), and are the *Sau*3A site that denotes one end of the *B. stearothermophilus* DNA cloned by Dowds *et al.* (1988). The *adh*2334 coding sequence is underlined, with the start codon (GTG) indicated in red and the stop codon (TGA) indicated in red and overlined in black. The first of two alternative start codons for a second (incomplete) open reading frame, downstream from the *adh*2334 coding sequence, overlaps with the third nucleotide of the previous stop codon and is also indicated in red.

Inverted repeat sequences, suggestive of transcription terminator sequences, are indicated in green (126 - 140 and 1338 - 1351). Putative -35 and -10 promoter elements (179 - 184 and 199 - 204, respectively) are indicated in magenta. Putative ribosome binding sites are indicated in blue (217 - 222 and 1242 - 1248).

1 GATCATTCATGCTTGAGCAAAAACTGAACACTATATTTTTAGCAGGTTTCCAACTGGAACAACCTGCTTTTTCTTTTCATGCAACAATTCTGCCAATATT 100
-35
101 TTGTTCAAGTCTGCCAAAAATTCGGCAAAATTCTATTTTGTATTTTTCTAATATTCCGCAAATTAAGCGCTTTCAATTTTGGTACGGTTCTTGCATGATA 200
-10 start
201 TAAAGTCGAAAAAAAATGAAAGGGGTTGATTTGAGTGAAAGCAGCAGTAGTTAACGAATTTAAAAAAGCACTTGAAATTAAAGAGGTGGAAAGACCCAAAC 300
M K A A V V N E F K K A L E I K E V E R P K
301 TAGAGGAAGGTGAAGTTTTAGTTAAGATCGAAGCATGCGGTGTGTGCCATACGGATTTACATGCAGCCCATGGTGACTGGCCGATTAAACCGAAATTGCC 400
L E E G E V L V K I E A C G V C H T D L H A A H G D W P I K P K L P
401 GTTAATTCCGGGGCATGAAGGAGTCGGAATTGTTGTAGAAGTCGCAAAGGGAGTTAAATCAATAAAAGTGGGAGATCGCGTCGGCATCCCATGGTTATAT 500
L I P G H E G V G I V V E V A K G V K S I K V G D R V G I P W L Y
501 TCTGCATGTGGTGAATGTGAATATTGTTTAACTGGTCAAGAAACACTTTGCCCGCATCAATTGAATGGCGGATATTCCGTTGATGGCGGATATGCGGAAT 600
S A C G E C E Y C L T G Q E T L C P H Q L N G G Y S V D G G Y A E
601 ACTGTAAAGCTCCCGCAGATTATGTGGCGAAAATTCCTGATAACTTAGATCCAGTTGAAGTAGCGCCAATTTTATGTGCAGGCGTAACGACATATAAAGC 700
Y C K A P A D Y V A K I P D N L D P V E V A P I L C A G V T T Y K A
701 GTTGAAAGTATCAGGTGCTCGGCCAGGTGAATGGGTGGCCATTTACGGTATTGGAGGACTTGGTCATATTGCATTACAATACGCCAAAGCGATGGGACTT 800
L K V S G A R P G E W V A I Y G I G G L G H I A L Q Y A K A M G L
801 AATGTGGTAGCTGTTGATATCAGTGATGAAAAATCAAAATTAGCCAAAGATTTAGGTGCAGATATTGCAATCAACGGTTTAAAAGAAGATCCTGTCAAAG 900
N V V A V D I S D E K S K L A K D L G A D I A I N G L K E D P V K
901 CGATACATGATCAAGTAGGCGGAGTTCATGCTGCTATTAGTGTAGCGGTGAATAAAAAAGCATTTGAACAGGCCTATCAGTCAGTGAAACGGGGAGGCAC 1000
A I H D Q V G G V H A A I S V A V N K K A F E Q A Y Q S V K R G G T
1001 ATTAGTTGTTGTCGGTCTGCCAAATGCTGATCTGCCTATTCCGATCTTCGATACAGTATTAAACGGCGTTTCCGTTAAAGGTTCCATCGTCGGTACAAGA 1100
L V V V G L P N A D L P I P I F D T V L N G V S V K G S I V G T R
1101 AAGGACATGCAGGAAGCATTAGATTTTGCTGCTCGCGGAAAAGTCCGTCCAATTGTTGAAACCGCTGAATTAGAAGAAATAAATGAAGTATTTGAAAGAA 1200
K D M Q E A L D F A A R G K V R P I V E T A E L E E I N E V F E R
1201 TGGAAAAAGGAAAAATCAACGGTCGTATCGTACTAAAATTGAAGGAGGATTGATGATGAGCTCAATTGCAGCCCCTAAATTAAAAGAAAAAGTTGAAAAG 1300
M E K G K I N G R I V L K L K E D M M S S I A A P K L K E K V E K
1301 TTTTTATCCGGCAAGAAAAAAATGTATATAAACGGCAGCTTTGTTGAAAGCGCTTCAGGAAAAACATTTGACACTCCTAACCCTGCAACGGGAGAAAGAT 1400
F L S G K K K M Y I N G S F V E S A S G K T F D T P N P A T G E R
1401 TGGCAACAGTTTATGAAGGAGATGCTGAAGATATCGACCGCGCTGTGAAGGCAGCCAGAGAAGCGTTTGAC 1471
L A T V Y E G D A E D I D R A V K A A R E A F D

Figure 3.4.1

The amino-acid sequence data obtained for ADH 2334, and for ADH 1503 (Chapter 2), the amino-acid sequence deduced from the *adh*1503 nucleotide sequence (Sakoda & Imanaka, 1992) and the amino-acid sequence deduced from the major open reading frame in pADH2 are combined in Figure 3.4.2. The first forty amino acids encoded by the major open reading frame correspond exactly to the amino-acid sequence of ADH 2334, determined using the purified protein, thus confirming that the major open reading frame did encode ADH 2334.

Because of the difficulty encountered in sequencing *adh*2334 in pADH1b, and as the nucleotide sequence of *adh*2334 was ultimately determined from DNA that had been excised from pADH1b, the data from nucleotide sequencing and from protein sequencing (Figure 3.4.2) were scrutinised to ensure that there was no evidence that the *adh*2334 nucleotide sequence was corrupted. The exact agreement between the N-terminal amino-acid sequence of ADH 2334 (obtained from the protein) and the amino-acid sequence deduced from the 5' end of the major open reading frame, confirmed that the 5' end of *adh*2334 was intact. The close identity between the C-terminal amino-acid sequence of ADH 1503 (obtained from the protein) and the C-terminal amino-acid sequence of ADH 2334 (deduced from the nucleotide sequence) strongly suggested that the 3' end of *adh*2334 was intact. The ADH 2334 sequence aligned with the ADH 1503 sequence without the introduction of any insertions or gaps. Furthermore, the amino-acid sequence deduced for ADH 2334 contained regions that were similar to, and aligned with, the internal peptides of ADH 1503 that had been confirmed by protein sequencing. It was concluded that the *adh*2334 coding sequence was intact in pADH2.

3.4.2.2 Confirmation that the Noncoding Upstream Nucleotide Sequence in pADH2 was Intact.

Plasmid pADH2 was constructed by inserting the 1.5-kb *Eco*RI-*Bam*HI fragment of pADH1b into pUC18. The insert in pADH2 was therefore expected to contain the multiple cloning site of pADH1b (Section 3.3.1; Figure 3.3.2). The sequence data for the sense strand of the *B. stearothermophilus* DNA in pADH2 (Section 3.3.3; Figure 3.3.10) includes the cloning site of pADH1b from (and including) the *Sma*I site (residues 9 - 14 inclusive) back to (and including) the GCTC (residues 1 - 4) of the GAGCTC *Sac*I site. The presence of the *Eco*RI site was demonstrated by restriction analysis (Section 3.3.2; Figure 3.3.5). Allowing for the two residues of the *Sac*I site that were outside the range of the sequence data, one can conclude that the sequence across the multiple cloning site of pADH1b was intact.

*adh*2334, as cloned by Dowds *et al.* (1988), was expressed from its own promoter, therefore, sequence motifs characteristic of promoter elements and a ribosome-binding site were expected upstream from the coding sequence, and these were present (Figure 3.4.1 and below). The *Sau*3A site, which denotes one end of

```
ADH 2334    1 MKAAVVNEFKKALEIKEVERPKLEEGEVLVKIEACGVCHTDLHAAHGDWP  50
ADH 1503    1 MKAAVVEQFKKPLQVKEVEKPKISYGEVLVRIKACGVCHTDLHAAHGDWP  50
              ******  *** *  **** ***  ***** * ******************

ADH 2334   51 IKPKLPLIPGHEGVGIVVEVAKGVKSIKVGDRVGIPWLYSACGECEYCLT 100
ADH 1503   51 VKPKLPLIPGHEGVGVIEEVGPGVTHLKVGDRVGIPWLYSACGHCDYCLS 100
               **************   **  **   ***************** * ***

ADH 2334  101 GQETLCPHQLNGGYSVDGGYAEYCKAPADYVAKIPDNLDPVEVAPILCAG 150
ADH 1503  101 GQETLCERQQNAGYSVDGGYAEYCRAAADYVVKIPDNLSFEEAAPIFCAG 150
              ******  * * ************ * **** ******   * *** ***

ADH 2334  151 VTTYKALKVSGARPGEWVAIYGIGGLGHIALQYAKAMGLNVVAVDISDEK 200
ADH 1503  151 VTTYKALKVTGAKPGEWVAIYGIGGLGHVAVQYAKAMGLNVVAVDLGDEK 200
              ********* ** *************** * **************  ***

ADH 2334  201 SKLAKDLGADIAINGLKEDPVKAIHDQVGGVHAAISVAVNKKAFEQAYQS 250
ADH 1503  201 LELAKQLGADLVVNPKHDDAAQWIKEKVGGVHATVVTAVSKAAFESAYKS 250
                *** ****   *    *    *   ******    ** * *** ** *

ADH 2334  251 VKRGGTLVVVGLPNADLPIPIFDTVLNGVSVKGSIVGTRKDMQEALDFAA 300
ADH 1503  251 IRRGGACVLVGLPPEEIPIPIFDTVLNGVKIIGSIVGTRKDLQEALQFAA 300
                ***  * ****    ************   ********* **** ***

ADH 2334  301 RGKVRPIVETAELEEINEVFERMEKGKINGRIVLKLKED            339
ADH 1503  301 EGKVKTIVEVQPLENINDVFDRMLKGQINGRVVLKVD..            337
               ***  ***   ** ** ** ** ** **** ***
```

Figure 3.4.2 **Alignment of the ADH 2334 and ADH 1503 Amino-Acid sequences**

The amino-acid sequence of ADH 2334 is that deduced from the major open reading frame (Figure 3.4.1). The amino-acid sequence of ADH 1503 was deduced from the gene sequence of Sakoda & Imanaka (1992). The amino-acid sequences that were confirmed by protein sequencing (Section 2.4; Figure 2.4.1) are underlined. Residues conserved between the two sequences are indicated by an asterisk.

the *B. stearothermophilus* DNA cloned by Dowds *et al.* (1988), is also present (legend to Figure 3.4.1). There is therefore no evidence that the nucleotide sequence between the cloning site and the coding DNA was corrupted.

3.4.2.3 Ribosome-Binding Sites

Putative ribosome-binding sites (Shine-Dalgarno sequences; Shine & Dalgarno, 1974) were identified by aligning the nucleotide sequences upstream from the *adh*2334 start codon, and from the *adh*2334 stop codon, with the complement of the RNA sequence of the 3' end of *B. stearothermophilus* 16S rRNA. The alignments were performed to give maximum Shine-Dalgarno complementarity (McLaughlin *et al.*, 1981). Two putative ribosome-binding sites were identified (Figure 3.4.3). The first was at nucleotides 217 - 222 inclusive, and was assumed to be the ribosome-binding site for *adh*2334. The second, at nucleotides 1242 - 1248 inclusive, was within the *adh*2334 open reading frame, and was assumed to be the ribosome-binding site for the second open reading frame. The positions of the two ribosome-binding sites, relative to the *adh*2334 coding sequence, are shown in Figure 3.4.1.

A search of *B. stearothermophilus* genes in the GenBank database revealed no spacing longer than ten bases for the distance between the *B. stearothermophilus* ribosome-binding site and the start codon. The spacing between the *adh*2334 ribosome-binding site and the *adh*2334 start codon (eleven bases) is, therefore, slightly longer than usual. The putative ribosome-binding site for the second open reading frame is shown relative to the first of the two possible start codons. The spacing between the selected ribosome-binding motif and the first start codon is short (four nucleotides) but there are four-nucleotide putative ribosome-binding site to start codon spacings in the literature, for example in *B. stearothermophilus dnaK* (Figure 3.4.3) and in the *B. stearothermophilus cel* operon (Lai & Ingram, 1993; Figure 3.4.3).

3.4.2.4 Transcription Promoter Elements

A consensus promoter sequence, deduced from the nucleotide sequences of the promoters of nine *B. stearothermophilus* genes, has been reported by Wu & Welker (1991). The consensus sequence of the -35 element is T_{75} T_{75} G_{75} A_{50} C_{50} t/c, where subscript 75 indicates that this base occurs in more than 75% of cases and subscript 50 indicates that this base occurs in 50 - 75% of cases. Weakly conserved positions (occurring in less than 50% of cases) are indicated by lower-case letters. A diagonal line indicates that the position can contain either residue in 50% of cases. The consensus sequence of the -10 element, which occurs 17 - 20 bases downstream from the -35 element, is T_{75} A_{75} T_{75} T_{50} a/c T_{50}.

The putative -10 and -35 promoter elements of *adh*2334 were identified by aligning the upstream noncoding nucleotide sequence with *B. stearothermophilus*

bstrna	1	......TAGAAAGGAGGTGATC.....	16
*adh*2334 RBS	209	AAAAAAATGAAAGGGGTTGATTTGAGTG.	236
*aldh*2334 RBS	1235	...AAAATTGAAGGAGGATTGATGATG..	1255
*adh*1503	447	.TTTGCCTAGAAGGAGGATTATAATCATG	474
nplt	85	.TTATATATGAAGGAGGAGAGGCCAAATG	112
glnqh	314	ATTTATGAAAAAGGAGGGGGGTCTATTG.	341
qcrA	564	....TGCATAGAGGGGGGAAGACGATG..	568
grpE	238	....CTTTAAAGGGAGGTGAAACCGATG.	261
celD	3961	.....AACTATAGGAGAGGGCGTTATG..	3982
dnaK	930	........TATAGGAGGTGACGTAATG..	948

Figure 3.4.3 **Identification of Ribosome-binding Sites**

The nucleotide sequence upstream from the *adh*2334 start codon (*adh*2334 RBS) and the nucleotide sequence preceding the second open reading frame (*aldh*2334 RBS) were aligned with the complement of the 3' end of the 16S ribosomal RNA of *B. stearothermophilus* (bstrna, Van Charldorp *et al.*, 1981). The sequences identified as putative ribosome-binding sites are indicated in blue. Start codons are indicated in red. For comparison, similar alignments are shown for *adh*1503 and for six other *B. stearothermophilus* sequences from the GenBank database. Ribosome-binding sites (blue) and start codons (red) are as identified in the original reference, and numbering (with the exception of bstrna) is the same as in the GenBank database. Abbreviations used are: *nplt*, the gene encoding neopullulanase (Kuriki & Imanaka, 1989); *glnqh*, the sequence immediately upstream from the first open reading frame of the glutamine transport operon (Wu & Walker, 1991); *qcrA*, the gene encoding the FeS protein of cytochrome reductase (Sone *et al.*, 1996); *celD*, the sequence immediately upstream from the fourth open reading frame of the cellobiose phosphotransferase system operon (Lai & Ingram, 1993); *grpE* and *dnaK*, the sequence, immediately upstream from the second and third open reading frames of the *dnaK* (heat-shock proteins) operon, respectively (Herbort *et al.*, 1996).

sequences that were designated as promoter elements in the GenBank database (Figure 3.4.4). The -10 element selected for *adh*2334 (TATAAA, nucleotides 199 - 204, magenta) is consistent with the Wu & Welker consensus sequence, and is also consistent with -10 elements identified for other *B. stearothermophilus* genes. Selection of a -35 sequence was difficult and the designation of nucleotides 77 - 82 (TTGGTA, magenta) as the -35 element is less secure than the designation of the -10 element. TTGGTA was selected because it contained the motif TTG and, therefore, seemed consistent with the consensus sequence and with other -35 elements presented in Figure 3.4.4. The distance between the promoter elements selected for *adh*2334 (fourteen bases) is shorter than that predicted by the Wu & Welker consensus sequence. Alternative selections for the *adh*2334 -35 element, which gave -35 to -10 spacings of seventeen, eighteen or nineteen bases, resulted in only single base identity with the -35 consensus sequence and, therefore, were disregarded. Selections for the -35 element that were based on -35 to -10 spacings longer than nineteen bases did not contain the TTG motif. Assuming that the inclusion of the TTG motif within the -35 element is a valid selection procedure, and given that the -10 element selected for *adh*2334 is consistent with the Wu & Welker consensus sequence and alignes with other *B. stearothermophilus* -10 elements, there does not appear to be an alternative to the arrangement shown in Figure 3.4.4. Unfortunately, a comparison with *adh*1503 was not possible, as Sakoda & Imanka (1992) did not find promoter elements upstream from the *adh*1503 coding sequence.

3.4.2.5 Transcription Termination Sites

Upstream from the -35 promoter sequence, there is a six-bp inverted repeat (Figure 3.4.1; nucleotides 126 - 140, green) followed by five consecutive Ts. Such a motif is suggestive of a rho-independent transcription terminator (Rosenberg & Court, 1979). Twelve *B. stearothermophilus* sequences annotated in the GenBank database as terminators, or potential terminators, were examined. The stem length of a hairpin loop formed from these inverted repeat sequences varied from ten to twenty bp, with the average stem length being fourteen bp.

Rosenberg & Court (1979) described twenty DNA sequences, of phage and bacterial origin, containing (or presumed to contain) transcription termination sites. The stem length of the hairpin loops formed from the inverted repeat sequences varied from three to eighteen bp, with an average stem length of eleven bp. A notable characteristic of the inverted repeat sequences described by Rosenberg & Court, and of the *B. stearothermophilus* sequences extracted from the database, is the presence, within the potential stem loop, of a least two contiguous G/C bp. The potential terminator sequence identified upstream from *adh*2334 appears, therefore, to be shorter than is usual, and does not contain contiguous G/C bp.

```
                                  -35                -10
adh2334  163 .AAGCGCTTTCAATTTTGGTACGGTTCTTGCATGATATAAAGTCGAAAAAA 214
mnsod    287 CAAGCAAGTCGATTGAAACATTGTGCCAACTTTGGTAAGCTA......... 330
trps     236 CAATAAATAGGGTGGCACCGCGGTCCTACCGTCCCTATCAGAGGGCTGGCA 288
bgab     377 ......TTGTTGTTGACAAATACTAAATTTTAACTTAATTTATAATTAAA. 422
nplt       9 ...CTACTGAATTTGACTTTTTTTCTCTCTTTTCGTACTCTTTAATCAGAG  58
npga     114 .TTGCCGTTCATTTTCCCAATACAATAAGGATGACTATTTTGGTAAAATTC 165
tnp       50 ....ACTCTATTTTTCCCAATACAAATACTGTAACTATTTTGGTAAAATTC  98
dac       29 ......TCAACTTTGACGGAGACATGTCCTCCGTCTATTTTTTTGG.....  70
cgt1       6 ..........CTTGACACTGAGTAATATACTGAATATTATAAAAATAAAG  47
mal       18 ...CAATAAGGGTTGTCTTTTTCCGGTTCATAGGATAACATTTGTACCGTA  67
```

Figure 3.4.4 **Identification of *adh*2334 Promoter Elements**

B. stearothermophilus gene upstream-sequences containing motifs designated as -10 and -35 elements (indicated in magenta) were selected from the GenBank database and arranged so that all -10 elements aligned. The upstream nucleotide sequence of *adh*2334 was then added, to obtain the best possible alignment with the maximum number of -10 elements. The -35 element for *adh*2334 was selected because it was the six amino-acid motif with the highest sequence identity to the Wu & Welker consensus sequence (Wu & Welker, 1991).

The -35 and -10 elements in the nucleotide sequences taken from the database are as identified in the original reference. Numbering is the same as in the GenBank database. Abbreviations used are: *mnsod*, the gene encoding manganese superoxide dismutase (Bowler *et al.*, 1990); *trps*, the gene encoding tryptophanyl synthetase (Barstow *et al.*, 1986); *bgab*, the gene encoding β-galactosidase 1 (Hirata *et al.*, 1986); *nplt*, the gene encoding neopullulanase (Kuriki & Imanaka, 1989); *npga*, the gene encoding neutral protease (Takagi *et al.*, 1985); *tnp*, the gene encoding thermolysin (Kubo & Imanaka, 1988); *dac*, the gene encoding D-alanine carboxypeptidase (Despreaux & Manning, 1993); *cgt1*, the gene encoding cyclodextrin glucanotransferase (Fujiwaara *et al.*, 1992); *mal*, the gene encoding malic enzyme (Kobayashi *et al.*, 1989).

The free energy of the stem loop structure that would form from nucleotides 126 - 140 was calculated, by the method described by Cantor & Schimmel (1980), to be +1.3 kcal mol^{-1}, which indicates that this nucleotide sequence cannot form a stem loop structure and is, therefore, not a transcription terminator. Both strands of the DNA upstream from the *adh*2334 start codon were translated, in their six possible open reading frames, and subjected to a BLAST homology search. No significant homology was detected between any one of the upstream translations and any protein in the GenBank, Swiss-Prot or PIR databases.

Downstream from *adh*2334, there is another inverted repeat sequence (Figure 3.4.1; nucleotides 1338 - 1351, green). This inverted repeat is not followed by consecutive Ts, indicating a rho-dependent transcription terminator (Rosenberg & Court, 1979). The free energy of the stem loop structure that would form from nucleotides 1338 - 1351 was calculated to be -1.5 kcal mol^{-1}. Furthermore, the sequence contains two contiguous G/C pairs and, therefore, fulfils the criteria for a rho-dependent transcription terminator.

3.4.2.6 Translation of the Second Open Reading Frame: Identification of a Downstream Aldehyde Dehydrogenase

The amino-acid sequence encoded by the second open reading frame in Figure 3.4.1 (initiated by the first ATG at 1253 - 1255, red) was submitted to a BLAST homology search of the GenBank database. The seventy-three amino-acid sequence aligned with the N-termini of aldehyde dehydrogenases (ALDHs; Figure 3.4.5) and, therefore, the second open reading frame in pADH2 was designated *aldh*2334.

The seventy-three amino-acid sequence of the putative *B. stearothermophilus* ALDH (ALDH 2334) has 46% identity with the N-terminus of human cytosolic ALDH, 36% identity with the N-terminus of human mitochondrial ALDH, 43% identity with the N-terminus of rat and mouse ALDH, but only 26% identity with the N-terminus of ALDH from *B. stearothermophilus* 1503.

In Figure 3.4.1, the initiation codon of the second open reading frame (nucleotides 1253 - 1255) overlaps with the *adh*2334 termination codon (1251 - 1253). An alternative open reading frame is that initiated at nucleotides 1256 - 1258. The alternative open reading frame eliminates the ALDH 2334 "extra" methionine in the alignment presented in Figure 3.4.5, however, that does not alter the argument that a gene encoding ALDH is situated immediately downstream from *adh*2334.

3.4.3 Genes Encoding Alcohol Dehydrogenase and Aldehyde Dehydrogenase Constitute Part of an Operon

The *E. coli trp* operon, in which the five genes involved in tryptophan biosynthesis are transcribed under the control of a single promoter, is the classic

```
ALDH 2334   1  MMSSIAAPKLKEKVEKFLSGKKKMYINGSFVESASGKTFDTPNPAT.GERLATVYEGDAEDIDRAVKAAREAFD  73
MALDH       1  .MSSPAQPRVPAPLADLKIQHTKIFINNEWHNSVSGKKFPVLNPAT.EEVICHVEEGDKADVDKAVKAARQAFQ  72
RALDH       1  .MSSPAQPAVPAPLANLKIQHTKIFINNEWHNSLNGKKFPVINPAT.EEVICHVEEGDKADVDKAVKAARQAFQ  72
HALDH1      1  .MSSSGTPDLPVLLTDLKIQYTKIFINNEWHDSVSGKKFPVFNPAT.EEELCQVEEGDKEDVDKAVKAARQAFQ  72
HALDH2      1  ..SAAATQAVPAPNQQPEVFCNQIFINNEWHDAVSRKTFPTVNPST.GEVICQVAEGDKEDVDKAVKAARAAFQ  71
ALDH 1503   1  ............MKVQTEIKTYFNYINGNWVSSVSNNVEPSINPANRHDIVGYVQRSTLEDVNEAVTAANEA..  60
                                **              **        *     *   ** **  *
```

Figure 3.4.5 **Alignment of Aldehyde Dehydrogenase Sequences**

The amino-acid sequence designated ALDH 2334 is that resulting from translation of the second open reading frame, initiated at 1253 - 1255, in Figure 3.4.1. The sequence is shown aligned with selected ALDH N-terminal amino-acid sequences from the GenBank database. Amino acids that are conserved in all six sequences are indicated by an asterisk. MALDH is mouse ALDH (Rongnoparut & Weaver, 1991), RLADH is rat ALDH (Dunn *et al.*, 1989), HALDH1 and HALDH2 are human cytoplasmic and human mitochondrial ALDH, respectively (Hsu & Chang, 1991), ALDH 1503 is *B. stearothermophilus* 1503 ALDH (Imanaka *et al.*, 1993).

example of prokaryotic polycistronic mRNA (reviewed by Yanofsky *et al.*, 1981). The nucleotide sequence across the overlap between *adh*2334 and *aldh*2334, and a selection of intercistronic nucleotide sequences from bacterial operons are presented in Figure 3.4.6.

A survey of *Bacillus* operons in the literature, including those presented in Figure 3.4.6, revealed considerable variation in the distance between the stop and start codons of adjacent genes, and variation in the relative positions of ribosome-binding sites and stop codons. Adjacent genes may be separated by as many as ninety-six noncoding nucleotides (the *B. stearothermophilus dnaK* operon; Herbort *et al.*, 1996), or by as few as three nucleotides (*B. stearothermophilus qcr* operon; Figure 3.4.6). The start codon and stop codon of adjacent genes may overlap, as in the *ptsHI* and the *sacPA* operons (Figure 3.4.6), or the coding regions of adjacent genes may overlap (*cysE* and *cysS* of *B. subtilis* and *B. stearothermophilus*, Gagnon *et al.*, 1994). In the *B. stearothermophilus ptsHI* operon, the stop codon of *ptsH* and the start codon of *ptsI* overlap in the sequence TAATG (Figure 3.4.6), which is similar to the overlap proposed for *adh*2334 and *aldh* 2334 (TGATG). Two contiguous ATG codons follow the stop codon at the junction between *trpD* and *trpC* in the *E. coli trp* operon (Figure 3.4.6). At the junction between *adh*2334 and *aldh*2334, two contiguous ATG codons also occur, but the codons overlap, rather than follow, the stop codon. N-terminal sequence analysis of the protein product of the *trpC* gene (Li *et al.*, 1975) indicated that only the second ATG codon is translated. This observation does not exclude the possibility that initiation of translation occurs at the first ATG, followed by removal of the first methionine residue. In the absence of protein sequence data for ALDH 2334, it is not possible to speculate further regarding the exact position of the *aldh*2334 start codon.

It is a general feature of bacterial operons that all the genes within an operon are preceded by ribosome-binding sites. Putative ribosome-binding sites were identified upstream from both the *adh*2334 and the *aldh*2334 coding sequence (Figures 3.4.1 and 3.4.3; blue). The putative ribosome-binding site for *aldh*2334 is situated entirely within the coding sequence of *adh*2334 (Figure 3.4.1), and this is consistent with what has been observed for other *B. stearothermophilus* operons. In operons in which start and stop codons overlap, or the intercistronic region is very short, the ribosome-binding site of the distal gene occurs within the coding region of the proximal gene (*B. stearothermophilus pts* and *qcr* operons; Figure 3.4.6), or overlaps with the proximal coding sequence stop codon (*B. stearothermophilus cel* operon; Figure 3.4.6).

Translational coupling, whereby the efficient translation of a downstream gene in an operon is dependent upon the translation of the 3' end of the gene preceding it, was originally described for the *E. coli trpD* gene by Oppenheim &

B. stearothermophilus adh-aldh AATTGAAGGAGGATTGATGATG

B. stearothermophilus ptsH-ptsI TTGGCGAAAGAGGGTCTTGCAGAATAATG

E. coli trpD-trpC GCGGCACGAGGGTAAATGATG

B. subtilis sacP-sacA GAGGAGAGGAAAAAATGACA

B. stearothermophilus qcrA-qcrB GAGGGGAGGCGTAATGCGTG

B. stearothermophilus celC-celD TAGGAGAGGGCGTTATG

Figure 3.4.6 **Intercistronic Regions of Some Bacterial Operons**

The nucleotide sequence at the junction between *adh*2334 and *aldh*2334 (*B. stearothermophilus adh-aldh*) was compared with the sequences across other bacterial intercistronic regions. The ribosome-binding site identified in the *adh-aldh* intercistronic region (Figure 3.4.3) is indicated in blue, as are other ribosome-binding sites. The sequences are arranged so that the start codons (red) of the distal genes are aligned. The stop codons of the proximal genes are underlined.

The *ptsHI* operon (Lai & Ingram, 1995) and the *cel* operon (Lai & Ingram, 1993) encode the genes required for cellobiose uptake by the phospho*enol*pyruvate-dependent phosphotransferase system. The *sacPA* operon (Fouet *et al.*, 1987) encodes enzymes of the sucrose phosphotransferase system. The authors did not identify a ribosome-binding site for the *sacA* gene. The *qcr* operon (Sone *et al.*, 1996) encodes the structural genes of cytochrome reductase. The intercistronic region between the second and third open reading frames of the *E. coli trp* operon (*trpD-trpC*) was taken from Yanofsky *et al.* (1981).

Yanofsky (1980). Subsequently, translational coupling has been demonstrated for many (but not all) downstream genes in *E. coli* operons where stop and start codons overlap, or where intercistronic regions are short (Schümperli *et al.*, 1982; Das & Yanofsky, 1984; Lindahl *et al.*, 1989; Little *et al.*, 1989). Translational coupling has been demonstrated in gene fusions in *B. subtilis* by Zaghloul *et al.* (1985) and by Sprengel *et al.* (1985).

The precise nature of translational coupling is not understood. The presence of the distal gene ribosome-binding site has been shown to be essential for translational coupling in some operons (*E. coli trpA* gene, Das & Yanofsky, 1984; *E. coli* threonine operon, Little *et al.*, 1989). However, translational coupling that was independent of the presence of a functional ribosome-binding site has been demonstrated in gene fusions in *B. subtilis* (Sprengel *et al.*, 1985). In that instance, translational coupling was dependent on the distance between the stop and start codons, being greatest when start and stop codons overlap in the sequence ATGA. A correlation between increased translational efficiency and decreased distance between stop and start codons has also been demonstrated by Schümperli *et al.* (1982) and by van de Guchte *et al.* (1991). Such observations have led to a model in which either the terminating ribosome (or at least the 30S subunit) reinitiates without being released from the RNA, or the terminating ribosome releases and leads to an increase in the local concentration of ribosomal subunits.

The overlap of the *adh* 2334 stop and the *aldh*2334 start codons, the position of the *aldh*2334 ribosome-binding site within the *adh*2334 open reading frame, and the short spacing between the *aldh*2334 ribosome-binding site and the *aldh*2334 start codon, are clearly consistent with what has been observed in other *Bacillus* operons, and are consistent with translational coupling between *adh* 2334 and *aldh*2334.

The arrangement of the *adh* and *aldh* genes in *B. stearothermophilus* 2334 is very different from that found for *adh* and *aldh* in *B. stearothermophilus* 1503. The *aldh* gene from *B. stearothermophilus* 1503 has its own promoter elements and is not adjacent to *adh*1503 (Imanaka *et al.*, 1993). A third gene encoding a *B. stearothermophilus* ADH (ADH 12403) has been cloned and sequenced (Cannio *et al.*, 1994) from the more thermophilic *B. stearothermophilus* NCIMB12403 (optimum growth temperature, 70 °C, compared to 55 °C for strains 2334 and 1503). The nucleotide sequence (794 bases) downstream from the coding sequence for ADH 12403 was extracted from the Swiss-Prot database, translated in all six reading frames, and the six amino-acid sequences were submitted to BLAST homology searchs to check for the presence of an ALDH. None of the sequences displayed significant homology with ALDH sequences.

Assuming that the function of the *B. stearothermophilus* 2334 *adh-aldh* operon is to direct the carbon skeleton of ethanol into the citric acid cycle, an enzyme (or enzymes) of acetate metabolism would be required to complete the operon. In bacteria (growing aerobically), the conversion of acetate to acetyl-CoA can be achieved either directly, *via* acetyl-CoA synthetase, or indirectly, *via* acetate kinase plus phosphotransacetylase (Figure 3.4.7). In *E. coli*, the major route of acetate conversion to acetyl-CoA has been shown to be *via* acetate kinase plus phosphotransacetylase. Acetyl-CoA synthetase is present in *E. coli*, but is dispensable for growth on acetate (Brown *et al.*, 1977; Kakuda *et al.*, 1994). In *B. subtilis*, however, acetyl-CoA synthetase is essential for growth on acetate [Grundy *et al.*, 1993(a)]. Acetate kinase and phosphotransacetylase are present in *B. subtilis*, but operate primarily in the direction of acetate synthesis during growth in the presence of excess carbohydrate [Grundy *et al.*, 1993(a); Grundy *et al.*, 1993(b)]. Mutants of *B. stearothermophilus* with defects in acetate metabolism have been isolated by Mallinder & Moir (1991). All of the mutants exhibited levels of acetate kinase and phosphotransacetylase that were equivalent to those levels found in wild type cells. The majority of the mutants proved to be defective in either the glyoxylate cycle, amino acid biosynthesis, or acetate uptake, leaving 35% of the total number of mutants unexplained. While these results allow the possibility that, in *B. stearothermophilus*, as in *B. subtilis*, acetate metabolism is mediated by acetyl-CoA synthetase, Mallinder & Moir do not appear to have assayed for that enzyme.

It is feasible that the gene encoding acetyl-CoA synthetase, or alternatively, the genes for acetate kinase and phosphotransacetylase, are situated downstream from *aldh*2334, thus completing a pathway from ethanol to acetyl-CoA (Figure 3.4.7). There do not appear to be any reports of such operons in the literature. In *Acinetobacter calcoaceticus*, a bacterium which can grow on benzyl alcohol as the sole carbon source, a benzaldehyde dehydrogenase and a benzyl alcohol dehydrogenase are coexpressed in response to either substrate (MacKintosh & Fewson, 1988). The enzyme activities result from two distinct proteins (MacKintosh & Fewson, 1988) that are encoded on an operon (Gillooly *et al.*, 1998). In the operon the gene encoding benzaldehyde dehydrogenase is the proximal gene, the coding sequences are separated by 33 bp, and no open reading frames were found either upstream from *aldh* or downstream from *adh*.

An alternative possibility is that the *B. stearothermophilus adh-aldh* operon contains only *adh* and *aldh* and the two enzymes are analogous to long-chain ADHs (Section 1.3.3) in that the ALDH is acetaldehyde-CoA dehydrogenase. Under aerobic conditions ADH 2334 and acetaldehyde-CoA dehydrogenase would catalyse the conversion of ethanol to acetyl-CoA directly, the reverse of the pathway catalysed by anaerobic long-chain ADH. The sequence homology between the N-terminus of

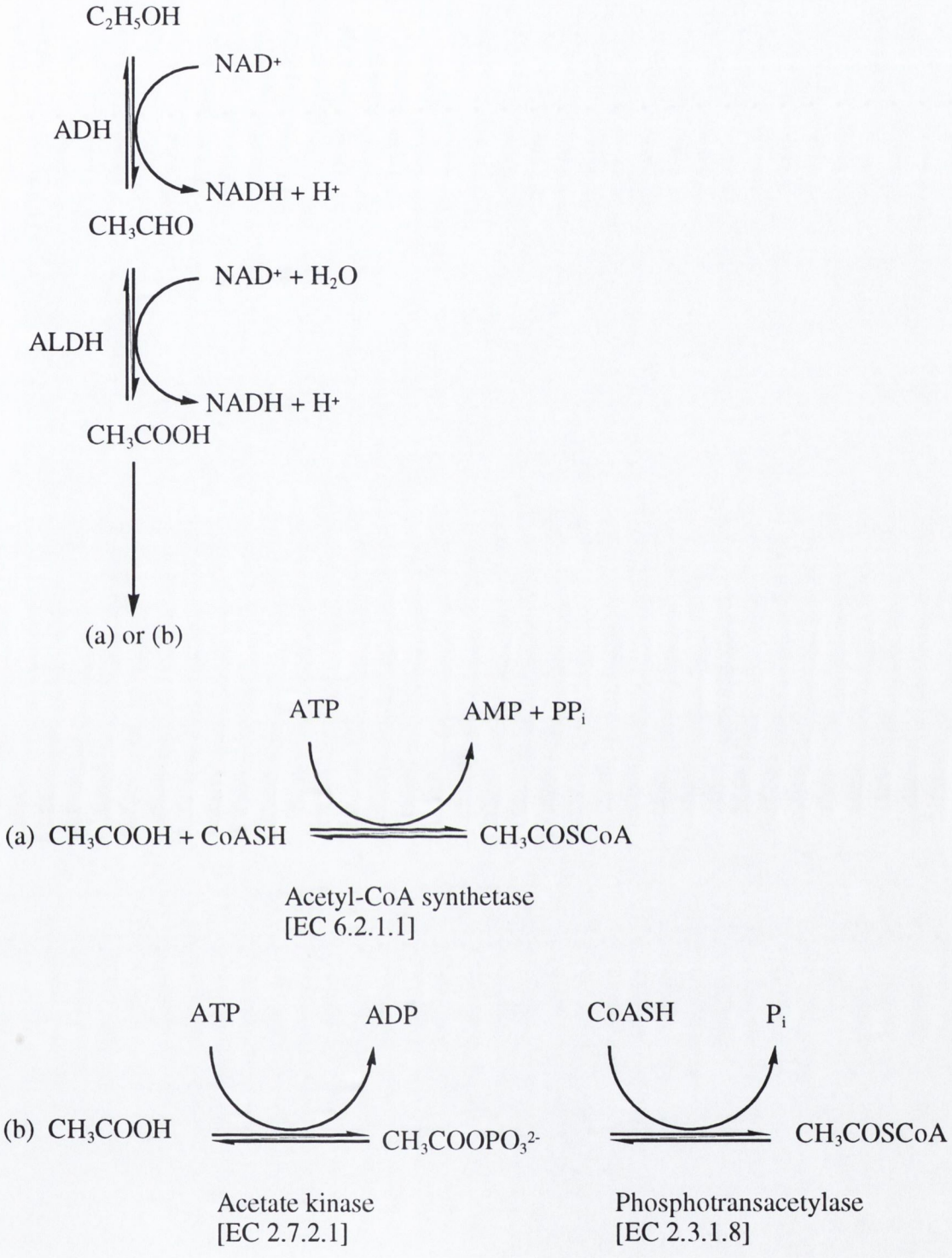


Figure 3.4.7 **Conversion of Ethanol to Acetyl-CoA**

ALDH 2334 and conventional mammalian ALDHs probably precludes this possibility.

3.4.4 Comparison of the Amino-Acid Sequences of *B. stearothermophilus* ADHs

When the amino-acid sequence of ADH 2334 was aligned with the ADH 1503 amino-acid sequence (Figure 3.4.2), the two enzymes were found to be identical at 244 positions (72% identity). The number of nonconserved amino acids (ninety-five) was therefore greater than the 38 +/- 15 predicted by the compositional analysis (Sheehan *et al.*, 1988). Consequently, the assignment of particular residues to a role in the differences in substrate specificity and kinetic mechanism between the two ADHs is a more difficult task than was originally anticipated.

The more thermophilic *B. stearothermophilus* ADH 12403 was mentioned in Section 3.4.3 and the ADH 12403 amino-acid sequence is included in an alignment of *B. stearothermophilus* ADH sequences in Figure 3.4.8. ADH 12403 has 91% identity with ADH 1503 (Cannio *et al.*, 1994). The identity between ADH 2334 and ADH 12403 is 72%.

Over the first 200 residues, the three ADHs are identical at 79% of the positions (158 residues). In the C-terminal region (residues 201-339), the three ADHs are identical at 57% of the positions (seventy-nine residues). A region of particularly low identity occurs at residues 211 - 227 inclusive, where only two residues, Asn-214 and Asp-219, are conserved between the three ADHs. Of the twenty-eight amino acid that differ between ADH 1503 and ADH 12403, six are clustered between nucleotides 211 and 227.

3.4.5 Alignment of *B. stearothermophilus* ADHs with HLADH and YADH1

The three *B. stearothermophilus* ADH sequences were aligned with the HLADH (E subunit) and the YADH1 sequences, as in Figure 3.4.9. Amino acid numbering in the text is that of HLADH, unless stated otherwise.

The *B. stearothermophilus* ADHs aligned with HLADH in a manner similar to YADH1. The gap in the yeast sequence, equivalent to a surface loop structure in HLADH (residues 119 - 139; Section 1.4.1), is present in each *B. stearothermophilus* sequence. There are insertions in the yeast sequence and in the *B. stearothermophilus* sequences adjacent to the substrate contact residues Leu-57 (two insertions) and Phe-93 (single insertion), and a single-residue gap in the yeast and *B. stearothermophilus* sequences adjacent to the substrate contact residue Thr-178. Alignment of the *B. stearothermophilus* sequences with HLADH and with YADH1 was difficult in two areas within the coenzyme-binding domain of HLADH (indicated in green in Figure 3.4.9). The region adjacent to Pro-243 corresponds to the region

```
ADH 2334      1 MKAAVVNEFKKALEIKEVERPKLEEGEVLVKIEACGVCHTDLHAAHGDWP  50
ADH 1503      1 MKAAVVEQFKKPLQVKEVEKPKISYGEVLVRIKACGVCHTDLHAAHGDWP  50
ADH 12403     1 MKAAVVEQFKEPLKIKEVEKPTISYGEVLVRIKACGVCHTDLHAAHGDWP  50
                ******  **  *  **** *    ***** * ****************

ADH 2334     51 IKPKLPLIPGHEGVGIVVEVAKGVKSIKVGDRVGIPWLYSACGECEYCLT 100
ADH 1503     51 VKPKLPLIPGHEGVGVIEEVGPGVTHLKVGDRVGIPWLYSACGHCDYCLS 100
ADH 12403    51 VKPKLPLIPGHEGVGIVEEVGPGVTHLKVGDRVGIPWLYSACGHCDYCLS 100
                **************   **  **   *************** * ***

ADH 2334    101 GQETLCPHQLNGGYSVDGGYAEYCKAPADYVAKIPDNLDPVEVAPILCAG 150
ADH 1503    101 GQETLCERQQNAGYSVDGGYAEYCRAAADYVVKIPDNLSFEEAAPIFCAG 150
ADH 12403   101 GQETLCEHQKNAGYSVDGGYAEYCRAAADYVVKIPDNLSFEEAAPIFCAG 150
                ******  * * ************ * **** ******   * *** ***

ADH 2334    151 VTTYKALKVSGARPGEWVAIYGIGGLGHIALQYAKAMGLNVVAVDISDEK 200
ADH 1503    151 VTTYKALKVTGAKPGEWVAIYGIGGLGHVAVQYAKAMGLNVVAVDLGDEK 200
ADH 12403   151 VTTYKALKVTGAKPGEWVAIYGIGGLGHVAVQYAKAMGLNVVAVDIGDEK 200
                ********* ** *************** * **************  ***

ADH 2334    201 SKLAKDLGADIAINGLKEDPVKAIHDQVGGVHAAISVAVNKKAFEQAYQS 250
ADH 1503    201 LELAKQLGADLVVNPKHDDAAQWIKEKVGGVHATVVTAVSKAAFESAYKS 250
ADH 12403   201 LELAKELGADLVVNPLKEDAAKFMKEKVGGVHAAVVTAVSKPAFQSAYNS 250
                  *** ****   *    *        ******    ** * **  ** *

ADH 2334    251 VKRGGTLVVVGLPNADLPIPIFDTVLNGVSVKGSIVGTRKDMQEALDFAA 300
ADH 1503    251 IRRGGACVLVGLPPEEIPIPIFDTVLNGVKIIGSIVGTRKDLQEALQFAA 300
ADH 12403   251 IRRGGACVLVGLPPEEMPIPIFDTVLNGIKIIGSIVGTRKDLQEALQFAA 300
                  ***  * ****    ***********    ********* **** ***

ADH 2334    301 RGKVRPIVETAELEEINEVFERMEKGKINGRIVLKLKED             339
ADH 1503    301 EGKVKTIVEVQPLENINDVFDRMLKGQINGRVVLKVD..            337
ADH 12403   301 EGKVKTIIEVQPLEKINEVFDRMLKGQINGRVVLTLEDK            339
                ***  * *   ** ** ** ** ** **** **
```

Figure 3.4.8 **Alignment of the Amino-Acid Sequences of *B. stearothermophilus* ADHs**

The amino acid sequences of ADH 2334 and ADH 1503 are as in Figure 3.4.2. The amino acid sequence of ADH 12403 is that deduced from the gene sequence (Cannio *et al.*, 1994). Residues conserved between the three ADHs are indicated by an asterisk. Three residues that have been associated with the higher thermostability of ADH 12403 compared to ADH 1503 (Cannio *et al.*, 1994) are indicated in cyan.

Figure 3.4.9 **Alignment of the Amino-Acid Sequences of *B. stearothermophilus* ADHs with HLADH and YADH1**

B. stearothermophilus sequences are as in Figure 3.4.8. The amino-acid sequence of YADH1 was deduced from the gene sequence of Bennetzen & Hall (1982). The amino-acid sequence of the E subunit of HLADH is that determined by Jörnvall [1970(a)]. The sequences were aligned, manually, in a manner which maximised conservation of the twenty-two residues that were identifed by Jörnvall *et al.* (1987) as being highly conserved between ADHs. The twenty-two residues are indicated by an asterisk under the YADH1 sequence. Residues that form ligands to the catalytic-zinc ion in HLADH and the equivalent residues in the other sequences are indicated in red, ligands to the second-zinc ion (four Cys residues, second line) are in blue. Residues identified as being associated with the substrate-binding pocket in HLADH (Eklund *et al.*, 1974; Eklund *et al.*,1982) and the equivalent residues in the other sequences are indicated in cyan. The coenzyme-binding domain of HLADH (Phe-176 - Ile-318 including any residues in cyan, Eklund *et al.*, 1974) is indicated in green with the ADP-binding βαβ-fold of HLADH, (residues 194 - 223 inclusive, Wierenga *et al.*, 1985; 1986), underlined.

ADH 2334 1 -----M---KAAVVNEFKKALEIKEVERPKLEEGEVLVKIEACGVCHTDLHAAHGDWPIK 52
ADH 1503 1 -----M---KAAVVEQFKKPLQVKEVEKPKISYGEVLVRIKACGVCHTDLHAAHGDWPVK 52
ADH 12403 1 M---KAAVVEQFKEPLKIKEVEKPTISYGEVLVRIKACGVCHTDLHAAHGDWPVK 52
HLADH 1 STAGKVIKCKAAVLWEEKKPFSIEEVEVAPPKAHEVRIKMVATGICRSDDHVVSGTL-V- 58
YADH1 1 ---SIPETQKGVIFYESHGKLEHKDIPVPKPKANELLINVKYSGVCHTDLHAWHGDWPLP 57
ADH 2334 53 PKLPLIPGHEGVGIVVEVAKGVKSIKVGDRVGIPWLYSACGECEYCLTGQETLCPHQLNG 112
ADH 1503 53 PKLPLIPGHEGVGVIEEVGPGVTHLKVGDRVGIPWLYSACGHCDYCLSGQETLCERQQNA 112
ADH 12403 53 PKLPLIPGHEGVGIVEEVGPGVTHLKVGDRVGIPWLYSACGHCDYCLSGQETLCEHQKNA 112
HLADH 59 TPLPVIAGHEAAGIVESIGEGVTTVRPGDKVIPLFTP-QCGKCRVCKHPEGNFCLKNDLS 117
YADH1 58 VKLPLVGGHEGAGVVVGMGENVKGWKIGDYAGIKWLNGSCMACEYCELGNESNCPHADLS 117
ADH 2334 113 G----------YSVDGGYAEYCKAPADYVAKIPDNLDPVEVAPILCAGV 151
ADH 1503 113 G----------YSVDGGYAEYCRAAADYVVKIPDNLSFEEAAPIFCAGV 151
ADH 12403 113 G YSVDGGYAEYCRAAADYVVKIPDNLSFEEAAPIFCAGV 151
HLADH 118 MPRGTMQDGTSRFTCRGKPIHHFLGTSTFSQYTVVDEISVAKIDAASPLEKVCLIGCGFS 177
YADH1 118 G----------YTHDGSFQQYATADAVQAAHIPQGTDLAEVAPILCAGI 156
ADH 2334 152 TTY-KALKVSGARPGEWVAIYGI-GGLGHIALQYAKAMGL-NVVAVDISDEKSKLAKDLG 208
ADH 1503 152 TTY-KALKVTGAKPGEWVAIYGI-GGLGHVAVQYAKAMGL-NVVAVDLGDEKLELAKQLG 208
ADH 12403 152 TTY-KALKVTGAKPGEWVAIYGI-GGLGHVAVQYAKAMGL-NVVAVDIGDEKLELAKELG 208
HLADH 178 TGYGSAVKVAKVTQGSTCAVFGL-GGVGLSVIMGCKAAGAARIIGVDINKDKFAKAKEVG 236
YADH1 157 TVY-KALKSANLMAGHWVAISGAAGGLGSLAVQYAKAMGY-RVLGIDGGEGKEELFRSIG 214
ADH 2334 209 ADIAINGLKE-DPVKAIHDQV--GGVHAAISVAVNKKAFEQAYQSVKRG-GTLVVVGLPN 264
ADH 1503 209 ADLVVNPKHD-DAAQWIKEKV--GGVHATVVTAVSKAAFESAYKSIRRG-GACVLVGLPP 264
ADH 12403 209 ADLVVNPLKE-DAAKFMKEKV--GGVHAAVVTAVSKPAFQSAYNSIRRG-GACVLVGLPP 264
HLADH 237 ATECVNPQDYKKPIQEVLTEMSNGGVDFSFEVIGRLDTMVTALSCCQEAYGVSVIVGVPP 296
YADH1 215 GEVFIDFTKEKDIVGAVLK-ATDGGAHGVINVSVSEAAIEASTRYVR-ANGTTVLVGMPA 272
ADH 2334 265 ADLPIP-IFDTVLNGVS----VKGSIVGTRKDMQEALDFAARGKVRP--IVETAELEEIN 317
ADH 1503 265 EEIPIP-IFDTVLNGV--K--IIGSIVGTRKDLQEALQFAAEGKVKT--IVEVQPLENIN 317
ADH 12403 265 EEMPIP-IFDTVLNGI--K--IIGSIVGTRKDLQEALQFAAEGKVKT--IIEVQPLEKIN 317
HLADH 297 DSQNLSMNPMLLLSGRTWKGAIFGGFKSKDSVPKLVADFMAKKFALDPLITHVLPFEKIN 356
YADH1 273 GAKCCSDVFNQVVK----SISIVGSYVGNRADTREALDFFAR--GLVKSPIKVVGLSTLP 326
ADH 2334 318 EVFERMEKGKINGRIVLKLKED 339
ADH 1503 318 DVFDRMLKGQINGRVVLKV--D 337
ADH 12403 318 EVFDRMLKGQINGRVVLTL-EDK 339
HLADH 357 EGFDLLRSGE-SIRTILTF--- 374
YADH1 327 EIYEKMEKGQIVGRYVVDTSK 347

Figure 3.4.9

of particularly low identity between the *B. stearothermophilus* ADHs noted previously (Section 3.4.4; Figure 3.4.8) and, as there is also low identity between the *B. stearothermophilus* sequences and HLADH, and between the *B. stearothermophilus* sequences and YADH1, it was difficult to position the gaps in the *B. stearothermophilus* sequences. Adjacent to Gly-320, there is high identity between the *B. stearothermophilus* sequences, however, gaps were necessary to maximise the alignments with the HLADH sequence.

3.4.5.1 ADH 2334 is an NAD$^+$-Dependent, Medium-Chain, Zinc-Dependent ADH

Alignments of the *B. stearothermophilus* ADH amino-acid sequences with the amino-acid sequence of HLADH and of YADH1 demonstrated that all three *B. stearothermophilus* enzymes conform to the requirements of the classical, NAD$^+$-dependent, medium-chain zinc-dependent ADHs. The three ligands to the HLADH catalytic-zinc ion (Cys-46, His-67 and Cys-174; indicated in red in Figure 3.4.9) are conserved in the *B. stearothermophilus* ADHs. The four ligands to the HLADH second-zinc ion (Cys-97, Cys-100, Cys-103 and Cys-111, indicated in blue) are also conserved in the *B. stearothermophilus* ADHs, as are the second-sphere carboxylic acid ligands Asp-49 and Glu-68. The twenty-two residues identified by Jörnvall *et al.* (1987) as being strictly conserved between zinc-dependent ADHs (Section 1.4.1) are indicated by an asterisk under the YADH1 sequence in Figure 3.4.9. All but two of the twenty-two residues are strictly conserved in the *B. stearothermophilus* ADHs. Pro-31 does not align with Pro residues in the *B. stearothermophilus* sequences, and Gly-77 is not conserved in ADH 2334. The NAD$^+$-specific residue, Asp-223, is conserved in the *B. stearothermophilus* ADHs.

Gly-77 falls within a region, about the catalytic-zinc His-ligand (His-67), that has been identified as a highly conserved region in microbial ADHs (Reid & Fewson, 1994). The conserved sequence motif is Gly-His-Glu-Xaa-Xaa-Gly-Xaa-Xaa-Xaa-Xaa-Xaa-Gly-Xaa-Xaa-Val, with the third Gly being equivalent to Gly-77. In the *B. stearothermophilus* ADH sequences, the conserved motif corresponds to residues 60 - 74 (*B. stearothermophilus* numbering; Figure 3.4.9). The motif is conserved in ADH 1503 and in ADH 12403 but, due to the substitution of Ala for Gly-77, it is not conserved in ADH 2334. The Reid & Fewson survey seemed biased towards fungal ADHs and, therefore, all of the bacterial, zinc-dependent ADHs (both NAD$^+$- and NADP-dependent) in the GenBank and PIR databases were examined. No bacterial ADH was found to contain a substitution at Gly-77. Apparently, ADH 2334 is the only bacterial ADH known to date in which the sequence motif about the catalytic-zinc ion His-ligand is not strictly conserved.

HLADH	YADH1	ADH 2334	ADH 1503	ADH 12403
Ser-48	Thr-45	Thr-40	Thr-40	Thr-40
Leu-57	Trp-54	Trp-49	Trp-49	Trp-49
Phe-93	Trp-92	Trp-87	Trp-87	Trp-87
Phe-110	Asn-110	Leu-105	Leu-105	Leu-105
Leu-116	Leu-116	Asn-111	Asn-111	Asn-111
Ser-117	Ser-117	Gly-112	Ala-112	Ala-112
Phe-140	Tyr-119	Tyr-114	Tyr-114	Tyr-114
Leu-141	Thr-120	Ser-115	Ser-115	Ser-115
Thr-178	Thr-157	Thr-152	Thr-152	Thr-152
Val-294	Met-270	Leu-262	Leu-262	Leu-262
Pro-296	Ala-272	Asn-264	Pro-264	Pro-264
Met-306	Asn-282	Asp-273	Asp-273	Asp-273
Leu-309	Val-285	Leu-276	Leu-276	Leu-276
Ser-310	Lys-286	Asn-277	Asn-277	Asn-277
Ile-318	Ile-290	Val-281	Ile-281	Ile-281

Table 3.4.1 **Amino Acids Associated with the Substrate-Binding Pocket of HLADH and YADH1, and the Equivalent Residues in the *B. stearothermophilus* ADHs**

3.4.5.2 Amino-Acid Residues Associated with Alcohol Substrate Binding in *B. stearothermophilus* ADHs

Amino-acid residues associated with alcohol binding in HLADH and in YADH1 were discussed in Chapter 1 (Section 1.4.2.3). HLADH and YADH1 residues that are associated with the substrate-binding pocket, and the equivalent residues in the *B. stearothermophilus* ADHs, are indicated in cyan in Figure 3.4.9 and are reproduced in Table 3.4.1. Thr-178, which is important for the correct positioning of the nicotinamide moiety with respect to alcohol, is the only amino acid that is conserved between the five ADHs.

There are difficulties involved in aligning the *B. stearothermophilus* sequences with the HLADH and YADH1 sequences at the C-terminus of the coenzyme-binding domain (Figure 3.4.9). Unfortunately, four of the HLADH substrate-binding residues occur in this region and, thus, it must be noted that the *B. stearothermophilus* residues listed as being equivalent to Met-306, Leu-309, Ser-310 and Ile-318 in Table 3.4.1 are somewhat speculative.

The residues postulated to be associated with substrate binding in ADH 1503 and ADH 12403 are identical. ADH 2334 and ADH 1503 differ at three positions, Gly and Ala interchange at the positions equivalent to Ser-117, Asn and Pro interchange at the positions equivalent to Pro-296 and Val and Ile interchange at the positions equivalent to Ile-318. Ser-117 and Ile-318 are situated in the wall of the substrate-binding pocket in HLADH, Pro-296 is located on a loop structure (residues 294 - 297) that undergoes a conformational change upon coenzyme binding. The conformational change shields the substrate-binding pocket from solution (Eklund *et al.*, 1981). Asn at 264 (*B. stearothermophilus* numbering) in ADH 2334 introduces an extra polar group and, hence, extra hydrogen-bonding potential, into the proposed substrate-binding pocket of ADH 2334, relative to ADH 1503.

The primary structure of the substrate-binding pockets of the *B. stearothermophilus* ADHs would seem to be more like that of YADH1 than that of HLADH. ADH 2334 has five exact matches with YADH1, but only two (Thr-178 and Leu-309) with HLADH. ADH 1503 has six exact matches with YADH1, and four (Thr-178, Pro-296, Leu-309 and Ile-318) with HLADH. In YADH1, and in the three *B. stearothermophilus* ADHs, bulky residues are substituted for Ser-48, Leu-57 and Phe-93; similarly, polar residues are substituted for Phe-140 and Leu-141. The similarity between the substrate-binding pockets of YADH1 and the *B. stearothermophilus* ADHs is surprising, given that ADH 2334 and ADH 1503, in contrast to YADH1, have the same pattern of decreasing K_m with increasing alkyl chain length as HLADH. Furthermore, ADH 2334 and ADH 1503 have K_m values for ethanol, propan-1-ol and butan-1-ol that are at least an order of magnitude less than the yeast enzyme (Sheehan *et al.*, 1988; Section 1.6.3). The dissimilarity

```
ADH 2334     38 CHTDLHAAHGDWPIKPKLPLIPGHEG
ADH 1503     38 CHTDLHAAHGDWPVKPKLPLIPGHEG
ADH 12403    38 CHTDLHAAHGDWPVKPKLPLIPGHEG
HLADH        46 CRSDDHVVSGTL-V-TPLPVIAGHEA
YADH1        43 CHTDLHAWHGDWPLPVKLPLVGGHEG
```

Alignment according to Jörnvall *et al.* (1978)

```
ADH 2334     38 CHTDLHAAHGDWPIKPKLPLIPGHEG
ADH 1503     38 CHTDLHAAHGDWPVKPKLPLIPGHEG
ADH 12403    38 CHTDLHAAHGDWPVKPKLPLIPGHEG
HLADH        46 CRSDDHVV-SGTLV-TPLPVIAGHEA
YADH1        43 CHTDLHAWHGDWPLPTKLPLVGGHEG
```

Alignment according to Sun & Plapp (1992)

Figure 3.4.10(a) Alternative alignments of substrate-binding residue Leu-57

```
ADH 2334   61 HEGVGIVVEVAKGVKSIKVGDRVGIPWLYSAC
ADH 1503   61 HEGVGVIEEVGPGVTHLKVGDRVGIPWLYSAC
ADH 12403  61 HEGVGIVEEVGPGVTHLKVGDRVGIPWLYSAC
HLADH      67 HEAAGIVESIGEGVTTVRPGDKVIPLFTP-QC
YADH1      66 HEGAGVVVGMGENVKGWKIGDYAGIKWLNGSC
```

Alignment according to Jörnvall *et al.* (1978)

```
ADH 2334   61 HEGVGIVVEVAKGVKSIKVGDRVGIPWLYSAC
ADH 1503   61 HEGVGVIEEVGPGVTHLKVGDRVGIPWLYSAC
ADH 12403  61 HEGVGIVEEVGPGVTHLKVGDRVGIPWLYSAC
HLADH      67 HEAAGIVESIGEGVTTVRPGDKV-IPLFTPQC
YADH1      66 HEGAGVVVGMGENVKGWKIGDYAGIKWLNGSC
```

Alignment according to Sun & Plapp (1992)

Figure 3.4.10(b) Alternative alignments of substrate-binding residue Phe-93

Figure 3.4.10 Alternative Alignments of Substrate-Binding Residues Leu-57 and Phe-93

Sequence alignments according to Jörnvall *et al.* (1978) were taken from Figure 3.4.9. Alignments according to Sun & Plapp (1992) show the published alignment of HLADH and YADH1 with the *B. stearothermophilus* ADHs added on. Ligands to the catalytic-zinc ion are in red, ligands to the second-zinc ion are in blue and residues assigned a role in substrate binding are in cyan.

between HLADH and ADH 2334 is particularly surprising as there are obvious similarities between the substrate specificities and kinetic parameters of the two enzymes. Both HLADH and ADH 2334 catalyse the oxidation of methanol, have identical kinetic mechanisms, and have similar k_{cat} and K_m values for ethanol, propan-1-ol and butan-1-ol (Sheehan *et al.*, 1988).

The three-dimensional model of YADH1, based on the known structure of HLADH, places Trp-54 at the top of the substrate-binding pocket, and Trp-92 at the bottom of the substrate-binding pocket. The role of these residues, and of Thr-45, in limiting the substrate specificity of YADH1 is currently a matter of debate. It has been postulated that Thr-45 and Trp-92 constrict the active-site of YADH1, due to their bulky side chains and thus impede the oxidation of molecules larger than ethanol. Experimental evidence for this hypothesis has been presented by Creaser *et al.* (1990). Experimental evidence in support of a role for Trp-92, and against a role for Trp-54, in limiting the substrate specificity of YADH1, has been reported by Green *et al.* (1993). There is no discussion in the literature as to why YADH1 cannot oxidise methanol.

Insertions in the YADH1 sequence (and in the *B. stearothermophilus* sequences), adjacent to the HLADH substrate-binding residues Leu-57 and Phe-93, were noted in Section 3.4.5. Sun & Plapp (1992) have pointed out that the assignment of substrate-binding residues in YADH1 is dependent upon where one positions the gaps in the HLADH sequence. In Figure 3.4.10 alignments proposed by Sun & Plapp (1992) are compared with those proposed by Jörnvall *et al.* (1978). The Sun and Plapp alignments result in Pro (rather than Trp) and Leu (rather than Trp) being the YADH1 substrate-binding residues equivalent to HLADH Leu-57 and Phe-93, respectively. The Sun & Plapp alignment in Figure 3.4.10(a) has been used to rationalize failed attempts to demonstrate unequivocally a role for Trp-54 in limiting the substrate specificity of YADH1 (Green *et al.*, 1993). The Sun & Plapp alignment results in a somewhat better local alignment, compared to that of Jörnvall *et al.* (1978), in the region containing Phe-93 [Figure 3.4.10(b)], and clearly results in a better local alignment of the *B. stearothermophilus* ADHs with HLADH. However, residue-93 is still conserved between YADH1 and the *B. stearothermophilus* ADHs. Leu in the *B. stearothermophilus* and yeast ADHs alignes with Phe-93 rather than Trp. The Sun & Plapp alignment is less convincing adjacent to Leu-57 [Figure 3.4.10(a)] in that a Gly residue in HLADH is not aligned with Gly in the other four sequences. Pro in the *B. stearothermophilus* and yeast ADHs alignes with Leu-57 rather than Trp.

The structure and flexibility of HLADH near Leu-116 have been shown to be important for substrate binding. In the S isoenzyme, the deletion of Asp-115 and the substitution of Leu for Phe-110 allows Leu-116 to move and, consequently,

relieves the steric hindrance that prevents the binding of steroid substrates in the E isoenzyme (Park & Plapp, 1991; 1992). In the *B. stearothermophilus* ADHs, Asn-111 (*B. stearothermophilus* numbering) is equivalent to Leu-116, and there are differences between ADH 2334 and ADH 1503 adjacent to Asn-111. These differences, for example, Pro at 107 in ADH 2334 but Glu at 107 in ADH 1503, may be relevant to the observed differences in substrate specificity between the two ADHs.

In the *B. stearothermophilus* ADHs the presence of Leu at the position equivalent to HLADH Val-294 may be significant. YADH1 has Met at this position (Met-270) whereas YADH2 has Leu, this substitution being the only difference in the substrate-binding pocket of the two isoenzymes. YADH2 has lower K_m values for the primary alcohols compared to YADH1 (Ganzhorn *et al.*, 1987). When Met-270 of YADH1 was changed to Leu (Ganzhorn *et al.*, 1987), the substitution resulted in a lowering (three- to five-fold) of the K_m values for butan-1-ol, pentan-1-ol and hexan-1-ol, and a significant increase in the catalytic efficiency for these alcohols (seven- to ten-fold). It was suggested that this may be due to either a decrease in steric hindrance or an increase in hydrophobicity, which resulted from the substitution of a methyl group for a thioether group. However, the substitution of Leu for Met did not result in a lowering of the K_m value, or an increase in the catalytic efficiency, for ethanol, indicating that substitutions outside of the substrate-binding pocket are also involved.

The comparison of putative substrate-binding residues did not reveal substitutions between ADH 2334 and ADH 1503 which can be invoked to explain the difference in the catalytic activity of ADH 2334 and ADH 1503 in the presence of methanol. The dissimilarity between the substrate contact residues of HLADH and the equivalent residues in ADH 2334 suggests that the ability to accept methanol as a substrate may be linked to the identity of amino acids which lie outside the substrate binding pocket.

3.4.5.3 Amino-Acid Residues Associated with Coenzyme-Binding in *B. stearothermophilus* ADHs

The method of Wierenga (Wierenga *et al.* 1985; 1986) was used to identify the ADP-binding βαβ-fold (NAD^+ diphosphate-binding site) in each *B. stearothermophilus* ADH (Figure 3.4.11). The three *B. stearothermophilus* ADH sequences and the ZADH1 sequence (Section 1.3.4) were aligned with the sequences taken from Wierenga *et al.* (1985) in such a way that Gly-201 and Asp-223, residues that are conserved in all NAD^+-dependent dehydrogenases (Wierenga *et al.*, 1985) were conserved. Asp-195 (*B. stearothermophilus*) corresponds to Asp-223 in HLADH. The selected sequences, residues 167 - 195 (*B. stearothermophilus* numbering), conform to the eleven-residue fingerprint sequence for a diphosphate-binding βαβ-fold at ten out of the eleven positions, and align with the ADP-binding

```
                    EEEEEE HHHHHHHHHHHHHH    EEEEEE
HLADH          194    TCAVFGLGGVGLSVIMGCKAAGAA-RIIGVD 223
YADH1          170 GHWVAISGAAGGLGSLAVQYAKAMGY--RVLGID 201
BSLDH            8    RVVVIGAGFVGASYVFALMNQGIADEIVLID  38
BSGAPDH          3    KVGINGFGRIGRNVFRAALKNPDIEVVAVND  33
ADH 2334       164 PGEWVAIYGIGGLGHIALQYAKAM--GLNVVAVD 195
ADH 1503       164 PGEWVAIYGIGGLGHVAVQYAKAM--GLNVVAVD 195
ADH 12403      164 PGEWVAIYGIGGLGHVAVQYAKAM--GLNVVAVD 195
ZADH1          161 PGQWLAIYGLGGLGNLALQYAKNVF-NAKVIAID 193
```

Figure 3.4.11 **Identification of βαβ Diphosphate-binding Domains in *B. stearothermophilus* ADHs**

The eleven-residue fingerprint identified by Wierenga *et al.* (1985) as being diagnostic of a diphosphate-binding site is indicated as follows: the N-terminal hydrophilic residue in magenta; the six hydrophobic residues in blue; three conserved glycine residues, in the motif Gly-Xaa-Gly-Xaa-Xaa-Gly, in green; the negativly charged C-terminal residue in red. The alignment of HLADH, YADH1, lactate dehydrogenase of *B. stearothermophilus* (BSLDH) and glyceraldehyde-3-phosphate dehydrogenase of *B. stearothermophilus* (BSGAPDH), and the assignment of residues in these sequences, were taken from Wierenga *et al.* (1985). The secondary structure of the diphosphate-binding domain of HLADH (Eklund, 1976) is indicated above the HLADH sequence by E (β-sheet) and H (α-helix).

site of HLADH (Figure 3.4.9; green and underlined). HLADH and glyceraldehyde-3-phosphate dehydrogenase of *B. stearothermophilus*, for which the diphosphate-binding folds have been confirmed by X-ray crystallography (Eklund *et al.*, 1981; Biesecker *et al.*, 1977) also give a score of ten against the fingerprint sequence. The deviant position is Trp-167, which lies at the N-terminus of the first α-helix, and is a position that is usually occupied by either a basic or a hydrophilic residue. ADH 1503 and ADH 12403 are identical over the entire sequence of the βαβ unit. In ADH 2334, Ile replaces Val at 179 and Leu replaces Val at 181.

A summary of the available information concerning residues that participate in coenzyme-binding in HLADH and YADH1, and the equivalent residues in the *B. stearothermophilus* ADHs, deduced from Figure 3.4.9, is presented in Table 3.4.2. Residues involved in adenine and adenosine ribose binding in HLADH and YADH1 were taken from Jörnvall *et al.* (1978). Diphosphate interactions with HLADH were taken from Wierenga *et al.* (1985), nicotinamide ribose and nicotinamide ring interactions with HLADH were taken from Eklund *et al.* (1981) and the equivalent residues in YADH1 were deduced from Figure 3.4.9.

Lys-228 and Arg-369 (Section 1.4.2.2) are conserved in the *B. stearothermophilus* sequences. Arg-47 is replaced by His in the *B. stearothermophilus* ADHs, and that replacement is usual in microbial ADHs (Reid & Fewson, 1994). With respect to the residues associated with the NAD^+-dependent proton-release step in HLADH (Section 1.5.2), the three *B. stearothermophilus* sequences conform to the pattern that has been identified for microbial ADHs (Reid & Fewson, 1994), namely, His-51 is conserved and Ser-48 is replaced by Thr (Figure 3.4.9).

When ADH 2334 and ADH 1503 are compared, the residues deduced to be nicotinamide-ring binding, nicotinamide-ribose binding and diphosphate binding are identical. However, residues that form the adenine-binding pocket are different at six of the nine positions. Two of the differences (equivalent to HLADH Pro-243 and Ile-250) fall within the region of lowest identity between the *B. stearothermophilus* ADHs (Figure 3.4.8; residues 211 - 227). There is one difference in the adenosine-ribose binding residues when ADH 2334 and ADH 1503 are compared. In Table 3.4.2 there are only two differences between ADH 1503 and ADH 12403 and these are in the adenine-binding pocket. X-ray crystallography studies of coenzyme binding to HLADH have shown that the adenine ring is situated in a hydrophobic pocket with the adenine N6 atom pointing out of the pocket towards the solution and close to the guanidinium group of Arg-271, which, in turn, is in a salt-linkage to Asp-273 (Eklund *et al.*, 1976). The residues in the *B. stearothermophilus* ADHs that are equivalent to the HLADH adenine-binding pocket residues are of a hydrophobic nature, but, as in YADH1, there is no clear equivalent of the Arg-271-Asp-273 salt-

	HLADH	YADH1	ADH 2334	ADH 1503	ADH 12403
Residues in the interior of the adenine-binding pocket	Phe-198 Val-222 Ile-224 Pro-243 Ile-250 Thr-274 Thr-277	Ser-176 Ile-200 Gly-202 Phe-221 Val-228 Ala-251 Ala-254	Tyr-171 Val-194 Ile-196 Gly-215 Val-221 Ala-243 Gln-246	Tyr-171 Val-194 Leu-196 Pro-215 Ala-221 Ala-243 Ser-246	Tyr-171 Val-194 Ile-196 Pro-215 Ala-221 Ala-243 Ser-246
Residues at the surface of the adenine-binding pocket	Arg-271 Asp-273	Ser-248 Ala-250	Asn-240 Lys-242	Ser-240 Ala-242	Ser-240 Pro-242
Residues associated with adenosine-ribose binding	Gly-199 Asp-223 Asn-225 Lys-228 Ile-269	Gly-177 Asp-201 Gly-203 Lys-206 Ser-246	Gly-172 Asp-195 Ser-197 Lys-200 Ala-238	Gly-172 Asp-195 Gly-197 Lys-200 Ala-238	Gly-172 Asp-195 Gly-197 Lys-200 Ala-238
Residues associated with diphosphate binding	Arg-47 Gly-202 Val-203 Arg-369	His-44 Gly-181 Leu-182 Arg-340	His-39 Gly-175 Leu-176 Arg-331	His-39 Gly-175 Leu-176 Arg-331	His-39 Gly-175 Leu-176 Arg-331
Residues associated with nicotinamide-ribose binding	His-51 Val-203 Ile-269 Gly-270 Gly-293 Val-294	His-48 Leu-182 Ser-246 Val-247 Gly-269 Met-270	His-43 Leu-176 Ala-238 Val-239 Gly-261 Leu-262	His-43 Leu-176 Ala-238 Val-239 Gly-261 Leu-262	His-43 Leu-176 Ala-238 Val-239 Gly-261 Leu-262
Residues in contact with the nicotinamide ring	Thr-178 Val-203 Val-292 Val-294	Thr-157 Leu-182 Val-268 Met-270	Thr-152 Leu-176 Val-260 Leu-262	Thr-152 Leu-176 Val-260 Leu-262	Thr-152 Leu-176 Val-260 Leu-262

Table 3.4.2 **Amino Acids Associated with Coenzyme Binding in HLADH and YADH1, and the Equivalent Residues in *B. stearothermophilus* ADHs**

bridge. There are, however, a Lys at 241 and Glu at 245 (*B. stearothermophilus* numbering), in both ADH 2334 and ADH 1503, that align close to Arg-271 and Asp-273 (Figure 3.4.9). In ADH 12403, Lys-241 is present but Glu-245 is replaced by Gln.

3.4.6 Identities Between *B. stearothermophilus* ADHs and other Medium-Chain, Zinc-Dependent ADHs

Pairwise identities between the *B. stearothermophilus* ADHs and other medium-chain, zinc-dependent ADHs are presented in Table 3.4.3. Abbreviations used in Table 3.4.3 are explained in Table 3.4.4. The *B. stearothermophilus* ADHs are the only NAD^+-dependent, medium-chain, zinc-dependent ADHs that, to date, have been identified from thermostable eubacteria. The other thermostable NAD^+-dependent ADH included in Table 3.4.3 (SSADH) is from the archaebacterium *S. solfataricus*.

The identities between the *B. stearothermophilus* ADHs and HLADH are approximately 30%, which is consistent with the identities between other microbial ADHs and HLADH (Table 3.4.3). Identities between NAD^+-dependent bacterial ADHs are, in general, low (Reid & Fewson, 1994) and, presumably, are related to phylogenetic differences, and to the diverse array of substrates and environmental conditions experienced by these enzymes *in vivo* (Table 3.4.4). The highest identity between ADH 2334 and an ADH of another species was 50% identity with NAD^+-dependent, medium-chain, zinc-dependent ADH from the mesophilic bacterium *Z. mobilis* (ZADH1). With the exception of their identity with ZADH1, the *B. stearothermophilus* ADHs have a somewhat higher identity with the fungal YADH1 (approximately 40%) than they do with other bacterial ADHs. The low identities between the *B. stearothermophilus* ADHs, which are ethanol dehydrogenases, and threonine dehydrogenase (ECTDH), sorbitol dehydrogenase (BSSDH) and the *Pseudomonas* ADHs (PSADH, PPBDH), which are not, is to be expected. The identities between the *B. stearothermophilus* ADHs and *A. eutrophus* ADH (AEADH) are consistent with the low identities between AEADH and the other ADHs in Table 3.4.3. AEADH has novel properties and lacks some of the highly conserved features of zinc-dependent ADHs (Steinbüchel & Schlegel, 1984; Jendrossek *et al.*, 1988).

3.4.7 Primary Structure of ADH 2334 in the Context of Thermostability

3.4.7.1 Amino-Acid Composition

B. stearothermophilus DSM 2334 is a thermophilic organism, albeit a moderate thermophile, therefore, the primary structure of ADH 2334 was examined for features that may convey thermostability to the enzyme. Analyses of the amino-acid composition of other thermostable enzymes, and of the frequency of various

	ADH 2334	ADH 1503	ADH 12403	ZADH1	AEADH	ECTDH	SSADH	TBADH	BSSDH	PSADH	PPBDH	HLADH	YADH1
ADH 2334	100	72	73	50	28	30	37	27	33	30	30	31	44
ADH 1503		100	92	54	32	29	35	27	37	26	32	32	41
ADH 12403			100	54	30	31	34	27	36	30	32	31	42
ZADH1				100	25	27	34	25	30	29	26	31	36
AEADH					100	30	26	37	32	27	25	25	26
ECTDH						100	24	27	32	27	27	27	30
SSADH							100	26	32	27	27	30	30
TBADH								100	31	28	27	29	21
BSSDH									100	26	34	31	25
PSADH										100	40	30	29
PPBDH											100	31	28
HLADH												100	26
YADH1													100

Table 3.4.3 **Pairwise Amino-acid Sequence Identities Among Zinc-Dependent ADHs**

Identities (%) were calculated using the GCG programme Bestfit, with a gap weight of 3 and a gap length of 0.1. Abbreviations are explained, and relevant references given, in Table 3.4.4.

Enzyme	Organism	Enzyme abbreviation used in the text	Coenzyme	Comments	Reference to amino-acid sequence
Ethanol dehydrogenase	*Zymomonas mobilis*	ZADH1	NAD^+	Obligate fermentative bacterium	Keshav *et al.* (1990)
L-Threonine dehydrogenase	*Escherichia coli*	ECTDH	NAD^+	Ligands to two zinc ions present, but zinc content uncertain	Aronson *et al.* (1989)
Sorbitol dehydrogenase	*Bacillus subtilis*	BSSDH	NAD^+		Ng *et al.* (1992)
Fermentative alcohol dehydrogenase	*Alcaligenes eutrophus* strain N9A	AEADH	NAD^+	No obvious second-zinc binding site	Jendrossek *et al.* (1988)
Alcohol dehydrogenase	*Thermoanaerobium brockii*	TBADH	$NADP^+$	Extreme thermophile Tolerates organic solvents	Peretz & Burstein (1989)
Secondary-alcohol dehydrogenase	*Thermoanaerobacter ethanolicus*	TEADH	$NADP^+$	Extreme thermophile One catalytic-zinc Cys-ligand replaced by Asp	Burdette *et al.* (1996)
Alcohol dehydrogenase	*Sulfolobus solfataricus* strain DSM1617	SSADH	NAD^+	Thermoacidophilic archaebacterium Tolerates organic solvents	Ammendola *et al.* (1992)
Alcohol dehydrogenase	*Pseudomonas* (species not identified)	PSADH	Not determined Contains a Glu equivalent to Asp-223	First enzyme in an operon involved in α-terpineol metabolism	Peterson *et al.* (1992)
Benzyl alcohol dehydrogenase	*Pseudomonas putida*	PPBDH	NAD^+	Gene is on a plasmid	Shaw *et al.* (1993)

Table 3.4.4 **Zinc-Dependent ADHs Discussed in the Text**

substitutions (mesophile to thermophile) suggest that some of the following features might be expected in the *B. stearothermophilus* ADHs: a high Arg/(Arg+Lys) ratio, resulting from a high frequency of Lys to Arg substitutions, mesophile to thermophile (Mrabet *et al.*, 1992); a low cysteine content (Ammendola *et al.*, 1992; Maras *et al.*, 1992); a high proline content (Matthews *et al.*, 1987; Watanabe *et al.*, 1994); Lys to Ala or Glu, Gly to Ala, Ser to Ala or Thr, Thr to Ala, Ile to Val, Glu to Arg and Asp to Arg substitutions, mesophile to thermophile (Menéndez-Arias & Argos, 1989). The Arg/(Arg+Lys) ratios (R/R+K) and the amino-acid content (% composition) of the *B. stearothermophilus* ADHs and other zinc-dependent ADHs are presented in Table 3.4.5. The frequency of some amino-acid substitutions between zinc-dependent mesophilic ADHs and ADH 2334 are presented in Table 3.4.6.

A high R/R+K is not a feature of the *B. stearothermophilus* ADHs. The values of R/R+K for the three *B. stearothermophilus* enzymes are lower than those of the three mesophilic ADHs in Table 3.4.5.

The low cysteine content of the three ADHs from extreme thermophiles (SSADH, TBADH, TEADH) relative to that of the other ADHs in Table 3.4.5 is striking. The *B. stearothermophilus* ADHs do have a low cysteine content compared to HLADH, but not compared to YADH1 or ZADH1. The cysteine content of an additional five mesophilic bacterial ADHs (ECTDH, BSSDH, AEADH, PSADH and PPBDH; Table 3.4.4) was calculated (data not presented) and the majority of these enzymes have a cysteine content that is equivalent to, or lower than, that of the *B. stearothermophilus* enzymes. A low cysteine content may reflect the microbial origin, rather than the thermostability, of the *B. stearothermophilus* enzymes.

ADH 2334 has a total of seventeen proline residues (5% composition), ADH from the more thermostable *B. stearothermophilus* 12403 has eighteen (5.3% composition) and the thermolabile ZADH1 has only nine proline residues. However, SSADH, an extremely thermophilic ADH, has fifteen (4.3% composition), whereas HLADH has nineteen (5.3% composition). It is not possible to make a meaningful correlation between total proline content and the thermostability of the ADHs shown in Table 3.4.5.

According to the statistical analysis of Menéndez-Arias & Argos (1989), a high alanine content is significant, in the context of thermostability, if it occurs at the expense of lysine, glycine, serine or threonine. The alanine content of the *B. stearothermophilus* ADHs is higher than that of HLADH, however, there is not a corresponding drop in the lysine or glycine content of the *B. stearothermophilus* ADHs compared to HLADH. The three *B. stearothermophilus* enzymes do have a lower serine content than the mesophilic enzymes shown in Table 3.4.5, and a lower threonine content. Inspection of the specific nature of the substitutions for serine and threonine in ADH 2334 revealed that the low serine and threonine content, relative to

	ADH 2334 T	ADH 1503 T	ADH 12403 T	ZADH1 M	HLADH M	YADH1 M	SSADH T	TEADH T	TBADH T
R/R+K	0.222	0.238	0.213	0.266	0.285	0.250	0.426	0.316	0.364
Cys	2.3	2.6	2.6	2.4	3.7	2.3	1.4	1.1	1.1
Pro	5.0	5.0	5.3	2.6	5.3	3.7	4.3	6.2	5.9
Ala	10.3	10.0	10.0	10.7	7.4	10	9.8	9.9	9.9
Lys	9.1	8.3	8.5	8.0	8.0	6.9	6.6	6.8	6.8
Gly	10.6	10.7	10.6	10.4	10.1	12.6	10.0	12.2	12.2
Ser	2.9	2.6	2.6	3.2	6.9	6.0	4.0	2.5	2.5
Thr	2.6	3.2	3.5	5.3	6.4	4.0	3.7	3.7	3.7
Val	12.0	13	11.8	11.5	10.4	10.3	10.3	10.5	10.5
Glu	8.2	6.8	7.9	7.7	5.6	5.7	7.2	5.9	5.9
Ile	7.0	5.9	6.5	5.6	6.4	6.0	6.9	7.4	7.4
Phe	1.5	2.0	2.3	3.2	4.8	2.3	1.7	3.9	3.9
Leu	7.9	7.7	7.9	6.5	6.6	6.9	9.2	6.5	6.5
Gln	2.0	3.5	2.6	2.6	2.1	2.6	2.8	1.1	0.8
Asp	5.3	5.0	4.4	5.6	4.5	4.6	4.3	5.6	5.6
His	2.3	2.6	2.6	2.6	1.8	3.1	2.0	2.8	2.8
Met	1.1	0.9	1.5	2.0	2.4	1.7	2.6	4.2	4.2
Arg	2.6	2.6	2.3	2.9	3.2	2.3	4.9	3.1	3.9
Trp	0.8	1.2	0.9	0.59	0.5	1.4	0.6	1.4	1.1
Asn	2.9	2.4	2.3	3.5	2.1	3.1	3.4	2.8	2.8
Tyr	2.9	3.2	3.2	2.3	1.0	3.7	3.7	1.7	1.7

Table 3.4.5 **Amino-Acid Compositions (%) of Mesophilic and Thermophilic ADHs**

The amino-acid compositions (number of residues) of mesophilic (M) and thermophilic (T), zinc-dependent, ADHs were obtained from the GenBank database using the GCG programme Composition and were then converted to % composition. Abbreviations and references are as in Table 3.4.4. Arg/(Arg+Lys) ratios (R/R+K) were calculated from the number of arginine and lysine residues in each protein.

	Ser to Xaa substitutions				Thr to Xaa substitutions			Xaa to Val substitutions				Xaa to Glu substitutions		
	Total Number	Different types	Ser to Ala	Ser to Thr	Total number	Different types	Thr to Ala	Total number	Different types	Ile to Val	Ala to Val	Total number	Different types	Lys to Glu
ZADH1 to ADH 2334	8	7	1	1	13	8	2	21	6	7	8	15	8	4
HLADH to ADH 2334	23	14	2	1	19	13	1	24	11	8	2	20	12	4
YADH1 to ADH 2334	16	13	3	1	9	6	0	22	9	8	4	15	9	4

Table 3.4.6 **Frequency of Selected Amino-Acid Substitutions; Mesophilic Enzyme to ADH 2334**

Substitutions (HLADH to ADH 2334 and YADH1 to ADH 2334) were determined from the alignment presented in Figure 3.4.9. Substitutions ZADH1 to ADH 2334 were determined from an alignment (not shown) obtained using the GCG programme Bestfit.

that in ZADH1, HLADH and YADH1, were due to a large number of different substitutions, rather than to specifically Ser to Ala or Thr, and Thr to Ala (Table 3.4.6).

The substitutions Ile to Val and Lys to Glu were noted as being among the most frequent substitutions between porcine heart cytoplasmic malate dehydrogenase and *Thermus flavus* malate dehydrogenase (Kelly *et al.*, 1993). The substitution Ile to Val was noted, by Korndörfer *et al.* (1995) as among the most frequent substitutions between *Homarus americanus* (lobster) and *B. stearothermophilus* glyceraldehyde-3-phosphate dehydrogenase. The substitution Ile to Val is the most common substitution of the type Xaa to Val, mesophile to thermophile, when one compares HLADH and ADH2334, and YADH1 and ADH 2334 (Table 3.4.6). Comparing ZADH1 and ADH 2334, the most common substitutions of the type Xaa to Val, mesophile to thermophile, are Ile to Val and Ala to Val, which together account for fifteen out of twenty-one substitutions. The substitution of the type Lys to Glu, mesophile to thermophile, is the most common substitution of the type Xaa to Glu when HLADH, YADH1 or ZADH1 is compared with ADH2334.

3.4.7.2 Specific Substitutions Involving Proline, Glutamic Acid and Lysine

Substitutions involving proline (rather than total proline content) have been invoked to explain the relative thermostabilities of ADH 12403 (half-life of 2 h at 70 °C) and ADH 1503 (half-life of 30 min at 70 °C), and to explain the relative thermostabilities of SSADH (half-life of 3 h at 85 °C) and ADH 12403 (Cannio *et al.*, 1994). In ADH 12403, compared to ADH 1503, Pro is substituted for Ala-242 (Figure 3.4.8). When the amino-acid sequence of SSADH is aligned with that of ADH 12403, there are six substitutions and three insertions involving Pro. ADH 12403, containing a point mutation in which Pro-242 was replaced by Ala, retained catalytic activity but showed reduced thermal stability (Fiorentino *et al.*, 1998).

There are nine Proline residues in the amino-acid sequence of ADH 2334 that are not conserved in HLADH or in YADH1 (Figure 3.4.9). Three of these residues, Pro-86, Pro-140 and Pro-306 (*B. stearothermophilus* numbering), were made to align with a proline in HLADH by simple adjustments to the alignment and, therefore, were not considered further. In a multiple sequence alignment of zinc-dependent bacterial ADHs, Pro-127 of ADH 2334 aligned with proline in AEADH, ECTDH and SSADH, and Pro-164 of the three *B. stearothermophilus* sequences aligned with proline in ZADH1 and BSSDH. The amino-acid sequences adjacent to the remaining proline residues were adjusted manually in an attempt to align the *B. stearothermophilus* proline residues with those in other sequences (Figure 3.4.12). Two proline residues, Pro-59 and Pro-268, appear to be unique to the *B.*

```
                                        *                                                                 *
ADH 2334    37 VCHTDLHAAHGDWPIKP.......KLPLIPGHEGVGIVVEVAKGVKSIKVG  80      253 RGGTLVVVGLPNADLPIP...IFDTVLNGVSVKGSIV.GT 288
ADH 1503    37 VCHTDLHAAHGDWPVKP.......KLPLIPGHEGVGVIEEVGPGVTHLKVG  80      253 RGGACVLVGLPPEEIPIP...IFDTVLNGVKIIGSIV.GT 288
ADH 12403   37 VCHTDLHAAHGDWPVKP.......KLPLIPGHEGVGIVEEVGPGVTHLKVG  80      253 RGGACVLVGLPPEEMPIP...IFDTVLNGIKIIGSIV.GT 288
HLADH       45 ICRSDDHVVSGTL.VTPL.........PVIAGHEAAGIVESIGEGVTTVRPG  86      285 AYGVSVIVGVPPDSQNLSMN..PMLLLSGRTWKGAIFGGF 322
YADH1       42 VCHTDLHAWHGDWPL.PV......KLPLVGGHEGAGVVVGMGENVKGWKIG  85      261 ANGTTVLVGMPAGAKCCS..DVFNQVVKSISIVGSYV.GN 297
ZADH1       36 VCHTDLHVKNGDFGDETGR.........ITGHEGIGIVKQVGEGVTSLKAG  77      251 AGGRVVAVGLPPEKMDLS...IPRLVLDGIEVLGSLV.GT 286
SSADH       37 VCHSDVHMRQGRFGNLRIVEDLGVKLPVTLGHEIAGKIEEVGDEVVGYSKG  87      263 KQGKYVMVGLFGADLHYH..APLIT.LSEIQFVGSLV.GN 298
TBADH       36 PCTSDIHTVFEGAIGERHNM........ILGHEAVGEVVEVGSEVKDFKPG  78      258 PGGTIANVNYFGEGEVLPVPRLEWGCGMAHKTIKGGLCPG 297
AEADH       40 ICGTDVHILKGEYPVA.........KGLTVGHEPVGIIEKLGSAVTGYREG  81      175 PGGTLSSLGVYSSDLTIPLSAFAAGLGD.HK.INTALCPG 212
PPBDH       39 LCHTDLVCRDQHYPV.PL.........PMVFGHEGAGVVERVGSAVKKVQPG  80      277 IGGEIGIVGAPPMGATVPVDIN..FLLFNRKLRGIVEGQS 314
PSADH       17 VCHTDLAIKDGV.PF.PL.........PAVLGHEGSGIVEAVGPGVKHLKPG  57      258 MRGQLAVVGVPPKLDATAAVSPLALIQKGLKLMGVIEGDS 297
BSSDH       44 ICGSDLHYYTNGRIG.....NYVVEKPFILGHECAGEIAAVGSSVDQFKVG  89      264 RGGKLAIVGL.PSQNEIPLNVPFIADNEI.DIYGIFRYA. 300
ECTDH       37 ICGTDVHIYNWD.....EWSQKTIPVPMVVGHEYVGEVVGIGQEVKGFKIG  82      255 HGGRIAMLGIPPSDMSIDWTKVIFKGLFIKGIYGREMF.. 292
```

Figure 3.4.12 **Identification of Proline Residues Unique to *B. stearothermophilus* ADHs**

Stretches of amino-acid sequence in which proline residues were not conserved between the *B. stearothermophilus* ADHs, HLADH and YADH1 were selected from Figure 3.4.9. Homologous stretches of sequence in bacterial zinc-dependent ADHs were identified using the GCG alignment programme Pileup. The alignments were adjusted, manually, to give the arrangement shown above. Ligands to the catalytic-zinc ion are indicated in red, amino acids which were identified as putative substrate-binding residues in *B. stearothermophilus* ADHs (Table 3.4.1) are indicated in magenta. Abbreviations and references are as in Table 3.4.4.

stearothermophilus ADHs. Pro-59 is adjacent to the catalytic-zinc His-ligand, and in *T. ethanolicus* secondary ADH, four putative thermostabilising prolines have been identified adjacent to catalytic-zinc ligands (Burdette *et al.*, 1996). Pro-268 is adjacent to a cluster of four amino acids that were designated as substrate binding residues in Table 3.4.1.

Cannio *et al.* (1994) have suggested that a salt-bridge between Glu-11 and Lys-14 in ADH 12403 contributes to the extra stability of ADH 12403 relative to ADH 1503, which has Glu at position 11 and Gln at position 14. ADH 12403, containing a point mutation in which Glu-11 was replaced by Lys, retained catalytic activity but showed reduced thermal stability (Fiorentino *et al.*, 1998). In ADH 2334 (from an organism whose thermostability appears to be identical to that of *B. stearothermophilus* 1503), a salt-bridge is possible between amino acid-11 (Lys) and amino acid-14 (Glu), but with a dipole moment orientated in the opposite direction to that in ADH 12403. The orientation of the Lys-Glu pair in ADH 2334, with the Lys nearest to the N-terminus, would be expected to be destabilising (Marqusee & Baldwin, 1987).

In conclusion, no single feature of the protein primary structure of ADH 2334 can be unequivocally linked to thermostability. Relative to homologous mesophilic ADHs, ADH 2334 does not have a low cysteine content, high proline content or a high Arg/(Arg+Lys) ratio. However, when selected mesophilic ADHs were compared to ADH 2334, two amino-acid substitutions that have been linked to thermostability, namely Ile to Val and Lys to Glu, were identified as being the most frequent substitutions of the type Xaa to Val and Xaa to Glu.

CHAPTER 4
THE CONSTRUCTION OF AN EXPRESSION SYSTEM FOR *adh*2334

4.1 INTRODUCTION

4.1.1 Expression of Recombinant Proteins in *E. coli*

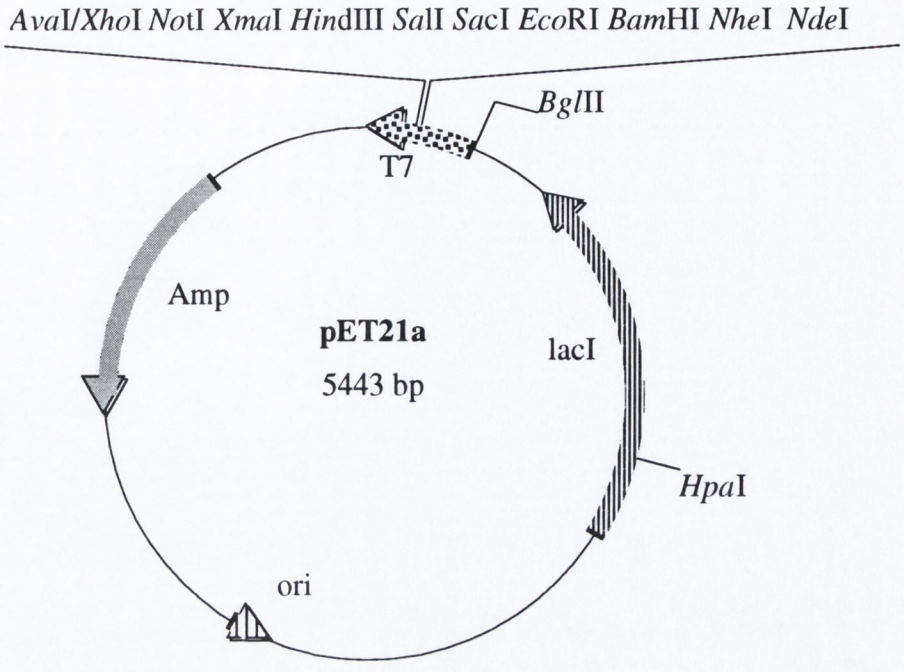
Methods for expressing large amounts of protein from a cloned gene, introduced into *E. coli*, have been invaluable in the purification and functional analysis of proteins. An overwhelming variety of methods, and plasmid vectors, are now available for the production of both fusion proteins and intact native proteins. The objective was the production of intact, functional recombinant ADH 2334 (rADH) and, therefore, the many popular methods for the production of fusion proteins (reviewed by Nygren *et al.*, 1994) are not considered here.

4.1.1.1 The pET Expression System

The pET expression system (plasmid for expression by T7 RNA polymerase) was developed by Tabor & Richardson (1985) and Rosenberg *et al.* (1987), and is reviewed by Studier *et al.* (1990). The system exploits the high efficiency and extreme selectivity of bacteriophage T7 RNA polymerase. T7 RNA polymerase is selective for specific promoters that are rarely encountered in DNA unrelated to T7 DNA. Furthermore, efficient termination signals for T7 RNA polymerase are also rare. Consequently, T7 RNA polymerase can make a complete transcript of almost any DNA that is placed under the control of a T7 promoter.

A plasmid map of a typical pET vector (pET21a), and the nucleotide sequence across the transcription and translation signals are shown in Figure 4.1.1. Bacteriophage T7 ϕ*10* promoter (T7 promoter), the T7 gene *10* ribosome-binding site (T7 RBS), an ATG translational start codon (within the *Nde*I site shown in Figure 4.1.1), the T7 Tϕ transcription termination signal (T7 terminator) and a multiple cloning site were inserted into plasmid pBR322. A coding sequence cloned into a pET vector, in such a manner that the second codon is fused to the plasmid ATG, will be transcribed from the T7 promoter and translated from the T7 ribosome-binding site. The pET21 series (developed by Novagen) carry the additional feature of a nucleotide sequence downstream from the multiple cloning site that encodes at least six consecutive His residues. A gene inserted in the correct reading frame can thus be expressed with a C-terminal fusion peptide (a His-Tag) that serves as an affinity tag for one-step purification by nickel-chelation chromatography.

The gene of interest is cloned into a pET vector in the absence of a source of T7 RNA polymerase and using a high transformation efficiency cloning host such as *E. coli* TG1. *E. coli* RNA polymerase mediated transcription from a T7 promoter is minimal, therefore, expression of the cloned gene does not occur at this stage. The recombinant plasmid is then transferred to a host containing a chromosomal copy of the gene for T7 RNA polymerase. *E. coli* BL21(DE3) (*E. coli* BL21), the host cell



*Bgl*II **T7 promoter** ***lac* operator**

AGATCTCGATCCCGCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAAC

T7 RBS *Nde*I

AATTCCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGGCTAG

*Sac*I

CATGACTGGTGACAGCAAATGGGTCGCGGATCCGAATTCGAGCTCCGTCGACAAGCTT

AvaI/XhoI His-Tag

GCGGCCGCACTCGAG**CACCACCACCACCACCAC**TGAGATCCGGCTGCTAACAAAGCCC

T7 terminator

GAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGG

GGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCC

Figure 4.1.1 Structure of Plasmid pET21a

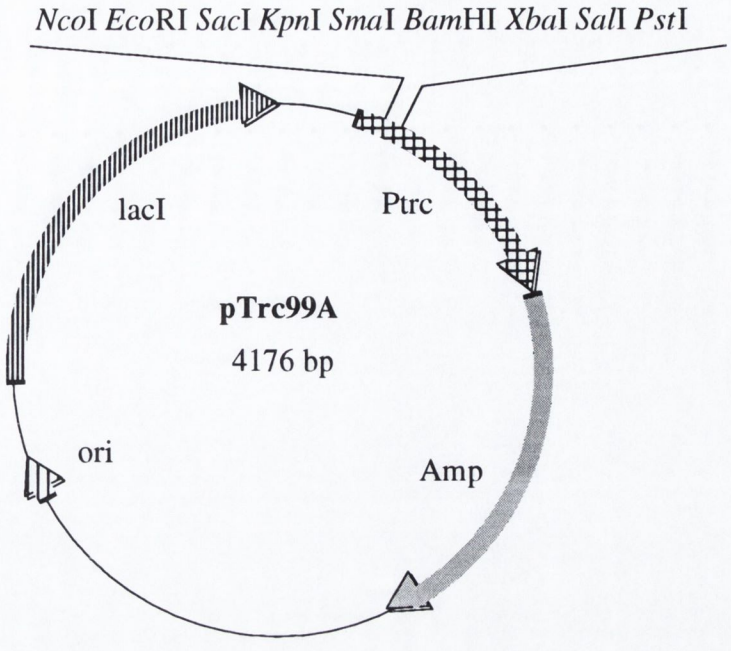
In the plasmid map, the notation T7 indicates the transcription and expression region of the vector. Amp indicates the ampicillin resistance gene, ori indicates the origin of plasmid replication, lacI indicates the *lacI* gene. Arrows indicate the direction of transcription. In the closely related vector pET21d, a *Nco*I restriction site is substituted for the *Nde*I site. The nucleotide sequence commences at, and includes, the *Bgl*II site shown in the plasmid map, and extends beyond the overlapping *Ava*I and *Xho*I sites. Other details are explained in the text.

commonly used for expression from pET plasmids, is a lysogen in which bacteriophage DE3 DNA, carrying the T7 RNA polymerase gene, the *lacI* gene [encoding the lactose operon repressor protein (*lac* repressor)] and the *lac*UV5 promoter are incorporated into the genome of *E. coli* BL21 (Studier & Moffatt, 1986). The *lac*UV5 promoter, which is repressed in the absence of the inducer IPTG, directs transcription of the T7 RNA polymerase gene. The addition of IPTG to growing cultures of transformed *E. coli* BL21 cells induces the synthesis of T7 RNA polymerase, which, in turn, transcribes the target DNA in the plasmid. *E. coli* BL21, an *E. coli* B strain, is deficient in some proteases (Grodberg & Dunn, 1988), which can be an advantage when expressing proteins that are particularly susceptible to proteolysis.

In the absence of IPTG, a basal level of transcription of the T7 RNA polymerase gene and, therefore, low-level expression of target DNA, does occur, consequently, genes whose products are toxic cannot be established in *E. coli* BL21. There have been two further refinements of the basic system. Firstly, in plasmids such as pET21a (Figure 4.1.1), a *lac* operator sequence is inserted downstream from the T7 promoter, and the plasmid also carries the *lacI* gene. The T7 promoter, as well as the *lac*UV5 promoter, is therefore under the control of the *lac* repressor. Secondly, a second series of expression hosts, the BL21(DE3)pLys hosts, were developed. These hosts carry the T7 lysozyme gene, and the gene encoding chloroamphenicol resistance, on a plasmid (Studier *et al.*, 1990). T7 lysozyme is a natural inhibitor of T7 RNA polymerase and thus prevents basal-level expression of the target gene in the absence of IPTG. The high levels of T7 RNA polymerase that are produced following induction by IPTG are sufficient to overwhelm inhibition by T7 lysozyme. T7 lysozyme is tolerated by *E. coli*, apparently because the protein is unable to pass through the inner membrane to reach the peptidoglycan layer.

4.1.1.2 The pTrc Expression System

The pTrc series of expression vectors were developed by Amann *et al.* (1988) and are derivatives of the pKK233-2 expression vector developed by Amann & Brosius (1985). A plasmid map of a typical pTrc vector (pTrc99A), and the nucleotide sequence across the transcription and translation signals, are shown in Figure 4.1.2. The essential features of pTrc vectors are the -35 promoter element of the *E. coli* tryptophan operon (-35 *trp*), the -10 element of the *E. coli lac* operon (-10 *lac*), the *E. coli lac* operator sequence (*lac* operator), the *E. coli lacZ* ribosome-binding site (*lacZ* RBS), a *Nco*I site, inserted immediately upstream from the multiple cloning site of pUC18, and the *rrnB* transcription terminator sequences (*rrnB* T1 and *rrnB* T2, Brosius *et al.*, 1981), inserted downstream from the multiple cloning site. The plasmid backbone is derived from pBR322. The plasmid also contains the *lacI* gene and, therefore, transcription is induced by the addition of IPTG in a manner



-35 *trp* **-10 *lac*** ***lac* operator**

TTGACAATTAATCATCCGGCTCGTATAATGTGTGGAATTGTGAGCGGATAACAAT

lacZ **RBS** *Nco*I *Sac*I

TCACACAGGAAACAGACCATGGAATTCGAGCTCGGTACCCGGGGATCCTCTAGAG

*Sal*I *Pst*I *rrnB* terminator

CGACCTGCAGGCATGCAAGCTTGGCTGTTTTGGC

Figure 4.1.2 **Structure of Plasmid pTrc99A**

In the plasmid map, the notation Ptrc indicates the transcription and expression region of the plasmid. All other notations are as in the legend to Figure 4.1.1. The details of the nucleotide sequence are explained in the text.

analogous to that described for the pET21 vectors. A coding sequence cloned into a pTrc vector in such a manner that the second codon is fused to the ATG of the *Nco*I site will be correctly spaced with respect to the *lacZ* ribosome-binding site and, in the presence of IPTG, will be transcribed from the hybrid -35 *trp*-10 *lac* promoter and translated from the *lacZ* ribosome-binding site.

The plasmid pTrc99A (Figure 4.1.2) was designed to utilise *E. coli* RNA polymerase. Induction of expression is controlled by the plasmid itself, therefore, a wide variety of *E. coli* strains can be used as host. The use of separate cloning and expression hosts (necessary for the pET system) is not a requirement of the pTrc system.

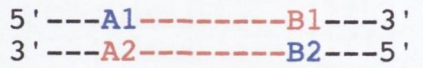
4.1.2 Polymerase Chain Reaction

The polymerase chain reaction (PCR, Saiki *et al.*, 1985; Mullis & Faloona, 1987) is used to amplify a segment of DNA that lies between two regions of known sequence. The principles of the procedure, as applied to circular plasmid DNA, are summarised in Figure 4.1.3. Two oligonucleotides are used as primers for a series of synthetic reactions catalysed by a DNA polymerase. The two priming sites lie on opposite strands of the template DNA, and flank the segment of DNA that is to be amplified. Template DNA is first denatured by heating in the presence of a large molar excess of the two primers and the four deoxynucleoside triphosphates. The reaction mixture is then cooled to allow the primers to anneal to their target sequences and the annealed primers are then extended by the addition of thermostable DNA polymerase. The cycle of denaturation, annealing and DNA synthesis is repeated many times with the products of one cycle serving as templates for the next. The product of the first cycle is a heterogeneous mixture of DNA molecules of undefined length. In the second cycle these molecules generate products of defined length [(d) in Figure 4.1.3] that act as templates in subsequent cycles and accumulate in an exponential fashion. Thus, the major product of PCR with a circular plasmid as template is a segment of linear double-stranded DNA whose termini are defined by the 5' termini of the oligonucleotide primers and whose length is defined by the distance between the primers. The primers are integrated into the product of the first cycle and are therefore part of the template for subsequent cycles, consequently the primers can be used to introduce base changes into the sequences flanking the amplified DNA.

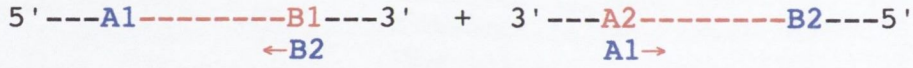
4.1.3 Selection of a Suitable Expression System for *adh*2334

The pET-*E. coli* BL21 system presented an attractive option for the production of rADH. The efficiency and selectivity of T7 polymerase can result in target protein levels as high as 50% of the total cell protein (Studier *et al.*, 1990). Production of such high levels of recombinant protein, at the expense of *E. coli*

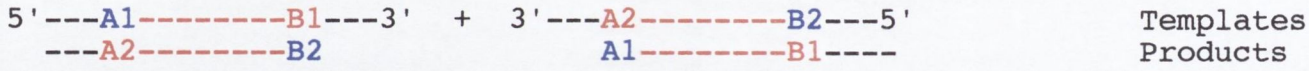
(a) For convenience only a portion of a plasmid is shown. The sequence to be amplified is indicated in red and is flanked by two regions of known sequence, A and B. The PCR primers have sequences equivalent to regions A1 and B2 (blue).



(b) DNA strands are separated by heating. The reaction mixture is cooled to hybridise primers A1 and B2 to single-strand templates. Primer extension occurs in the direction of the arrows.



(c) The products of the first cycle are undefined at one end.



(d) The products of the first cycle act as templates in the second cycle, thus the products of the second cycle are defined at both ends.

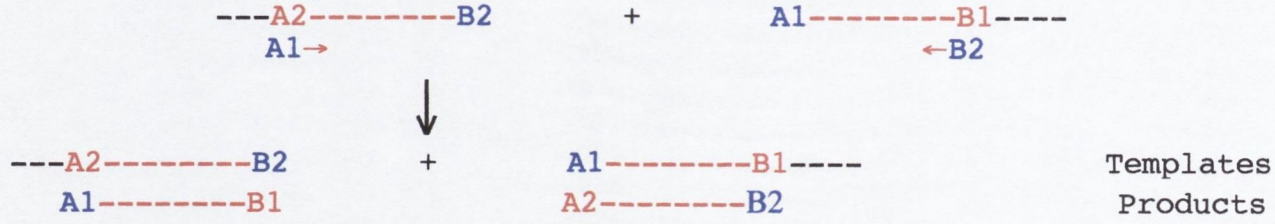


Figure 4.1.3 PCR on a Plasmid Template

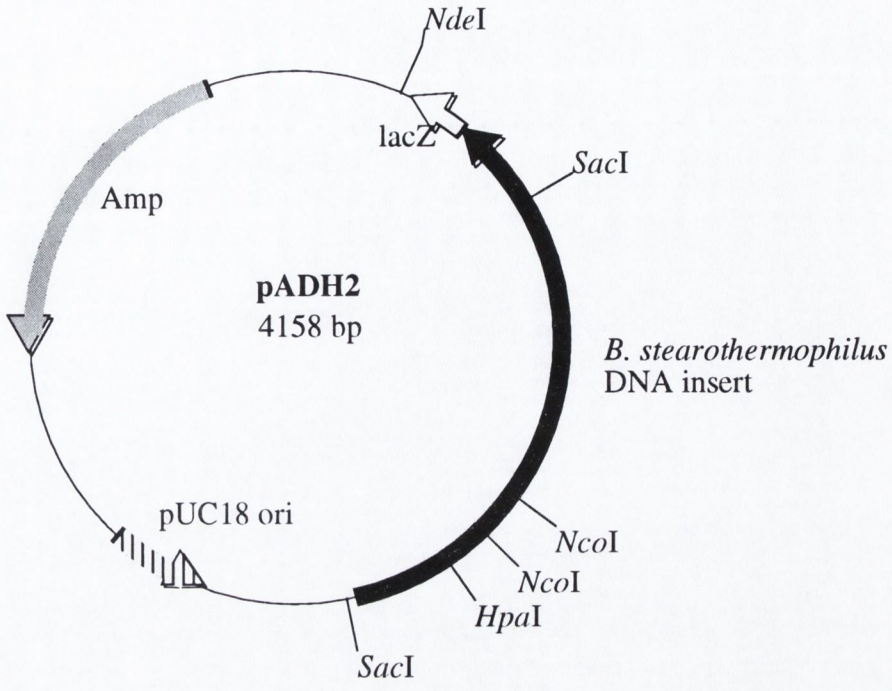
proteins, greatly simplifies purification of the desired protein, but does introduce the risk of the formation of inclusion bodies (Hockney, 1994).

To facilitate discussion of a strategy for designing an expression system for *adh*2334, a simplified plasmid map of pADH2, and details of the nucleotide sequences across the 5' end and the 3' end of *adh*2334 are shown in Figure 4.1.4. Plasmid pADH2 contained the promoter elements, translation elements and the entire coding sequence of *adh*2334 (Chapter 3). The *Sac*I site downstream from the *adh*2334 stop codon was conveniently placed for cloning the complete coding sequence (including stop codon) into either a pET21 vector or into pTrc99A. The 5' end of *adh*2334 was somewhat more problematic as the GTG start codon was not compatible with either the *Nco*I, or the *Nde*I, sequence.

The 5' *Hpa*I site (unique in pADH2) was conveniently placed for removal of the *B. stearothermophilus* promoter sequences, ribosome-binding site and the 5' end of the coding sequence. A scheme involving restriction of pADH2 with *Hpa*I and *Sac*I, to generate a 1-kb *Hpa*I-*Sac*I fragment containing the *adh*2334 coding sequence minus the first eighteen 5' nucleotides (six codons), and ligation of this fragment into pET21a in the presence of a synthetic double-stranded *Nde*I-*Hpa*I adapter, is shown in Figure 4.1.5. The synthetic adapter provides a *Nde*I-compatible end, replaces the GTG start codon with ATG, restores the *adh*2334 coding sequence in the correct reading frame, and places the coding sequence in the correct position relative to the pET21a transcription and translation signals. The *Nde*I and the *Sac*I cloning sites are restored and are unique in the final construct, thus, the scheme allows for the removal of the *adh*2334 coding sequence (including start and stop codons) to another vector. The scheme also allows for the insertion (*via* the *Sac*I site) of the *aldh*2334 coding sequence downstream from the *adh*2334 coding sequence. The scheme does not allow for the translation of the His-Tag, as translation terminates at the *adh*2334 stop codon.

An alternative scheme, involving the use of PCR to change the GTG start codon to ATG and to introduce restriction sites at the 5' and 3' ends of the *adh*2334 coding sequence, is outlined in Figure 4.1.6. The PCR primer on the left of Figure 4.1.6(a) introduces three base-changes at the 5' end of the *adh*2334 coding sequence, thus introducing a *Bsp*HI site, changing the GTG start codon to ATG, but otherwise leaving the coding sequence intact. The PCR primer on the right introduces two base-changes downstream from the *adh*2334 stop codon, thereby introducing an *Xho*I site. The PCR product, after amplification, is a linear, double-stranded molecule, ~1 kb long with a *Bsp*HI site at the 5' end of the coding sequence and an *Xho*I site at the 3' end.

Restriction at a *Bsp*HI site produces a *Nco*I compatible end and restriction at *Xho*I produces an end compatible with either an *Xho*I end or a *Sal*I end. Therefore



RBS start *Hpa*I

217 GAAAGGGGTTGATTTGA**GTG**AAAGCAGCAGTAGTTAACGAATTTAAAAAA 267

M K A A V V N E F K K

5'-end of *adh*2334 in plasmid pADH2 (sense strand)

stop *Sac*I

1247 GAGGAT**TGA**TGATGAGCTCAATTGCAG 1272

E D

3'-end of *adh*2334 in plasmid pADH2 (sense strand)

Figure 4.1.4 Restriction Sites in pADH2 Relevant to Cloning *adh*2334 into an Expression Vector

*adh*2334 extends from, and includes, fifteen bases upstream from the *Hpa*I site (underlined) to, and excluding, four bases upstream from the *B. stearothermophilus Sac*I site (underlined). Numbering is the same as in Figure 3.4.1. The start and stop codons of *adh*2334 are shown in bold. The *adh*2334 ribosome binding site (RBS) is underlined. All other notations are defined in the legend to Figure 3.1.1.

1. Cleave pADH2 (Figure 4.1.3) with *Hpa*I and *Sac*I to produce a 1 kb *Hpa*I-*Sac*I fragment carrying the *adh*2334 coding sequence minus the first six codons. The nucleotide sequence, and translation, at each end of the 1 kb fragment are shown.

```
      HpaI                                        stop      SacI
5'-AACGAATTTAAAAAA____________GAGGATTGATGATGAGCT-3'
3'-TTGCTTAAATTTTTT____________CTCCTAACTACTAC-5'
     N  E  F  K  K                E  D
```

2. A double-stranded adapter with *Nde*I-*Hpa*I compatible ends.

```
      NdeI          HpaI
5'-TATGAAAGCAGCAGTAGTT-3'
3'-ACTTTCGTCGTCATCAA-5'
```

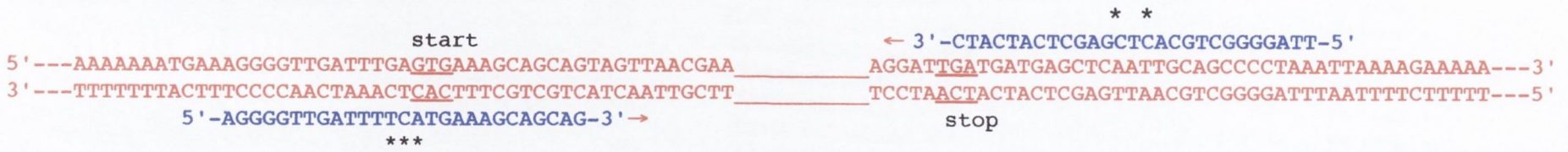
3. Ligate the adapter to the 1 kb *Hpa*I-*Sac*I *adh* fragment, to restore the coding sequence.

```
  NdeI              HpaI                                  stop      SacI
5'-TATGAAAGCAGCAGTAGTTAACGAATTTAAAAAA____________GAGGATTGATGATGAGCT-3'
  3'-ACTTTCGTCGTCATCAATTGCTTAAATTTTTT____________CTCCTAACTACTAC-5'
     M  K  A  A  V  V  N  E  F  K  K                E  D
```

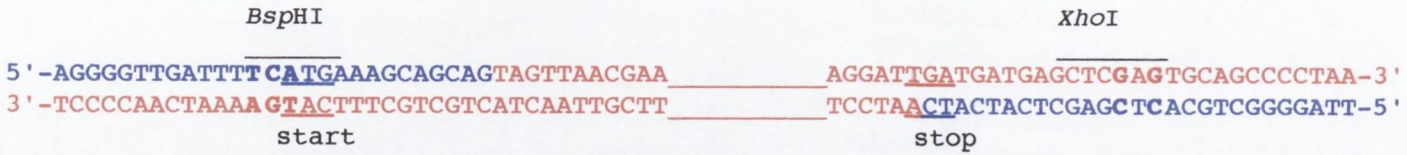
4. The *Nde*I-*Sac*I adapter-*adh* construct will ligate into *Nde*I-*Sac*I pET21a (Figure 4.1.1).

Figure 4.1.5 A Strategy for Cloning the *adh*2334 Coding Sequence into pET21a

(a) The nucleotide sequences adjacent to the start and stop codons (underlined) of *adh*2334 in plasmid pADH2 are indicated in red. Two PCR primers are indicated in blue with the directions of primer extensions indicated by arrows. The positions of mismatched bases in the primer sequences are indicated by an asterisk.



(b) After PCR amplification the major product would be a linear DNA molecule containing the *adh*2334 coding sequence flanked by a *Bsp*HI site and an *Xho*I site (overlined) that were created by the base changes (indicated in bold type) introduced by the mismatched primers. The GTG start codon has been changed to ATG.



(c) Restriction of the PCR product with *Bsp*HI produces a *Nco*I compatible end. Restriction with *Xho*I produces an end compatible with either an *Xho*I or a *Sal*I end. The *Bsp*HI-*Xho*I PCR product will ligate into *Nco*I-*Xho*I pET21d and into *Nco*I-*Sal*I pTrc99A.

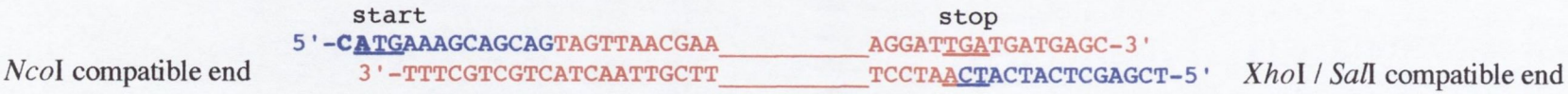


Figure 4.1.6 Strategy, Using PCR, for Cloning the *adh*2334 Coding Sequence into pET21d or pTrc99A

restriction of the PCR product with *Bsp*HI and *Xho*I renders the product suitable for cloning into either *Nco*I-*Xho*I pET21d or into *Nco*I-*Sal*I pTrc99A. The *Nco*I site is not recreated in the final construct.

4.2 MATERIALS AND METHODS

4.2.1 Cloning Hosts, Expression Hosts and Plasmids

Cloning hosts and expression hosts are listed in Table 4.2.1 and plasmids are listed in Table 4.2.2. The *E. coli* strains SURE, BL21 and BL21(DE3)pLysS (pLysS) were propagated from single colonies obtained from Dr. Mark Carrington (University of Cambridge). *E. coli* strain DH5α was propagated from a frozen glycerol stock obtained from Dr. Larry Vaughan (Trinity College, Dublin). Plasmid pADH2 DNA was that used in the determination of the nucleotide sequence of *adh*2334 in Chapter 3. Plasmid pET21a, a 490 ng μl^{-1} solution, was obtained from Novagen, transformed into *E. coli* TG1 and purified from a large-scale culture. Plasmid pTrc99A, a 200 ng μl^{-1} solution, was obtained from Dr. Orla Ennis (Trinity College, Dublin) transformed into *E. coli* SURE and purified from a large-scale culture. Plasmid pET21d DNA was obtained from Dr. Mark Carrington , as a solution (100 ng μl^{-1}) of plasmid DNA that had been purified by equilibrium centrifugation in a CsCl-ethidium bromide gradient, and was used directly.

4.2.2 Working Stocks of Cells, Culture Media and Culture Conditions

The procedures described in Chapter 3 (Section 3.2.4) were used to prepare working stocks of *E. coli* on agar plates: strains TG1 and SURE were stored on M9 agar; working stocks of other strains and all transformed *E. coli* were stored on LB agar or 2TY agar, with or without antibiotics, as appropriate. Culture media and culture conditions were as described in Chapter 3. *E. coli* strains TG1, DH5α and SURE were grown in LB medium, and strains BL21 and pLysS were grown in 2TY medium. Cells transformed with pET vectors, or with pTrc99A, were grown in medium that contained ampicillin at a final concentration of 0.1 mg ml^{-1}. Cells transformed with pADH2 were grown in medium that contained ampicillin at a final concentration of 0.2 mg ml^{-1}. The presence of the pLysS plasmid is verified by resistance to chloramphenicol, therefore, *E. coli* pLysS cells were always grown in the presence of chloramphenicol (final concentration in media, 34 μg ml^{-1}). Chloramphenicol was stored at -20 °C as a solution in ethanol (34 mg ml^{-1}).

4.2.3 Storage of *E. coli* as Cryocultures

Laboratory stocks of *E. coli*, and *E. coli* transformed with plasmid DNA, were stored indefinitely as cryocultures at -70 °C. Cryocultures were prepared by mixing 0.7 ml of a chilled stationary-phase culture with 0.3 ml of an ice-cold 50% (w/v) solution of glycerol, the suspension was snap-frozen in liquid nitrogen and stored at -70 °C.

E. coli strain	Reference and genotype	Function
TG1	Section 3.2.2	Cloning
DH5α	Hanahan (1983) F'/*endA1 hsdR17*($r^-_K m^-_K$) *supE44 thi-1 recA1 gyrA* (Nal^r) *relA1* Δ(*lacZYAargF*)*U169deoR* (ϕ80*dlac* Δ(*lacZ*)*M15*) recombination-deficient	Cloning
SURE	Greener (1990) e14$^-$ (McrA$^-$)Δ(*mcrCB-hsdSMR-mrr*)*171 endA1 supE44 thi-1 gyrA96 relA1 lac recB recJ sbcC umuC::*Tn*5*(Kan^r) *uvrC*[F' *proAB lacI*^{q}Z*ΔM15* Tn*10*(Tet^r) Amy Cam^r]^{c,d} recombination-deficient	Cloning
BL21(DE3)	Studier & Moffatt (1986) F'*ompThsdS*$_B$($r^-_B m^-_B$; an *E. coli* B strain) a lysogen carrying the T7 RNA polymerase gene	Expression
BL21(DE3)pLysS	Studier *et al.* (1990) as BL21(DE3) with the T7 lysozyme gene and the gene encoding resistance to chloramphenicol on a plasmid	Expression

Table 4.2.1 ***E. coli* Host Strains**

Plasmid	Reference	Function
pADH2	Chapter 3	Source of the *adh*2334 coding sequence
pET21a	A derivative (by Novagen) of the original pET vectors described by Studier *et al.* (1990) Figure 4.1.1	Expression vector
pET21d	Differs from pET21a only in that the *Nde*I restriction site of pET21a is replaced by a *Nco*I restriction site	Expression vector
pTrc99A	Amann *et al.* (1988) Figure 4.1.2	Expression vector

Table 4.2.2 **Cloning and Expression Plasmids**

4.2.4 Synthetic Oligonucleotides

The nucleotide sequences, and functions, of the synthetic oligonucleotides A_s, A_a, PB_s and PX_a are presented in Table 4.2.3. The oligonucleotides A_s and A_a were synthesised (40 nmol scale) at the Department of Genetics, Trinity College, Dublin, using an Applied Biosystems model 391 PCR MATE DNA synthetiser and the phosphite-triester method. A_s and A_a were obtained from the DNA synthetiser in solution in ammonia, with the dimethoxy trityl protecting groups removed, and required further purification. The solutions were incubated for 16 h at 55 °C and the ammonia was then removed under vacuum. The solid oligonucleotides were dissolved in H_2O (40 µl) and centrifuged (10,000 ***g***, 5 min). The clarified oligonucleotide solutions were removed into clean tubes containing 160 µl of H_2O, chloroform was added (200 µl), the tubes were vortex-mixed, centrifuged (10,000 ***g***, 10 min) and the oligonucleotides were recovered from the aqueous layers by precipitation in the presence of NaCl and ethanol (Section 3.2.13). The final oligonucleotide pellets were dissolved in 20 µl of H_2O. The concentration of solutions of A_s and A_a was estimated by comparing the fluorescence observed when an aliquot of oligonucleotide solution and an aliquot of a DNA sequencing primer (Section 3.2.19.1) of exactly the same length, and of known concentration, were electrophoresed through polyacrylamide (denaturing conditions; Section 4.2.5).

The PCR primers, PB_s and PX_a (0.2 µmol of each), were synthesised at the Protein and Nucleic Acid Chemistry Facility, Department of Biochemistry, University of Cambridge, using the phosphite-triester methodology. The PCR primers were obtained from the DNA synthetiser as salt-free powders and with the dimethoxy trityl protecting groups removed. Each primer was dissolved in 1 ml of sterile H_2O and the concentration was estimated by the spectrophotometric method (Section 3.2.19.1). The estimated concentration of each oligonucleotide solution was approximately 1 µg μl^{-1}. The solutions were stored at -20 °C and used without further purification.

4.2.5 Polyacrylamide-Gel Electrophoresis of Oligonucleotides

Polyacrylamide-gel electrophoresis of oligonucleotides, under both denaturing and nondenaturing conditions, was performed in the vertical slab-gel apparatus described in Chapter 2 (Section 2.1.4), with TBE as the anode and the cathode buffer. Agarose gel loading-buffer (Section 3.2.16) was used as the sample buffer and as tracker-dye. Following electrophoresis, gels were soaked for 10 min in a solution of ethidium bromide (5 µg ml^{-1}) and viewed using a transilluminator.

Single-stranded synthetic oligonucleotides were analysed under denaturing conditions by electrophoresis through gels comprising 19% (w/v) acrylamide, 1% (w/v) bisacrylamide, 42% (w/v) urea. The acrylamide-bisacrylamide

Designation	Function	Nucleotide sequence
A_S	The sense strand of a *Nde*I - *Hpa*I adapter (see below)	5'-TATGAAAGCAGCAGTAGTT-3'
A_a	As for A_S, but the antisense strand	5'-AACTACTGCTGCTTTCA-3'
*Nde*I - *Hpa*I adapter	A double-stranded adapter constructed by annealing together A_S and A_a. The adapter introduced a *Nde*I site at the 5' end of *adh*2334 and consequently altered the GTG start codon to ATG. Bases that contribute to the *Hpa*I site (blunt end) and the *Nde*I site (staggered end) are underlined and the *adh* start codon (ATG) is overlined.	start 5'-TATGAAAGCAGCAGTAGTT-3' 3'-ACTTTCGTCGTCATCAA-5'
PB_S	A PCR primer used to introduce a *Bsp*HI site (underlined) at the 5' end of *adh*2334 and, consequently, to alter the GTG start codon to ATG (overlined). Bases in bold type are base changes that were introduced into *adh* 2334 by the PCR. The primer directed the synthesis of the sense strand of the *adh* gene, with plasmid pADH2 as the template.	start 5'-AGGGGTTGATTT**TCA**TGAAAGCAGCAG-3'
PX_a	A PCR primer used to introduce a *Xho*I site (underlined) downstream from the *adh* 2334 stop codon. Bases that contribute to the *adh* stop codon are overlined and bases in bold type are base changes introduced into the *adh* gene by the PCR. The primer directed the synthesis of the antisense strand of the *adh* gene, with plasmid pADH2 as the template.	stop 3'-CTACTACTCGAG**CTC**ACGTCGGGGATT-5'

Table 4.2.3 **Nucleotide Sequences and Functions of Synthetic Oligonucleotides**

stock solution was 38% (w/v) acrylamide, 2% (w/v) *N*, *N*'-methylenebisacrylamide, and was deionised (Section 3.2.19.4) before use. Gels were prepared by mixing 5 ml of acrylamide-bisacrylamide stock solution, 1 ml of ten-times concentrated TBE and 4.2 g of urea, the volume was adjusted to 10 ml and 10 µl of *N*,*N*,*N'*,*N*'-tetramethylethylenediamine and 30 µl of 25% (w/v) ammonium persulfate were added. Gels were heated to 40 °C (444 V, 27 mA, 1 h) before loading the samples.

Aliquots of oligonucleotide solutions (up to 5 µl) were mixed with 4 µl of formamide, incubated at 75 °C for 2 min, and then loaded onto the gel. The progress of electrophoresis was monitored by observing the migration of bromophenol blue (4 µl of agarose gel loading-buffer) that had been loaded in one lane of the gel. Gels were maintained at 40 °C throughout electrophoresis (444 V, 27 mA), and electrophoresis was terminated when the bromophenol blue came within 2 cm of the bottom of the gel.

Electrophoresis of oligonucleotides under nondenaturing conditions was through 19% (w/v) acrylamide, 0.66% (w/v) bisacrylamide. The acrylamide-bisacrylamide stock solution was 29% (w/v) acrylamide, 1% (w/v) *N*, *N*'-methylenebisacrylamide. Gels were prepared by mixing 6.66 ml of acrylamide-bisacrylamide stock solution, 2.27 ml of H_2O, 1 ml of ten-times-concentrated TBE, 100 µl of 10% (w/v) ammonium persulfate and 10 µl of *N*,*N*,*N'*,*N*'-tetramethylethylenediamine. Samples of interest (200 - 600 ng, up to 5 µl) were mixed with 2 µl of agarose gel loading-buffer and were loaded onto the gel. Electrophoresis was for 90 min at 72 V.

4.2.6 Phosphorylating and Annealing Synthetic Oligonucleotides

Kinase (8 U μl^{-1}) was obtained from Promega. The ATP used was a 1 mM solution in 10 mM Tris-HCl, pH 8, and was stored as small aliquots at -20 °C. Kinase buffer was 500 mM Tris-HCl, 100 mM $MgCl_2$, 50 mM dithiotretiol, 1 mM Na_2EDTA, pH 7.6.

The two synthetic oligonucleotides, which were used to construct the double-stranded *Nde*I-*Hpa*I adapter, were phosphorylated, separately, in a final volume of 10 µl, and with approximately 1 U of kinase per pmol of 5' ends. A volume of oligonucleotide, equivalent to approximately 16 pmol, was mixed with 1 µl of kinase buffer, 1.5 µl of 1 mM ATP, 16 U of kinase and the appropriate volume of H_2O. The solutions were incubated at 37 °C for 2 h, and then at 68 °C for 10 min. The two oligonucleotides were then annealed by mixing the two solutions. The annealing mixture was incubated at 70 °C for 3 min and then at 37 °C for 30 min. The annealed oligonucleotides were stored at -20 °C.

4.2.7 Purification of Plasmid DNA

Small-scale purification of plasmid DNA from 2-ml cultures was by the 'boiled-lysis' method described in Chapter 3 (Section 3.2.13.2). Recovery of DNA from agarose was as described in Section 3.2.13.4.

Large-scale purification of plasmid DNA was by the alkali-lysis method developed by Birnboim & Doly (1979) and Ish-Horowicz & Burke (1981). The potassium acetate buffer used to neutralise the alkaline lysates was prepared by mixing 5 M potassium acetate (60 ml), acetic acid (11.5 ml) and H_2O (28.5 ml).

Transformed *E. coli* cells were harvested from 200-ml stationary-phase cultures (2000 ***g***, 4 °C, 15 min), resuspended in 8 ml of TES-sucrose, transferred to 50-ml centrifuge tubes and lysed by the addition of 8 ml of a freshly prepared solution of 2% (w/v) SDS in 0.4 M NaOH. The resulting lysates were neutralised by adding 9.6 ml of potassium acetate buffer and were then placed on ice for 5 min. Cell debris and chromosomal DNA were removed by centrifugation (43,000 ***g***, 10 min, 4 °C). The cleared supernatant was removed into a 50-ml centrifuge tube, 0.6 volumes of propan-2-ol were added and the tube was left at room temperature for at least 30 min. The plasmid DNA precipitate was collected by centrifugation (43,000 ***g***, 15 min, 4 °C) and washed once (without resuspension) by adding 1 ml of 70% (v/v) ethanol and recentrifuging as before. All visible traces of ethanol were removed by leaving the tube to drain in an inverted position on a pad of tissue paper (10 min, room temperature). The dry pellet of plasmid DNA was resuspended in 1 ml of TE and incubated at 50 °C in the presence of RNase A (100 µl of a 10 mg ml^{-1} solution, 20 min). Protein was removed from the RNase-treated plasmid DNA by extraction into phenol-chloroform-isoamyl alcohol, and the DNA was recovered by precipitation in the presence of NaCl and ethanol. The pellet of plasmid DNA was resuspended in a minimum volume of TE and subjected to equilibrium centrifugation in a CsCl-ethidium bromide gradient, to remove contaminating chromosomal DNA (Section 3.2.13.3).

4.2.8 Electrophoresis of Plasmid DNA

Electrophoresis and subsequent visualisation of plasmid DNA were exactly as described in Section 3.2.16. Agarose concentrations (final) were either 0.8 or 1% (w/v). The DNA size markers used are listed in Tables 4.2.4 and 4.2.5. The *Eco*RI digest of phage SPPI DNA (Table 4.2.4) was obtained from Boehringer and the *Bst*EII digest of phage λ DNA (Table 4.2.5) was obtained from Dr. Maria Redpath (University of Cambridge). The *Bst*EII phage λ size markers were incubated at 60 °C for 5 min before electrophoresis in an (unsuccessful) attempt to prevent hybridisation of the 8454-bp fragment to the 5686-bp fragment. Yields of restriction fragments were estimated by comparing the fluorescence of the fragment of interest

Fragment number	Amount of DNA (ng) per 0.5 μg of digest	Size of fragment kb
1	97.5	8.51
2	84.3	7.35
3	70.0	6.11
4	55.5	4.84
5	41.2	3.59
6	32.2	2.81
7	22.4	1.95
8	21.3	1.86
9	17.3	1.51
10	15.9	1.39
11	13.3	1.16
12	11.2	0.98
13	8.3	0.72
14	5.5	0.48
15	4.1	0.36

Table 4.2.4 DNA Size Markers, *Eco*RI Digest of Phage SppI DNA

Fragment number	Size of fragment bp
1	8454
2	7242
3	6369
4	5686
5	4822
6	4324
7	3675
8	2323
9	1929
10	1371
11	1264
12	702
13	224
14	117

Table 4.2.5 DNA Size Markers, *Bst*EII Digest of Phage λ DNA

with that of the fragment closest in size in the phage SppI DNA size markers (Table 4.2.4).

The concentration of double-stranded DNA was estimated by comparing the fluorescence observed when a restriction digest of the DNA of interest and a known amount of *Eco*RI phage SPPI size markers (Table 4.2.4) were electrophoresed through 1% agarose and visualised using ethidium bromide.

4.2.9 Recovery of DNA in Centricon Concentrators

Centricon concentrators (Centricon YM-100, membrane impervious to 100,000-M_r) were obtained from Amicon and were used to remove residual traces of phenol and PCR primers from solutions of DNA. The solution of DNA (less than 500 µl) was added to 2 ml of H_2O, which was contained in the top chamber of the Centricon and the device was centrifuged (2000 ***g***, 30 min, 10 °C). The concentrated DNA solution (less than 100 µl) was left in the Centricon, 1 ml of H_2O was added, and the concentration step was repeated as before, with the DNA being recovered in a volume of approximately 35 µl.

4.2.10 Polymerase Chain Reaction

PCR was performed with plasmid pADH2 as the template DNA, and with PB_s and PX_a (Table 4.2.3) as PCR primers. PCR reactions were performed in an "Allsopp" programmable PCR machine. PCR buffer was 500 mM KCl, 100 mM Tris-HCl, 1% (v/v) Triton X-100, pH 9.0. dXTP mix comprised 10 mM each of the four deoxynucleoside triphosphates (dATP, dCTP, dGTP and dTTP). *Thermus aquaticus* Taq polymerase (2 U μl^{-1}) was recombinant protein purified from *E. coli* by the method of Engelke *et al.* (1990), and was obtained from Dr. Mark Carrington. Solutions of PB_s and PX_a were 200 pmol μl^{-1}.

Plasmid pADH2 DNA (1 ng) was mixed with 10 µl of PCR buffer, 2.5 µl of dXTP mix, 5 µl of 50 mM $MgCl_2$, 100 pmol of each of the PCR primers, PB_s and PX_a, and the volume was adjusted to 100 µl with H_2O. The reaction mixture was overlaid with 200 µl of paraffin oil and maintained at 96 °C for 6 min (to denature template DNA). The mixture was cooled to 50 °C, Taq polymerase (2 U) was added and the reaction mixture was subjected to twenty cycles of amplification (extension at 72 °C for 0.8 min, increasing to 1.4 min over twenty cycles, denaturation at 93 °C for 0.5 min). The programme was terminated after a final extension for 10 min.

The paraffin oil was removed from the PCR reaction mixture by adding 200 µl of chloroform, mixing vigorously and centrifuging at 10,000 ***g*** for 5 min. The paraffin-free (upper) aqueous layer was removed into a clean tube and extracted twice with phenol-chloroform-isoamyl alcohol. The final aqueous layer, containing the PCR product (*adh* PCR product), was removed into 2 ml of H_2O, which were

contained in a Centricon concentrator, and the residual phenol-chloroform-isoamyl alcohol and the primers were removed as described in Section 4.2.9. The *adh* PCR product was recovered in a volume of 35 µl and at a DNA concentration of approximately 200 ng ml^{-1}.

4.2.11 Restriction of DNA

The restriction enzymes *Sac*I, *Hpa*I, *Nde*I, *Nco*I and *Sal*I were obtained from Boehringer as 10 U μl^{-1} solutions. *Xho*I (20 U μl^{-1}) and *Bsp*HI (5 U μl^{-1}) were obtained from New England BioLabs. The buffer systems used with the various enzymes and combinations of enzymes are given in Table 4.2.6. There was a problem associated with obtaining complete restriction in a reasonable time-scale and that appeared to be due to a change in the quality of the laboratory water. Water from various sources was surveyed and the most complete restriction in the shortest time was found to occur in h.p.l.c. grade water (Section 2.1.2). Autoclaved h.p.l.c. grade water was therefore used when preparing restriction digest reaction mixtures.

Analytical and preparative restriction digests of plasmid DNA (with a single enzyme or with two enzymes simultaneously) and subsequent analysis of the restriction fragments were, in general, as described in Chapter 3 (Section 3.2.17). When h.p.l.c. grade water was used, typical analytical digests contained 400 ng of DNA, 5 U of any particular enzyme and the time-scale was 1 - 2 h.

4.2.12 Dephosphorylation of Restricted Plasmid DNA

Calf intestine alkaline phosphatase (1 U μl^{-1}) was obtained from Boehringer and from Promega. Restricted DNA (typically 1 µg, 0.057 pmol of 5'-phosphorylated ends) that had been recovered from agarose gels was mixed with 5 µl of Promega dephosphorylation buffer (500 mM Tris-HCl, 10 mM $MgCl_2$, 1 mM $ZnCl_2$, 10 mM spermidine, pH 9.0), 0.5 U of alkaline phosphatase and the appropriate volume of H_2O to give a final volume of 50 µl. The mixture was incubated at 37 °C for 15 min and then at 56 °C for 15 min. A further 0.1 U of alkaline phosphatase was added, the mixture was incubated at 37 °C for 15 min and then at 56 °C for 15 min. Alkaline phosphatase was removed from the reaction mixture by extraction into phenol-chloroform-isoamyl alcohol and the dephosphorylated DNA was recovered from the final aqueous layer by precipitation from solution in the presence of NaCl and ethanol.

In an alternative procedure, restricted DNA was dephosphorylated by adding the alkaline phosphatase directly to the restriction mixture. Plasmid DNA (5 µg) that had been restricted with *Xho*I and *Nco*I (pET21d) or with *Sal*I and *Nco*I (pTrc99A) was dephosphorylated by adding 5 U of calf intestine alkaline phosphatase directly to the restriction mixture and incubating at 37 °C for 15 min. The

Enzyme(s)	Buffer system (final concentration in digest)	Source of buffer and supplier's designation
*Sac*I	33 mM Tris-acetate, 10 mM magnesium acetate, 66 mM potassium acetate, 0.5 mM dithiothreitol, pH 7.9	Boehringer Buffer A
*Hpa*I	As for *Sac*I	
*Sac*I plus *Hpa*I	As for *Sac*I	
*Nde*I	10 mM Tris-HCl, 10 mM $MgCl_2$, 50 mM NaCl, 1mM dithioerythritol, pH 7.5	Boehringer Buffer M
*Nde*I plus *Sac*I	As for *Nde*I	
*Nco*I	50 mM Tris-HCl, 100 mM NaCl, 10 mM $MgCl_2$,1 mM dithioerythritol, pH 7.5	Boehringer Buffer H
*Xho*I	As for *Nco*I	
*Nco*I plus *Xho*I	As for *Nco*I	
*Xho*I plus *Bsp*HI	20 mM Tris-acetate, 10 mM magnesium acetate, 50 mM potassium acetate, 1 mM dithiothreitol, pH 7.9	New England BioLabs Buffer 4
*Sal*I	As for *Nco*I	

Table 4.2.6 **Buffer Systems Used for Restriction Enzymes**

dephosphorylated DNA was recovered, as described above, and resuspended in 50 µl of H_2O. The final DNA concentration was approximately 100 ng μl^{-1}.

4.2.13 Ligation Reactions

T4 DNA ligase (1 Weiss unit μl^{-1}) and ligation buffers were purchased from Gibco BRL and from Boehringer. Ligation buffers were stored as small aliquots at -20 °C and were discarded after a single use. Boehringer ligation buffer was 660 mM Tris-HCl, 50 mM $MgCl_2$, 10 mM dithioerythritol, 10 mM ATP, pH 7.5, and was diluted ten-fold in ligation reaction mixtures. Gibco BRL ligation buffer was 250 mM Tris-HCl, 50 mM $MgCl_2$, 5 mM ATP, 5 mM dithiothreitol, 25% (w/v) polyethylene glycol-8000, pH 7.6, and was diluted five-fold in ligation reaction mixtures. Ligation reactions were performed in a final volume of 10 or 20 µl, at 18 °C, and typically for 16 h, but occasionally for 6 h. 1 Weiss unit of ligase was used for up to 300 ng of total DNA present in the reaction mixture.

In ligations of the *Nde*I-*Hpa*I adapter to the 6.4-kb *adh*-pET21a linear construct, the adapter was present in a twenty-fold molar excess. In ligations of the 1-kb *Hpa*I-*Sac*I *adh* fragment to linear *Nde*I-*Sac*I pET21a, the 1-kb fragment was present in a three-fold molar excess. When the *Nde*I-*Hpa*I adapter was ligated to the 1-kb *Hpa*I-*Sac*I *adh* fragment, a sixteen-fold molar excess of adapter was achieved by the cumulative addition of three aliquots of adapter over a time period of 4.5 h.

In ligation of the *Bsp*HI-*Xho*I *adh* PCR product into pET21d and pTrc99A reaction mixtures contained 0.051 pmol (33.3 ng) of *Bsp*HI-*Xho*I *adh* PCR product, 1 Weiss unit of T4 DNA ligase and 0.024 pmol (100 ng) of dephosphorylated *Nco*I-*Xho*I pET21d or 0.0276 pmol (75 ng) of dephosphorylated *Nco*I-*Sal*I pTrc99A in a final volume of 10 µl. Positive ligation controls contained 100 ng of dephosphorylated *Nco*I-*Xho*I pET21d or 75 ng of dephosphorylated *Nco*I-*Sal*I pTrc99A and 1 Weiss unit of T4 DNA ligase in a final volume of 10 µl. The ligation reaction mixtures were incubated for 6 h at 18 °C and were then transformed (method of Hanahan; Section 4.2.15) into competent *E. coli* cells.

4.2.14 Preparation of Competent *E. coli* Cells

E. coli cells used for the purification of plasmid DNA vectors, *E. coli* cells used in attempts to construct an expression system in pET21a, *E. coli* BL21 and *E. coli* pLysS cells were made competent by the method of Morrison (1979). The method of Hanahan (1983) was used to prepare competent *E. coli* cells for transformation with the ligation products from the subcloning of the *adh* PCR product into pET21d and pTrc99A.

4.2.14.1 Preparation of Competent *E. coli* Cells by the Method of Morrison

E. coli (100-ml culture) were grown until the A_{600} value reached 0.4 - 0.5 and were then chilled by immersion of the culture flask in ice. The bacteria were harvested by centrifugation (2000 ***g***, 10 min, 4 °C), resuspended (very gently) in 15 ml of ice-cold $MgCl_2$ (100 mM) and harvested again. The cell pellet was resuspended in 50 ml of ice-cold $CaCl_2$ (100 mM) and was placed on ice for 1 h. Competent cells, other than *E. coli* BL21 and *E. coli* pLysS, were recovered by centrifugation (2600 ***g***, 10 min, 4 °C) and resuspended in 10 ml of ice-cold $CaCl_2$ (100 mM). Centrifugation at 5800 ***g*** was necessary to form a cell pellet of *E. coli* BL21 and *E. coli* pLysS. To compensate for the very low transformation efficiency of *E. coli* BL21 and *E. coli* pLysS, the final cell pellets (from starting cultures of 100 ml) were resuspended in 1 ml of $CaCl_2$. Competent cells were stored on ice for up to 18 h before being transformed or, alternatively, were stored at -70 °C as glycerol suspensions. Glycerol suspensions were prepared by mixing 10 ml of competent cells with 1.5 ml of ice-cold glycerol (80% w/v), the mixture was dispensed into pre-chilled microcentrifuge tubes, snap-frozen by immersion in liquid nitrogen and stored at -70 °C. Competent cells stored at -70 °C displayed no loss in transformation efficiency over a time-scale of six months.

4.2.14.2 Preparation of Competent *E. coli* Cells by the Method of Hanahan

SOB medium was Bacto-tryptone (20 g) and bacto-yeast extract (5 g), dissolved in 1 l of 10 mM NaCl, 2.5 mM KCl. Immediately before use, $MgCl_2$ (10 ml of a 1 M solution) and $MgSO_4$ (10 ml of a 1 M solution) were added. TFB solution (stored at 4 °C) was 10 mM 4-morpholine-ethanesulfonic acid (Mes), 100 mM KCl, 45 mM $MnCl_2$, 3 mM hexamine cobalt (III) chloride, 10 mM $CaCl_2$, pH 6.2, and was prepared using a 100 mM stock solution of Mes that had been adjusted to pH 6.2 with 10 M KOH. DND solution was prepared by mixing 1.53 g dithiothreitol, 9 ml of dimethylsulfoxide, 100 µl of 1 M potassium acetate, pH 7.5, and adjusting the volume to 10 ml.

A single *E. coli* colony was used to inoculate 10 ml of SOB medium and the cultures were grown until the A_{600} value reached 0.3 - 0.5. Cultures were chilled on ice for 10 min, bacteria were harvested by centrifugation (2000 ***g***, 10 min, 4 °C), resuspended in 3 ml of TFB, incubated on ice for 15 min and harvested again. The cell pellet was resuspended in 0.8 ml of TFB and 28 µl of DND were added. The cells were incubated on ice for 10 min, a further 28 µl of DND were added, the cells were incubated on ice for a further 10 min and then transformed with ligation reaction mixtures.

4.2.15 Transformation of *E. coli* cells

Cells made competent by the method of Morrison (Section 4.2.14.1) were transformed as described in Chapter 3 (Section 3.2.9). Cells made competent by the method of Hanahan (Section 4.2.14.2) were transformed as described below (method of Hanahan, 1983).

Ligation reaction mixtures, corresponding to 20 ng of vector DNA (typically 2 - 3 µl), were mixed with 210 µl of competent cells, the tubes were incubated on ice for 30 min, incubated at 42 °C for 2 min and then returned to the ice. An aliquot of each transformation mixture (50 µl) was then spread on LB plus ampicillin agar plates and the plates were incubated for 16 h at 37 °C.

4.2.16 Screening *E. coli* Transformants for IPTG-Dependent Expression

Small-scale cultures (3 ml of 2TY plus ampicillin) were inoculated with single colonies of transformed bacteria. Cultures were grown at 37 °C in an orbital shaker (250 revolutions min^{-1}) until the A_{600} value reached approximately 0.6 (2 - 3 h), 0.5 ml (pre-induction sample) was removed from each culture and stored on ice, the remaining culture being made 1 mM in IPTG (addition of 1 M IPTG) and returned to the incubator. The A_{600} values of the cultures were determined 2.5 h after the addition of IPTG, and 0.5 ml (post-induction sample) was removed from each culture.

Cells were harvested from the pre-induction and post-induction samples by centrifugation (10,000 ***g***, 10 min, room temperature) and the cell pellets were resuspended in 2% (w/v) SDS, 0.025 M Tris-HCl, pH 6.8, 2.5% (w/v) glycerol, 0.003% (w/v) bromophenol blue, 1% (v/v) 2-mercaptoethanol (SDS-PAGE sample buffer). The volume of SDS-PAGE sample buffer to be added to each cell pellet was calculated according to Equation 7 where A_{600} is the A_{600} value of the culture at the time the cells were harvested.

$$\mathbf{A_{600} \div 10 = ml\ of\ SDS\text{-}PAGE\ sample\ buffer} \qquad \mathbf{(7)}$$

Resupended pellets were incubated at 100 °C for 5 min, centrifuged (10,000 ***g***, 10 min), to remove cell debris, and 10 µl of each cleared lysate were analysed by polyacrylamide-gel electrophoresis (Section 4.2.17).

4.2.17. Polyacrylamide-Gel Electrophoresis

E. coli cell lysates were analysed using the sodium dodecyl sulfate-polyacrylamide-gel electrophoresis (SDS-PAGE) system developed by Laemmli (1970). Electrophoresis was performed in the vertical slab-gel apparatus described in

Chapter 2 (Section 2.1.4) and with 6% (w/v) Tris, 2% (w/v) SDS, 28.8% (w/v) glycine as the anode and the cathode buffer. Acrylamide, bisacrylamide and Tris-HCl stock solutions were filtered (3.0 μm pore-size Millipore filters) and stored at 4 °C. Stock solutions of acrylamide were deionised (Section 3.2.19.4).

Proteins used as subunit apparent M_r markers were obtained from Sigma and, unless otherwise specified, were as listed in Table 4.2.7. The M_r markers were dissolved in SDS-PAGE sample buffer to give a standard mixture that provided 2 mg ml^{-1} of each protein. The standard mixture was incubated at 100 °C for 5 min, dispensed into small aliquots and stored at -20°C.

Resolving gels were 15.0% (w/v) acrylamide, 0.086% (w/v) bisacrylamide, 0.375 M Tris-HCl pH 8.7, 0.1 % (w/v) SDS, and were prepared by mixing 1.1 ml H_2O, 3.75 ml of 30% (w/v) acrylamide, 0.65 ml of 1% (w/v) bisacrylamide, 1.87 ml of 1.5 M Tris-HCl, pH 8.7, 37 μl of 20% (w/v) SDS, 50 μl of a freshly prepared solution of 10% (w/v) ammonium persulfate and 5 μl of *N,N,N',N'*-tetramethylethylenediamine. Stacking gels were 5.7% (w/v) acrylamide, 0.15% (w/v) bisacrylamide, 0.143 M Tris-HCl, pH 6.8, 0.115% (w/v) SDS and were prepared by mixing 4.4 ml of H_2O, 1.67 ml of 30% (w/v) acrylamide, 1.3 ml 1% (w/v) bisacrylamide, 1.25 ml of 1 M Tris-HCl, pH 6.8, 50 μl of 20% (w/v) SDS, 50 μl of a freshly prepared solution of 10% (w/v) ammonium persulfate and 5 μl of *N,N,N',N'*-tetramethylethylenediamine.

Electrophoresis was performed at 50 - 75 V (constant voltage) and at room temperature until the bromophenol blue tracker-dye migrated to within 2 cm of the anode end of the gel. Following electrophoresis, gels were fixed and stained simultaneously by shaking for 30 min in 1.25% (w/v) Coomassie Brilliant Blue R-250, 50% (v/v) methanol, 7% (v/v) acetic acid. Background staining was removed by shaking the gels in multiple changes of 30% (w/v) methanol, 10% (v/v) acetic acid.

Protein	**Subunit M_r**
Bovine serum albumin	66,000
Ovalbumin	45,000
Glyceraldehyde-3-phosphate dehydrogenase	36,000
Carbonic anhydrase	29,000
α-Lactalbumin	14,000

Table 4.2.7 Sigma SDS Apparent M_r Markers

4.2.18 N-Terminal Amino-Acid Sequence Analysis of Recombinant Protein

E. coli cell lysates that had been subjected to SDS-PAGE were transferred onto ProBlott (Applied Biosystems PVDF membrane) using the "wet blotting" technique and with 10 mM CAPS, 10% (v/v) methanol, pH 11, as the transfer buffer (as previously described for electroblotting onto Immobilon (Section 2.1.5). Transfer was achieved by applying a constant current of 0.7 A for 45 min. Protein transferred onto ProBlott was visualised by staining (1 min) with 0.1% (w/v) Amido black in 40% (v/v) methanol, 1% (v/v) acetic acid. Background staining was then removed by soaking the ProBlott in multiple changes of water. The blot was left to dry at room temperature, "bands" of interest were excised from the dry blot and subjected to N-terminal amino-acid sequence analysis (Section 2.1.6).

4.3 RESULTS

4.3.1 Restriction Analyses of Plasmids pADH2 and pET21a and Isolation of Relevant Restriction Fragments

Restriction of plasmid pADH2 with *Sac*I alone yielded two *Sac*I-*Sac*I fragments; the 1.3-kb fragment carrying the entire *adh* gene and the 2.8-kb fragment corresponding approximately to the pUC18 vector that had been used in the construction of pADH2 (Figure 4.3.1; Lane 2). Restriction of pADH2 with *Hpa*I alone yielded a single *Hpa*I-*Hpa*I fragment (~4.1 kb long) corresponding to the linear plasmid (Figure 4.3.1; Lane 3). When the plasmid was doubly restricted with *Sac*I and *Hpa*I, the 1.3-kb *Sac*I-*Sac*I fragment yielded two *Hpa*I-*Sac*I fragments, a 1-kb fragment corresponding to *adh*2334 minus the 5' end (Figure 4.3.1; Lane 4) and a 0.260-kb fragment, which is not visible in Figure 4.3.1. Plasmid pADH2 DNA was restricted with *Sac*I and *Hpa*I simultaneously, the restriction fragments were separated by electrophoresis through agarose and the 1-kb *Hpa*I-*Sac*I *adh* fragment was excised from the gel and purified from the agarose.

Restriction of the expression vector pET21a with *Sac*I, with *Nde*I, and with *Sac*I and *Nde*I together, resulted in linear plasmid DNA (5.4 kb long; Figure 4.3.2). The linear *Nde*I-*Sac*I pET21a DNA was purified from agarose after agarose gel-electrophoresis of the restriction digest mixture.

4.3.2 Monitoring Construction of the Double-Stranded *Nde*I-*Hpa*I Adapter

The double-stranded *Nde*I-*Hpa*I adapter was constructed by annealing the two single-stranded synthetic oligonucleotides A_s (nineteen nucleotides) and A_a (seventeen nucleotides). The sizes of the two single-stranded oligonucleotides were confirmed by comparing their mobilities through polyacrylamide (denaturing conditions) with that of DNA sequencing primers of known length. A_s comigrated with a nineteen-nucleotide sequencing primer and A_a comigrated with a seventeen-nucleotide sequencing primer [Figure 4.3.3(a)]. When the annealing reaction mixture was analysed by electrophoresis through polyacrylamide (nondenaturing conditions), only one band was observed [Figure 4.3.3(b); Lane 1], and that band had an electrophoretic mobility that was different from that of either of the single-stranded oligonucleotides [Figure 4.3.3(b)].

4.3.3 Ligation of the *Nde*I-*Hpa*I Adapter to the 1-kb *Hpa*I-*Sac*I *adh* Fragment, Cloning of the Construct into *Nde*I-*Sac*I pET21a

The *Nde*I-*Hpa*I adapter was ligated to the purified 1-kb *Hpa*I-*Sac*I *adh* fragment. The ligation reaction mixture was restricted with *Nde*I, to remove all but

one adapter molecule, and with *Sac*I to eliminate possible *adh* dimers. Oligonucleotide fragments were removed from the reaction mixture by centrifugation in a Centricon and the enzymes were removed by phenol extraction. The 1-kb *Nde*I-*Sac*I construct was then ligated into dephosphorylated 5.4-kb *Nde*I-*Sac*I pET21a. When the ligation reaction mixture was transformed into competent *E. coli* TG1, and the transformants were grown on LB plus ampicillin agar, six colonies per ng of vector DNA were obtained. A dephosphorylated *Nde*I-*Sac*I pET21a plus ligase control yielded two colonies per ng of vector DNA. Colonies (200 in total) were transferred to nitrocellulose and screened for ADH antigenic activity. Fifty-three colonies cross-reacted with anti-ADH 2334 antiserum (ADH-positive colonies). Plasmid DNA was purified from eighteen ADH-positive colonies, but no plasmid of the correct size (6.4 kb) was found and no plasmid contained a 1-kb *Nde*I-*Sac*I insert. All plasmids gave three fragments, 2.4, 1.3 and 0.45 kb, when restricted with *Nde*I and *Sac*I simultaneously, rather than the expected 5.4-kb and 1-kb fragments. Plasmid DNA was also purified from two colonies that had not cross-reacted with antibody (ADH-negative colonies). That DNA yielded two fragments, 2.4 and 0.45 kb, when restricted with *Nde*I and *Sac*I simultaneously, which indicated that the antigenic activity was due to expression from the 1.3-kb fragment. A more detailed analysis of one of the plasmid DNA preparations from an ADH-positive colony (Figure 4.3.4) revealed a 4-kb *Nde*I-*Nde*I fragment (Lane 2), two *Sac*I-*Sac*I fragments (1.3 and 2.8 kb; Lane 3), two *Hpa*I-*Sac*I fragments (0.3 kb and 1 kb long; Lane 4) and two *Hpa*I-*Nde*I fragments (1.5 and 2.4 kb long; Lane 5).

4.3.4 Ligation of the 1-kb *Hpa*I-*Sac*I *adh* Fragment to Linear *Nde*I-*Sac*I pET21a, Formation of a 6.4-kb *adh*-pET21a Linear Construct

The 1-kb *Hpa*I-*Sac*I *adh* fragment was ligated to the *Sac*I end of the 5.4-kb, linear, *Nde*I-*Sac*I pET21a to create a 6.4-kb *adh*-pET21a linear construct. The major products of the ligation were constructs with electrophoretic mobilities corresponding to linear DNA of approximately 2, 6.4 and 10 kb (Figure 4.3.5; Lane 3). A control ligation of the 1-kb *Hpa*I-*Sac*I *adh* fragment indicated that the ligation product of 2 kb resulted from dimerisation of the 1-kb *adh* fragment (Figure 4.3.5; Lane 6). A control ligation of the linear *Nde*I-*Sac*I pET21a indicated that the ligation product of ~ 10 kb resulted from dimerisation of the linear pET21a DNA (Figure 4.3.5; Lane 5). The 6.4-kb construct was assumed to be a linear *Hpa*I-*Nde*I construct, resulting from ligation of the 1-kb *adh* fragment to the *Sac*I end of the linear pET21a DNA. The construct was isolated by excising the 6.4-kb 'band' from an agarose gel.

4.3.5 Ligation of the *Nde*I-*Hpa*I Adapter to the 6.4-kb *adh*-pET21a Linear Construct

The *Nde*I-*Hpa*I adapter was ligated to the purified 6.4-kb *Hpa*I-*Nde*I *adh*-pET21a linear construct and the ligation reaction mixture was transformed into *E. coli* TG1. Two ampicillin-resistant colonies per ng of linear construct DNA were obtained. Plasmid DNA was isolated from twenty transformants and restricted with *Nde*I. Each plasmid preparation yielded a single restriction fragment, either 5.4 or 3 kb long.

In a second procedure, dephosphorylated *Nde*I-*Sac*I pET21a, the 1-kb *Hpa*I-*Sac*I *adh* fragment and the *Nde*I-*Hpa*I adapter were simply mixed together and incubated in the presence of ligase for 16 h. In a third procedure, dephosphorylated *Nde*I-*Sac*I pET21a and the 1-kb *Hpa*I-*Sac*I *adh* fragment were incubated in the presence of ligase for 7 h before the addition of adapter (and additional ligase), and continuation of the incubation for a further 14 h. Both procedures resulted in approximately the same yield of ampicillin-resistant colonies (six colonies per ng of pET21a DNA). Arbitrarily selected colonies (fifty from each procedure) were transferred to nitrocellulose and screened for ADH antigenic activity. Nine colonies from the procedure where all components of the ligation were simply mixed together cross-reacted with anti-ADH 2334 antibody. One colony from the procedure where *Nde*I-*Sac*I pET21a and the *Hpa*I-*Sac*I *adh* fragment were preincubated, cross-reacted with antibody. Plasmid DNA was isolated from the ten ADH-positive colonies and, as before, no plasmid of the correct size was found. The ten plasmids were identical (a single 4-kb *Hpa*I-*Hpa*I fragment, a single 4-kb *Nde*1-*Nde*I fragment and two *Sac*I-*Sac*I fragments (2.8-kb and 1.3-kb) and were apparently identical to the constructs described in Section 4.3.3. Plasmid DNA isolated from an ADH-negative colony gave a single 5.4-kb fragment when restricted with *Nde*1.

4.3.6 Insertion of *Nde*I-*Hpa*I Adapter into pADH2, Construction of pADH2a

Plasmid pADH2 was restricted with *Hpa*I, the linear 4.1-kb plasmid DNA was isolated from an agarose gel and was dephosphorylated and ligated to the *Nde*I-*Hpa*I adapter. The ligation reaction mixture was transformed into *E. coli* TG1, and the plasmid DNA isolated from nine of the resulting ampicillin-resistant transformants was screened for the presence of a 1-kb *Nde*I-*Sac*I fragment that is diagnostic for the incorporation of at least two adapter molecules into the linear *Hpa*I-*Hpa*I pADH2. One plasmid containing a 1-kb *Nde*I-*Sac*I fragment was identified. This plasmid was designated pADH2a and a restriction analysis and plasmid map are presented in Figure 4.3.6.

4.3.7 Cloning of the 1-kb *Nde*I-*Sac*I Fragment of pADH2a into pET21a

The 1-kb *Nde*I-*Sac*I fragment of pADH2a was isolated, after electrophoresis of a large-scale digest through agarose, and ligated into pET21a that had been restricted with *Nde*I and *Sac*I, isolated from agarose and dephosphorylated. Aliquots of the ligation reaction mixture were transformed into *E. coli* TG1, DH5α and SURE cells. The yields of transformants were (approximately) forty-two TG1 colonies per ng of DNA, eleven DH5α colonies per ng of DNA and seventy-eight SURE colonies per ng of DNA. The background, due to self-ligation of the vector, was measured by transforming dephosphorylated *Nde*I-*Sac*I pET21a DNA plus ligase into each strain and was approximately two colonies per ng of DNA for each strain. Plasmid DNA was isolated from forty ampicillin-resistant transformants (twenty TG1 colonies, ten DH5α colonies and ten SURE colonies) and was screened for the presence of a 1-kb *Nde*I-*Sac*I insert. No plasmid was found that yielded the correct pattern of 1-kb plus 5.4-kb fragments (insert plus plasmid backbone). The typical restriction fragment pattern obtained when plasmids isolated from DH5α and TG1 transformants were restricted with *Nde*I and *Sac*I (simultaneously) were 2.4, 1.3 and 0.45 kb. The *Nde*I plus *Sac*I restriction analyses of the plasmids isolated from *E. coli* DH5α are presented in Figure 4.3.7. The plasmids isolated from SURE transformants were restricted with *Sac*I alone and gave a fragment pattern of either 2.8 kb plus 1.3 kb, or 2.8 kb alone. Plasmid DNA isolated from a colony that resulted from transformation of the dephosphorylated *Nde*I-*Sac*I pET21a DNA plus ligase into TG1 was analysed after restriction with *Sac*I and was found to give the 2.8 kb plus 1.3 kb pattern.

4.3.8 Subcloning of an *adh* PCR Product into pET21d and into pTrc99A

PCR was used to introduce a *Bsp*HI site at the 5' end of *adh*2334 and to introduce an *Xho*I site downstream from the *adh*2334 stop codon. An agarose gel-electrophoresis analysis of the *adh* PCR product is presented in Figure 4.3.8. The PCR product was approximately 1 kb in length, and approximately 7 μg were obtained from 1 ng of pADH2 template.

Restriction of pET21d with *Xho*I alone, and with *Xho*I and *Nco*I simultaneously, yielded the expected 5.4 kb linear DNA (Figure 4.3.9).

The restriction fragments obtained when pTrc99A was restricted with *Nco*I and *Sal*I simultaneously included a 3.6-kb and a 0.7-kb fragment, which were assumed to be the result of *Nco*I star activity. When pTrc99A was incubated with *Nco*I alone for 1 h, the expected 4.1-kb linear fragment, resulting from cleavage at the single *Nco*I site, was observed (Figure 4.3.10; Lane 2). The star activity was

avoided by restricting pTrc99A with *Nco*I alone for 1 h, removing the *Nco*I by extraction into phenol-chloroform-isoamyl alcohol and then adding *Sal*I (Figure 4.3.10).

The *adh* PCR product (5 μg) was restricted with *Bsp*HI and with *Xho*I simultaneously for 1 h. The restriction enzymes were then removed from solution by extraction into phenol-chloroform-isoamyl alcohol and the restricted DNA was ligated into *Nco*I-*Xho*I pET21d. Aliquots of the ligation reaction mixture were transformed into *E. coli* strains TG1 and SURE. In a parallel experiment, the *Bsp*HI-*Xho*I *adh* PCR product was ligated into *Nco*I-*Sal*I pTrc99A, and aliquots of that ligation reaction mixture were also transformed into *E. coli* strains TG1 and SURE. There were thus four different clones possible: the construct pET21d plus *adh* in *E. coli* TG1 and the construct pET21d plus *adh* in *E. coli* SURE (pETADH TG1 and pETADH SURE); the construct pTrc99A plus *adh* in *E. coli* TG1 and the construct pTrc99A plus *adh* in *E. coli* SURE (pTrcADH TG1 and pTrcADH SURE).

The *E. coli* TG1 transformants yielded eight ampicillin-resistant colonies per ng of vector DNA with both the pET21d construct and the pTrc99A construct. The *E. coli* SURE transformants yielded approximately fifty ampicillin-resistant colonies per ng of vector DNA with both constructs. Plasmid DNA was purified from small-scale cultures of five of each of the four different types of colony. The yield of plasmid DNA from the five cultures of pTrcADH SURE clones was too low to work with and these clones were not investigated further.

Plasmid DNA isolated from the five pETADH SURE clones and from four of the five pETADH TG1 clones yielded a single DNA fragment of 6.4 kb when restricted with *Xho*I (Figure 4.3.11). Plasmid DNA purified from the pTrcADH TG1 clones was analysed by restriction with *Nco*I. The pTrc99A *Nco*I cloning site is not recreated upon ligation to a *Bsp*HI site, however, the *adh* gene contains two *Nco*I sites (Section 4.1; Figure 4.1.4). Two *Nco*I-*Nco*I fragments (4.9 kb and 0.122 kb) were therefore diagnostic of the correct construct. A fragment of approximately 5 kb was visible in the five digests (Figure 4.3.12).

4.3.9 IPTG-Dependent Expression in pETADH and in pTrcADH Clones

Aliquots of the nine 6.4-kb pETADH plasmid DNA preparations were transformed into the expression host *E. coli* BL21 and small-scale cultures of the transformants were screened for IPTG-dependent expression (Figure 4.3.13). In the five cultures of clones that had been isolated from *E. coli* SURE [Figure 4.3.13(a)] expression was unequivocal. In four of the cultures, the product of expression was a protein with an apparent subunit M_r that was correct for rADH (approximately 36,000). Expression of a 36,000-M_r protein was also unequivocal in three of the four

cultures of clones isolated from TG1, however, one of these was a doublet (Figure 4.3.13(b); Clone 15). *E. coli* BL21 cells transformed with pET21d (no insert) did not exhibit IPTG-dependent expression (Figure 4.3.13(b); C).

Small-scale cultures of three of the pTrcADH TG1 clones were screened for IPTG-dependent expression and, while IPTG-dependent expression did occur in the three cultures, it did not result in the production of a protein with an apparent subunit M_r of 36,000. The product of expression was a protein with an apparent subunit M_r of 30,000 in one of the cultures and a protein with an apparent subunit M_r of 50,000 in the other two cultures (Figure 4.3.14).

4.3.10 Confirmation that the Product of Expression was Recombinant ADH 2334

One of the pETADH constructs isolated from *E. coli* SURE (Figure 4.3.13(a); Clone 1) was selected for further characterisation and was designated plasmid pETADH. A cell lysate of *E. coli* pLysS cells, transformed with pETADH and grown in the presence of IPTG for 6 h, was subjected to SDS-PAGE and electroblotted onto the protein sequencing membrane Problott. The band corresponding to a subunit M_r of 36,000 was excised from the blot and submitted to N-terminal amino-acid sequencing. The twenty-two-residue amino-acid sequence obtained was: M K A A V V N E F X X A L E I X E V E X P A, where X corresponds to 'unknown'. The sequence was homogeneous, the data were unambiguous at those positions where assignments were made and, allowing for the 'unknowns', corresponded exactly to the N-terminus of ADH 2334.

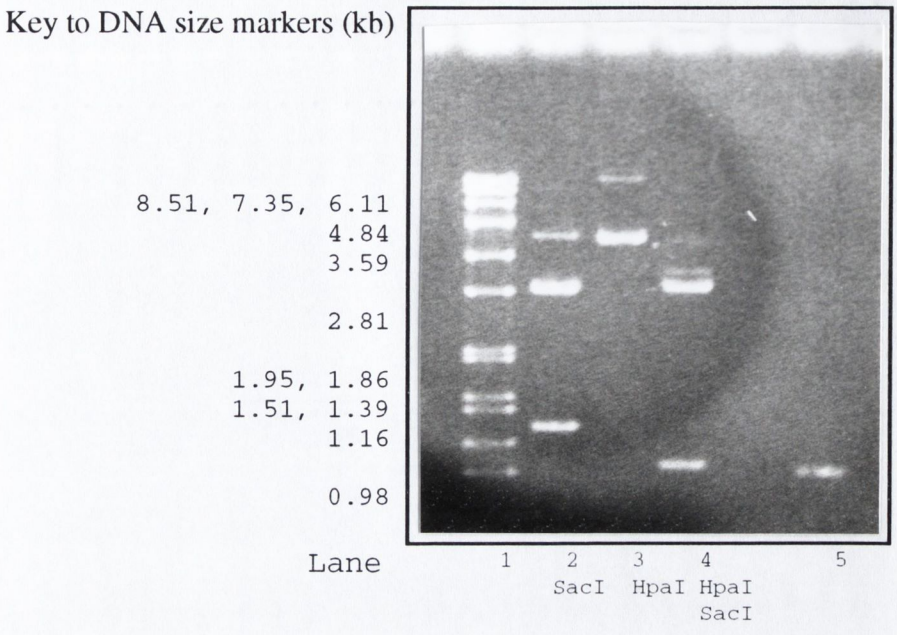


Figure 4.3.1 **Restriction of pADH2 with *Sac*I and *Hpa*I, Isolation of the 1-kb *Hpa*I-*Sac*I *adh* Fragment**

Lane 1 contained DNA size markers: phage SPPI DNA (0.5 μg) restricted with *Eco*RI. Lane 2 contained plasmid pADH2 (4.1 kb; Section 4.1.3; Figure 4.1.4) that was restricted with *Sac*I. Lane 3 contained pADH2 that was restricted with *Hpa*I. Lane 4 contained pADH2 that was restricted with *Sac*I and *Hpa*I simultaneously. Lane 5 contained a sample [10% (v/v) of the final yield from 2 μg of pADH2 DNA] of the purified 1-kb *Hpa*I-*Sac*I fragment.

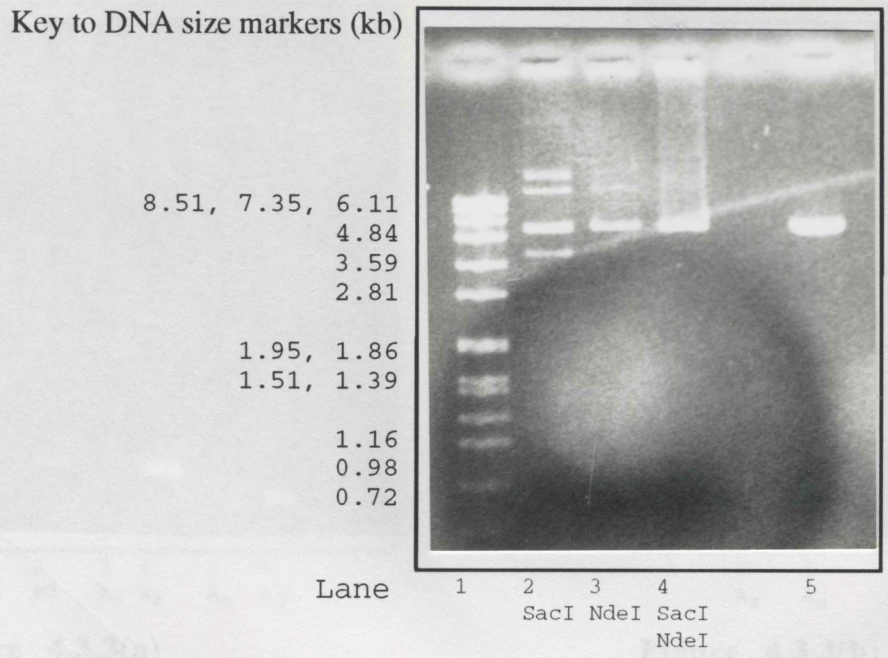


Figure 4.3.2 **Restriction of pET21a with *Sac*I and *Nde*I, Isolation of the Linear *Nde*I-*Sac*I pET21a**

Lane 1 contained DNA size markers: phage SPPI DNA (0.5 μg) restricted with *Eco*RI. Lane 2 contained plasmid pET21a (5.4 kb; Section 4.1.1.1; Figure 4.1.1) that was restricted with *Sac*I. Lane 3 contained pET21a that was restricted with *Nde*I. Lane 4 contained pET21a that was restricted with *Sac*I and *Nde*I simultaneously. Lane 5 contained a sample [10% (v/v) of the final yield from 1 μg of pET21a DNA] of the purified doubly restricted pET21a DNA.

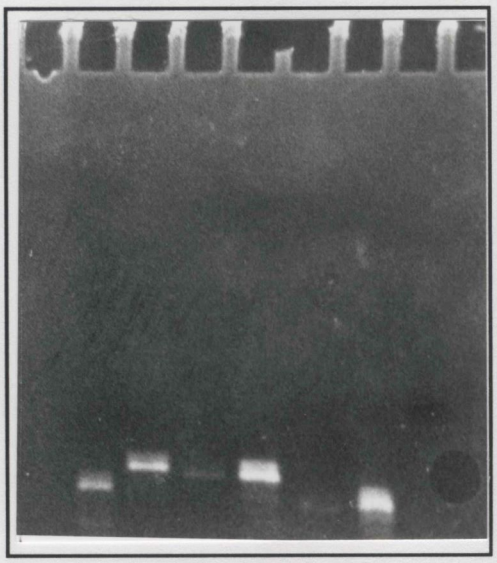


Figure 4.3.3(a)

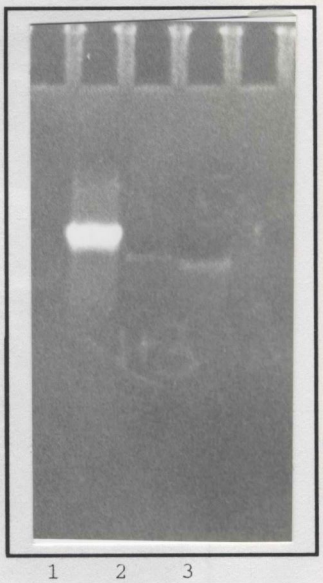


Figure 4.3.3(b)

Figure 4.3.3(a) **Denaturing Polyacrylamide-Gel Electrophoresis of Single-Stranded Oligonucleotides**

The DNA sequencing primers P4 (seventeen bases; 320 ng; Lane 1) and P7 (nineteen bases; 544 ng; Lane 2) were those described in Chapter 3 (Section 3.2.19.1; Figure 3.3.7). Lanes 3 and 4 contained the single-stranded oligonucleotide A_s (1 and 5 µl of a ten-fold-dilution of the stock solution; Section 4.2.4). Lanes 5 and 6 contained the single-stranded oligonucleotide A_a (1 and 5 µl of a ten-fold-dilution of the stock solution). Electrophoresis was through 19% (w/v) polyacrylamide (denaturing conditions; Section 4.2.5).

Figure 4.3.3(b) **Nondenaturing Polyacrylamide-Gel Electrophoresis of the Double-Stranded *Nde*I-*Hpa*I Adapter**

Lane 1 contained an aliquot (equivalent to 664 ng) of the double-stranded *Nde*I-*Hpa*I adapter. Lane 2 contained the single-stranded oligonucleotide A_s (2 µl of a ten-fold dilution of the stock solution). Lane 3 contained A_a (2 µl of a ten-fold dilution of the stock solution) Electrophoresis was through 19% (w/v) polyacrylamide (nondenaturing conditions; Section 4.2.5).

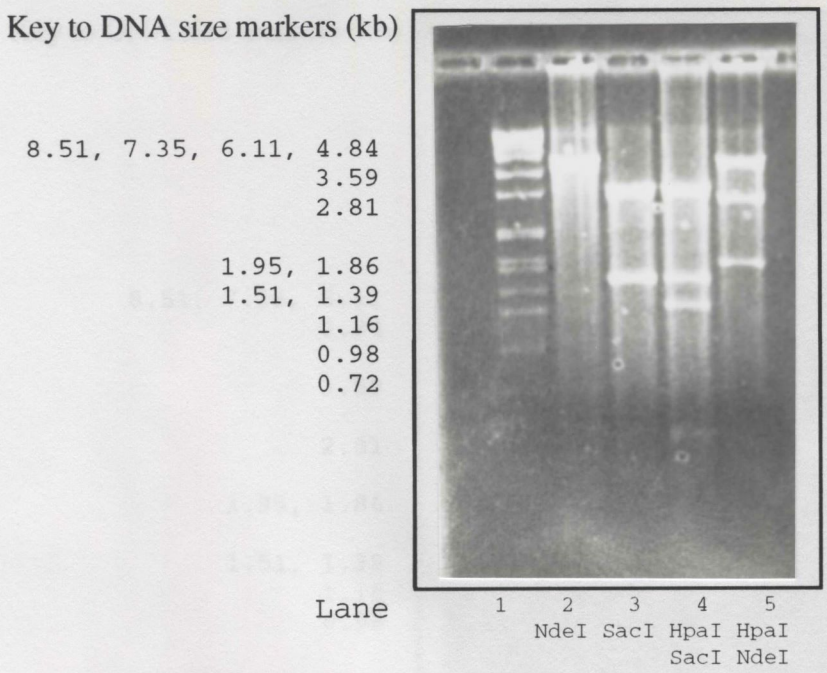


Figure 4.3.4 **Restriction Analyses of Plasmid DNA Purified from an *E. coli* Transformant that had Stained Positive for ADH Antigenic Activity**

The restriction analyses are those of one of eighteen colonies selected on the basis of cross-reactivity with anti-ADH2334 antibody. Lane 1 contained DNA size markers: phage SPPI DNA (0.5 μg) restricted with *Eco*RI. Lane 2 contained plasmid DNA restricted with *Nde*I alone. Lane 3 contained DNA restricted with *Sac*I alone. Lane 4 contained DNA restricted with *Hpa*I and *Sac*I simultaneously, and Lane 5 contained DNA restricted with *Hpa*I and *Nde*I simultaneously.

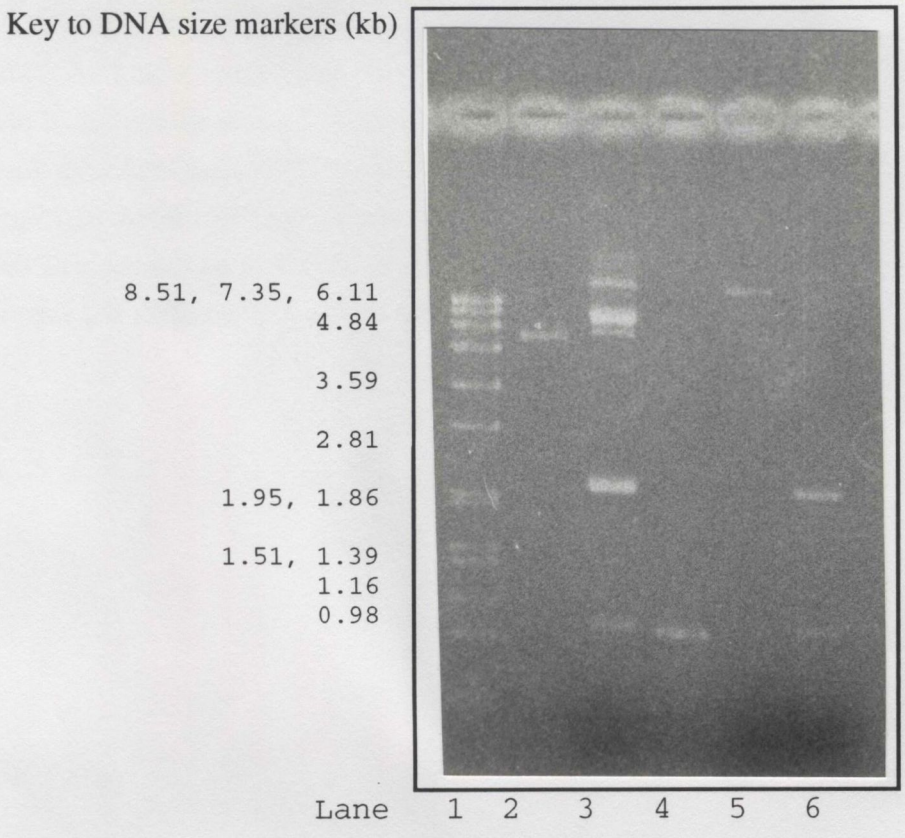


Figure 4.3.5 **Construction of a 6.4-kb *adh*-pET21a Linear Construct**

Linear *Nde*I-*Sac*I pET21a (5.4 kb; Lane 2) and the 1-kb *Hpa*I-*Sac*I *adh* fragment of pADH2 (Lane 4) were ligated. The total ligation reaction mixture was electrophoresed through agarose (Lane 3). A pET21a control ligation reaction mixture (Lane 5) contained 300 ng of the linear *Nde*I-*Sac*I pET21a DNA and 0.3 Weiss units of ligase. An *adh* control ligation reaction mixture (Lane 6) contained 58.6 ng of the 1-kb *Hpa*I-*Sac*I *adh* fragment and 0.3 Weiss units of ligase. Lane 1 contained DNA size markers: phage SPPI DNA (1 μg) restricted with *Eco*RI.

Figure 4.3.6 **Restriction Analyses and Plasmid Map of pADH2a**

The photograph (a) is a composite of two gels. Lane 2 contained plasmid DNA restricted (incompletely) with *Sac*I. Lane 3 contained plasmid DNA restricted with *Nde*I and *Sac*I. Lane 5 contained a second aliquot of plasmid DNA restricted with *Sac*I. Lane 6 contained plasmid DNA restricted with *Sac*I and *Hpa*I. Lanes 1 and 4 contained DNA size markers: phage SPPI DNA (0.75 μg) restricted with *Eco*RI. The plasmid map in (b) was constructed on the basis of the restriction analyses presented in (a). The notations are as explained in the legend to Figure 3.1.1.

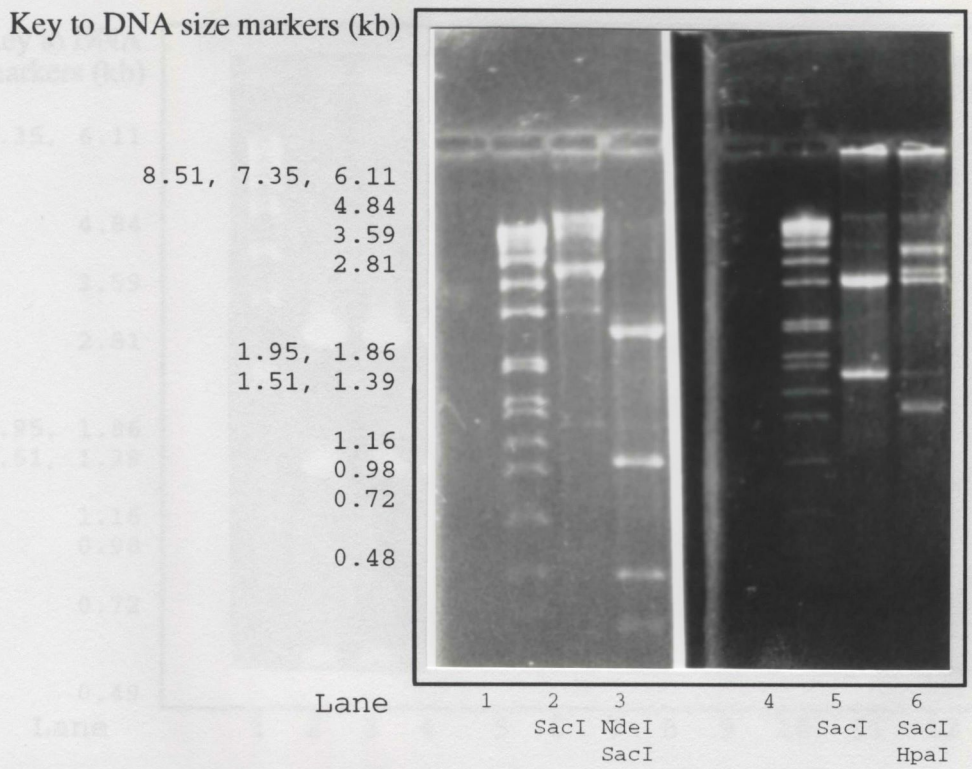


Figure 4.3.6(a)

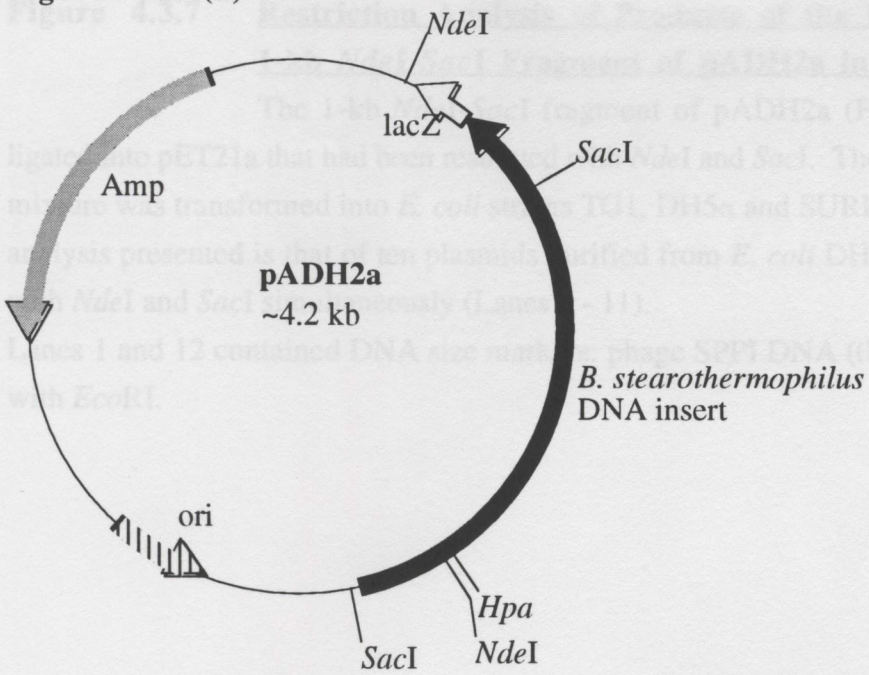


Figure 4.3.6(b)

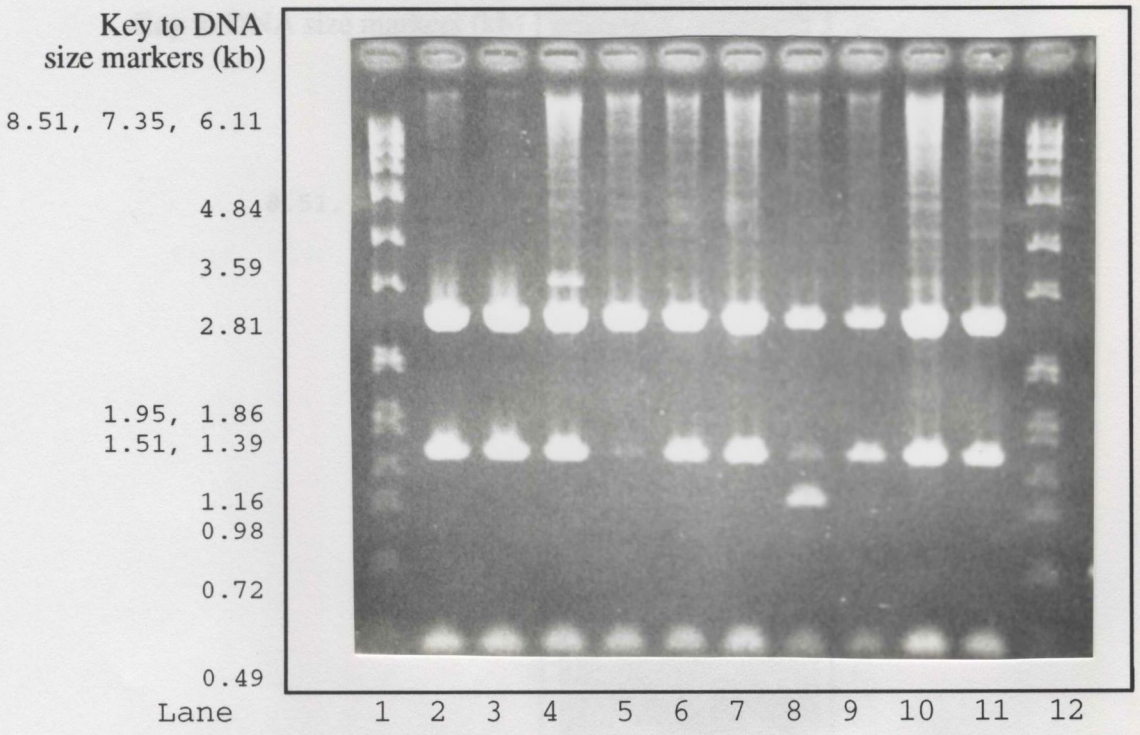


Figure 4.3.7 **Restriction Analysis of Products of the Ligation of the 1-kb *Nde*I-*Sac*I Fragment of pADH2a into pET21a**

The 1-kb *Nde*I-*Sac*I fragment of pADH2a (Figure 4.3.6) was ligated into pET21a that had been restricted with *Nde*I and *Sac*I. The ligation reaction mixture was transformed into *E. coli* strains TG1, DH5α and SURE. The restriction analysis presented is that of ten plasmids purified from *E. coli* DH5α, and restricted with *Nde*I and *Sac*I simultaneously (Lanes 2 - 11).

Lanes 1 and 12 contained DNA size markers: phage SPPI DNA (0.75 μg) restricted with *Eco*RI.

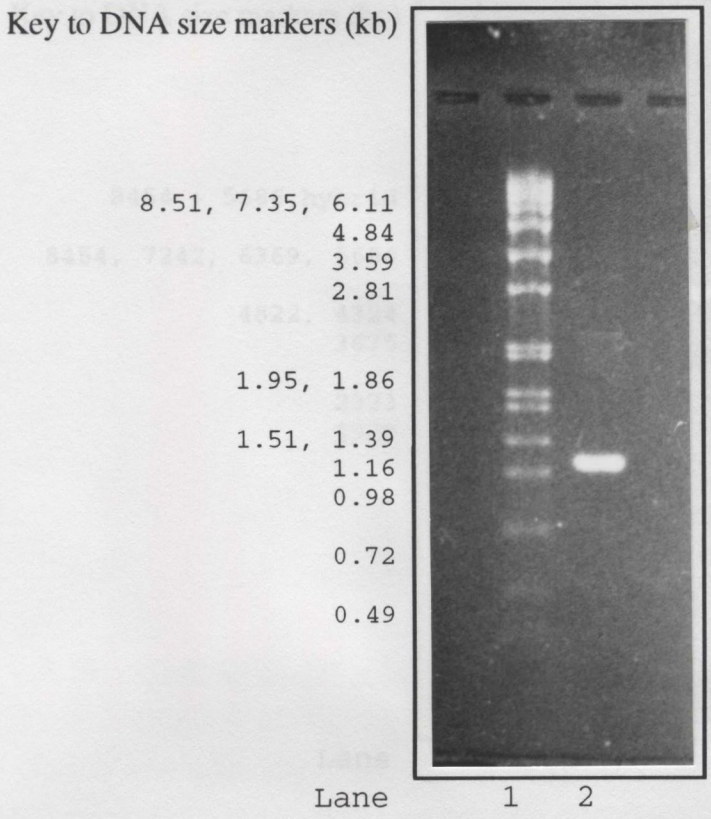


Figure 4.3.8 **Analysis of the *adh* PCR Product**

PCR with the primers PX_a and PB_s (100 pmoles of each) and plasmid pADH2 (1 ng) as the template DNA was described in Section 4.2.10. Lane 1 contained DNA size markers: phage SPPI DNA (0.75 µg) restricted with *Eco*RI. Lane 2 contained the *adh* PCR product: an aliquot [5.7% (v/v)] of the material recovered after the final centrifugation in a Centricon concentrator.

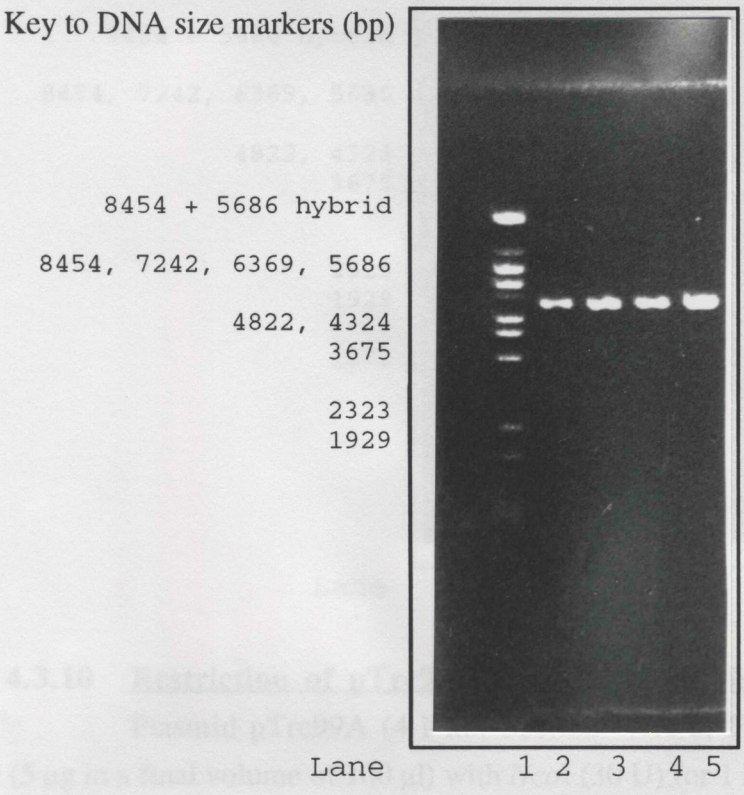


Figure 4.3.9 **Restriction of pET21d with *Xho*I and *Nco*I**

Plasmid pET21d (5.4 kb, identical to pET21a, except that the *Nde*I site is replaced by *Nco*I) was restricted (5 μg in a final volume of 100 μl) with *Xho*I for 1 h (Lane 2; 2 μl of the total). *Nco*I was added and the restriction continued for a further 1 h (Lane 3; 2 μl of the total). The doubly restricted DNA was dephosphorylated (Lane 4; 2 μl of the *Xho*I-*Nco*I pET21d dephosphorylation reaction mixture) and recovered by precipitation in the presence of NaCl and ethanol [Lane 5; 8% (v/v) of the material recovered]. Lane 1 contained DNA size markers: phage λ DNA (0.5 μg) restricted with *Bst*EII.

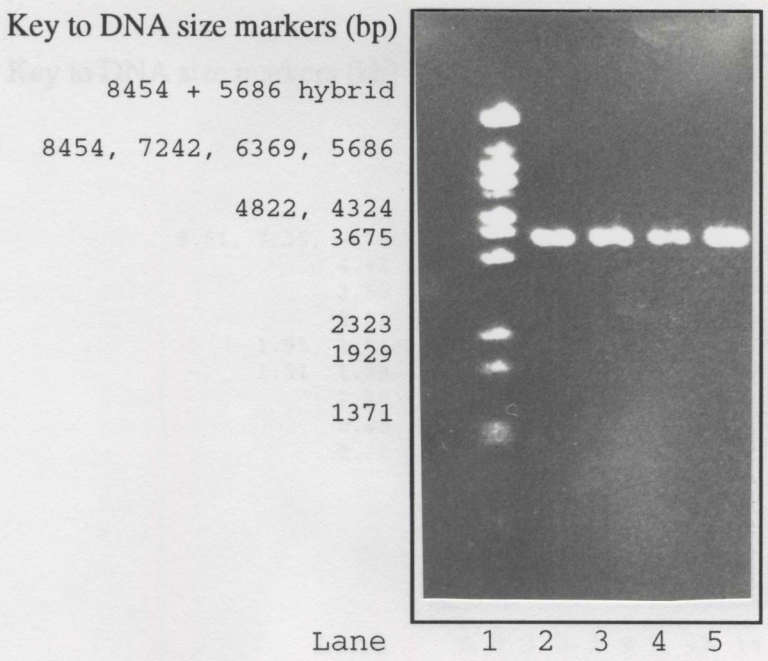


Figure 4.3.10 **Restriction of pTrc99A with *Nco*I and *Sal*I**

Plasmid pTrc99A (4.1 kb; Section 4.1.1.2; Figure 4.1.2) was restricted (5 μg in a final volume of 100 μl) with *Nco*I (30 U) for 1 h (Lane 2; 2 μl of the total). The enzyme was removed by extraction into phenol-chloroform-isoamyl alcohol and the restricted DNA was recovered by precipitation in the presence of NaCl and ethanol. The final DNA pellet was resuspended in 87 μl of H_2O, 10 μl of ten-times concentrated Boehringer Buffer H and 30 U of *Sal*I were added, and the mixture was incubated for 1 h at 37 °C (Lane 3; 4 μl of the total). The doubly restricted DNA was dephosphorylated (Lane 4; 2 μl of the total) and recovered by precipitation in the presence of NaCl and ethanol [Lane 5; 8% (v/v) of the material recovered]. Lane 1 contained DNA size markers: phage λ DNA (0.5 μg) restricted with *Bst*EII.

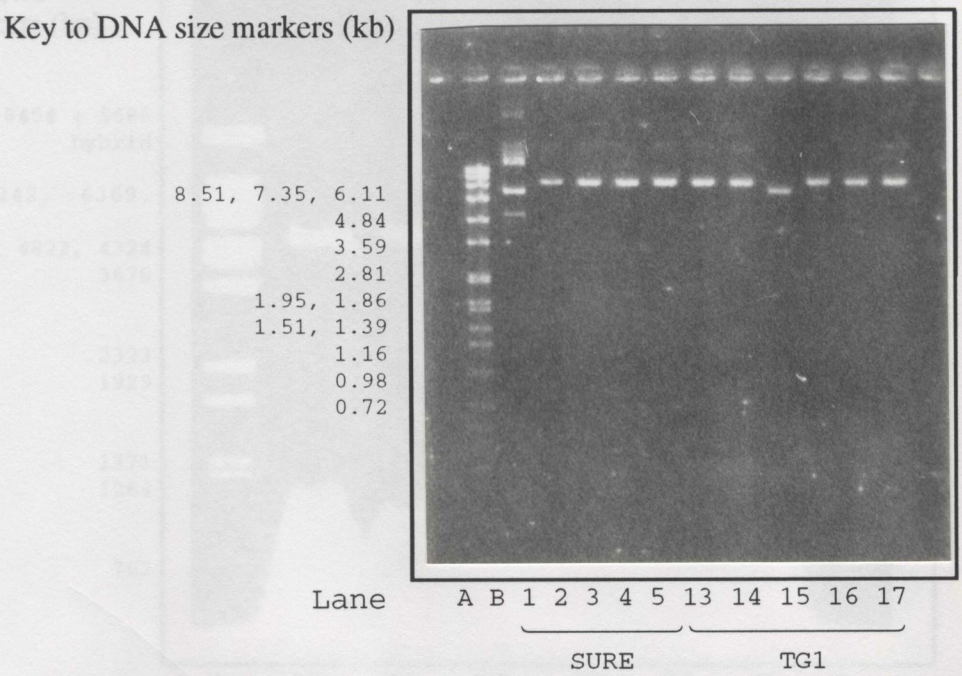


Figure 4.3.11 **Restriction Analysis of the Products of the Ligation of the *Bsp*HI-*Xho*I *adh* PCR Product into pET21d**

Lane A contained DNA size markers: phage SPP1 DNA (0.5 µg) restricted with *Eco*RI. Lane B contained the vector pET21d DNA restricted with *Xho*I. Lanes 1 - 5 contained the plasmid constructs purified from *E. coli* SURE and restricted with *Xho*I.

Lanes 13 - 17 contained the plasmid constructs purified from *E. coli* TG1 and restricted *Xho*I.

The numbers 1 - 5 and 13 - 17 were those originally assigned to colonies of transformed bacteria and subsequently to the plasmid DNA isolated from those ten colonies. The same numbering system is used in the text and in all subsequent Figures.

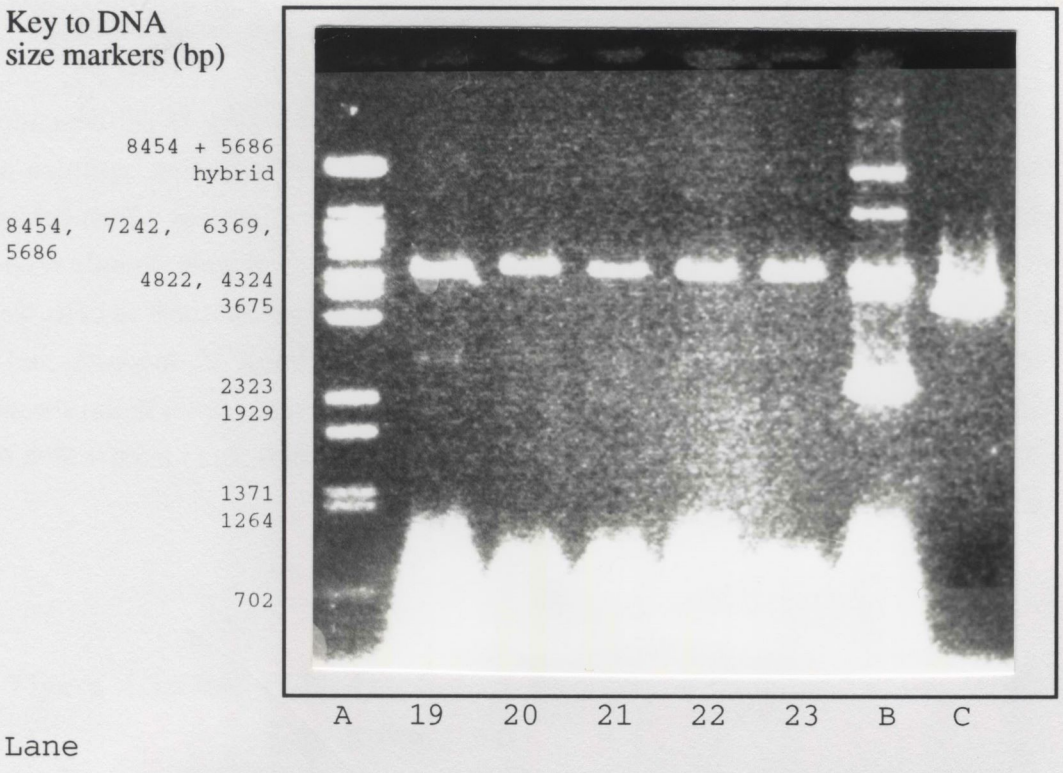


Figure 4.3.12 **Restriction Analysis of the Products of the Ligation of the *Bsp*HI-*Xho*I *adh* PCR Product into pTrc99A**

Lane A contained DNA size markers: phage λ DNA (0.5 μg) restricted with *Bst*EII. Lane B contained a *Nco*I restriction of DNA purified from a colony obtained from transformation of a dephosphorylation control ligation (no insert) into *E. coli* TG1. Lane C contained vector pTrc99A DNA (500 ng) restricted with 5 units of *Nco*I. Lanes 19 - 23 contained the five plasmid constructs restricted with *Nco*I. The smear across the bottom of the gel is RNA.

Figure 4.3.13 **IPTG-Dependent Expression from pETADH Constructs**

E. coli BL21 transformants were screened for IPTG-dependent expression exactly as described in Section 4.2.16. Pre-induction samples are indicated by -, post-induction samples by +. The final volumes in which the cell pellets from pre- and post-induction samples were resuspended were directly related to cell number, as explained in Section 4.2.16, and equivalent numbers of cells were therefore loaded in each lane of the gel. M indicates subunit M_r markers, and C indicates *E. coli* BL21 tranformed with pET21d (no insert) and grown in the absence and presence of IPTG. Numbers indicate the clones (Figure 4.3.11) from which the recombinant DNA was purified.

Key to subunit M_r markers

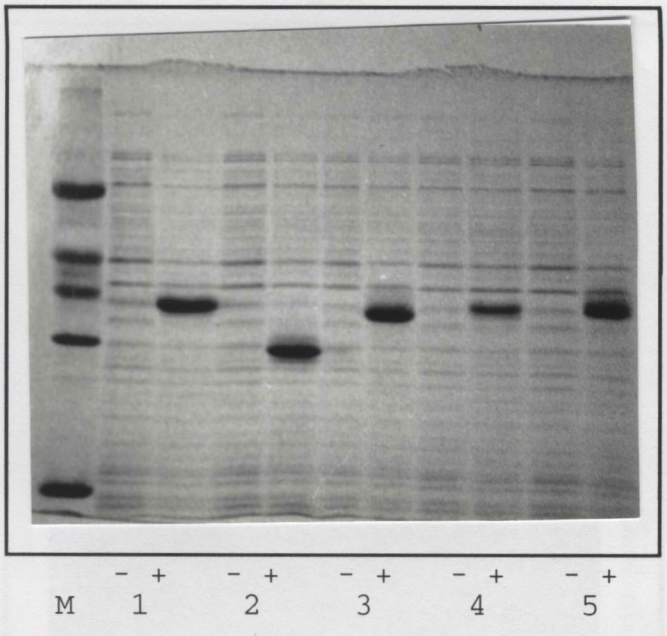


Figure 4.3.13(a) IPTG-Dependent Expression from pETADH SURE Constructs 1 - 5

Key to subunit M_r markers

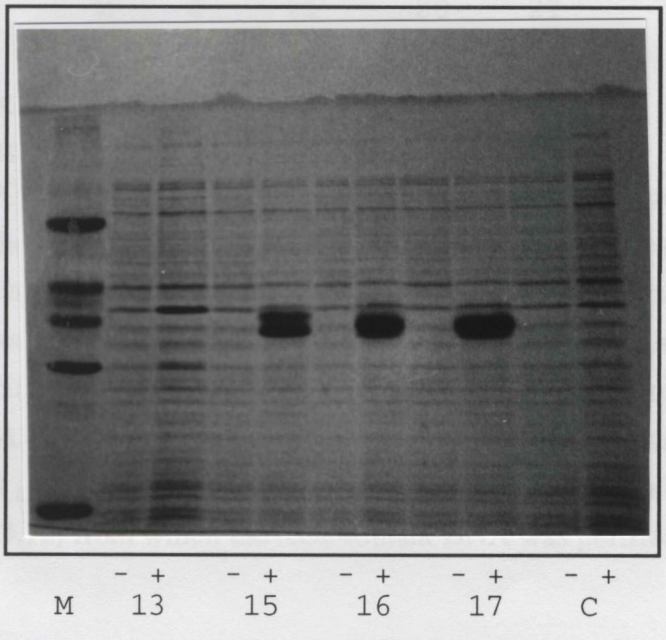


Figure 4.3.13(b) IPTG-Dependent Expression from pETADH TG1 Constructs 13, 15, 16 and 17

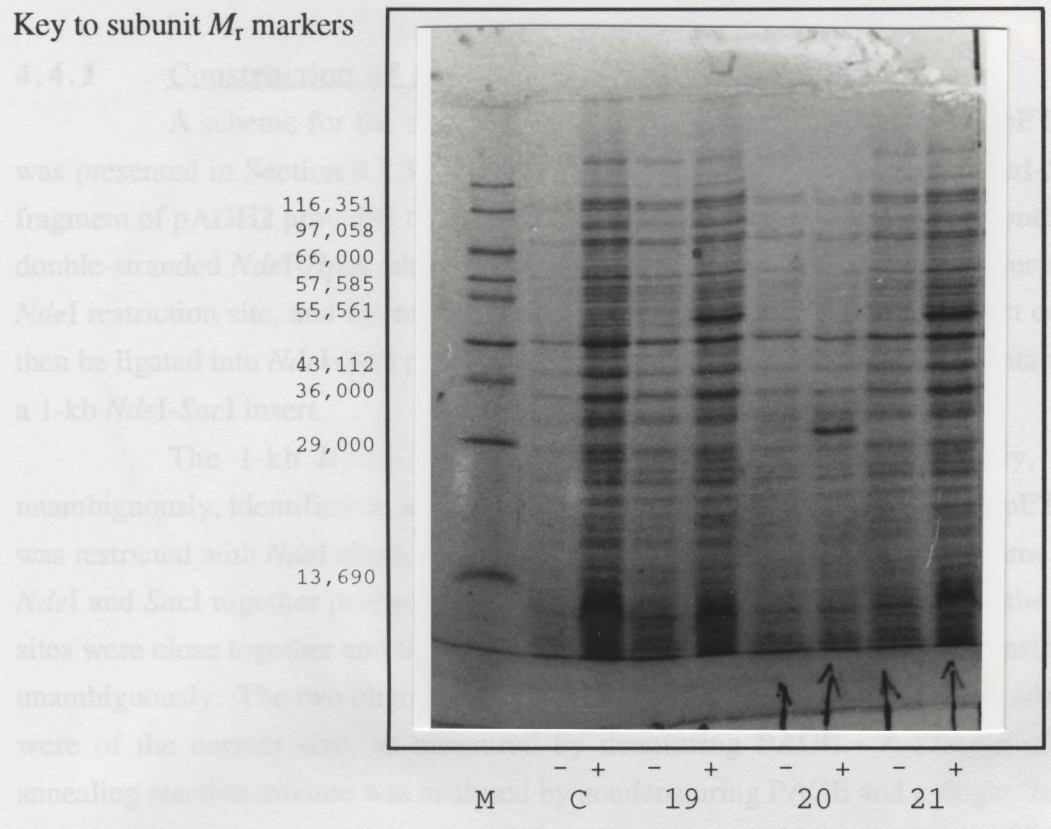


Figure 4.3.14 **IPTG-Dependent Expression from pTrcADH TG1 Constructs**

E. coli TG1 transformants were screened for IPTG-dependent expression as described in the legend to Figure 4.3.13. Pre-induction samples are indicated by -, post-induction samples by +. M indicates subunit M_r markers: B-galactosidase, phosphorylase B, bovine serum albumin, catalase, glutamate dehydrogenase, creatine kinase, glyceraldehyde-3-phosphate dehydrogenase, carbonic anhydrase, RNaseA, 5 µg of each. C indicates *E. coli* TG1 transformed with pTrc99A (no insert) and grown in the absence and presence of IPTG. Numbers indicate the clones (Figure 4.3.12) from which the recombinant DNA was purified.

4.4 DISCUSSION

4.4.1 Construction of an *adh* Expression System in pET21a

A scheme for the construction of an *adh* expression system in pET21a was presented in Section 4.1.3 (Figure 4.1.5). In that scheme, the 1-kb *Hpa*I-*Sac*I fragment of pADH2 provided the *adh*2334 coding sequence minus 18 bp, a synthetic double-stranded *Nde*I-*Hpa*I adapter restored the 18 bp, simultaneously introducing a *Nde*I restriction site, and the resulting linear *Nde*I-*Sac*I adapter-*adh* construct could then be ligated into *Nde*I-*Sac*I pET21a, to yield a 6.4-kb plasmid construct containing a 1-kb *Nde*I-*Sac*I insert.

The 1-kb *Hpa*I-*Sac*I *adh* fragment of pADH2 was easily, and unambiguously, identified on agarose gels. Linear DNA was obtained when pET21a was restricted with *Nde*I alone, or with *Sac*I alone, and it is reasonable to assume that *Nde*I and *Sac*I together produced linear *Nde*I-*Sac*I vector DNA. However, the two sites were close together and therefore, double restriction could not be demonstrated unambiguously. The two oligonucleotides used to create the double-stranded adapter were of the correct size, as measured by denaturing PAGE. A sample of the annealing reaction mixture was analysed by nondenaturing PAGE and a single "band" with an electrophoretic mobility that differed from that of either oligonucleotide was observed, indicating that the oligonucleotides did anneal.

The immunoblotting procedure provided a rapid method for screening large numbers of transformants and had been used successfully during the construction of pADH2 (Chapter 3). The method was used to screen transformants resulting from ligation of the *Nde*I-*Sac*I adapter-*adh* construct into *Nde*I-*Sac*I pET21a, and two distinct populations of transformants were identified; those that expressed ADH antigen, and those that did not. Plasmids isolated from ADH-positive colonies were 4 kb (or some multiple of 4 kb) in size, rather than 6.4 kb. Assuming a plasmid size of 4 kb, the ADH positive plasmids contained one *Nde*I site, two *Sac*I sites, one *Hpa*I site and a 1-kb *Hpa*I-*Sac*I fragment, but did not contain a 1-kb *Nde*I-*Sac*I fragment. Allowing for the approximate nature of fragment sizes, as determined by electrophoresis, the plasmid map one could construct for the ADH-positive plasmids (Figure 4.4.1) is strikingly similar to that of pADH2 (4158 bp; Figure 4.1.4). Ligation reactions could not have been contaminated with partially restricted or unrestricted pADH2 plasmid backbone because the 1-kb *Hpa*I-*Sac*I fragment of pADH2 had been excised from an agarose gel after electrophoretic separation of the components of a restriction-digest mixture. The linear *Nde*I-*Sac*I pET21a DNA used was also excised from agarose and cannot have been contaminated with pADH2. One must conclude that the similarity between the 4-kb construct and pADH2 was a coincidence. The antigenic activity was shown to be due to a 1.3 kb

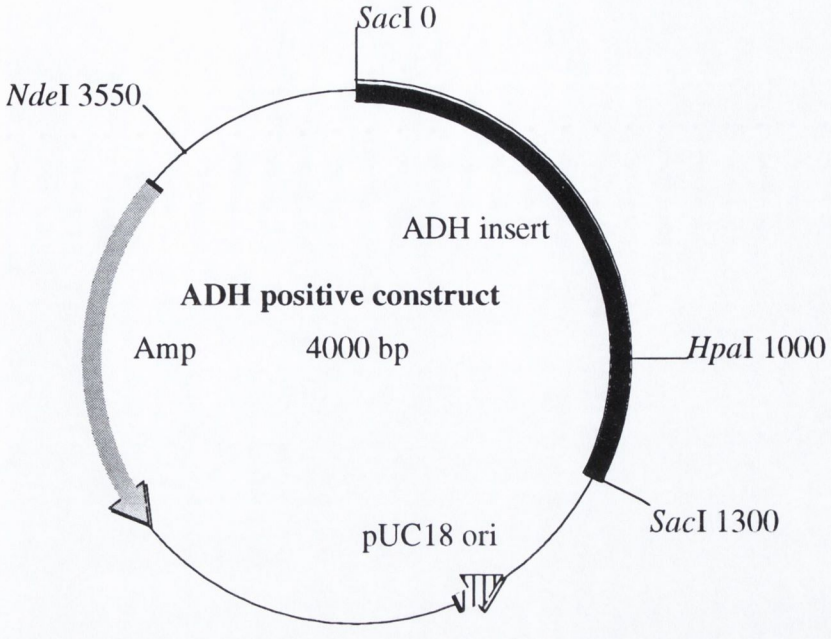


Figure 4.4.1 **Plasmid Map of the 4-kb ADH Positive Construct**

The map was constructed on the basis of the restriction analysis presented in Figure 4.3.4 and discussed in Section 4.3.3.

*Sac*I-*Sac*I insert, which did not include a *Nde*I site. The 1.3 kb *Sac*I-*Sac*I insert did include a *Hpa*I site, indicating that the original 1-kb *Hpa*I-*Sac*I *adh* fragment was intact. Assuming that the final construct was 4 kb, a fragment approximately 2.4 kb long, including the pET21a *Hpa*I site, was deleted from the vector backbone, and a *Sac*I site was introduced.

Two further attempts were made to construct an *adh* expression system in pET21a. In the second procedure, the 1-kb *Hpa*I-*Sac*I *adh* fragment was ligated to linear *Nde*I-*Sac*I pET21a, to produce a linear 6.4-kb *Hpa*I-*Nde*I *adh*-pET21a construct. The 6.4-kb construct was then circularised by adding the *Hpa*I-*Nde*I adapter. The *Hpa*I-*Nde*I *adh*-pET21a linear construct was identified on an agarose gel and purified, demonstrating that the 1-kb *Hpa*I-*Sac*I *adh* fragment was capable of ligation to *Nde*I-*Sac*I pET21a, and demonstrating that the 6.4-kb linear construct was stable *in vitro*. However, the construct was never recovered when it was ligated to adapter and transformed into *E. coli* TG1. Furthermore, the plasmids purified from the transformants were identical to the 4-kb constructs obtained using the original scheme. That result demonstrated that the deletion and recombination events that led to the production of the 4-kb construct, and the introduction of a *Sac*I site, occurred after transformation into *E. coli.*

In a third procedure, pADH2 was restricted with *Hpa*I, to produce linear *Hpa*I-*Hpa*I DNA and the linear DNA was circularised by the addition of the *Nde*I-*Hpa*I adapter. Circularisation was possible upon the incorporation of any multiple of two adapter molecules, the simplest situation being the ligation (*Hpa*I to *Hpa*I) of a single adapter molecule at each end of the linear pADH2, followed by circularisation (*Nde*I to *Nde*I). Plasmid pADH2 contained a *Nde*I site, originating from pUC18, and situated approximately 218 bp downstream from the pUC18 multiple cloning site. In pADH2, the pUC18 *Nde*I site was situated approximately 0.45 kb downstream from the *B. stearothermophilus* internal *Sac*I site (Section 4.1.3; Figure 4.1.4). Simple recirularisation of pADH2, without incorporation of adapter, was expected to give a 1.3-kb *Sac*I-*Sac*I restriction fragment and two *Nde*I-*Sac*I restriction fragments (0.45 and 2.4 kb long), and the majority of the nine transformants analysed did give this restriction pattern. The presence of a 1-kb *Nde*I-*Sac*I fragment was diagnostic for the incorporation of at least two adapter molecules into pADH2, and one plasmid (pADH2a; Section 4.3.6; Figure 4.3.6) was identified with that restriction pattern. Ligation of the *Nde*I-*Hpa*I adapter to either the purified 1-kb *Hpa*I-*Sac*I *adh* fragment, or to the purified 6.4-kb *Hpa*I-*Nde*I *adh*-pET21a linear construct, could not be monitored directly, but, the construction of pADH2a demonstrated that the adapter did contain *Nde*I-*Hpa*I compatible ends and that those ends were phosphorylated.

The purified 1-kb *Nde*I-*Sac*I fragment of pADH2a was ligated into *Nde*I-*Sac*I pET21a. The predicted result was a 6.4-kb plasmid containing a single *Nde*I site

and a single *Sac*I site. The ligation mixture was transformed into three different hosts (*E. coli* strains TG1, DH5α and SURE) to eliminate the possibility that the problem was related to the use of *E. coli* TG1 as host. Forty transformants were screened and no plasmid contained the intact 5.4-kb pET21a backbone. A single plasmid was found that gave a fragment pattern of 1, 2.4 and 0.45 kb when restricted with *Nde*I and *Sac*I simultaneously (Section 4.3.7; Figure 4.3.7); as this plasmid clearly did not contain the intact pET21a backbone, it was disregarded.

4.4.2 Construction of an Expression System Containing an *adh* PCR Product

Cloning of the 1-kb *Hpa*I-*Sac*I *adh* fragment into the *Nco*I and *Sac*I restriction sites of the expression vector pET21d was possible using a *Nco*I-*Hpa*I double-stranded adapter and any of the approaches described for pET21a. However, as these approaches had all failed, and for reasons that remained unclear, a procedure that did not involve a synthetic adapter, a *Nde*I site or a *Sac*I site was adopted for the construction of an expression system in pET21d.

PCR, with plasmid pADH2 as template DNA, was used to introduce an *Xho*I site (that partially overlapped the *Sac*I site) downstream from the *adh* stop codon and a *Bsp*HI site at the 5' end of the coding sequence according to the scheme outlined in Figure 4.1.6.

The production of the 1-kb *adh* PCR product was straightforward and minor contamination was barely visible on agarose gels (Section 4.3.8; Figure 4.3.8). When the PCR product was restricted with *Bsp*HI and *Xho*I, ligated into *Nco*I-*Xho*I pET21d and the ligation mixture transformed into *E. coli* TG1 and into *E. coli* SURE, nine transformants yielded plasmid DNA of the correct size (6.4 kb) and contained a unique *Xho*I site. The presence of the *Xho*I site indicated that the correct base changes had been introduced at the 3' end of the *adh* coding sequence. Four of the five plasmids isolated from *E. coli* SURE directed expression of a protein with an apparent subunit M_r that was correct for ADH 2334, namely 36,000. The fifth pETADH SURE construct directed the expression of a protein with an apparent subunit M_r of approximately 30,000 (Figure 4.3.13(a); Clone 2), suggesting either the introduction of an in-frame stop codon during the PCR, or a deletion of approximately 180 bp from the insert. Three of the four 6.4-kb plasmids isolated from *E. coli* TG1 directed IPTG-dependent expression of a 36,000-M_r protein. One of these gave a 'double-band' pattern (Figure 4.3.13(b); Clone 15), presumably due to proteolysis of the 36,000-M_r protein.

The motivation for cloning the *adh* PCR product into pTrc99A was a desire to maximise the number of possible constructs, given the previous failure to establish an expression system in pET21a. It quickly became apparent that the

expression system could be established using pET21d and only a small number of pTrcADH constructs were characterised. This small sample may not have been representative of the whole.

The yields of ampicillin-resistant transformants obtained when the pTrc99A ligation reaction mixtures were transformed into *E. coli* TG1, and into *E. coli* SURE, were equivalent to those obtained for the pET21d ligation reaction mixtures. On several occasions during the course of this work, it was observed that, compared to *E. coli* TG1 transformants, *E. coli* SURE transformants grew very slowly and the time-scale of culture had been adjusted accordingly, to ensure reasonable yields of plasmid DNA. In the case of the pTrcADH SURE transformants, slow growth appeared to be exaggerated and a working amount of plasmid DNA was not obtained. The growth rates of the pTrcADH TG1 transformants were equivalent to those normally observed for *E. coli* TG1 transformants in liquid culture. Five pTrcADH TG1 plasmids were screened by restriction with *Nco*I. The five plasmids appeared to be identical and were definitely larger than the parent vector. The *adh*2334 coding sequence contained two internal *Nco*I sites and therefore the insertion of the complete *adh* PCR product into pTrc99A was predicted to give a *Nco*I restriction fragment pattern of 4.9 kb and 0.122 kb. A fragment of approximately 5 kb was detected in each of the five plasmids. However, the plasmid preparations were contaminated with RNA, which obscured the lower portion of the gel, and the presence of the 0.122-kb fragment has to be assumed.

The ease with which the expression system was established in pET21d was in stark contrast to the difficulties encountered when attempting to construct an expression system in the closely related vector pET21a. The correct pET21d construct was identified from a screen of only five *E. coli* SURE transformants that had been selected, in a totally random manner, from a pool of 250 colonies. SDS-PAGE analysis indicated that a further three of the five pETADH SURE constructs were also correct. The numbers of ampicillin-resistant transformants per ng of vector DNA obtained in the pET21a procedures were, in general, similar to those obtained in the pET21d procedure. When attempting to construct the expression system in pET21a, at least ten, and often twenty, transformants were screened for every ligation. Furthermore, the selection procedure was not arbitrary; frequently, those colonies that did not react with anti-ADH 2334 antibody were removed from the total pool of colonies before plasmid size was investigated. Therefore, it seems extremely unlikely that the correct pET21a construct was present but was not detected.

The N-terminal amino-acid sequence of the 36,000-M_r protein expressed from one of the plasmids (pETADH SURE construct, Clone 1) was determined. The N-terminal sequence of the recombinant protein corresponded to the N-terminal sequence determined for the wild-type protein. It was concluded therefore that the

correct base changes had been introduced at the 5' end of the coding sequence, that the 36,000-M_r protein was rADH, and that the pETADH SURE construct clone 1 (designated pETADH) was a suitable expression plasmid for the recombinant *adh* 2334 gene.

CHAPTER 5
PURIFICATION AND CHARACTERISATION OF RECOMBINANT ADH 2334

5.1 INTRODUCTION

Following over-expression of recombinant ADH 2334 in *E. coli*, the logical next-step was purification of the enzyme and confirmation that it possessed NAD^+-dependent ethanol dehydrogenase activity together with kinetic parameters that were comparable to those reported for the wild-type enzyme purified by Sheehan *et al.* (1988). Confirmation that a cloned gene is expressed and that the recombinant protein has the correct N-terminal amino-acid sequence and subunit M_r is not a guarantee that soluble, biologically-active protein will be obtained following purification. The classic experiments on the refolding of ribonuclease *in vitro* (Anfinsen *et al.*, 1961) did demonstrate that all the information necessary to determine the final conformation of a protein can reside in the amino-acid sequence. However, subsequent studies on multidomain oligomeric proteins have demonstrated that spontaneous acquisition of a biologically-active conformation is not universal among proteins (Jaenicke, 1991).

Growing polypeptide chains, as they emerge from the ribosome, may participate in premature interactions with other intra- or inter-polypeptide domains that lead to the formation of insoluble, biologically inactive, protein aggregates. The formation of inappropriate folding interactions is prevented *in vivo* by accessory proteins that stabilise, temporarily, unfolded or partially folded protein conformations (Fisher & Schmid, 1990; Jaenicke, 1991; Georgopoulos, 1992; Bukau & Horwich, 1998; Morshauser *et al.*, 1999). Accessory proteins fall into two classes: enzymes, 'foldases', that catalyse specific isomerisation steps, and chaperone proteins. Foldases in *E. coli* include DsbA, a periplasmic enzyme involved in protein disulfide-bond formation (Bardwell *et al.*, 1991; Kamitani *et al.*, 1992), its oxidoreductase DsbB (Bardwell *et al.*, 1993) and peptidyl prolyl *cis-trans* isomerase, the catalyst for the isomerisation of Xaa-Pro peptide bonds (Liu & Walsh, 1990). Chaperone proteins stabilise unfolded or partially folded domains of their target proteins. Relatively few chaperones tend to assist in the folding and assembly of numerous different proteins, binding to unfolded proteins in a nonspecific manner. *E. coli* chaperone proteins include the heat shock proteins DnaK (Wang *et al.*, 1998), GroEL and GroES (Gething & Sambrook, 1992; Braig, 1998; Sigler *et al.*, 1998).

The formation of inappropriate folding interactions becomes increasingly likely at high protein concentration and temperature (Jaenicke, 1991; Mitraki *et al.*, 1991). Over-expression of recombinant proteins in *E. coli* frequently results in a dramatic increase in intracellular protein concentration. Insoluble protein aggregates, that derive from partially-folded intermediates of recombinant gene products, were first observed, by Williams *et al.* (1982), in *E. coli* cells expressing recombinant human insulin genes. Subsequently, such aggregates, described as inclusion bodies,

have been a major obstacle to the purification of functional recombinant proteins (Hockney, 1994). Inclusion bodies accumulate in the cytoplasm, can be seen with a phase-contrast microscope, and are not disrupted by sonication, or lysis under high pressure.

In Chapter 4 expression from pETADH was confirmed by SDS-PAGE analyses of whole-cell lysates. Inclusion bodies are lysed by boiling in SDS and therefore that method cannot reveal the presence of inclusion bodies in an expression host. The formation of inclusion bodies is revealed by low concentrations (or absence) of the recombinant protein in the soluble fraction collected after sonication or lysis under high pressure. Biologically-active protein can sometimes be recovered from insoluble inclusion bodies by denaturation and refolding *in vivo*. However, the exact procedure for a given protein has to be determined empirically (Grisshammer & Nagai, 1995).

The final consideration in relation to the production of biologically-active recombinant proteins in *E. coli* is the potential requirement for post-translational covalent modification of the protein primary structure. Post-translational covalent modifications of prokaryotic proteins include removal of N-terminal signal sequences, glycosylation, disulfide bond formation, and phosphorylation. N-terminal signal sequences occur only in proteins that are exported from the cytoplasm (Schatz & Dobberstein, 1996). ADH 2334 is made in, and functions in, the cytoplasm and therefore a N-terminal signal sequence is not required. Glycosylation of proteins in prokaryotes is usually associated with proteins that form part of the peptidoglycan of the cell wall (Drickamer & Taylor, 1998) and is not a characteristic of prokaryotic cytoplasmic proteins. Disulfide linkages will not occur in ADH 2334 because the reducing environment of the cytoplasm prevents disulfide bond formation (Freedman, 1984). Phosphorylation in prokaryotes is associated with proteins involved in signal transduction and gene expression (Hess *et al.*, 1988; Stock *et al.*, 1990) and not with proteins, such as ADH, which have a metabolic role.

5.2 MATERIALS AND METHODS

5.2.1 Source of Recombinant ADH

Recombinant ADH was purified from *E. coli* pLysS transformants that carried the plasmid pETADH (Figure 5.2.1). Cells transformed with pETADH were stored as cryocultures (Section 4.2.3).

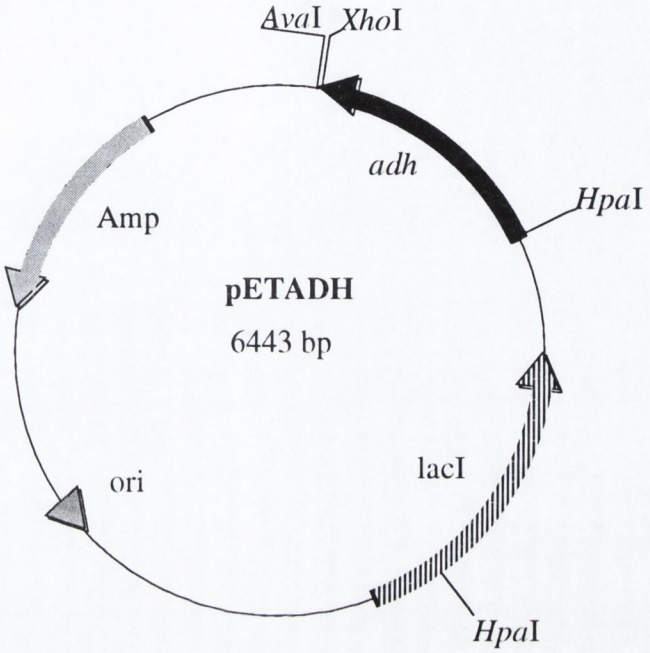


Figure 5.2.1 The Expression Plasmid pETADH

The plasmid was constructed by inserting a 1 kb *Bsp*HI-*Xho*I *adh* PCR product (*adh*) into pET21d which had been restricted with *Nco*I and *Xho*I (Chapter 4). The *Xho*I site was recreated in the ligation product but the *Bsp*HI and *Nco*I sites were not. The notation Amp indicates the ampicillin resistance gene, ori is the origin of plasmid replication, lacI is the *lacI* gene and arrows indicate the direction of transcription.

5.2.2 Culture Medium, Culture Conditions and Estimation of Cell Number

Culture medium was 2TY medium (Section 3.2.3) containing 100 μg ml^{-1} ampicillin and 34 μg ml^{-1} chloramphenicol. Cultures were typically 100 - 300 ml and were inoculated with cryocultures, 1.3 ml of cryoculture (equivalent to 1 ml of an overnight culture), being added to 100 ml of medium. When cells were grown in media supplemented with $ZnSO_4$ the appropriate volume of sterile 10 mM, 100 mM or 1M $ZnSO_4.7H_2O$ was added to the medium prior to the addition of inoculum. Cultures were grown, at 37 °C and at 250 revolutions min^{-1}, until the A_{600} value reached approximately 1.0 (3 - 3.5 h). Cultures were then adjusted to 1 mM IPTG (addition of 1 M IPTG) and returned to the incubator for a further 4 h. The number of cells in cultures was estimated using the approximation that an A_{600} value of 1 was equivalent to 10^9 cells ml^{-1} (Mandelstam & McQuillen, 1973). Determinations of A_{600} values were made on aliquots of cultures that had been diluted thirty-fold.

5.2.3 Potassium Phosphate Buffers

Potassium phosphate buffers were prepared by appropriate dilution, at room temperature, of sterile 1 M stock solutions. Stock solutions were prepared by mixing 1 M $K_2HPO_4.3\ H_2O$ and 1 M KH_2PO_4 to give the required pH.

5.2.4 Protease Inhibitors, Dithiothreitol and DNase I

Protease inhibitors were obtained from Sigma. Solutions of leupeptin (10 mg ml^{-1} in H_2O) and pepstatin (10 mg ml^{-1} in dimethyl sulphoxide) were stored at -20 °C. Solutions of phenylmethane sulphonyl fluoride (100 mM in acetone) were prepared immediately before use and were added to buffer that had been pre-warmed to 40 °C. Solutions of dithiothreitol (0.5 M in sterile H_2O) and DNase I (2 mg ml^{-1} in sterile H_2O) were stored as small aliquots at -20 °C.

5.2.5 *E. coli* Lysis-Buffers

Two different lysis-buffers were used. In initial experiments cells were lysed into 100 mM potassium phosphate, 2 mM EDTA, 0.02% (v/v) 2-mercaptoethanol, 0.1 mM phenylmethane sulphonyl fluoride, 3 μg ml^{-1} DNase, pH 6.1 (EDTA lysis-buffer). This buffer was described by Sheehan *et al.* (1988) and used by the authors during the purification of wild-type ADH 2334 from *B. stearothermophilus*. The EDTA lysis-buffer was modified following concern about the affect of EDTA on the zinc content of rADH, and to eliminate the possibility of mercaptoethanol acting as a substrate for ADH. The composition of the modified lysis-buffer, which was used in the purification scheme for rADH, was 100 mM potassium phosphate pH 6.1, 0.5 mM dithiothreitol, 0.1 mM phenylmethane

sulphonyl fluoride, 6 μg ml^{-1} DNase I, 0.5 μg ml^{-1} leupeptin, 1 μg ml^{-1} pepstatin. The modified lysis-buffer was prepared immediately before use by mixing 3 ml of a sterile stock solution of 1 M potassium phosphate pH 6.1 with 27 ml of sterile H_2O, the solution was warmed to 40 °C, 30 μl of 100 mM phenylmethane sulphonyl fluoride were added, the solution was cooled to 4 °C and dithiothreitol (30 μl of 0.5 M), DNase I (90 μl of 2 mg ml^{-1}), leupeptin (1.5 μl of 10 mg ml^{-1}) and pepstatin (3 μl of 10 mg ml^{-1}) were added.

5.2.6 Saturated Ammonium Sulfate Solution

Ammonium sulfate was obtained from Merck (product number 1.01217.1000) and according to the manufacturer's specifications was 0.0001% (w/w) in Zn^{2+}. A saturated solution of ammonium sulfate in potassium phosphate buffer was prepared by adding 100 g of ammonium sulfate to 150 ml of 100 mM potassium phosphate pH 7.0. The mixture was stirred overnight (room temperature) and then filtered through Whatman filter paper. The filtrate, when stored at 4 °C, had a pH of about 6.1. When necessary, the pH of the cold filtrate was adjusted to 6.1 by the addition of 5 M KOH.

5.2.7 Lysis of *E. coli* Cells

Cell pellets were frozen at -80 °C for at least 30 min (frequently overnight) and then thawed. Cell pellets that originated from small-scale (5 - 10 ml) cultures were dispersed in 1 ml of lysis-buffer (Section 5.2.5) per ml of original culture, cells originating from large-scale (200 - 300 ml) cultures were dispersed in 15 - 22 ml of lysis-buffer, depending on the size of the pellets. The cell suspensions were placed on ice for 10 min, at which time an obvious increase in the viscosity of the suspension was evident. The suspension was then either passed through a French pressure cell (Aminco) at an internal cell pressure of 165.6 x 10^3 kPa, or sonicated (six 30 s bursts at 120 W). Sonication was ineffective in achieving complete lysis of large volumes of cells and was only used for volumes less than 20 ml. Following sonication, or passage through the French pressure cell, lysates were centrifuged to remove cell-debris (43,000 ***g***, 20 min, 4°C), and the clear supernatants were transferred to sterile containers.

5.2.8 Ammonium Sulfate Fractionation

All solutions were maintained at 4 °C. The required % saturation ammonium sulfate was achieved by slow addition (with stirring) of ice-cold 100% saturated ammonium sulfate (Section 5.2.6). The volume (V) of saturated ammonium sulfate solution (in ml) required to adjust 100 ml of solution of initial saturation S_1 to a final saturation of S_2 was calculated using the Equation 8 where S_1

and S_2 are expressed as fractions of the saturated solution, e.g. $S_1 = 0.5$ and $S_2 = 0.7$ when adjusting from 50% to 70% saturation (Dawson *et al.*, 1986).

$$\mathbf{V = 100\ (S_2 - S_1)\ /\ 1 - S_2} \qquad \mathbf{(8)}$$

After the addition of ammonium sulfate, the solution was stirred on ice for 45 min and then centrifuged (43,000 ***g*** , 15 min, 4 °C) to collect the protein precipitate.

5.2.9 Purification of Recombinant ADH

Recombinant ADH was purified from two separate 300-ml cultures of *E. coli* pLysS transformants. One culture was 500 μM in supplementary $ZnSO_4$, the other contained no supplementary $ZnSO_4$. Aseptic techniques (Section 3.2.1) were adhered to during the growth and harvesting of *E. coli* transformants and were maintained throughout the purification procedure, in as far as was practical.

The cells were harvested by centrifugation (2,700 ***g***, 15 min, 4 °C) and the medium was decanted. The cell-pellet from each culture was washed three times by adding 100 ml of sterile 100 mM potassium phosphate pH 7.0 to the pellet, replacing the centrifuge bottle in the rotor at an orientation of 180 ° to the previous, and centrifuging as before. Phosphate buffer was removed from the final cell-pellets by standing the centrifuge bottles in an inverted position on pads of tissue paper. Cell-pellets were then stored overnight at -80 °C. Frozen cells were thawed by standing at room temperature for approximately 10 min and then dispersed into ice-cold modified lysis-buffer (Section 5.2.5), 22 ml of lysis-buffer per pellet. The cell suspensions were placed on ice for 10 min and then lysed by passage through a French pressure cell. The lysate was centrifuged to remove cell-debris (43,000 ***g***, 20 min, 4°C).

The clarified lysates were adjusted to 70% saturation ammonium sulfate by the addition of saturated ammonium sulfate solution (2.33 ml of 100% saturated ammonium sulfate per ml of lysate), stirred on ice for 45 min and then centrifuged (43,000 ***g*** , 15 min, 4 °C). The supernatants (which contained rADH) were transferred to clean beakers and adjusted to 85% saturation ammonium sulfate by addition of a saturated solution of ammonium sulfate (1 ml ml^{-1}). The precipitates of rADH were collected by centrifugation (as above), the bulk of the supernatant was removed from each centrifuge tube with a Pasteur pipette and the tubes were left to drain in an inverted position on a pad of tissue paper. The precipitates were dissolved in minimal volumes of sterile filtered 100 mM potassium phosphate pH 7.0 (1 ml per 100 ml of original culture), dispensed in small aliquots into sterile tubes stored at -20 °C.

5.2.10 Measurement of Protein Concentration

The method of Lowry *et al.* (1951), modified as described by Hartree (1972), was used to measure the protein concentration of cell lysates, of fractions from various stages of the purification of rADH and of pure rADH. The composition and preparation of all reagents and the technical details of the assay were exactly as described by Hartree (1972). All determinations were in triplicate. Standard solutions of bovine serum albumin, in the concentration range 10 - 100 μg ml^{-1} and in a final volume of 1 ml, were prepared by appropriate dilution, into H_2O, of a 1 mg ml^{-1} stock solution. A reagent blank, containing 1 ml of H_2O and the assay reagents, was processed in exactly the same way as the standards, and was used to set zero absorbance at 650 nm. A typical standard curve is presented in Figure 5.2.2.

Approximate estimates of protein concentration were made by measuring the absorbance of protein solutions at 280 nm. It was assumed that an A_{280} of 1 was equivalent to a protein concentration of 1 mg ml^{-1}.

5.2.11 Polyacrylamide-Gel Electrophoresis

E. coli cell lysates and samples from various stages of the purification of rADH were analysed using SDS-PAGE. Reagents, conditions and subunit M_r markers were as in Section 4.2.18. All resolving gels were 15.0% (w/v) acrylamide and unless stated otherwise the gel-mould dimensions were 90 x 80 x 1 mm. Prior to electrophoresis protein was precipitated from solution by the addition of trichloroacetic acid (Section 2.1.2.1). Precipitated protein was dissolved, to a concentration of approximately 5 μg μl^{-1}, in SDS-PAGE sample buffer (plus 2-mercaptoethanol), boiled for 5 min and either analysed immediately or stored at -20 °C. The usual loadings on a gel were volumes equivalent to 30 μg of protein.

5.2.12 Determination of ADH Enzyme Activity by Spectrophotometric Assay

The ethanol-dependent enzyme activity of rADH in cell lysates and in ammonium sulfate fractions was determined using a continuous spectrophotometric assay in which the production of NADH was monitored. Cell lysates and ammonium sulfate fractions were stored on crushed ice and were assayed within four hours of preparation. The conditions used for the assay were those described by Sheehan *et al.* (1988) for the assay of wild-type ADH 2334. The instrumentation used was an Ultraspec 11 spectrophotometer (LKB) equipped with a constant-temperature cuvette chamber and connected to a chart-recorder. All assays were performed at 37 °C and in quartz cuvettes with a 1 cm path length. The assay-buffer was 100 mM potassium phosphate pH 7.0 and was filtered (3.0 μm pore-size Millipore filters) before use. The disodium salt of NAD^+ was obtained from Boehringer, solutions were freshly

prepared each day and were 100 mg ml^{-1} (150 mM) in 100 mM potassium phosphate pH 7.0. The final concentration of NAD^+ in the cuvette was 0.51 mM. Laboratory-grade ethanol was used and was assumed to be 97.5% (w/w), 10 µl in a 3 ml assay therefore corresponding to a final concentration of approximately 56 mM. The protein concentration in the cuvette was 0.5 - 2 µg ml^{-1}.

Assay-buffer (typically 2.97 ml), ethanol (10 µl), and NAD^+ (10 µl) were mixed in the cuvette and the absorbance at 340 nm was monitored for 1 min to ensure the absence of a blank-rate. Protein solution (typically 10 µl) was then added and the increase in absorbance was monitored for at least 3 min. The rate of change of reaction velocity was estimated manually from the initial, linear, portion of the recorder trace of each assay. Reaction rates were linear for approximately 3 min and typical changes in absorbance were in the range 0.005 - 0.01 min^{-1}.

Control experiments demonstrated that the increase in absorbance at 340 nm was both enzyme dependent and substrate dependent. Ethanol and NAD^+ together, or separately, resulted in no observable increase in absorbance over 3 min. Enzyme alone, enzyme plus ethanol, and enzyme plus NAD^+ resulted in no observable increase in absorbance. The rate observed when enzyme plus NAD^+, or enzyme plus ethanol, were preincubated in the cuvette for three minutes and the second substrate added was reduced to approximately one third of that observed when enzyme was added last. Addition of further enzyme to an assay mixture exhibiting such a low level of activity resulted in stimulation of the reaction rate to the expected level indicating that the enzyme, but not either substrate, was unstable in the cuvette.

5.2.13 Determination of ADH Enzyme Activity by Fluorimetric Assay

Determinations of initial reaction velocity (v_o) that were used to determine K_m, maximum reaction velocity (V_{max}) and k_{cat} values of purified rADH were made by determining the rate of increase of fluorescence emission intensity at 465 nm (excitation at 340 nm) due to ethanol-dependent NADH production. Purified rADH that had been stored at -20 °C (Section 5.2.9) was used in the determinations of the kinetic constants. Frozen protein was allowed to thaw slowly on crushed ice and was kept on crushed ice for the duration of the measurements (3 - 6 h). Enzyme was added last to the (prewarmed) assay mixture. Individual aliquots were discarded at the end of the day and a fresh aliquot used for any subsequent assays.

The instrumentation used was a Perkin Elmer LS-50B fluorimeter equipped with a constant-temperature cuvette chamber, an automatic mixing facility for the cuvette contents and connected to a printer. The assay-buffer and NAD^+ solutions were as described for the spectrophotometric assay (Section 5.2.12).

Ethanol used in the kinetic studies was obtained from James Burroughs. The density of the undiluted ethanol, determined by measuring the mass of known volumes on an analytical balance (Mettler), was 0.797 g ml^{-1} at 20 °C, equivalent to an ethanol concentration of 97.5% (w/w) (Hodgman, 1955). A primary stock (1.68 M) was prepared by diluting 0.794 g of ethanol to 10 ml in a volumetric flask. Working stocks (168 mM and 16.8 mM) were prepared by diluting the primary stock as appropriate.

All assays were performed in triplicate, at 37 °C, pH 7.0, in quartz cuvettes with a 1 cm path length, in a final volume of 2.93 ml, at a fixed final concentration of NAD^+ of 0.51 mM and with ethanol at concentrations ranging from 5.76 µM - 1153 µM. Potassium phosphate buffer (100 mM, pH 7.0), NAD^+ (150 mM, 10 µl) and the appropriate volume of 168 mM or 16.8 mM ethanol were mixed in a cuvette and the fluorescence at 465 nm was monitored for at least 20 s to ensure the absence of a blank rate. Enzyme was added to the cuvette using a Hamilton syringe and the increase in fluorescence intensity was monitored for approximately 100 s. Amounts of enzyme added to the cuvette were 0.8 µg (0.5 µl) of +Zn rADH and 0.7 µg (1 µl) of -Zn rADH. The rate of change of reaction velocity was estimated manually from the initial, linear, portion of the printed record of each reaction-progress curve.

The control experiments described for the spectrophotometric assay (Section 5.2.12) were performed for the fluorimetric assay. There was no increase in fluorescence intensity unless all components of the assay were present in the cuvette and the increase in fluorescence was both enzyme dependent and substrate dependent. With enzyme added last and ethanol concentration in excess of 56 µM reaction-progress curves were linear for at least 30 s. Reaction-progress curves were linear over a time-scale of at least 20 s when ethanol was at, or in excess of, 17 µM. Below 17 µM ethanol reaction-progress curves were linear over a time-scale of 10 s.

5.2.14 Calibration Curves for Measurement of NADH Concentration

A primary stock solution of approximately 1 mM NADH was prepared by dissolving 7 mg of solid NADH (disodium salt, Boehringer) in 10 ml of 100 mM potassium phosphate pH 7. A secondary stock was then prepared by ten-fold dilution of the primary stock and the concentration of the secondary stock was accurately determined by measuring the absorbance at 340 nm. A fluorescence standard was prepared by ten-fold dilution of the secondary standard. The fluorescence intensity of appropriate dilutions of the fluorescence standard (in 100 mM potassium phosphate buffer pH 7.0) was measured at 456 nm (excitation at 340 nm). An example of a calibration curve is presented in Figure 5.2.3.

5.2.15 Calculation of Specific Activity

A Unit (U) of enzyme activity was defined as that catalysing the formation of 1 µmol of NADH per min. Specific activity was defined as the number of Units of enzyme activity per mg of protein (U mg^{-1}). Formation of NADH was monitored by the spectrophotometric method and absorbance values were converted to µmol of NADH using a value of 6.22 x 10^3 M^{-1} cm^{-1} for the molar absorption coefficient of NADH. Protein concentrations for specific activity estimation were measured (in triplicate) using the Lowry assay, the mean value being used in the calculation of specific activity.

5.2.16 Determination of Kinetic Parameters

Values for K_m (apparent) and V_{max} were obtained by analysing the experimental data with the non-linear regression data analysis programme Enzfitter (Leatherbarrow, 1987). Theoretical values of initial velocity (v_o) at a range of theoretical substrate concentrations were generated from the Michaelis-Menten equation (Michaelis & Menten, 1913; Equation 9) using the values of K_m and V_{max} obtained by non-linear regression data analysis. The theoretical values of initial velocity were than used to generate theoretical Michaelis-Menten plots [v_o as a function of substrate concentration ([S])] and theoretical Hanes plots ([S] / v_o as a function of [S]; Hanes, 1932; Equation 10].

$$v_o = V_{max}[S] / ([S] + K_m) \quad (9)$$

$$[S] / V_{max} = K_m / V_{max} + 1 / V_{max} \text{ x } [S] \quad (10)$$

The parameter k_{cat} was defined as the number of mol of NADH formed per second per mol of subunit rADH (V_{max} / [E], where [E] was total concentration of enzyme). The protein concentrations of the rADH preparations used to determine k_{cat} values were measured using the Lowry assay (Section 5.2.10) and the molar subunit concentration was calculated using a subunit M_r of 36,205 (Section 3.4.2.1). Values of V_{max} were converted to mol of NADH formed per second using calibration curves relating fluorescence intensity to accurately known concentrations of NADH (Section 5.2.14; Figure 5.2.3).

5.2.17 Measurement of Zinc Ion Concentration

The Zn^{2+} concentration of bacteriological medium, of phosphate buffer, of ammonium sulfate solutions and of solutions of purified rADH was measured by atomic absorption spectroscopy using a Pye Unicam atomic absorption

spectrophotometer. Zero absorbance (at 213.9 nm) was set with distilled water. Samples were diluted into, and all standards were prepared in, distilled water collected from the still at the same time as that used to set zero absorbance. All samples and standards were prepared in, and stored in, polyethylene plastic containers that had been washed extensively in distilled water.

The primary Zn^{2+} standard was 10 mM and was prepared in an analytical manner using a mass of $ZnSO_4.7H_2O$ that had been accurately determined using an analytical balance (Mettler). Secondary standards (40 μM and 400 μM) were prepared by dilution of the primary standard. Working standards covering the concentration range 0.64 μM - 3.2 μM Zn^{2+} were prepared by appropriate dilutions of the 40 μM standard and working standards covering the concentration range 3.2 μM - 16.0 μM Zn^{2+} were prepared by appropriate dilutions of the 400 μM standard. Absorbance values were determined by passing sample through the flame until the detector registered a constant value, typically 5 ml of sample were necessary. Examples of standard curves for the two concentration ranges are presented in Figure 5.2.4.

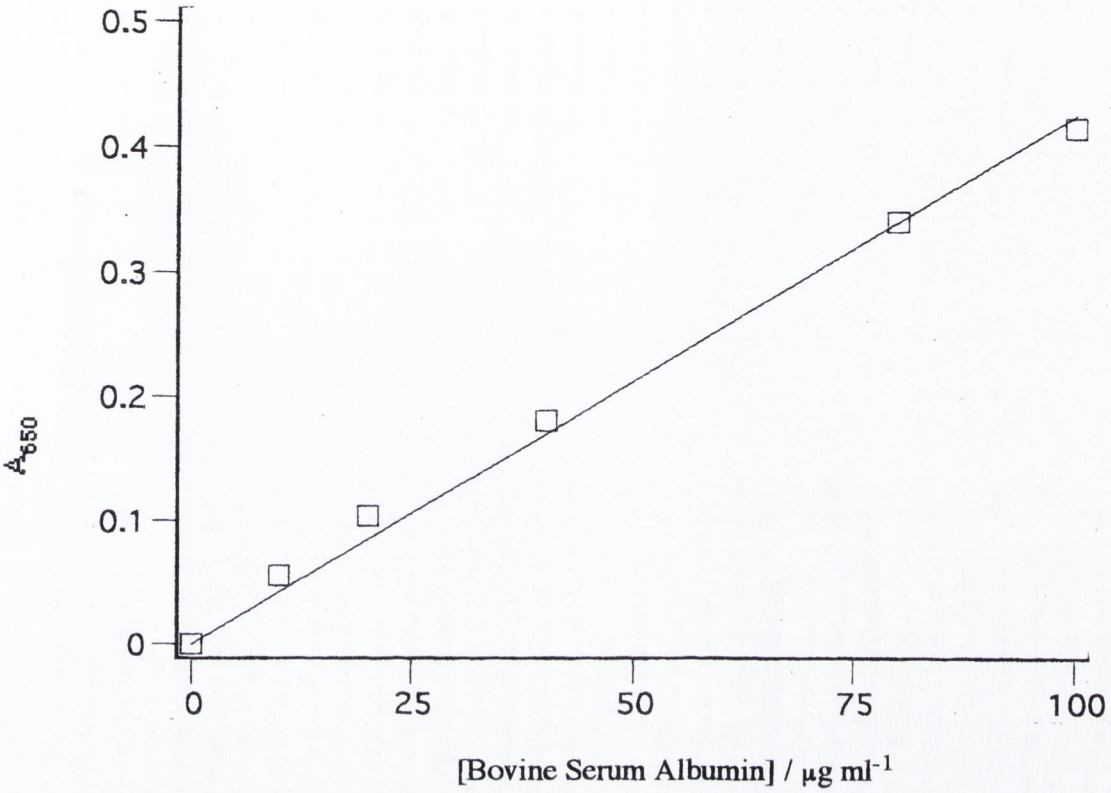


Figure 5.2.2 **An Example of a Standard Curve Used for Measurement of Protein Concentration by the Lowry Assay**

Each result represents the mean ± S.D. of a determination performed in triplicate. In all cases the standard error of the mean was less than the representation of the point.

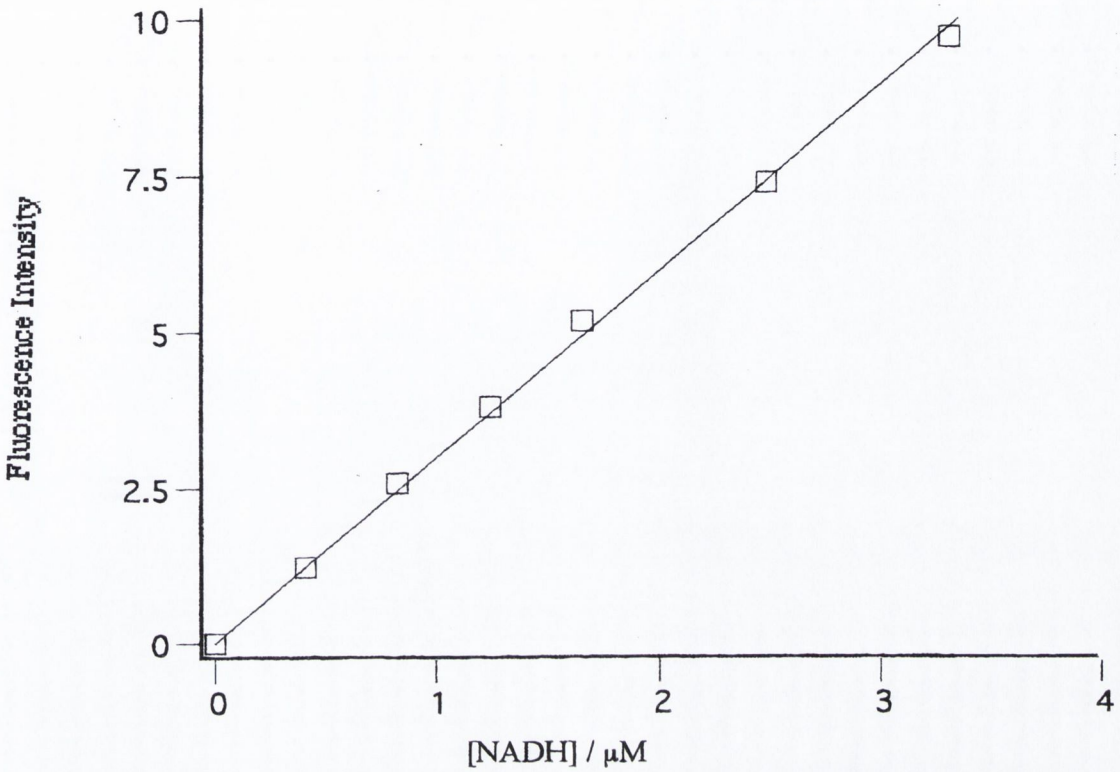


Figure 5.2.3 **An Example of a Calibration Curve Used to Relate Fluorescence Intensity to NADH Concentration**

A fluorescence standard (8.26 μM in this example) was prepared, by dilution of primary and secondary standards, and the fluorescence of dilutions of the standard were determined as described in Section 5.2.14. Each result represents a single determination.

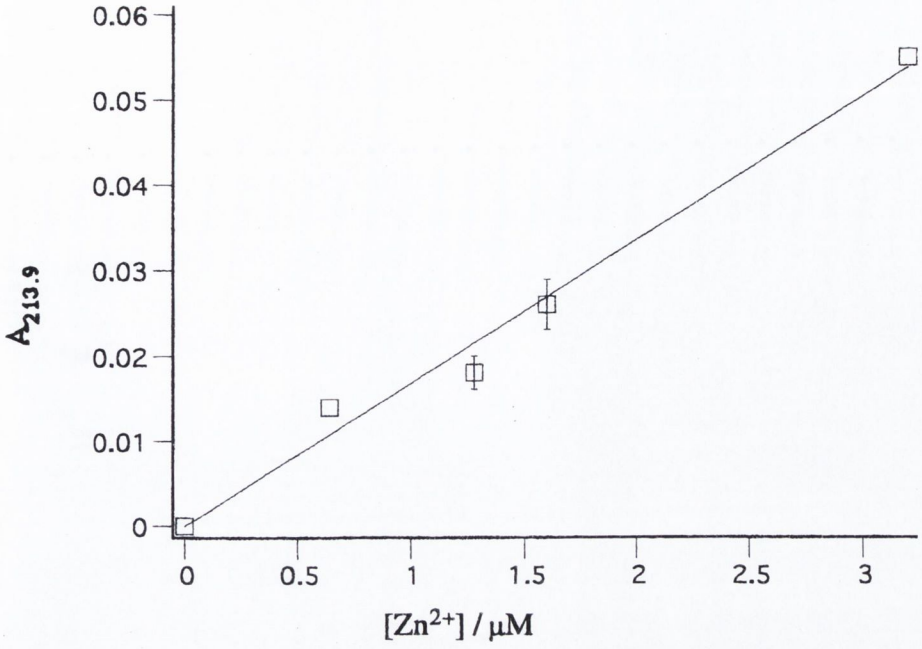


Figure 5.2.4(a) **An Example of a Standard Curve for Measurement of Zinc Ion Concentration in the Range 0.64 μM - 3.2 μM**

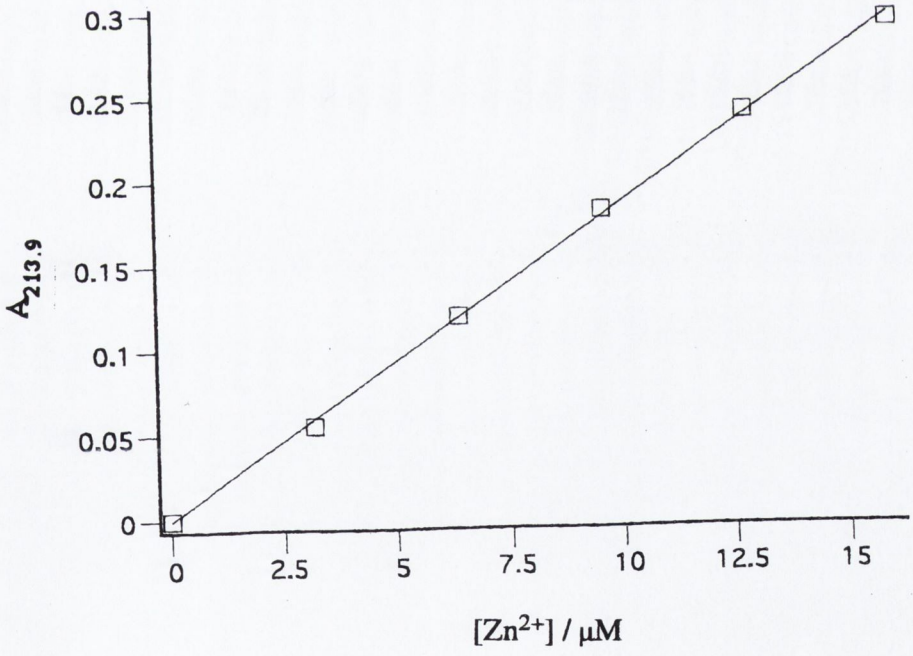


Figure 5.2.4(b) **An Example of a Standard Curve for Measurement of Zinc Ion Concentration in the Range 3.2 μM - 16 μM**

The results presented represent the mean ± S.D. of three determinations made on each solution. Where no error bar is shown the standard error of the mean is less than the representation of the point.

5.3 RESULTS

5.3.1 Optimum Conditions for IPTG-Dependent Expression

The pETADH plasmid was not maintained in the expression host *E. coli* BL21, when transformants were stored on agar plates. In fact, when *E. coli* BL21 transformants carrying the original nine pETADH constructs (Section 4.3.9; Figure 4.3.13) were screened for IPTG-dependent expression, after storage on agar plates for fifteen days, expression was evident only in cells carrying constructs 2, 4, and 16. The pETADH plasmid was maintained, for at least six months, in the expression host *E. coli* pLysS, when transformants were stored as cryocultures.

The growth rate and cell yield of the BL21 and the pLysS strain were compared. Growth curves for nontransformed cells and for cells transformed with pETADH are presented in Figure 5.3.1. In liquid culture for 8 h, at 37 °C and in the presence of IPTG, both strains of transformed cells attained a cell density equivalent to an A_{600} value of approximately 1.5 [Figure 5.3.1(b)]. For both strains, the final cell density attained by the transformed cells was reduced, compared to the final cell density of nontransformed cells.

To compare expression from pETADH in the two strains, and to determine the time necessary to achieve optimum levels of expression, transformed cells were grown in liquid culture, cells were harvested before, and at various times after, the addition of IPTG and cell lysates were prepared. The growth curve and corresponding SDS-PAGE analysis of *E. coli* BL21 transformants are presented in Figure 5.3.2. IPTG was added at 4 h (Lane 1) and the maximum level of expression was achieved 3 h later (Lane 4). The growth curve and corresponding SDS-PAGE analysis of *E. coli* pLysS transformants are presented in Figure 5.3.3. IPTG was added at 2.5 h (Lane 2) and the maximum level of expression was achieved 4 h later (Lane 6). The pLysS cells were left in culture for 23 h after the addition of IPTG and that resulted in proteolysis of the recombinant protein (Lane 8). In the SDS-PAGE analyses presented in Figures 5.3.2 and 5.3.3 equivalent numbers of cells were loaded in each lane and therefore the data show that there was little difference between the two host cells in terms of the yield of recombinant protein obtained 4 h after addition of IPTG. The Coomassie Blue staining intensity of the gels indicated that both hosts produced the recombinant protein at the level of at least 30% of the total cell protein.

The *E. coli* pLysS host contains the T7 lysozyme gene (Section 4.1.1.1). This is potentially useful in that simple procedures that rupture the inner cell-membrane (such as freeze-thawing) allow endogenous lysozyme access to the peptidoglycan layer and greatly facilitate quick and complete cell lysis (Studier *et al.*, 1990). The pLysS host was selected on the basis that it was potentially useful to be

able to lyse the cells quickly. There was little difference between the two hosts when the yields of recombinant protein were monitored.

5.3.2 Ammonium Sulfate Fractionation of a Cell Lysate and Specific Activity of Recombinant ADH in the Fractions

Transformed *E. coli* pLysS cells (200-ml culture) were grown in the presence of IPTG for 3 h, harvested, and lysed by sonication into 12.5 ml of EDTA lysis-buffer (Section 5.2.5). Insoluble material was removed by centrifugation and the soluble fraction (post-sonication supernatant) was subjected to ammonium sulfate fractionation. Five ammonium sulfate pellets were collected: 0 - 30% saturation; 30 - 40% saturation; 40 - 50% saturation; 50 - 65% saturation and 65 - 80% saturation. Each ammonium sulfate pellet was resuspended in 2 ml of 50 mM potassium phosphate buffer pH 7, and dialysed against the same buffer.

An SDS-PAGE analysis of a sample of the post-sonication supernatant, and SDS-PAGE analyses of the resuspended ammonium sulfate pellets are presented in Figure 5.3.4, 30 μg of total protein were loaded in each lane. The post-sonication supernatant (PSS) was in Lane 2: the presence of the intensely stained 36,000-M_r protein in that lane demonstrated that rADH was expressed as soluble protein and at a level far in excess of any one soluble endogenous *E. coli* protein. Comparison of the staining intensity of the 36,000-M_r protein in Lane 2 with that of 5 μg of the 36,000-M_r marker in Lane 1 suggests that approximately 15 μg of the 30 μg loaded onto the gel were rADH. The 65 - 80% ammonium sulfate pellet (Lane 7) was 'almost pure' 36,000-M_r protein. However, the 36,000-M_r protein was also present in all the other ammonium sulfate pellets (Lanes 3 - 6). The staining intensity of the 36,000-M_r protein varies across Lanes 3 - 6, however, it is never less than 25% of that in Lane 7 and is considerably more in Lane 3. One expects to see some spread of a given protein across ammonium sulfate precipitates, due to mechanical trapping of the supernatant in the pellets, but the relative amounts of the 36,000-M_r protein present in each fraction are difficult to rationalise on that basis.

The post-sonication supernatant and the five resuspended ammonium sulfate pellets were assayed for ADH enzyme activity and the results, presented as specific activities, are presented in Table 5.3.1, along with protein concentrations and values for total protein and total enzyme activity. The total ADH enzyme activity of the post-sonication supernatant, which represented the total yield from 200 ml of culture, was 2 U. Approximately 50% of the total activity was recovered in the 65 - 80% ammonium sulfate pellet. The SDS-PAGE analysis (Figure 5.3.4) shows that the 65-80% saturation ammonium sulfate pellet was 'almost pure' 36,000-M_r protein, the presence of ADH enzyme activity in that fraction confirms that the 36,000-M_r protein was rADH. The specific activity of the 'almost pure' 65-80% ammonium

sulfate pellet was 0.09 U mg^{-1}, comparing unfavourably with the value of 2.1 U mg^{-1} reported for the purified, salt-free, wild-type enzyme (Sheehan *et al.*, 1988). The 0 - 30% ammonium sulfate pellet contained no ADH enzyme activity and that was surprising because 7 mg of protein were recovered in the 0 - 30% ammonium sulfate pellet and about 50% of that protein was the 36,000-M_r protein. The three pellets collected in the 30 - 65% saturation range did contain ADH enzyme activity. Values for the specific activity and the total amount of activity of the 30 - 40% saturation pellet were approximately 20% of those recorded for the 65 - 80% saturation pellet. Values for specific activity and total amount of activity of the other two ammonium sulfate pellets were approximately 10% of those recorded for the 65 - 80% saturation pellet. Of the total 2 U present in the post-sonication supernatant 0.74 U were not accounted for.

The total protein content of the post-sonication supernatant was 101 mg, a total of 42 mg were recovered in the five ammonium sulfate pellets indicating that approximately 60 mg of protein remained in solution in the 65 - 80% saturation ammonium sulfate supernatant. The recovered protein was distributed approximately equally between the five pellets.

5.3.3 Addition of Zinc Ions to Culture Medium Resulted in Increased Specific Activity of Recombinant ADH

The use of the modified lysis-buffer, which did not contain Na_2EDTA (Section 5.2.5), resulted in no increase in the specific activity of rADH present in the post-sonication supernatant. Some increase in the specific activity of rADH present in the post-sonication supernatant was achieved by reducing the time elapsing between cell lysis and the enzyme assay and subsequent experiments (Section 5.3.6) demonstrated that enzyme activity was reduced by dialysis. Undialysed, crude lysate, prepared from cells grown for 4 h in the presence of IPTG and assayed immediately, had a specific activity of 0.18 U mg^{-1}.

The addition of $ZnSO_4$ to the culture medium caused an increase in the specific activity of rADH in post-sonication supernatants. There was variability between initial experiments as to the magnitude of the effect achieved by including 500 μM Zn^{2+} in the culture medium. The increase in specific activity ranged from three to ten-fold, the variability being due to variability in the values obtained for rADH activity in the lysates derived from cells that had not been exposed to supplementary Zn^{2+}. That variability was related to the time elapsing between cell lysis and the assay. When a regimen that minimised the time-period between cell lysis and enzyme assay was strictly adhered to, the maximum fold-increase in the specific activity of the crude lysate prepared from transformants grown in the

presence of 500 μM supplementary Zn^{2+} was consistently in the range three to four-fold.

Two 5-ml cultures of *E. coli* pLysS transformants were grown in 2TY medium, one of the cultures was supplemented with 500 μM $ZnSO_4$ prior to the addition of inoculum, the other contained no supplementary Zn^{2+}. Cells were grown for 2.5 h before, and for 2.5 h after, the addition of IPTG. Cells were harvested from 3 ml of each culture and the cells were lysed, by sonication, into 3 ml of modified lysis-buffer. The protein concentration of each post-sonication supernatant was measured and each supernatant was assayed for ADH enzyme activity. The specific activity of rADH in the soluble fraction of each lysate is presented in Table 5.3.2(a). Growth of the expression host in the presence of 500 μM supplementary $ZnSO_4$ resulted in a 3.3-fold increase in the specific activity of rADH in the soluble fraction. Comparing the concentration of cells in the two cultures, the cell concentration was lowest in the culture that had been supplemented with Zn^{2+}. However, the protein concentration was highest in the soluble fraction derived from the culture grown in the presence of supplementary Zn^{2+} and, therefore, *E. coli* grown in the presence of supplementary Zn^{2+} produced more soluble protein per cell than *E. coli* grown in standard 2TY medium. The increase in soluble protein was less than two-fold. Increasing the length of time cells were exposed to both IPTG and supplementary Zn^{2+} did not result in a greater than 3.3-fold increase in specific activity. Cells were grown, in the absence and in the presence of supplementary $ZnSO_4$, for 2 h before and for 4 h after the addition of IPTG and the fold increase in the specific activity of rADH in the lysate prepared from transformants grown in the presence of supplementary Zn^{2+} was again 3.3-fold [Table 5.3.2(b)]. In this experiment the amount of soluble protein produced per cell was again higher when *E. coli* was grown in the presence of supplementary Zn^{2+}.

Control experiments were performed to check that including Zn^{2+} in the culture medium did not influence expression of an endogenous *E. coli* ADH activity. Two cultures of nontransformed pLysS cells, one containing no supplementary Zn^{2+} the other containing 750 μM supplementary $ZnSO_4$ were grown for 3 h, IPTG was added to both cultures and growth was continued for a further 3 h. Cells were harvested and lysed, and the lysates were assayed for ADH activity. No ADH activity was detected in the cells derived from either culture, confirming that the increase in specific activity observed when pETADH was expressed in the presence of supplementary Zn^{2+} was due to increased enzyme activity of rADH, or alternatively, increased expression of soluble rADH, or a combination of the two. SDS-PAGE analyses of the nontransformed cells are presented in Figure 5.3.5 (Lanes 2 - 5). A comparison between Lanes 3 and 5 confirms the result of the enzyme assays in that, there is no evidence for the expression of a 36,000-M_r *E. coli*

protein in medium supplemented with Zn^{2+}. Figure 5.3.5 also shows a comparison of expression from pETADH in the absence and presence of 750 μM supplementary Zn^{2+} (Lanes 6 - 9; equivalent numbers of cells in each lane). The greater amount of rADH present in Lane 9, compared to Lane 7, indicates that addition of Zn^{2+} to the medium resulted in increased expression of pETADH.

Investigations into the affect of supplementary Zn^{2+} on the growth of transformed cells are presented in Figure 5.3.6. Cells did not grow in medium containing 2000 μM supplementary Zn^{2+} and final cell yield in the presence of 1500 μM Zn^{2+} was reduced compared to that achieved in the absence of supplementary Zn^{2+} [Figure 5.3.6(a)]. Cell density at 3 h and at 6 h as a function of supplementary Zn^{2+} concentration are presented in Figure 5.3.6(b). Cell yield increased with increasing supplementary Zn^{2+} concentration up to 100 μM Zn^{2+}. Over the range 100 - 1000 μM Zn^{2+} increasing the concentration of Zn^{2+} resulted in a small reduction in cell yield over the first three hours of growth. A more definite reduction in cell yield was evident when the cells were grown for 6 h in the concentration range 100 - 1000 μM Zn^{2+}.

The relationship between the specific activity of rADH in cell lysates and the concentration of supplementary Zn^{2+} in the growth medium is presented in Figure 5.3.7. Cells were grown in the presence of different concentrations of supplementary Zn^{2+}, cell lysates were prepared and the specific activity of rADH in each lysate was determined. Over the concentration range 0 - 500 μM supplementary Zn^{2+}, the specific activity of rADH increased as the concentration of Zn^{2+} in the medium increased. The maximum effect, which in Figure 5.3.7 is a 3.3-fold increase in rADH activity, was achieved at 500 μM Zn^{2+}. Increasing the concentration of supplementary Zn^{2+} above 500 μM was not accompanied by a corresponding increase in rADH activity. The affect of increasing concentrations of supplementary Zn^{2+}, in the culture media, on the stability of the enzyme activity, in cell lysates stored at 4 °C, is presented in Figure 5.3.8. When cells were grown in the absence of supplementary Zn^{2+} only 17.4% of the original activity remained after four days, whereas 56.4% of the original activity was detected in the lysate prepared from cells grown in the presence of 500 μM supplementary Zn^{2+}.

5.3.4 Depletion of Zinc Ions from Culture Medium During Expression

E. coli pLysS cells transformed with pETADH were grown in standard 2TY medium. Samples were removed at various times over a total time-period of 24 h. The samples were centrifuged to remove the cells and the concentration of Zn^{2+} in the cell-free aliquots of media was measured (Figure 5.3.9). A second culture

contained *E. coli* pLysS cells transformed with the vector pET21d as a control for the utilisation of Zn^{2+} for cell growth, as distinct from the synthesis of rADH.

Over 24 h the Zn^{2+} concentration in the medium containing cells that produced rADH did decrease, from 25.2 μM to 21.9 μM. However, over the time period during which rADH was expressed (3 - 7 h after inoculation) the concentration of Zn^{2+} in the medium remained approximately constant. Over the same time period the Zn^{2+} concentration in the medium containing cells that did not produce rADH decreased from 25.8 μM to 23.1 μM.

5.3.5 Ammmonium Sulfate Fractionation of a Cell Lysate Prepared from Cells Grown in the Presence of Zinc Ions

Transformants were grown as a 500-ml culture in medium that had been supplemented with 750 μM Zn^{2+}. Expression was induced for 4 h, cells were lysed by sonication , and the entire post-sonication supernatant (+Zn lysate), was subjected to ammonium sulfate fractionation. Four ammonium sulfate pellets were collected; 0 - 50% saturation, 50 - 60% saturation, 60 - 70% saturation, and 70 - 85% saturation. A control culture, also 500 ml, contained no added Zn^{2+} and was fractionated with ammonium sulfate in exactly the same manner (-Zn lysate). SDS-PAGE analyses of the two ammonium sulfate fractionation schemes are presented in Figure 5.3.10. The behavior, during ammonium sulfate fractionation, of the +Zn lysate 36,000-M_r protein was very different from that of the -Zn lysate 36,000-M_r protein. The +Zn lysate 36,000-M_r protein remained in solution until the ammonium sulfate concentration exceeded 70% saturation. In contrast, the -Zn lysate 36,000-M_r protein was present in all the ammonium sulfate pellets collected and that was reflected in the high protein concentration of the -Zn lysate ammonium sulfate pellets (0 - 70% saturation range), compared to the equivalent pellets derived from the +Zn lysate. The results presented in Figure 5.3.10 for the -Zn lysate are not directly comparable to those presented previously in Figure 5.3.4 and Table 5.3.1. The protein precipitates were not collected at exactly the same % saturation ammonium sulfate. However, in general terms the behavior of the 36,000-M_r protein in the fractionation scheme presented in Figure 5.3.10 was the same as that presented in Figure 5.3.4, the 36,000-M_r protein was present in all the ammonium sulfate pellets.

In Figure 5.3.10, the final supernatants are included in the analyses and in both the -Zn lysate and the +Zn lysate the 36,000-M_r protein was present in the supernatant. The difference in the distribution of the 36,000-M_r protein in the two fractionation schemes in Figure 5.3.10 cannot be attributed to some inadvertent difference in experimental conditions as that would have resulted in a concomitant difference in the distribution of *E. coli* proteins. With the exception of a 66,000-M_r protein, the behavior of endogenous *E. coli* proteins was the same in the two

fractionations. One must conclude that the difference in the behavior of the two rADH preparations reflected a genuine difference between the rADH expressed in the presence and absence of supplementary Zn^{2+}.

5.3.6 Purification of Recombinant ADH

When a cell lysate was adjusted to 70% saturation ammonium sulfate in a single step and centrifuged the resulting supernatant contained pure rADH (Figure 5.3.11). The enzyme precipitated from solution when the 70% saturation ammonium sulfate supernatant was adjusted to 85% saturation ammonium sulfate.

A summary of the yield and specific activity of rADH at each stage of the final purification scheme is presented in Table 5.3.3. The enzyme purified from the culture that contained supplementary Zn^{2+} was designated +Zn rADH, the enzyme purified from the culture that contained no supplementary Zn^{2+} was designated -Zn rADH. SDS-PAGE analyses of the final 85% saturation ammonium sulfate pellets are shown in Figure 5.3.12. The recovery of +Zn rADH enzyme activity in the 70% saturation supernatant was 81% (70 U; 35 mg). Given that the 70% supernatant contained only rADH, the total amount of +Zn rADH produced by the 300-ml culture can be calculated to be 43 mg, equivalent to a yield of 143 mg l^{-1}. Expression of rADH in the presence of supplementary Zn^{2+} was therefore at the level of 33% of the total soluble protein produced by *E. coli*.

A value of 113% was obtained for the recovery of -Zn rADH in the 70% saturation supernatant. That probably resulted from the fact that the French pressure cell lysate was not assayed immediately upon preparation but many hours later. The enzyme activity of the $-Zn^{2+}$ lysate was unstable (Section 5.3.3) and therefore the value obtained for the total activity of the -Zn French pressure cell lysate may be less than the true value. The amount of -Zn rADH recovered in the 70% supernatant was 26 mg. Because a reliable figure was not available for the recovery of -Zn rADH in the 70% supernatant it is not possible to say more than that expression of rADH in the absence of supplementary Zn^{2+} was at the level of at least 87 mg l^{-1}.

The specific activity of the resuspended 85% saturation ammonium sulfate pellet of +Zn rADH was 2.29 U mg^{-1}, which represents a four-fold increase in specific activity relative to the specific activity of the 85% saturation ammonium sulfate pellet of -Zn rADH.

Recovery of rADH from the 70% saturation supernatant was incomplete, approximately 46% of the +Zn rADH and 20% of the -Zn rADH was recovered as a 85% saturation ammonium sulfate precipitate. Recovery of +Zn rADH from the 70% supernatant by dialysis and by gel-filtration was attempted and both methods resulted in significant loss of enzyme activity (but not of protein). The loss of specific activity upon dialysis ranged from 35 - 84%. The loss of specific activity upon gel-filtration

on Sephadex G-25 in the presence of dithiothreitol, phenylmethane sulphonyl fluoride and 1% (v/v) ethanol was ~76%. Ammonium sulfate precipitation was used to recover the enzyme from the 70% saturation supernatant because the enzyme activity was stable under those conditions.

The SDS-PAGE analysis of the purified +Zn rADH and -Zn rADH was performed using a "long-format" SDS gel to check for any difference in apparent subunit M_r that might have gone undetected using the 'mini-gel' system. Electrophoresis was continued until the tracker-dye ran off the end of the gel and no difference in electrophoretic mobility was detected (Figure 5.3.12; Lanes 3 and 4).

5.3.7 The Zinc Content of Purified Recombinant ADH

Atomic absorption spectroscopy was used to measure the subunit : Zn^{2+} stoichiometry of the two preparations of rADH. Examples of standard curves, used for the measurements of Zn^{2+} concentration, were presented in Figure 5.2.4. In the concentration range 3.2 μM - 16 μM Zn^{2+} the standard curves were linear but the standard curves for concentrations below 3.2 μM Zn^{2+} were not linear. All of the measurements, with the exception of those made on the -Zn rADH preparation, were made using the 3.2 μM - 16 μM Zn^{2+} standard curve. The concentration of Zn^{2+} in the various reagents used in the purification of rADH and the concentration of Zn^{2+} in the final preparations of +Zn rADH and -Zn rADH are presented in Table 5.3.4.

The values presented for the Zn^{2+} concentration of the enzyme preparations and the molar ratios of ADH : Zn^{2+} were not corrected for contamination by Zn^{2+} introduced by reagents used during the purification procedure, or for carry-over from the medium. The Zn^{2+} concentration of the saturated ammonium sulfate solution was 12% and 3.6% of the concentration in solutions of -Zn rADH and +Zn rADH, respectively. However, the enzymes were precipitated out of ammonium sulfate solution and resuspended in phosphate buffer. The values obtained for the Zn^{2+} concentration of the phosphate buffers were less than the uncertainty in the measurement of Zn^{2+} concentration in either of the enzyme preparations.

Cell lysates derived from cultures of *E. coli* pLysS cells transformed with pET21d and, therefore, not producing rADH, were used to measure the magnitude of the carry-over of Zn^{2+} from the supplemented medium to the cell-pellet. *E. coli* pLysS cells transformed with pETADH and cells transformed with pET21d were grown as 200-ml cultures in medium supplemented with 500 μM Zn^{2+} (pETADH +Zn and pET21d +Zn). Both transformants were also grown in unsupplemented medium (pETADH -Zn and pET21d -Zn). IPTG was added to all four cultures after 3 h and growth was continued for a further 4 h. Each cell pellet was washed three times with phosphate buffer as described in Section 5.2.9 and lysed into 15 ml of

modified lysis-buffer. The Zn^{2+} content of the cell lysates, expressed as nmol of Zn^{2+} mg^{-1} of protein are presented in Table 5.3.5.

The Zn^{2+} content of the cell lysate derived from the pET21d +Zn culture was 3.5 nmol mg^{-1}. That value was approximately double that of the lysate derived from the pET21d -Zn culture, demonstrating that carry-over of Zn^{2+} from supplemented medium to the cell pellet did occur. Comparing the Zn^{2+} content of the cell lysates derived from the two pET21d cultures, carry-over was calculated to be 1.7 nmol of Zn^{2+} mg^{-1} of protein, therefore, 15% of Zn^{2+} content of the pETADH +Zn lysate could be attributed to carry over. Carry-over at that level could not have had a significant affect on the Zn^{2+} content of the final preparation of +Zn rADH because the pETADH +Zn lysate was diluted a total of six-fold during the purification procedure and furthermore, the protein was precipitated out of solution and redissolved in phosphate buffer.

The molar ratio of subunit rADH to zinc for the -Zn rADH preparation was 1 : 1.0 and the molar ratio of subunit rADH to zinc obtained for the +Zn rADH was 1 : 1.5.

5.3.8 Determination of the Kinetic Parameters of Recombinant ADH Using a Fluorimetric Assay

Initial experiments to determine the value of K_m for ethanol using the -Zn rADH preparation and the spectrophotometric assay indicated that V_{max} was approached at 46 μM ethanol. Reaction progress curves were linear for about 1 min at 46 μM ethanol, the time-scale of linearity decreasing to 30 s as the ethanol concentration was decreased below 46 μM. The variation in triplicate measurements about the mean was unacceptably large. When the fluorimetric assay was used to measure v_o, reaction progress curves were linear for no more than 30 s at 56 μM ethanol and it was apparent that rates determined using the spectrophotometric assay and low concentrations of ethanol were frequently not true initial rates. The large variability in triplicate measurements observed with the spectrophotometric assay at low substrate concentrations could therefore be accounted for.

The experimental data obtained for the two enzyme preparations are displayed as Michaelis-Menten plots in Figure 5.3.13. The data obtained over the lower range of ethanol concentrations (5.76 μM - 173 μM) are presented as Hanes plots in Figure 5.3.14. The K_m and k_{cat} values that were determined for the two enzyme preparations are presented in Table 5.3.6. The values of K_m are apparent and were determined at a fixed NAD^+ concentration of 0.51 mM.

5.3.9 Addition of Zinc Ions to Purified -Zn Recombinant ADH

Zinc ions were added to purified -Zn rADH in an attempt to stimulate enzyme activity *in vitro*. In all experiments the enzyme was assayed (spectrophotometric assay) in 100 mM potassium phosphate buffer and the intrinsic Zn^{2+} concentration of that buffer was assumed to be 0.76 μM (from Table 5.3.4). In one procedure purified -Zn rADH (stored at -20 °C) was assayed and then reassayed under identical conditions but with $ZnSO_4$ added to the assay mixture. In these experiments the rADH concentration (subunit) in the cuvette was 0.4 μM and the supplementary Zn^{2+} concentration in the cuvette was 0.6 μM or 33 μM, achieved by the addition of 1 mM and 10 mM $ZnSO_4$, respectively. No effect on enzyme activity was observed.

In a different procedure, enzyme was added to the cuvette before the addition of Zn^{2+}. The increase in absorbance due to rADH activity was monitored continuously as the concentration of Zn^{2+} was increased by the cumulative addition of $ZnSO_4$ to the cuvette. A facsimile of the recorder trace from one of these experiments is presented in Figure 5.3.15. In the example presented the rADH concentration (subunit) in the cuvette was 37.5 nM and the added Zn^{2+} concentration was increased by the cumulative addition of aliquots of 0.1 mM, 1 mM or 10 mM $ZnSO_4$, each addition increasing the supplementary Zn^{2+} concentration ten-fold up to a final concentration of 33.3 μM supplementary Zn^{2+}. It is clear that the addition of Zn^{2+} to the cuvette did not cause a stimulation in rADH activity.

5.3.10 The Effect of Temperature on Recombinant ADH

-Zn rADH (2 μl aliquots of the purified enzyme) were incubated (10 min) at 0 °C, at room temperature and at 37 °C in the presence of different concentrations of added Zn^{2+}. The incubated enzyme was then assayed (spectrophotometric assay) at 37 °C. The enzyme subunit concentration in all incubations was 19 μM and the supplementary $ZnSO_4$ concentration in the incubation mixtures ranged from 0 to 166.6 μM. At any given incubation temperature the inclusion of Zn^{2+} in the incubation had no affect on the enzyme activity measured at 37 °C. However, the incubation temperature itself did influence the enzyme activity observed at 37 °C (Figure 5.3.16). When the enzyme was incubated at 37 °C the catalytic activity was increased by approximately a factor of four relative to the activity observed when the enzyme was incubated at 0 °C.

The effect of incubation temperature on the enzyme activity of the +Zn rADH preparation was investigated. The enzyme was incubated on ice, at 18 °C and at 37 °C with dithiothreitol present (0.5 mM final concentration) and without dithiothreitol present. Enzyme that had been incubated in the presence of dithiothreitol was assayed in the presence of 0.5 mM dithiothreitol, and the incubation

mixtures which did not contain dithiothreitol were assayed in the absence of dithiothreitol. The enzyme activity, determined at 37 °C, of both sets of incubation mixtures increased as the incubation temperature was increased. The enzyme activity observed for the 37 °C incubation was increased by about a factor of four, relative to that of the 0 °C incubation (Figure 5.3.17).

The affect of the incubation temperature on the activity of +Zn rADH was reversible. The enzyme was incubated on ice, assayed, moved to a 37 °C water-bath and assayed again. The activity observed in the second assay was increased relative to that observed in the first assay. The activity decreased when the enzyme was returned to the ice bucket but increased again when the enzyme was returned to the 37 °C bath (Figure 5.3.18). In view of the unexpected nature of these results it is emphasised that all experiments involving incubations were carried out on enzyme preparations that had been stored at -20 °C, thawed on crushed ice and held on ice for time periods ranging from 10 min to 4 h before the incubations were set up.

Figure 5.3.1 **Growth of Expression Hosts in 2TY Medium**

Nontransformed cells, (a), were grown as 200 ml liquid cultures and with no IPTG added to the media. Inoculation of each culture was with 2 ml of a stationary-phase culture. Cells transformed with pETADH, (b), were grown as 100 ml liquid cultures with IPTG added at the times indicated. Inoculation of each culture was with 1 ml of a stationary-phase culture. Cell density was monitored by removing 1 ml samples at various times and determining the A_{600} value of a three-fold dilution (into 2TY medium) of the sample.

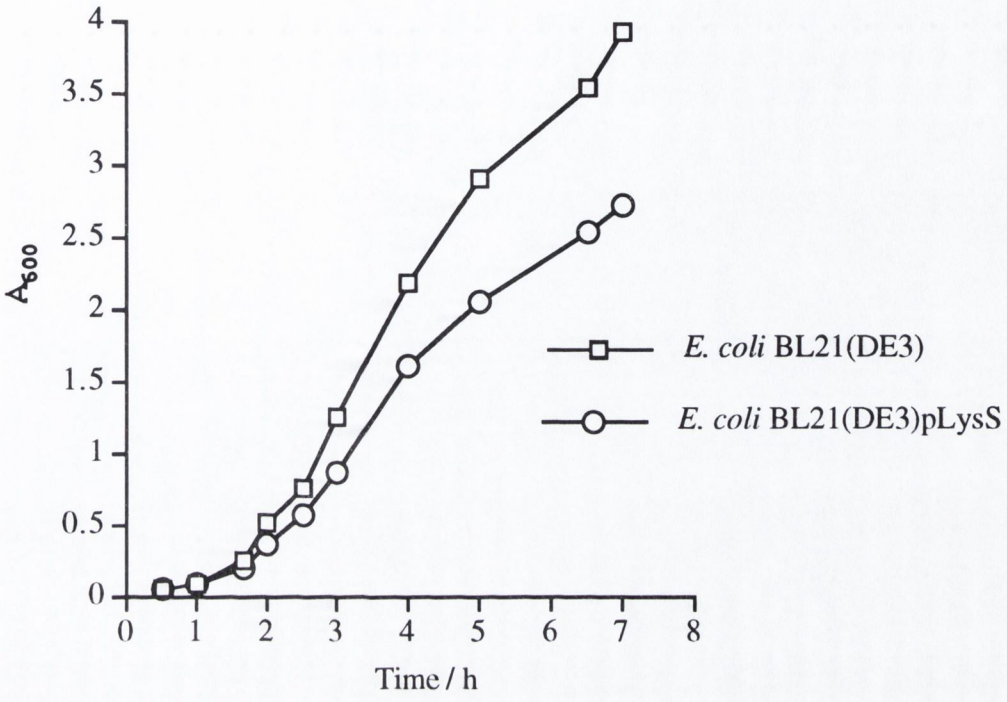


Figure 5.3.1(a) Nontransformed Expression Hosts

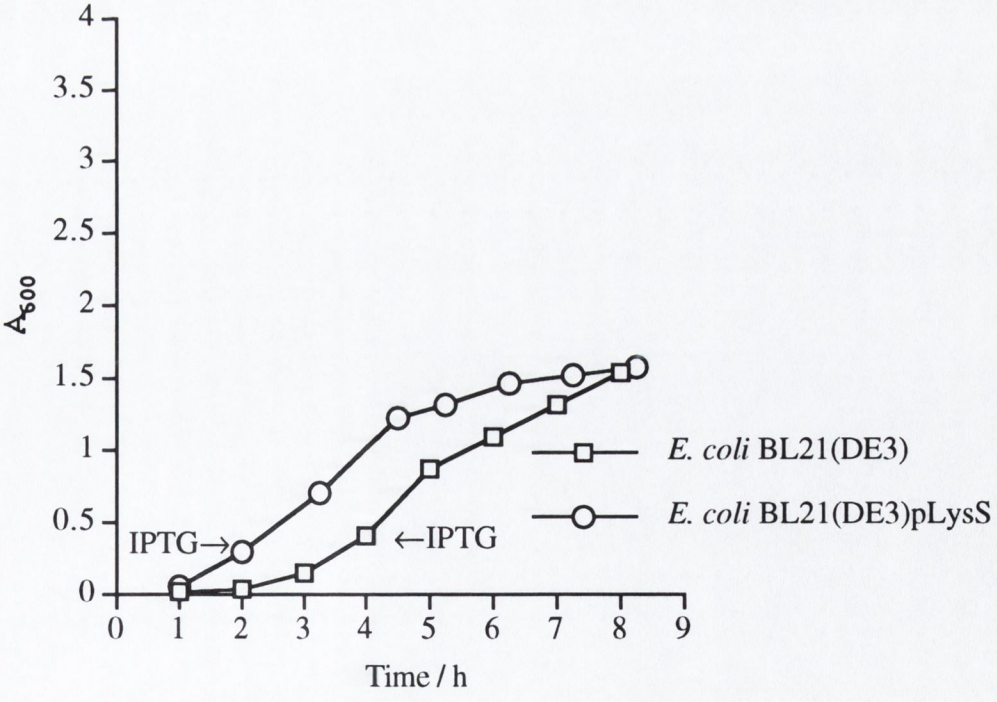


Figure 5.3.1(b) Expression Hosts Transformed with pETADH

Figure 5.3.2 **IPTG-Dependent Expression in *E. coli* BL21**

E. coli BL21 transformants were grown as a 100 ml liquid culture. The inoculum was 1 ml of a stationary-phase culture and IPTG was added at 4 h. Cell density was monitored as described in the legend to Figure 5.3.1 and the growth curve is presented in (a). An SDS-PAGE analysis of cells removed from the culture at various times is presented in (b). Cells were harvested from 1 ml samples removed immediately before (Lane 1) and at hourly intervals after (Lanes 2 - 5) the addition of IPTG. Cell pellets were lysed by suspension in SDS-PAGE sample buffer, the volume (ml) of sample buffer added to a pellet being A_{600} value ÷ 5.0, equivalent numbers of cells were thus analysed when 10 µl of each lysate were subjected to electrophoresis. Lane 6 contained subunit M_r markers, 5 µg of each.

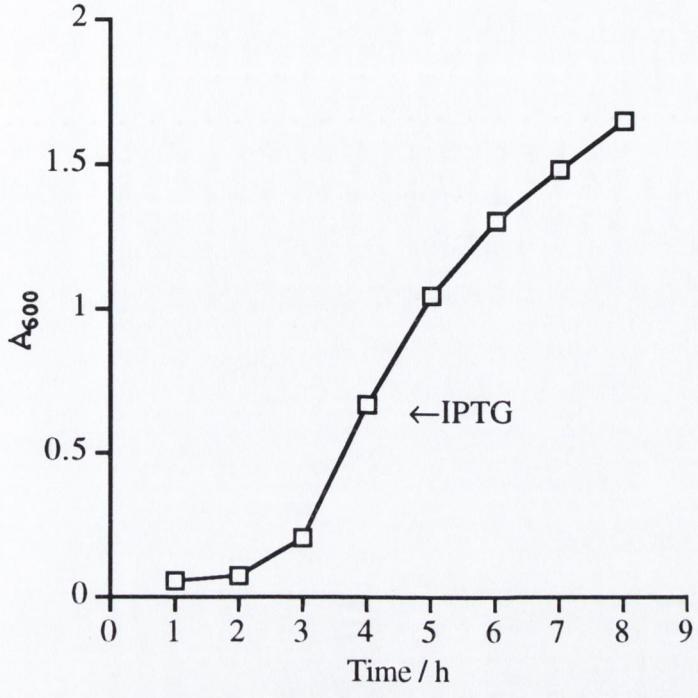


Figure 5.3.2(a)

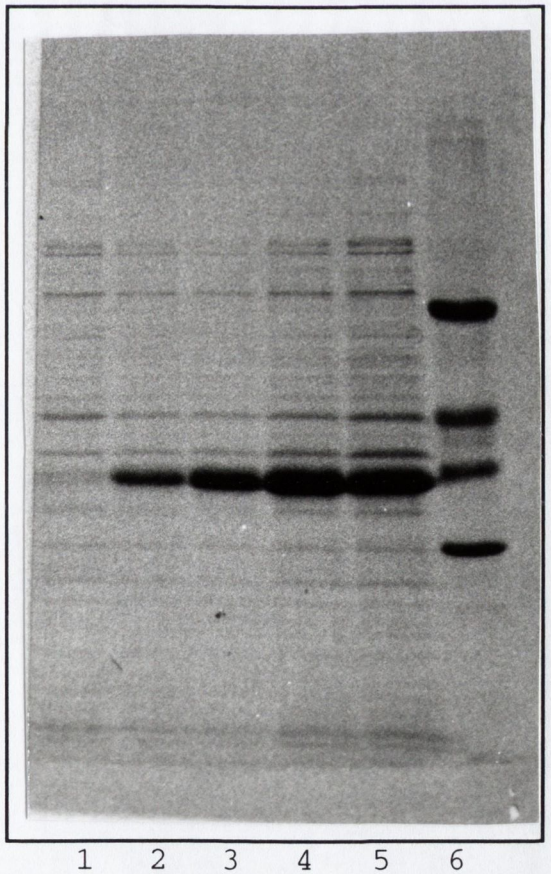


Figure 5.3.2(b)

Figure 5.3.3 **IPTG-Dependent Expression in *E. coli* pLysS**

E. coli pLysS transformants were grown as a 200 ml liquid culture. The inoculum was 2.6 ml of cryoculture and IPTG was added at 2.5 h. The growth curve is presented in (a) and the SDS-PAGE analysis in (b). Cells were harvested from 1 ml samples removed immediately before (Lane 2) and at hourly intervals after the addition of IPTG (Lanes 3 - 8). All other details as in the legend to Figure 5.3.2. Lane 1 contained subunit M_r markers, 5 μg of each.

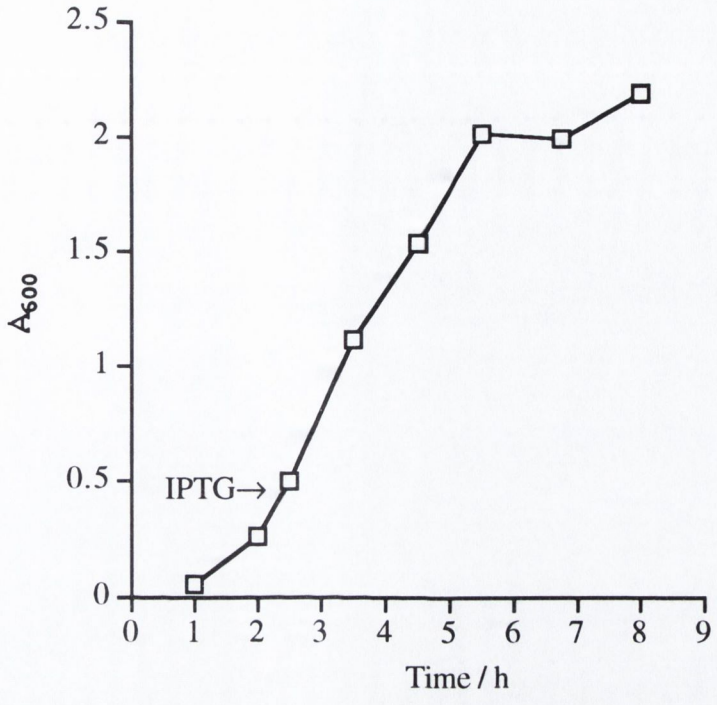


Figure 5.3.3(a)

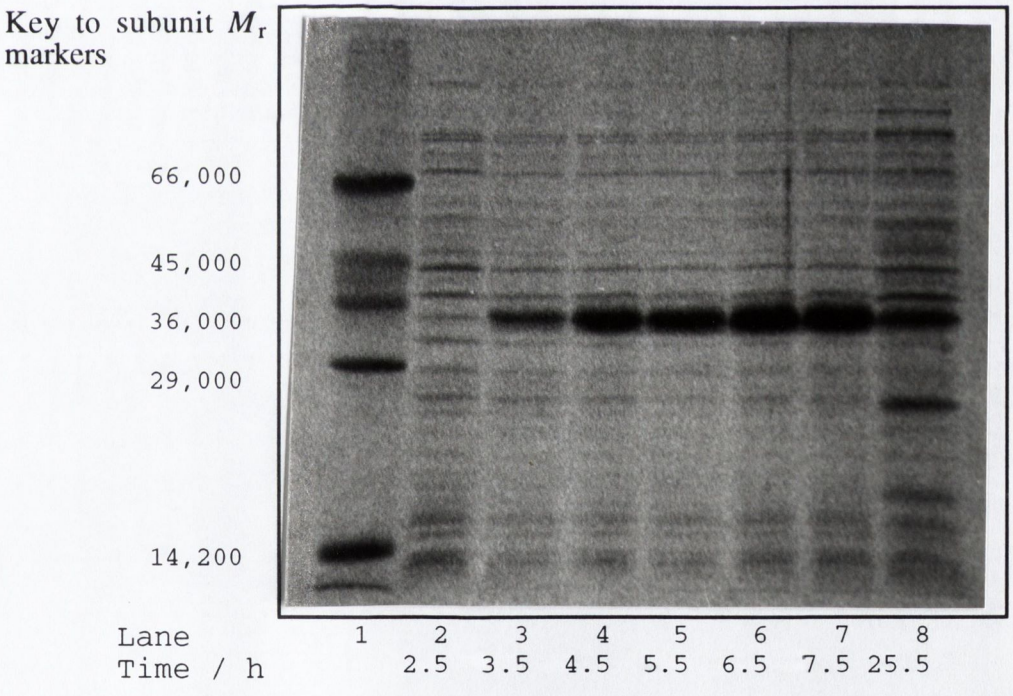


Figure 5.3.3(b)

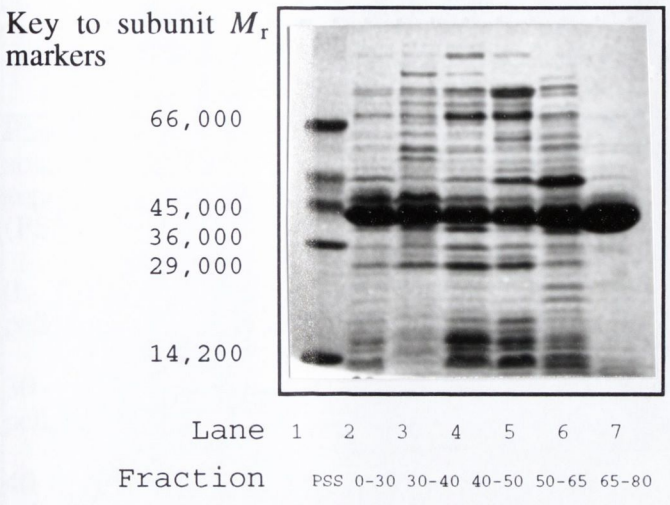


Figure 5.3.4 **Ammonium Sulfate Fractionation of an *E. coli* Cell Lysate**

Ammonium sulfate fractionation of a lysate of *E. coli* pLysS transformants is described in the text (Section 5.3.2). The protein concentration of the post-sonication supernatant (before ammonium sulfate fractionation) and of each dialysed ammonium sulfate precipitate was determined by the Lowry assay and a volume of each fraction, equivalent to 30 μg of total protein, was analysed by SDS-PAGE. Lane 1 contained subunit M_r markers, 5 μg of each. Lane 2 contained post-sonication supernatant (PSS). Lanes 3 - 7 contained the resuspended ammonium sulfate pellets and 0 - 30 *etc.* refers to the % saturation ammonium sulfate.

Fraction	Volume ml	[Protein] mg ml^{-1}	Total protein mg	Specific activity U mg^{-1}	Total activity U
Post-sonication supernatant (PSS)	12.5	8.1	101	0.02	2
0 - 30% pellet	2	3.4	7	0	0
30 - 40% pellet	2	4.6	9	0.02	0.2
40 - 50% pellet	2	4.4	9	0.01	0.09
50 - 65% pellet	2	3.4	7	0.01	0.07
65 - 80% pellet	2	5.0	10	0.09	0.9

Table 5.3.1 **Protein Concentration and ADH Enzyme Activity of the Post-Sonication Supernatant and of each Ammonium Sulfate Precipitate Presented in Figure 5.3.4**

[$ZnSO_4$] μM	[Cell] 10^9 cells ml^{-1}	[Protein] mg ml^{-1}	Total protein mg	Soluble protein per cell 10^{-10} mg	Specific activity U mg^{-1}	Fold increase in specific activity
0	3.93	0.91	2.73	2.3	0.08	
500	3.57	1.1	3.3	3.1	0.26	3.3

(a) Cells Harvested 2.5 h after Addition of IPTG

[$ZnSO_4$] μM	[Cell] 10^9 cells ml^{-1}	[Protein] mg ml^{-1}	Total protein mg	Soluble protein per cell 10^{-10} mg	Specific activity U mg^{-1}	Fold increase in specific activity
0	2.91	0.64	1.92	2.2	0.18	
500	2.42	0.62	1.86	2.6	0.59	3.3

(b) Cells Harvested 4 h after Addition of IPTG

Table 5.3.2 Growth of *E. coli* pLysS Transformants in the Presence and Absence of 500 μM Supplementary Zn Ions

Cultures (5 ml) of *E. coli* pLysS transformants were grown in 2TY medium and in 2TY medium supplemented with 500 μM $ZnSO_4$ as described in Section 5.3.3. The number of cells in a ml of each culture were calculated using the approximation A_{600} of 1 is equivalent to 10^9 cells ml^{-1} (Mandelstam & McQuillen, 1973). Total protein represents the total amount of soluble protein obtained from 3 ml of lysate.

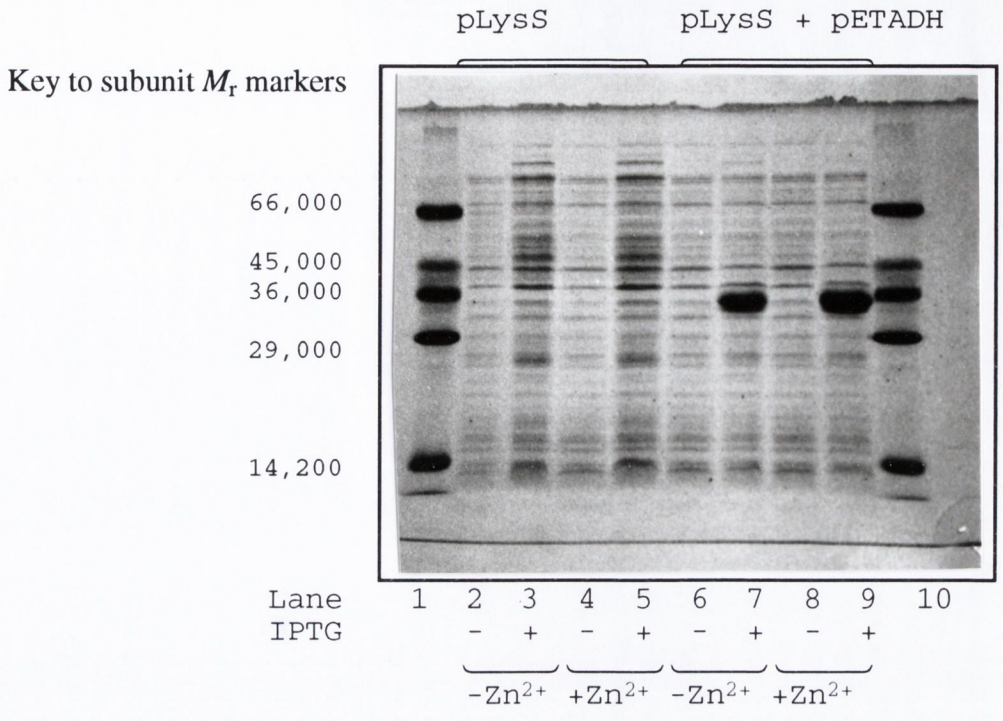


Figure 5.3.5 **Expression in Nontransformed and Transformed *E. coli* Cells in the Absence and in the Presence of $ZnSO_4$**

Nontransformed *E. coli* pLysS cells (Lanes 2 - 5) were grown as 5 ml cultures in the absence ($-Zn^{2+}$) and in the presence ($+Zn^{2+}$) of 750 μM supplementary Zn^{2+} for 3 h, samples (250 μl) were removed from each culture (-), IPTG was added to the remainder and the cells were grown for a further 3 h and a second sample was removed from each culure (+). *E. coli* pLysS cells transformed with pETADH (Lanes 6 - 9) were cultured, and samples removed, in exactly the same manner. Equivalent numbers of cells were loaded in each lane of the gel, as explained in the legend to Figure 5.3.2. Lanes 1 and 10 contained subunit M_r markers, 5 μg of each.

Figure 5.3.6 **Growth of *E. coli* Transformants in the Presence of Supplementary Zinc Ions**

Medium was inoculated with a cryoculture of *E. coli* pLysS transformants, dispensed (6 ml aliquots) into 25 ml tubes and each tube was supplemented with a different concentration of $ZnSO_4$. Cultures were grown for 3 h, adjusted to 1 mM IPTG and returned to the incubator for a further 3 h. Cell growth was monitored at (approximately) hourly intervals by withdrawing 300 µl samples from each culture and determining the A_{600} value of ten-fold dilutions (in 2TY medium) of the samples. A selection of the growth curves are presented in (a) and cell density at 3 h and at 6 h as a function of supplementary Zn^{2+} concentration is presented in (b).

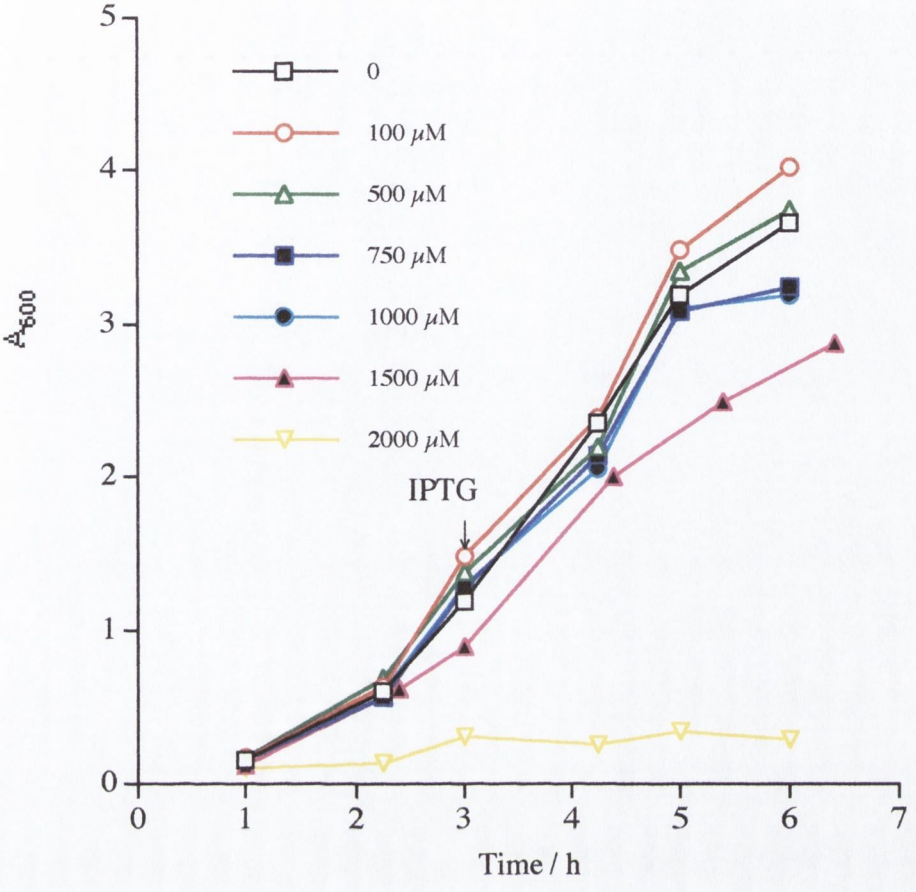


Figure 5.3.6(a)

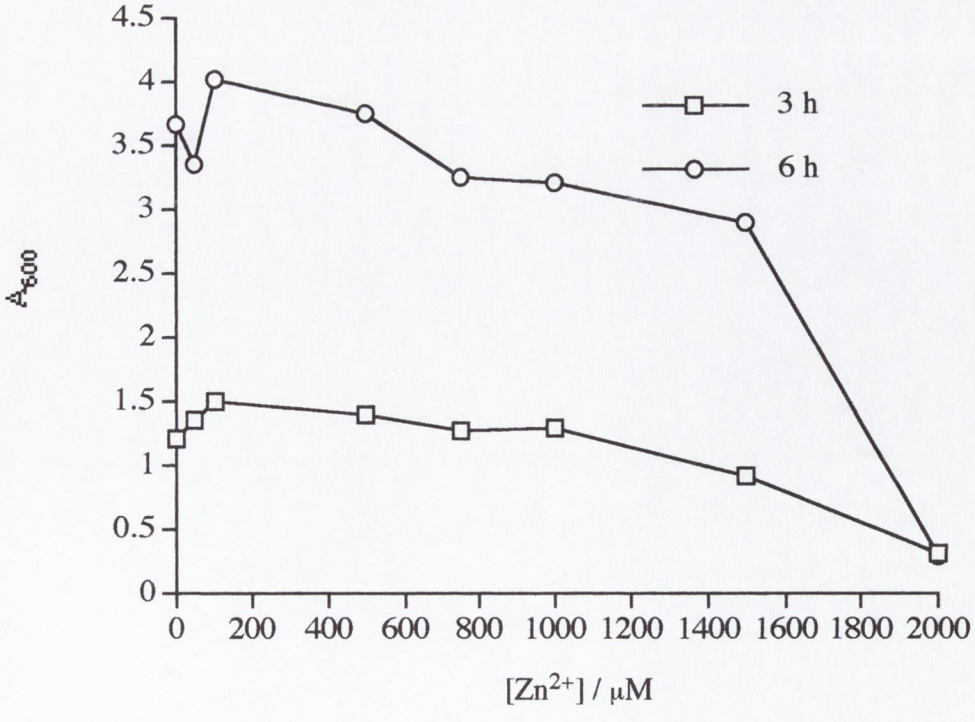


Figure 5.3.6(b)

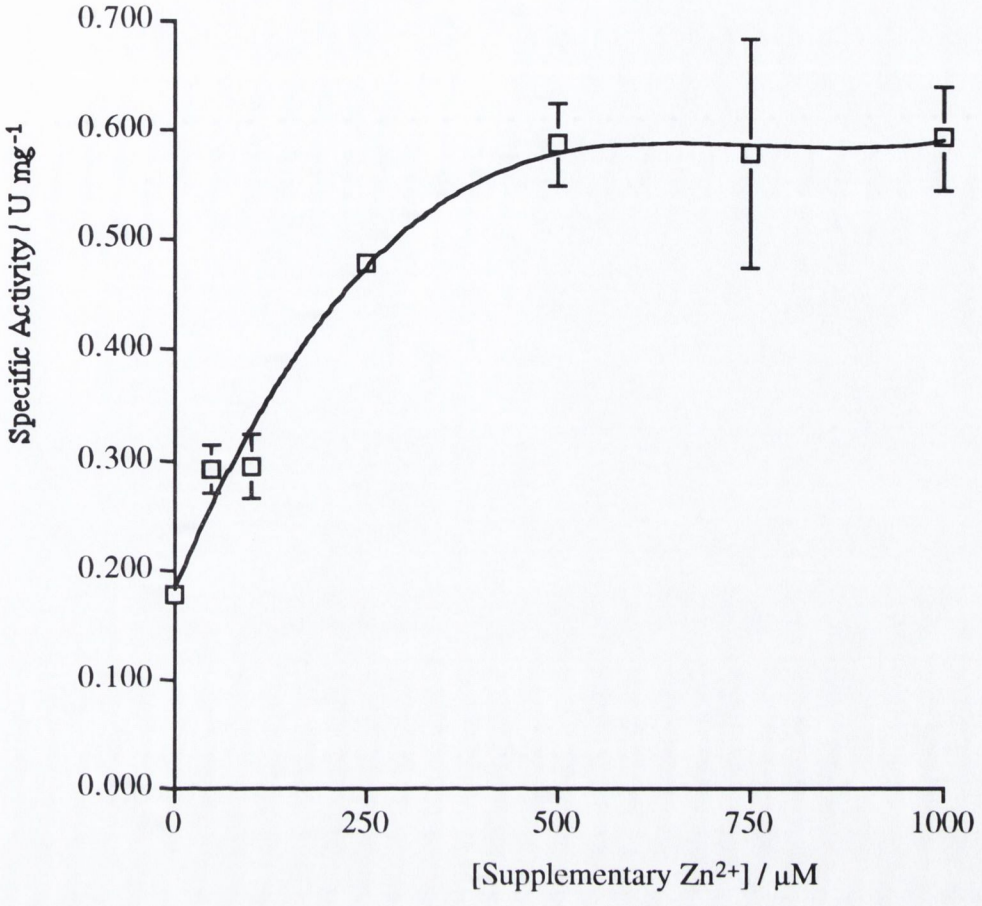


Figure 5.3.7 **Specific Activity of ADH in Cell lysates as a Function of the Concentration of Supplementary Zinc Ions in the Culture Medium**

Medium was inoculated with a cryoculture of *E. coli* pLysS transformants and dispensed (5 ml aliquots) into 25 ml tubes. Each tube was supplemented with a different concentration of $ZnSO_4$. The cultures were grown for 2 h before and for 4 h after the addition of IPTG. Cells were harvested from 3 ml aliquots of each culture, lysed and the specific activity of each post-sonication supernanant was determined immediately. Each result represents the mean $\pm$ S.D. of three enzyme assays performed on each lysate. Where no error bar is shown the standard error of the mean is less than the representation of the point.

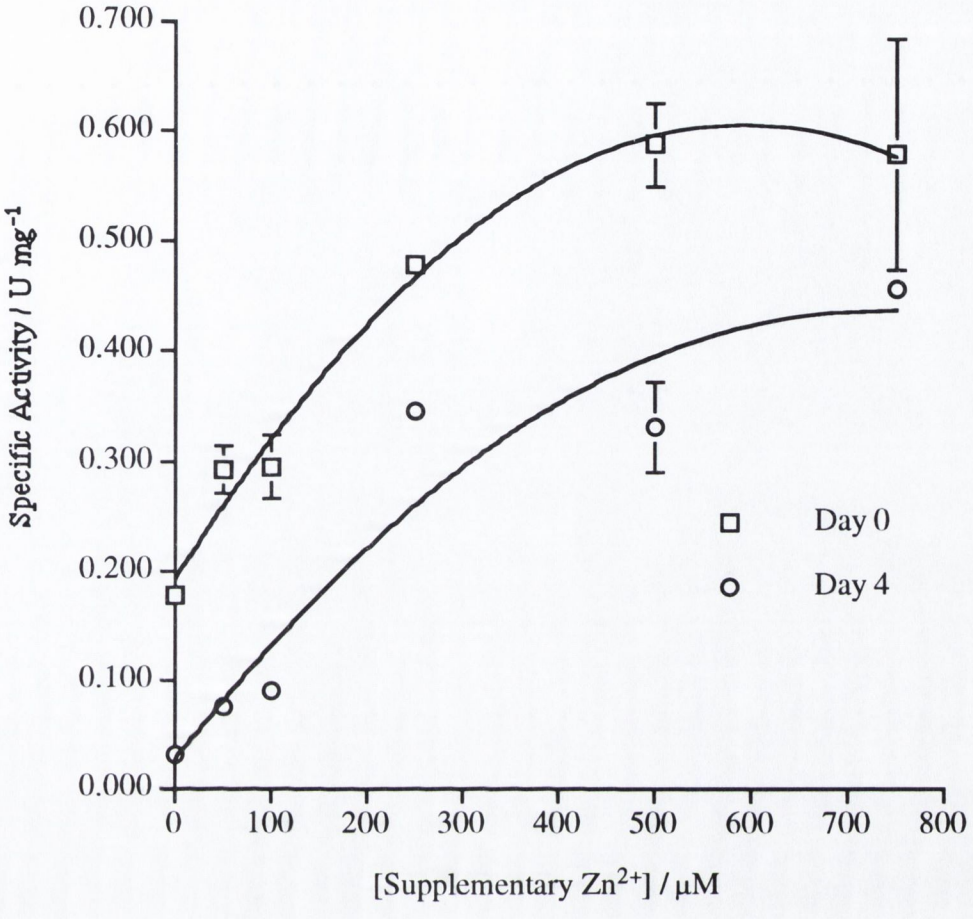


Figure 5.3.8 **<u>The Affect of Supplementary Zinc Ions of the Stability of ADH Enzyme Activity</u>**

Cultures were grown in the presence and absence of various concentrations of supplementary $ZnSO_4$, cells were harvested, lysed and the specific activity of each lysate was determined as described in the legend to Figure 5.3.7. The lysates were stored at 4 °C for five days and reassayed. Each result represents the mean ± S.D. of three enzyme assays performed on each lysate. Where no error bar is shown the standard error of the mean is less than the representation of the point.

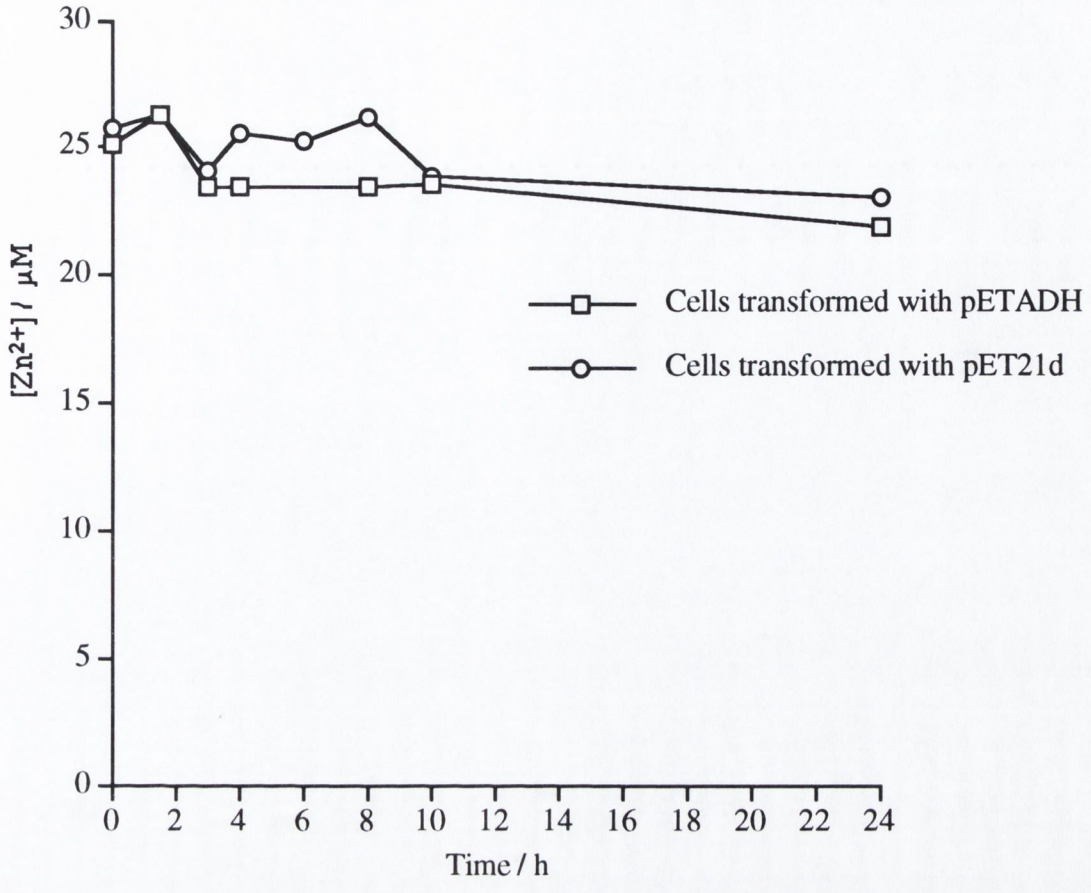
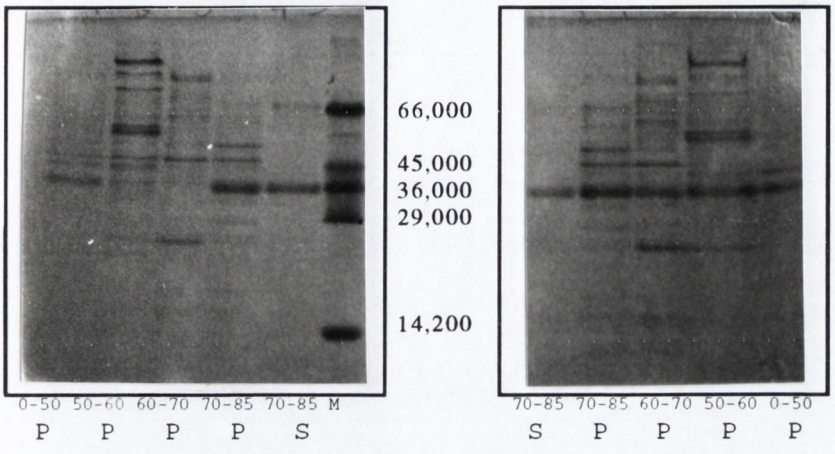


Figure 5.3.9 **The Concentration of Zn Ions in Aliquots of Medium Removed from Cultures of *E. coli* Transformants**

E. coli pLys cells transformed with pETADH or with pET21d were grown as 300 ml cultures in 2TY medium. IPTG was added to both cultures at 3 h. Samples (10 ml) were removed from both culture flasks immediately after the addition of inoculum (zero time) and at various other times over a total time-period of 24 h by pouring directly into polyethylene capped centrifuge tubes. Cells were removed from the media by centrifugation (2,700 ***g***, 20 min, 4 °C), the cell-free supernatants were decanted into polyethylene tubes and the concentration of Zn^{2+} was measured by atomic absorption spectroscopy.

Figure 5.3.10 **Ammonium Sulfate Fractionations of Cell Lysates Prepared from *E. coli* Transformants Grown in the Presence and Absence of Supplementary Zinc Ions**

Transformants (500 ml cultures) were induced to express rADH in the presence (+Zn lysate) and in the absence (-Zn lysate) of 750 μM supplementary Zn^{2+}. The cells from each culture were resuspended in 26 ml of 100 mM potassium phosphate pH 6.1, lysed by sonication, and the entire post-sonication supernatant was subjected to ammonium sulfate fractionation, (0 - 50%, 50 - 60% *etc.*). The precipitate at each stage of the fractionation was resuspended (volumes indicated in the table) in 50 mM potassium phosphate, pH 7, and the entire supernatant was carried through to the next step. The protein concentration of each resuspended precipitate and final supernatant was estimated by determining the A_{280} value. A volume of each fraction, equivalent to 30 μg of total protein, was analysed by SDS-PAGE. Ammonium sulfate precipitates are indicated by P, supernatants by S, M indicates subunit M_r markers (5 μg of each), n.d. indicates values that were not determined.



+Zn lysate

-Zn lysate

Fraction	Volume ml	[Protein] mg ml^{-1}
Post-sonication supernatant	27	n.d.
0 - 50% pellet	1	7.9
50 - 60% pellet	4	7.1
60 - 70% pellet	10	12.24
70 - 85% pellet	5	31.54
70 - 85% supernatant	100	2.53

Fraction	Volume ml	[Protein] mg ml^{-1}
Post-sonication supernatant	30	n.d.
0 - 50% pellet	1	13.5
50 - 60% pellet	4	13.84
60 - 70% pellet	10	35.42
70 - 85% pellet	5	35.1
70 - 85% supernatant	100	2.56

Figure 5.3.10

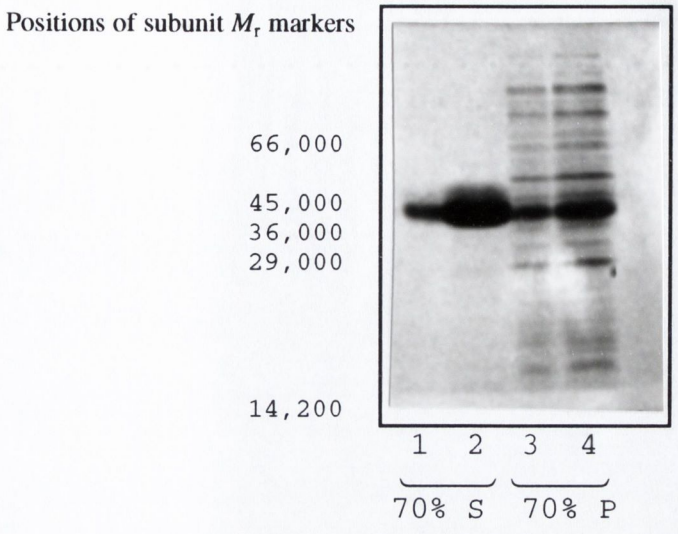


Figure 5.3.11 **Purification of Recombinant ADH by Ammonium Sulfate Fractionation**

Cells from a 200 ml culture (750 μM supplementary Zn^{2+}) were lysed by passage through a French pressure cell. The entire post-lysis supernatant (15 ml) was adjusted to 70% saturation ammonium sulfate by the addition of 35 ml of 100% saturated ammonium sulfate. The solution was stirred on ice for 45 min and centrifuged (43,000 ***g***, 15 min, 4 °C). The pellet was resuspended in 8 ml of 50 mM potassium phosphate, pH 7.0, the volume of the supernatant was 43 ml. The protein concentration of each fraction was measured by the Lowry assay. Lanes 1 and 2 contained aliquots of the 70% supernatant, 4.6 μg and 23 μg respectively. Lanes 3 and 4 contained aliquots of the resuspended pellets, 20 μg and 40 μg respectively.

Table 5.3.3 **Summary of the Purification of Recombinant ADH by Ammonium Sulfate Fractionation**

Purification of rADH by ammonium sulfate fractionation was exactly as described in Section 5.2.5. Results presented for protein concentrations represent the mean ± S.D. of three Lowry assays performed on each sample. Results presented for specific activities represent the mean ± S.D. of three enzyme assays performed on each lysate. Recoveries are expressed as % recovery of the specific activity of the French pressure cell lysates.

Fraction	Volume ml	[Protein] mg ml^{-1}	Total protein mg	Specific activity U mg^{-1}	Total activity U	Recovery %
Cell lysate	20	6.5 ± 0.39	130	0.66 ± 0.06	86	100
70% saturation ammonium sulfate supernatant	59	0.59 ± 0.02	35	2.0 ± 0.14	70	81
85% saturation ammonium sulfate pellet	3	4.7 ± 0.14	14	2.3 ± 0.20	32	37

Table 5.3.3(a) +Zn Culture

Fraction	Volume ml	[Protein] mg ml^{-1}	Total protein mg	Specific activity U mg^{-1}	Total activity U	Recovery %
Cell lysate	20	8.3 ± 0.29	166	0.09 ± 0.00	15.33	100
70% saturation ammonium sulfate supernatant	60	0.44 ± 0.02	26	0.66 ± 0.06	17.33	113
85% saturation ammonium sulfate pellet	3	2.0 ± 0.10	6	0.57 ± 0.02	3	20

Table 5.3.3(b) -Zn Culture

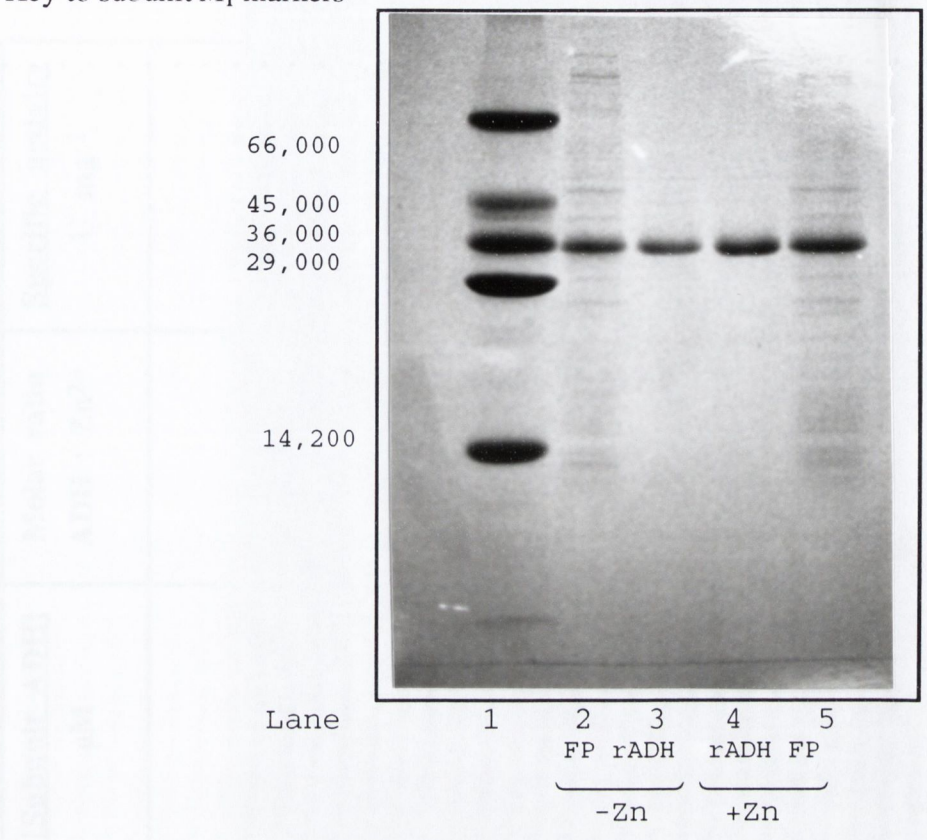


Figure 5.3.12 **SDS-PAGE Analysis of Purified Recombinant ADH**

Purifed -Zn rADH (Lane 3; 4 μg) and purifed +Zn rADH (Lane 4; 6 μg) were the resuspended 85% saturation ammonium sulfate pellets described in Table 5.3.3. Lanes 2 and 5 contained aliquots of the clarified supernatants obtained after the French pressure cell step (FP). SDS-PAGE analysis on a 15% (w/v) acrylamide resolving gel was essentially as described in Section 5.2.11 however, a 110 mm long gel was used (gel mould dimensions 110 x 140 x 1 mm) and electrophoresis was continued until the tracker dye ran off the end of the gel. Lane 1 contained subunit M_r markers, 10 μg of each.

Solution or enzyme preparation	[Zn ion] μM	[Protein] mg ml^{-1}	[Subunit ADH] μM	Molar ratio ADH : Zn^{2+}	Specific activity U mg^{-1}
Cell pellet wash buffer (100 mM potassium phosphate, pH 7.0)	0.76 ± 0.09				
Modified lysis-buffer	0.75 ± 0.14				
Saturated $(NH_4)_2SO_4$ solution	7.0 ± 0.7				
Final protein pellet resuspension buffer (100 mM potassium phosphate, pH 7.0)	0.76 ± 0.09				
-Zn rADH	57 ± 1	2.0 ± 0.10	56	1 : 1.0	0.57 ± 0.02
+Zn rADH	195 ± 3.8	4.7 ± 0.14	129	1 : 1.50	2.3 ± 0.20

Table 5.3.4 **The Zinc Content of Recombinant ADH and of Solutions used in the Purification**

The molar concentration of subunit rADH was determined using a subunit M_r of 36,205 (Section 3.4.2.1). Protein concentration was determined by the Lowry assay. Saturated $(NH_4)_2SO_4$ solution was diluted two-fold, the preparation of -Zn rADH was diluted twenty-fold and the preparation of +Zn rADH was diluted fifty-fold for the atomic absorption measurements. All dilutions were into distilled water. Each result represents the mean ± S.D. of three determinations performed on each sample.

Culture	[Zn ion] μM	[Protein] mg ml^{-1}	Zn ion content nmol mg^{-1}
pET21d $-Zn^{2+}$	14.3	7.9	1.8
pET21d $+Zn^{2+}$	41.7	11.8	3.5
pETADH $-Zn^{2+}$	65	12	5.5
pETADH $+Zn^{2+}$	74	6.5	11.4

Table 5.3.5 **Zinc Ion Content of *E. coli* Cell Lysates**

Figure 5.3.13 **Michaelis-Menten Plots of Initial Velocity Data for Recombinant ADH**

Initial reaction velocity was determined by monitoring Δ fluorescence intensity s^{-1} with assay conditions as described in Section 5.2.13. Values of K_m and V_{max} were obtained for each enzyme preparation by non-linear regression data analysis of the experimental data. The values K_m and V_{max} were then used to generate theoretical Michaelis-Menten plots. The experimental data are displayed superimposed on the theoretical Michaelis-Menten plots. Each result represents the mean initial velocity $\pm$ S.D. calculated from the results of three assays performed at each substrate concentration. Where no error bar is shown the error was less than the representation of the point.

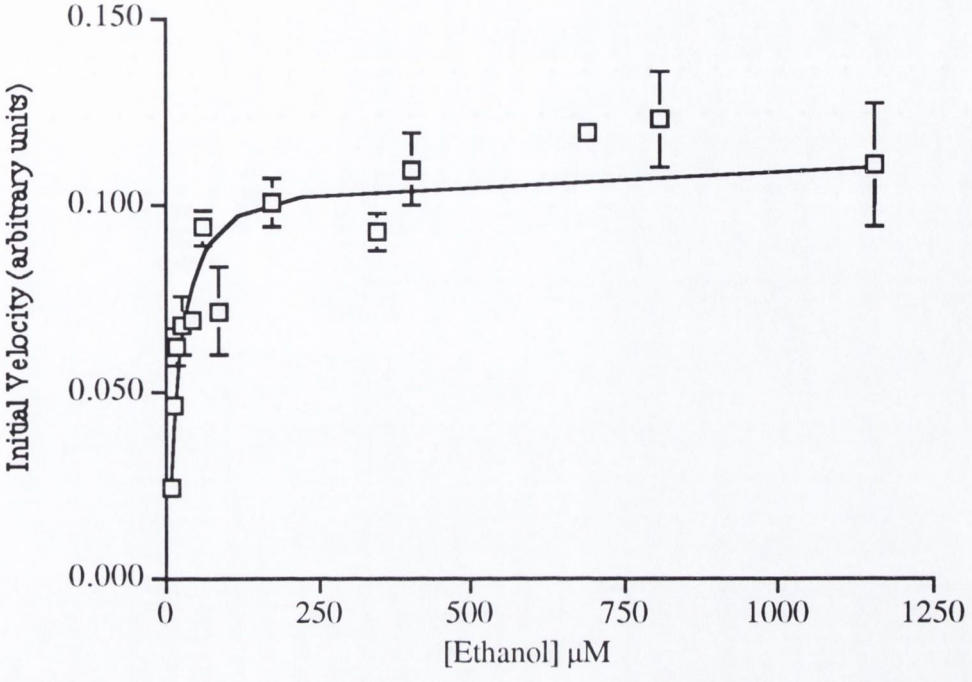


Figure 5.3.13(a) **+Zn rADH**

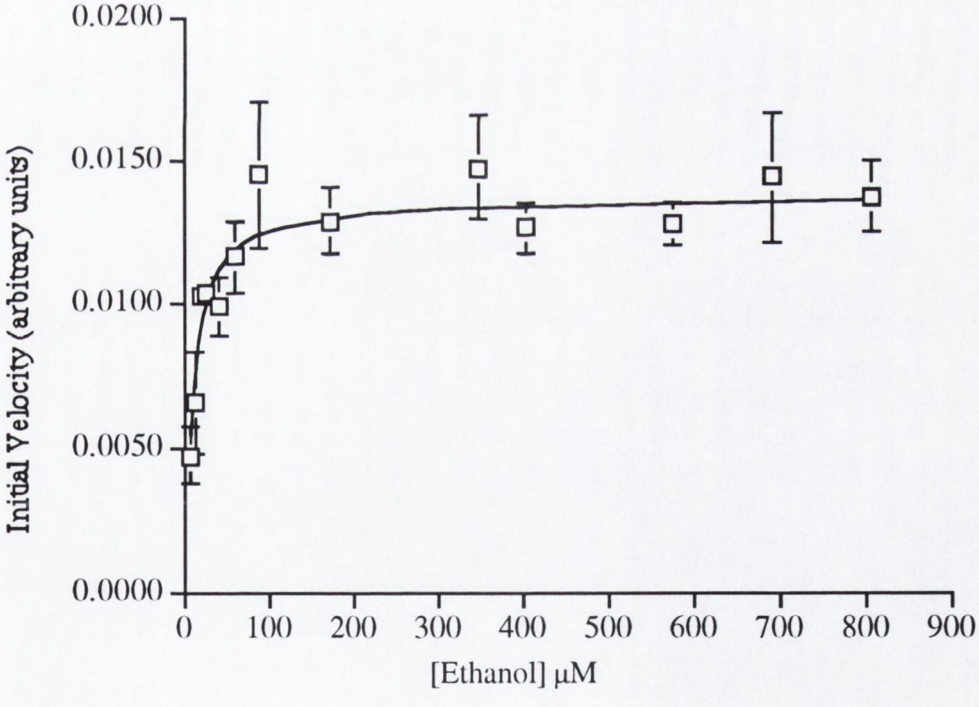


Figure 5.3.13(b) **-Zn rADH**

Figure 5.3.14 **Hanes Plots of Initial Velocity Data for Recombinant ADH**

Initial reaction velocity was determined, and values of K_m and V_{max} were obtained as described in the legend to Figure 5.3.13. The values K_m and V_{max} were then used to generate theoretical Hanes plots and the experimental data are displayed superimposed on the theoretical Hanes plots. Each result represents the mean initial velocity $\pm$ S.D. calculated from the results of three assays performed at each substrate concentration. Where no error bar is shown the error was less than the representation of the point.

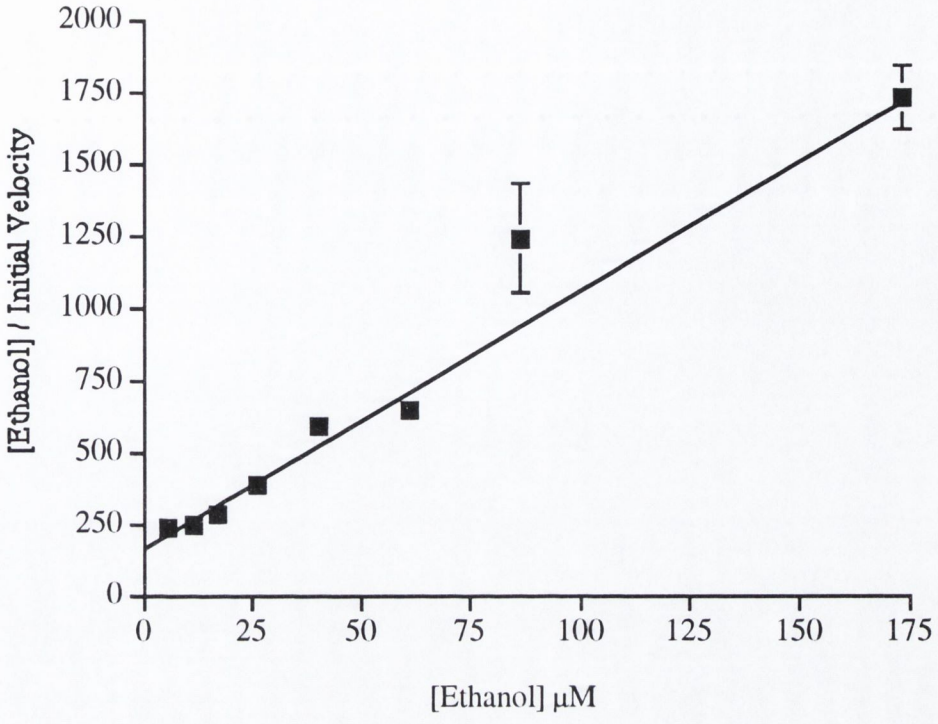


Figure 5.3.14(a) **+Zn rADH**

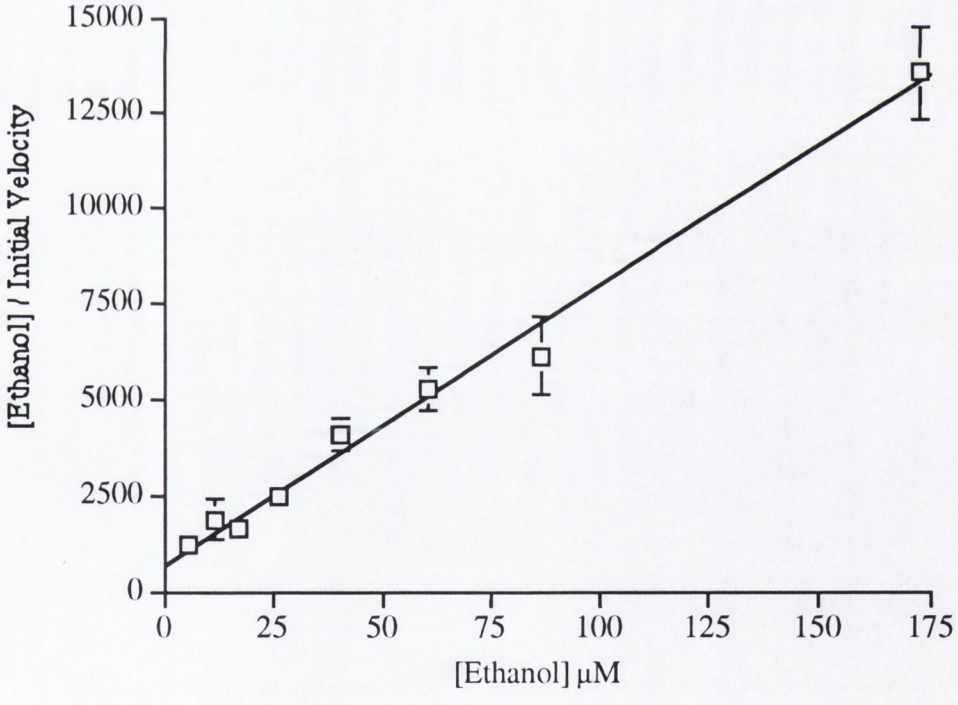


Figure 5.3.14(b) **-Zn rADH**

Enzyme	K_m (μM)	V_{max} (Δ I s^{-1} μl^{-1})	k_{cat} (s^{-1})
-Zn rADH	9.6 ± 1.4	0.0137 ± 0.0003	0.197
+Zn rADH	18.4 ± 2.4	0.113 ± 0.003	1.68
ADH 2334	82		1.3

Table 5.3.6 **Kinetic Parameters for rADH with Ethanol as Substrate**

The kinetic parameters for the wild-type enzyme ADH 2334 are those determined by Sheehan *et al.* (1988). The values of K_m and V_{max} for the two recombinant enzyme preparations -Zn rADH and +Zn rADH were obtained from the experimental data displayed in Figure 5.3.13. Values of k_{cat} for -Zn rADH and +Zn rADH were determined as described in Section 5.2.16.

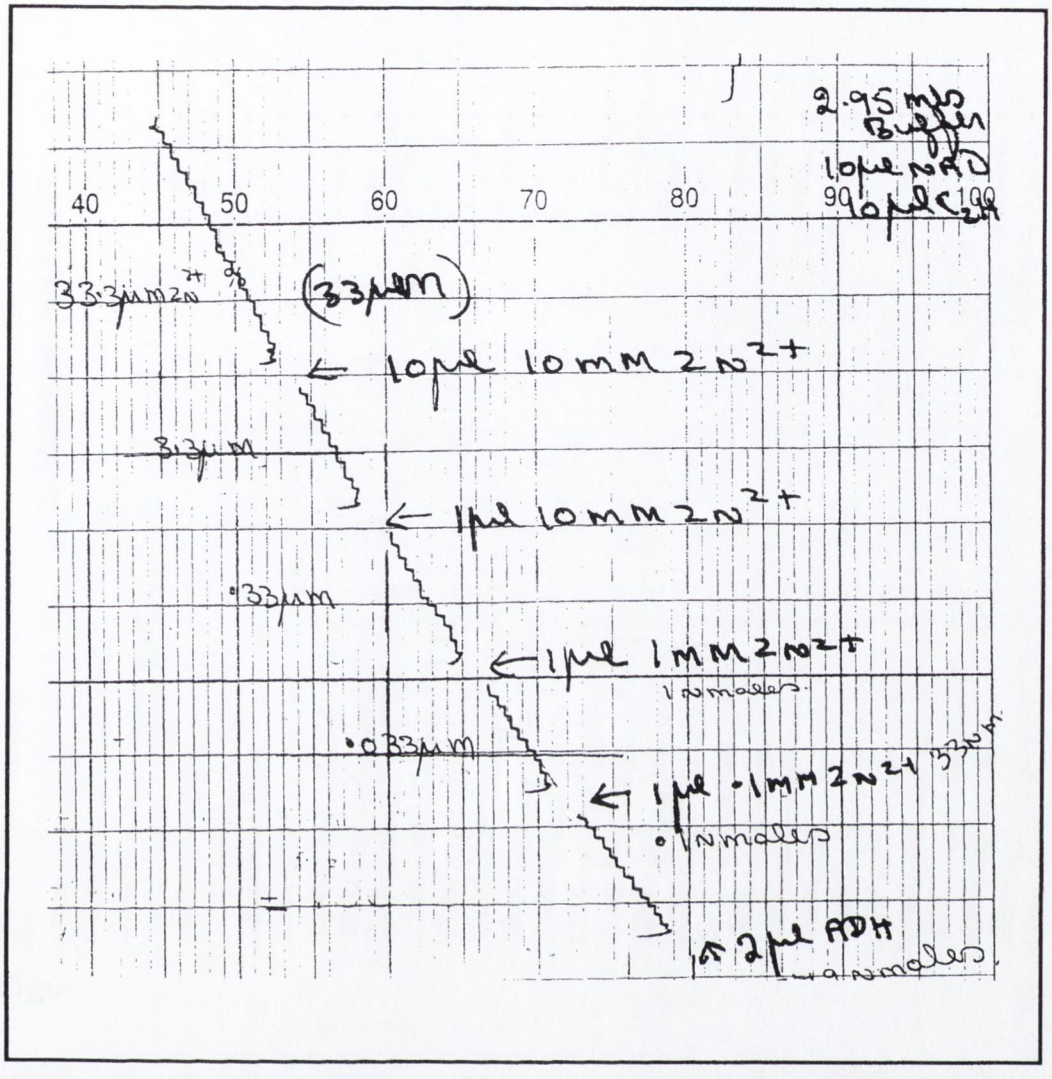


Figure 5.3.15 **The Addition of Zinc Ions to the -Zn rADH Assay Mixture**

The assay mixture was potassium phosphate buffer (100 mM, pH 7.0, 2.95 ml), NAD^+ (150 mM, 10 µl) and ethanol (10 µl). Assay temperature was 37 °C. Enzyme (2 µl ADH) was added and the increase in absorbance at 340 nm was continuously monitored as the concentration of Zn^{2+} in the cuvette was increased by the cumulative addition of $ZnSO_4$. On the facsimile 1 cm in the vertical direction corresponds to 1 min and 1 cm in the horizontal direction corresponds to Δ absorbance of 0.01.

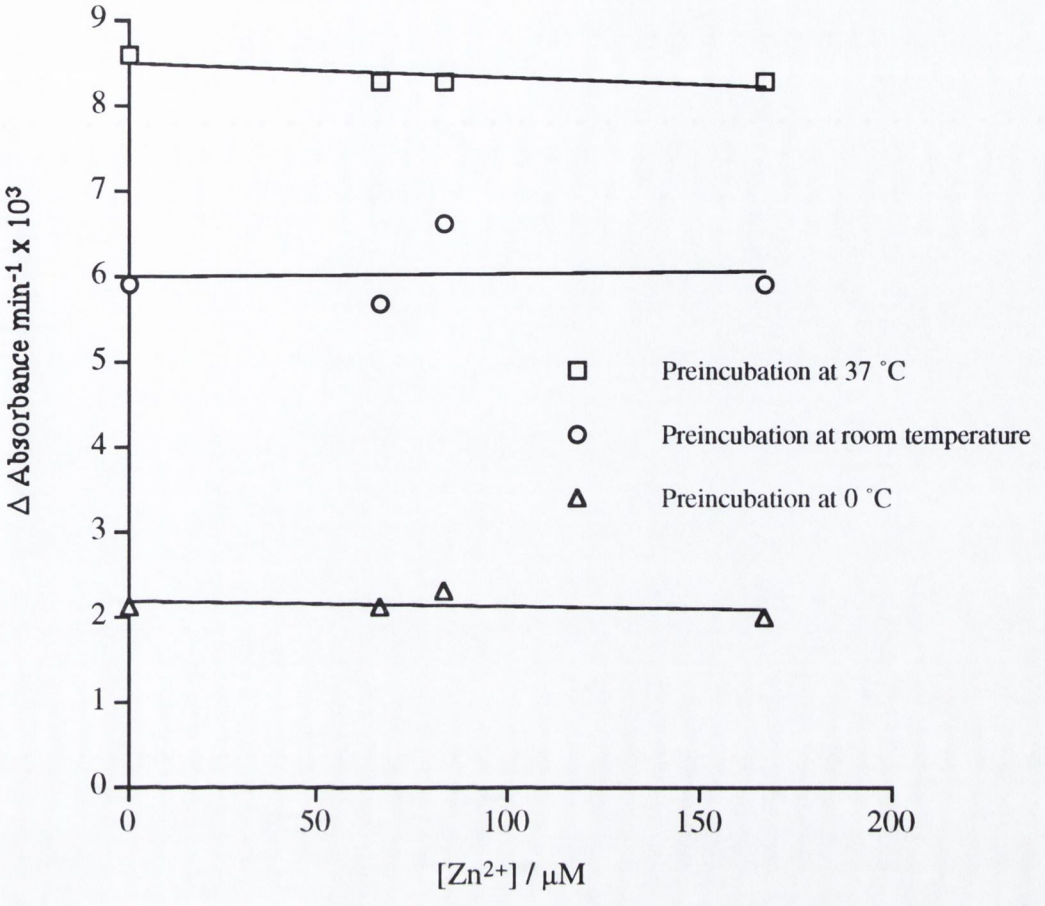


Figure 5.3.16 **Enzyme Activity of -Zn rADH After Preincubation of the Enzyme with Zinc Ions**

Aliquots of -Zn rADH were incubated in the presence of different concentrations of $ZnSO_4$ and at three different temperatures as described in Section 5.3.10. Each incubation contained 2 µl of -Zn rADH (113 pmol of subunit rADH), 1 - 4 µl of 0.1 mM or 0.5 mM $ZnSO_4$ and the appropriate volume of H_2O for a final volume of 6 µl. The enzyme activity of 5 µl of each preincubation mixture was determined at 37 °C and is presented as a function of the concentration of Zn^{2+} in the preincubation mixture.

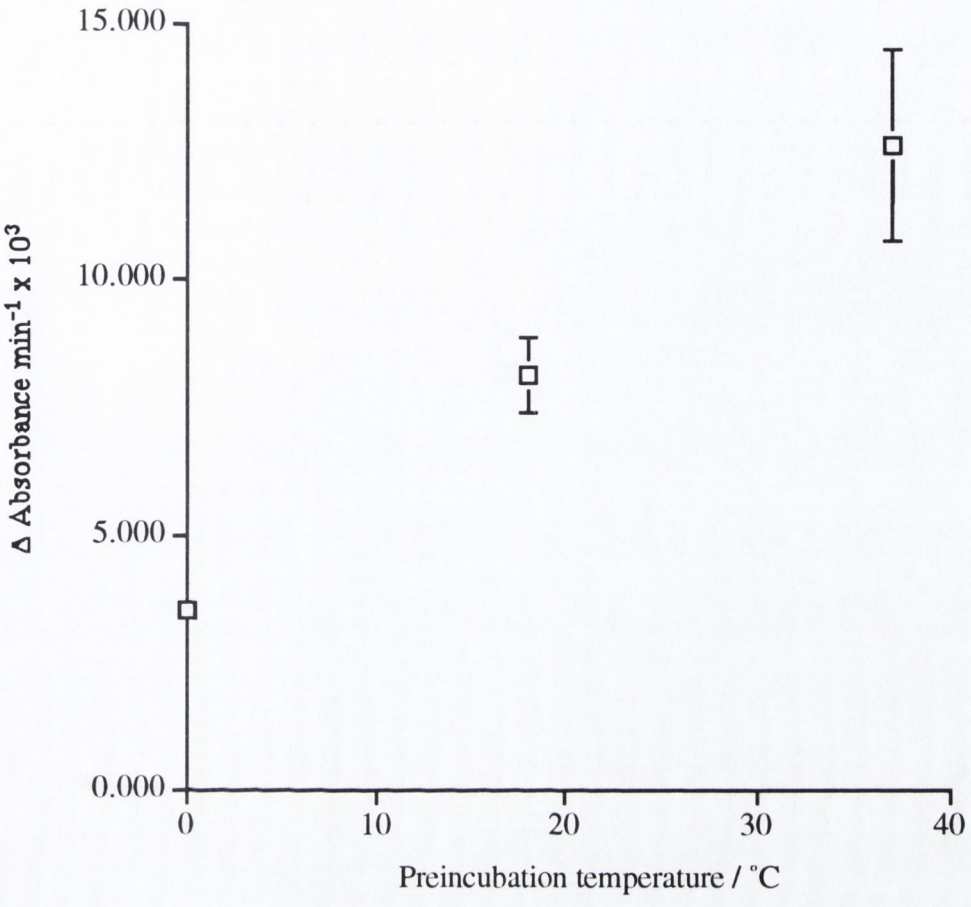


Figure 5.3.17 **Enzyme Activity of +Zn rADH After Preincubation of the Enzyme at Different Temperatures**

Aliquots of +Zn rADH were preincubated for 10 min at three different temperatures (on ice, at 18 °C and at 37 °C). Each incubation contained 4 µl (23.2 µg) of +Zn rADH and 8 µl of 100 mM potassium phosphate, 0.5 mM dithiothreitol, pH 7.0 that had been pre-equilibrated to the desired temperature. The enzyme activity of 10 µl of each preincubation mixture was determined at 37 °C, in a final volume of 3 ml. The assay buffer was 0.5 mM in dithiothreitol. Each result represents the mean ± S.D. of the results of three preincubations at each temperature. Where no error bar is shown the standard error of the mean is less than the representation of the point.

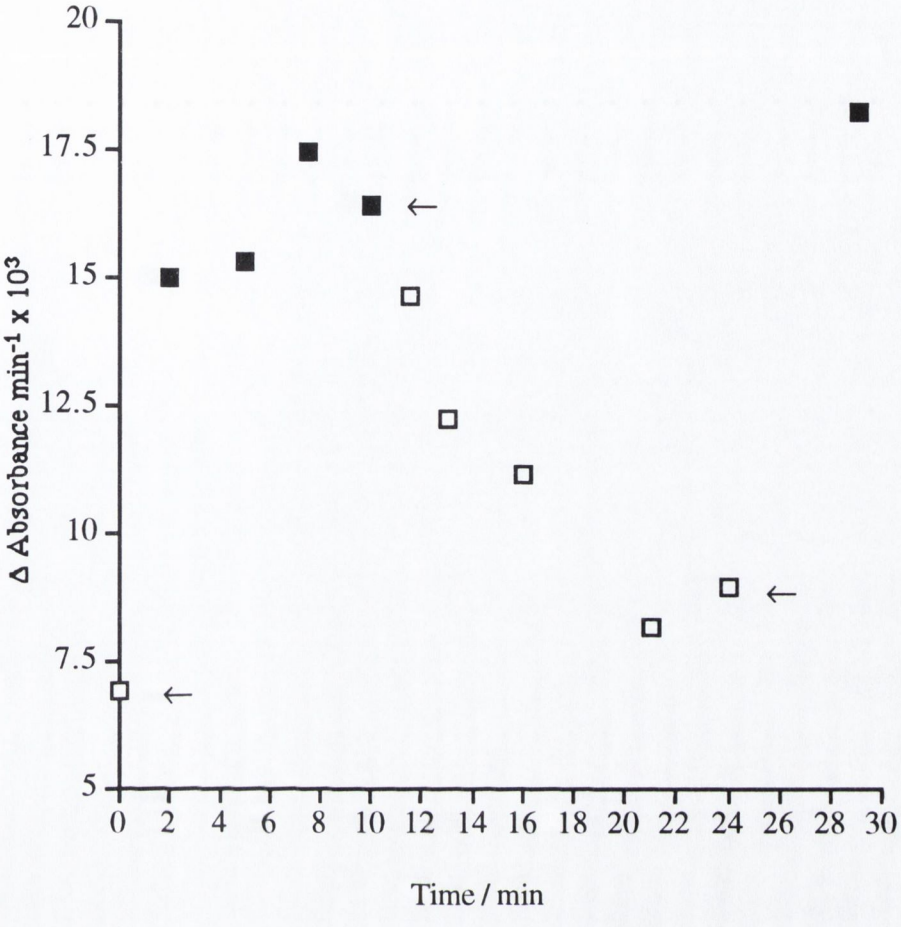


Figure 5.3.18 **<u>The Change in Enzyme Activity of +Zn rADH as the Preincubation Temperature was Varied</u>**

A single aliquot of +Zn rADH (40 µl, 232 µg) was mixed with 80 µl of ice-cold 100 mM potassium phosphate, 0.5 mM dithiothreitol, pH 7.0 and preincubated on ice for 14 min, 10 µl was removed and assayed at 37 °C (zero time, arrowed). The preincubation mixture was immediately moved to a 37 °C water-bath and 10 µl aliquots were removed and assayed at various times (filled squares). After 10 min (arrowed) the preincubation was returned to ice and the activity was monitored for a further 14 min (open squares). Finally, the preincubation was returned to the 37 °C water-bath (arrowed) and assayed again after 5 min (filled square). All assays were at 37 °C, in a final volume of 3 ml and the assay buffer was 0.5 mM in dithiothreitol.

5.4 DISCUSSION

5.4.1 Expression from pETADH in the Absence and Presence of Supplementary Zinc Ions, Specific Activity of the Enzymes Produced

The objective of the work described in this chapter was the production of intact, functional rADH, and the expression system had been carefully designed to that end. The results of initial experiments to establish optimum conditions for IPTG-dependent expression were dramatic demonstrations of the efficiency and selectivity of the T7 polymerase expression system. Four hours after the addition of IPTG at least 30% of the protein produced by *E. coli* was recombinant protein. SDS-PAGE analysis of a post-sonication supernatant demonstrated that the rADH produced, or at least a large amount of it, was soluble protein. SDS-PAGE analysis of an ammonium sulfate fractionation of the post-sonication supernatant revealed that the 65 - 80% ammonium sulfate pellet contained rADH with only very minor contamination from *E. coli* protein, 10 mg of 'almost pure' protein being obtained from a 200-ml culture.

When the expression host was grown in standard 2TY medium the high yield of recombinant protein was not accompanied by a high yield of enzyme activity. The specific activity of 'almost pure' rADH in the 65 - 80% ammonium sulfate pellet was 0.09 U mg^{-1}, approximately one twentieth of the value reported for the wild-type protein. The *adh*2334 nucleotide sequence data had confirmed the presence of both the catalytic-zinc and the second-zinc ligands (Chapter 3). rADH was expected to be a zinc-dependent enzyme with a subunit : Zn^{2+} stoichiometry of 1 : 2, therefore, the possibility that rADH enzyme activity was limited by the concentration of available Zn^{2+} was investigated. The effect of supplementing culture media with Zn^{2+} was unequivocal: the specific activity of rADH in cell lysates increased, up to a constant value, as the concentration of Zn^{2+} in the culture medium was increased. The maximum effect of Zn^{2+} supplementation was achieved at 500 μM supplementary Zn^{2+}.

There are a number of possibilities as to how the observed increase in the specific activity of rADH in cell lysates could occur. Firstly, Zn^{2+} supplementation could increase the intrinsic biological activity of the enzyme, without affecting either the concentration of rADH protein, or the concentration of any other protein. Secondly, Zn^{2+} supplementation could increase the expression of rADH, without affecting the biological activity of the enzyme. Thirdly, the intrinsic biological activity of the enzyme was independent of Zn^{2+} supplementation, but high concentrations of Zn^{2+} in the medium prevented degradation of the enzyme. In the absence of supplementation, the specific activity of the lysate was lower because the active-

enzyme concentration was lower. Finally, Zn^{2+} supplementation could increase the specific activity of rADH by suppressing the synthesis of other proteins.

When the recombinant protein was purified to homogeneity from a culture that was 500 μM in supplementary Zn^{2+} the specific activity of the pure protein was increased four-fold relative to the protein purified from a culture that contained no supplementary Zn^{2+}. That result confirmed that the inclusion of Zn^{2+} in the culture medium did cause an increase in the enzyme activity of rADH, independent of any other affect that Zn^{2+} may have had on expression levels, protein degradation or protein synthesis.

When discussing expression levels (and ignoring protein degradation), a distinction must be made between the total amount of recombinant protein expressed and the amount of recombinant protein present in the soluble fraction. The two will not be equivalent if inclusion bodies are formed. SDS-PAGE and Coomassie Blue staining of lysates, which were prepared by boiling equivalent numbers of cells in SDS demonstrated, in a qualitative manner, that expression of rADH was increased when the medium was supplemented with Zn^{2+}. Comparisons, on a per cell basis, of the total amount of soluble protein produced in the presence and absence of supplementary Zn^{2+} demonstrated that more soluble protein was produced in the presence of supplementary Zn^{2+}. The increase was small, 1.2 - 1.3-fold.

The total amount of (soluble) rADH produced in the presence of 500 μM supplementary Zn^{2+} was calculated to be 143 mg per litre of culture. The available data for expression of soluble rADH in the absence of supplementary Zn^{2+} (at least 87 mg per litre of culture) represented the amount of rADH recovered in the 70% ammonium sulfate supernatant rather than the total amount present in the cell lysate and the two figures are not therefore directly comparable.

If the higher specific activity of the +Zn lysate relative to the -Zn lysate was due only to the fact that high concentrations of Zn^{2+} in the medium protected the enzyme against degradation the +Zn enzyme and the -Zn enzyme would be expected to have the same specific activity when purified, however, that was not what was observed. With respect to Zn^{2+} suppressing the synthesis of endogenous *E. coli* proteins, it seems unlikely that *E. coli* that had experienced a 3.3-fold reduction in cellular protein could grow and reproduce normally.

5.4.2 Heterogeneity in the Recombinant Protein Produced in the Absence of Supplementary Zinc Ions

The 36,000-M_r protein, present in crude cell lysates derived from cultures grown in the absence of supplementary Zn^{2+}, precipitated at a number of different ammonium sulfate concentrations. This anomalous behavior, which was not observed when a cell lysate derived from a culture grown in the presence of

supplementary Zn^{2+} was subjected to the same fractionation scheme, suggested that the recombinant protein produced in the absence of supplementary Zn^{2+} was not homogeneous. The determination of the N-terminal amino-acid sequence of the recombinant protein expressed from pETADH was reported in Chapter 4. The recombinant protein that was sequenced was from an electroblot of a cell lysate derived from *E. coli* transformants grown in the absence of supplementary Zn^{2+} and was, therefore, a mixture of the material subsequently found to fractionate at different ammonium sulfate concentrations. The sequence data gave a single, homogeneous, amino-acid sequence, confirming that the 36,000-M_r proteins that precipitated at different ammonium sulfate concentrations had identical electrophoretic mobilities in the presence of SDS and identical N-terminal amino-acid sequences.

The distribution of -Zn rADH across all ammonium sulfate fractions may be a reflection of heterogeneity at the level of metal ion content. ADH activity was detected in all of the fractions that contained the 36,000-M_r protein with the exception of the 0 - 30% pellet. The 0 - 30% pellet may therefore represent rADH that is completely devoid of Zn^{2+}, the other fractions containing rADH molecules with differing Zn^{2+} contents. Differing Zn^{2+} contents may result in different protein conformations and/or different isoelectric points. The solubility of a protein in ammonium sulfate solution is dependent, in part, on the isoelectric point of the protein, solubility approaching a minimum at a pH equivalent to the isoelectric point of the protein (Dixon & Webb, 1961). Varying the stoichiometry of Zn^{2+} in the protein may cause a shift in the isoelectric point which causes protein molecules with differing Zn^{2+} stoichiometry to precipitate at different ammonium sulfate concentrations. If this is the case it may be possible to separate the different 'forms' of rADH by ion-exchange chromatography.

The literature contains reports of heterogeneity observed during the expression of iron-dependent proteins. Over-expression of *Desulfovibrio gigas* desulforedoxin resulted in two forms of the enzyme that separated during anion-exchange chromatography (Czaga *et al.*, 1995). Both enzymes were dimers, with identical subunit M_rs and identical N-terminal amino-acid sequences. The heterogeneity was at the level of the metal ion content. In one form the metal-binding sites of both subunits were occupied with Fe^{3+}, this being the only wild-type form that has been identified to date whereas the other was a mixed-metal species that contained half an equivalent each of iron and zinc per subunit. Over-expression of recombinant *Clostridium pasteurianum* rubredoxin (Eidsness *et al.*, 1992) also resulted in a mixture of species. 30% of the expressed protein had the expected wild-type subunit : Fe stoichiometry of 1 : 1 whilst the remainder contained no iron but displayed a subunit : Zn^{2+} stoichiometry of 1 : 1.

5.4.3 The Zinc Content of Recombinant ADH Produced in the Absence and Presence of Supplementary Zinc Ions

The specific activity of the enzyme purified from medium supplemented with Zn^{2+} was the 'correct' specific activity in that it was the same as the specific activity of the wild-type enzyme. The problem that had to be addressed was why the enzyme produced in the absence of added Zn^{2+} had a low specific activity. The most obvious explanation was that the enzyme produced by cells grown in unsupplemented medium was deficient in Zn^{2+}. The subunit rADH : Zn^{2+} stoichiometry of the -Zn rADH preparation was determined by atomic absorption spectroscopy to be 1 : 1.0, a value consistent with the production, by *E. coli*, of zinc-deficient rADH. The subunit rADH : Zn^{2+} stoichiometry determined for the +Zn rADH preparation was 1 : 1.5, a value higher than that obtained for the -Zn rADH preparation. The different specific activities of the two enzymes can therefore be rationalised in terms of the different Zn^{2+} contents of the two enzymes.

Values reported for the Zn^{2+} content of HLADH determined by atomic absorption spectroscopy range from 1.5 - 2 Zn^{2+} ions per subunit [Drum *et al.*, 1969(b)] and the value obtained for the +Zn rADH preparation was within that range. Values obtained for the Zn^{2+} content of enzymes in solution vary with the purification procedure used and with the method of determination of protein concentration (Drum *et al.*, 1969(b); Bosron *et al.*, 1975). Bosron *et al.* (1975) report a 25% - 30% reduction in the Zn^{2+} content of alkaline phosphatase when ammonium sulfate is included in the purification procedure and suggest that loss of Zn^{2+} from the protein is due to the formation of a $Zn(NH_3)_4^{2+}$ complex. The -Zn rADH and +Zn rADH preparations were prepared in an identical manner and therefore any removal of Zn^{2+} by ammonium sulfate is unlikely to have affected the relative difference in the Zn^{2+} contents of the two preparations. Phosphate buffer (0.1M, pH 7.0) has been shown not to affect the Zn^{2+} content of purified horse liver ADH (Gerschitz *et al.*, 1978).

The Zn^{2+} contents of the rADH preparations were based on protein concentrations that had been determined by the Lowry assay. A more reliable method of calculating protein concentrations uses values of absorbance at 280 nm and the extinction coefficient of the protein determined from the amino-acid composition (Gill & von Hippel, 1989), and this was attempted. The extinction coefficient of rADH, calculated on the basis of the tyrosine and tryptophan content and with the assumption that rADH contains no disulfide bonds, was 29,870 M^{-1} cm^{-1}. Unfortunately it was not possible to use values of absorbance at 280 nm to make accurate determinations of protein concentration because both protein preparations exhibited a broad absorption maximum in the region 240 nm - 290 nm (λ_{max} 260 nm), presumably due to contamination by nucleic acid.

5.4.4 Extracellular and Intracellular Concentrations of Zinc Ions During Expression from pETADH

Expression of rADH in the absence of supplementary Zn^{2+} was at the level of at least 87 mg l^{-1} and was definitely not higher than 143 mg l^{-1} (the level of expression in the presence of supplementary Zn^{2+}). The amount of Zn^{2+} necessary to achieve the correct stoichiometry in the recombinant protein was calculated assuming that -Zn rADH was expressed at the high level of 143 mg l^{-1}. Assuming expression at 143 mg l^{-1} the amount of -Zn rADH produced by a 300-ml culture was 43 mg (1.2 μmol subunit rADH). The *E. coli* culture from which -Zn rADH was purified had an A_{600} value of 3.7 (determined with medium that was diluted thirty-fold). Using the approximation A_{600} value of 1 is equivalent to 10^9 cells ml^{-1} (Mandelstam & McQuillen, 1973) the 300-ml culture contained 1.1 x 10^{12} cells. Therefore, it can be calculated that 1.1 x 10^{-18} mol of rADH was produced per cell, implying a need for 2.2 x 10^{-18} mol of Zn^{2+} per cell. The concentration of Zn^{2+} in unsupplemented medium was determined by atomic absorption spectroscopy to be approximately 25 μM, therefore the amount of Zn^{2+} available (in the medium) to each cell grown in 300 ml of unsupplemented medium can be calculated to be 6.8 x 10^{-18} mol, enough zinc to supply all of the -Zn rADH produced. *E. coli* itself does not require zinc in measurable quantities for growth as large amounts of cells can be grown at Zn^{2+} concentrations < 10^{-8} M (Harris & Coleman, 1968). Given that the production of fully active rADH was not limited by the amount of Zn^{2+} in unsupplemented medium *per se*, the amount of Zn^{2+} inside cells had to be the factor that limited the production of rADH containing the correct amount of zinc.

An average value for the cytoplasmic volume of an *E. coli* cell grown in Tryptone medium is 2.1 x 10^{-15} l (Woldringh & Nanninga, 1985). The amount of rADH produced per cell was 1.1 x 10^{-18} mol and therefore the concentration of rADH in an *E. coli* cell, when the cells were harvested, was 520 μM. Assuming all enzyme molecules were fully saturated with Zn^{2+} the resulting bound-zinc concentration would be 1040 μM

The metal ion content of cells in synchronously growing cultures of *E. coli* B/r has been investigated by Kung *et al.* (1976). Uptake of K^+, Ca^{2+} and Mg^{2+} were observed to occur throughout the cell cycle, therefore, the intracellular concentrations of those ions remained approximately constant. In contrast, Zn^{2+} was taken up by cells before, but not during, log phase. Cell number per ml of culture increased during log phase while the amount of Zn^{2+} per ml of culture remained constant. It follows that intracellular Zn^{2+} concentration must have decreased during log phase. The zinc content of mammalian fibroblast cells has also been shown to be linked to the cell-cycle, with cells taking up Zn^{2+} during proliferation but not during differentiation (Schmidt & Beyersmann, 1999).

Expression of rADH was initiated during log phase and the enzyme was purified from cells that were harvested before the end of log phase. From the data presented by Kung *et al.* (1976) the concentration of Zn^{2+} in a cell entering log phase can be calculated to be 640 μM (1.3×10^{-18} moles of Zn^{2+} per cell) with approximately 20% of the Zn^{2+} being present as free Zn^{2+}. The synthesis of rADH created a demand for 2.2×10^{-18} mol of Zn^{2+} per cell at a time when, according to the results of Kung *et al.* (1976), *E. coli* does not normally take up Zn^{2+} and furthermore, intracellular Zn^{2+} concentration is decreasing. When the extracellular concentration of Zn^{2+} was monitored, during growth of *E. coli* transformants carrying the expression plasmid, extracellular Zn^{2+} concentration remained constant over the time period during which expression of rADH occurs. That result, which is in agreement with the results of Kung *et al.* (1976), demonstrates that the *E. coli* expression host did not respond to the demand for Zn^{2+} caused by the expression of rADH.

Very little is known about mechanisms for Zn^{2+} accumulation and transport in bacteria (Silver & Walderhaug, 1992). With respect to trace-metal ion transport in general an extensive literature exists only for iron (reviewed by Guerinot, 1994). A key factor in the uptake of iron by aerobic bacteria is the production, by the organism, of siderophores, Fe^{3+}-specific chelators that sequester Fe^{3+} from the extracellular environment (Neilands, 1995; Moeck & Coulton, 1998). The possible role of metallothionein in supplying Zn^{2+} to apoforms of mammalian metalloenzymes is discussed by Kägi (1991). Metallothionein has been purified from prokaryotes (Olafson *et al.*, 1988) but as yet its exact role remains to be established. Transport of Zn^{2+}, *via* systems that primarily function as Mg^{2+} transporters, has been demonstrated in *Aerobacter aerogenes* by Webb (1970) and in *Alcaligenes eutrophus* by Nies & Silver (1989). Translocation of Zn^{2+} *via* a high affinity Mn^{2+} system in the cyanobacterium *Synechocystis* has been reported by Bartsevich & Pakrasi (1996).

Energy-dependent transport of $^{65}Zn^{2+}$ in *E. coli* was reported by Bucheder & Broda (1974). The system exhibited saturation kinetics (K_m of 20 μM) and was competitively inhibited by Cd^{2+}. It was argued (Silver, 1978) that a K_m of 20 μM, being comparable to that observed for Mg^{2+} and two orders of magnitude higher than that observed for the trace element Mn^{2+}, was too high to be reasonable for specific transport of a trace ion and represented nonspecific uptake of Zn^{2+} by the Mg^{2+} system. Recently (Patzer & Hantke, 1998) an operon has been identified in *E. coli* that encodes a Zn^{2+}-binding protein, an ATPase and a membrane-binding protein. The regulator protein for the operon was also identified. The *E. coli* Zn^{2+}-binding protein is analogous to a Zn^{2+}-specific transport protein identified in *Haemophilus influenzae* by Lu *et al.* (1997). The *H. influenzae* and *E. coli* Zn^{2+} transport systems, a Mn^{2+} transporter of *Synechocystis* sp. (Bartsevich & Pakrasi, 1995), and

a Zn^{2+} transport system described for *Streptococcus pneumoniae* (Dinthilhac *et al.*, 1997), have been classified as a new family of trace-cation transport systems (Dinthilhac *et al.*, 1997; Patzer & Hantke, 1998).

5.4.5 Addition of Zinc Ions to Purified -Zn Recombinant ADH

The addition of $ZnSO_4$ to the ADH assay-mixture did not cause a stimulation in the enzyme activity of pure -Zn rADH. Similarly, incubation of purified -Zn rADH with $ZnSO_4$, at a range of temperatures, caused no stimulation in enzyme activity that could be attributed to the presence of added Zn^{2+}. The preparation of apoenzymes, by deliberate exposure of purified enzymes to chelators, has been described many times. Among the zinc-dependent enzymes the technique has been described for carbonic anhydrase (Lindskog & Malmström, 1962), alkaline phosphatase (Simpson & Vallee, 1968) and for selective removal, under anaerobic conditions, of the catalytic-site Zn^{2+} from crystalline HLADH (Maret *et al.*, 1979; Schneider *et al.*, 1983). In all of these experiments the removal of Zn^{2+} resulted in loss of enzyme activity, but not gross conformational changes in protein structure. Furthermore, and in all cases, the addition of stoichiometric amounts of Zn^{2+} to the zinc-free enzyme resulted in instantaneous restoration of full biological activity.

The biosynthesis of (wild-type) *E. coli* apo-alkaline phosphatase under zinc-free conditions ($< 10^{-8}$ M Zn^{2+}), was described by Harris & Coleman (1968). The purified apoenzyme was zinc-free and inactive but had a conformation that did not differ significantly from that of the biologically-active species. Addition of Zn^{2+} to the purified apoenzyme, *in vitro*, resulted in instantaneous and complete activation and demonstrated, conclusively, that Zn^{2+} was not involved in the formation of the correct alkaline phosphatase conformation *in vivo*. The reactivation, in solution, of denatured, zinc-free HLADH has been examined in detail by Gerschitz *et al.* (1978) and Rudolph *et al.* (1978). The results of those experiments demonstrated that, in contrast to alkaline phosphatase, ADH requires Zn^{2+} for both catalytic activity and at least *in vitro*, for formation of the conformation leading to the biologically-active enzyme.

The relevance of factors that affect *in vitro* reconstitution of denatured proteins to the *in vivo* protein folding process is a matter of debate (Jaenicke, 1987; Kolb *et al.*, 1995). Experiments *in vitro* necessarily involve refolding that commences from the complete, one dimensional, amino-acid chain and takes place in the absence of the ribosome and accessory proteins. Folding *in vivo* may be co-translational, accompanying the growth of the polypeptide chain from the N- to the C-terminal end. The addition of Zn^{2+} to -Zn rADH differs, in a number of respects, to the activation of HLADH as described by Gerschitz *et al.* (1978) and Rudolph *et al.* (1978). Firstly, -Zn rADH was not denatured, secondly, -Zn rADH is not zinc-free

and thirdly, ADH 2334, being a bacterial ADH, is probably a tetramer rather than a dimer. However, the conditions under which denatured zinc-free HLADH can be reactivated *in vitro* are the only guide available for assessing the reasons why -Zn rADH was not activated by the addition of Zn^{2+}.

Reconstitution of totally unfolded, zinc-free HLADH involves a consecutive reaction sequence (Gerschitz *et al.,* 1978; Rudolph *et al.*, 1978). The first step is a first-order transconformation that results (over 10 min) in the formation of a monomeric, inactive intermediate. Addition of stoichiometric amounts of Zn^{2+} initiates a second-order association of monomers that leads to the dimeric, biologically-active protein. The extent of reactivation upon the addition of Zn^{2+} to inactive monomers was dependent on the timing of the addition of Zn^{2+}. When the addition of Zn^{2+} was delayed (beyond 10 min) monomers were inactivated by an uncharacterised side-reaction, which was not prevented by the addition of SH protecting agents and which resulted in reduced yield of dimer when Zn^{2+} was eventually added. The extent of reactivation was also dependent on the concentration of (denatured) enzyme and on the concentration of added Zn^{2+}. The upper limit of 70% reactivation was achieved in 144 h, at a protein concentration of 10 - 30 μg ml^{-1} (0.25 - 0.75 μM) and stoichiometric amounts of Zn^{2+}. At 20 °C and over 15 min, with all other conditions optimum, 20% of the maximum possible reactivation was achieved. Reactivation was inhibited by concentrations of Zn^{2+} in excess of stoichiometric amounts, an observation that was interpreted as due to the formation of an inactive complex of Zn^{2+} and monomer.

In the experiments with -Zn rADH, the conditions of ionic strength and pH were those described by Gerschitz *et al.*(1978) as optimum (100 mM phosphate, pH 7.0). Stoichiometric concentrations of Zn^{2+} were included in all the experiments. However, the protein concentrations were not always optimum and the time scale of the experiments may have been too short. In the experiments where Zn^{2+} was added directly to the cuvette, the -Zn rADH concentration was 0.4 μM or 37.5 nM, and activity was monitored for not more than 10 min. For HLADH, 0.4 μM was within the optimum protein concentration range but only 10% reactivation was observed in 10 min (at 20 °C). At a protein concentration of 37.5 nM only 5% reactivation of HLADH was observed.

The subunit M_r of +Zn rADH and of -Zn rADH was determined by electrospray mass spectrometry (personal communication, C. J. Bailey). The value obtained for the +Zn preparation was 36,184 ± 4, the value obtained for the -Zn preparation was 36,182 ± 4 (the value determined on the basis of the amino-acid sequence was 36,205; Section 3.4.2). The results of the mass spectrometry determinations established that the two enzyme preparations were identical in terms of their amino-acid composition and eliminated the possibility that unoccupied Zn^{2+}

ligands in -Zn rADH underwent a gross chemical modification *in vivo* which prevented the binding of Zn^{2+} *in vitro*. The possibility remains that partially active -Zn rADH is trapped, *in vivo*, in a conformation that cannot bind Zn^{2+} *in vitro* and therefore cannot be activated. The *E. coli* intracellular environment is a reducing environment and therefore the formation of disulfide linkages between unoccupied Cys ligands is not a realistic possibility.

Finally, adventitious metal-binding to unoccupied Zn^{2+} ligands in -Zn rADH must be considered. Stringent measures to prevent adventitious metal-binding to the apoenzyme were a feature of the classical studies on the role of Zn^{2+} in enzyme catalysis. Such measures included the use of double-deionised distilled water, extraction of all reagents and growth media with dithizone, soaking glass ware for 6 weeks in nitric acid (Thiers, 1957). Failure to stimulate *in vitro* the activity of an iron-deficient recombinant protein, which was attributed to adventitious metal-binding during storage, has been described by Ray & Bauerle (1991). Over-expression of recombinant tryptophan-sensitive 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase isoenzyme resulted in an enzyme with a subunit : Fe^{2+} stoichiometry of 1 : 0.3 and variable amounts of Zn^{2+}. The nature of the metal activator *in vivo* has not been established unequivocally. However, the evidence available suggests that Fe^{2+} is the preferred metal cofactor (Stephens & Bauerle, 1991; Akowski & Bauerle, 1997). The catalytic activity of the freshly purified recombinant enzyme was stimulated three-fold by the addition of Fe^{2+}, stimulation by added Zn^{2+} being about 30%. The magnitude of the response to added Fe^{2+} was inversely proportional to the age of the purified enzyme, being maximal in freshly prepared enzyme and declining during extended storage. Increasing the concentration of Fe^{2+} in the assay mixture did not compensate for the loss of activation. Enzyme that had been stored and that demonstrated a reduced response to the addition of Fe^{2+} was treated with *o*-phenanthroline and then with Fe^{2+}. The addition of iron completely restored activity implying that the enzyme adventitiously acquired nonfunctional metal ions, at unoccupied metal-binding sites, during storage. With respect to -Zn rADH no measures were taken to prevent contamination with metal ions, either during cell growth, during the purification procedure, or during storage. It is therefore possible that the failure of added Zn^{2+} to stimulate the enzymatic activity of the purified -Zn rADH preparation was due to adventitious metal-occupation of Zn^{2+} binding sites.

5.4.6 Zinc Limitation May Be a Feature of High-Level Expression of Recombinant Zinc-Dependent Enzymes

The literature relating to problems encountered during the expression of recombinant zinc-dependent proteins is summarised in Table 5.4.1. Recombinant phosphomannose isomerase (Proudfoot *et al.*, 1996) was expressed in *E. coli* at the

DNA construct	In the absence of supplementary Zn^{2+}	In the presence of supplementary Zn^{2+}	Reference
Candida albicans phosphomannose isomerase	90% of the protein was expressed into inclusion bodies	Four-fold increase in the yield of soluble protein in the presence of 1 mM Zn^{2+}	Proudfoot *et al.* (1996)
Truncated coding sequence for bovine link protein in fusion with *mal* E	50% of the protein was expressed into inclusion bodies	Yield of soluble protein increased in the presence of 500 μM supplementary Zn^{2+}. Biological activity of the soluble protein increased.	Varelas *et al.* (1997)
Entire coding sequence of *Schistosoma mansoni* cytosolic Cu/Zn superoxide dismutase	Specific activity of the purified enzyme was 60% of that of the wild-type enzyme. 10% increase in activity when purified enzyme was incubated with 100 μM Cu^{2+} and Zn^{2+}.	Not described	
Truncated coding sequence for human placental alkaline phosphatase (magesium- and zinc-dependent)	Most of the protein was expressed into inclusion bodies.	Yield of soluble activity reduced in the presence of 100 μM Zn^{2+}, but increased three-fold in the presence of 1 mM Mg^{2+}	Beck & Burtscher (1994)
Entire coding sequence for human stomach ADH in fusion with *gst*	50% of the purified enzyme was inactive.	Not described	Kedishvili *et al.* (1995)
Entire coding sequence of human glyoxalase I (including eight silent mutations)	The purified enzyme had low activity and an incorrect protein : metal stoichiometry	Specific activity of purified enzyme equivalent to that of the wild-type enzyme. The purified enzyme had the correct protein : metal stoichiometry.	Ridderström & Mannervik (1996)

Table 5.4.1 **Expression, in *E. coli*, of Zinc-Dependent, Recombinant Proteins**
mal E is the gene encoding maltose-binding protein, *gst* is the gene encoding glutathione *S*-transferase

level of 30% of total protein, but only 10% of the recombinant protein produced was soluble. When the expression host was grown in the presence of 1 mM supplementary Zn^{2+} there was an increase in the specific activity of the soluble fraction, which resulted from an increase in the amount of recombinant protein present in the soluble fraction. The total amount of phosphomannose isomerase produced (soluble plus insoluble) was independent of the addition of Zn^{2+} to the culture medium. Proudfoot *et al.* (1996) suggest that one of the factors driving the expression of recombinant phosphomannose isomerase into inclusion bodies is incorrect folding of the enzyme due to metal-limitation. However no data are presented for the metal content of the inclusion body protein, or for the metal content of the protein expressed into the soluble fraction in the absence of supplementary Zn^{2+}. No data are reported with respect to the specific activity of purified phosphomanose isomerase expressed into the soluble fraction in the absence of supplementary Zn^{2+}.

Varelas *et al.* (1997) have shown that expression of recombinant bovine link protein in the absence of Zn^{2+} supplement results in the formation of inclusion bodies and the production of soluble protein, in approximately equal quantities. The inclusion body protein could not be renatured using conventional techniques and the soluble protein exhibited minimal biological activity. When the medium was supplemented with Zn^{2+} the amount of soluble recombinant protein was increased, and the biological activity of the soluble protein increased to a level comparable to that of the wild-type protein. Those results demonstrate that the situation is not simply one in which the available Zn^{2+} is used to produce correctly folded, soluble, biologically-active protein and when the Zn^{2+} is used up any subsequent metal-deficient protein produced is driven into inclusion bodies.

In the case of recombinant *adh*2334 it was not established unequivocally that no protein was expressed into inclusion bodies in the absence of supplementary Zn^{2+}. However, samples removed from stationary-phase cultures of *E. coli* pLysS transformants, which expressed rADH, were examined under a phase-contrast microscope. The *E. coli* cells were clearly visible, but there was no visual evidence of inclusion bodies (data not shown).

Recombinant alkaline phosphatase was expressed in *E. coli* grown in media supplemented with either Zn^{2+} or Mg^{2+}. Relative to the activity observed when cells were grown in the absence of metal ion supplement, the enzyme activity in the soluble fraction was decreased when the medium was supplemented with Zn^{2+} but was increased when the media was supplemented with Mg^{2+} (Beck & Burtscher, 1994). From the data presented it is not clear whether the metal ions were affecting the intrinsic activity of the soluble enzyme, or the yield of the soluble enzyme, or both simultaneously. Only one reference to an inactive recombinant ADH was found

Recombinant DNA construct	**Level of expression mg l^{-1} of medium**	**Reference**
Complete coding sequence of *B. stearothermophilus adh*2334	143 (in the presence of supplementary Zn^{2+})	This work
Complete coding sequence of *B. stearothermophilus adh*1503 (expressed in *B. subtilis*)	not given	Sakoda & Imanaka (1992)
Complete coding sequence of *B. stearothermophilus adh*12403	1.25	Cannio *et al.* (1994)
Complete coding sequences for both isoenzymes of HLADH	28	Park & Plapp (1991)
Entamoeba histolytica $NADP^+$-dependent ADH in fusion with *gst*	not given	Kumar *et al.* (1992)
Entire coding sequence for human stomach ADH in fusion with *gst*	3	Kedishvili *et al.* (1995)
Entire coding sequence for *Thermoanaerobacter ethanolicus* 39E secondary ADH	not given in mg l^{-1} given as 1 - 5% of total protein	Burdette *et al.* (1996)
Entire coding sequence of *Sulfolobus solfataricus* ADH	38	Cannio *et al.* (1996)

Table 5.4.2 **Expression of Recombinant Zinc-dependent ADHs**

All of the constructs, with the exception of *adh*1503, were expressed in *E. coli*.

histolytica enzyme has only one Zn^{2+}-binding site, therefore, the zinc requirement will be less than that for other zinc-dependent ADHs.

This is the first report of the expression of a fully functional, recombinant ADH being limited by the concentration of Zn^{2+} in the bacteriological culture medium. The phenomenon is almost certainly linked to the extraordinary high levels of expression attained in this case and it is likely that zinc-limitation is a general feature of high-level expression of recombinant zinc-dependent enzymes in *E. coli*.

5.4.7 Kinetic Parameters of Recombinant ADH Produced in the Absence and Presence of Supplementary Zinc Ions

The specific activity of the wild-type enzyme, which was purified by ammonium sulfate fractionation followed by hydrophobic chromatography, chromatography on hydroxyapatite and gel-filtration, was 2.1 U mg^{-1} (Sheehan *et al.*, 1988). The value of K_m (apparent) for ethanol and the wild-type enzyme, obtained at a fixed NAD^+ concentration of 0.51 mM, and using a spectrophotometric assay, was 82 μM (Sheehan *et al.*, 1988). The K_m for NAD^+ and the wild-type enzyme was 39 μM (Sheehan *et al.*, 1988).

Conditions of temperature, pH, ionic strength and NAD^+ concentration under which the kinetic parameters of the two rADH preparations were determined were the same as those used for the wild-type enzyme. The data obtained for both rADH preparations, using a fluorimetric assay, conformed to the Michaelis-Menten equation and to the Hanes equation. The values obtained for the K_m (apparent) for ethanol and the +Zn rADH and the -Zn rADH preparation were 18.4 μM ± 2.4 and 9.6 μM ± 1.5, respectively.

A fluorimetric assay was used to generate the experimental data for the recombinant enzymes because of the difficulties associated with obtaining reproducible data for the -Zn rADH preparation at low ethanol concentrations. When the fluorimetric assay was used it became apparent that rates determined for -Zn rADH at low ethanol concentrations using the spectrophotometric method were frequently not true v_0. It was possible to obtain reliable data for +Zn rADH using the spectrophotometric assay and those data gave a value of 7 μM for the K_m for ethanol and +Zn rADH. Therefore, the discrepancy in the values of K_m for ethanol obtained for the wild-type enzyme and +Zn rADH cannot be attributed to the use of the fluorimetric assay for the recombinant enzyme.

Substrate inhibition of ADH 2334 at ethanol concentrations in excess of 50 mM was demonstrated by Sheehan *et al.* (1988). The data conformed to the partial inhibition mechanism described by Dalziel & Dickinson (1966) for HLADH (Section 1.5.1). The concentrations of ethanol used to determine K_m values for the rADHs did

not exceed 1.25 mM and there was no evidence of substrate inhibition operating over the concentration range used for the determination of K_m values.

The k_{cat} value obtained for the +Zn rADH preparation was 1.68 s^{-1}, similar to the value reported for the wild-type enzyme (1.3 s^{-1}, Sheehan *et al.*, 1988) and consistent with the close agreement in the specific activity values of the two preparations (2.3 U mg^{-1} for +Zn rADH and 2.1 U mg^{-1} for wild-type). The k_{cat} value obtained for -Zn rADH was 0.197 s^{-1}, which represented an 8.5-fold difference when compared to the value obtained for +Zn rADH and that was not consistent with the four-fold difference observed in the specific activity values of the freshly prepared enzymes (Table 5.3.3). The inconsistency is attributed to the -Zn rADH preparation losing enzyme activity during storage. The k_{cat} values of the two preparations were determined five months after the enzymes were purified.

The low value of k_{cat} for -Zn rADH compared to +Zn rADH is consistent with a simple picture of the composition of the -Zn rADH preparation, namely a mixture containing completely inactive and completely active rADH molecules. According to this simple picture, the completely inactive molecules are assumed to be totally devoid of Zn^{2+}. The completely active molecules are assumed to contain their full complement of Zn^{2+} and to have a three-dimensional configuration and a chemical activity that was identical to that of the +Zn rADH preparation. If the situation was as just described one might expect that the K_m values for the two preparations would be the same and they are not. There is a two-fold difference in the K_m values obtained for the two recombinant enzyme preparations, with the -Zn rADH preparation having the lower value. It is very difficult to assess the significance of a two-fold difference in K_m values determined on two different enzyme preparations. The difference may not be significant. The expression for K_m contains k_{cat} ($K_m = k_{-1} + k_2 / k_1$; $k_2 = k_{cat}$) therefore, the value of k_{cat} will influence the value of K_m. The low value of k_{cat} for -Zn rADH will lower the value of K_m for -Zn rADH assuming k_{-1} and k_1 are constant and k_{-1} is not very much larger than k_2.

5.4.8 Reversible Cold-Inactivation of rADH

When -Zn rADH was preincubated at different temperatures and subsequently assayed at 37 °C, the enzyme activity was dependent on the preincubation temperature. The activity increased approximately four-fold when the preincubation temperature was increased from 0 °C to 37 °C. The dilution factor, when the preincubation mixtures were added to the prewarmed assay mixtures, was 1 in 600, therefore, the different activities can not be attributed to different temperatures in the cuvettes. This 'preincubation affect' was also demonstrated with +Zn rADH and so, was an intrinsic property of the recombinant enzyme. The deactivation of +Zn rADH by reduction in temperature was reversed when the temperature was

increased, similarly, the activation of rADH when the temperature was increased was reversed when the temperature was reduced. rADH is a cold-inactivated enzyme and the cold-inactivation is reversible. It is reasonable to assume that the wild-type enzyme is also cold-inactivated.

Numerous cold-labile *Bacillus* glutamate dehydrogenases have been reported by Jahns (1992) and the enzyme from the mesophile *B. cereus* has been studied in detail (Jahns & Kaltwasser, 1993). Other cold-inactivated enzymes include mutant forms of *E. coli* inorganic pyrophosphatase (Velichko *et al.*, 1995), *E. coli* tryptophanase (Erez *et al.*, 1998), human estradiol 17 β-dehydrogenase (Li & Lin, 1996) and spinach chloroplast F_1-ATPase (Hightower & McCarty, 1996). In the *B. cereus* and *E. coli* enzymes cold-inactivation occurs at 0 °C, is reversible over time-scales (15 min) comparable to those in which cold-inactivation of rADH was reversed and involves dissociation and reassociation of subunits. The mechanism of cold-inactivation of *E. coli* tryptophanase has been studied in detail and shown to occur in two steps. In the first step, exposure to the cold causes the covalent bond between the pyridoxal 5'-phosphate prosthetic group and the enzyme to break. This step is partially inhibited by K^+ and by a high concentration of protein. In the second step, the tetrameric apoenzyme dissociates into inactive dimers. The dissociation is partially prevented by salting-out anions, particularly HPO_4^{2-}, also SO_4^{2-}, and is independent of the presence of K^+. Full reactivation requires the addition of pyridoxal 5'-phosphate to the warmed enzyme. Addition of the cofactor Mg^{2+} to mutant inorganic pyrophosphatase did not protect against dissociation of the hexamer into inactive trimers, but did decrease the time-scale of reassociation of the trimers. Cold-inactivation of *B. cereus* glutamate dehydrogenase is not prevented by SO_4^{2-}, phosphate anions or the addition of substrates, and is independent of protein concentration. Cold-inactivation of estradiol 17 β-dehydrogenase involves aggregation, rather than dissociation, of oligomers (Jarabak *et al.*, 1966).

During the determination of the kinetic parameters of rADH the enzyme was always stored on ice between assays and was therefore inadvertently preincubated at 0 °C. A lag period, corresponding to reactivation of the enzyme, might be expected when the cold enzyme was added to the cuvette. No lag was observed. In Figure 5.3.18 when the cold enzyme was moved to 37 °C maximum activity was restored in approximately two min. However the enzyme was maintained at approximately 2 mg ml^{-1} when the temperature was raised. When the cold enzyme was added to the warm cuvette during the measurement of the kinetic parameters the enzyme was diluted 3000-fold. Assuming that cold-inactivation involves dissociation of ADH subunits, reassociation would be concentration dependent and that would explain why no lag was observed over the time period of the kinetic measurements. In experiments where the affect of preincubation

temperature on the enzyme activity of rADH was determined the ethanol concentration in the cuvette was approximately 56 mM and therefore V_{max} was being measured. The data (Figures 5.3.16 and 5.3.17) demonstrate that V_{max} (and hence k_{cat}) decreases with decreasing preincubation temperature.

A number of observations concerning the behavior of wild-type ADH 2334 and the behavior of the recombinant enzyme can be explained in terms of cold-inactivation. Active ADH 2334 was extremely difficult to purify (Sheehan *et al.*, 1988; C. Bailey, personal communication). When the wild-type enzyme was purified (at 4 °C), ammonium sulfate fractionation was followed by hydrophobic chromatography on octyl Sepharose CL 4B. Large losses of enzyme activity occurred unless 1% (v/v) ethanol was included in the elution buffer and that can be interpreted as substrate protection against cold-inactivation. During initial attempts to purify rADH, chromatography on octyl Sepharose CL 4B was included in the purification scheme (data not shown) and large losses of enzyme activity were observed when the elution buffer did not contain ethanol. Almost complete retention of enzyme activity (both +Zn rADH and -Zn rADH) occurred, when ethanol was included in the elution buffer.

Removal of ammonium sulfate from the +Zn rADH preparation by dialysis, or by gel filtration, consistently resulted in very large losses in enzyme activity (Section 5.3.6). The loss of activity during dialysis and during gel-filtration can interpreted as cold-inactivation caused by either removal of SO_4^{2-}, or reduction in protein concentration, or a combination of the two. The total amount of -Zn rADH activity recovered in the 70% saturation ammonium sulfate supernatant was greater than that detected in the cell lysate [Table 5.3.3(b)]. As was noted in Section 5.3.6, the cell-lysate was not assayed immediately after it was prepared, therefore, the most likely explanation for the discrepancy is that the enzyme in the cell lysate was cold inactivated. However, on two other occasions during trial purifications of +Zn rADH, and when the cell lysates was assayed immediately after the cells were lysed, the total activities of the 70% ammonium sulfate supernatants were higher than those of the cell lysates, as much as 53% higher on one occasion. The increase in activity observed when ammonium sulfate was added to rADH suggests that SO_4^{2-} may be involved in reactivating cold-inactivated rADH.

The enzyme activity of -Zn rADH in post-sonication supernatants was particularly unstable relative to the activity of the +Zn rADH (Section 5.3.3). It is possible that cold-inactivation is more severe in the case of the zinc-deficient enzyme. Purified +Zn rADH and -Zn rADH, which had been stored for five years at -20 °C at a protein concentration of 5.8 mg ml^{-1} and 2.3 mg ml^{-1}, respectively, were thawed and assayed. Approximately 10% of the original activity of the +Zn rADH preparation remained, the -Zn rADH activity was barely detectable.

Given that ADH 2334 is a cold-inactivated enzyme the manner in which the wild-type and recombinant enzymes were purified and stored, and their behaviour after prolonged storage, becomes an issue when determining and comparing kinetic parameters. The discrepancy in the values of K_m for ethanol obtained for the wild-type enzyme and +Zn rADH may be due to differences in the conditions underwhich the enzymes were purified, a difference in the age of the enzyme preparations and differences in the conditions under which the enzymes were stored. The inactivation of the wild-type enzyme that occurred during purification was attributed, at the time, to proteolysis and the purification scheme was designed to minimise proteolysis (Sheehan *et al.*, 1988). It may be that the difficulties encountered when handling this enzyme are entirely due to cold-inactivation and any further work on this enzyme must take cold-inactivation into account.

REFERENCES

REFERENCES

Akowski, J.P. & Bauerle, R. (1997) *Biochemistry* **36**, 15817-15822

Altschul, S.F., Gish, W., Miller, W., Myres, E.W. & Lipman, D.J. (1990) *J. Mol. Biol.* **215**, 403-410

Amann, E. & Brosius, J. (1985) *Gene* **40**, 183-190

Amann, E., Ochs, B. & Abel, K.-J. (1988) *Gene* **69**, 301-315

Ameyama, M., Matsushita, K., Shinagawa, E. & Adachi, O. (1991) *Vitamins and Hormones* **46**, 229-270

Ammendola, S., Raia, C.A., Caruso, C., Camardella, L., D'Auria, S., de Rosa, M. & Rossi, M. (1992) *Biochemistry* **31**, 12514-12523

Amons, R. (1987) *FEBS Lett.* **212**, 68-72

Andersson, P., Kvassman, J., Oldén, B. & Pettersson, G. (1984) *Eur. J. Biochem.* **139**, 519-527

Anfinsen, C.B., Haber, E., Sela, M. & White, F.H. (1961) *Proc. Natl. Acad. Sci. U. S. A.* **47**, 1309-1314

Anthony, C. (1996) *Biochem. J.* **320**, 697-711

Antoine, E., Rolland, J.-L., Raffin, J.-P. & Dietrich, J. (1999) *Eur. J. Biochem.* **264**, 880-889

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

Atkinson, T., Banks, G.T., Bruton, C.J., Comer, M.J., Jakes, R., Kamalagharan, T., Whitaker, A.R. & Winter, G.P. (1979) *J. Appl. Biochem.* **1**, 247-258

von Bahr-Lindstrom, H., Höög, J.O. & Jörnvall, H. (1989) *Prog. Clin. Biol. Res.* **290**, 167-180

Baldock, C., Rafferty, J.B., Stuitje, A.R., Slabas, A.R. & Rice, D.W. (1998) *J. Mol. Biol.* **284**, 1529-1546

Bardwell, J.C.A., McGovern, K. & Beckwith, J. (1991) *Cell* **67**, 581-589

Bardwell, J.C.A., Lee, J.-O., Jander, G., Martin, N., Belin, D. & Beckwith, J. (1993) *Proc. Natl. Acad. Sci. U. S. A.* **90**, 1038-1042

Barstow, D.A., Clarke, A.R., Chia, W.N., Wigley, D., Sharman, A.F., Holbrook, J.J., Atkinson, T. & Minton, N.P. (1986) *Gene* **46**, 47-55

Bartsevich, V.V. & Pakrasi, H.B. (1996) *J. Biol. Chem.* **271**, 26057-26061

Beard, S.J., Hughes, M.N. & Poole, R.K. (1995) *FEMS Microbiol. Lett.* **131**, 205-210

Beck, R. & Burtscher, H. (1994) *Protein Expr. Purif.* **5**, 192-197

Benach, J., Atrian, S., Gonzàlez-Duarte, R. & Ladenstein, R. (1998) *J. Mol. Biol.* **282**, 383-399

Bennetzen, J.L. & Hall, B.D. (1982) *J. Biol. Chem.* **257**, 3018-3025

Bergler, H., Fuchsbichler, S., Hogenauer, G. & Turnowsky, F. (1996) *Eur. J. Biochem.* **242**, 689-694

Biesecker, G., Harris, J.I., Thierry, J.C., Walker, J.E. & Wonacott, A.J. (1977) *Nature* **266**, 328-333

Biggin, M.D., Gibson, T.J. & Hong, G.F. (1983) *Proc. Natl. Acad. Sci. U. S. A.* **80**, 3963-3965

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

Boleda, M.D., Saubi, N., Farrés, J. & Parés, X. (1993) *Arch. Biochem. Biophys.* **307**, 85-90

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

Bosron, W.F. & Prairie, R.L. (1972) *J. Biol. Chem.* **247**, 4480-4485

Bosron, W.F., Kennedy, F.S. & Vallee, B.L. (1975) *Biochemistry* **14**, 2275-2282

Bosron, W.F. & Li, T.-K. (1987) *Enzyme.* **37**, 19-28

Braig, K. (1998) *Curr. Opin. Struct. Biol.* **8**, 159-165

Bridgen, J., Kolb, E. & Harris, J.I. (1973) *FEBS Lett.* **33**, 1-3

Bright, J.R., Byrom, D., Danson, M.J., Hough, D.W. & Towner, P. (1993) *Eur. J. Biochem.* **211**, 549-554

Brosius, J., Dull, T.D., Sleeter, D.D. & Noller, H.F. (1981) *J. Mol. Biol.* **148**, 107-127

Brown, T.D.K., Jones-Mortimer, M.C. & Kornberg, H.L. (1977) *J. Gen. Microbiol.* **102**, 327-336

Bruchhaus, I. & Tannich, E. (1994) *Biochem. J.* **303**, 743-748

Brändén, C.-I., Eklund, H., Nordström, B., Boiwe, T., Söderlund, G., Zeppezauer, E., Ohlsson, I. & Åkeson, Å. (1973) *Proc. Natl. Acad. Sci. U. S. A.* **70**, 2439-2442

Brändén, C.-I., Jörnvall, H., Eklund, H. & Furugren, B. (1975) in *The Enzymes* (Boyer, P.D. ed.), pp. 103-190, Academic Press, London

Bucheder, F. & Broda, E. (1974) *Eur. J. Biochem.* **45**, 555-559

Bukau, B. & Horwich, A.L. (1998) *Cell* **92**, 351-366

Burdette, D.S., Vieille, C. & Zeikus, J.G. (1996) *Biochem. J.* **316**, 115-122
Cabiscol, E., Aguilar, J. & Ros, J. (1994) *J. Biol. Chem.* **269**, 6592-6597

Cameron, A.D., Roper, D.I., Moreton, K.M., Muirhead, H., Holbrook, J.J. & Wigley, D.B. (1994) *J. Mol. Biol.* **238**, 615-625

Cameron, E.J. & Esty, J.R. (1926) *J. Infect. Dis.* **39**, 89-105

Cannio, R., Rossi, M. & Bartolucci, S. (1994) *Eur. J. Biochem.* **222**, 345-352

Cannio, R., Fiorentino, G., Carpinelli, P., Rossi, M. & Bartolucci, S. (1996) *J. Bacteriol* **178**, 301-305

Cantor, C.R. & Schimmel, P.R. (1980) *Biophysical Chemistry Part III*, W. H. Freeman & Co., San Francisco, pp1183-1264

Carter, P., Bedouelle, H. & Winter, G. (1985) *Nucleic Acids Res.* **13**, 4431-4443

Cederlund, E., Peralba, J.M., Parés, X. & Jörnvall, H. (1991) *Biochemistry* **30**, 2811-2816

Chicken, E., Spode, J.A. & Jones, C.W. (1981) *FEMS. Microbiol. Lett.* **11**, 181-185

Churchill, P., Hempel, J., Romovacek, H., Zhang, W.-W., Brennan, M. & Churchill, S. (1992) *Biochemistry* **31**, 3793-3799

Ciriacy, M. (1975) *Mol. Gen. Genet.* **138**, 157-164

Clark, D.P. (1989) *FEMS Microbiol. Rev.* **63**, 223-234

Clarke, A.R., Atkinson, T., Campbell, J.W. & Holbrook, J.J. (1985) *Biochim. Biophys. Acta* **829**, 387-396

Cohen, S.N., Change, A.C.Y. & Hsu, L. (1972) *Proc. Natl. Acad. Sci. U. S. A.* **69**, 2110-2114

Colonna-Cesari, F., Perahia, D., Karplus, M., Eklund, H., Brändén, C.I. & Tapia, O. (1986) *J. Biol. Chem.* **261**, 15273-15280

Conway, T., Sewell, G.W., Osman, Y.A. & Ingram, L.O. (1987) *J. Bacteriol.* **169**, 2591-2597

Conway, T. & Ingram, L.O. (1989) *J. Bacteriol.* **171**, 3754-3759

Creaser, E.H., Murali, C. & Britt, K.A. (1990) *Protein. Eng.* **3**, 523-526

Czaja, C., Litwiller, R., Tomlinson, A.J., Naylor, S., Tavares, P., LeGall, J., Moura, J.J.G., Moura, I. & Rusnak, F. (1995) *J. Biol. Chem.* **270**, 20273-20277

Dalziel, K. (1962) *Biochem. J.* **84**, 244-254

Dalziel, K. & Dickinson, F.M. (1966) *Biochem. J.* **100**, 34-46

Dalziel, K. (1975) in *The Enzymes* (Boyer, P.D. ed.), pp. 1-60, Academic Press, London

Danielsson, O. & Jörnvall, H. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89**, 9247-9251

Das, A. & Yanofsky, C. (1984) *Nucleic Acids Res.* **12**, 4757-4768

Davies, G.J., Gamblin, S.J., Littlechild, J.A. & Watson, H.C. (1992) *J. Mol. Biol.* **227**, 1263-1264

Davis, G.J., Bosron, W.F., Stone, C.L., Owusu-Dekyi, K. & Hurly, T.D. (1996) *J. Biol. Chem.* **271**, 17057-17061

Dawson, R.M.C., Elliott, D.C., Elliott, W.H. & Jones, K.M. (1986) *Data for Biochemical Research*, Clarendon Press, Oxford, England, 3rd Ed., pp537-539

Day, C.P., Bashir, R., James, O.F.W., Bassindine, M.F., Crabb, D.W., Thomasson, H.R., Li, T.-K. & Edenberg, H.J. (1991) *Hepatology* **14**, 798-801

DeTraglia, M.C., Schmidt, J., Dunn, M.F. & Mc Farland, J.T. (1977) *J. Biol. Chem.* **252**, 3493-3500

Despreaux, C.W. & Manning, R.F. (1993) *Gene* **131**, 35-41

Devereux, J., Haeberli, P. & Smithies, O. (1984) *Nucleic Acids Res.* **12**, 387-395

Devine, K.M., Hogan, S.T., Higgins, D.G.. & McConnell, D.J. (1989) *J. Bacteriol.* **171**, 1166-1172

Dickinson, F.M. & Dalziel, K. [1967(a)] *Nature* **214**, 31-33

Dickinson, F.M. & Dalziel, K. [1967(b)] *Biochem. J.* **104**, 165-172

Dickinson, F.M. & Monger, G.P. (1973) *Biochem. J.* **131**, 261-270

Dintilhac, A., Alloing, G., Granadel, C. & Claverys, J.P. (1997) *Mol. Microbiol.* **25**, 727-739

Dixon, D. & Webb, E.C. (1961) *Advances in Protein Chemistry* **16**, 197-219

Donk, P.J. (1920) *J. Bacteriol.* **5**, 373-374

Dothie, J.M., Giglio, J.R., Moore, C.B., Taylor, S.S. & Hartley, B.S. (1985) *Biochem. J.* **230**, 569-578

Dowds, B.C., Sheehan, M.C., Bailey, C.J. & McConnell, D.J. (1988) *Gene* **68**, 11-22

Drewke, C. & Ciriacy, M. (1988) *Biochim. Biophys. Acta* **950**, 54-60

Drickamer, K. & Taylor, M.E. (1998) *Trends Biochem. Sci.* **23**, 321-324

Drum, D.E., Li, T.-K. & Vallee, B.L. [1969(a)] *Biochemistry* **8**, 3792-3797

Drum, D.E., Li, T.-K. & Vallee, B.L. [1969(b)] *Biochemistry* **8**, 3783-3791

Duester, G., Jörnvall, H. & Hatfield, G.W. (1986) *Nucleic Acids Res.* **14**, 1931-1941

Duester, G., Farrés, J., Felder, M.R., Holmes, R.S., Höög, J.-O., Parés, X., Plapp, B.V., Yin, S.-J. & Jörnvall, H. (1999) *Biochem. Pharmacol.* **58**, 389-395

Dunn, T.J., Koleske, A.J., Lindahl, R. & Pitot, H.C. (1989) *J. Biol. Chem.* **264**, 13057-13065

Edenberg, H.J., Dailey, T.L. & Zhang, K. (1989) *Prog. Clin. Biol. Res.* **290**, 181-192

Edman, P. & Begg, G. (1967) *Eur. J. Biochem.* **1**, 80-91

Edwards, K.J., Barton, J.D., Rossjohn, J., Thorn, J.M., Taylor, G.L. & Ollis, D.L. (1996) *Arch. Biochem. Biophys.* **328**, 173-183

Ehrig, T., Bosron, W.F. & Li, T.-K. (1990) *Alcohol Alcohol.* **25**, 105-116

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

Eidsness, M.K., O'Dell, S.E., Kurtz, D.M., Robson, R.L. & Scott, R.A. (1992) *Protein Eng.* **5**, 367-371

Eklund, H., Nordström, B., Zeppezauer, E., Söderlund, G., Ohlsson, I., Boiwe, T. & Brändén, C.-I. (1974) *FEBS Lett.* **44**, 200-204

Eklund, H., Nordström, B., Zeppezauer, E., Söderlund, G., Ohlsson, I., Boiwe, T., Söderberg, B.-O., Tapia, O., Brändén, C.-I. & Åkeson, Å. (1976) *J. Mol. Biol.* **102**, 27-59

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

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

Eklund, H. & Brändén, C.-I. (1983) in *Zinc Enzymes* (Spiro, T.G. ed.), pp. 123-152, John Wiley & Sons, New York

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

Engelke, D.R., Kirkos, A., Bruck, M.E. & Ginsburg, D. (1990) *Anal. Biochem.* **191**, 396-400

Epstein, I. & Grossowicz, N. (1969) *J. Bacteriol.* **99**, 414-417

Epstein, I. & Grossowicz, N. (1975) *J. Bacteriol.* **122**, 1257-1264

Erez, T., Gdalevsky, G.Ya., Torchinsky, Y.M., Philips, R.S. & Parola, A.H. (1998) *Biochim. Biophys. Acta* **1384**, 365-372

Estonius, M., Svensson, S. & Höög, J.-O. (1996) *FEBS Lett.* **397**, 338-342

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

Fiorentino, G., Cannio, R., Rossi, M. & Bartolucci, S. (1998) *Protein Eng.* **11**, 925-930

Fisher, G. & Schmid, F.X. (1990) *Biochemistry* **29**, 2205-2212

Flynn, T.G. (1982) *Biochem. Pharmacol.* **31**, 2705-2712

Fontana, A. & Gross, E. (1986) in *Practical Protein Chemistry A Handbook* (Darbre, A. ed.), pp. 67-120, John Wiley & Sons, Chichester

Fouet, A., Arnauld, M., Klier, A. & Rapoport, G. (1987) *Proc. Natl. Acad. Sci. U. S. A.* **84**, 8773-8777

Freedman, R.B. (1984) *Trends Biochem. Sci.* **9**, 438-441

Friedman, M., Krull, L.H. & Cavins, J.F. (1970) *J. Biol. Chem.* **245**, 3868-3871

Fujiwara, S., Kakihara, H., Woo, K.B., Lejeune, A., Kanemoto, M., Sakaguchi, K. & Imanaka, T. (1992) *Appl. Environ. Microbiol.* **58**, 4016-4025

Gagnon, Y., Breton, R., Putzer, H., Pelchat, M., Grunberg-Manago, M. & Lapointe, J. (1994) *J. Biol. Chem.* **269**, 7473-7482

Ganzhorn, A.J., Green, D.W., Hershey, A.D., Gould, R.M. & Plapp, B.V. (1987) *J. Biol. Chem.* **262**, 3754-3761

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

Geisow, M.J. & Aitken, A. (1989) in *Protein Sequencing A Practical Approach* (Findlay, J.B.C. & Geisow, M.J. eds.), pp. 85-98, IRL Press, Oxford

Georgopoulos, C. (1992) *Trends Biochem. Sci.* **17**, 295-299

Gerschitz, J., Rudolph, R. & Jaenicke, R. (1978) *Eur. J. Biochem.* **87**, 591-599

Gething, M-J. & Sambrook, J. (1992) *Nature* **355**, 33-45

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

Gillooly, D.J., Robertson, A.G. & Fewson, C.A. (1998) *Biochem. J.* **330**, 1375-1381

Godson, G.N. & Sinsheimer, R.L. (1967) *Biochim. Biophys. Acta* **149**, 489-495

Goodlove, P.E., Cunningham, P.R., Parker, J. & Clark, D.P. (1989) *Gene* **85**, 209-214

Goodwin, P.M. & Anthony, C. (1998) *Adv. Microb. Physiol.* **40**, 3-80

Gordon, E.J., Stella, M.B. & Sawyer, L. (1992) *J. Mol. Biol.* **227**, 356-358

Gordon, R.E. & Smith, N.R. (1949) *J. Bacteriol.* **58**, 327-341

Gordon, R.E. (1981) in *The Aerobic Endospore-Forming Bacteria: Classification and Identification* (Berkeley, R.C.W. & Goodfellow, M. eds.), pp. 1-15, Academic Press, London

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

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

Greener, A. (1990) *Strategies* **3**, 5-6

Grisshammer, R. & Nagai, K. (1995) in *DNA Cloning: A Practical Approach Volume II* (Glover, D.M. & Hames, B.D. eds.), pp. 59-92, IRL Press, Oxford

Grodberg, J. & Dunn, J.J. (1988) *J. Bacteriol.* **170**, 1245-1253

Groen, B., Frank, J., Duine, J. & Duine, J.A. (1984) *Biochem. J.* **223**, 921-924

Grundy, F.J., Waters, D.A., Takova, T.Y. & Henkin, T.M. [1993(a)] *Mol. Microbiol.* **10**, 259-271

Grundy, F.J., Waters, D.A., Allen, H.G. & Henkin, T.M. [1993(b)] *J. Bacteriol.* **175**, 7348-7355

Gryczan, T., Shivakunar, A.G. & Dubnau, D. (1980) *J. Bacteriol.* **141**, 246-253

Guan, K.-L. & Dixon, J.E. (1991) *Anal. Biochem.* **192**, 262-267

van de Guchte, M., Kok, J. & Venema, G. (1991) *Mol. Gen. Genet.* **227**, 65-71

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

Han, C.-L., Liao, C.-S., Wu, C.-W., Hwong, C.-L., Lee, A.-R. & Yin, S.-J. (1998) *Eur. J. Biochem.* **254**, 25-31

Hanahan, D. (1983) *J. Mol. Biol.* **166**, 557-580

Hanes, C.S. (1932) *Biochem. J.* **26**, 1406-1421

Hardy, F., Vriend, G., Veltman, O.R., van der Vinne, B., Venema, G. & Eijsink, V.G.H. (1993) *FEBS Lett.* **317**, 89-92

Harlow, E. & Lane, D. (1988) *Antibodies A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, pp471-510

Harris, M.I. & Coleman, J.E. (1968) *J. Biol. Chem.* **243**, 5063-5073

Hartree, E.F. (1972) *Anal. Biochem.* **48**, 422-427

Helfman, D.M., Feramisco, J.R., Fiddes, J.C., Thomas, G.P. & Hughes, S.H. (1983) *Proc. Natl. Acad. Sci. U. S. A.* **80**, 31-35

Hemmingsen, L., Bauer, R., Bjerrum, M.J., Adolph, H.W., Zeppezauer, M. & Cedergren-Zeppezauer, E. (1996) *Eur. J. Biochem.* **241**, 546-551

Henderson, C.E., Perham, R.N. & Finch, J.T. (1979) *Cell* **17**, 85-93

Henehan, G.T., Chang, S.H. & Oppenheimer, N.J. (1995) *Biochemistry* **34**, 12294--32768

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

Herbort, M., Schon, U., Angermann, K., Lang, J. & Schumann, W. (1996) *Gene* **170**, 81-84

Hess, J.F., Bourret, R.B. & Simon, M.I. (1988) *Nature* **336**, 139-143

Hewick, R.M., Hunkapiller, M.W., Hood, L.E. & Dreyer, W.J. (1981) *J. Biol. Chem.* **256**, 7990-7997

Hightower, K.E. & McCarty, R.E. (1996) *Biochemistry* **35**, 4852-4857

Hirata, H., Fukazawa, T., Negoro, S. & Okada, H. (1986) *J. Bacteriol.* **166**, 722-727

Hockney, R.C. (1994) *TIBTECH* **12**, 456-463

Hodgman, C.D. (1955) *Handbook of Chemistry and Physics*, Chemical Rubber Publishing Co., Cleveland, Ohio, 37th Ed.,

Hol, W.G., Van Duijnen, P.T. & Berendsen, H.J. (1978) *Nature* **273**, 443-446

Holmes, D.S. & Quigley, M. (1981) *Anal. Biochem.* **114**, 193-197

Holmquist, B. & Vallee, B.L. (1991) *Biochem. Biophys. Res. Commun.* **178**, 1371-1377

Hong, Z., LoVerde, P.T., Hammarskjöld, M.-L. & Rekosh, D. (1992) *Exper. Parasitol.* **75**, 308-322

Horwitz, J.P., Chua, J., Curby, R.J., Tomson, A.J., Da Rooge, M.A., Fisher, B.E., Mauricio, J. & Klundt, I. (1964) *J. Med. Chem.* **7**, 574

Howard, M.J., Chauham, H.J., Domingo, G.J., Fuller, C. & Perham, R.N. (2000) *J. Mol. Biol.* **295**, 1023-1037

Hsu, L.C. & Chang, W.-C. (1991) *J. Biol. Chem.* **266**, 12257-12265

Hubacek, J. & Glover, S.W. (1970) *J. Mol. Biol.* **50**, 111-127

Hughes, M.N. & Poole, R.K. (1989) *Metals and Micro-organisms*, Chapman and Hall, London,

Hunkapiller, M.W. (1985) *Applied Biosystems User Bulletin* **14**,

Hunkapiller, T., Kaiser, R.J., Koop, B.F. & Hood, L. (1991) *Science* **254**, 59-67

Hurley, T.D., Edenberg, H.J. & Bosron, W.F. (1990) *J. Biol. Chem.* **265**, 16366-16372

Hurley, T.D., Bosron, W.F., Hamilton, J.A. & Amzel, L.M. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88**, 8149-8153

Imanaka, T., Ohta, T., Sakoda, H., Widhyastuti, N. & Matsuoka, M. (1993) *J. Ferment. Bioeng.* **76**, 161-167

Inoue, T., Sunagawa, M., Mori, A., Imai, C., Fukuda, M., Takagi, M. & Yano, K. (1989) *J. Bacteriol.* **171**, 3115-3122

Ish-Horowicz, D. & Burke, J.F. (1981) *Nucleic Acids Res.* **9**, 2989-2998

Ismaiel, A.A., Zhu, C.X., Colby, G.D. & Chen, J.S. (1993) *J. Bacteriol.* **175**, 5097-5105

Itakura, K., Rossi, J.J. & Wallace, R.B. (1984) *Ann. Rev. Biochem.* 323-356

Ito, K., Takashashi, M., Yoshimoto, T. & Tsuru, D. (1994) *J. Bacteriol.* **176**, 2483-2491

Jaenicke, R. (1974) *Eur. J. Biochem.* **46**, 149-155

Jaenicke, R. (1987) *Prog. Biophys. molec. Biol.* **49**, 117-237

Jaenicke, R. (1991) *Biochemistry* **30**, 3147-3161

Jahns, T. (1992) *FEMS Microbiol. Lett.* **96**, 187-192

Jahns, T. & Kaltwasser, H. (1993) *J. Gen. Microbiol.* **139**, 775-780

Jany, K.-D., Ulmer, W., Fröschle, M. & Pfleiderer, G. (1984) *FEBS Lett.* **165**, 6-10

Jarabak, J., Seeds, A.E. & Talalay, P. (1966) *Biochemistry* **5**, 1269-1278

Jeffery, J., Chesters, J., Mills, C., Sadler, P.J. & Jörnvall, H. (1984) *EMBO J.* **3**, 357-360

Jeloková, J., Karlsson, C., Estonius, M., Jörnvall, H. & Höög, J.-O. (1994) *Eur. J. Biochem.* **225**, 1015-1019

Jendrossek, D., Steinbüchel, A. & Schlegel, H.G. (1988) *J. Bacteriol.* **170**, 5248-5256

Jendrossek, D., Kruger, N. & Steinbuchel, A. (1990) *J. Bacteriol.* **172**, 4844-4851

John, J., Crennell, S.J., Hough, D.W., Danson, M.J. & Taylor, G.L. (1994) *Structure* **2**, 385-393

Johns, D. (1989) in *HPLC of Macromolecules a Practical Approach* (Oliver, R.W.A. ed.), pp. 1-42, IRL Press, Oxford

Johnson, A.R., Chen, Y.-W. & Dekker, E.E. (1998) *Arch. Biochem. Biophys.* **358**, 211-221

Juliá, P., Farres, J. & Parés, X. (1987) *Eur. J. Biochem.* **162**, 179-189

Jörnvall, H. [1970(a)] *Eur. J. Biochem.* **16**, 25-40

Jörnvall, H. [1970(b)] *Eur. J. Biochem.* **16**, 41-49

Jörnvall, H. (1977) *Eur. J. Biochem.* **72**, 425-442

Jörnvall, H., Eklund, H. & Brändén, C.-I. (1978) *J. Biol. Chem.* **253**, 8414-8419

Jörnvall, H., Persson, M. & Jeffery, J. (1981) *Proc. Natl. Acad. Sci. U. S. A.* **78**, 4226-4230

Jörnvall, H., von Bahr-Lindstrom, H. & Jeffery, J. (1984) *Eur. J. Biochem.* **140**, 17-23

Jörnvall, H., von Bahr-Lindström, H., Jany, K.-D., Ulmer, W. & Fröschle, M. (1984) *FEBS Lett.* **165**, 190-196

Jörnvall, H., Persson, B. & Jeffery, J. (1987) *Eur. J. Biochem.* **167**, 195-201

Kaiser, R., Holmquist, B., Vallee, B.L. & Jörnvall, H. (1989) *Biochemistry* **28**, 8432-8438

Kaiser, R., Fernández, M.R., Parés, X. & Jörnvall, H. (1993) *Proc. Natl. Acad. Sci. U. S. A.* **90**, 11222-11226

Kakuda, H., Hosono, K., Shiroishi, K. & Ichihara, S. (1994) *J. Biochem.* **116**, 916-922

Kamitani, S., Akiyama, Y. & Ito, K. (1992) *EMBO J.* **11**, 57-62

Karlsson, C. & Höög, J.-O. (1993) *Eur. J. Biochem.* **216**, 103-107

Kasahara, M. & Anraku, Y. (1974) *J. Biochem.* **76**, 967-976

Kedishvili, N.Y., Bosron, W.F., Stone, C.L., Hurley, T.D., Peggs, C.F., Thomasson, H.R., Popov, K.M., Carr, L.G., Edenburg, H.J. & Li, T.-K. (1995) *J. Biol. Chem.* **270**, 3625-3630

Kedishvili, N.Y., Gough, W.H., Chernoff, E.A.G., Hurley, T.D., Stone, C.L., Bowman, K.D., Popov, K.M., Bosron, W.F. & Li, T.-K. (1997) *J. Biol. Chem.* **272**, 7494-7500

Kelly, C.A., Nishiyama, M., Ohnishi, Y., Beppu, T. & Birktoft, J.J. (1993) *Biochemistry* **32**, 3913-3922

Keshav, K.F., Yomano, L.P., An, H.J. & Ingram, L.O. (1990) *J. Bacteriol.* **172**, 2491-2497

Kessler, D., Leibrecht, I. & Knappe, J. (1991) *FEBS Lett.* **281**, 59-63

Kessler, D., Herth, W. & Knappe, J. (1992) *J. Biol. Chem.* **267**, 18073-18079

Klapper, D.G., Wilde, C.E. & Capra, J.D. (1978) *Anal. Biochem.* **85**, 126-131

Klinman, J.P. (1981) *CRC. Crit. Rev. Biochem.* **10**, 39-78

Knappe, J. & Sawers, G. (1990) *FEMS Microbiol. Rev.* **75**, 383-398

Kobayashi, K., Doi, S., Negoro, S., Urabe, I. & Okada, H. (1989) *J. Biol. Chem.* **264**, 3200-3205

Koivusalo, M., Baumann, M. & Uotila, L. (1989) *FEBS Lett.* **257**, 105-109

Kolb, E. & Harris, J.I. (1971) *Biochem. J.* **124**, 76P

Kolb, V.A., Makeyev, E.V., Kommer, A. & Spirin, A.S. (1995) *Biochem. Cell Biol.* **73**, 1217-1220

Korndörfer, I., Steipe, B., Huber, R., Tomschy, A. & Jaenicke, R. (1995) *J. Mol. Biol.* **246**, 511-521

Krebs, H.A. & Perkins, J.R. (1970) *Biochem. J.* **118**, 635-644

Krook, M., Marekov, L. & Jörnvall, H. (1990) *Biochemistry* **29**, 738-743

Kubo, N. & Imanaki, T. (1988) *J. Gen. Microbiol.* **134**, 1883-1892

Kumar, A., Shen, P-S., Descoteaux, S., Pohl, J., Bailey, G. & Samuelson, J. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89**, 10188-10192

Kung, F.-C., Raymond, J. & Glaser, D.A. (1976) *J. Bacteriol.* **126**, 1089-1095

Kuriki, T. & Imanaka, T. (1989) *J. Gen. Microbiol.* **135**, 1521-1528

Kvassman, J. & Pettersson, G. (1980) *Eur. J. Biochem.* **103**, 565-575

Kvassman, J., Larsson, A. & Pettersson, G. (1981) *Eur. J. Biochem.* **114**, 555-563

Kägi, J.H. (1991) *Methods Enzymol.* **205**, 613-626

Kärst, U., Schütte, H., Baydoun, H. & Tsai, H. (1900) *J. Gen. Microbiol.* **135**, 1305-1313

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

Lai, X. & Ingram, L.O. (1993) *J. Bacteriol.* **175**, 6441-6450

Lai, X. & Ingram, L.O. (1995) *Microbiol.* **141**, 1443-1449

Leatherbarrow, R.J. (1987) *Manual for a Non-linear Regression Data Analysis Program for the IBM PC*, Elsevier Science Publishers BV, Amsterdam,

Levy, H.R. & Vennesland, B. (1957) *J. Biol. Chem.* **228**, 85-96

Li, B. & Lin, S.-X. (1996) *Eur. J. Biochem.* **235**, 180-186

Li, D. & Stevenson, K.J. (1997) *J. Bacteriol.* **179**, 4433-4437

Li, H., Hallows, W.H., Punzi, J.S., Pankiewicz, K.W., Watanabe, K.A. & Goldstein, B.M. (1994) *Biochemistry* **33**, 11734-11744

Li, S.-L., Hanlon, J. & Yanofsky, C. (1975) *J. Bacteriol.* **123**, 761-764

Light, D.R., Dennis, M.S., Forsythe, I.J., Liu, C.-C., Green, D.W., Kratzer, D.A. & Plapp, B.V. (1992) *J. Biol. Chem.* **267**, 12592-12599

Lindahl, L., Archer, R.H., McCormick, J.R., Freedman, L.P. & Zengel, J.M. (1989) *J. Bacteriol* **171**, 2639-2645

Lindskog, S. & Malmström, B.G. (1962) *J. Biol. Chem.* **237**, 1129-1137

Little, S., Hyde, S., Campbell, C.J., Lilley, R.J. & Robinson, M.K. (1989) *J. Bacteriol.* **171**, 3518-3522

Liu, J. & Walsh, C.T. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87**, 4028-4032

Ljungdahl, L.G. (1979) *Adv. Microb. Physiol.* **19**, 149-243

Loviny, T., Norton, P.M. & Hartley, B.S. (1985) *Biochem. J.* **230**, 579-585

Lowry, O.H., Rosebrough, N.J., Farr, A.L. & Randall, R.J. (1951) *J. Biol. Chem.* **193**, 265-275

Lu, D., Boyd, B. & Lingwood, C.A. (1997) *J. Biol. Chem.* **272**, 29033-29038

Ma, K., Robb, F.T. & Adams, M.W.W. (1994) *Appl. Environ. Microbiol.* **60**, 562-568

Ma, K., Loessner, H., Heider, J., Johnson, M.K. & Adams, M.W.W. (1995) *J. Bacteriol.* **177**, 4748-4756

Ma, K. & Adams, M.W. (1999) *J. Bacteriol.* **181**, 1163-1170

Mackenzie, K.F., Conway, T., Aldrich, H.C. & Ingram, L.O. (1989) *J. Bacteriol.* **171**, 4577-4582

MacKintosh, R.W. & Fewson, C.A. (1988) *Biochem. J.* **250**, 743-751

Mallinder, P.R & Moir, A. (1991) *J. Gen. Microbiol.* **137**, 779-785

Mallinder, P.R., Pritchard, A. & Moir, A. (1992) *Gene* **110**, 9-16

Mandel, M. & Higa, A. (1970) *J. Mol. Biol.* **53**, 159-162

Mandelstam, J. & McQuillen, K. (1973) in *Biochemistry of Bacterial Growth* (Mandelstam, J. & McQuillen, K. eds.), pp. 20-25, Blackwell Scientific Publications, Oxford

Maras, B., Consalvi, V., Chiaraluce, R., Politi, L., de Rosa, M., Bossa, F., Scandurra, R. & Barra, D. (1992) *Eur. J. Biochem.* **203**, 81-87

Maret, W., Andersson, I., Dietrich, H., Schneider-Bernlöhr, H. & Einarsson, R. (1979) *Eur. J. Biochem.* **98**, 501-512

Marqusee, S. & Baldwin, R.L. (1987) *Proc. Natl. Acad. Sci. U. S. A.* **84**, 8898-8902

Maser, E. & Netter, K.J. (1989) *Biochem. Pharmacol.* **38**, 3049-3054

Maser, E., Oppermann, U.C.T., Bannenberg, G. & Netter, K.J. (1992) *FEBS Lett.* **297**, 196-200

Matsushita, K., Takahashi, K. & Adachi, O. (1993) *Biochemistry* **32**, 5576-5582

Matsushita, K., Yukushi, T., Toyama, H., Adachi, O., Miyoshi, H., Tagami, E. & Sakamoto, K. (1999) *Biochim. Biophys. Acta* **1409**, 154-164

Matthews, B.W., Nicholson, H. & Becktel, W.J (1987) *Proc. Natl. Acad. Sci. U. S. A.* **84**, 6663-6667

Matthews, B.W. (1993) *Annu. Rev. Biochem.* **62**, 139-160

McEvily, A.J., Holmquist, B., Auld, D.S. & Vallee, B.L. (1988) *Biochemistry* **27**, 4284-4288

McLaughlin, J.R., Murray, C.L. & Rabinowitz, J.C. (1981) *J. Biol. Chem.* **256**, 11283-11291

Menéndez-Arias, L. & Argos, P. (1989) *J. Mol. Biol.* **206**, 397-406

Michaelis, L. & Menten, M.L. (1913) *Biochem. Z.* **49**, 333-369

Mitraki, A., Fane, B., Haase-Pettingell, C., Sturtevant, J. & King, J. (1991) *Science* **253**, 54-58

Moeck, G.S. & Coulton, J.W. (1998) *Mol. Microbiol.* **28**, 675-681

Moreno, A. & Parés, X. (1991) *J. Biol. Chem.* **266**, 1128-1133

Morrison, D.A. (1979) *Methods Enzymol.* **68**, 326-331

Morshauser, R.C., Hu, W., Wang, H., Pang, Y., Flynn, G.C. & Zuiderweg, E.R.P. (1999) *J. Mol. Biol.* **289**, 1387-1403

Mrabet, N.T., Van den Broeck, A., Van den brande, I., Stanssens, P., Laroche, Y., Lambeir, A.-M., Matthijssens, G., Jenkins, J., Chiadmi, M., van Tilbeurgh, H., Rey, F., Janin, J., Quax, W.J., Lasters, I., De Maeyer, M. & Wodak, S.J. (1992) *Biochemistry* **31**, 2239-2253

Mullis, K.B. & Faloona, F.A. (1987) *Methods Enzymol.* **155**, 335-350

Murali, C. & Creaser, E.H. (1986) *Protein Eng.* **1**, 55-57

Murphy, G. & Ward, E.S. (1989) in *Nucleic Acids Sequencing a Practical Approach* (Howe, C.J. & Ward, E.S. eds.), pp. 99-115, IRL Press, Oxford

Mårdh, G., Dingley, A.L., Auld, D.S. & Vallee, B.L. (1986) *Proc. Natl. Acad. Sci. U. S. A.* **83**, 8908-8912

Mårdh, G. & Vallee, B.L. (1986) *Biochemistry* **25**, 7279-7282

Nair, R.V., Bennett, G.N. & Papoutsakis, E.T. (1994) *J. Bacteriol.* **176**, 871-885

Napoli, J.L. (1999) *Biochim. Biophys. Acta* **1440**, 139-162

Neale, A.D., Scopes, R.K., Kelly, J.M. & Wettenhall, R.E.H. (1986) *Eur. J. Biochem.* **154**, 119-124

Neilands, J.B. (1995) *J. Biol. Chem.* **270**, 26723-26726

Ng, K., Ye, R., Wu, X.-C. & Wong, S.-L. (1992) *J. Biol. Chem.* **267**, 24989-24994

Ng, T. & Kenealy, W.R. (1986) in *Thermophiles: General, Molecular and Applied Microbiology* (Brock, T.D. ed.), pp. 197-215, Wiley, New York

Nies, D.H. & Silver, S. (1989) *J. Bacteriol.* **171**, 4073-4075

Norrander, J., Kempe, T. & Messing, J. (1983) *Gene* **27**, 181-186

Nygren, P-A., Stähl, S. & Uhlén, M. (1994) *TIBTECH* **12**, 184-188

O'Connor, T., Ireland, L.S., Harrison, D.J. & Hayes, J.D. (1999) *Biochem. J.* **343**, 487-504

Obradors, N., Cabiscol, E., Aguilar, J. & Ros, J. (1998) *Eur. J. Biochem.* **258**, 207-213

Ohlsson, I., Nordström, B. & Brändén, C.-I. (1974) *J. Mol. Biol.* **89**, 339-354

Olafson, R.W., Mc Cublin, W.D. & Kay, C.M. (1988) *Biochem. J.* **251**, 691-699

Omburo, G.A., Kuo, J.M., Mullins, L.S. & Raushel, F.M. (1992) *J. Biol. Chem.* **267**, 13278-13283

Oppenheim, D.S. & Yanofsky, C. (1980) *Genetics* **95**, 785-795

Ortlepp, S.A., Ollington, J.F. & McConnell, D.J. (1983) *Gene* **23**, 267-276

Park, D.-H. & Plapp, B.V. (1991) *J. Biol. Chem.* **266**, 13296-13302

Park, D.-H. & Plapp, B.V. (1992) *J. Biol. Chem.* **267**, 5527-5533

Patzer, S.I. & Hantke, K. (1998) *Mol. Microbiol.* **28**, 1199-1210

Pearl, L.H., Demasi, D., Hemmings, A.M., Mazzarella, F.S.L., Raia, C.A., D'Auria, S. & Rossi, M. (1993) *J. Mol. Biol.* **229**, 782-784

Pelmont, J., Tournesac, C., Mliki, A., Barrelle, M. & Beguin, C. (1989) *FEMS. Microbiol. Lett.* **48**, 109-113

Peretz, M. & Burstein, Y. (1989) *Biochemistry* **28**, 6549-6555

Persson, B., Krook, M. & Jörnvall, H. (1991) *Eur. J. Biochem* **200**, 537-543

Persson, B., Bergman, T., Keung, W.M., Waldenström, U. & Holmquist, B. (1993) *Eur. J. Biochem.* **216**, 49-56

Persson, B., Zigler, J.S. & Jörnvall, H. (1994) *Eur. J. Biochem.* **226**, 15-22

Perutz, M.F. & Raidt, H. (1975) *Nature* **255**, 256-259

Peterson, J.A., Lu, J.-Y., Geisselsoder, J., Graham-Lorence, S., Carmona, C., Witney, F. & Lorence, M.C. (1992) *J. Biol. Chem.* **267**, 14193-14203

Pettersson, G. (1987) *CRC. Crit. Rev. Biochem.* **21**, 349-389

Proudfoot, A.E.I., Goffin, L., Payton, M.A., Wells, T.N.C. & Bernard, A.R. (1996) *Biochem. J.* **318**, 437-442

Qulali, M., Wolfla, C.E., Ross, R.A. & Crabb, D.W. (1989) *Prog. Clin. Biol. Res.* **290**, 143-153

Ramaswamy, S., Eklund, H. & Plapp, B.V. [1994(a)] *Biochemistry* **33**, 5230-5237

Ramaswamy, S., Kratzer, D.A., Hershey, A.D., Rogers, P.H., Arnone, A., Eklund, H. & Plapp, B.V. [1994(b)] *J. Mol. Biol.* **235**, 777-779

Ramaswamy, S., el Ahmad, M., Danielsson, O., Jörnvall, H. & Eklund, H. (1996) *Protein Sci.* **5**, 663-671

Ramaswamy, S., Park, D.-H. & Plapp, B.V. (1999) *Biochemistry* **38**, 13951-13959

Rao, P.V., Krishna, C.M. & Zigler, J.S. (1992) *J. Biol. Chem.* **267**, 96-102

Ray, J.M. & Bauerle, R. (1991) *J. Bacteriol.* **173**, 1894-1901

Reaney, S.K., Bungard, S.J. & Guest, J.R. (1993) *J. Gen. Microbiol.* **139**, 403-416

Reid, M.F. & Fewson, C.A. (1994) *Crit. Rev. Microbiol.* **20**, 13-56

Rensing, R., Mitra, B. & Rosen, B.P. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94**, 14326-14331

Ridderström, M. & Mannervik, B. (1996) *Biochem. J.* **314**, 463-467

Robinson, G.A., Bailey, C.J., Dowds, B.C.A., (1994) *Biochim. Biophys. Acta* **1218** 432-434

Rodokanaki, A., Holmes, R.K. & Borrás, T. (1989) *Gene* **78**, 215-224

Roitel, O., Sergienko, E. & Branlant, G. (1999) *Biochemistry* **38**, 16084-16091

Rongnoparut, P. & Weaver, S. (1991) *Gene* **101**, 261-265

Rosenberg, A.H., Lade, B.N., Chui, D., Lin, S., Dunn, J.J. & Studier, F.W. (1987) *Gene* **56**, 125-135

Rosenberg, M. & Court, D. (1979) *Ann. Rev. Genet.* **13**, 319-353

Rosenthal, B., Mai, Z., Caplivski, D., Ghosh, S., De La Vega, H., Graf, T. & Samuelson, J. (1997) *J. Bacteriol.* **179**, 3736-3745

Rossmann, M.G., Liljas, A., Brändén, C.-I. & Banaszak, L.J. (1975) in *the Enzymes* (Boyer, P.D. ed.), pp. 61-102, Academic Press, London

Rudolph, R., Gerschitz, J. & Jaenicke, R (1978) *Eur. J. Biochem.* **87**, 601-606

Runswick, M.J. & Harris, J.I. (1978) *FEBS Lett.* **92**, 365-367

Russell, D.W., Smith, M., Williamson, V.M. & Young, E.T. (1983) *J. Biol. Chem.* **258**, 2674-2682

Saiki, R.K., Scharf, S., Faloona, F., Mullis, K.B., Horn, G.T., Erlich, H.A. & Arnheim, N. (1985) *Science* **230**, 1350-1354

Sakoda, H. & Imanaka, T. (1992) *J. Bacteriol.* **174**, 1397-1402

Sambrook, J., Fristsch, E.F. & Maniatis, T. (1989) *Molecular Cloning A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 2nd Ed.,

Sanger, F., Nicklen, S. & Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74**, 5463-5467

Schaaff, I., Heinisch, J. & Zimmermann, F.K. (1989) *Yeast.* **5**, 285-290

Schatz, G. & Dobberstein, B. (1996) *Science* **271**, 1519-1526

Schläpfer, B.S. & Zuber, H. (1992) *Gene* **122**, 53-62

Schmidt, C. & Beyersmann, D. (1999) *Arch. Biochem. Biophys.* **364**, 91-98

Schneider, G., Eklund, H., Cedergren-Zeppezauer, E. & Zeppezauer, M. (1983) *Proc. Natl. Acad. Sci. U. S. A.* **80**, 5289-5293

Schrover, J.M.J., Frank, J., Van Wielink, J.E. & Duine, J.A. (1993) *Biochem. J.* **290**, 123-127

Schägger, H. & von Jagow, G. (1987) *Anal. Biochem.* **166**, 368-379

Schümperli, D., McKenney, K., Sobieski, D.A. & Rosenberg, M. (1982) *Cell* **30**, 865-871

Sekhar, V.C. & Plapp, B.V. (1990) *Biochemistry* **29**, 4289-4295

Sellin, S., Holmquist, B., Mannervik, B. & Vallee, B.L. (1991) *Biochemistry* **30**, 2514-2518

Serrano, L., Horovitz, A., Avron, B., Bycroft, M. & Fersht, A.R. (1990) *Biochemistry* **29**, 9343-9352

Shaw, J.P., Rekik, M., Schwager, F. & Harayama, S. (1993) *J. Biol. Chem.* **268**, 10842-10850

Shearer, G.L., Kim, K., Lee, K.M., Wand, C.K. & Plapp, B.V. (1993) *Biochemistry* **32**, 11186-11194

Sheehan, M.C., Bailey, C.J., Dowds, B.C. & McConnell, D.J. (1988) *Biochem. J.* **252**, 661-666

Shine, J. & Dalgarno, L. (1974) *Proc. Natl. Acad. Sci. U. S. A.* **71**, 1342-1346

Sidman, K.E., George, D.G., Barker, W.C. & Hunt, L.T. (1988) *Nucleic Acids Res.* **16**, 1869-1871

Sigler, P.B., Xu, Z., Rye, H.S., Burston, S.G., Fenton, W.A. & Horwich, A.L. (1998) *Ann. Rev. Biochem.* **67**, 581-608

Silver, S. (1978) in *Bacterial Transport* (Rosen, B.P. ed.), Marcel Dekker, Inc., New York

Silver, S. & Walderhaug, M. (1992) *Microbiol. Rev.* **56**, 195-228

Silverstein, E. & Boyer, P.D. (1964) *J. Biol. Chem.* **239**, 3908-3914

Simpson, R.T. & Vallee, B.L. (1968) *Biochemistry* **7**, 4343-4350

Sofer, W. & Ursprung, H. (1968) *J. Biol. Chem.* **243**, 3110-3115

Sone, N., Tsuchiya, N., Inoue, M. & Noguchi, S. (1996) *J. Biol. Chem.* **271**, 12457-12462

Sprengel, R., Reiss, B. & Schaller, H. (1985) *Nucleic Acids Res.* **13**, 893-909

Sridhara, S., Wu, T.T., Chused, T.M. & Lin, E.C.C. (1969) *J. Bacteriol.* **98**, 87-95

Steinbüchel, A. & Schlegel, H.G. (1984) *Eur. J. Biochem.* **141**, 555-564

Stephens, C.M. & Bauerle, R. (1991) *J. Biol. Chem.* **266**, 20810-20817

Stewart, M.J., McBride, M.S., Winter, L.A. & Duester, G. (1990) *Gene* **90**, 271-279

Stock, J.B., Stock, A.M. & Mottonen, J.M. (1990) *Nature* **344**, 395-400

Stone, C.L., Burnell, J.C., Li, T.-K. & Bosron, W.F. (1989) *Prog. Clin. Biol. Res.* **290**, 155-165

Stone, C.L., Li, T.-K. & Bosron, W.F. (1989) *J. Biol. Chem.* **264**, 11112-11116

Studier, F.W. & Moffatt, B.A. (1986) *J. Mol. Biol.* **189**, 113-130

Studier, F.W., Rosenberg, A.H., Dunn, J.J. & Dubendorff, J.W. (1990) *Methods Enzymol.* **185**, 60-89

Sun, H.-W. & Plapp, B.V. (1992) *J. Mol. Evol.* **34**, 522-535

Sund, H. & Theorell, H. (1963) in *The Enzymes Vol. 7* (Boyer, P.D., Lardy, H. & Myrbäck, K. eds.), pp. 25-83, Academic Press , London

Svensson, S., Hedberg, J.J. & Höög, J.-O. (1998) *Eur. J. Biochem.* **251**, 236-243

Sytkowski, A.J. & Vallee, B.L. (1976) *Proc. Natl. Acad. Sci. U. S. A.* **73**, 344-348

Tabakoff, B. & Erwin, V.G. (1970) *J. Biol. Chem.* **245**, 3263-3268

Tabor, S. & Richardson, C.C. (1985) *Proc. Natl. Acad. Sci. U. S. A.* **82**, 1074-1078

Tabor, S. & Richardson, C.C. (1989) *J. Biol. Chem.* **264**, 6447-6458

Takagi, M., Imanaka, T. & Aiba, S. (1985) *J. Bacteriol.* **163**, 824-831

Takemura, H., Kondo, K., Horinouchi, S. & Beppu, T. (1993) *J. Bacteriol.* **175**, 6857-6866

Tamarit, J., Cabiscol, E., Aguilar, J. & Ros, J. (1997) *J. Bacteriol.* **179**, 1102-1104

Tarr, G.E., Beecher, J.F., Bell, M. & McKean, D.J. (1978) *Anal. Biochem.* **84**, 622-627

Teipel, J.W. & Koshland, D.E. (1971) *Biochemistry* **10**, 792-805

Thatcher, D.R. (1980) *Biochem. J.* **187**, 875-883

Theorell, H. & Chance, B. (1951) *Acta Chem. Scand.* **5**, 1127-1144

Thiers, R.E. (1957) in *Methods of Biochemical Analysis Vol. 5* (Glick, D. ed.), pp. 273-335, Interscience Publishers, Inc., New York

Thorn, J.M., Barton, J.D., Dixon, N.E., Ollis, D.L. & Edwards, K.J. (1995) *J. Mol. Biol.* **249**, 785-799

Udovikova, E.A. & Wojtczak, L. (1998) *Int. J. Biochem. Cell Biol.* **30**, 597-608

Ullmann, A., Jacob, F. & Monod, J. (1967) *J. Mol. Biol.* **24**, 339-343

Vallee, B.L. & Auld, D.S. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87**, 220-224

Van Charldorp, R., Van Kimmenade, A.M.A. & Van Knippenberg, P.H. (1981) *Nucleic Acids Res.* **9**, 4909-4917

Varelas, J.B., Roy, C. & Hering, T.M. (1997) *Arch. Biochem. Biophys.* **347**, 1-8

Velichko, I.V., Volk, S.E., Dudarenkov, V.Yu., Magretova, N.N., Chernyak, V.Ya., Goldman, A., Cooperman, B.S., Lahti, R. & Baykov, A.A. (1995) *FEBS Lett.* **359**, 20-22

Vieille, C., Burdette, D.S. & Zeikus, J.G. (1996) *Biotech. Ann. Rev.* **2**, 1-83

Villaroya, A., Juan, E., Egestad, B. & Jörnvall, H. (1989) *Eur. J. Biochem.* **180**, 191-197

Vonck, J., Arfman, N., de Vries, G.E., Van Beeumen, J., Van Bruggen, E.F.J. & Dijkhuizen, L. (1991) *J. Biol. Chem.* **266**, 3949-3954

de Vries, G.E., Arfman, N., Terpstra, P. & Dijkhuizen, L. (1992) *J. Bacteriol.* **174**, 5346-5353

Wagner, F.W., Burger, A.R. & Vallee, B.L. (1983) *Biochemistry* **22**, 1857-1863

Walter, K.A., Bennett, G.N. & Papoutsakis, E.T. (1992) *J. Bacteriol.* **174**, 7149-7158

Wang, H., Kurochkin, A.V., Pang, Y., Hu, W., Flynn, G.C. & Zuiderweg, E.R.P. (1998) *Biochemistry* **37**, 7929-7940

Watanabe, K., Chishiro, K., Kitamura, K. & Suzuki, Y. (1991) *J. Biol. Chem.* **266**, 24287-24294

Watanabe, K., Masuda, T., Ohashi, H., Mihara, H. & Suzuki, Y. (1994) *Eur. J. Biochem* **226**, 277-283

Webb, M. (1970) *Biochim. Biophys. Acta* **222**, 428-439

Wehner, E.P., Rao, E. & Brendel, M. (1993) *Mol. Gen. Genet.* **237**, 351-358

Weinhold, E.G. & Benner, S.A. (1995) *Protein Eng.* **8**, 457-461

Weiss, B., Jacquemin-Sablon, A., Live, T.R., Fareed, G.C. & Richardson, C.C. (1968) *J. Biol. Chem.* **243**, 4543-4555

Wierenga, R.K., De Maeyer, M.C.H. & Hol, W.G.J. (1985) *Biochemistry* **24**, 1346-1357

Wierenga, R.K., Terpstra, P. & Hol, W.G.J. (1986) *J. Mol. Biol.* **187**, 101-107

Williams, D.C., Van Frank, R.M., Muth, J.W. & Burnett, G.L. (1982) *Science* **215**, 687-689

Williams, N. & Frank, P.H. (1990) *Mol. Biochem. Parasitol.* **43**, 125-132

Williamson, V.M. & Paquin, C.E. (1987) *Mol. Gen. Genet.* **209**, 374-381

Wills, C. (1976) *Nature* **261**, 26-29

Wills, C. & Jörnvall, H. (1979) *Eur. J. Biochem.* **99**, 323-331

Wilson, K.J. & Yuan, P.M. (1989) in *Protein Sequencing a Practical Approach* (Findlay, J.B.C. & Geisow, M.J. eds.), pp. 1-41, IRL Press, Oxford

Winberg, J.-O., Hovik, R., McKinley-McKee, J.S., Juan, E. & Gonzalez-Duarte, R. (1986) *Biochem. J.* **235**, 481-490

Winberg, J.-O. & McKinley-McKee, J.S. (1998) *Biochem. J.* **329**, 561-570

Witzmann, S. & Bisswanger, H. (1998) *Biochim. Biophys. Acta* **1385**, 341-352

Woldringh, C.L. & Nanninga, N. (1985) *Molecular Cytology of Escherichia coli*, Academic Press, London, pp161-197

Wolf, J. & Sharp, R.J. (1981) in *The Aerobic Endospore-Forming Bacteria: Classification and Identification* (Berkeley, R.C.W. & Goodfellow, M. eds.), pp. 251-296, Academic Press, London
Wratten, C.C. & Cleland, W.W. (1965) *Biochemistry* **4**, 2442-2451

Wu, C.Y., Mote, J.Jr. & Brennan, M.D. (1990) *Genetics* **125**, 599-610

Wu, L. & Welker, N.E. (1991) *J. Bacteriol.* **173**, 4877-4888

Xia, Z-x., Dai, W-w., Zhang, Y., White, S.A., Boyd, G.D. & Mathews, F.S. (1996) *J. Mol. Biol.* **259**, 480-501

Xie, P., Parsons, S.H., Speckhard, D.C., Bosron, W.F. & Hurley, T.D. (1997) *J. Biol. Chem.* **272**, 18558-18563

Xu, Q.-Y. & Shively, J.E. (1988) *Anal. Biochem.* **170**, 19-30

Yanisch-Perron, C., Viera, J. & Messing, J. (1985) *Gene* **33**, 103-119

Yanofsky, C., Platt, T., Crawford, I.P., Nichols, B.P., Christie, G.E., Horowitz, H. & Van Cleemput, M. (1981) *Nucleic Acids Res.* **9**, 6647-6668

Yoon, K.H. & Pack, M.Y. (1990) *Nucleic Acids Res.* **18**, 187-187

Yoshida, A. (1990) *Prog. Clin. Biol. Res.* **344**, 327-340

Young, E.T. & Pilgrim, D. (1985) *Mol. Cell Biol.* **5**, 3024-3034

Youngleson, J.S., Jones, W.A., Jones, D.T. & Woods, D.R. (1989) *Gene* **78**, 355-364

Yu, J., Donoviel, M.S. & Young, E.T. (1989) *Mol. Cell Biol.* **9**, 34-42

Zaghloul, T.I., Kawamura, F. & Doi, R.H. (1985) *J. Bacteriol.* **164**, 550-555

Zhu, X., Byrnes, M., Nelson, J.W. & Chang, S.H. (1995) *Biochemistry* **34**, 2560-2565