https://theses.gla.ac.uk/

Theses Digitisation:

https://www.gla.ac.uk/myglasgow/research/enlighten/theses/digitisation/

This is a digitised version of the original print thesis.

Copyright and moral rights for this work are retained by the author

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

This work cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Enlighten: Theses
https://theses.gla.ac.uk/
research-enlighten@glasgow.ac.uk

Biochemical and Molecular Characterisation of Dihydrolipoamide Dehydrogenase from Potato Mitochondria

A thesis submitted for the degree of

DOCTOR OF PHILOSOPHY

Saiqa Saeed Khan, B.Sc. (Hons.)

Division of Biochemistry and Molecular Biology,

December, 1996.

UNIVERSITY
of
GLASGOW

ProQuest Number: 10391510

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.

ProQuest 10391510

Published by ProQuest LLC (2017). Copyright of the Dissertation is held by the Author.

All rights reserved.
This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

Thesis
10822
Copy 2

GLASGOW
UNIVERSITY
LIBRARY

For

Mum and Dad

Abstract

Dihydrolipoamide dehydrogenase (E3) is structurally and functionally related to the group of enzymes known as the pyridine nucleotide-disulphide oxidoreductases. It is a homodimer with a M_r of approx. 2 x 50 000, each subunit containing one non-covalently bound FAD molecule and a redox active disulphide. The importance of E3 lies in its function as an integral component of the 2-oxoacid dehydrogenase complexes and the glycine decarboxylase complex.

In mammals and some prokaryotes, a common E3 enzyme forms part of all the complexes; however, in recent years a variety of complex specific isoforms have been isolated from certain prokaryotes. Additionally, E3 enzymes which are not associated with 2-oxoacid dehydrogenase complexes have been discovered suggesting an as yet unknown function. Most plants investigated to date contain two organelle specific isoforms of E3; one found in mitochondria and the other in chloroplasts. In pea, a single E3 enzyme complements the activities of all the complexes in mitochondria. The discovery in this laboratory of three novel isoforms of E3 in the mitochondria of potato tubers and subsequently barley leaves suggests a possible physiological role in which each functions preferentially or exclusively with one of the multienzyme complexes. The isoforms (α_2, $\alpha\beta$ and β_2) appear to arise as a result of the various combinations of two closely related polypeptides. A modified procedure for purification of the potato tuber E3 isoforms enabled the isolation of sufficient protein for the determination of their precise molecular masses. By electrospray mass spectrometry it was found that the α polypeptide had a M_r of 49 446 compared to 49 561 for the β polypeptide. Although the α polypeptide was the smaller form, it migrated more slowly than the β polypeptide through SDS-PAGE.

Reconstitution assays analysed the possible complex specific roles of the E3 mitochondrial isoforms from potato tubers. Bovine heart OGDC was employed due to the inherent problems in the purification of the complexes from plants. All three isoforms were unable to promote the activity of OGDC stripped of its native

E3 component; however competitive binding assays demonstrated that they interacted weakly with the E1/E2 subcomplex of OGDC. Preferential binding of one or more isoforms to the subcomplex was not shown and valid conclusions could not be drawn. Of interest was the comparison of reconstitution ability of a variety of E3s. Bovine mucosal and porcine heart E3 were found to give poor reconstitution with bovine heart E1/E2 subcomplex (15-20%) and yeast E3 did not promote any complex activity although they exhibited tight binding. Inspite of the high sequence conservation between these E3s, subtle differences in catalytic and/or binding orientations between bovine heart E1/E2 and E3 from different species are evident. Similar results were obtained with PDC by other researchers.

Cloning the possible genes for the α and β polypeptides was carried out as a means to establishing the physiological roles of the isoforms. All the clones obtained by amplification of cDNA from potato leaf RNA and screening a potato leaf cDNA library had identity with the β polypeptide by their deduced amino acid sequences and predicted M_r values. This represented only the second E3 to be cloned from plants. The open reading frame consisted of 1512 bp encoding a protein with 470 residues and a 34 residue leader sequence. Expression of a cDNA, truncated at the 5′ end, produced insoluble protein. This was possibly due to the absence of the N-terminal region causing incorrect folding and lack of FAD binding.

The relationship of the α polypeptide to the β polypeptide remains to be established; however, they are expected to have a high degree of sequence identity. The suggestion that the isoforms have distinct properties would imply that the α polypeptide is a distinct gene product. However, the similarities between the polypeptides at the N-terminus, M_r values and their strong cross reaction to yeast E3 antisera would suggest that the two forms have arisen as a result of alternate splicing or an in-frame deletion of one internal amino acid by another event. A post-translational modification involving the addition of a chemical group was eliminated as the cDNA encoding the larger polypeptide (β) had been cloned.

Acknowledgements

I wish to express sincere thanks to my supervisor, Professor Gordon Lindsay, for his guidance, encouragement and time while I was a postgraduate student. I acknowledge the BBSRC for financial support during three years of my Ph.D.

I would like to thank Roger Adams, John Coggins, Ian Hamilton, John Kusel and Sanya Sanderson for their added advice and critical assessment while writing my thesis. I am additionally grateful to Roger for providing me with extra bench space in his laboratory during the final year of my project.

I am indebted to Sriharsa Pradhan who was responsible for giving me a crash course in molecular biology techniques. I thank him for his enthusiasm and giving me his time unselfishly. Thanks also to Janet Greenwood, Andy Elwell who never failed to solve my problems with computers, Tino Krell for the electrospray mass data and Jason Ward for figures on page 4.

My flatmates, who have all lived with me during my Ph.D, are thanked from the bottom of my heart for their friendship and emotional support over the years: Maureen Mackenzie, Friederike Seydel, Sanya Sanderson, Emily Macdonald and Jackie Paterson.

My warmest thanks also to the following friends who have made my seven years in Glasgow, including my years in the lab, unforgettable and full of fun: Stefan Beil, Belen Calvo, Ruth Fullerton, Colin Johnson, Frank Junker, Tino Krell, Tim Lever, Heather Lindsay, Graham McCartney, Alirio and Anneka Melendez, Clare Miller, Gillian Muir, Romena Qazi, Jaqueline Rice, Karin Rossnagel, Karin Schmidl, Anette Sommer and Jason Ward.

I am deeply indebted to my parents for their emotional and financial support throughout my education. I wish to give them and the rest of my family, including my brothers Rehan and Adil, very special thanks for their unwavering support.

Table of Contents

Abstract		i
Acknowledgements		iii
Table of Contents		iv
List of Figures		xi
List of Tables		xvi
Abbreviations		xvii
Chapter 1	**Introduction**	1
1.1	Dihydrolipoamide Dehydrogenase	2
1.2	The 2-Oxoacid Dehydrogenase Complexes	2
1.2.1	Enzyme Reaction	3
1.2.2	Quaternary Structure	5
1.3	Constituent Enzymes of the 2-Oxoacid Dehydrogenase Complexes	8
1.3.1	2-Oxoacid Dehydrogenase (E1)	8
1.3.2	Dihydrolipoamide Acyltransferase (E2)	9
1.3.2.1	Domain Structure of E2	10
1.3.2.2	Peripheral Subunit-binding Domain	12
1.3.2.3	Lipoyl Domains	12
1.3.2.4	3-Dimensional Structure Determination	14
1.3.3	Protein X	16
1.3.4	Dihydrolipoamide Dehydrogenase (E3)	18
1.3.4.1	Domain Structure	21
1.4	Regulation	23
1.5	Disease States Associated with the 2-Oxoacid Dehydrogenase Complexes	27
1.6	Plant 2-Oxoacid Dehydrogenase Complexes	29
1.7	Plant Mitochondrial PDC	30

1.7.1	Regulation of Mitochondrial PDC	32
1.7.2	Effect of Light and Photosynthesis on PDC Activity	36
1.8	Plant Plastid PDC	37
1.8.1	Regulation Of Plastid PDC	38
1.9	Plant OGDC	39
1.10	Plant BCDC	40
1.11	Glycine Decarboxylase (GDC)	41
1.12	Aims of this Thesis	47
Chapter 2	**Materials and Methods**	**49**
2.1	Materials	50
2.1.1	Plants	50
2.1.2	Animals	50
2.1.3	cDNA Library	50
2.1.4	Bacterial and Bacteriophage Strains	50
2.1.5	Plasmid Vectors	51
2.1.6	Chemicals	56
2.1.7	Enzymes	56
2.1.8	Molecular Size and Weight Markers	57
2.1.9	Synthetic Oligonucleotides	57
2.1.10	Photographic Materials	57
2.1.11	Miscellaneous	57
2.2	Methods	58
2.2.1	Growth of Plants	58
2.2.2	Determination of Protein Concentration	58
2.2.3	Dialysis of Samples	58
2.2.4	Concentration of samples	59
2.2.5	SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE)	59
2.2.6	Staining and Destaining Gels	60

2.2.7 Western Blotting and Immunological Detection of Proteins by Enhanced Chemiluminescence (ECL) 60
2.2.8 Preparation of Antiserum 62
2.2.9 Purification of Enzymes 62
2.2.9.1 Purification of Pyruvate Dehydrogenase Complex (PDC) and 2-Oxoglutarate Dehydrogenase Complex (OGDC) from Bovine Heart. 62
2.2.9.2 Separation of E3 from E1/E2 Core of OGDC 64
2.2.9.3 Preparation of a Crude Mitochondrial Dihydrolipoamide Dehydrogenase Extract from Potato Tubers 64
2.2.9.4 Preparation of Diethylaminoethyl Cellulose (DEAE-52) Ion Exchange Chromatography Columns 65
2.2.9.5 Mitochondrial Isolation from Potato Tubers 65
2.2.9.6 Mitochondrial Isolation from Cauliflower Florets 66
2.2.9.7 Purification of E3 from a Variety of Sources 66
2.2.10 Enzyme Assays 67
2.2.10.1 2-Oxoacid Dehydrogenase Complexes 67
2.2.10.2 Dihydrolipoamide Dehydrogenase 67
2.2.10.3 Preparation of Dihydrolipoamide 68
2.2.10.4 Reconstitution Studies on Bovine Heart OGDC 68
2.2.11 Media for the Growth of Bacteria and Bacteriophage 69
2.2.12 Storage of DNA 70
2.2.13 Phenol/Chloroform Extraction 70
2.2.14 Ethanol or Isopropanol Precipitation of DNA 70
2.2.15 Restriction Mapping 71
2.2.16 DNA Electrophoresis in Agarose Gels 71
2.2.17 Electrophoresis of RNA using Formaldehyde Gels 72
2.2.18 Recovery of DNA from Low Melting Point Agarose 72

2.2.19	Recovery of DNA from Non-Denaturing Polyacrylamide Gels	73
2.2.20	Small Scale Preparation of Plasmid DNA	73
2.2.21	Large Scale Preparation of Plasmid DNA	74
2.2.22	The Polymerase Chain Reaction (PCR)	74
2.2.23	First-Strand cDNA Synthesis and PCR Amplification using Degenerate and Nested Primers	75
2.2.23.1	First-Strand cDNA Synthesis from Plant RNA	75
2.2.23.2	Second-Strand Synthesis and Amplification of DNA by PCR	76
2.2.24	Quantification of Nucleic Acids	76
2.2.24.1	Absorption at 260 nm	76
2.2.24.2	Fluorometry	77
2.2.25	Dephosphorylation of Plasmid DNA	77
2.2.26	Ligations	78
2.2.27	Preparation and Transformation of Competent Cells	78
2.2.27.1	Hanahan Method	78
2.2.27.2	Calcium Chloride	78
2.2.28	Extraction and Purification of Total RNA from Plant Tissue	79
2.2.29	Transfer and Detection of Nucleic Acids	80
2.2.29.1	Radiolabelling of DNA Fragments	81
2.2.29.2	Southern Blotting and Hybridisation	81
2.2.29.3	Northern Blotting and Hybridisation	82
2.2.30	Screening a Lambda ZAP II cDNA Library	82
2.2.30.1	Plating and Titering the Library	82
2.2.30.2	Plaque Lifts	83
2.2.30.3	Selection of Bacteriophage Lambda Plaques	84
2.2.30.4	Preparation of Plate Lysate Stocks	84
2.2.30.5	Purification of Recombinant Bacteriophage DNA	85

2.2.31 *In Vivo* Excision of the pBluescript Phagemid from the Lambda ZAP II Vector 85
2.2.32 Generation of Clones for Sequencing by Directed Deletions 86
2.2.32.1 Restriction Digestion of Plasmid DNA 86
2.2.32.2 Exo III Digestion, Ligation and Transformation 87
2.2.33 Automated Dyedeoxy Terminator Cycle Sequencing 87
2.2.34 Expression of Heterologous Proteins 88
2.2.34.1 Subcloning the Target Gene into an Expression Vector 88
2.2.34.2 Transformation of the Expression Vector to an Expression Host Strain 89
2.2.34.3 Expression and Purification of Heterologous Protein 90
2.2.35 Molecular Mass Determination by Electrospray Mass Spectrometry 90
Chapter 3 Purification and M_r Analysis of Dihydrolipoamide Dehydrogenase (E3) Isoforms from Potato Tubers 92
3.1 Introduction 93
3.2 Results 97
3.2.1 Preparation of a Crude Organellar Extract from Potato Tubers 97
3.2.2 Elution of E3 from a DEAE-52 Ion Exchange Column 99
3.2.3 Resolution of the E3 Enzyme by Source 15Q Ion Exchange Chromatography 103
3.2.4 Determination of the Subunit Molecular Mass of E3 by Electrospray Mass Spectrometry (Es-Ms) 106
3.3 Discussion 111

Chapter 4	**Analysis of the Possible Complex Specific Roles of the Potato Tuber E3 Isoforms by Reconstitution of Intact OGDC Activity**	**114**
4.1	Introduction	115
4.2	Results	118
4.2.1	Dissociation of Bovine Heart OGDC	118
4.2.2	Time Course for Reconstitution of OGDC Activity	122
4.2.3	Reconstitution of OGDC Activity by Homologous and Heterologous E3 Enzymes	125
4.2.4	E3 Binding Analysis	127
4.2.5	Partial Purification of OGDC from Potato Tubers	129
4.2.6	Partial Purification of OGDC from Cauliflower Mitochondria	131
4.3	Discussion	135
Chapter 5	**Cloning the Potato Mitochondrial Dihydrolipoamide Dehydrogenase (E3) Gene(s)**	**138**
5.1	Introduction	139
5.2	Strategy for Cloning the E3 Gene	142
5.3	Results	143
5.3.1	RNA Preparation from Potato Leaves	143
5.3.2	First-Strand cDNA Synthesis	145
5.3.3	PCR Amplification of the 5´-Region of the ORF of the Potato Mitochondrial E3 Gene	145
5.3.3.1	Amplification of First-Strand cDNA	146
5.3.3.2	Cloning and sequencing the 80 Base Pair Fragment	148
5.3.4	Nested-PCR Amplification of the Full Length cDNA Encoding the Potato Mitochondrial E3	150
5.3.4.1	Design of the Nested-PCR Primers	150
5.3.4.2	PCR Amplification of First-Strand cDNA	151

5.3.4.3 Nested-PCR Of The Amplified DNA Products 153
5.3.4.4 Cloning the Nested PCR Product 156
5.3.4.5 Confirmation of Identity by Partial Sequencing and Restriction Mapping of the Clones Containing the *Taq* Amplified Insert 159
5.3.5 Generation Of Clones For Sequencing By Directed Deletions 162
5.3.5.1 DNA Sequence the pTaq22 and pTaq16 Subclones 164
5.4 Discussion 174
Chapter 6 Heterologous Expression of the Potato Mitochondrial Dihydrolipoamide Dehydrogenase Gene(s) in *Escherichia coli* **177**
6.1 Introduction 178
6.2 Results 179
6.2.1 Subcloning the Taq Inserts into pET28c Expression Vectors 179
6.2.2 Time Course of Expression from the pET28c Clones 181
6.2.3 Purification of the Expressed Dihydrolipoamide Dehydrogenase from BL21(DE3) Cells 188
6.3 Discussion 192
Chapter 7 Screening a Potato Leaf cDNA Library for Mitochondrial Dihydrolipoamide Dehydrogenase Genes **195**
7.1 Introduction 196
7.2 Results 197
7.2.1 Titre of the Potato Leaf cDNA Library 197
7.2.2 Probe Preparation 197
7.2.3 Screening the Potato cDNA Library for the E3 Gene 198
7.2.4 PCR Analysis of the Isolated Library Clones 201
7.2.5 *In Vivo* Excision of the pBluescript Phagemid from Lambda ZAP II 205

7.2.6 Restriction Digest Analysis 206

7.2.7 Sequence Analysis of the Potato Leaf cDNA Clones 208

7.3 Discussion 213

Chapter 8 Discussion and Future Research 222

8.1 Dihydrolipoamide Dehydrogenase (E3) from Potato Tuber Mitochondria 223

8.1.1 Purification and M_r Analysis of Potato Mitochondrial E3 223

8.1.2 Complex Specificity of the Mitochondrial E3 isoforms 224

8.2 Cloning the Potato Mitochondrial E3 gene(s) 226

8.2.1 Cloning by PCR Amplification of E3 cDNA 226

8.2.2 Isolation of E3 clones from a Potato leaf cDNA Library 227

8.3 Future Research 230

References 231

List of Figures

Chapter 1 Introduction

1.1 Location of the 2-oxoacid dehydrogenase complexes in central metabolism 4

1.2 Reaction scheme for the oxidative decarboxylation of the 2-oxoacids by respective dehydrogenase complexes 4

1.3 Interpretative models of the quaternary structure of dihydrolipoamide acyltransferase (E2) 6

1.4 Schematic representation of the domain structure of the E2 and protein X components 11

1.5 Three-dimensional structure of the *B. stearothermophilus* lipoyl domain 15

1.6 Comparison of the chain fold of human glutathione reductase and human E3 22

1.7 Schematic representation of the covalent modification of pyruvate dehydrogenase and its control by effectors 26

1.8 Model illustrating effectors of the reversible phosphorylation of pea leaf mitochondrial PDC 34

1.9 Interactions of the photosynthetic, photorespiratory and citric acid cycles 43

1.10 The reaction mechanism of glycine decarboxylase multienzyme complex 45

Chapter 2 Materials and Methods

2.1 pBluescript II KS (+/-) phagemid 52

2.2 pET-28a-c (+) vectors 53

2.3 pBluescript SK (+/-) phagemid 54

2.4 Lambda ZAP II vector 55

Chapter 3 Purification and M_r Analysis of Dihydrolipoamide Dehydrogenase (E3) Isoforms from Potato Tubers

3.1 Elution of E3 activity from the DEAE-52 ion exchange column 101

3.2 Analysis of purification of potato tuber E3 from pure mitochondria by affinity chromatography 102

3.3 Resolution of the DEAE-52 purified potato extract by Source 15Q ion exchange chromatography 104

3.4 Purification of the E3 isoforms from a crude organellar pellet from potato tubers, using the modified protocol 105

3.5A Electrospray mass spectra of the $\alpha\beta$ heterodimeric E3 isoenzyme from potato tuber mitochondria 108

3.5B MaxEnt deconvolution of electrospray mass spectra of the $\alpha\beta$ heterodimeric E3 isoenzyme from potato tuber mitochondria 108

3.6A MaxEnt deconvolution of electrospray mass spectra of the β_2 E3 isoenzyme from potato tuber mitochondria 110

3.6B MaxEnt deconvolution of electrospray mass spectra of the α_2 E3 isoenzyme from potato tuber mitochondria 110

Chapter 4 Complex Specificity of Potato Tuber E3 Isoforms

4.1 Dissociation of bovine heart OGDC 119

4.2 Dissociation of bovine heart OGDC under high salt (2M NaCl) at alkaline pH (pH9) 120

4.3 Dissociation of bovine heart OGDC at 1M NaCl at neutral pH 121

4.4 Time course of reconstitution of bovine heart E1/E2 core with native E3 123

4.5 Time course of reconstitution of bovine heart PDC activity by re-assembly of E2/X and E1/E3 subcomplexes 124

4.6 Comparison of reconstitution of bovine heart OGDC activity by re-assembly of native and heterologous E3s with the E1/E2 subcomplex 126

4.7 Percent decrease in reconstituted OGDC activity resulting from competition between the native E3 and potato tuber E3 isoforms 128

4.8 Elution profile from a Superdex 200 gel filtration column of PDC, OGDC and E3 activities from a potato mitochondrial extract 130

4.9 Stimulation of cauliflower mitochondrial OGDC and PDC by reconstitution with increasing amounts of porcine heart E3 132

4.10 Reconstitution of OGDC activity from cauliflower mitochondria with increasing amounts of pig heart E3 134

Chapter 5 Cloning the Potato Mitochondrial Dihydrolipoamide Dehydrogenase (E3) Gene(s)

5.1 Analysis of the RNA preparation from potato leaves 144

5.2 N-terminal sequence obtained for the potato mitochondrial α and β polypeptides, showing positions of the degenerate primers A and B 145

5.3 Analysis of PCR products obtained by *Taq* polymerase amplification of first-strand cDNA 147

5.4 Analysis of clones containing the 80 bp insert 149

5.5 DNA sequence of clones pBkC, pBkD and pBkF representing amino acids 5-28 of the mature E3 protein 148

5.6 PCR amplification of first-strand cDNA with *Pfu* (native and recombinant) DNA polymerase 152

5.7 Analysis of the potato leaf cDNA re-amplified by nested-PCR using native and recombinant *Pfu* polymerase 154

5.8 Analysis of the potato leaf cDNA re-amplified by nested-PCR using *Taq* DNA polymerase 155

5.9 Gel purified PCR products amplified with *Pfu* or *Taq* DNA polymerase for cloning purposes 157

5.10 Analysis of clones obtained by transformation with the recombinant *Taq* and the *Pfu* PCR products 157

5.11 Analysis of clones obtained by transformation of SURE *E. coli* cells with the recombinant *Pfu* PCR product 158

5.12 Restriction digests of the recombinant plasmids containing the *Taq* PCR product 160

5.13a Restriction map of clone pTaq22 showing orientation of insert in pBluescript KS+ 161

5.13b Restriction map of clone pTaq15 showing orientation of insert in pBluescript KS+ 161

5.14 Schematic for the generation of nested-deletions of pTaq16 163

5.15a Time course of Exo III deletions of clone pTaq16 produced at 30 s intervals 165

5.15b Subclones of clone pTaq16 containing a decreasing ladder of inserts 165

5.16 DNA sequence comparison of the clones pTaq22, pTaq23 and pTaq16 169

Chapter 6 Heterologous Expression of the Potato Mitochondrial Dihydrolipoamide Dehydrogenase Gene(s) in *Escherichia coli*

6.1 Reading frame for expression of the pET28c expression construct, containing the E3 gene 180

6.2 Analysis of the *Taq* amplified inserts subcloned from pBluescript into the pET28c expression vector 182

6.3 Analysis of the time course of expression, at 37°C from clones pTaqE13 and pTaqE15 in *E. coli* Bl21(DE3) or BL21(DE3)pLysS cells 185

6.4a Time course of expression at 37°C, from clones pTaqE23, pTaqE13, pTaqE14 and pTaqE16 constructs in *E. coli* BL21(DE3) cells 186

6.4b Time course of expression at 37°C, from clones pTaq23 and pTaq22 constructs in *E. coli* BL21(DE3) cells 187

6.5 Immunological identification of the protein expressed from clones pTaqE13, pTaqE14 and pTaqE16 189

6.6 Analysis of the insoluble, expressed protein 190

6.7 Time course of expression from pTaqE22 over 30 min at 30°C and 37°C 190

Chapter 7 Screening a Potato Leaf cDNA Library for Mitochondrial Dihydrolipoamide Dehydrogenase Genes

7.1 DNA probe (900 bp) prepared from a *Eco*RI/*Kpn*I digest of pTaq22 199

7.2 Autoradiographs showing the progress of screening the potato leaf cDNA library for a clone containing the putative E3 gene 200

7.3 PCR analysis of the 5′ region of the library clones using the degenerate primers designed from the N-terminal region of the E3 enzyme 203

7.4 PCR amplification of the 5′ region of the library clones - priming with T3 and an E3 gene specific primer (AGSP) 204

7.5a Restriction map of clones pB9-3, pB10-1, pB9-2, pB6-1 and pB8-1 207

7.5a Restriction map of clone pB3-1 207

7.6 Analysis of the purified plasmids pB3-1, pB9-3 and pB10-1 used for sequencing the E3 gene from potato 209

7.7 Nucleotide sequence of potato E3 and its deduced amino acid sequence 210

7.8 Primary amino acid sequence comparison of potato E3 (clone pB3-1) and E3s from other sources 218

Chapter 8 Discussion and Future Research

8.1 Unrooted evolutionary tree of potato E3 in relation to other E3s, based on amino acid sequence 229

Appendix

A DNA sequence comparison of the clones pB3-1, pB9-3 and the PCR derived clones pTaq22, pTaq23 and pTaq16

B Amino acid comparison of the library clones pB3-1 and pB9-3 with the *Taq* amplified clones pTaq22, pTaq23 and pTaq16

List of Tables

1.1 Examples of complexes showing cubic morphology (octahedral symmetry) or dodecahedral morphology (icosahedral symmetry) about the E2 core 7

3.1 Purification of potato mitochondrial E3 98

3.2 A comparison of the M_r values from a variety of E3 subunits from their deduced amino acid sequences 113

5.1 Sequence of the forward and reverse primers used to amplify the 5′ region of the potato mitochondrial E3 gene 146

5.2 Sequence of the nested forward primers used to amplify the E3 cDNA from first-strand cDNA 150

5.3 Sequencing primers designed from the consensus sequence of the sense strand of pTaq22 166

6.1 Results from the induction of expression of the Taq clones 183

7.1 Primer sequences used for the amplification of the 5′ upstream sequences of the isolated library clones 205

7.2 Percent identity and similarity of several E3s with potato E3 217

Abbreviations

amp	ampicillin
approx.	approximately
BCDC	branched-chain 2-oxoacid dehydrogenase complex
bp	base pair
BSA	bovine serum albumin
CTAB	hexadecyltrimethyl-ammonium bromide
Da	daltons
DEPC	diethyl pyrocarbonate
DNA	deoxyribonucleic acid
DTT	dithiothreitol
E1	2-oxoacid dehydrogenase
E2	dihydrolipoamide acyltransferase
E3	dihydrolipoamide dehydrogenase
EDTA	ethylenediaminetetra-acetic acid
EGTA	ethylene glycol-bis(b-aminoethyl ether)
Es-Ms	electrospray mass spectrometry
Fig.	figure
GDC	glycine decarboxylase complex
GR	glutathione reductase
IPTG	isopropyl-thio-galactopyranoside
kb	kilobase
LMP	low melting point agarose
M_r	relative molecular mass
MSUD	maple syrup urine disease
NAD^+	nicotinamide adenine dinucleotide, oxidised form
NADH	nicotinamide adenine dinucleotide, reduced form
$NADP^+$	nicotinamide adenine dinucleotide phosphate, oxidised form

OGDC	2-oxoglutarate dehydrogenase complex
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PBC	primary biliary cirrhosis
PCR	polymerase chain reaction
PDC	pyruvate dehydrogenase complex
PEG	polyethylene glycol
pfu	plaque forming unit
PMSF	phenylmethylsulphonylfluoride
psi	pounds per square inch
PVP	polyvinylpyrrolidone
RNA	ribonucleic acid
SHMT	serine hydroxymethyltransferase
SDS	sodium dodecyl sulphate
TCA	tricarboxylic acid cycle
TEMED	N,N,N',N'-tetramethylethylenediamine
TES	(N-tris[hydroxymethyl]-methyl-2-aminoethane-sulphonic acid
tet	tetracyclin
ThDP	thiamine diphosphate
THF	tetrahydrofolate
Tris	2-amino-2-(hydroxymethyl)-1,3-propandiol
Tween 20	polyoxyethylenesorbitan monolaureate
TX-100	Triton X-100
vol	volume(s)
v/v	volume to volume
w/v	weight to volume
X	protein X
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

Chapter 1

Introduction

1.1 Dihydrolipoamide Dehydrogenase

Dihydrolipoamide dehydrogenase (E3) belongs to the pyridine nucleotide-disulphide oxidoreductase group of enzymes comprising glutathione reductase, mercuric reductase and trypanothione reductase (Carothers *et al*, 1989). These enzymes are homodimers containing one flavin adenine dinucleotide (FAD) and a redox active disulphide per subunit. The importance of E3 lies in its function as a component of the family of 2-oxoacid dehydrogenase complexes and the related glycine decarboxylase where it reoxidises the lipoamide group, transferring reducing equivalents onto NAD^+ via its FAD cofactor. E3 will be discussed more fully in section 1.3.4. This chapter will begin with an update on the current knowledge of the mammalian and bacterial complexes and their component enzymes since they have been investigated more thoroughly than plant complexes.

1.2 The 2-Oxoacid Dehydrogenase Complexes

E3 is the flavoprotein component of the ubiquitous family of 2-oxoacid dehydrogenase complexes comprising the pyruvate dehydrogenase complex (PDC), 2-oxoglutarate dehydrogenase complex (OGDC) and branched-chain 2-oxoacid dehydrogenase complex (BCDC). Each mammalian complex is located within the mitochondrial matrix, probably loosely associated with the mitochondrial inner membrane. This family of enzymes has been studied extensively in bacteria and mammals because they are model examples of multienzyme complexes, which are defined as self-assembling aggregates of non-covalently interacting proteins, catalysing successive steps in a metabolic pathway. (Perham, 1991, Reed, 1974). Interesting questions arise concerning the assembly of the complexes and the mechanisms for channelling substrates and intermediates between multiple active sites.

Of prime importance is the key regulatory position in central metabolism of the complexes in catalysing essentially irreversible reactions (Yeaman, 1986). A

scheme of the metabolic interaction of the complexes is represented in Fig. 1.1. PDC oxidatively decarboxylates pyruvate to acetyl CoA, committing the entry of carbohydrate into the TCA cycle for subsequent oxidation, with the concomitant generation of ATP. It also has a role in supplying acetyl CoA for a range of biosynthetic pathways. OGDC is a component of the TCA cycle converting 2-oxoglutarate to succinyl CoA and controls the flux of carbon atoms around the latter stages of the cycle. Additionally, OGDC supplies succinyl CoA for the biosynthesis of porphyrins (in mammalian cells), lysine and methionine. BCDC is the only complex with a broad substrate specificity, catalysing the committed step in the degradation of branched-chain amino acids leucine, isoleucine and valine. It is also involved in the catabolism of methionine and threonine. In mammals BCDC has nutritional signficance in controlling the levels of branched-chain amino acids, converting excesses into acyl CoA derivatives and thereby recycling carbon atoms back into central metabolism.

1.2.1 Enzyme Reaction

Each complex is composed of multiple copies of three component enzymes termed E1, E2 and E3. E1 is a substrate specific 2-oxoacid dehydrogenase which has thiamine diphosphate (ThDP) as an essential cofactor. E2 is a complex specific dihydrolipoamide acyltransferase possessing a covalently attached lipoic acid cofactor, and E3 is the FAD-containing dihydrolipoamide dehydrogenase. These three enzymes utilise the same reaction mechanism in each complex to catalyse the co-ordinated oxidative decarboxylation of their respective 2-oxoacid substrates, releasing CO_2 and generating the relevant acyl CoA (Fig. 1.2). E1 catalyses the overall rate-limiting step in the reaction: the decarboxylation of the 2-oxoacid substrate with the formation of a hydroxyacyl-ThDP intermediate (reaction 1). The E1 enzyme then catalyses the reductive acylation of the S^8-thiols of the lipoamide cofactor, covalently bound to the lipoyl domains of the oligomeric E2 enzyme to form an S^8-acyllipoamide intermediate (reaction 2). The acyl group is transferred at the E2 active site to the CoA acceptor (reaction 3), leaving the lipoic

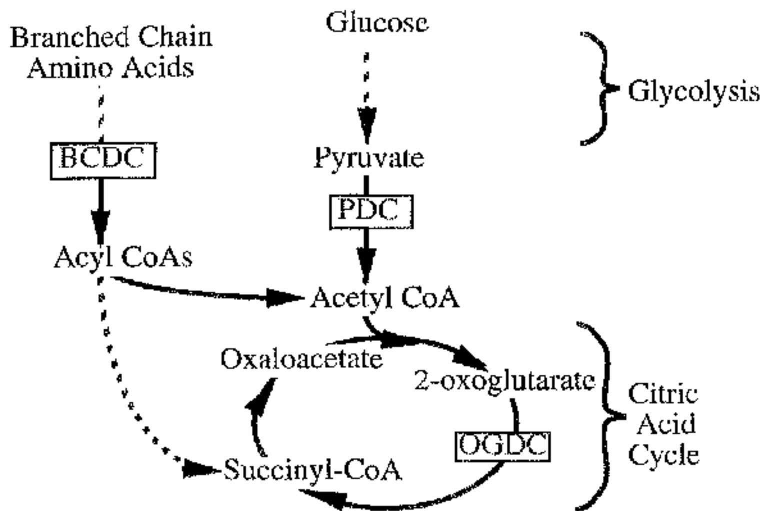
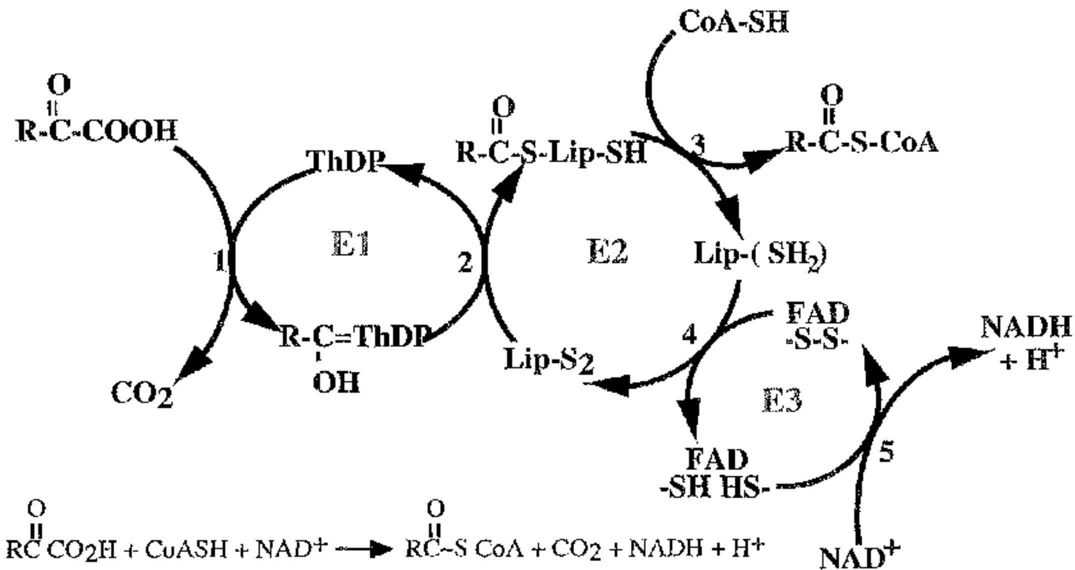


Fig. 1.1: Location of the 2-oxoacid dehydrogenase complexes in central metabolism

PDC, pyruvate dehydrogenase complex; OGDC, 2-oxoglutarate dehydrogenase complex; BCDC, Branched-chain 2-oxoacid dehydrogenase complex. Broken lines indicate that several steps are involved.



$$RCO_2H + CoASH + NAD^+ \longrightarrow RC{\sim}S\,CoA + CO_2 + NADH + H^+$$

Fig. 1.2: Reaction scheme for the oxidative decarboxylation of the 2-oxoacids by respective 2-oxoacid dehydrogenase complexes

E1, 2-Oxoacid Dehydrogenase; E2, Dihydrolipoamide Acyltransferase; E3, Dihydrolipoamide Dehydrogenase; TPP, Thiamine pyrophosphate; Lip, Lipoic acid; R = CH_3 (for pyruvate), $COOH\text{-}CH_2\text{-}CH_2$ (for 2-oxoglutarate) and $(CH_3)_2CHCH_2$ (for 2-ketoisocaproic acid derived from leucine).

acid in the reduced state. This cofactor is re-oxidised by the FAD-containing E3 component (reaction 4) which in turn becomes re-oxidised by NAD^+ as the final electron acceptor (reaction 5). The E3 component, carrying out the same reaction might be expected to be an identical enzyme in all three complexes but this is not always the case and will be discussed later in this chapter.

1.2.2 Quaternary Structure

The 2-oxoacid dehydrogenases have M_r values of several million (5-10 $\times 10^{-6}$) and are seen as particles of 30-40 nm in the electron microscope (Henderson *et al.*, 1979; Oliver & Reed, 1982), being larger than ribosomes. The oligomeric E2 component forms a central, structural core around which multiple copies of E1 and E3 are bound tightly, but non-covalently. It was apparent by electron microscopy that the relationship of the E2 subunits formed two types of distinct polyhedral morphology: the cube and the pentagonal dodecahedron. The cubic morphology is formed by 24 copies of E2 arranged with octahedral symmetry and the dodecahedron by 60 E2 copies arranged with icosahedral symmetry (Fig. 1.3). Table 1.1 provides a summary of the E2 morphology found in complexes from different species. Structural determination of the cubic core of *Azotobacter vinelandii* suggested that the building blocks were eight trimers of E2, forming a hollow cavity, connected to the outside by large channels (Mattevi *et al.*, 1992b).

The E1 of mammalian PDC, BCDC and bacterial (gram-positive) PDC are tetramers consisting of two copies of two non-identical subunits, α and β, whereas in mammalian OGDC and bacterial (gram-negative) OGDC and PDC, E1 exists as a homodimer. For example, bovine heart PDC has about 30 E1 heterotetramers ($\alpha_2\beta_2$) and 6 E3 homodimers which are thought to bind near the 30 edges and to the 12 faces of the pentagonal dodecahedron E2 core, respectively (Reed & Hackert, 1990; Wu & Reed, 1984). *E. coli* PDC has 12 E1 homodimers arranged on the 12 edges of the cube-like E2 core and 6 E3 dimers organised in the 6 faces.

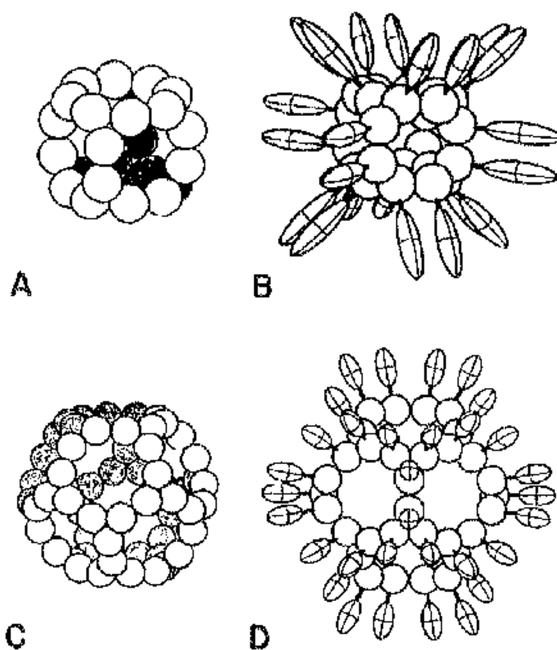


Fig 1.3: Interpretative models of the quaternary structure of dihydrolipoamide acyltransferase (E2)

A) Model of those acyltransferases consisting of 24 subunits arranged in groups of 3 about the 8 vertices of a cube.

B) Model of the 24-subunit acyltransferases illustrating the proposed domain structure. Each of the 24 acyltransferase subunits is represented by one sphere and its attached ellipsoid. The spheres represent the compact, inner core domains, and the ellipsoids represent the extended lipoyl domain.

C) Model of those acyltransferases consisting of 60 subunits arranged in groups of three about the 20 vertices of a pentagonal dodecahedron.

D) Model of the 60-subunit acyltransferases illustrating the proposed domain structure. The figure is viewed down a 2-fold axis of symmetry.

Reproduced from Reed *et al.* (1989).

Table 1.1

Examples of complexes showing cubic morphology (octahedral symmetry) or dodecahedral morphology (icosahedral symmetry) about the E2 core

Morphology	Source of 2-oxoacid dehydrogenase	Reference
24 E2 (octahedral symmetry)	*Escherichia coli* (gram negative) PDC &OGDC	Reed (1974) Danson *et al.* (1979)
and	*Azotobacter vinelandii* PDC & OGDC	Hanemaaijer *et al.* (1988)
Homodimeric E1	Mammalian OGDC	Reed (1974)
	BCDC*	Griffin *et al.* (1988) Hackert *et al.* (1989)
60 E2 (icosahedral symmetry)	Mammalian PDC	Reed (1974)
	Bacillus stearothermophilus PDC	Henderson and Perham (1980)
and	*Bacillus subtilis* PDC	Lowe *et al.* (1983)
Heterotetrameric E1	*Streptococcus faecalis* PDC	Reed (1974)
	Saccharomyces cerevisiae PDC	Keha *et al.* (1982)

* BCDC is an exception as it has a heterotetrameric E1 component

An additional component, termed protein X is found tightly associated with the E2 cores of mammalian and yeast PDC (De Marcucci & Lindsay, 1985; Jilka *et al.*, 1986), but has not been identified in any other 2-oxoacid dehydrogenase complexes (e.g. OGDC and BCDC). It does not appear to have an essential role in catalysis but is thought to involve binding of the E3 to the E2 core (Gopalkrishnan *et al.*, 1989). Protein X was thought to be present in 6 copies in the bovine kidney PDC (Jilka *et al.*, 1986), however, Maeng *et al.* (1994) provided evidence for 12 copies of protein X attached to 12 E3 homodimers in *Saccharomyces cerevisiae*. The presence of 12 copies of X was also recently established in bovine heart PDC by Sanderson *et al.* (1996a), which is consistent with one X subunit with each face of the E2 pentagonal dodecahedron. Cross-linking data of the lipoyl domains from bovine heart PDC tentatively suggested that 6 protein X dimers bind 6 dimers of E3. In addition, bovine heart PDC possesses a tightly bound kinase (1-3 molecules/complex in bovine kidney PDC) and a loosely associated phosphatase, requiring Ca^{2+} ions (approx. 5 molecules/ complex, Yeaman, 1989).

1.3 Constituent Enzymes of the 2-Oxoacid Dehydrogenase Complexes

1.3.1 2-Oxoacid Dehydrogenase (E1)

E1 catalyses the irreversible and rate-limiting step in the oxidative decarboxylation of the 2-oxoacids and thus determines the kinetic parameters of the reaction. It is the main regulatory centre of the complexes via dephosphorylation (activation) and phosphorylation (inactivation) cycles of the E1α component, involving a specific phosphatase and a kinase (Linn *et al.*, 1969).

The E1 can be divided into three groups: E1 from OGDC, E1 from PDC of Gram-negative bacteria and E1 from eukaryotic PDC and BCDC. Sequence homologies between these groups are minimal or non-existent, and even E1 enzymes from OGDC and PDC from the same organisms (e.g *E. coli*) are quite

dissimilar (Darlinson *et al.*, 1984). In contrast, the E2 components of both complexes have striking homologies. The E1 of bacterial and mammalian OGDC and bacterial PDC exist as homodimers (110 000 kDa/polypeptide). However, the E1 components of mammalian PDC and BCDC are heterotetramers ($\alpha_2\beta_2$) consisting of two copies of two non-identical subunits termed α (41 kDa) and β (36 kDa) (Reed, 1974; Perham, 1996). The β subunit anchors E1 to the subunit binding domain of E2.

A common ThDP binding sequence motif exists in E1α which was found by comparisons with other ThDP requiring enzymes, including pyruvate decarboxylase and pyruvate oxidase and transketolase (Hawkins *et al.*, 1989). The characterisation of individually expressed E1α and E1β subunits have shown that both are required for catalytic activity and coenzyme binding (Korotchkina *et al.*, 1995; Jeng *et al.*, 1994; see Korotchkina *et al.*, 1996, for review). Robinson and Chun (1993) proposed that the mode of ThDP binding by human E1 was similar to that of the three ThDP-requiring enzymes named above. The mechanism of ThDP binding by these three enzymes has been determined by analysis of their crystal structures, which confirms that the conserved ThDP binding structure is a common feature. They are all composed of identical subunits (2 for TK; 4 for PD and POX). Two different domains are required to bind the coenzyme, one binding the pyrophosphate group and the other the pyrimidine ring. Each subunit contains both domains, however ThDP is bound in a cleft between subunits so that the pyrophosphate domain belongs to one subunit and the pyrimidine domain to the other subunit. The function of the ThDP binding motif found by Hawkins *et al.* (1989) is to bind the pyrophosphate moiety.

1.3.2 Dihydrolipoamide Acyltransferase (E2)

The E2 component fulfils a number of essential roles in the 2-oxoacid dehydrogenases: 1) it forms the symmetrical core around which multiple copies of E1 and E3 are co-ordinated, 2) it functions as an acyltransferase, catalysing the

formation of the acyl-CoA product, 3) it provides the attachment site for the lipoic acid cofactor which interacts with the E1 and E3 active sites (Yeaman, 1989).

1.3.2.1 Domain Structure of E2

Limited proteolysis studies, mainly of the E2 from *E. coli* PDC and OGDC, revealed the complex, functional domain structure of this component (Bleile *et al.*, 1979; Packman & Perham, 1986) and confirmation was provided when Stephens *et al.* (1983) and Spencer *et al.* (1984) obtained the gene sequences of the E2 enzymes from *E. coli* PDC (*ace*F) and OGDC (*suc*B), respectively (Fig. 1.4). The N-terminal half of *E. coli* PDC E2 consists of three, highly homologous lipoyl domains of about 80 amino acids each. A structurally distinct domain of about 50 amino acids (peripheral subunit-binding domain) follows the lipoyl domains and is involved in the binding of E3 (Packman & Perham, 1986). Removal of these domains by trypsin, leaves a large (29kDa), catalytic domain at the C-terminal end which promotes the acyltransferase reaction, is involved in E1 binding and aggregates to form the core of the complex. The E2 polypeptide of OGDC is highly homologous with that of PDC E2 but has only one lipoyl domain. The independent units making up E2 are connected by flexible hinge regions, rich in alanine, proline and charged amino acids. Proteolytic cleavage sites for trypsin and *Staphylococcus aureus* V8 protease are clustered in these polypeptide segments, defining the limits of the domains and facilitating their purification from each other.

Comparisons of E2 amino acids sequences from a variety of organisms found that they presented the same segmented structures and were highly homologous (Guest *et al.*, 1989; Fig. 1.4). The degree of sequence conservation in the peripheral subunit-binding domains is consistent with the fact that E3 subunits from one source will often complement the overall enzyme complex activity of E1 and E2 subunits from a heterologous source and can also form hybrid high-M_r complexes. The most sustained homologies are in the catalytic domain. The conserved His-602 in *E. coli* PDC E2 is an important active-site residue and is flanked by other conserved residues in a region that closely resembles the

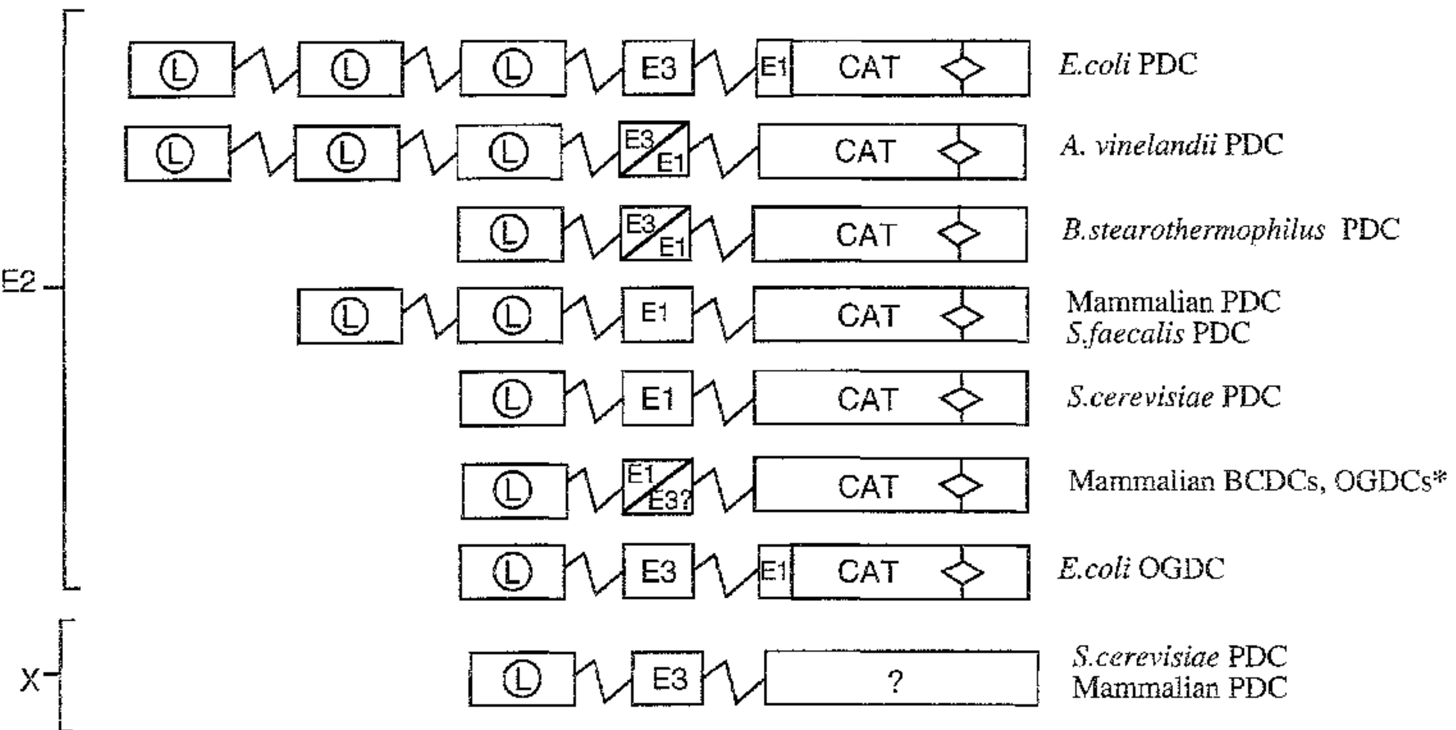


Fig. 1.4: Schematic representation of the domain structure of the E2 and protein X components

L, Lipoate-containing domain; E3 /E1, E3 and/or E1 binding domain; CAT, catalytic core domain;

Zigzags indicate the linkers regions and diamonds indicate the approx. position of the putative active sites.

The domain structure of protein X is shown for comparison.

*Mammalian OGDC has been shown to lack the E3 binding domain. E3 binding sequences are found on E1 and the E1 binding sites are unknown as yet (Nakano *et al.*, 1991).

Reproduced from Sanderson *et al.* (1996)

conserved active-site region of chloramphenicol acetyltransferases (Guest, 1987; Russell & Guest, 1990).

1.3.2.2 Peripheral Subunit-binding Domain

The binding of E1 and E3 to the E2 cores are found to differ between prokaryotic and eukaryotic PDCs, resulting from the presence, in the latter complex, of the tightly bound protein X. As mentioned earlier, *E. coli* E2 (from PDC and OGDC) possesses an E3 binding domain between the outer lipoyl and the inner catalytic domains. A major part of the E1 binding site resides in the E2 catalytic domain. However, in some cases, *Azotobacter vinelandii* (Hanemaaijer *et al.*, 1988), *B. stearothermophilus* E2 (Packman *et al.*, 1988) and mammalian BCDC (Lau *et al.*, 1992), E1 binding is also associated with the E3 binding domain. It was found that E1 and E3 do not bind simultaneously but compete for space on the binding domain of E2 in these organisms (Lessard & Perham, 1995). In contrast, the similar subunit-binding domain of mammalian PDC and OGDC E2 was thought to bind E1 only (Rahmatullah *et al.*, 1989). The information presented above is summarised in Fig. 1.4 and the binding of E3 to mammalian PDC and OGDC will be discussed later.

Unexpectedly, the binding of E3 to the E2 binding domain of *B. stearothermophilus* was found to be at a molar ratio of 1:1 (i.e one E3 dimer binds to only one E2 binding domain) (Hipps *et al.*, 1994). The crystal structures of the E3 complexed with the E2 binding domain revealed that the site of attachment was in close proximity to the two fold axis of the E3 dimer and that both subunits of the dimer provided interacting residues (Mande *et al.*, 1996).

1.3.2.3 Lipoyl Domains

The lipoyl domains each carry a lipoic acid prosthetic group, covalently bound to a lysine side chain via an amide linkage. This post-translational modification is carried out by specific ligases, which recognise not only the domain, but also the specific lysine residue (Brookfield *et al.*, 1991). It is not

known whether these enzymes are protein specific or generally recognise the lipoylation sites of E2. Lipoylation causes no conformational changes, consistent with the lipoyl-lysine being a swinging arm, free in solution (Dardel *et al.*, 1990; Perham, 1996). The number of domains is variable and there appears to be no correlation with the source of the complex. Thus, prokaryotic PDCs from *E. coli* and *A. vinelandii* (octahedral) have three lipoyl domains whereas the prokaryotic *S. faecalis* (icosahedral) and mammalian (eukaryotic) PDC have two (Hodgson *et al.*, 1988) . In contrast, yeast PDC has a single lipoyl domain. BCDC and OGDC from all sources have been found to contain a single lipoyl domain per E2 polypeptide (Fig. 1.4).

It is not clear why *E. coli* E2 contains three lipoyl domains because two can be removed without affecting catalytic activity (Guest *et al.*, 1985); the insertion of extra domains to a genetically engineered complex containing a single domain is increasingly detrimental but not inactivating (Machado *et al.*, 1992). Each domain is known to be able to function independently of the other lipoyl domains (Allen *et al.*, 1989). Although the number of lipoyl domains does not appear to affect catalytic activity, it has been shown that wild-type PDC (containing three lipoyl domains) confers a growth advantage on *E. coli* compared to strains expressing PDC with fewer lipoyl domains (Guest *et al.*, 1996).

The attachment of the lipoyl cofactor to protein domains is an important structural feature for catalysis in all complexes. Its ability to act as a substrate for E1 is enhanced for attached lipoic acid, as judged by the K_{cat}/K_m factor, which is 10 000 fold higher than for the free substrate (Graham *et al.*, 1989). Lipoyl domains proteolytically cleaved from the PDC of bovine heart were efficiently reductively acetylated by the E1 component whereas free lipoic acid or lipoamide was not (Bleile *et al.*, 1981). Furthermore, a lipoylated decapeptide with an amino acid sequence identical with that surrounding each of the three lipoyl-lysine residues in the *E. coli* E2 chain of PDC was ineffective as a substrate indicating that a folded lipoyl domain is essential to the reaction (Graham *et al.*, 1989). Additionally, the

E1 from OGDC only recognises the lipoyl domains from the same complex and *vice versa*. The DKA motif, found on either side of the lipoylated lysine is not required for recognition by the ligase enzymes, but is important for E1 recognition since replacement of the aspartic and alanine residues caused a decrease in the rate of reductive acetylation (Wallis & Perham, 1994)

The lipoamide cofactor, covalently bound to a lysine residue, is very flexible, forming 14Å long 'swinging arms' which facilitate interaction between the three active sites (Reed, 1974). ^{1}H NMR spectroscopy of E2s from *E. coli* PDC and OGDC has shown that extra flexibility is provided by the mobile linker regions, which give the lipoyl domains their extended conformation and provides the means for active-site coupling (Perham *et al.*, 1981). In the intact complex, the lipoyl domains and the subunit-binding domain extend outwards from the inner E2 core, interdigitating between the E1 and the E3 components. The E1 and E3 catalysed reactions take place outside the E2 core, whereas the acyltransferase reaction takes place in the inner E2 core (Perham, 1996). The lipoyl-lysine residues can interact with each other by means of an extensive network of intramolecular coupling reactions that allow the channelling of the acyl groups between different E2 subunits in the enzyme core (Bates *et al.*, 1977). It has been shown for the PDC complex that any one E1 component can be visited by more than one lipoyl group. Thus there is no need to have the three active sites in close proximity to each other and distant sites can be coupled (Perham, 1996).

1.3.2.4 3-Dimensional Structure Determination

The mobility of the lipoyl domains has hindered the growth of well ordered crystals (Fuller *et al.*, 1979); thus, three dimensional structural determination has concentrated on separate domains. NMR spectroscopy has been used to solve the structures of the small E3 binding domain (35 amino acids) on the E2 component of *E. coli* OGDC (Robien *et al.*, 1992) and *B. stearothermophilus* PDC (Kalia *et al.*, 1993). The same technique was used in the determination of the 3-D structures of the lipoyl domain from *B. stearothermophilus* (Dardel *et al.*, 1993) and *E. coli*

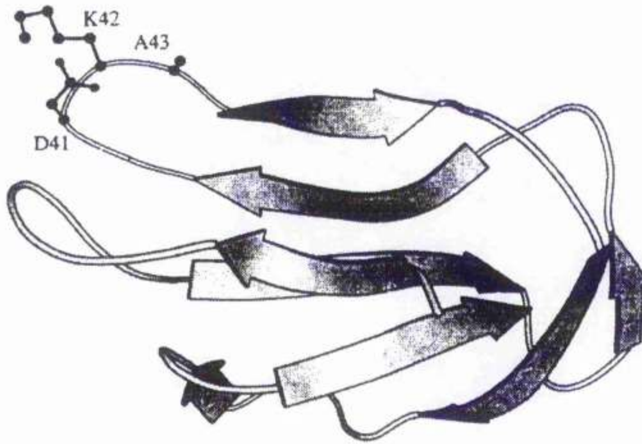


Fig. 1.5: Three-dimensional structure of the *B. stearothermophilus* lipoyl domain

The exposed β-turn containing the lipoyl-lysine residue is shown, indicating the positions of the conserved DKA motif.
Reproduced from Dardel *et al*. (1993).

PDC (Green *et al.*, 1995). This domain is composed of two β-sheets making a flattened barrel forming a core of hydrophobic residues with the lipoyl-lysine perched at the tip of an exposed β-turn (Fig. 1.5), a feature shared with the biotinyl-lysine swinging arm in the biotinyl domain of pyruvate carboxylase (Perham, 1996). The structure of the catalytic domain of E2 from *Azotobacter vinelandii* PDC was determined by X-ray crystallography and structure analysis has demonstrated a similarity to chloramphenicol acetyltransferase (CAT) (Mattevi *et al.*, 1992b; Schulze *et al.*, 1991).

1.3.3 Protein X

Protein X, unique to PDC from mammals and yeast, was originally thought to be a proteolytic fragment of E2. However, De Marcucci and Lindsay (1985) first showed that it was a distinct protein, tightly bound to E2, by the use of specific antisera and peptide mapping. Since X is so tightly bound, it was originally suggested that it integrates into the core system during the assembly of the E2 core (Rahmatullah *et al.*, 1989), but Sanderson *et al.* (1996a) have postulated a post-assembly interaction in mammals which is consistent with genetic evidence from yeast which reveals that the oligomeric E2 can form in the absence of protein X (Lawson *et al.*, 1991a). Protein X has a similar domain structure to E2 (Fig. 1.4), deduced by limited proteolysis of the bovine heart PDC (Neagle & Lindsay, 1991; Rahmatullah *et al.*, 1989). Sequence analysis of the distinct gene encoding the yeast protein X (Behal *et al.*, 1989) indicated that the single lipoyl domain and the adjacent subunit binding domain was homologous with that of yeast E2 but the rest of the structure had no sequence similarity with any other protein. Significantly, the C-terminal domain had no acetyltransferase active site motif. Very recently PDC, purified from *Crithidia fasciculata*, a primitive eukaryote, has been found to contain protein X but unusually it contained multiple lipoyl domains (Diaz & Komuniecki, 1995).

The essential role of protein X was found to be in binding and positioning E3 to the E2 core (Rahmatullah *et al.*, 1989) in a specific manner that is essential for functional PDC. E3 was found to protect the protein X component from degradation by proteases, trypsin and arg C (Gopalkrishnan *et al.*, 1989; Neagle & Lindsay, 1991). Deletion of the subunit binding domain of protein X led to lowered affinity for E3 and subsequent loss in PDC activity (Lawson *et al.*, 1991a). Moreover, studies with a protein X deficient *S. cerevisiae* mutant indicate that, in the absence of protein X, the E2 core assembles properly but does not bind E3 (Lawson *et al.*, 1991a).

The lipoyl groups of protein X and mammalian E2 have the unique property of undergoing diacetylation, whereby both S^6,S^8-thiols incorporate acetyl groups in the presence of pyruvate or acetyl-CoA (Hodgson *et al.*, 1986). It is not known if this is a catalytic function of X, but it has been proposed that diacetylation could permit PDC to act as a reservoir for acetyl groups under conditions of excess acetyl CoA and NADH. Moreover, protein X has been shown to be able to substitute for the lipoyl domains of E2, allowing 10-15% of residual activity of E2 lacking lipoyl domains (Lawson *et al.*, 1991b). It is thought that the lipoyl domains of protein X can service the E2 active site (Sanderson *et al.*, 1996a).

A puzzling aspect is the absence of an equivalent E3 binding protein in the mammalian OGDC and BCDC. Interestingly, the sequence of the human and rat E2 gene of OGDC showed that it lacked the highly conserved E3 and/or E1 binding domain found in all other E2 (OGDC) sequences to date (Nakano *et al.*, 1991; Nakano *et al.*, 1994). Proteolytic analysis and sequence comparisons by Rice *et al.* (1992) revealed that the N-terminus of OGDC E1 from bovine heart was involved in promoting the stable interaction of E3 with the assembled complex. 'Domain shuffling' was apparent in this case as sequences related to the lipoyl domain of protein X and its putative E3 binding domain are situated in the N-terminal region of the E1 enzyme. The E1 component probably binds to an as yet unidentified sequence in the inner catalytic domain in the eukaryotic OGDCs.

1.3.4 Dihydrolipoamide Dehydrogenase (E3)

E3 is the FAD-containing enzyme of the three 2-oxoacid dehydrogenase complexes, the distantly related glycine decarboxylase (GDC) where it is named L-protein and the bacterial acetoin dehydrogenase system involved in acetoin utilisation. It reoxidises the dihydrolipoamide moieties, covalently linked to a specific lysine residue in the E2 components of the 2-oxoacid dehydrogenases and the dihydrolipoamide of the hydrogen-carrier protein of the glycine cleavage system, utilising NAD^+ as the final electron acceptor. Sequence comparison has shown that all E3s from prokaryotic and eukaryotic sources exhibit a high degree of homology. Porcine heart and human E3 have 96% identity with each other and 44% identity with *E. coli* E3 whereas *A. vinelandii* has 40% identity with *E. coli* E3. Since E3 carries out the same reaction in all the complexes it would be reasonable to expect it to complement the activities of all the 2-oxoacid dehydrogenases and GDC within the same species. However, with the explosion in purification and cloning studies of E3 in recent years, from a variety of organisms, the situation has become less clear cut.

Only one E3 gene (*lpd*) has been characterised in *E. coli*, placed adjacent to the *ace*EF operon genes encoding the E1 and E2 genes of PDC (Stephens *et al.*, 1983). The *suc*AB genes encoding E1 and E2 of OGDC are found on a separate codon and are complemented by the *lpd* gene product to form an intact complex. This is facilitated by differential transcription of the *ace*EF-*lpd* genes and the *lpd* gene from two separate promoters, as determined by quantification of mRNA (Quail *et al.*, 1994; Spencer & Guest, 1985). Most of the E3 for PDC is synthesised as read through transcripts from the 6415-base pair *ace*EF-*lpd* genes, but a 1670-base lpd transcript supplies most of the E3 for OGDC. A second distinct E3 enzyme, present in low amounts, was purified by Richarme (1989) from an *lpd*⁻ mutant of *E. coli* , deficient in the PDC and OGDC E3. It was smaller than the already characterised E3, being 46 000 Da in comparison to 50 554 Da. The role of this new E3 is still unknown, but the author suggested an involvement in

binding protein-dependent transport of maltose and galactose. E3 is present in single copy in *A. vinelandii* (Westphal & Kok, 1988) and *P. fluorescens*; however, unlike *E. coli* the genes were found on the OGDC operon (Westphal & Kok, 1990). The E3 gene can also be transcribed independently of the OGDC E1 and E2 genes in these organisms.

A second important function of E3 is as a component of BCDC. This complex is not only found in eukaryotes, but also in bacteria which can utilise branched-chain amino acids for growth e.g. *Pseudomonas putida* and *P. aeruginosa*. They are known to contain a specific E3 for BCDC (LPD-val), found in the BCDC operon, which is different from the PDC, OGDC and GDC E3 (LPD-glc). LPD-val was induced in media containing branched-chain amino acids as a carbon source (McCully *et al.*, 1986; Sokatch *et al.*, 1981) and LPD-glc was induced in glucose synthetic medium (Sokatch & Burns, 1984; Sokatch *et al.*, 1983). A third E3 (LPD-3) discovered in *P. putida* by Burns *et al.* (1989a) was not part of a 2-oxoacid dehydrogenase operon and appeared to be able to replace LPD-glc as the component for PDC and OGDC but its precise role is unknown. It has been suggested that it may be part of the newly discovered acetoin dehydrogenase complex (Oppermann & Steinbuchel, 1994). Primary sequence comparisons of LPD-3 found it to be more closely related to the eukaryotic E3. Freudenberg *et al.* (1989) have recently identified an atypically small E3 subunit, associated with GDC from the anaerobic bacterium *Eubacterium acidaminophilum*. This appears to be the only GDC specific E3 characterised to date.

Recently, cloning of the PDC and the OGDC operons in the strictly respiratory bacterium, *Alcaligenes eutrophus*, revealed a new example of multiple isoforms of E3. In contrast to *E. coli* and *A. vinelandii*, where the E3 gene is found on the PDC operon, or on the OGDC operon, respectively, *A. eutrophus* uniquely possesses genetically distinct E3s for both PDC and OGDC (Hein & Steinbuchel, 1996). Both E3s appear not to be closely related, but it remains to be elucidated whether they are interchangeable. The PDC E3 was found to be a new

type of enzyme, containing a lipoyl domain at its N-terminus which was separated from the rest of the enzyme by a flexible hinge region, as in E2 (Hein & Steinbuchel, 1994). This unique property was shared only with the E3 component (*aco*L gene product) of the acetion dehydrogenase enzyme system of the strictly fermentative, gram-positive bacterium *Clostridium magnum.*

Unusually, uncomplexed E3 has been discovered in the absence of the 2-oxoacid dehydrogenase complexes, both in prokaryotic species (halophilic and thermophilic archaebacteria, Danson *et al.*, 1984) and in the bloodstream form of the eukaryotic African parasite, *Trypanosoma brucei* (Danson *et al.*, 1987). In *T. brucei* it is closely associated with the plasma membrane but it remains to be established whether it performs a similar role in transport, suggested for the second E3 enzyme of *E. coli.*

The situation in eukaryotes is similar to that found in *E. coli*. A mutant in the yeast *lpd1* gene abolishes all PDC and OGDC activity (Dickinson *et al.*, 1986) and in mammals, immunological data (Matuda & Saheki, 1985) and reconstitutional studies (Sakurai *et al.*, 1970) all indicate that PDC, OGDC and BCDC use the same E3 subunit. The best evidence for a single copy of E3 comes from studies of lactic acidosis in man where the genetic lesion affects the production of lipoamide dehydrogenase. PDC, OGDC and BCDC are all affected since pyruvate, 2-oxoglutarate and branched-chain amino acids accumulate in the serum of these patients (Robinson *et al.*, 1977). It has been assumed that the E3 subunit also promotes GDC activity in mammals, although an intact complex has never been isolated intact from this source. Cloning of the E3 gene for glycine decarboxylase (termed L-protein) from pea leaf mitochondria (Bourguignon *et al.*, 1992; Turner *et al.*, 1992), and subsequent analysis of PDC E3 from the same organelle has confirmed the presence of a single E3 complementing all the mitochondrial 2-oxoacid dehydrogenase complexes (Bourguignon *et al.*, 1996). Efforts to detect multiple isoforms in mammals have yielded what appear to be conformational isomers of the same protein (Williams, 1976). There is also some isolated evidence

that rat liver mitochondria contain two immunologically distinct E3, one being specific to GDC (Carothers *et al.*, 1987).

1.3.4.1 Domain Structure

E3 is related, structurally and mechanistically, to the pyridine nucleotide-disulphide oxidoreductases - a family of flavoenzymes, which are distinguished by containing an redox active disulphide. Other members include glutathione reductase (Krauth-Siegel *et al.*, 1982; Schulz *et al.*, 1978), mercuric reductase (Schiering *et al.*, 1991), thioredoxin reductase (Kuriyan *et al.*, 1991a) and trypanothione reductase (Kuriyan *et al.*, 1991b). They are all active as homodimers (approx. 2 x 50 kDa), a FAD cofactor is non-covalently bound to each of their subunits, they always catalyse an electron transfer between NAD(P)H and a disulphide/dithiol and they have significant sequence identity (>20%).

A complete amino acid sequence was first obtained for glutathione reductase (GR) and the complete tertiary structure by X-ray diffraction analysis (Thieme *et al.*, 1981). The structural and mechanistic similarities between glutathione reductase and E3 have been recognised for many years (Williams, 1976) and since there is considerable homology at the primary amino acid sequence between GR and human E3 (33%), GR has been used as a model to predict the secondary structure of all E3s cloned to date. The first example of this was when the sequence of *E. coli* E3 was fitted into the 3 dimensional structure of GR (Rice *et al.*, 1984). The well defined domain structure of GR, namely the FAD-binding domain (residues 1-157), NADPH binding domain (residues 158-293), a central domain (residues 294-364) and a dimer interface domain (residues 365-478), is found in all E3s. There is striking homology of the sequences around the nucleotide binding sites, the active site cysteines found in the FAD domain and in the contact region between the dimers. Distinctive chain folds around the nucleotide binding sites are evident. Fig. 1.6 shows a comparison of the overall secondary structure of GR and human E3.

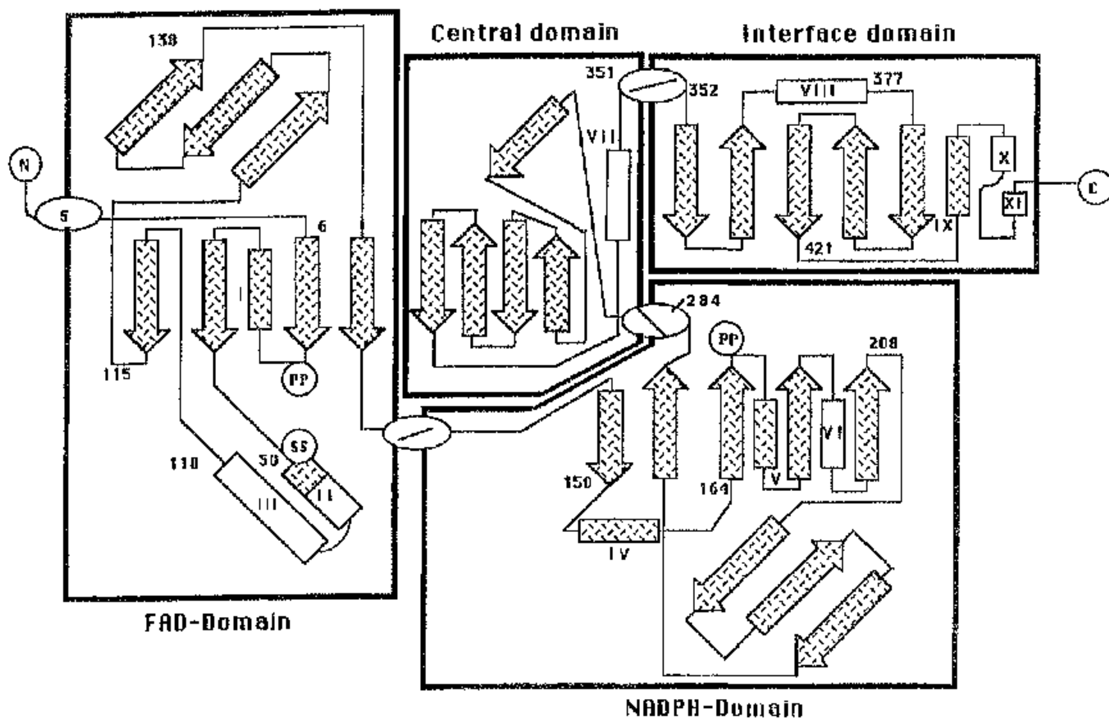


Fig. 1.6: Comparison of the chain fold of human glutathione reductase and human E3

Strands of beta-sheets are shown as arrows and alpha-helices as rectangles. The open rectangles indicate helices with little sequence conservation between E3 and GR. The four domains are separately identified by enclosure in boxes. Numbers indicate amino acid residues of human E3 and roman numerals identify alpha helices. The residues in ovals denote domain boundaries.

PP, pyrophosphate moieties of FAD and NAD; **S-S**, the redox-active dithiol (Cys-Cys of E3); **N** and **C**, amino and carboxy terminus, respectively. Reproduced from Thekkumkara *et al.* (1989).

The crystal structures of E3 from *A. vinelandii*, *P. putida* and *P. fluoresens* (Mattevi *et al.*, 1992a; Mattevi *et al.*, 1993; Mattevi *et al.*, 1991) have been elucidated in recent years and have confirmed many postulates of the catalytic mechanisms and active site residues. The reason for the arrangement of E3 as dimers became obvious when the crystal structures were analysed. In each enzyme the active sites are located at the dimer interface and essential amino acids are provided by each subunit. The active site is composed of three elements; the flavin, disulphide and a histidine which acts as the base. Thus, the monomer can not be catalytically active.

1.4 Regulation

PDC is the link in the metabolism of glucose between glycolysis and the TCA cycle and since pyruvate can also be used for gluconeogenesis, it is subject to short-term (end product inhibition, covalent modification, response to hormonal changes) and long-term regulation by nutritional changes. In addition, the tight control of glucose metabolism and catabolism of branched-chain amino acids assures continuous availability for energy production, biosynthetic intermediates and protein synthesis and disposal of excess substrates for prevention of the toxic effects of high levels of these substrates (Patel & Harris, 1995).

Eukaryotic PDC and BCDC are the only mitochondrial complexes found to be regulated by reversible phosphorylation (James *et al.*, 1995; Linn *et al.*, 1969). The E1α from PDC and BCDC becomes phosphorylated at three and two specific serine residues, respectively, leading to a dramatic decrease in the V_{max} and therefore, inactivation of the complex. Inactivation correlates with phosphorylation at only the first site (Cook *et al.*, 1983; Davis *et al.*, 1977; Yeaman *et al.*, 1978; Zhao *et al.*, 1994) and it is thought that phosphorylation at the additional sites may play a regulatory role in inhibiting dephosphorylation of the inactivating site (Sugden *et al.*, 1978). Zhao *et al.* (1994) confirmed, by site directed mutagenesis of BCDC, that site 1 was indeed responsible for kinase-mediated inactivation of E1

but phosphorylation at site 2 played no role in lowering the binding affinity of the E1 for E2, previously suggested by Cook *et al.* (1985). Analysis of the amino acid sequence of mammalian PDC and BCDC, surrounding the principal phosphorylation site (site 1) reveals considerable homology. There is no evidence for the phosphorylation of OGDC, although ATP alters kinetic parameters of the complex by causing an increase in the K_m for 2-oxoglutarate (Hunter & Lindsay, 1986). The major regulator of OGDC activity is Ca^{2+}.

Control of PDC activity is manifested in the control of the phosphorylation/dephosphorylation cycle itself. The Mg^{2+} requiring PDC kinase is inhibited by the 2-oxoacid substrates and stimulated by the reaction products (i.e. NADH, ATP and the acyl CoA), whereas BCDC kinase is insensitive to control by the reaction products (Paxton & Harris, 1984). PDC kinase is also inhibited by ThDP as a result of binding at the catalytic site of E1 and thereby altering the conformation about phosphorylation site 1 so that the serine is less accessible to the kinase.

Specific kinases of PDC and BCDC are tightly bound to their E2 components and co-purify with them. The K1 and K2 isoforms of E1 kinase of PDC as well as the kinase component of the BCDC bind to the inner lipoyl domain region of E2. Interestingly, in preparations of bovine PDC, as few as 1 molecule of kinase are bound to the E2 core but a large number of E1 (20-30 E1 tetramers) can be rapidly phosphorylated. This has been explained by evidence suggesting efficient transfer of the kinase between the lipoyl domains of different E2 components in the core (Ono *et al.*, 1993). While Stepp *et al.* (1983) suggested that the PDC kinase was a heterodimer, there is very little evidence for this conclusion. The K1 subunit can bind to E2 and function as a kinase alone. However, patterns in native gel electrophoresis are consistent with the following dimer forms for two kinase subunits - $[K1]_2$, [K1K2] and $[K2]_2$ (see Roche *et al.* (1996) for review). Such a dimer state would help explain movement at the surface of the E2 core by a

"hand over hand" mechanism involving continuous partial dissociation by one subunit followed by interchange to another lipoyl domain (Liu *et al.*, 1995).

The BCDC and PDC kinase genes have been cloned from rat liver and heart cDNA libraries (Popov *et al.*, 1992; Popov *et al.*, 1993; Popov *et al.*, 1994) and surprisingly they did not reveal significant sequence similarity with any known eukaryotic serine or threonine protein kinase. However, considerable sequence similarity exists between the kinase for BCDC and PDC and the prokaryotic family of histidine protein kinases. This work established a new family of mitochondrial protein kinases in eukaryotes.

Much less is known about the complex-specific phosphatases. They are known to be less tightly bound to the complexes and can be purified as separate enzymes. PDC phosphatase consists of two subunits of M_r 97 000 and 50 000, with catalytic activity residing in the smaller subunit. The significance of the FAD molecule in the larger subunit is unclear (Teague *et al.*, 1982). The enzyme requires Mg^{2+} for activity and is stimulated by Ca^{2+} ions which reduce the K_m for Mg^{2+} and its E1 substrate (Pettit *et al.*, 1972; Reed & Yeaman, 1987). It has been proposed that the phosphatase has one intrinsic Ca^{2+} binding site and a second one is generated when the phosphatase is complexed with E2. It is not known whether the second site is present on the phosphatase or on E2 but it may facilitate Mg^{2+} dependent dephosphorylation by positioning the phosphatase and phosphorylated E1 favourably on E2. The other effectors of phosphatase activity are NADH, which inhibits phosphatase activity and the polyamine, spermine, which stimulates by lowering the K_m for Mg^{2+} (Roche *et al.*, 1989). Effectors of PDC kinase and phosphatase are shown in Fig. 1.7.

Acute control of mammalian complexes by hormones has been studied extensively. The effects of adrenaline in the heart and vasopressin in the liver are mediated by increased Ca^{2+} levels in the cytosol and hence in the mitochondrial matrix. Direct binding of Ca^{2+} to the OGDC causes allosteric activation by reducing the K_m for its substrate, 2-oxoglutarate (McCormack & Denton, 1979).

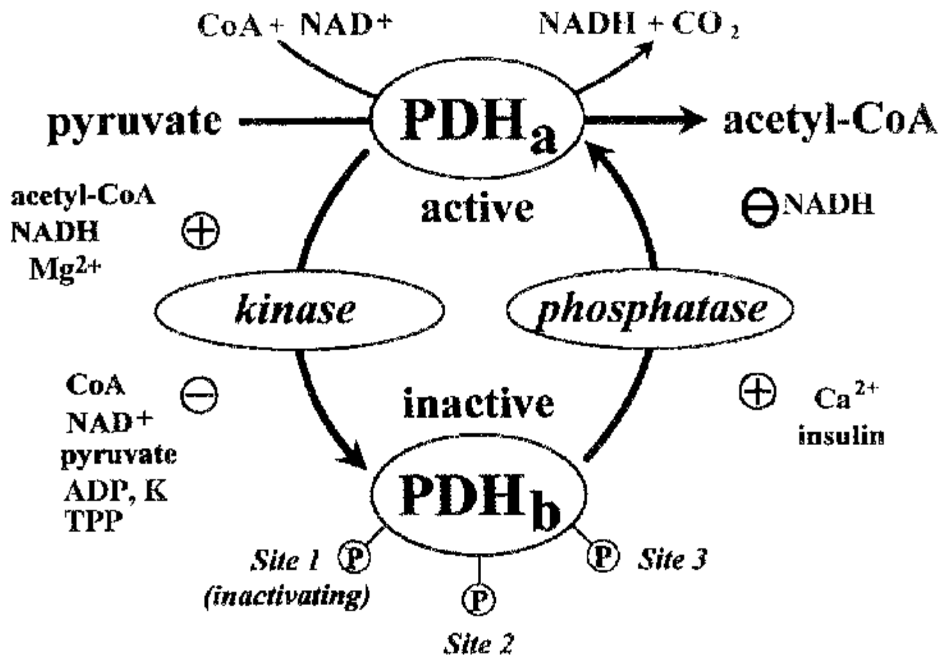


Fig. 1.7: Schematic representation of the covalent modification of pyruvate dehydrogenase and its control by effectors

Reproduced from Sugden and Holness (1994).

This is antagonised by ATP which increases the K_m for the substrate. PDC is also activated by Ca^{2+} *in vitro*, but by a different mechanism. In this case it appears that Ca^{2+} stimulates the phosphatase by the mechanisms described above (Denton & McCormack, 1985). Insulin is known to promote the dephosphorylation of PDC, enhancing its activity in adipose tissue. The second messenger activating the phosphatase has not been verified since Marshall *et al.* (1984) showed that the effects of insulin were insensitive to Ca^{2+} concentrations. BCDC is insensitive to Ca^{2+} levels and there is a lack of evidence for hormonal control of this complex. A major mechanism for acute control of BCDC is via inhibition of BCDC kinase by the 2-oxoacid substrates (Jones & Yeaman, 1986) and high circulating levels of the 2-oxoacids in animals fed a high protein diet are likely to inhibit the kinase, maintaining the complex in its active form.

1.5 Disease States Associated with the 2-Oxoacid Dehydrogenase Complexes

As the 2-oxoacid dehydrogenase complexes play a central role in metabolism, genetic defects in the PDC and BCDC result in metabolic aberrations causing a build up of toxic levels of their substrates. The availability of the gene sequences for most of the subunits and knowledge about their gene structure, organisation and regulation have facilitated the identification of inherited disorders of PDC and BCDC metabolism at the genetic level in recent years.

Human PDC deficiency is a potentially severe inborn error of oxidative metabolism. Owing to the multisubunit structure of PDC, a variety of genetic defects are possible. Heterogeneity in PDC deficiency has been shown to occur at both protein and mRNA levels. In almost all cases, the basic defect appears to be in the E1 component, and in particular E1α (Ho *et al.*, 1989). However, there are cases resulting from defects of E2, E3 or protein X (Marsac *et al.*, 1993; Robinson *et al.*, 1990). Characteristic features are metabolic acidosis and neurological disorders, distinguished by significant structural abnormalities of the central

nervous system. There is a wide range of severity of symptoms, beginning with severe lethal lactic acidosis developing a few hours after birth to a mild form with no significant acidosis. All forms of PDC deficiency present with neurological disorders (Brown *et al.*, 1989)), largely resulting from the dependence of the brain on glucose oxidation for energy production..

Inherited disorders of BCDC, described over 30 years ago, lead to maple syrup urine disease (MSUD). This disease is characterised by large increases in the concentration of branched-chain amino acids and their respective 2-oxoacids in the blood, tissues and urine (Patel & Harris, 1995). Again, there are varying degrees of severity of this disease, from physical and mental impairment and early death to a milder form later in life. The genetic lesions are more heterogeneous than with PDC errors as defects involve any of the subunits of BCDC.

Primary biliary cirrhosis (PBC) is a chronic autoimmune disease of the liver in which the epithelial cells of the small bile duct are destroyed by inflammation, leading to liver cell damage and cirrhosis (Kaplan, 1987; Yeaman *et al.*, 1989). In Europe it is the commonest indication for liver transplantation. PBC is characterised by the presence of circulating autoantibodies targeted against mitochondrial antigens, mainly located on the inner mitochondrial membrane. PBC-specific mitochondrial autoantigens have been characterised (Berg & Klein, 1988), with antibodies to one (M2a antigen) predominating in 85-95% of patients with PBC. The predicted protein sequence from the M2a cDNA (Coppel *et al.*, 1988) and immunoblotting the sera from PBC patients with purified E2 protein identified the autoantigen as E2 of mammalian PDC. The majority of the serum samples (95%) were found to cross-react positively with PDC E2 (Yeaman *et al.*, 1988). Subsequently it was found that component X from PDC, E2 from OGDC and BCDC (Fussey *et al.*, 1988) and E1α and E1β from PDC (Fregeau *et al.*, 1989) all cross reacted with antisera from PBC patients. However, the E2 component is the main autoantigen and the epitope has been mapped to the inner lipoyl domain of the molecule (Surh *et al.*, 1990). The mechanism of how

autoantibodies in PBC sera are generated against intracellular mitochondrial antigens and why target tissues are limited to the bile ducts of the liver are largely unknown. Furthermore, the role of these autoantibodies in the pathogenesis of PBC and the fine specificity of the antibodies specific to PDC E2 is still unclear (Gershwin & Mackay, 1991).

1.6 Plant 2-Oxoacid Dehydrogenase Complexes

It is apparent from the evidence presented above, that a vast amount of knowledge has been gained on the 2-oxoacid dehydrogenases from a variety of prokaryotic and eukaryotic sources. In stark contrast, little information is available on the equivalent plant complexes, owing to the low abundance of mitochondria per fresh weight of tissue. A frequently cited example is the purification of a mere 1mg of mitochondrial PDC from 500 kg of broccoli florets (Rubin & Randall, 1977). Investigations of the plant complexes not only enhances the overall knowledge of the 2-oxoacid dehydrogenases, but also provides a comparison of the metabolic processes between animals and plants. Considerable metabolic similarities exist between plants and animals; however, the different anatomy, physiology and organismal requirements have led to an interesting diversity between these two classes of eukaryotes. PDC in plant cells illustrates both the similarities in metabolic processes and the differences dictated by the need to respond to diverse external stimuli (Randall *et al.*, 1989). Uniquely plants contain two distinct and spacially separate PDCs, one located in the mitochondrial matrix and the other in the plastid stroma. Each PDC isoform has its own structural, catalyic and regulatory properties (Miernyk *et al.*, 1985).

A main characteristic of plant cells is the greater complexity of subcellular organisation in comparison to animal cells. Fig. 1.9 illustrates the interaction of metabolic pathways found in chloroplasts, peroxisomes and mitochondria of plant cells. All *de novo* fatty acid and the latter steps in isoprenoid biosynthesis by plant cells occurs within the plastids (Ohlrogge *et al.*, 1979), and as such a separate

glycolytic pathway exists in plastids for the production of acetyl CoA for this process (Dennis & Miernyk, 1982). Examples of plastids are chloroplasts in light-grown leaves, chromoplasts in fruits and floral tissues and the leucoplasts and amyloplasts in seeds and other non-green tissues. A second, classical glycolytic pathway is found in the cytoplasm of plant cells. Another feature, unique to plant cells, is the existence of two separate systems for electron transport-coupled to synthesis of ATP. One is oxidative phosphorylation in mitochondria, operative in the dark and photophosphorylation in chloroplasts, operative in the light. Mitochondrial PDC occupies an ideal position for regulation of carbon flow into the TCA cycle during photosynthesis (Randall *et al.*, 1989).

The following discussion will centre around PDC, since it is the only complex from plants to have been purified and characterised. Of the components, E3 has received a great deal of attention in our laboratory, with its purification to homogeneity from a variety of plants and the cloning and sequencing of the mitochondrial gene from potato leaves presented in this thesis.

1.7 Plant Mitochondrial PDC

Mitochondrial PDC serves as an entry point for carbon into the TCA cycle. Based upon sedimentation analysis, mitochondrial PDCs are large aggregates, similar to their mammalian counterparts (Rubin & Randall, 1977). Kinetic characterisation of the purified PDC from broccoli florets found that cofactor requirements (NAD^+, ThDP, CoA, Mg^{2+} and Ca^{2+}) were similar to that of *E. coli* and yeast. Initial analysis of pea mitochondrial PDC by western blotting with antibodies raised against PDC from broccoli florets (Rubin & Randall, 1977) revealed subunits of M_r 97 700, 67 400, 58 100, 43 300 and 37 000 (Camp & Randall, 1985). The 58 100 Da subunit seemed to be E3 since it also cross reacted with anti-porcine E3 and the 43 300 and 37 000 Da species corresponded to E1α and β, respectively. Further analysis of pea mitochondrial PDC was carried out by Taylor *et al.* (1992) using antibodies to specific components of bovine heart PDC.

There was no appreciable cross reaction with mammalian anti-E1; however, it did cross-react with yeast anti-E1, highlighting a band of 41 000 Da, which is thought to be E1α. This corresponds to the 43 300 Da polypeptide identified by Camp *et al.* (1988) and Luethy *et al.* (1995) who used a monoclonal antibody to maize E1α. The anti-bovine E2 cross reacted with a 50 000 Da polypeptide which is lower than mammalian E2 (70 000 Da) but similar in size to the yeast protein and is consistent with the presence of a single lipoyl domain (Taylor *et al.*, 1992). However, recent immunological blotting of E2 expressed from a clone isolated from *Arabidopsis thaliania* found that its M_r was close to 80 kDa and suggested that the 50 kDa band may represent protein X (Guan *et al.*, 1995). E3 was thought to be a 67 000 Da species but since this work was carried out, purified pea mitochondrial E3 has been shown to be 58 000 Da (Conner *et al.*, 1996). This is in close agreement with previously reported values (Bourguignon *et al.*, 1992; Turner *et al.*, 1992).

In recent years, the PDC E1α gene (Luethy *et al.*, 1995), the E1β gene (Luethy *et al.*, 1994) and the E2 gene (Guan *et al.*, 1995) have been cloned from *Arabidopsis thaliana* and have revealed the high sequence homologies between plants and other eukaryotes and prokaryotes. A common feature of all these genes is the presence of a mitochondrial targeting pre-sequence. E1α gene has an open reading frame of 1176 bp, encoding a protein of 389 amino acids. Overall homology with other eukaryotic E1α sequences falls within the range of 47-51% identity and 63-69% sequence similarity. The sequence surrounding the ThDP binding site and phosphorylation sites 1 and 2 are highly conserved; however, a serine is replaced by a threonine at site 2. The E1β gene has an open reading frame of 1230 bp and the deduced amino acid sequence (363) has 65% sequence identity with *S. cerevisiae* E1β. *Arabidopsis* E2 contains regions which are highly homologous to the human and yeast E2 and it appears that the E2 contains two lipoyl domains, similar to human E2 (Perham, 1991). The E3 and/or E1 binding site shows lower sequence identity with human E2.

The only E3 to be cloned to date is that of pea leaf L-protein from glycine decarboxylase (GDC) by Turner *et al.* (1992) and Bourguignon *et al.* (1992). The gene has a 1503 bp open reading frame, encoding 501 amino acids, with a M_r of 49 721. The amino acid sequence was found to be highly conserved between human, porcine and yeast E3, particularly within the functional domains (FAD- and NAD-binding) and the authors predict the pea E3 shares a similar secondary and tertiary structure with other E3s. This protein was subsequently found to be the same as the E3 for PDC in pea leaf mitochondria using a number of biochemical criteria, including precise subunit M_r comparisons using electrospray mass spectrometry (Bourguignon *et al.*, 1996).

1.7.1 Regulation of Mitochondrial PDC

Since PDC occupies an important branch point of various metabolic routes, it is a logical site for regulation. The many layers of regulation of plant mitochondrial PDC are very similar to mammalian mechanisms. If the energy demands of the cell could be met by some means other than the TCA cycle, the PDC is the site for regulatory mechanisms reducing input into the cycle and conserving reduced carbon. Such a condition may exist during photosynthesis and photorespiration when, for example, glycine is converted to serine with concomitant production of NADH in C_3 plants (Randall *et al.*, 1990b).

Fine control is mediated through competitive inhibition by the reaction products, NADH and acetyl CoA, with inhibition constants in the micromolar range (Rubin *et al.*, 1978). The complex from plants was observed to be more sensitive to the *in vitro* NADH:NAD^+ ratio than to the acetyl CoA:CoA ratio (Rubin *et al.*, 1978). The possible regulation by ThDP has been suggested because of its very rapid dissociation from the broccoli complex.

The suggestion that plant PDCs could be regulated by covalent modification, in the same manner as the mammalian PDCs, was put forward with the demonstration that PDC could be inactivated in an ATP-dependent manner, and subsequently reactivated through a Mg^{2+}-stimulated process (Randall & Rubin,

1977; Rao & Randall, 1980). Phosphorylation was verified by ^{32}P incorporation from [^{32}P]ATP of a single subunit of M_r 43 100 (Randall *et al.*, 1981). This subunit was shown to be E1α by precipitation of the bound radioactivity using PDC-specific antibodies (Miernyk *et al.*, 1985).

The site for phosphorylation appears to be significantly different from that of bovine, porcine or yeast PDC since antibodies to a synthetic peptide containing phosphorylation sites 1 and 2 of bovine E1α did not cross react with plant mitochondrial or plastid PDC. In addition, the synthetic peptide does not serve as a substrate for the plant phosphorylation system unless the serine is substituted for a threonine. These data conflict with *in situ* evidence, which suggests that serine is the phosphorylation site (Randall *et al.*, 1989). Covalent modification has been found to be a mechanism used by a host of mitochondrial PDCs, for example spinach leaves (Rao & Randall, 1980), pea (Randall *et al.*, 1981) and the nongreen endosperm of developing castor oil seeds (Rapp & Randall, 1980).

The activity of mitochondrial PDC is under steady state control (Budde & Randall, 1988b) in which the activity of PDC is dependent upon the ratio of the kinase to phosphatase activity. The steady state favours inactivation of the complex since *in vitro* rates of the phosphatase are usually six times lower than the kinase. Therefore, it seems that in the absence of effectors the *in vivo* complex would be phosphorylated and inactive all the time (Miernyk & Randall, 1987). Adjustments to the steady state are made by inhibiting inactivation and stimulating reactivation (Budde *et al.*, 1988a). Fig. 1.8 provides a model summarising the information about the control of reversible phosphorylation of pea leaf mitochondrial PDC.

Plant PDC kinase has been difficult to study due to its low abundance and the fact that it dissociates from the complex during purification from pea seedlings (Randall *et al.*, 1989). Thus, all *in vitro* studies in regulation have used partially purified PDC or mitochondrial extracts. *In situ* experiments, measuring the various respiratory states of intact mitochondria from green pea seedlings with the oxygen electrode have verified most of the *in vitro* results and helped to establish that

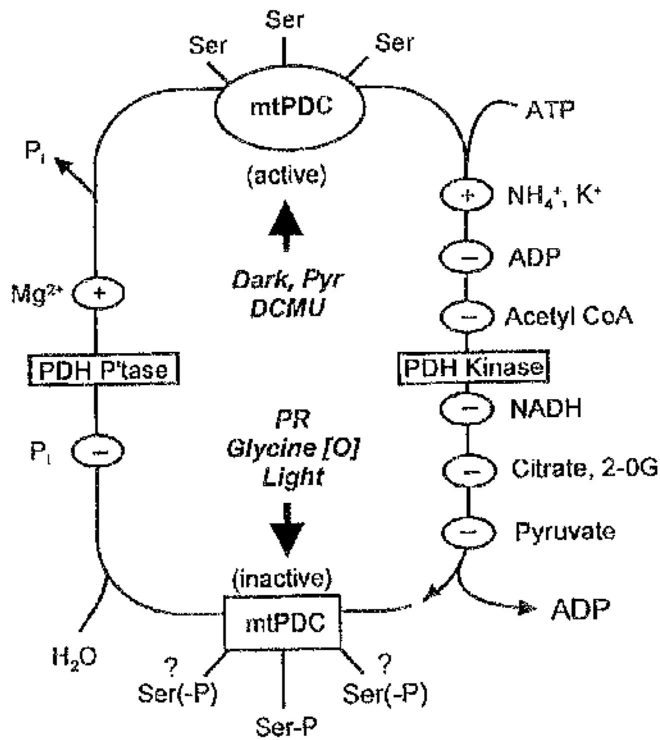


Fig. 1.8: Model illustrating effectors of the reversible phosphorylation of pea leaf mitochondrial PDC

Reproduced from Leuthy *et al.* (1996).

reversible phosphorylation of PDC does take place under physiological conditions. There were some differences in the degree of regulation and order of importance of various effectors *in situ*, compared with the *in vitro* situation as described in the following paragraphs.

In vitro activity is maximal at pH 7.5, with Mg-ATP the preferred phosphoryl donor. The K_m for ATP is low (2.5 μM) in comparison with the K_m for Mg^{2+} (3.8 mM), so PDC will always be in the inactive (phosphorylated) state if the kinase is not regulated. In contrast to the mammalian PDC and *in vitro* observations of the plant complex, changes in the ATP:ADP ratio had no effect on the activation state of the complex (Budde & Randall, 1987) *in situ*, whereas ATP added exogenously caused inactivation of PDC. Inactivation lasted until ATP was removed by ATP-scavenging with excess glucose and required 20-40 fold less Mg^{2+} (0.5 mM) than was needed for *in vitro* inactivation. (Budde *et al.*, 1988a). This concentration of Mg^{2+} is found to be physiologically relevant in inhibiting the kinase.

The phosphorylation state of PDC is mainly under the control of pyruvate concentration. Inhibition of the kinase by pyruvate is uncompetitive with respect to ATP (Schuller & Randall, 1989). However, the stimulatory effect of K^+ and NH_4^+ ions, reflecting oxidation of glycine, (Schuller & Randall, 1989) seems to override the inhibitory effect of pyruvate and ThDP *in vitro*. The mechanism of stimulation by these monovalent ions is to increase the affinity of the kinase for ATP (Schuller *et al.*, 1993). Stimulation by NH_4^+, a product of photorespiratory carbon metabolism, provides a mechanism for regulating the citric acid cycle during photosynthesis and will be discussed later under control of PDC by photorespiration in the light.

Confusing *in vitro* observations were made about the effect of acetyl-CoA and NADH on kinase activity. As these products inhibit PDC activity, they would be expected to stimulate kinase activity; however an inhibition of kinase activity was observed. Similarly, citrate inhibition of the kinase (causing PDC to remain active)

is unexpected since increased citrate levels would indicate a decreased demand for acetyl CoA. It is possible that there are alternative uses for citrate and acetyl CoA outside the mitochondrion, for example, in fatty acid and isoprenoid synthesis which takes place in the chloroplast (Randall *et al.*, 1989).

The phosphatase requires a divalent cation (Mg^{2+}) for activity, with activation by Mg^{2+}>Mn^{2+}>Co^{2+} having K_m values of 3.8, 1.7, and 1.4 mM, respectively (Miernyk & Randall, 1987). Interestingly, while the mammalian PDC phosphatase is stimulated by Ca^{2+}, the pea enzyme was found to be inhibited *in vitro* by micromolar levels of this ion. This result contradicted the *in situ* result that Ca^{2+} had no effect on PDC activity (Budde & Randall, 1988b).

1.7.2 Effect of Light and Photosynthesis on PDC Activity

Protein phosphorylation tends to occur in response to extracellular signals, therefore regulation was investigated *in planta* by Budde and Randall (1990) to determine the effect of light on PDC activity. Upon illumination of pea leaves, there was a decrease in steady-state mitochondrial PDC activity to 10-20% of the levels exhibited by dark adapted leaves, which could be reversed within minutes of returning plants to the dark. In experiments where photosynthesis and/or photorespiration was chemically inhibited, the light-dependent inactivation of PDC was also prevented, indicating that both these processes are essential for regulation (Budde & Randall, 1990; Gemel & Randall, 1992). It can be concluded from these experiments that when mitochondria are able to oxidise substrates other than pyruvate, PDC is quickly inactivated by phosphorylation, conserving carbohydrate when the cell is photosynthetically active, (Gemel & Randall, 1992; Randall *et al.*, 1990b). An example is the photorespiratory metabolite, glycine, which is the preferred substrate for ATP production in mitochondria during illumination and is the only photorespiratory intermediate that is metabolised in the mitochondria during photosynthesis (Fig. 1.9). PDC inactivation when mitochondria are oxidizing glycine reflects the decreased pyruvate levels characteristic of photosynthetically active tissues, the increased NADH levels due to glycine

oxidation and increased ATP levels due to oxidative phosphorylation. This inactivation would be enhanced by NH_4^+ ions (by-product of glycine oxidation) on the kinase. Mitochondrial ATP production is then supported by glycine oxidation and conserves carbohydrates that provide pyruvate.

1.8 Plant Plastid PDC

Plastid PDC provides acetyl CoA for lipid, isoprenoid and amino acid synthesis and is the only known source of NADH in chloroplasts. This role for plastid PDC has been well established in developing castor oil seeds, where large quantities of fatty acids are synthesised for storage (Dennis & Mierynk, 1982) and in chloroplasts of green tissues for *de novo* synthesis of fatty acids (Williams & Randall, 1979). To date, PDC has been found in plastids isolated from every plant source investigated with the exception of the leucoplasts from germinating castor oil seeds (Rapp & Randall, 1980). Plastid PDC catalyses the same overall reaction and has the same kinetic mechanism as the mitochondrial PDC; however, it is structurally different and has higher pH and Mg^{2+} requirements (Camp & Randall, 1985). The higher Mg^{2+} (10 mM) and alkaline pH (pH 8) is in accordance with indirect regulation by light/dark transitions.

The complex appears to be less tightly associated than the mitochondrial form and the only subunit to be identified was the 58 000 Da E3 component (Camp & Randall, 1985). Taylor *et al.* (1992) carried out an immunological investigation into the plastid PDC from pea leaves, using antibodies to specific components of bovine heart and yeast PDC. The E1 component seemed to be an immunologically distinct protein since no cross reaction was observed with the E1 antisera. In contrast, the pea mitochondrial protein exhibited a band of 41 000 Da. The E2 polypeptide appeared to be a 50 000 Da protein; however, following the work of Guan *et al.* (1995) with the pea mitochondrial complexes it may possibly represent protein X. The molecular weight is consistent with one lipoyl domain. Mammalian E3 antisera cross reacted with a protein of 52 000 Da, smaller than the 58 000 Da

identified by Camp and Randall (1985). Organelle specific isoforms of E3 were suggested by Taylor *et al.* (1992) and closer inspection confirmed that this was indeed the case. Resolution of total pea leaf extracts on Mono Q ion exchange chromatography revealed two peaks of E3 activity. The first was found to correspond to mitochondrial E3 with a M_r of 58 000 and the second smaller peak was a related plastidic E3 with a M_r of 52 000 (Conner *et al.*, 1996). Both proteins were found to be structurally distinct.

1.8.1 Regulation Of Plastid PDC

Plastid PDC resembles mitochondrial PDC in that it is subject to competitive, end product inhibition by NADH and acetyl CoA. However, a major difference between their regulatory mechanism is the lack of reversible phosphorylation. In this sense they resemble bacterial complexes (Randall *et al.*, 1989). Phosphorylation would be counter-productive since fatty acid biosynthesis, which is light driven, demands acetyl CoA and NADH. Moreover, ATP is being synthesised at its highest rate by photophosphorylation. The only metabolites specifically affecting PDC activity were 50 μM oleic acid (57% inhibition) and palmitic acid (36% stimulation) (Camp *et al.*, 1988).

Evidence suggests that the main form of regulation of plastid PDCs is light-dark transition and the changes it brings about in the physical environment. The stroma of non-illuminated plastids is at neutral pH and there is a low concentration of divalent ions. Upon illumination the pH rises to 8 and the concentrations of free Mg^{2+} and Ca^{2+} ions increase substantially (Hind *et al.*, 1974). Plastidic PDCs are well suited to the conditions found within the stroma during illumination. Their K_m for pyruvate is lowest at pH 8 and greater concentrations of divalent ions (1 mM) are required than mitochondrial PDCs (0.4 mM) for maximal activity (Camp & Randall, 1985). The complex is also relatively insensitive to inhibition by ATP and NADPH, which would increase in concentration during illumination (Camp *et al.*, 1988).

The source of acetyl CoA for fatty acid biosynthesis has been under debate for a number of years. Two pathways exist for the generation of acetyl CoA: 1) action of plastid PDC via the decarboxylation of pyruvate; 2) action of acetyl CoA synthetase which converts acetate to acetyl CoA. The acetate is proposed to originate in the mitochondria, having been converted from acetyl CoA by acetyl CoA hydrolase (Zeiher & Randall, 1990). Kang and Rawsthorne (1994) found that pyruvate was five times better than acetate as a substrate for fatty acid synthesis in plastids from developing embryos of oil seed rape. However, acetate has been shown to be a better substrate in other cases (Springer & Heise, 1989). On the other hand, malate has been reported to be better than both pyruvate and acetate, but malate conversion to fatty acids involves PDC (Smith *et al.*, 1992). The conclusion drawn from these observations was that the species to species variations in substrate incorporation into fatty acids may be dependent on the relative levels of the various enzymes present in the plastid. For example, castor oil endosperm have a higher level of NADP-malic enzyme relative to other plastids allowing for higher rates of malate incorporation into fatty acids (Smith *et al.*, 1992). The levels of plastid PDC relative to mitochondrial PDC are also species dependent. In pea leaves, 75-90% of the total cellular PDC is mitochondrial (Camp & Randall, 1985), whereas in barley leaves 60% of the PDC is plastidic (Kromer *et al.*, 1994). In general plastid PDC is the major source for acetyl CoA, but additional acetyl CoA could be supplied by mitochondrial PDC.

1.9 Plant OGDC

OGDC from plants has not been studied as extensively as PDC since its activity is masked by high levels of NADH oxidase activity and has proven to be impossible to isolate as an intact complex. The E3 component readily dissociates from the complex; hence Poulsen and Wedding (1970) reported the isolation of the E1/E2 subcomplex from cauliflower. They subsequently obtained intact OGDC activity by reconstitution with cauliflower or porcine heart E3. A higher level of

activity was obtained with the cauliflower E3 in comparison with the porcine E3, suggesting that the mode of binding of the two E3s was different. The determination of kinetic parameters was carried out using porcine heart E3 because of the difficulties encountered in the purification of cauliflower E3 and the inhibition of the overall reaction produced by impurities in the purified samples. Direct molecular mass measurements were not carried out but a complex with a high M_r was suggested by sedimentation analysis and elution of the complex in the void volume from a Sephadex G-100 column. OGDC has only been reported in plant mitochondria and this location was confirmed by A. Carmichael in pea (PhD thesis, Glasgow University, 1994). However, due to the poor activity of OGDC in plant extracts, its presence in other organelles cannot be ruled out.

The reconstituted complex was very similar to bacterial and mammalian complexes for the following reasons. It was found to be totally dependent on 2-oxoglutarate, CoA, ThDP and NAD^+ for activity, with a pH optimum of 6.9, and underwent product inhibition with succinyl-CoA. As with mammalian and plant PDC, the addition of Mg^{2+} or Ca^{2+} ions restored activity of the complex, diminished by the addition of 1 μM EDTA. However, in contrast to mammalian PDC but similar to bacterial OGDC, 5′-AMP is found to be the preferred adenine nucleotide for activation by reducing the K_m of the enzyme for 2-oxoglutarate, causing a 10-fold activation (Wedding & Black, 1971). Craig and Wedding (1980) discovered later that this effect was pH dependent and at pH values above 8, 5′-AMP inhibited by increasing the K_m. The effect of AMP is thought to be in the tighter binding of Mg-ThDP to the E1 component. In addition, there seems to be an interaction in the regulation of OGDC and PDC within plant mitochondria as OGDC was shown to be extremely sensitive to pyruvate inhibition (Dry & Wiskich, 1985).

1.10 Plant BCDC

Gerbling and Gerhardt (1988) demonstrated the presence of BCDC in peroxisomes from mung bean hypocotyls. It catalysed, in a CoASH- and NAD-

dependent reaction, the release of CO_2 from branched-chain 2-oxoacids and the concomitant formation of NADH and the acyl-CoA, which contains one carbon atom less than the substrate. Fatty acid degradation occurs in the peroxisomes of plant cells and the substrate in these non-fatty tissues results from the degradation of branched chain amino acids (leucine, isoleucine and valine). Intermediates of their catabolism are branched-chain 2-oxoacids which are activated by BCDC (Gerbling & Gerhardt, 1987). The branched-chain 2-oxoacids are then converted to propionyl-CoA by a series of steps, which are then metabolised by modified β-oxidation to acetyl-CoA. The physiological role of peroxisomes has yet to be determined, although the analogous organelles, glyoxysomes, in fatty tissue are involved in the conversion of lipid to sucrose during germination (Gerbling & Gerhardt, 1989).

BCDC has not been purified and so regulatory mechanisms have not been determined. A. Carmichael (Ph.D. thesis, Glasgow University, 1994) reported a possible BCDC location in pea and potato mitochondria. Therefore, the absence of BCDC activity in the past may be due to inactivation of the complex by phosphorylation or the developmental synthesis of BCDC.

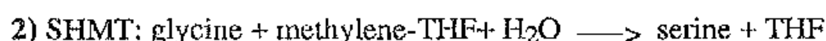
1.11 Glycine Decarboxylase (GDC)

GDC is a loosely associated multienzyme complex located in the mitochondria which catalyses the oxidative decarboxylation of glycine (Walker & Oliver, 1986a). It is particularly worth mentioning as it has many similarities with the 2-oxoacid dehydrogenase complexes. Additionally, the activities of GDC determine the regulation of plant PDC in the light, as discussed above.

GDC has been found in plants (Sarojini & Oliver, 1983), animals (Hiraga & Kikuchi, 1980) and bacteria (Klein & Sagers, 1966). In mammalian tissues, GDC converts glycine derived from protein catabolism into one-carbon fragments that can be used in the anabolic functions of the mitochondria and serine that can enter three-carbon metabolism and glycolysis (Oliver, 1994). The complex is found at its

highest levels in the mitochondria of photosynthetic tissue and its physiological relevance is that it carries out a key step in photorespiration. The photorespiratory cycle diverges from the photosynthetic carbon fixation reactions with the oxygenation of ribulose 1,5-bisphosphate by the oxygenase reaction of ribulose-1,5-bisphosphate carboxylase/ oxygenase (Rubisco) in the chloroplast (Husic *et al.*, 1987). The recycling of the phosphoglycolate into 3-phosphoglycerate for the Calvin cycle proceeds in leaf cells via a long chain of reactions in which O_2 and ATP are consumed while CO_2 is released and spans three different organelles (Fig. 1.9). The product of the reaction, phosphoglycolate, is hydrolysed to glycolate and converted to glycine by transamination in the peroxisome. Glycine enters the mitochondria where it is cleaved by GDC to CO_2, NH_3 and N^5,N^{10}-methylene-THF with the reduction of NAD^+ to NADH. The latter compound reacts with a second molecule of glycine to form serine in a reaction catalysed by serine hydroxymethyltransferase (SHMT). Serine then returns to the peroxisomes where it is converted to glycerate for re-entry to the chloroplast. Ribulose 1,5-bisphosphate is regenerated from the glycerate via glycerate 3-phosphate.

1) GDC: glycine + NAD^+ + THF ——> methylene-THF + CO_2 + NH_4^+ + NADH


2) SHMT: glycine + methylene-THF+ H_2O ——> serine + THF

overall reaction: 2 glycine + NAD^+ ——> serine + CO_2 + NH_4^+

In C_3 plants, the flux of carbon through the photorespiratory pathway is about equal on a molar basis to the flux through the photosynthetic C_3 cycle, thus making it a major metabolic pathway in leaves. The conversion of glycine to serine in green leaf mitochondria is currently considered to be the major source of CO_2 released during photorespiration (Douce *et al.*, 1994). The CO_2 released is about 25% of the amount fixed by Rubisco, and so this pathway decreases the carbon that is available to support plant growth. The rate of NH_4^+ release is nearly 10 times

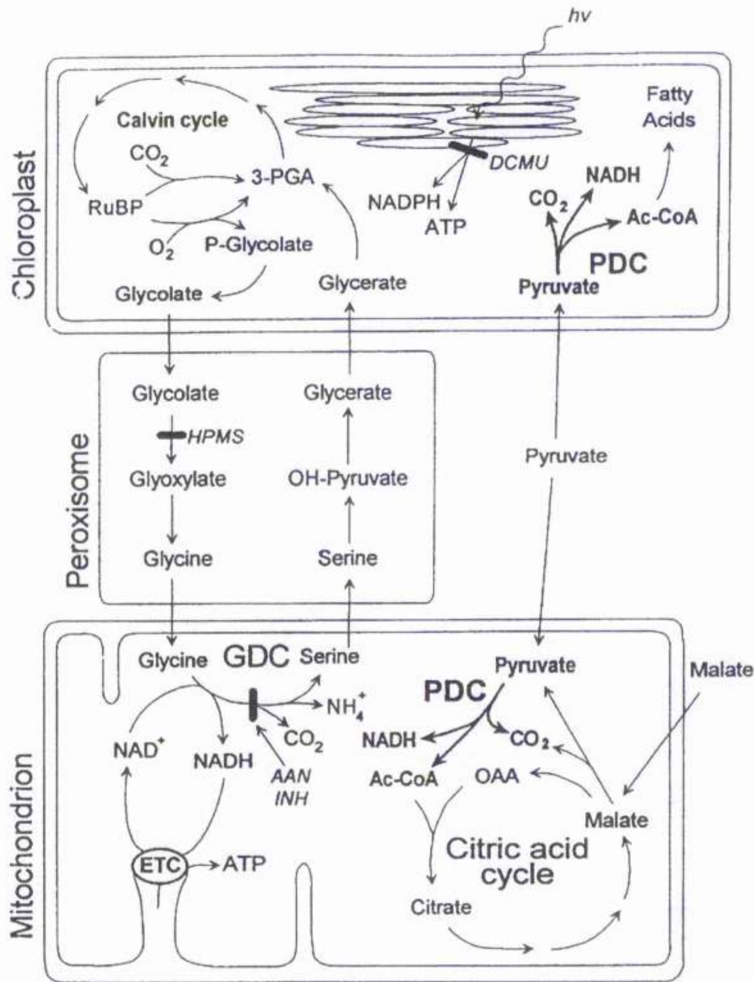


Fig. 1.9: Interactions of the photosynthetic, photorespiratory and citric acid cycles

Not all the intermediates are shown and stoichiometry has not been maintained. GDC: Glycine decarboxylase complex, PDC: Pyruvate dehydrogenase complex. The photorespiratory cycle takes place in all three organelles. Reproduced from Leuthy *et al*. (1996).

higher than the rate of primary NH_4^+ fixation, and the released NH_4^+ must be recycled with good efficiency to prevent the loss of nitrogen (Oliver, 1994). Although the NADH produced is available to the mitochondrial electron transport chain, the photorespiratory cycle results in net energy consumption

The reaction carried out by GDC requires the sequential action of four enzymes: 1) P-protein (97 kDa homodimer): a pyridoxal 5-phosphate (PLP)-dependent amino acid decarboxylase; 2) H-protein (13.9 kDa monomer): a carrier protein with a bound lipoamide cofactor; 3) T-protein (45 kDa monomer): a tetrahydrofolate (THF) transferase; 4) L-protein (50 kDa homodimer): a flavoprotein dihydrolipoamide dehydrogenase. The reaction, shown in Fig. 1.10, begins with the α-amino group of glycine forming a Schiff base with pyridoxal 5-phosphate at the active site of the P-protein. The lipoamide containing H-protein reacts with glycine bound to the P-protein to form a methylamine-H-protein intermediate. During the reaction the α-carboxyl of glycine is lost as CO_2. The T-protein catalyses the transfer of the carbon of the methylamine group to tetrahydrofolate (THF), with the formation of N^5,N^{10}-methylene-THF and the release of the amino group of glycine as NH_3. The reduced lipoamide from the H-protein, resulting from this transfer migrates to the active site of L-protein and where it is reoxidised back to lipoamide with the sequential reduction of FAD and NAD^+.

The dihydrolipoamide acyltransferase function of E2 has been divided between the H-protein and the T-protein of GDC. The long flexible arm of the lipoamide attached to a lysine residue of H-protein moves reaction intermediates and reducing equivalents between the reactive sites located on the other three subunits of the complex in a similar manner to E2 (Walker & Oliver, 1986a). The transferase domain is located in T-protein and is a tetrahydrofolate-based methyltransferase instead of the CoASH-dependent acyl transferase used by PDC, OGDC and BCDC. L-protein is the only component to be involved with other complexes

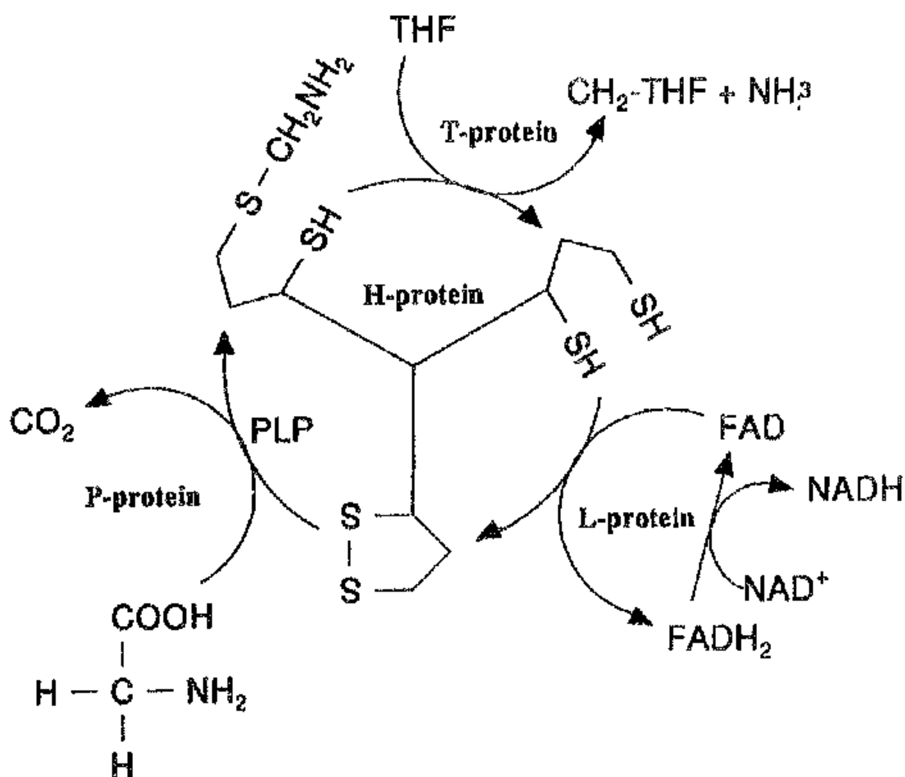


Fig. 1.10: The reaction mechanism of glycine decarboxylase multienzyme complex

The GDC reaction requires all four component proteins of the multienzyme complex. The lipoamide cofactor on the H-protein carries reaction intermediates between the reactive sites of the P-protein, T-protein and L-protein.

Reproduced from Oliver (1994).

(PDC and OGDC) in pea leaves and is encoded by a single gene (Bourguignon *et al.*, 1992; Bourguignon *et al.*, 1996).

Mitochondria isolated from etiolated pea shoots and potato tubers oxidise glycine poorly, whereas mitochondria from green leaves show rapid rates of glycine oxidation (Oliver, 1994). GDC is present at a very high concentration in plant mitochondria, where it comprises 30-50% of the soluble proteins present in the matrix (Oliver *et al.*, 1990). This is necessary to cope with the rapid glycine release during the course of photorespiration which can be as much as 50% of the photosynthetic rate and ten times the rate of the TCA cycle (Douce *et al.*, 1994). This is in contrast with mammalian mitochondria where GDC represents a minor fraction of the total matrix protein. The high protein concentration influences the formation of a loose multienzyme complex in pea leaf mitochondria with an approximate subunit ratio of 2 P-dimers:27 H-protein monomers: 9 T-protein monomers:1 L-protein dimer (Oliver *et al.*, 1990). Upon dilution the complex tends to dissociate into its component enzymes. This has precluded purification of the intact complex by gel filtration or ultracentrifugation. The only method that yields intact complex is the isolation of the separate enzymes followed by reconstitution *in vitro* (Bourguignon *et al.*, 1988; Walker & Oliver, 1986a)

As with other photosynthetic and photorespiratory enzymes, the synthesis of GDC is light dependent. The component enzymes are present in low amounts in etiolated leaves and increases dramatically (about 10-fold) upon exposure to light (Walker & Oliver, 1986b). The increase in activity results from *de novo* protein synthesis, largely regulated at the transcriptional level (Kim & Oliver, 1990; Bourguignon *et al.*, 1988). Day *et al.* (1985) and Bourguignon *et al.* (1988) have shown that this accumulation led to an increase in the density of mitochondria on Percoll gradients. There is also evidence for spatial regulation in the expression of GDC activity. In spinach, GDC is abundant in mitochondria from photosynthetic tissues whereas it is expressed at much lower levels in the mitochondria of roots, stalks and leaf veins (Gardestrom *et al.*, 1980)

GDC is subject to feedback inhibition by the products of the reaction, NADH and serine (Bourguignon *et al.*, 1988) . However this is prevented by unique substrate transporters in plant mitochondria. NADH is shuttled out of the mitochondria by an oxaloacetic acid/malate transporter in conjunction with malate dehydrogenase activity in the matrix and cytosol (Oliver & Walker, 1984). Glycine/ serine exchange proteins in the inner mitochondrial membrane prevent inhibition by serine.

1.12 Aims of this Thesis

As discussed earlier, E3 is a single gene product in the mammalian and pea 2-oxoacid dehydrogenase complexes. Evidence in prokaryotes for complex specific isoforms and for isoforms unrelated to any complexes has been increasing in recent years, dispelling the idea that E3 is a single gene product in all organisms. The primary aim was to investigate the possible complex specific roles of three E3 isoforms, termed α_2, $\alpha\beta$ and β_2, discovered in potato tuber mitochondria by Fullerton *et al.* (1996). The idea was based on previous work which suggested tissue specific expression of the isoforms. The α_2 form was found to predominate in leaf mitochondria where GDC is the major complex and the β_2 form was found in greater amounts in tuber mitochondria and was postulated to be specific for PDC and OGDC. A complex specific role of the three isoforms from potato mitochondria seemed more plausible with the discovery of an organelle specific E3 isoform in chloroplasts.

Initially, complex specificity was to be addressed by determining if any of the potato E3 isoforms preferentially restored intact complex activity to plant and mammalian OGDC E1/E2 subcomplexes which had been stripped of their native E3 components. The results of these experiments were to be compared with similar experiments carried out on PDC and GDC by others. Prior to the reconstitution experiments being carried out, the yield of the three potato isoforms obtained by the existing purification procedure required improvement.

The latter approach employed was the cloning and expression of the potato E3 gene(s) in order to determine the relationship between the α and β polypeptides. The added significance of this work was that the E3 gene(s) from potato would represent only the second to be cloned from a plant source to date.

Chapter 2

Materials and Methods

2.1 Materials

2.1.1 Plants

Potato tubers (*Solanum tuberosum*, cv. Maris Piper) were purchased from Glasgow Fruit Market.

2.1.2 Animals

Bovine hearts were obtained from Paisley Abbatoir, Sandyford Road, Paisley. New Zealand White Rabbits for antisera production were bought at 3 months old from MRC accredited resources.

2.1.3 cDNA Library

The potato leaf cDNA library was a kind gift of Dr. H. Hesse, Max Planck Institute for Plant Molecular Physiology, Golm, Germany. It was made using the Uni-ZAP XR vector kit containing Lambda ZAP II, a lambda insertion type cDNA cloning vector (Stratagene, USA).

2.1.4 Bacterial and Bacteriophage Strains

Escherichia coli **XL1-Blue MRF´** (Stratagene Ltd., Cambridge): a restriction and recombination deficient host strain for propagating plasmid DNA, plating Lambda ZAP II cDNA libraries and double-stranded sequencing.

Genotype: Δ(*mcrA*)183, Δ(*mcrCB-hsdSMR-mrr*) 173, *endA*1, *supE*44, *thi*-1 *recA*1, *gyrA*96, *relA*1, Lac [F´*proAB*, *lacI*^{q}ZΔM15, TnI0(tet^{r})] (Bullock *et al.*, 1987)

Escherichia coli **DH5α**: a recombination deficient, suppressing strain for the propagation of plasmid DNA.

Genotype: *supE*44 Δ*lacU*169 (ϕ80 *lac*ZΔM15) *hsdR*17 *recA*1 *endA*1 *gyrA*96 *thi*-1 *relA*1 (Hanahan, 1983).

Escherichia coli **BL21(DE3)** (Novagen, USA): employed for the high level expression of genes cloned into expression vectors containing the bacteriophage T7 promoter.

Genotype: *hsd*S gal (λ*cIts*857 *ind*1 *Sam*7 *nin*5 *lacUV*5-T7 gene 1) (Studier & Moffatt, 1986)

Escherichia coli **SOLR** (Stratgene Ltd, Cambridge): a non-suppressing strain used with VCSM13 helper phage (Stratagene Ltd, Cambridge) for the *in vivo* excision of Bluescript SK- from Lambda ZAP II vectors.

Genotype: SOLR: e14-(*mcrA*) Δ(*mcrCB-hsd*SMR-mr^r)171 *sbcC recJ uvrC umuC* :: Tn5(kan^r) *lac gyrA*96 *relA*1 *thi*-1 *endA*1 λ^R [F′ *proAB lacI*^{q}ZΔM15] Su^- (Hay & Short, 1992)

2.1.5 Plasmid Vectors

The vectors used for cloning in this thesis are tabulated below and presented Figs. 2.1 and 2.2. The maps of pBluescript SK +/- (Fig. 2.3) and Lambda ZAP II (Fig. 2.4) are presented for reference.

Name	Characteristics	Purpose
pBluescript II KS+ (Stratagene)	Ampicillin resistance. Lac Z gene for blue/white colour selection. F′ origin of replication for single-strand rescue by helper phage. M13 forward and reverse primer sites for sequencing.	General cloning vector
pET28c (Novagen)	Kanamycin resistance. Ribosome-binding site and initiating codon for translation. Histidine tag.	Heterologous expression of foreign genes in *E. coli*

pBluescript® II KS (+/-) phagemid

The pBluescript® II KS (+/-) phagemid is a 2961-bp phagemid derived from pUC19. The KS designation indicates the polylinker is oriented such that *lacZ* transcription proceeds from *Kpn* I to *Sac* I.

f1 (+) origin: (3-459 bp) f1 filamentous phage origin of replication allowing recovery of the sense strand of the *lacZ* gene when a host strain containing the pBluescript II phagemid is co-infected with helper phage.

f1 (-) origin: (3-459 bp) f1 filamentous phage origin of replication allowing recovery of the antisense strand of the *lacZ* gene when a host strain containing the pBluescript II phagemid is co-infected with helper phage.

ColE1 origin: (1032-1972 bp) Plasmid origin of replication used in the absence of helper phage.

***lacZ* gene:** (*lac* promoter: 816-938 bp) This portion of the *lacZ* gene provides α-complementation for blue/white color selection of recombinart phagemids. An inducible *lac* promoter upstream from the *lacZ* gene permits fusion protein expression with the β-galactosidase gene product.

MCS: (657-759 bp) Multiple cloning site flanked by T3 and T7 RNA promoters (please see the polylinker sequence below).

Ampicillin: (1975-2832 bp) Ampicillin-resistance gene (Amp^r) for antibiotic selection of the phagemid vector.

Please Note: The upper strand is designated the (+) strand and the lower strand is designated the (-) strand

GenBank® # X52327 [KS(+)] & # X52329 [KS(-)]

Nae I 131
f1 (-) origin
Ssp I 442
Ssp I 2850
Ssp I 19
Nae I 330
f1 (+) origin
Xmn I 2645
Sca I 2526
Pvu I 2416
Ampicillin
Pvu I 500
Pvu II 529
lacZ
BssH II 619
Sac I 657
T7
Kpn I 759
BssH II 792
T3
pBluescript® II KS (+/-)
2961 bp
Pvu II 977
ColE1 origin
Afl III 1153

Reverse primer
5' GGAAACAGCTATGACCATG 3'

T3 primer
5' AATTAACCCTCACTAAAGGG 3'

KS Primer
5' TCGAGGTCGACGGTATC 3'

MET · BssH II · T3 promoter +1 · Kpn I · Eco0109 I / Dra II / Apa I · Xho I · Hinc II / Acc I / Sal I · Bsp106 I / Cla I · Hind III · EcoR V · EcoR I

```
5' GGAAACAGCTATGACCATGATTACGCCAAGCGCGCAATTAACCCTCACTAAAGGGAACAAAAGCTGGGTACCGGGCCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATCG
3' CCTTTGTCGATACTGGTACTAATGCGGTTCGCGCGTTAATTGGGAGTGATTTCCCTTGTTTTCGACCCATGCGCCGGGGGGGAGCTCCAGCTGCCATAGCTATTCGAACTATAGCTTA
```

816 · 792 · 759

β Galactosidase

Pst I · Sma I · BamH I · Spe I · Xba I · Eag I / Not I · BstX I · Sac II · Sac I · BssH II

```
AATTCCTGCAGCCCGGGGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAATTCGCCCTATAGTGAGTCGTATTACGCGCGCTCACTGGCCGTCGTTTTACAA 3' (+
```

657 · +1 T7 promoter · 619

3' CTAGGTGATCAAGATCTCGC 5'
SK Primer

3' CGGGATATCACTCAGCATAATG 5'
T7 Primer

3' TGACCGGCAGCAAAATG 5'
M13 -20 Primer

pET-28a-c(+) Vectors

The pET-28a-c(+) vectors carry an N-terminal His•Tag®/thrombin/T7•Tag® configuration plus an optional C-terminal His•Tag sequence. Unique sites are shown on the circle map. Note that the sequence is numbered by the pBR322 convention, so the T7 expression region is reversed on the circular map. The cloning/expression region of the coding strand transcribed by T7 RNA polymerase is shown below. The f1 origin is oriented so that infection with helper phage will produce virions containing single-stranded DNA that corresponds to the coding strand. Therefore, single-stranded sequencing should be performed using the T7 terminator primer.

pET-28a(+) sequence landmarks	
T7 promoter	370-386
T7 transcription start	369
His•Tag coding sequence	270-287
T7•Tag coding sequence	207-239
Multiple cloning sites	
(*Bam*H I - *Xho* I)	158-203
His•Tag coding sequence	140-157
T7 terminator	26-72
lacI coding sequence	773-1852
pBR322 origin	3286
Kan coding sequence	3995-4807
f1 origin	4903-5358

The maps for pET-28b(+) and pET-28c(+) are the same as pET-28a(+) (shown) with the following exceptions: pET-28b(+) is a 5368bp plasmid; subtract 1bp from each site beyond *Bam*H I at 198. pET-28c(+) is a 5367bp plasmid; subtract 2bp from each site beyond *Bam*H I at 198.

pET-28a-c(+) cloning/expression region

pBluescript® SK (+/-) phagemid

The pBluescript® SK (+/-) phagemid is a 2958 bp phagemid derived from pUC19. The SK designation indicates the polylinker is oriented such that *lacZ* transcription proceeds from *Sac* I to *Kpn* I.

f1 (+) origin: (6-462 bp) f1 filamentous phage origin of replication allowing recovery of the sense strand of the *lacZ* gene when a host strain containing the pBluescript phagemid is co-infected with helper phage.

f1 (-) origin: (6-462 bp) f1 filamentous phage origin of replication allowing recovery of the antisense strand of the *lacZ* gene when a host strain containing the pBluescript phagemid is co-infected with helper phage.

ColE1 origin: (1032-1912 bp) Plasmid origin of replication used in the absence of helper phage.

***lacZ* gene:** (*lac* promoter: 816-938 bp) This portion of the *lacZ* gene provides α-complementation for blue/white color selection of recombinant phagemids. An inducible *lac* promoter upstream from the *lacZ* gene permits fusion protein expression with the β-galactosidase gene product.

MCS: (657-759 bp) Multiple cloning site flanked by T3 and T7 RNA promoters (please see the polylinker sequence below).

Ampicillin: (1975-2832 bp) Ampicillin resistance gene (Amp[r]) for antibiotic selection of the phagemid vector.

Please Note: The upper strand is designated the (+) strand and the lower strand is designated the (-) strand.

GenBank® #52325 [SK(+)] & #52324 [SK(-)]

Lambda ZAP® II Vector*

40,820 bp

Short, J.M., et al. (1988) *Nucl. Acids Res.* 16: 7583-7600.

* U.S. Patent No. 5,128,256

β-Galactosidase →

T3 → ← T7

Left End 0
Bgl II 0.42
Mlu I 0.46
Mlu I 5.55
Apa I 10.09
Pvu I 11.93
SnaB I 12.19
Mlu I 15.37
Kpn I 17.05
Mlu I 17.79
Kpn I 18.56
Sac I 22.03
Not I 22.05
Xba I 22.06
Spe I 22.06
BamH I 22.07
Sma I 22.07
EcoR I 22.09
Hind III 22.10
Xho I 22.12
Kpn I 22.14
Hind III 22.48
BamH I 22.97
Sma I 26.62
BamH I 29.50
Bgl II 30.71
Hind III 31.90
Hind III 32.47
Bgl II 33.10
Bgl II 33.75
Bgl II 33.81
Sma I 34.89
Hind III 36.46
Nar I 38.00
Right End 40.82

A — J
Amp^r
ColE1
att
int
xis
cI857 (nin5)

2.1.6 Chemicals

The following reagents were obtained from Sigma Chemical Co., Poole, Dorset, UK: substrates and coenzymes for enzymatic assays, phenylmethylsulphonyl fluoride (PMSF), benzamidine-HCl, Coomassie Brilliant Blue type R250, antifoam A concentrate, Percoll, thio-octic acid amide, 3-(N-Morpholino) propane-sulphonic acid (MOPS), polyvinylpyrrolidone (PVP), N-tris[hydroxymethyl]-methyl-2-aminoethane-sulphonic acid (TES), L-ascorbic acid and diethyl pyrocarbonate (DEPC).

Polyethylene glycol 6000 grade, Triton X-100 and Tris-buffered phenol were obtained from Fisons, Loughborough, England. Pyronin Y dye was obtained from George T. Gurr Ltd., London. Folin and Ciocalteu's phenol reagent was from FSA Laboratory supplies, Loughborough, England. 2-Mercaptoethanol was purchased from Prolabo, Paris, France. Leupeptin was from the Marketing Association, Herts, England. DTT was obtained from Boehringer Mannheim GmbH, Germany. CTAB, sarcosyl and guanidine thiocyanate were from Fluka, Germany. Growth media for bacterial cultures were purchased from Mikrobiologie. All other chemicals including reagents for SDS-polyacrylamide gel electrophoresis (SDS-PAGE) were of Analar grade from BDH Chemicals Ltd., Poole, UK. or were of the highest grade commercially available.

2.1.7 Enzymes

Bovine serum albumin (BSA), protein A (Staphylococcus aureus, Cowan 1 strain), dihydrolipoamide dehydrogenase from porcine heart, bovine mucosa and yeast were obtained from Sigma Chemical Co., Poole, Dorset.

DNA restriction enzymes, DNA ligase, calf intestinal alkaline phosphatase, proteinase K, *Pfu* and *Taq* polymerases were supplied by Promega Ltd., Southampton or Boehringer Mannheim Ltd., Lewes, UK.

2.1.8 Molecular Size and Weight Markers

Low molecular weight marker proteins for M_r determinations by SDS-PAGE were purchased from Pharmacia Ltd., Milton Keynes. The standards comprised phosphorylase b (M_r 94 000), albumin (M_r 67 000), ovalbumin (M_r 43 000), carbonic anhydrase (M_r 31 000), soya bean trypsin inhibitor (M_r 20 100) and lysozyme (M_r 14 000). The following standards were used as size markers for the analysis of nucleic acids by gel electrophoresis:

100 bp ladder (GIBCO/BRL)

RNA ladder (GIBCO/BRL)

DRIgest III,λ DNA-*Hind*III/φX-174 RF DNA-*Hae*III digest (Pharmacia)

2.1.9 Synthetic Oligonucleotides

Synthetic oligonuleotides for PCR were made in the Division of Biochemistry and Molecular Biology by Dr. V. Math, using an Applied Biosystems 381A DNA synthesiser for the phosphite-triester method.

M13 Universal and M13 Reverse primers were purchased from Pharmacia Ltd., Milton Keynes, England.

2.1.10 Photographic Materials

The X-Omat-100 processor and X-Omat S film was obtained from Kodak Ltd., Dallimore Road, Manchester, England. High performance Hyperfilm-ECL was supplied by Amersham International, Bucks, UK.

2.1.11 Miscellaneous

FPLC columns, gel chromatography materials and CNBr-activated Sepharose 4B were purchased from Pharmacia Ltd., Milton Keynes. Nitrocellulose filters (0.22μm) and sterile Millex-gv syringe filters were from Millipore S.A., Molsheim, France. Rat serum, Freund's complete and incomplete adjuvant were obtained from Sigma Chemical Co., Poole, Dorset. Hybond-C extra, Hybond N+, the ECL immuno blotting kit and [α-^{32}P]-dCTP (3000

Ci/mmol) were bought from Amersham International, Bucks., U.K. Non-immune donkey serum was supplied by the Scottish Antibody Production Unit (SAPU), Lanarkshire, Scotland. The Micro BCA assay kit was obtained from Pierce, Illinois, USA. Plast-X autoradiography cassettes were from Anthony Monk (England) Ltd., Sutton-in-Ashfield, U.K. Centricon-10, 30 and Centriprep-10 and 30 concentrators were bought from Amicon Ltd., Stonehouse, Gloucestershire. Chromaspin-30 and 100, equilibrated in TE were purchased from Clonetech, USA. Screw cap, Oak Ridge centrifuge tubes (PPCO) were from Nalgene, USA.

2.2 Methods

2.2.1 Growth of Plants

Potato tubers were allowed to sprout by storing in the dark, at room temperature, for approx. two weeks. Sprouts were removed with approx. 3-4 cm of tuber tissue, planted in pots of compost and grown at 200 $mE.m^{-2}.s^{-1}$, in a growth chamber with a 12 h photoperiod (22°C light/18°C dark). After the primary leaves had emerged the plants were re-potted and leaf tissue was collected for harvesting approx. 3 weeks later.

2.2.2 Determination of Protein Concentration

Protein concentrations were determined by the method of Lowry *et al.* (1951), modified by Markwell *et al.* (1976). For very low protein concentrations (0.5 $\mu g.ml^{-1}$ - 20 $\mu g.ml^{-1}$) the Micro BCA Assay was used. Protein standards were constructed with bovine serum albumin and absorbances were read at 660 nm and 562 nm respectively.

2.2.3 Dialysis of Samples

Dialysis was carried out in Visking Tubing at 4°C, with enough buffer changes to dilute out salts or exchange buffers. The tubing was pre-treated by boiling for 5 min in 1 mM EDTA (pH 8.0) and 10 mM sodium bicarbonate. It was

then rinsed and boiled in distilled water for a further 5 min and stored in 100% ethanol.

2.2.4 Concentration of samples

Protein samples were concentrated in dialysis tubing by covering with dry poly(ethylene glycol) (PEG 6000) at 4°C until the desired volume had been reached. The other method employed was ultrafiltration using Amicon Centricon or Centriprep concentrators with a molecular weight cut-off of 10 000 or 30 000 Da. The samples were spun for 20-40 min at 5000 g. They were useful for concentrating and desalting in a single step.

2.2.5 SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE)

Proteins were resolved using the discontinuous buffer system according to the method of Laemmli (1970) under denaturing conditions. Stock solutions were used to polymerise resolving gels of 7-15% (w/v) acrylamide containing 0.368 M Tris-HCl, pH 8.8, 0.1% (w/v) SDS (sodium dodecyl sulphate), 0.1% (w/v) ammonium persulphate and 0.001% (v/v) TEMED (N,N,N',N'-tetramethylethylenediamine) with stacking gels of 4.4% (w/v) acrylamide containing 0.12 M Tris-HCl, pH 6.8, 0.1% SDS (w/v), 0.1% (w/v) ammonium persulphate and 0.001% (v/v) TEMED. Solutions were filtered through 0.2 μm filters, degassed prior to use and gels stored at 4°C.

Stock solutions:

Acrylamide solution:	29.2% (w/v) acrylamide
	0.8% (w/v) bisacrylamide
Resolving gel buffer:	0.735 M Tris-HCl, pH 8.8
	0.2% (w/v) SDS
Stacking gel buffer:	0.17 M Tris-HCl, pH 6.8
	0.14% (w/v) SDS
Ammonium persulphate:	10% (w/v)

Protein samples were boiled in an equal vol of Laemmli sample buffer (containing 0.062 M Tris-HCl, pH 6.8, 2% (w/v) SDS, 10% (w/v) sucrose, 10 mM DTT and a trace of pyroninY) for 2-3 min before loading on to the gel.

Gels were run using vertical electrophoresis kits from Bethesda Research Laboratories (BRL) at a constant current of 60 mA. Electrode buffer contained 0.025 M Tris-HCl pH 8.3, 0.2 M glycine and 1% (w/v) SDS.

2.2.6 Staining and Destaining Gels

Staining was carried out using 0.1% (w/v) Coomassie Brilliant Blue R-200 dissolved in 10% (v/v) acetic acid, 50% (v/v) methanol. Gels were destained in 20% (v/v) methanol, 10% (w/v) acetic acid. Gels to be dried for storage were first allowed to soak overnight in 35% (v/v) ethanol and 2% (v/v) glycerol before being dried in an Easy Breeze apparatus (Hoefer Scientific Instruments, USA).

Silver staining was used as an alternative to Coomassie Blue staining when nanogram quantities of protein/band were to be analysed. The procedure was carried out using the method of Wray *et al.* (1981). Protein was fixed by immersing the gel in 50% (v/v) methanol for 8 h at room temperature. The gel was then placed for 20 min in the staining solution consisting of 0.8% (w/v) silver nitrate, 0.12 M ammonia solution and 0.08% (w/v) NaOH with constant, gentle shaking. The gel was washed with several changes of distilled water over the course of an hour and a developer (0.24 mM citric acid and 0.02% (v/v) formaldehyde), prepared not more than 5 min before use, was then added until all protein bands could be visualised. The gel was washed again with distilled water to remove all traces of developer.

2.2.7 Western Blotting and Immunological Detection of Proteins by Enhanced Chemiluminescence (ECL)

Following separation by gel electrophoresis, proteins were electrophoretically transferred onto Hybond-C nitrocellulose paper in the presence of buffer containing 192 mM glycine, 0.02% (w/v) SDS, 20% (v/v) methanol and

25 mM Tris-HCl (pH 8.3) as described by Towbin *et al.* (1979) and Batteiger *et al.* (1982). This was carried out in a BioRad Trans-Blot cell either at 40 mA overnight or at 400 mA for 2.5 h. The nitrocellulose was briefly stained with Ponceau S solution which allowed visualisation of transferred protein bands. The membrane was washed free of this stain with distilled water and excess binding sites were blocked by immersing it for 1-2 h with agitation, at 37°C in a solution containing 20 mM Tris-HCl, pH 7.2, 15 mM NaCl, 5% (w/v) non-fat dried milk, 5% (v/v) normal donkey serum and 0.2% (v/v) Tween-20.

Excess blocking reagent was then washed from the membrane at room temperature, with a solution containing 20 mM Tris-HCl, pH 7.2, 15 mM NaCl, 1% (w/v) non-fat dried milk and 1% (v/v) normal donkey serum. Washing was carried out twice briefly, once for 15 min and then twice for 5 min using 200 ml of wash buffer each time. The membrane was then incubated overnight, at 4°C in 100 ml of solution containing 20 mM Tris-HCl, pH 7.2, 1% (w/v) non-fat dried milk, 1% (v/v) normal donkey serum, 0.1% (v/v) Tween-20 and the appropriate dilution of the primary antibody.

The membrane was washed again as detailed above and incubated for 2 h at room temperature with a 1:1000 dilution of horseradish peroxidase labelled secondary antibody. This was carried out in approx. 10 ml of buffer containing 20 mM Tris-HCl, pH 7.2, 150 mM NaCl, 1% (w/v) non-fat dried milk, 1% (v/v) normal donkey serum. In order to keep the volumes of the antibody solutions as small as possible, incubation was sometimes carried out in a sealed bag. Following incubation in secondary antibody the membrane was given three 30 min washes in wash buffer and one 30 min wash with a buffer containing 20 mM Tris-HCl, pH 7.2 and 150 mM NaCl. The detection steps were carried out in a dark room following the protocol outlined in the ECL kit manual. Exposure times were typically 15 s to 1 min.

2.2.8 Preparation of Antiserum

Antisera to the native 2-oxoacid dehydrogenase complexes isolated from bovine heart were prepared as in the method reported by De Marcucci *et al.* (1985). Subunit specific antisera were prepared as follows: 0.25-1.0 mg of desired antigen, dissolved in 0.75 ml of 0.9% (w/v) NaCl and mixed with 0.75 ml of complete Freund's adjuvant was injected subcutaneously into rabbits at various sites in the neck, back and thighs. A booster dose was then administered four weeks later consisting of 0.1-0.5 mg antigen in 0.75 ml of 0.9% (w/v) NaCl and 0.75 ml of Freund's incomplete adjuvant. The rabbits received a further boost 2 weeks prior to bleeding. Antiserum was collected by bleeding the rabbit from an ear vein. The blood was allowed to clot overnight at 4°C and the antiserum was removed using a Pasteur pipette. Antiserum was centrifuged at 700 g for 5 min to remove any remaining red blood cells and stored in 1 ml aliquots at -80°C. Further collections of antisera were made at monthly intervals, 2 weeks after additional booster injections, using the regime described above.

2.2.9 Purification of Enzymes

2.2.9.1 Purification of Pyruvate Dehydrogenase Complex (PDC) and 2-Oxoglutarate Dehydrogenase Complex (OGDC) from Bovine Heart.

PDC and OGDC were purified according to the method of Stanley and Perham (1980) with the following modifications (De Marcucci & Lindsay, 1985). All procedures were performed at 4°C.

Bovine heart was cleared of fat and connective tissue and stored at -80°C until required. Routinely, 600 g of tissue was homogenised for 5 min in 1 litre of 50 mM MOPS buffer, pH 7.0 containing 2.7 mM EDTA, 0.1 mM DTT, 3% (v/v) Triton X-100, 1 mM PMSF, 1 mM benzamidine-HCl, silicone antifoam (0.5 $ml.l^{-1}$). The homogenate was then diluted with an equal vol of the same buffer and centrifuged at 10 000 g for 20 min. The pellets were discarded and the pH of the

supernatant adjusted to 6.45 with 10% (v/v) acetic acid. Protein was precipitated with the addition of 0.12 vol of 35% (w/v) polyethylene glycol 6000 (PEG 6000) and the solution was left stirring on ice for 30 min followed by centrifugation at 18 000 g for 15 min. The resulting pellets were resuspended by homogenisation with a loose-fitting Teflon glass homogeniser in 400 ml of 50 mM MOPS buffer, pH 6.8 containing 2.7 mM EDTA, 0.1 mM DTT, 1% (v/v) Triton X-100, 0.15 μM leupeptin, 1 mM PMSF, 1 mM benzamidine-HCl before being centrifuged at 25 000 g for 40 min. The supernatant was filtered through muslin to remove fat particles and 0.013 vol of 1 M $MgCl_2$ and 0.05 vol of 1 M sodium phosphate buffer (pH 6.3) was added with the dropwise addition of 0.5 M NaOH to keep the pH above 6.8. The pH was then adjusted to 6.45 with 10% (v/v) acetic acid for a second addition of 0.12 vol PEG 6000 (35% w/v). The supernatant was left to stir on ice for 30 min after which it was centrifuged at 25 000 g for 10 min. The pellets were resuspended in 160 ml of 2.7 mM EDTA, 0.1 mM DTT, 1% (v/v) Triton X-100, 1 mM PMSF, 1 mM benzamidine-HCl, 50 mM MOPS buffer, pH 6.8 and 0.5% (v/v) rat serum. Rat serum was added to the buffer as a source of protease inhibitors to prevent losses of up to 50% OGDC activity from occurring overnight (Wieland, 1975).

The following day, the suspension of partially purified enzyme was homogenised and centrifuged at 25 000 g for 1 h to remove any denatured protein. A third, differential PEG 6000 precipitation was then carried out by which PDC and OGDC were separated from each other. The pH was adjusted again to 6.45 with 10% (v/v) acetic acid and 0.06 vol of 35% (w/v) PEG 6000 was added to the stirring solution for 30 min. The amount of PEG added is variable at this point depending on the purity of PDC or OGDC required. It was found that 0.05 vol of PEG would give purer PDC and 0.04 vol for purer OGDC. This solution was centrifuged at 25 000 g for 10 min to pellet the OGDC. PDC was to be found in the supernatant and was pelleted by centrifuging at 100 000 g for 150 min. Both the PDC and OGDC pellets were resuspended in 4-6 ml of 50 mM MOPS buffer, pH

6.8 containing 2.7 mM EDTA, 0.1 mM DTT, 1% (v/v) Triton X-100, 0.15 μM leupeptin, 1 mM PMSF, 1 mM benzamidine-HCl and stored at 4°C.

2.2.9.2 Separation of E3 from E1/E2 Core of OGDC

Approx. 20-30 mg of OGDC were dissociated by a 20 min incubation on ice in the following buffer: 50 mM potassium phosphate buffer, pH 7.6, 1 M NaCl, 1 mM DTT and 0.01% (v/v) Triton X-100, 1 mM $MgCl_2$ and 0.2 mM ThDP. The E1/E2 and E3 components were separated by gel filtration using a Superose 12 column (HR 16/50, approx. 100 ml bed vol) which was attached to a Pharmacia FPLC system. The column was pre-injected with 2 M NaCl to maintain a dissociating front and was run in the same buffer as above. Peak fractions were dialysed in 50 mM potassium phosphate buffer (pH 7.6) containing 1 mM $MgCl_2$, 0.2 mM ThDP. The E3 fraction was heat-treated at 65°C for 10 min and the heat labile proteins removed by pelleting at high speed in a microfuge.

2.2.9.3 Preparation of a Crude Mitochondrial Dihydrolipoamide Dehydrogenase Extract from Potato Tubers

Potato tuber tissue was ground using a Waring commercial blender at high speed for five second bursts approx. five times. In each case 3-4 vol of grinding buffer were used containing 0.3 M sucrose, 50 mM MOPS buffer (3-[N-Morpholino]propane-sulphonic acid), pH 7.6, 2 mM EDTA, 2 mM $MgCl_2$, 1% (w/v) polyvinylpyrrolidone and 2 mM glycine. Fresh Iso-ascorbic acid (30 mM) was added on the day of preparation and the pH adjusted to 7.6.

The homogenate was filtered through four layers of pre-wetted muslin into cold centrifuge buckets containing protease inhibitors (1 mM phenylmethylsulphonyl fluoride (PMSF), 1 mM benzamidine-HCl and 0.15 μM leupeptin) and centrifuged at 1000 g for 5 min. The supernatant fluid was retained and centrifuged at 12000 g for 20 min. The resulting crude mitochondrial pellet was solubilised, with the aid of a hand homogeniser, in 200 ml of extraction buffer (50 mM potassium phosphate buffer, pH 6.8, 1 mM EDTA, 1 mM DTT, 0.1%

(v/v) Triton X-100, 1 mM phenylmethylsulphonyl fluoride (PMSF), 1 mM benzamidine-HCl and 10 μM leupeptin). The extract was heat-treated at 65°C for 30 min followed by pelleting of denatured protein at 12000 g for 20 min. The heat-treated extract was dialysed into 20 mM potassium phosphate buffer, pH 6.8 containing 1 mM DTT and the protease inhibitors in preparation for application to a DEAE-52 ion exchange column.

2.2.9.4 Preparation of Diethylaminoethyl Cellulose (DEAE-52) Ion Exchange Chromatography Columns

The dry cellulose was stirred in an excess of 20 mM Tris, pH 8 for 30 min after which it was left to settle. The buffer was replaced before pouring columns (55 ml bed vol) which were packed by gravity flow of the buffer through the material. The column was equilibrated with 20 mM potassium phosphate buffer, pH 6.8. Small columns (1 ml) were poured in Pasteur pipettes, plugged loosely at the bottom end with some glass wool.

2.2.9.5 Mitochondrial Isolation from Potato Tubers

Mitochondria from potato tubers were isolated by a process described by Day *et al.* (1988) with all procedures carried out at 4°C. Tubers were homogenised and filtered as described in section 2.2.9.3. The homogenate was centrifuged at 1000 g for 5 min, the supernatant retained and re-centrifuged at 12 000 g for 20 min. The resulting crude mitochondrial pellet was resuspended using a hand homogeniser in 30 ml of wash buffer containing 0.3 M sucrose, 0.1% (w/v) BSA, 1 mM glycine and 10 mM (N-tris[hydroxymethyl]-methyl-2-aminoethane-sulphonic acid (TES), pH 7.2. Both spins were repeated and the resultant mitochondrial pellet resuspended in approx. 2 ml of wash buffer. The mitochondrial extract was then layered onto a self generating gradient comprising 10 mM potassium phosphate, pH 7.2, 30% (v/v) Percoll, 1 mM EDTA, 0.1% (w/v) BSA and centrifuged at 40 000 g for 35 min. The mitochondria formed a distinct broad band near the bottom of the gradient. This could be removed using a Pasteur pipette

carefully so as not to disturb the upper regions of the gradient which contain the membrane fractions.

The Percoll was then removed from the mitochondrial fraction by diluting it in 5 vol of wash buffer minus glycine prior to centrifugation at 12 000 g for 15 min. The upper section of the supernatant fraction was removed and the pellet further diluted with wash buffer minus glycine and BSA and again centrifuged at 12 000 g for 15 min. This last step was repeated until a firm mitochondrial pellet was formed which was then resuspended in a minimal vol of wash buffer minus glycine and BSA.

2.2.9.6 Mitochondrial Isolation from Cauliflower Florets

Approximately 1.5 kg of cauliflower florets were homogenised using a Waring commercial blender at high speed for five second bursts approx. five times in 1.5 l of grinding buffer (50 mM MOPS, pH 7.4, 0.6 M sucrose, 1 mM KCl, 5 mM EDTA, 10 mM $MgCl_2$, 0.05% (w/v) BSA, 0.1% (w/v) PVP-40, 30 mM Iso-ascorbic acid). The homogenate was filtered through six layers of pre-wetted muslin into cold centrifuge buckets containing protease inhibitors (1 mM phenylmethylsulphonyl fluoride (PMSF), 1 mM benzamidine-HCl and 0.15 μM leupeptin). The homogenate was centrifuged at 6000 g for 15 min and the resulting pellet was gently teased into suspension with a paintbrush into 60 ml wash buffer containing 50 mM MOPS, pH 7.0, 0.4 M sucrose, 1 mM EDTA, and 10 mM $MgCl_2$. The supernatant was retained from the next centrifugation at 600 g for 10 min. It was further centrifuged at 12 000 g for 10 min and the pellets resuspended into 60 ml wash buffer (minus sucrose), containing 0.1% (v/v) Triton X-100. Solubilisation of the crude mitochondrial pellet was allowed to continue at 4°C for 30 min after which time it was centrifuged at 12 000 g for 10 min.

2.2.9.7 Purification of E3 from a Variety of Sources

Native E3 from bovine heart OGDC and PDC was isolated by a 20 min incubation at 4°C in the presence of 20 mM potassium phosphate buffer, pH 7.6

and 1 M NaCl followed by heat treatment at 65°C for 10 min to remove heat labile proteins (pelleted at 10 000 g, 15 min). The bovine heart E3s and commercial preparations of porcine heart (Type III) and *Candida utilis* yeast E3 (Type IV) were dialysed against 50 mM potassium phosphate buffer, pH 7.6, 1 mM $MgCl_2$ and 0.2 mM ThDP and concentrated on Centricon-30 centrifugal concentrators before use. The potato E3 isoforms were purified as described in section 2.2.9.3 and Chapter 3 (for Mono Q separation of the individual isoforms) and dialysed against the same buffer.

2.2.10 Enzyme Assays

2.2.10.1 2-Oxoacid Dehydrogenase Complexes

The overall activities of the 2-oxoglutarate dehydrogenase and pyruvate dehydrogenase complexes were assayed according to the method of Brown and Perham (1976). The reaction mixture used was 50 mM potassium phosphate buffer (pH 7.6) containing 3 mM NAD^+, 1 mM $MgCl_2$, 0.2 mM ThDP, 2.6 mM cysteine-HCl, 0.13 mM coenzyme A, 2 mM 2-oxoglutarate (or 2 mM pyruvate when assaying for PDC activity). Assays were performed at 30°C in a final vol of 0.7 ml. The assay was initiated by the addition of enzyme and the reaction followed by monitoring the formation of NADH spectrophotometrically at 340 nm (Danson *et al.*, 1978).

2.2.10.2 Dihydrolipoamide Dehydrogenase

Assays were performed according to the method of Jackman *et al.* (1990) at 30°C in a final vol of 0.7 ml. The assay mixture contained 50 mM potassium phosphate buffer, pH 7.6, 3 mM NAD^+, 1 mM $MgCl_2$, 0.2 mM ThDP and 20 μg of dihydrolipoamide (for preparation see 2.2.10.3). The reaction was initiated by addition of enzyme and activity followed by monitoring NADH formation at 340 nm. All enzymatic data presented in this thesis represent an average of duplicate determinations differing by less than 5%.

2.2.10.3 Preparation of Dihydrolipoamide

Dihydrolipoamide was prepared according to the method described by Reed *et al.* (1958). 400 mg of lipoamide (thio-octic acid amide) was dissolved in 80% (v/v) methanol at 4°C. 400 mg of sodium borohydride was dissolved in 2 ml of distilled water, also at 4°C, and then added to the lipoamide/methanol. The solution was stirred at room temperature until it became clear (approx. 45 min). The pH of the solution was then lowered to 2.0 with 0.25 M HCl and extracted with 40 ml of chloroform. The lower layer containing the dihydrolipoamide was removed and dried down under N_2. This material was re-dissolved by stirring on a hot plate in toluene/heptane (2.5:1.0). After drying down once more under N_2 until a minimal vol of liquid remained, the dihydrolipoamide was allowed to crystallise naturally and stored at -20°C. For use, approx. 20.7 mg of dihydrolipoamide was dissolved in 1 ml of ethanol.

2.2.10.4 Reconstitution Studies on Bovine Heart OGDC

Reconstitution of bovine heart E1/E2 core with the E3 component was carried out in Eppendorf tubes by incubating increasing amounts of E3 (0-50 μg) with a fixed amount of E1/E2 core (7 μg) at 30°C. The final vol of each reconstitution was kept constant with the addition of 50 mM potassium phosphate buffer, pH 7.6, 1 mM $MgCl_2$ and 0.2 mM ThDP. After 12 min a sample (equivalent to 5 μg of E1/E2) was added to an assay cuvette containing solutions A, B and C required for the assay of OGDC activity, performed as in section 2.2.10.1.

The determination of the time course of reconstitution was carried out in one Eppendorf tube by adding 12 x the assay amount of E1/E2 (60 μg) and a saturating amount of E3. Aliquots equivalent to 5 μg of E1/E2 were removed at specified intervals after incubation and assayed for activity.

2.2.11 Media for the Growth of Bacteria and Bacteriophage

Standard growth media, as described by Sambrook *et al.* (1989), were used throughout this work. They were sterilised by autoclaving for 20 min at 15 psi, or if heat labile, were filter sterilised.

LB medium: 1% (w/v) bactotryptone, 0.5% (w/v) yeast extract, 0.5% (w/v) NaCl, pH 7.5 (adjusted with NaOH).

LB Agar: As LB medium plus the addition of 15 g of agar per litre.

SOC medium: 1% (w/v) bactotryptone, 0.5% (w/v) yeast extract, 0.5% (w/v) NaCl, 10 mM $MgSO_4.7H_2O$, 10 mM $MgCl_2$, pH 7.0. Filter sterilised glucose solution (1 M) to a final concentration of 20 mM added after autoclaving.

NZY Bottom Agar: 1% (w/v) casein hydrolysate, 0.5% (w/v) NaCl, 0.5% (w/v) yeast extract, 0.2% (w/v) $MgSO_4.7H_2O$, pH 7.5, 15 g of agar per litre.

NZY Top Agarose: 1% (w/v) casein hydrolysate, 0.5% (w/v) NaCl, 0.5% (w/v) yeast extract, 0.2% (w/v) $MgSO_4.7H_2O$, pH 7.5, 7 g agarose per litre.

Ampicillin and kanamycin (Sigma) were made up at 50 $mg.ml^{-1}$ in water and filter sterilised. They were used at a concentration of 0.075 $mg.ml^{-1}$. Carbenicillin (Sigma) was used in place of ampicillin in most cases as it is a more stable analogue. It was used at a concentration of 0.1 $mg.ml^{-1}$. Tetracycline was made up at a concentration of 10 $mg.ml^{-1}$ in ethanol and used at a concentration of 0.012 $mg.ml^{-1}$. The long term storage of the antibiotics was at -20°C.

Escherichia coli (*E. coli*) bacterial cultures were grown at 37°C in Luria-Bertani media (LB) supplemented with the appropriate antibiotics in an orbital shaker at 225 rpm. Strains for short term storage were maintained on LB plates at 4°C and those for long term storage were mixed with glycerol to a final concentration of 15% (v/v) and stored at -70°C. Cultures used for the preparation of competent cells by the Hanahan method (2.2.27.1) were grown in SOC medium.

Bacteria required to produce a lawn for subsequent bacteriophage infection were grown in LB supplemented with 0.2% (w/v) maltose and 10 mM $MgSO_4$. Bacteriophage were stored as phage suspensions in SM buffer (100 mM NaCl, 10 mM $MgSO_4.H_2O$, 50 mM Tris-HCl at pH 7.5 and 0.01% (w/v) gelatine) supplemented with 0.01% (v/v) chloroform at 4°C.

2.2.12 Storage of DNA

DNA vectors, recombinants and oligonucleotides were routinely stored in TE buffer (10 mM Tris-HCl, pH 8.0 and 1 mM EDTA) at –20°C. DNA stored in this way remains stable for several years. Long-term storage of DNA was under 100% ethanol at –70°C.

2.2.13 Phenol/Chloroform Extraction

Phenol, equilibrated with TE buffer (10 mM Tris-HCl, pH 8.0, 1 mM EDTA) and containing 8-hydroxyquinoline at 0.1% (w/v) as an antioxidant, was mixed with chloroform/isoamylalcohol (24:1) at a ratio of 1:1. Extraction of an aqueous DNA solution was carried out by adding an equal vol of phenol/chloroform and vortexing the mixture. The aqueous layer was collected after a 5 min spin in a microfuge and the organic layer was re-extracted with an equal vol of TE buffer. The aqueous layers were pooled and the extraction was repeated once with water saturated chloroform.

2.2.14 Ethanol or Isopropanol Precipitation of DNA

A salt solution was added to an aqueous solution of DNA (either Na-acetate, pH 5.2 to 0.3 M or NaCl to 0.2 M) and the DNA precipitated with 2.5 vol of 98% ethanol (v/v) or 0.7 vol of isopropanol at -70°C for 15 min. The precipitate was recovered by centrifuging at 10 000 g for 15 min at 4°C and then washed with 70% (v/v) ethanol to remove any salt that may have co-precipitated.

2.2.15 Restriction Mapping

Restriction digests were routinely carried out for 1-5 h at 37°C (or the temperature recommended by the enzyme supplier) in 20-100 µl aliquots which contained a suitable amount of DNA, 1-10 units per µg DNA of the appropriate restriction enzyme and 1x restriction buffer supplied with the enzyme. One unit of restriction enzyme is defined as the amount of enzyme required to digest 1 µg of DNA to completion in one hour. In practice a several-fold excess was usually added to digest the DNA. Genomic DNA was incubated overnight using 8 units of restriction enzyme per µg of DNA. The completion of the digest was monitored by electrophoresis of an aliquot through an agarose gel.

2.2.16 DNA Electrophoresis in Agarose Gels

Agarose gels were routinely used to separate DNA fragments in the size range 0.1-20 kb. Slab gels were prepared by dissolving agarose (BRL ultraPURE) in TBE buffer, pH 8.3 (89 mM Tris, 89 mM boric acid and 2 mM EDTA) at the desired concentration (0.7-1.2% [w/v]). TAE pH 8.0 (40 mM Tris, 20 mM acetic acid and 1 mM EDTA) was the buffer of choice when gel purification of the DNA was required. Ethidium bromide was added at a concentration of 0.1 $\mu g.ml^{-1}$ to allow visualisation of the DNA in the gel. Samples were prepared for loading by the addition of 0.2 vol of 6x loading buffer (0.25% (w/v) bromophenol blue, 40% (w/v) sucrose and 50 mM EDTA). The gel was placed in a tank containing TBE or TAE buffer and the samples were loaded into the submerged wells. Electrophoresis was carried out at 2-5 volts/cm. DNA was visualised under 320 nm UV light and photographed with a UVP imager.

For DNA in the range of 50-400 bp, 12 % (w/v) non-denaturing TBE-polyacrylamide mini-gels were made using the protocol in section 2.2.5. The SDS was eliminated from the mixture and 10x TBE was used as the buffer component.

2.2.17 Electrophoresis of RNA using Formaldehyde Gels

A 1% (w/v) formaldehyde agarose gel was made by adding 1 g of agarose and 10 ml 10x MOPS buffer (0.2 M MOPS, pH 7.0, 0.5 M sodium acetate, pH 7.0 and 0.5 M EDTA) to 73 ml of water. After the agarose had dissolved by heating in a microwave it was cooled to 55°C and 17 ml of formaldehyde was added. The gel was cast in a fume hood, allowed to set for 30 min and submerged in 1x MOPS electrophoresis buffer. RNA samples were prepared (10-30 μg RNA, 12.5 μl of deionised formamide, 2.5 μl 10x MOPS, 4 μl formaldehyde and water to 25 μl), denatured at 65°C for 5 min and kept on ice. They were mixed with loading dye (2.2.16) and loaded into the wells. Electrophoresis was carried out at 80 - 100 volts until the bromophenol blue dye was 1 cm from the bottom of the gel. The gel was directly processed for Northern blotting or was stained with ethidium bromide (0.5 $\mu g.ml^{-1}$ in 10 mM ammonium acetate) for 30 min followed by destaining in sterile water.

2.2.18 Recovery of DNA from Low Melting Point Agarose

Low melting point agarose gels were prepared in the same way as standard agarose gels (using TAE buffer [section 2.2.16]) except that they were run in the cold room. The desired band was excised from the gel with a scalpel and transferred to an Eppendorf tube. Three vol of TE buffer were added to the agarose gel slice and the sample was placed at 65°C for 5 min to melt the agarose. The sample was extracted once with phenol equilibrated with TE as described in section 2.2.13. The aqueous phase was re-extracted with phenol/chloroform and twice with water saturated chloroform before it was ethanol precipitated.

The Geneclean II Kit (BIO101 Inc. USA) was used as a more efficient alternative to phenol/chloroform extraction. The protocol in the manufacturer's instruction manual was followed. This method gives clean DNA for ligations and has a working range of between 1-5 μg.

2.2.19 Recovery of DNA from Non-Denaturing Polyacrylamide Gels

The 'crush and soak' method was used for the purification of low molecular weight PCR products from polyacrylamide gels. The PCR reaction was cleaned by phenol/chloroform extraction and ethanol precipitated. The entire sample was separated on a 12% TBE-polyacrylamide gel (section 2.2.16) and the bands visualised with ethidium bromide. The band of interest was cut out of the gel, transferred to a microfuge tube and crushed using a yellow tip. The crushed gel was rotated at 37°C for 4 h (room temperature for 10-16 h) in 1-2 vol of elution buffer (0.5 M ammonium acetate, 10 mM magnesium acetate, 1 mM EDTA, pH 8.0 and 0.1% (w/v) SDS). The gel pieces were pelleted by centrifugation at 12 000 rpm for 5 min in a microfuge. Remaining gel pieces were removed by spinning the sample through glass wool and the DNA was then precipitated overnight at -20°C with 10 mM $MgCl_2$ and 3 vol of 100% ethanol. The sample was re-precipitated with 0.3 M sodium acetate and 2.5 vol of 100% ethanol. The final pellet was dissolved in 10 µl of water or TE buffer.

2.2.20 Small Scale Preparation of Plasmid DNA

A 3 ml overnight culture was prepared from a single colony of transformed bacteria in LB medium supplemented with the appropriate antibiotics. The bacterial suspension (1.5 ml) was transferred to a microfuge tube and the cells pelleted at 5000 rpm for 5 min at 4°C. The pellet was washed in 150 µl of cold STE (0.1 M NaCl, 10 mM Tris-HCl, 1 mM EDTA pH 8.0) followed by resuspension in 38 µl cold solution 1 (50 mM glucose, 25 mM Tris-HCl, pH 8.0 and 1 mM EDTA) and 75 µl solution 2 (0.2 M NaOH, 1% (w/v) SDS). After a 10 min incubation on ice 110 µl of cold solution 3 (4 vol of 5 M potassium acetate, 0.6 vol acetic acid and 0.4 vol distilled H_2O) was added and incubated on ice for a further 10 min. The bacterial debris was pelleted in a microfuge at high speed for 10 min at 4°C. An equal vol of isopropanol was added to the supernatant and plasmid DNA allowed to precipitate at room temperature for 10 min. The

precipitate was pelleted in a microfuge at high speed for 5 min, dried and resuspended in 50 µl TE-RNase A (1 $\mu g.\mu l^{-1}$). Following a 15 min incubation at 37°C, the RNA free plasmid DNA was recovered by ethanol precipitation.

2.2.21 Large Scale Preparation of Plasmid DNA

The Qiagen Midi Prep Kit (Qiagen GmbH, Germany) was used to obtain up to 150 µg of plasmid DNA from a 30 ml bacterial culture. Manufacturer's instructions were followed carefully to purify plasmid DNA suitable for ligations and sequencing.

Alternatively, the same method described in section 2.2.20 could be used; however, the volumes of all the solutions were scaled up to allow for an overnight transformed bacterial culture of 130 ml in LB supplemented with antibiotics. The cells were harvested by centrifugation for 15 min at 6000 rpm in a Beckman J2-21 centrifuge using a JA14 rotor and washed in 13 ml cold STE (0.1 M NaCl, 1 mM EDTA, 10 mM Tris-HCl, pH 8.0). Solution 1 (section 2.2.20) contained 5 mg lysozyme to aid the breakdown of cell walls. Prior to ethanol precipitation the plasmid DNA was cleaned by phenol/chloroform extraction (section 2.2.13). Routinely 130-150 µg of DNA were recovered.

2.2.22 The Polymerase Chain Reaction (PCR)

The polymerase chain reaction is a technique for the *in vitro* amplification of specific DNA sequences by the simultaneous primer extension of complementary strands of DNA. Standard PCR amplification reactions were set up in a total vol of 50 µl. The reaction mix contained: 5 µl of Promega 10x *Taq* DNA polymerase buffer (500 mM KCl, 100 mM Tris-HCl, pH 9.0, 100 mM NaCl, 0.1 mM EDTA, 5 mM DTT, 1.0% (v/v) Triton X-100), 0.2 mM each dATP/dGTP/dCTP/dTTP, 2 mM $MgCl_2$, 1 unit *Taq* DNA polymerase, 10 pmol each of the two oligonucleotide primers used (or 1000 pmol of degenerate primers) and a suitable amount of target DNA. Optimum $MgCl_2$ concentrations can vary for different primer pairs between

0.5 mM and 2.5 mM; however the lowest possible optimum concentration was used as this led to increased fidelity of the *Taq* enzyme (Brown, 1992)].

Pfu DNA polymerase (Stratagene) was used where proof-reading and high fidelity were required. The basic components of the reaction mix were: 0.6 mM nucleotide mix, 10 pmol primers, 1x *Pfu* buffer and 2.5 units of either native or cloned *Pfu* DNA polymerase. The 'hot start' technique was used with *Pfu* PCR reactions. Typically all reactants, except the polymerase, were mixed and overlaid with mineral oil. After the tubes had stabilised at 95°C for 1 min, the polymerase was added below the mineral oil and normal cycling continued.

The reaction was overlaid with 100 µl of mineral oil to prevent evaporation during heating and the reaction carried out in a Techne PHC-3 thermal cycler. The exact conditions used for amplification varied with the nature and length of the target sequence and the primers. The stages involved in the PCR were an initial denat-uration at 95°C, followed by repeated cycles of denaturation (94°C), annealing of primers (50-60°C), extension from primer (65°C). The cycles were followed by a final extension time at 72°C. Finally the reactions were cooled to 4°C and the PCR amplified products analysed on agarose or polyacrylamide gels (for products in the range of 50-200 bp).

PCR products to be purified and cloned were extracted with phenol /chloroform (section 2.2.13) before being ethanol precipitated (section 2.2.14). The products were gel purified using the Gene Clean kit, described in section 2.2.18, following restriction digestion with the appropriate enzymes.

2.2.23 First-Strand cDNA Synthesis and PCR Amplification using Degenerate and Nested Primers

2.2.23.1 First-Strand cDNA Synthesis from Plant RNA

The First-Strand cDNA Synthesis Kit was obtained from Pharmacia Biotech to generate full-length first-strand cDNA from an RNA template, catalysed

by CLONED FPLC*pure* Moloney Murine Leukemia Virus (M-MuLV) reverse transcriptase. The source of RNA used for this work was potato leaf RNA.

The RNA sample (5 μg) was made up to a total vol of 20 μl with RNAse-free water, heat denatured at 65°C for 10 min and then chilled on ice. The *Not*I-d(T)$_{18}$ primer was diluted 25x in RNAse-free water. The following components were combined in a sterile tube: 11 μl of bulk first-strand reaction mix (containing the reverse transcriptase and dNTP mix), 1 μl of DTT, 1 μl of the diluted *Not*I-d(T)$_{18}$ primer (0.2 μg) and the heat denatured RNA. The reaction was allowed to proceed at 37°C for 1 h.

2.2.23.2 Second-Strand Synthesis and Amplification of DNA by PCR

The resulting RNA:cDNA heteroduplex from the first-strand reaction (section) was used directly for second-strand synthesis by PCR. Degenerate primers or gene specific upstream primers, used in conjunction with the downstream *Not*I-d(T)$_{18}$ primer, were designed to amplify the cDNA encoding potato mitochondrial E3. A portion of the first-strand reaction (5 μl) was heated at 90°C for 5 min to denature the RNA:cDNA duplex and inactivate the reverse transcriptase and then chilled on ice. A PCR amplification reaction was set up as described in section 2.2.22 using 5 μl first-strand reaction. Cycling parameters used for the PCR amplification of the first-strand cDNA are detailed in chapter 5.

2.2.24 Quantification of Nucleic Acids

2.2.24.1 Absorbance at 260 nm

The absorbance of various dilutions of DNA samples was measured using 1 ml quartz cuvettes in a spectrophotometer at 260 nm. An absorbance unit of 1 corresponds to approx. 50 μg.ml^{-1} for double-stranded DNA, 40 μg.ml^{-1} for single-stranded DNA and RNA and 30 μg.ml^{-1} for oligonucleotides (Sambrook *et al.*, 1989). For a quantitative assay of more dilute solutions, the GIBCO DNA Mass Ladder was used. It consisted of 6 blunt ended fragments ranging in size

from 100-2000 bp. Electrophoresis of 2 μl of DNA Mass Ladder resulted in bands containing 100, 60, 40, 20 and 5 ng of DNA.

2.2.24.2 Fluorometry

Accurate measurements of DNA concentrations, down to 10 $ng.ml^{-1}$, were determined using a TKO 100 DNA minifluorometer (Hoefer Scientific Instruments, USA). DNA solutions were added to 2.0 ml of a working dye solution (see below). Aliquots of the working dye solution, of final vol 100 ml, were replaced daily and stored in the dark at room temperature. Each aliquot consisted of: 90 ml double-distilled, filter sterilised water (to remove particulate contaminants), 10 ml TNE, 10x (100 mM Tris-HCl pH 8.0, 10 mM EDTA, 1.0 M NaCl; adjusted to pH 7.4 with conc. HCl) and Hoechst H33258 dye solution to a final concentration of 100 $ng.ml^{-1}$. Readings were performed as described in the manufacturer's instructions. Serial dilutions of calf thymus DNA, from 10–500 $ng.ml^{-1}$ final concentration, were used to construct standard curves from which the DNA concentration of a fixed vol of sample DNA was determined.

2.2.25 Dephosphorylation of Plasmid DNA

In cases where only one enzyme was used to cut the vector, the 5′-phosphate groups of the linear plasmid were removed by calf intestinal alkaline phosphatase (CIP) to avoid self-ligation of the vector. The reaction mix containing 1 μg of linearised vector DNA, CIP buffer (1 mM $MgCl_2$, 1 mM $ZnCl_2$, 1 mM spermidine and 50 mM Tris-HCl, pH 8.4) and 1 unit of calf intestinal alkaline phosphatase, in a total vol of 50 μl, was incubated at 37°C for 30 min. SDS and EDTA were added to final concentrations of 0.5% (w/v) and 5 mM respectively. Proteinase K was added to a final concentration of 100 $\mu g.ml^{-1}$ and incubation continued for 30 min at 37°C. The reaction was subjected to a phenol and phenol/chloroform extraction and the DNA recovered by ethanol precipitation.

2.2.26 Ligations

Routinely, 200 ng of linearised vector DNA were used in a ligation in a total vol of 10 μl. The vector:insert ratio was 1:3. 1 μl of 10x ligation buffer (50 mM $MgCl_2$, 50 mM DTT, 500 $\mu g.ml^{-1}$ BSA and 200 mM Tris-HCl, pH 7.4) and 1 μl of 10 mM ATP were added to the DNA and the ligation initiated by the addition of 1 Weiss unit of bacteriophage T4 DNA ligase. The sample was incubated for 8-16 h at 16°C.

2.2.27 Preparation and Transformation of Competent Cells

2.2.27.1 Hanahan Method

E. coli XL1Blue and DH5α cells were made competent by the method developed by Hanahan (1983) as described by Sambrook *et al.* (1989). This procedure allowed the preparation of highly competent cells ($1x10^7$-10^8 transformants μg^{-1} of plasmid DNA) which were stored frozen in 100μl aliquots for future use.

To transform the cells, 100 μl were allowed to thaw on ice and plasmid DNA (no more than 50 ng) was added. After a 30 min incubation on ice, the tubes were transferred to a 42°C water bath, incubated for 90 s and rapidly transferred to an ice bath. After 5 min, 800 μl of SOC media (section 2.2.11) was added and incubated for 45 min at 37°C in an orbital shaker to allow the bacteria to express the antibiotic resistance. LB agar plates, containing the appropriate antibiotic, were prepared for colour selection by spreading 4 μl of filter sterilised, 20% (w/v in water) IPTG on each one. The plates were allowed to dry and then spread with 40 μl of a 2% solution (w/v in dimethyl formamide) of the chromogenic substrate, X-gal. Dilutions of the bacterial suspension were made and spread evenly on the LB plates.

2.2.27.2 Calcium Chloride

The *E. coli* strains BL21(DE3) and BL21(DE3)pLysS, used for the expression of heterologous protein, were made competent by the $CaCl_2$ method.

These competent cells were usually made freshly on the day. An overnight culture was used to inoculate 30 ml of LB medium and cells were allowed to grow to a density of 0.3 (A_{600}). The pelleted cells were resuspended in half the volume with ice cold $CaCl_2$ (100 mM), incubated on ice for 20 min and centrifuged again at 3000 g for 3 min. The final pellet was resuspended in one tenth of the original volume with ice cold $CaCl_2$. The cells were now ready to be transformed.

Plasmid DNA was diluted to a concentration of 2 ng.μl^{-1} in TEN buffer (50 mM Tris-HCl, pH 7.5, 1 mM EDTA, 50 mM NaCl). For transformation, 100 ng of DNA was added to 100 µl of cells and incubated on ice for 30 min. The cells were heat shocked at 37°C for 90 s and then returned back to incubate on ice for 30 min. LB medium (1 ml) was added followed by incubation at 37°C in an orbital shaker for 45 min. The transformed cells were resuspended in a final vol of 200 µl of LB and 10-30 µl aliquots were plated on LB kanamycin plates, prepared with IPTG and X-gal as in section 2.2.27.1.

2.2.28 Extraction and Purification of Total RNA from Plant Tissue

Potato leaves were grown as described in section 2.2.1. They were harvested, snap frozen in liquid nitrogen in batches of 5 g and stored at -80°C until required.

This procedure required elimination of ribonucleases from all glassware, pipette, tips and solutions. All chemicals were used exclusively for "RNA work", glassware and spatulas were rinsed with chloroform and baked overnight at 180°C, tips were soaked in 0.1% (v/v) diethyl pyrocarbonate (DEPC) and solutions were made with sterile Milli Q water previously treated with 0.1% (v/v) DEPC. DEPC is a non-specific inhibitor of ribonucleases; however it is a potent alkylating agent of RNA and must be destroyed completely before use by autoclaving.

Potato leaves (10 g) were ground to a fine powder in the presence of liquid nitrogen and transferred to a 50 ml Nunc disposable tube. This was followed by the addition of 1 ml of solution D for every gram of starting material (4 M guanidinium-isothiocyanate, 25 mM sodium citrate, pH 7.0, 0.5% (w/v) sarcosyl,

0.1 M 2-mercaptoethanol). The following solutions were added per gram of tissue in a sequential manner, shaking the tube after each addition: 0.2 $ml.g^{-1}$ of 2 M sodium acetate (pH 4), 1.0 ml g^{-1} of water saturated phenol and 0.4 $ml.g^{-1}$ of chloroform:iso-amyl alcohol (24:1). After vigorous shaking the material was kept in ice for 15 min. The mixture was centrifuged at 10 000g for 20 min at 4°C in screw cap Oak Ridge tubes and then 1 vol of isopropanol was added to the resulting supernatant. The precipitated RNA was spun down at 10 000 g for 25 min and the pellet was given at least three washes with 2.5 M sodium acetate (pH 5.2). The final pellet was washed with 70% ethanol and air dried. The RNA pellet was dissolved in sterile, DEPC treated water and stored at -80°C.

2.2.29 Transfer and Detection of Nucleic Acids

The following solutions were required for Northern and Southern blots:

20x SSC (pH 7.0) 3 M NaCl
0.3 M Na_3 citrate

20x SSPE (pH 7.4) 3.6 M NaCl
0.2 M sodium phosphate
0.02 M EDTA

100x Denhardt's 2% (w/v) BSA
2% (w/v) Ficoll
2% (w/v) PVP

10 $\mu g.ml^{-1}$ sonicated salmon sperm DNA: dissolve in water overnight with stirring and then sonicate the DNA to an average length of 400-800 bp as judged by agarose gel electrophoresis.

2.2.29.1 Radiolabelling of DNA Fragments

The Megaprime DNA labelling system from Amersham was used for the radioactive labelling of probes for hybridization. DNA required for labelling was diluted in water to a final concentration of 5 ng.μl^{-1}. A total of 5 μl (25ng) of DNA were denatured for 5 min in a boiling water bath in the presence of random sequence nonameric primers (5 μl). The DNA was cooled to room temperature and 10 μl of the labelling buffer (containing dGTP, dATP, dTTP, $MgCl_2$, 2-mercaptoethanol and Tris-HCl buffer, pH 7.5), 23 μl of water, 2 μl Klenow enzyme and 5 μl [α-^{32}P]dCTP (3 000 Ci/mmol) were added. The reaction was incubated at 37°C for 30 min.

Chromaspin-10 columns were used to separate the unincorporated free nucleotides from labelled DNA. Chromaspin columns were shaken to mix the resin and spun at 1700 g for 3 min to pack the resin. The 50 μl sample was applied to the centre of the bed and was spun at 1700 g for 4 min to collect the product.

2.2.29.2 Southern Blotting and Hybridisation

Alkali blotting was the method employed to transfer resolved DNA fragments from agarose gels to positively charged nylon membranes (Hybond-N+). After electrophoresis, the DNA was depurinated by placing the agarose gel in 0.25 M HCl for 10 min after the bromophenol blue dye had changed colour (blue to bright yellow). This treatment was essential for the transfer of DNA fragments larger than 5 kb. The gel was rinsed in distilled water and the DNA denatured by soaking the gel in 0.4 M NaOH for 30 min. A capillary blot was set up as described by Sambrook *et al.* (1989) using 0.4 M NaOH as the transfer buffer and the Hybond-N+ membrane. Transfer was allowed to proceed overnight after which time the blot was disassembled. The wells were marked on the membrane with a pencil and it was washed with 2x SSC. The membrane could then be stored dry, between clean Whatman 3 MM paper, or it could be processed for hybridisation.

A pre-hybridisation mix was made up in a final vol of 25 ml as follows: 5x SSPE, 5x Denhardt's solution, 0.5% (w/v) SDS and 100 $\mu g.ml^{-1}$ heat

denatured, sonicated salmon sperm DNA. The membrane was pre-hybridised in a Mini 10 hybridisation oven (Hybaid LTD, UK) at 65°C for a minimum of 2 h. The heat denatured probe (1-5 pg.ml^{-1} hybridisation buffer, 10^8-10^9 cpm.μg^{-1}) was snap-cooled on ice and then added to the pre-hybridisation mix and further incubated for 12-16 h at 65°C. The membrane was washed once in 2x SSC, 0.1% (w/v) SDS for 15 min at 65°C and subsequently once in 1x SSPE, 0.1% (w/v) SDS for 20 min and finally in 0.1x SSPE, 0.1% (w/v) SDS for 15 min. The membrane was then dried and autoradiographed for the desired time. If the membrane was to be re-washed or stripped and re-hybridized then it was kept damp at all times.

2.2.29.3 Northern Blotting and Hybridisation

RNA from a formaldehyde gel was transferred by the capillary blotting method using 20x SSPE as the transfer buffer and Hybond-N membranes. The gel was not pre-stained with ethidium bromide as this would reduce the efficiency of RNA transfer. Following an overnight transfer, blots were washed gently with 2x SSPE and air-dried. RNA was fixed to the membrane by baking at 80°C for 2 h. The hybridisation and washing protocols were identical to the southern blot.

2.2.30 Screening a Lambda ZAP II cDNA Library

2.2.30.1 Plating and Titering the Library

LB medium (25 ml), supplemented with 0.2% (w/v) maltose (to induce the *lam*B gene coding for the bacteriophage λ receptor) and 10 mM $MgSO_4$ was innoculated with a single colony of freshly plated *E. coli* XL1Blue MRF′. The culture was grown at 37°C at 225 rpm in an orbital shaker until it had an A_{600} of 0.5. The bacterial culture was pelleted at 2000 rpm for 10 min and the cells were resuspended and diluted to a density of 1.0 at A_{600} in 10 mM $MgSO_4$. These cells were stored at 4°C overnight.

A serial dilution of the library was made with SM buffer (100 mM NaCl, 10 mM $MgSO_4.7H_2O$, 50 mM Tris-HCl, pH 7.5 and 0.01% (w/v) gelatine) and 10 μl was mixed with 990 μl of the bacterial host cells. The cells were incubated at

37°C for 20 min for the phage to infect the host. The entire content of the tube was mixed with 8 ml of top agar at 42°C and plated on two day old bottom agar plates (140 mm diameter), previously equilibrated to 37°C. The inverted plate was incubated at 37°C for 8 h or until the plaques were clearly visible, but not confluent and the titre of the library was calculated as follows:

$$\frac{\text{total no of plaques} \times \text{Dilution factor} \times 1000}{\text{volume of phage plated } (10\mu l)} = \text{number of plaque forming units per ml } (pfu.ml^{-1})$$

2.2.30.2 Plaque Lifts

Plaques were prepared as outlined in 2.2.30.1. For the primary screen 500 000 phage, at a density of 60 000 plaques per plate, leading to confluent growth. For the secondary and tertiary screens each positively hybridising phage sample was plated at a density of 50-500 plaques so that individual plaques could be distinguished. Plates were chilled at 4°C for 60 min to allow the top agar to harden before plaque lifts were attempted. Meanwhile, the appropriate number of charged nylon membranes (Hybond-N+) were labelled. A membrane was laid carefully onto each plate from the centre outward, so that it made direct contact with the plate without smearing the plaques. Several asymmetric points were marked with a needle to record the orientation of the membrane on the plate and after 60 s it was peeled off using blunt forceps. When duplicate lifts were made, the second membrane was left on the plate for a 2 min duration.

Each membrane was placed, DNA side up, for 4 min on two sheets of Whatmann 3 MM paper soaked in denaturation solution (0.5 M NaOH, 1.5 M NaCl). The membranes were drained of excess denaturant and transferred to two sheets of 3 MM paper soaked in neutralisation solution (0.5 M Tris-HCl, pH 7.5 and 1.5 M NaCl) for a further 4 min. Finally, the membranes were rinsed in 2x SSC for 5 min and left to dry on 3 MM paper. The membranes were then alkali fixed with 0.05 M NaOH for 20 min, rinsed with 5x SSPE and then air-dried.

The membranes were hybridised with a [^{32}P]-labelled homologous probe as described in section 2.2.29.1 and 2. For homologous screening the protocol from Sambrook *et al.* (1989) was followed. Just before autoradiography the reference marks on the membrane were spotted with radioactive ink (500 μl ink and 1 μl [^{32}P]-dCTP).

2.2.30.3 Selection of Bacteriophage Lambda Plaques

The pattern of spots produced on the film, by hybridisation of plaque DNA with a probe, were aligned with their corresponding position on the plate with the help of the reference marks on the membrane and plate. Plaques of interest were removed from the plate as plugs using a cut 1 ml disposable tip. The agar plug was transferred to a 1.5 ml microfuge tube which contained 1 ml of SM buffer and 50 μl of chloroform, vortexed briefly and left at room temperature for 1 h to allow the phage to diffuse out into the solution. The phage suspension was titred as described in section 2.2.30.1 and the process of screening was repeated with a lower plaque density per plate until plaque purity was attained. This typically took three rounds of screening.

2.2.30.4 Preparation of Plate Lysate Stocks

Once a single hybridising plaque had been isolated and titred, 6×10^4 pfu were used to produce confluent lysis of a bacterial lawn grown on a 140 mm diameter plate. Phage were eluted from the agar by the addition of 8 ml of SM buffer to the plate, which was left shaking gently at 4°C overnight.

The next day the bacterial suspension was recovered into a sterile polypropylene tube and the plate rinsed with 2 ml of SM buffer. A volume of 0.3% (v/v) chloroform was added to the suspension, vortexed and incubated at room temperature for 15 min. Bacterial cell debris was removed from the suspension by centrifugation at 3000 rpm in a JA20 rotor (Beckman) for 10 min. The supernatant was transferred to a fresh tube and chloroform was added to 0.3% (v/v) to prevent any bacterial growth. Phage lysates were stored at 4°C.

2.2.30.5 Purification of Recombinant Bacteriophage DNA

Wizard Lambda DNA Purification System from Promega was used to prepare recombinant phage DNA. The protocol was followed as described by the manufacturer.

2.2.31 *In Vivo* Excision of the pBluescript Phagemid from the Lambda ZAP II Vector

The potato leaf cDNA library was made using a lambda insertion type cDNA cloning vector, Lambda ZAP II. This vector contains the phagemid, pBluescript SK- (Short *et al.* 1988) which can be excised by the M13 based helper phage, VCSM13 (Stratagene).

The following components were combined in an Eppendrof tube: 200 µl of XL1 Blue MRF′ cells (cultured in LB supplemented with 0.2% (w/v) maltose and 10 mM $MgSO_4$ and stored in 10 mM $MgSO_4$) at an A_{600} of 1.0, 250 µl phage stock (>$1x10^5$ phage particles in SM buffer) and 1 µl of the VCSM13 helper phage (>$1x10^6$ pfu.μl^{-1}). The tube was incubated at 37°C for 15 min, 3 ml of LB was added and incubated for a further 3 h with shaking and during this time the phagemid containing the insert was excised in the presence of the gene II protein provided by the helper phage. The single-stranded pBluescript phagemids were circularised and then packaged into filamentous phage particles with the proteins provided by the helper phage. Heat treatment at 65°C was carried out for 20 min to inactivate the parent lambda phage vector, kill the bacterial cells and release the phage particles into the supernatant after which the debris was pelleted at 1000 g for 15 min. The supernatant, containing the excised pBluescript phagemid packaged as filamentous phage particles, was decanted into a fresh Eppendorf tube and was stable at 4°C for 1-2 months.

E. coli SOLR cells (Stratagene) were cultured in LB and stored for a maximum of 24 h in 10 mM $MgSO_4$ at an A_{600} of 1.0. To rescue the excised phagemids, 200 µl of the bacterial cells were added to 100 µl and 10 µl respectively of the phage supernatant. The samples were incubated at 37°C for 15 min before

being plated on LB-ampicillin plates (200 μl per 140 mm plate). The plates were incubated at 37°C, overnight and colonies appearing the next day contained the pBluescript double-stranded plasmid. Single colonies were suspended in 100 μl of sterile water and 5 μl was used for screening by PCR amplification. The rest of the sample was used to inoculate a 3 ml overnight culture for the preparation of glycerol stocks and for mini preparations of the cloned DNA.

2.2.32 Generation of Clones for Sequencing by Directed Deletions

The Erase-A-Base system from Promega was routinely used for generating nested deletions from one end of the cloned DNA. The system is based on the procedure developed by Henikoff *et al.* (1984) in which Exonuclease III (Exo III) is used to specifically digest insert DNA from a 5′ protruding or blunt restriction site. The adjacent sequencing primer binding site is protected by a 4-base 3′ overhang restriction site as it is resistant to digestion by Exo III.

2.2.32.1 Restriction Digestion of Plasmid DNA

Typically 5-7 μg of the plasmid was restricted with two restriction enzymes from one end of the insert which were known not to cut within the cloned insert. One enzyme produces a 5′ overhang and the other a 3′ overhang. The 3′ overhang must be toward the vector, thus protecting the primer binding site and the other enzyme must be close to the insert. Examples of appropriate enzymes are listed in the manual accompanying the kit. A reaction volume of 50 μl was always preferred with two enzymes having compatible reaction buffers and incubation temperatures. If the enzymes have different salt concentration requirements then the lower salt buffer was used first with its enzyme and then the salt concentration was adjusted for the second enzyme. After an incubation of 2-3 h the digestion of the plasmid was checked on a gel. The rest of the digest was isolated with phenol/chloroform and ethanol precipitated (section 2.2.13).

2.2.32.2 Exo III Digestion, Ligation and Transformation

The plasmid pellet was thoroughly resuspended in Exo III 1x buffer (66 mM Tris-HCl, pH 8.0 and 0.66 mM $MgCl_2$). In the mean time 7.5 µl of S1 nuclease mix was added to the appropriate numbers of tubes which were kept on ice. The DNA was warmed up to 37°C in a water bath and 500-600 units of Exo III was mixed rapidly. After a 20 s lag period, 2.5 µl of the sample was transferred into a S1 nuclease mix tube at 30 s intervals and mixed. When all the samples had been transferred, the S1 nuclease tubes were incubated at room temperature for 30 min. The nuclease reaction was stopped by the addition of 1 µl of S1 stop buffer to each tube. The S1 was heat inactivated at 70°C for 10 min. At this point the Exo III digestion was checked by electrophoresis of 2 µl of the above samples on a 1% agarose TBE gel. To each sample 1 µl of Klenow mix was added and the samples incubated for 5 min at 37°C. Then 1 µl of dNTP mix was added to each tube and incubation was continued for another 10 min. The DNA was then ligated by the addition of 40 µl of ligase mix to each tube and incubation at 16°C overnight. The following day, *E. coli* cells were transformed by the Hanahan method, as described in section 2.2.27.1.

2.2.33 Automated Dyedeoxy Terminator Cycle Sequencing

Automated sequencing was carried out on double-stranded plasmid template (Qiagen purified) using the Taq Dyedeoxy Terminator Cycle Sequencing kit (Applied Biosystems Inc, USA). A 20x reaction pre-mix contained 80 µl of 5x TACS buffer (400 mM Tris-HCl, 10 mM $MgCl_2$, 100 mM $(NH_4)_2SO_4$, pH 9.0), 20 µl dNTP mix, 20 µl of each of the Dyedeoxy A, T, G and C terminators and 10 µl AmpliTaq DNA polymerase. A typical sequencing reaction consisted of 1 µg of the plasmid DNA, 5 pmol of primer (M13 universal or reverse or gene specific primers), 9.5 µl of dyedeoxy reaction pre-mix and sterile water to a final vol of 20 µl. Each reaction mix was overlayed with mineral oil and the cycle sequencing reaction carried out in a Perkin-Elmer Cetus thermal cycler, model 480 with the

following cycling parameters: rapid thermal ramp to 96°C for 2 min, 25x [96°C for 30 sec, 50°C for 15 sec, 60°C for 4 min], soak at 4°C.

The products were purified by a modified phenol/chloroform extraction and ethanol precipitation, as detailed in the manual. The dried pellet was stored at -70°C until used.

Later sequencing reactions were carried out using the new ABI PRISM kit, which contained a modified AmpliTaq DNA polymerase. The reactions contained 0.5 μg plasmid DNA, 4-6 pmol of sequencing primer, 8 μl of the reaction mix (containing AmpliTaq, the four dye labelled dNTPs) and water to a final vol of 20 μl. The cycling parameters were exactly the same as above. As smaller concentrations of the dyes were used in comparison to the Taq Dyedeoxy Terminator Cycle Sequencing kit, only the ethanol precipitation step was required.

Sequagel-6 from National Diagnostics, USA was routinely used to cast the gel for DNA sequencing. The gel was pre-run with 1x TBE buffer (section 2.2.16). The samples were mixed with 3 μl of formamide and denatured at 90°C for 2 min. They were chilled on ice before being loaded on the gel.

Sequence data were analysed using the Genetic Computer Group (GCG), Wisconsin sequence analysis package, version 8.1.

2.2.34 Expression of Heterologous Proteins

The pET system (Novagen) was chosen to subclone and express eukaryotic genes in *E. coli* under the control of strong bacteriophage T7 transcription and translation signals.

2.2.34.1 Subcloning the Target Gene into an Expression Vector

The recombinant plasmids (pBluescript) were purified using a Qiagen Midi Prep kit (section 2.2.21) and digested with suitable restriction enzymes to cut out the inserts. The restriction sites had to be chosen with care so that the inserts could be subcloned into the expression vector in the sense orientation. Digested inserts were gel purified using the Gene Clean kit (section 2.2.18) and ligated to

either the pET28a, b or c translation vector, depending on the reading frame of expression relative to the *Bam*H1 site as described in section 2.2.26.

The ligated plasmids were initially cloned into a host not possessing the T7 RNA polymerase gene, thereby maintaining them in a transcriptionally silent form. The host strain chosen for transformation was *E. coli* DH5α, made competent by the Hanahan method (section 2.2.27.1). Cells were plated on LB supplemented with kanamycin (75 $\mu g.ml^{-1}$) and putative clones were isolated after an overnight incubation at 37°C. Each colony was cultured in 3 ml LB media with kanamycin and 1.5 ml used for a small scale plasmid preparation (section 2.2.20). The recombinant plasmids were analysed for the correct insert by digesting the vector with the restriction enzymes used for cloning. Positive clones were stored as glycerol stocks until required for expression.

2.2.34.2 Transformation of the Expression Vector to an Expression Host Strain

In order to be expressed, the recombinant plasmid had to be transferred to suitable expression hosts, *E. coli* BL21(DE3) and BL21(DE3)pLysS. They contain a chromosomal copy of the T7 RNA polymerase gene which is induced by the addition of isopropyl-β-D-thiogalactopyranoside (IPTG).

LB kanamycin plates were streaked with the transformed DH5α glycerol stock and a colony was used to inoculate 50 ml of LB media also supplemented with kanamycin. Plasmids, prepared using the Qiagen Midi Prep kit, were transformed into BL21(DE3) cells and BL21(DE)plysS made competent by the calcium chloride method (section 2.2.27.2). The transformed cells were finally resuspended in 200 μl of LB and 10, 20 and 30 μl of these cells were plated on LB kanamycin for BL21(DE3) cells or LB kanamycin/chloramphenicol for BL21(DE3)plysS. Separate colonies were transferred to LB plates, sectioned in a grid form, the following day and allowed to grow overnight at 37°C. The colonies were stored in this form at 4°C for up to a month.

2.2.34.3 Expression and Purification of Heterologous Protein

One colony on the gridded plate was touched briefly with a yellow tip and this was used to inoculate a mini culture of LB supplemented with kanamycin or kanamycin/chloamphenicol. Following overnight growth, 1 ml of this culture was inoculated into 100 ml LB with the appropriate antibiotic, and the cultures grown at 225 rpm to an A_{600} of about 0.5 at 30°C or 37°C. At this point the cells were in mid log phase and ready to be induced with 0.4 mM of freshly prepared IPTG.

Samples of 1 ml were removed at 0, 1, 2 and 3 h after induction (for analysis on SDS-PAGE gels) and their A_{600} was noted. The cells were pelleted in a microfuge, resuspended in 10 µl per 0.1 of an absorbance unit of Laemmli buffer containing 10 mM DTT and 5-20 µl were loaded on a 10% SDS-polyacrylamide gel as described in section 2.2.5.

After incubation for 3 h, the cultures were spun at 6000 rpm in a JA14 rotor (Beckman) for 15 min and resulting pellets were washed with STE (0.1 M NaCl, 1 mM EDTA, 10 mM Tris-HCl, pH 8.0). The final pellets were resuspended in 15 ml of 50 mM potassium phosphate buffer, pH 7.6 containing the protease inhibitors, benzamidine and leupeptin. The samples were subjected to 3-4 passages through a French press, keeping them on ice at all times. Any precipitate was spun down and the samples assayed for E3 activity (section 2.2.10.2). Untransformed BL21(DE3) cells were used as a control for the background activity of bacterial E3.

2.2.35 Molecular Mass Determination by Electrospray Mass Spectrometry

Protein samples required for mass spectrometry were washed to remove all salts by diluting 50-fold with HPLC grade water and reconcentrated using Centricon-10 centrifugal concentrators. Mass spectrometry was performed on a VG Platform quadruple mass spectrometer fitted with a pneumatically assisted electrospray (ionspray) source and controlled via the VG MassLynx software (VG Biotech Ltd., Altrincham, Cheshire, UK). Carrier solvent (1:1 (v/v)

acetonitrile/water, 0.2% formic acid) infusion was controlled at 10 $\mu l.ml^{-1}$ using a Harvard syringe pump (Harvard Apparatus, South Natic, MA, USA.). Protein samples were dissolved in carrier solvent at a concentration of 20 $pmol.\mu l^{-1}$, centrifuged at 5000 g for 2 min and then 10-20 μl samples injected directly into the carrier stream. MaxEnt deconvolution (Ferrige *et al.*, 1992) was applied for quantitative analysis of the raw data using a 1.0 Da peak width and 1.0 Da per channel resolution.

Chapter 3

Purification and M_r Analysis of Dihydrolipoamide Dehydrogenase (E3) Isoforms from Potato Tubers

3.1 Introduction

Dihydrolipoamide dehydrogenase (E3) has been purified from a variety of organisms and characterised at the biochemical and genetic level. Analysis of the plant 2-oxoacid dehydrogenase complexes and their component enzymes has lagged behind their bacterial and mammalian counterparts because of the low abundance of mitochondria per fresh weight of plant material (Randall & Miernyk, 1990a). Characterisation has concentrated on PDC and GDC, mainly from pea and broccoli. GDC is distantly related to the family of 2-oxoacid dehydrogenase complexes and has been especially well studied from pea leaf mitochondria (Bourguignon *et al.*, 1988; Walker & Oliver, 1986a). While GDC has been purified from a number of plant and animal sources, it is found at its highest level in mitochondria of C_3 plant leaves where it comprises about a third of the soluble matrix protein (Oliver *et al.*, 1990). The L-protein of GDC is a dihydrolipoamide dehydrogenase which has been shown to be a common constituent of GDC, PDC, and OGDC in pea by copy number analysis of the cloned gene from the mitochondria of pea leaves (Bourguignon *et al.*, 1992; Turner *et al.*, 1992). Further evidence to support this conclusion was provided by a comparison of the precise masses (obtained by electrospray mass spectrometry) of L-protein and PDC E3, also purified from the mitochondria of pea leaves, which were found to be identical (Bourguignon *et al.*, 1996).

In contrast to the detailed structural and mechanistic information available on GDC, the plant 2-oxoacid dehydrogenase complexes have yet to be fully characterised. Although the mammalian PDC complexes are confined to the mitochondrion, two distinct, spacially separate PDCs are found in plants, one located in mitochondria and the other in chloroplasts (Reid *et al.*, 1977; Williams & Randall, 1979; Rubin & Randall, 1977). The mitochondrial PDC serves as a primary entry point for carbon into the citric acid cycle and related metabolism. Chloroplast PDC, on the other hand, provides the acetyl-CoA and NADH required

for fatty acid and isoprenoid biosynthesis in photosynthetic tissues (Camp & Randall, 1985; Williams & Randall, 1979).

To date, it has not been possible to isolate intact OGDC from plants owing to its low abundance in comparison to PDC. Additionally, Poulsen and Wedding (1970) found that the E3 component readily dissociated from the complex during purification procedures. The only documented partial purification of OGDC was by Poulsen and Wedding (1970), who concentrated on the isolation of the E1/E2 subcomplex of OGDC from cauliflower florets. Intact OGDC activity was obtained by these authors following reconstitution with saturating levels of porcine heart E3. The third 2-oxoacid dehydrogenase complex, BCDC, has been detected in the peroxisomes of mung bean hypocotyls but it has not, as yet, been purified (Gerbling & Gerhardt, 1988; Gerbling & Gerhardt, 1989).

Information on the subunit composition of the plant complexes is limited. Immunological studies, carried out by Camp and Randall (1985) and Taylor *et al.* (1992), revealed differences in the banding pattern between chloroplastic and mitochondrial PDC indicating the possible presence of organelle-specific subunits. Indeed, a distinct pea chloroplastic E3 has been isolated recently (Conner *et al.*, 1996) and is the first component of plastidic PDC to be identified to date. The N-terminal sequence of this E3 was compared with that of the pea mitochondrial E3, reported by Bourguignon *et al.* (1996) and was found to exhibit only 25-40% identity at the amino acid level, suggesting that it is encoded by a separate gene.

The novel isolation of three isoforms of E3 from the mitochondria of potato tubers (*Solanum tuberosum*, cv. Maris Piper) was first reported by this laboratory (Fullerton *et al.*, 1996). The purification procedure included the use of an affinity column employing the E2/X core proteins from bovine heart PDC. Total E3 from a mitochondrial extract was found to bind to this column, eluting in a single peak of activity with high salt (1 M NaCl). This single peak of activity was then resolved into three peaks by Mono Q ion exchange chromatography in the ratio of 1:3:5. SDS-PAGE analysis of each peak revealed the presence of α_2, $\alpha\beta$ and β_2 dimers,

arising from the three possible combinations of two polypeptide chains with subunit M_r values of approx. 58 000 (α) and 56 000 (β). The presence of three E3 enzymes was not caused by differential aggregation of a single form of E3 because re-application of the separate peaks to the Mono Q column resulted in their elution at the same point in the gradient. Moreover, proteolytic degradation was eliminated as a possibility since the same three E3 isoenzymes were observed from a preparation carried out in the absence of protease inhibitors (Fullerton *et al.*, 1996). The tetraploid nature of *Solanum tuberosum* was found not to be a contributing factor to the presence of isoforms because three E3 isoenzymes were also obtained from a diploid strain of potatoes (*Solanum phoreta*). Additionally, the recent discovery of three E3 isoenzymes in barley leaves (*Hordeum vulgare* cv. Hart), another diploid strain, dispels the notion that this is a unique feature of potatoes (Fullerton *et al.*, 1996). However, further characterisation of mitochondrial E3s from other plants such as turnip (*Brassica campestris*), horseradish (*Armoracia rusticana*) and pea (*Pisum sativum*) indicate the presence of only one form of the enzyme in mitochondria (Fullerton *et al.*, 1996).

The heterodimeric nature of the peak 2 isoform is highly unusual as all E3s characterised to date are present as homodimers. The similarities between the α and β polypeptides were highlighted by the fact that both cross-reacted strongly with antisera raised to yeast E3 and that they exhibited identical N-terminal sequences. Notably, the N-termini were identical with that of the pea leaf mitochondrial E3 sequence (Bourguignon *et al.*, 1992; Turner *et al.*, 1992) with the exception of one amino acid at residue 11 where an isoleucine was replaced by a valine.

Although the three isoforms from potato tubers were purified to homogeneity, the yield in terms of protein was poor. Less than 0.2 mg of total E3 protein from 5 kg of potatoes (R. Fullerton, Ph.D Thesis, Glasgow University, 1995) was obtained after the affinity chromatography step, leading to only 0.03-0.06 mg of each isoform resolved by Mono Q resin. After resolution on the Mono Q column, approx. 0.5-2.5 units of each enzyme were obtained. Direct kinetic

experiments were feasible on this small amount of enzyme as the assay for E3 was highly sensitive but experiments requiring more protein, such as investigation into the complex specific roles of the isoenzymes or the determination of the M_r values of the α and β polypeptides by electrospray mass spectrometry, were impossible to carry out.

Clearly there was potential for improvement of the potato preparation in terms of final yield of E3. This chapter deals with attempts to increase the final yield of each of the E3 isoforms, for their further characterisation. Isolation of highly purified mitochondria generally led to 70-80% loss in E3 activity, prior to initiating the purification, owing to low yields of organelle recovery. Therefore, the rationale was to start with greater amounts of E3 by employing a crude mitochondrial fraction as the starting material. However, as the initial extract was to be obtained from a crude pellet and not highly pure mitochondria the purity of the E3 isoforms was expected to be compromised. Additionally, it was realised that scaling up the original procedure using crude mitochondrial pellets was likely to require amendments to the original purification scheme.

3.2 Results

3.2.1 Preparation of a Crude Organellar Extract from Potato Tubers

The method for the preparation of the tuber mitochondrial extract, described by R. Fullerton (Ph.D Thesis, Glasgow University, 1995), was modified to reduce the number of steps which were routinely responsible for huge losses of the E3 enzyme. These included purification of mitochondria from a crude pellet and the affinity purification step using the E2/X column mentioned in the introduction to this chapter. The modified method is described in detail in Materials and Methods (section 2.2.9.3).

Firstly, the homogenisation buffer was modified from 10 mM potassium phosphate (pH 7.6) to 50 mM MOPS (pH 7.6), to improve the buffering capacity of the medium and preserve the intactness of the organelles. Secondly, BSA (bovine serum albumin) was removed from all buffers as it tended to co-purify with the potato E3. The function of BSA was two fold: to bind free fatty acids which could have a detergent effect on mitochondria and to bind some alkaloids and polyphenols released from the vacuoles (e.g flavonoids and tannins). Two pilot purifications from potato tubers were carried out in the presence and absence of BSA, and the difference in the total number of units extracted from the initial organellar pellets and the overall yields of mitochondrial E3 were very similar. This indicated that BSA was not necessary for maintaining organellar integrity or protecting the E3 enzyme from alkaloids and polyphenols in this method.

The isolation of mitochondria, by centrifugation through self-generating silica gradients (Percoll), was eliminated to avoid the 70-90% loss of this organelle during the long isolation procedure. The organelles were ruptured to release their contents by the use of 0.1% (v/v) Triton X-100 as described in Materials and Methods (section 2.2.9.3). A higher concentration of Triton X-100 was avoided

Table 3.1

Purification of Potato Mitochondrial E3

E3 assays and protein concentrations were determind as described in 2.2.10.2 and 2.2.2. A unit of enzyme activity is defined as the amount of enzyme required for the formation of 1 μmol NADH/min.

Sample	Volume (ml)	$mg.ml^{-1}$	Total mg	Total units	Percent recovery	Specific activity (units/mg)	Purifi-cation
Initial Triton X-100 extract	211.0	2.6	548.6	715.0	100	1.3	1.0
Heat-treated extract	182.0	0.8	145.6	554.0	77	3.8	3.0
Dialysed extract (for DEAE-52)	198.0	0.7	138.6	569.0	79	4.1	3.2
DEAE-52 purified extract	50.0	0.2	10.0	477.0	67	47.7	36.6
Mono Q purified peak 1	0.36	3	1.2	60		50.0	-
Mono Q purified peak 2	0.4	6	2.5	185	43	74	-
Mono Q purified peak 3	0.38	5	1.9	65		34.2	-

since it is difficult to remove from the final, purified enzymes. Routinely, using crude organellar pellets, it has been possible to obtain 20 to 30-fold more E3 enzyme at this stage than the original method. The initial specific activity of the extract obtained from the pure mitochondrial pellet was 2.8 units/mg (R. Fullerton, Ph.D Thesis, Glasgow University, 1995) compared with 1.3 units/mg from the crude organellar pellet (Table 3.1), suggesting that the crude pellet was approx. half as pure as the mitochondrial pellet. This surprisingly small difference in the purity of the crude pellet and the pure mitochondrial pellet obtained by R. Fullerton (Ph.D Thesis, Glasgow University, 1995) can be explained by the fact that although there are more contaminating organelles in the crude pellet, it is greatly enriched for mitochondria since large scale disruption of this organelle on Percoll gradients was avoided.

Heat-treatment at 65°C was retained in the protocol as a purification step, removing heat labile proteins and releasing any heat stable E3 still bound to the complexes. However, the time of heat-treatment was increased to 30 min from the original 10 min because it allowed a 3-fold purification (Table 3.1) compared with a 1.5-fold purification at 10 min. Whereas R. Fullerton (Ph.D Thesis, Glasgow University, 1995) reported a loss of only 10% of the total E3, the modification resulted in an overall 20% loss. The crude, heat treated extract was bright yellow owing to the release of polyphenolic and other oxidising compounds from the potato tubers.

3.2.2 Elution of E3 from a DEAE-52 Ion Exchange Column

Further purification of the extract on the bovine E2/X affinity column, as described by R. Fullerton (Ph.D Thesis, Glasgow University, 1995), was not performed since the binding capacity of this column was low (approx. 15 units of total E3); efforts to scale up the columns were difficult owing to the large amounts of E2/X required and the expense of the CNBr-activated Sepharose 4B. A further disadvantage was that preparation of the column was time consuming. However, it was desirable to add a chromatography step after heat treatment of the extract since

direct application of the heat treated extract to the Mono Q column led to poor purification of the E3 isoforms. The method of choice was ion exchange chromatography using diethylaminoethyl cellulose (DEAE-52) since the scale-up of the columns was relatively cost effective. This chromatography step had to fulfil two more criteria: to remove the yellow polyphenolic compounds from the extract which tend to cause inactivation of enzymes and to concentrate the 200 ml extract to a manageable volume for Mono Q ion exchange chromatography. Use of Centriprep-30 concentrators was avoided because the concentration procedure led to significant loss in E3 activity.

Small scale experiments revealed that the best recoveries (up to 70-80%) were obtained from DEAE-52 columns (1 ml) eluting with 260 mM NaCl in 20 mM potassium phosphate buffer, pH 6.8. Larger DEAE-52 columns (55 ml bed vol) were prepared as described in Materials and Methods (section 2.2.9.4) and chromatography performed as described in Fig. 3.1. One peak of E3 activity was eluted (Fig. 3.1) which was now pale in colour and transparent. A bright yellow band was retained on the column which suggested that the chromatography step was successful in removing most of the polyphenols. Furthermore, concentration of the sample from approx. 200 ml after heat treatment and dialysis to 50 ml after chromatography was achieved without excessive loss in activity as recoveries were routinely in the 60-70% range (Table 3.1).

In comparison with the original preparation of total E3 isolated from mitochondria and purified by affinity chromatography (R. Fullerton, Ph.D. Thesis, Glasgow University, 1995), this modified protocol recovered 20 to 30-fold more units of E3 enzyme. In total, approx. 10 mg of E3 were obtained in comparison to 0.2 mg by the original method. Both the affinity chromatography step of the original method and the DEAE-52 ion exchange chromatography performed in the modified method produced a 30 to 35-fold purification. Although the E3 from the affinity chromatography step was extremely pure (Fig. 3.2, lane 5), the DEAE-52 column sample was not completely homogeneous. An indication of the difference

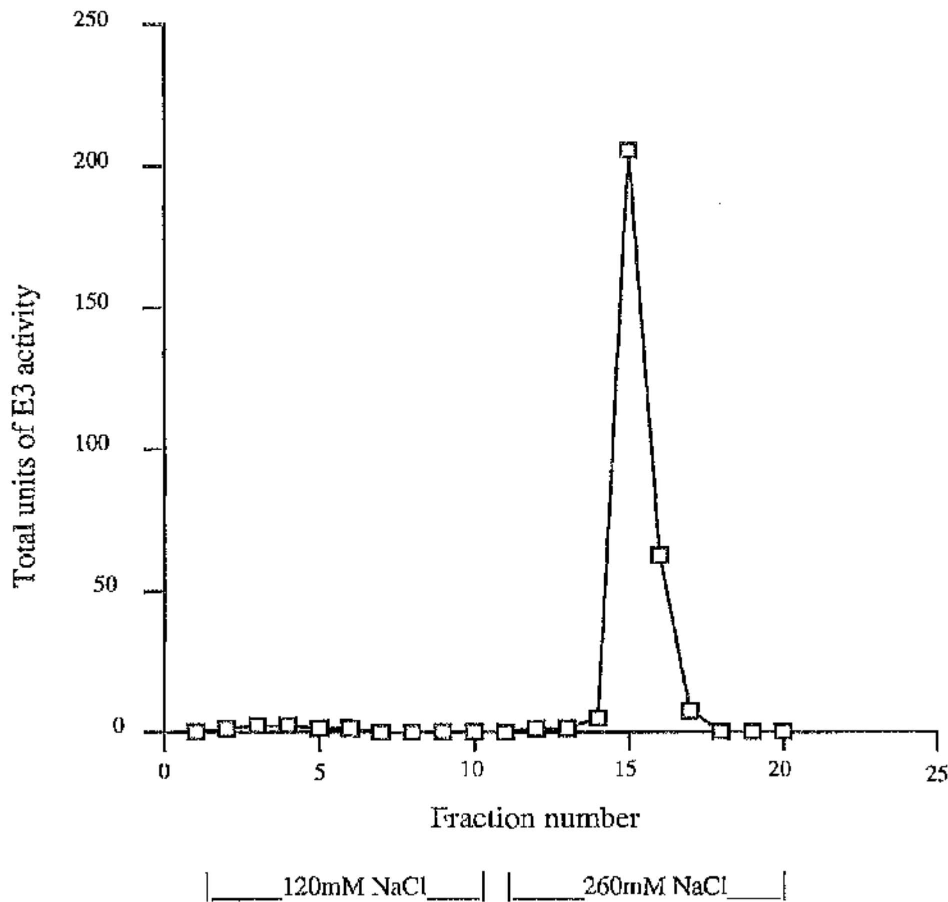


Fig. 3.1: Elution of E3 activity from the DEAE-52 ion exchange column
Heat treated potato tuber extract was dialysed into 20mM potassium phosphate buffer, pH 6.8 (containing 1 mM DTT, 1 mM EDTA, 1 mM PMSF and 1 mM benzamidine-HCl). The sample was divided in half (2 x 100 ml) and applied to two separate columns packed with DEAE-52 (55 ml bed volume) which had been equilibrated with 20 mM potassium phosphate buffer, pH 6.8. Unbound protein was washed through with 1 vol of 120 mM NaCl and E3 was eluted with 260 mM NaCl (both prepared in 20 mM potassium phosphate buffer, pH 6.8). Fractions (9 ml) were collected and and assayed for E3 activity (2.2.10.2). Total units of activity were plotted against fraction number.

(Total units = μmol of NADH formation/min/9ml)

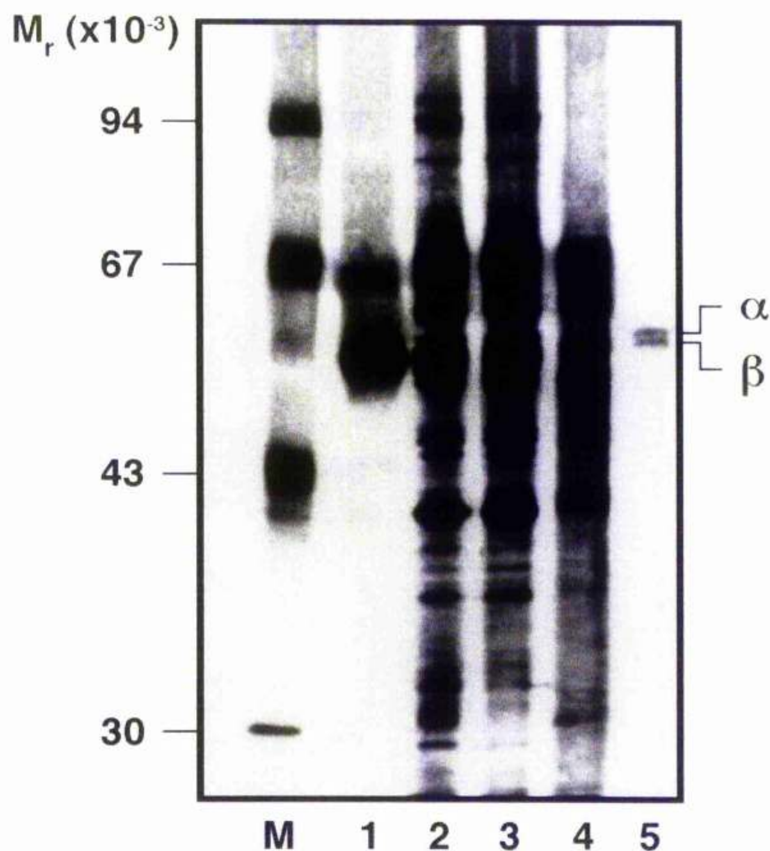


Fig. 3.2: Analysis of purification of potato tuber E3 by E2/X affinity chromatography

Protein from various steps in E3 purification on the E2/X affinity column from potato tuber mitochondria was electrophoresed on 10% (w/v) SDS-PAGE and silver stained. M: molecular weight markers, 1: purified porcine heart E3, 2: crude potato mitochondrial extract, 3: heat-treated mitochondrial extract (65°C/10 min), 4: unbound protein, 5: peak E3 activity fraction from eluant

Reproduced from Fullerton *et al.* (1996)

in purity of the two samples is provided by a comparison of the specific activities, which were approx. 100 units/mg (purification to homogeneity by affinity chromatography, R. Fullerton, Ph.D Thesis, Glasgow University, 1995) and 42 units/mg (purification by DEAE-52 chromatography, Table 3.1). Additionally, the specific activity suggests that the E3 should be about 40% of the total sample after purification by DEAE-52 chromatography. Nevertheless, the enrichment achieved allowed the purification of the E3 isoforms to near homogeneity in the subsequent steps.

3.2.3 Resolution of the E3 Enzyme by Source 15Q Ion Exchange Chromatography

The final stage involved further purification and resolution of the three peaks of E3 activity on an ion exchange chromatography column attached to an FPLC apparatus. The 1 ml Mono Q anion exchange chromatography column used in the original method was replaced with an 8 ml Source 15Q column to increase the binding capacity for the E3 enzyme. This strong anion exchanger is identical to Mono Q except that the bead size is larger. The sample, eluted from the DEAE-52 column, was applied to the Source 15Q column as described in Fig. 3.3. E3 activity assays (section 2.2.10.2) confirmed the elution of three peaks of E3 activity (Fig. 3.3) in agreement with the results of Fullerton *et al.* (1996). However, it was noticed that the peak ratios were somewhat different. On average the peak ratios were 1:3:2 in comparison with 1:3:5 reported by R. Fullerton (Ph.D Thesis, Glasgow University, 1995). A single column run was sufficient to resolve the entire sample, allowing nearly 40-50% of the E3 enzyme to be recovered (Table 3.1). The units of activity were in the range of 60 units for peak 1, 185 units for peak 2 and 65 units for peak 3; nearly 20-fold more enzyme per peak compared with purification from isolated mitochondria. The protein concentrations were in the range of 1-3 mg per peak in comparison with 0.03-0.06 mg of E3 obtained per peak with the original method. The purification factor was difficult to determine since the single E3 sample was being separated into three distinct activities.

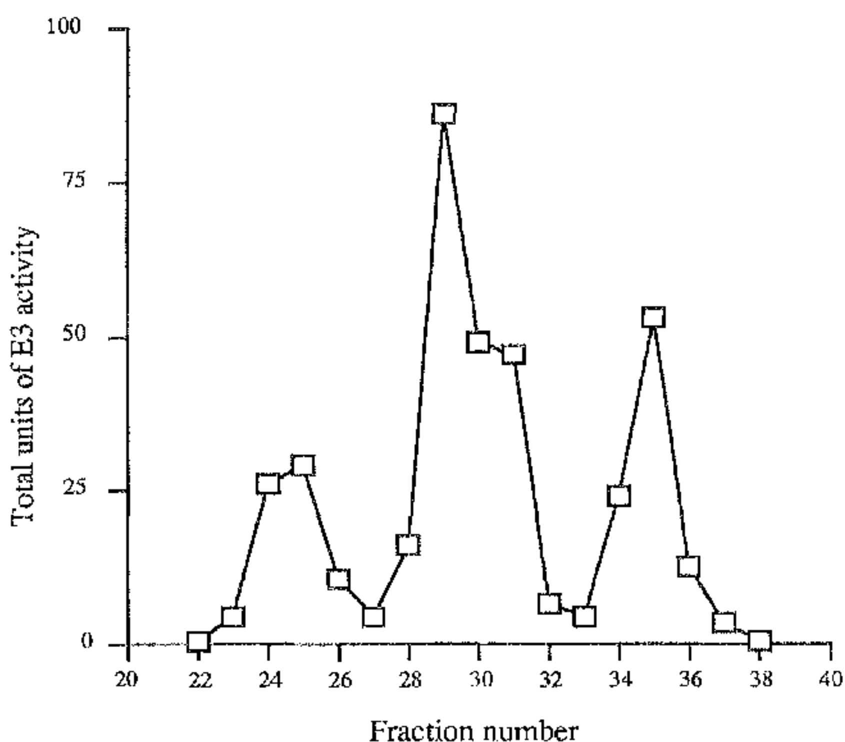


Fig. 3.3: Resolution of the DEAE-52 purified potato extract by Source 15Q ion exchange chromatography

The peak fractions eluted from the DEAE-52 column were pooled and dialysed into 50 mM imidazole buffer containing 1 mM DTT before being applied to the Source 15Q column (8 ml) in buffer A (1 mM EGTA, 1 mM β-mercaptoethanol, 10 mM potassium phosphate buffer, pH 6.8) at a flow rate of 1 ml.min^{-1}. Proteins were eluted with a linear gradient (240 ml) of increasing potassium phosphate (10-400 mM in buffer A) at a flow rate of 2 ml.min^{-1}. Fractions (2 ml) were collected for E3 activity assays (section 2.2.10.2).

□ Units of E3 activity per fraction. (μmol NADH.min^{-1}.2ml^{-1})

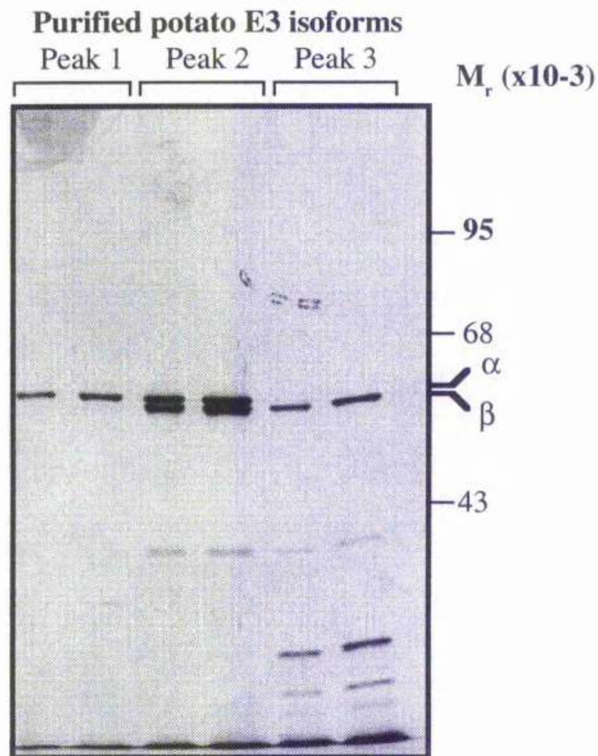


Fig. 3.4: Purification of the E3 isoforms from a crude organellar pellet from potato tubers, using the modified protocol

E3 was purified from potato tubers as described in the text. Samples (1 ml) were collected at each stage and resolved on a 10% (w/v) SDS-PAGE gel (in duplicate) as described in Materials and Methods, section 2.2.5 and then subjected to silver staining (section 2.2.6). Peak 1: $\alpha 2$ isoform Peak 2: $\alpha\beta$ isoform Peak 3: $\beta 2$ isoform

Analysis of the protein in each peak by SDS-PAGE and silver staining (Fig. 3.4) revealed that the purity of each E3 isoform was close to homogeneity apart from two or three low M_r contaminating bands. In contrast, the isoforms reported by R. Fullerton (Ph.D Thesis, Glasgow University, 1995) had been purified to complete homogeneity. Therefore, the scale-up procedure increased the final yield of enzyme at the cost of slightly reduced purity. The gel showed that peaks 1 and 3 represent the α_2 and β_2 homodimers respectively and peak 2 the $\alpha\beta$ heterodimer. The α and β polypeptides are clearly distinguishable in the peak 2 sample.

3.2.4 Determination of the Subunit Molecular Mass of E3 by Electrospray Mass Spectrometry (Es-Ms)

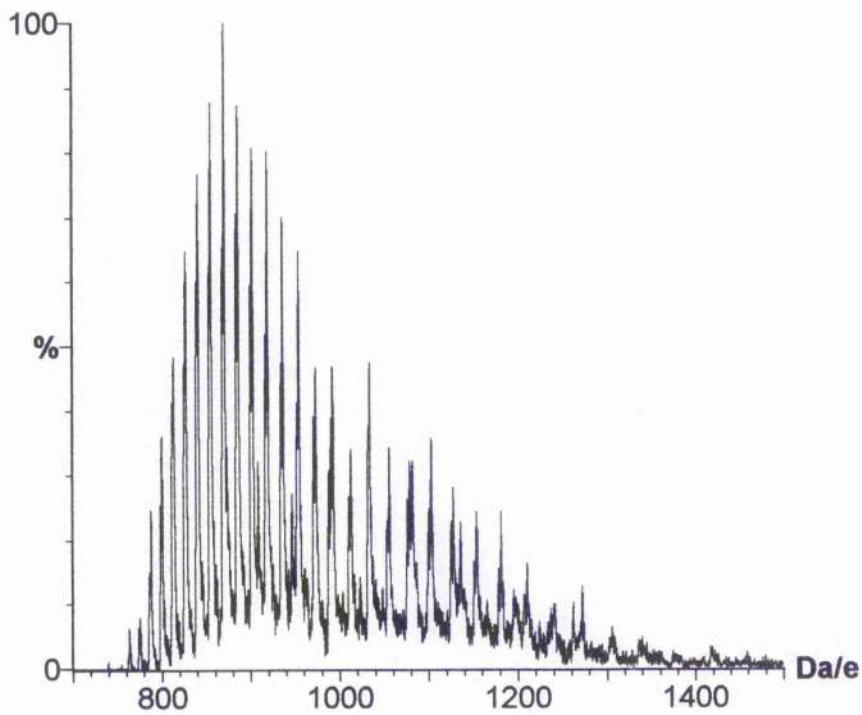
An important consequence of the increased yield of the E3 isoenzymes was that it provided sufficient protein for the precise determination of subunit M_r values by ES-MS. All traces of salt were removed from the E3 isoforms as described in Materials and Methods (section 2.2.35) to ensure low background signals from contaminating ions. ES-MS was performed by T. Krell at the facility based at Strathclyde University.

Fig. 3.5A presents the electrospray mass spectra of the $\alpha\beta$ heterodimer (peak 2 sample) and Fig. 3.5B is the MaxEnt deconvolution of the spectra showing the M_r of the α and β polypeptides. The M_r for the α polypeptide was calculated to be 49 446 $\pm$ 12.28 and the M_r for the β polypeptide 49 562 $\pm$ 6.83; a difference of approx. 116 Da. The small difference in mass between the two polypeptides present in this sample was clearly discernible by ES-MS, a measure of the accuracy of this technique. A confirmation of the two values was obtained from two separate samples by analysing the α_2 homodimer from peak 1 (Fig. 3.6A) and the β_2 homodimer from peak 3 (Fig. 3.6B). As can be seen from MaxEnt deconvolutions of the spectra, the M_r of the two major peaks in the sample agree very well with the M_r of the α and β polypeptides cited above. Interestingly, these M_r values were lower than the ones obtained by SDS-PAGE (Fig. 3.4) which were approx. 58 000 for the α polypeptide and 56 000 for the β polypeptide.

Fig. 3.5A: Electrospray mass spectra of the $\alpha\beta$ heterodimeric E3 isoenzyme from potato tuber mitochondria

Fig. 3.5B: MaxEnt deconvolution of electrospray mass spectra of the $\alpha\beta$ heterodimeric E3 isoenzyme from potato tuber mitochondria

The M_r of the α polypeptide is 49 446 $\pm$ 12.28 and the M_r for the β polypeptide is 49 565 $\pm$ 6.83.


100
%
0
Da/e
800
1000
1200
1400

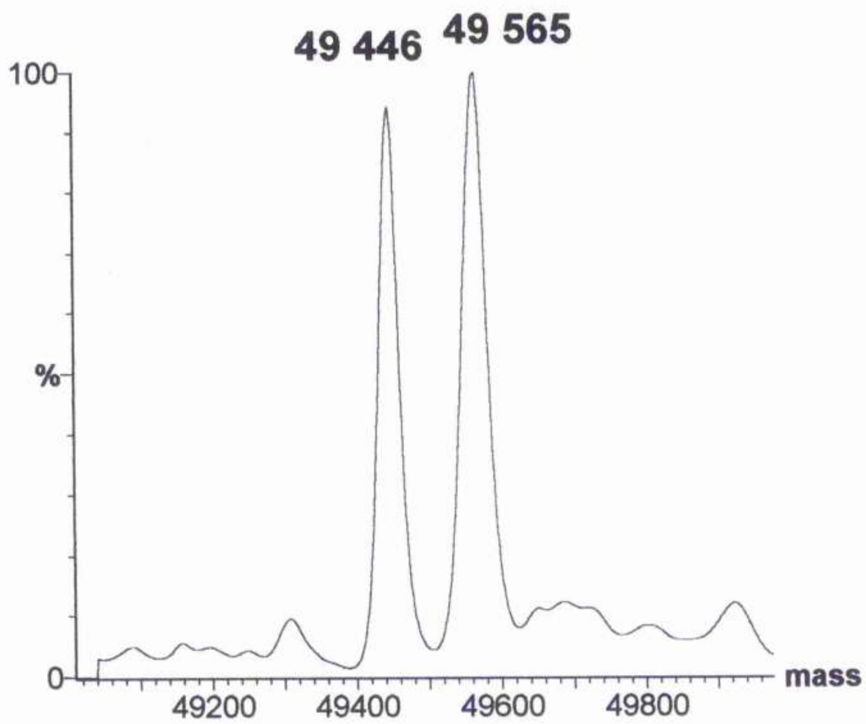
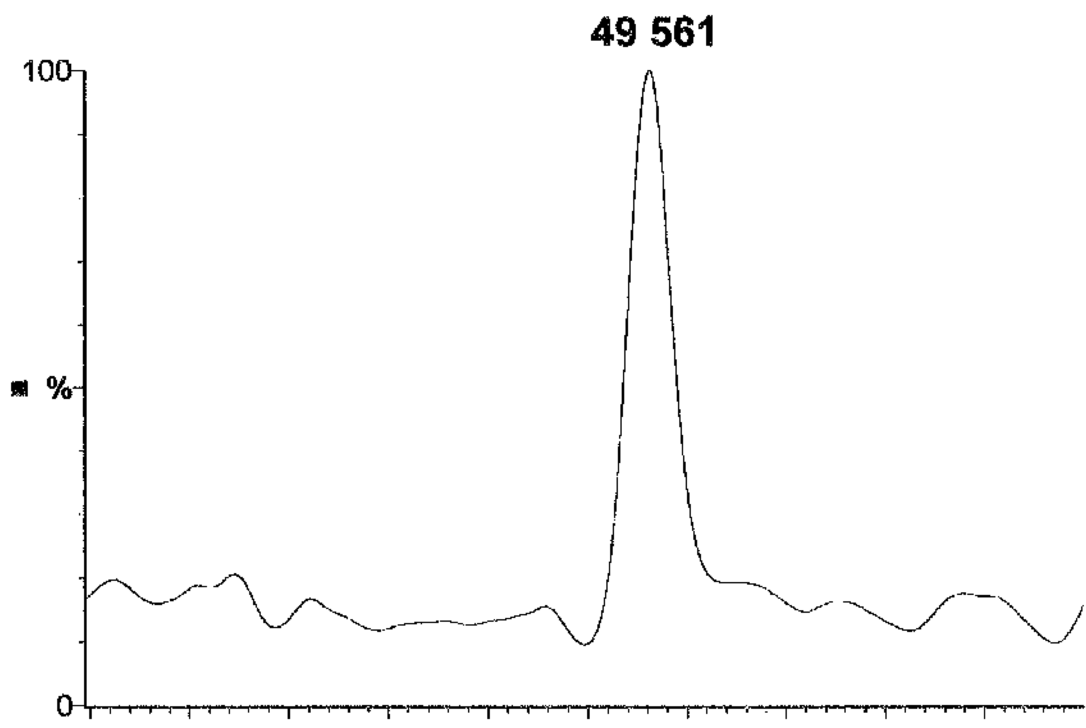
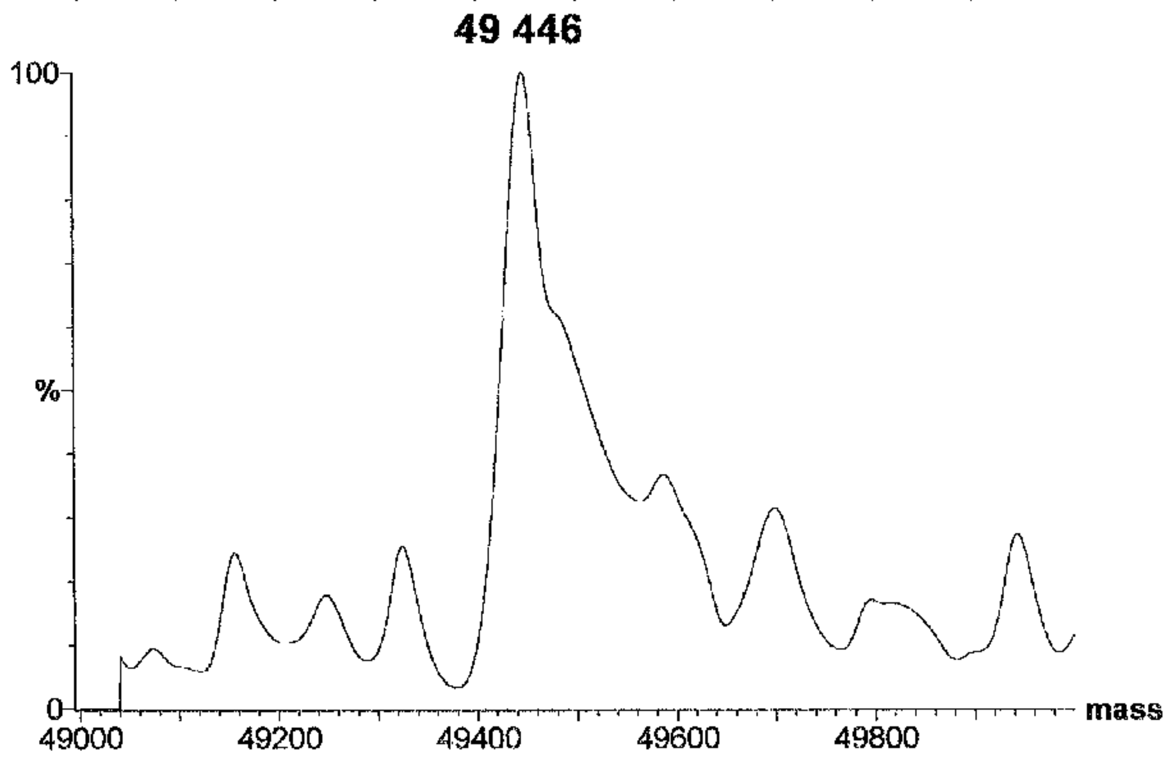

49 446
49 565
100
%
0
mass
49200
49400
49600
49800

Fig. 3.6A: MaxEnt deconvolution of electrospray mass spectra of the β_2 E3 isoenzyme from potato tuber mitochondria; M_r for the β polypeptide

Fig. 3.6B: MaxEnt deconvolution of electrospray mass spectra of the α_2 E3 isoenzyme from potato tuber mitochondria; M_r for the α polypeptide


49 561
100
%
0


49 446
100
%
0
mass
49000
49200
49400
49600
49800

3.3 Discussion

The aim of scaling up the original method for the purification of the E3 isoenzymes was completed successfully since the yield of each isoform was increased 20-30 fold from the original method. The Source 15Q step was effective in resolving the three isoforms (Fig. 3.5). Approx. 1-3 mg of protein/peak were obtained in the final sample compared with only 0.03-0.06 mg/peak obtained by purifying mitochondria prior to purification. However, in comparison with R. Fullerton (Ph.D Thesis, Glasgow University, 1995) who had reported completely pure E3s, the isoforms did contain some contaminating bands of low M_r value (Fig. 3.4). This is a reflection of the extraction of a crude (2-fold) organellar pellet as opposed to a pure mitochondrial pellet which was required for the scale-up procedure. Hence, the increased yield was at the cost of slightly reduced purity. The addition of the DEAE-52 column had the advantage that it removed most of the polyphenols from the sample at an early stage. This was important because these oxidising compounds are known to inhibit plant enzymes. Moreover, it provided a quick method of purification (36-fold) and concentration in one step, with minimal loss in E3 activity.

An important observation, with regard to the crude nature of the extract, was that no E3 species other than the three original mitochondrial isoforms were detected. This is strongly suggestive of E3s with a mitochondrial origin as they were directly comparable to the isoenzymes described by R. Fullerton (Ph.D Thesis, Glasgow University, 1995). However, very similar E3 isoforms existing in other organelles (i.e storage plastids of potato tubers) may not be distinguished by this method of purification. The ratio of the third peak (compared to peaks 1 and 2) eluting from the Mono Q column was found to be half the value reported in the original method. An explanation was provided subsequently by R. Cook (unpublished observations) who examined the heat stabilities of the three isoforms. Peak 3 (β_2 isoform) was found to undergo 50% inactivation over a 30 min

incubation at 67°C. This result correlated well with the apparent reduction in activity of this isoform, presumably resulting from the heat treatment step.

The scaled-up purification procedure provided enough material for a precise M_r determination of the α and β polypeptides by ES-MS. Previous attempts had failed because insufficient enzyme was available for the process. The M_r values were found to be 49 446 (α) and 49 562 (β). Thus, the M_r values from SDS-PAGE were over-estimated by approx. 12-15%. The discovery of the very small difference in mass between the two polypeptides was unexpected, especially as they could be clearly distinguished by SDS-PAGE. In addition, recent preliminary evidence from our laboratory (R. Cook, unpublished observations) has suggested that each isoform exhibits distinct biochemical properties. They were found to have differing K_m values for NAD^+ (120 μM [α_2], 200 μM [$\alpha\beta$] and 500 μM [β_2]) and for dihydrolipoamide (80 μM [α_2], 220 μM [$\alpha\beta$] and 240 μM [β_2]). Similarly, they had differing heat stabilities, with the α_2 exhibiting the greatest stability over a 30 min incubation at 80°C. Their sensitivity to high ionic strength varied with the α_2 form, again being more resistant to inactivation with NaCl (R. Cook, unpublished observations).

It was unclear why the polypeptides ran anomalously through a denaturing polyacryamide gel as the larger β form travelled faster than the smaller α form. The determined M_r values of the α and β polypeptides are the same range found for E3 subunits from a variety of sources. A comparison of M_r values, as predicted by sequence data, is presented in Table 3.2. In all cases, the M_r found by SDS-PAGE was approx. 6-8 000 more than the precise M_r determinations.

Of considerable interest is whether the two polypeptides are the products of two different genes or whether they are the products of differential splicing. Only sequence analysis will be able to disclose their relationship with each other. As there have been no reported post-translational modifications of E3s in any system, it is difficult to attribute post-translational modifications as a reason for the mass difference, or indeed the biochemical variations between the isoforms. Although

the first 28 amino acids were shown to be identical by N-terminal sequencing (Fullerton *et al.*, 1996), it is possible that the small difference in size is due, co-incidentally, to substitutions of amino acids very similar in mass in view of the differing enzymatic and protein-chemical characteristics of the three mitochondrial isoforms.

Table 3.2

A comparison of the M_r values from a variety of E3 subunits from their deduced amino acid sequences

Source of E3	M_r / subunit	Reference
Human liver	50 216	Pons *et al.* (1988)
Saccharomyces cerevisiae	54 010	Ross *et al.* (1988)
Pseudomonas putida (LPD-Val)	48 949	Burns *et al.* (1989b)
Azobacter vinelandii	49 436	Westphal and De Kok (1988)
Pea leaf mitochondria	49 721	Bourguignon *et al.* (1992)
Potato tuber mitochondria (α)	49 446	This study
Potato tuber mitochondria (β)	49 562	This study

Chapter 4

Analysis of the Possible Complex Specific Roles of the Potato Tuber E3 Isoforms by Reconstitution of Intact OGDC Activity

4.1 Introduction

Considerable interest has focused on the question of whether all mitochondrial 2-oxoacid dehydrogenase complexes share a common E3 or whether each complex uses a different E3 subunit, each encoded by separate genes. E3s from a variety of organisms are found to have extensive identity at the amino acid level (human liver E3 shows 96% identity with that of pig heart E3) and remarkable similarity in their tertiary structures (Carothers *et al.*, 1989). They are all organised into four domains; the FAD-binding domain, the NAD-binding domain, the central domain and the interface domain, with a high degree of sequence conservation occurring within these domains.

Immunological data (Matuda & Saheki, 1985) and observations on human genetic disorders (Stansbie *et al.*, 1986) have suggested that a single E3 for PDC, OGDC, BCDC and probably GDC (Kochi *et al.*, 1986; O'Brien, 1978) exists in mammals. However, there is some isolated evidence conflicting with this general view. Immunological evidence suggests the existence of two isoenzymes of E3 in rat liver; one possibly being specific for GDC (Carothers *et al.*, 1987). In support of this observation, patients with E3 deficiencies have excesses of the 2-oxoacids although they display normal glycine metabolism (Carothers *et al.*, 1987).

In the bacterial organisms, *Azotobacter vinelandii* and *Pseudomonas fluorescens*, one E3 gene product is present. Other species provide considerable evidence in support of distinct E3s complementing the 2-oxoacid dehydrogenases. *Escherichia coli*, was always thought to have a single E3 gene (*lpd*), found on the PDC operon, which provided E3 for PDC and OGDC (Spencer & Guest, 1985). Incredibly, a previously unknown form of E3 was purified from a *lpd*$^-$ *E. coli* mutant (Richarme, 1989). This distinct gene product was nearly 10 000 Da smaller than the normal *E. coli* enzyme, with a M_r of 46 000 Da and had been overlooked as it was expressed at low levels in wild-type *E. coli*. The function of this

additional E3 is unclear but it was proposed to be involved in the binding protein-dependent galactose transport system.

Pseudomonas putida can use branched-chain amino acids for growth, and consequently possesses an E3 (LPD-Val) specific to BCDC (Sokatch *et al.*, 1981) in addition to the distinct E3 (LPD-Glc) component for OGDC, PDC (Burns *et al.*, 1989b) and GDC (Sokatch & Burns, 1984). A third E3 (LPD-3) was discovered by Burns *et al.* (1989a) which was related to LPD-Glc. *Alcaligenes eutrophus* also has two distinctly different E3 components, one for PDC and the other for OGDC but it is not yet clear whether they are interchangeable. Freudenberg *et al.* (1989) reported the first purification of a specific GDC E3 from the anaerobic, glycine utilising, *Eubacterium acidaminophilum* which was unusually small (35 kDa).

E3s have been found in the bloodstream form of the eukaryotic African parasite, *Trypanosoma brucei* (Danson *et al.*, 1987) and halophilic archaebacteria (Danson *et al.*, 1984). This is an unusual situation as these organisms are known to lack the 2-oxoacid dehydrogenases, suggesting a cellular function for E3 which is yet to be elucidated.

In plants the situation also requires further clarification. However, initially Turner *et al.* (1992) and Bourguignon *et al.* (1992 & 1996), both publishing the sequence for GDC E3 from pea mitochondria, simultaneously concluded that the same E3 was shared between PDC, OGDC and GDC. Uniquely in plants, a distinct PDC is found in chloroplasts and as a reflection of this, a distinct E3 was recently purified from this organelle (Conner *et al.*, 1996).

Whilst it seems that the organisation of the E3 enzyme is unpredictable, it is surprising that three different enzymes have been discovered in potato tubers as Bourguignon *et al.* (1996) had confirmed the presence of one E3 in pea leaf mitochondria. Furthermore, analysis of other legumes (horseradish and turnip) in our laboratory revealed the presence of only one E3. More recently three isoenzymes were found in barley leaf mitochondria (Fullerton *et al.*, 1996), indicating that multi-isoforms of E3 are not unique to potatoes. As detailed in

chapter 3, the isoenzymes in potato have arisen as a consequence of the existence of two polypeptide subunits: α and β, which have the ability to combine into distinct α_2, $\alpha\beta$ and β_2 dimeric species. Their physiological roles in potato and barley are not known but it is seems, at least in the potato, that they may have differing biochemical properties (R. Cook, unpublished observations).

Since the existence of complex specific isoforms has been reported in some prokaryotes, the presence of multiple isoforms may suggest that they are playing similar roles in potatoes. Reconstitution studies were designed to determine if one isoform would preferentially promote the activity of OGDC stripped of its E3 component. OGDC was chosen because it is possible to dissociate only the E3 component from the E1/E2 core, unlike PDC where both the E1 and the E3 components dissociate together. Initially, bovine heart OGDC E1/E2 core was used in consideration of its ready availability in the laboratory. It had been demonstrated that potato E3 has some affinity to the mammalian core since the enzyme could bind to an affinity column prepared with the E1/E2 core from bovine heart OGDC, similar to the E2/X column developed by Fullerton *et al.* (1996). Upon separation on a Mono Q column, the elution profile of the total E3 purified on this affinity column, showed three peaks of E3 activity which were identical to the potato isoforms described in Fig. 3.4 of chapter 3 (data not shown). Once the conditions for reconstitution were established, the experiments were to be extended to OGDC from potato tubers. The ability of E3 from heterologous sources (pig heart, bovine mucosa and *Candida utilis* [yeast]) to reconstitute the activity of bovine heart OGDC was to be compared to that of the potato isoforms. Analysis of PDC reconstitution by the potato isoforms was performed in conjunction with other researchers in the laboratory.

4.2 Results

4.2.1 Dissociation of Bovine Heart OGDC

The first step was to dissociate and separate the E1/E2 subcomplex from the E3 component of bovine heart OGDC, prepared as described in Materials and Methods (section 2.2.9.1). Dissociation was first attempted using high salt (2 M NaCl) and alkaline conditions used previously in the laboratory for the dissociation of PDC into E2/X and E1/E3 subcomplexes. OGDC was incubated for 1 h as described in Fig. 4.2 (legend), after which the E1/E2 subcomplex was separated from the dissociated E3 component by gel filtration on a Superose 12 column (bed volume-100 ml) attached to a FPLC apparatus. Protein eluted in two peaks (Fig. 4.1), at 32 ml (void volume) and approx. 66 ml. SDS-PAGE analysis showed that the high M_r E1/E2 subcomplex eluted in the void volume (Fig. 4.2, peak 1); however alkaline pH treatment was found to be too harsh for the OGDC since the E1 component was also being partially stripped from the core and co-eluting with the E3 fraction (Fig. 4.2, peak 2).

Dissociation was modified to a 20 min incubation on ice in the presence of lower salt (1 M NaCl) at neutral pH. Additionally, 1 mM $MgCl_2$ and 0.2 mM ThDP were included to maintain the stability of the E1 component. Again, the protein eluted in a similar profile from the gel filtration column as was illustrated in Fig. 4.1. SDS-PAGE analysis of the peak fractions (Fig. 4.3, peak 1) confirmed that the E1/E2 core was intact with no E3 contamination. Surprisingly, residual background OGDC activities were measurable in the E3 fraction which would interfere with reconstitution activities. The problem was solved by heat treatment of the E3 fraction at 65°C for 10 min followed by pelleting the denatured protein, which eliminated background activity without loss in E3 activity (Fig. 4.3). There is a heat stable contaminant (M_r 45 000) in the E3 sample which appears to co-purify with the OGDC sample on occasions. Attempts to identify it by N-terminal protein sequencing were unsuccessful.

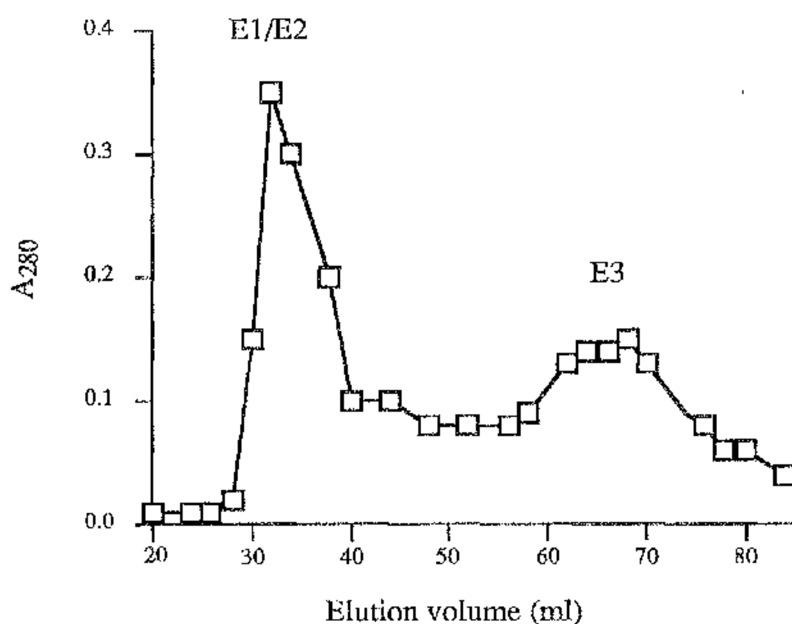


Fig. 4.1: Dissociation of bovine heart OGDC

OGDC (15 mg) was treated with 1 M NaCl as described in section 2.2.9.2 prior to loading onto a Superose 12 FPLC column (100 ml bed volume). Running buffer (1 mM DTT,1 M NaCl, 0.01% (v/v) Triton X-100 and 50 mM potassium phosphate buffer, pH 7.6) was pumped through the column at a flow rate of 1 $ml.min^{-1}$. Fractions (2 ml) were collected and their absorbances measured at 280 nm.

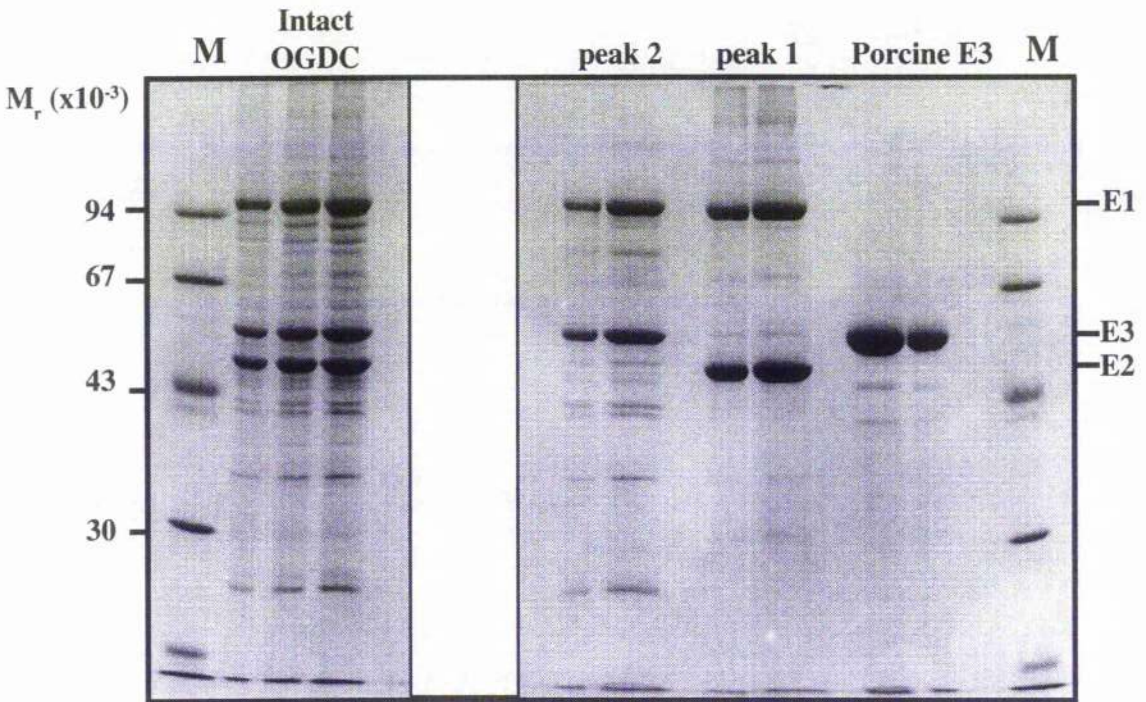


Fig. 4.2: Dissociation of bovine heart OGDC under high salt (2 M NaCl) at alkaline pH (pH 9)

Bovine heart OGDC (20 mg) was incubated at 4°C for 1 h in 50 mM Tris-HCl (pH 9), 2 M NaCl, 1 mM DTT, 0.01% (v/v) Triton X-100. The dissociated E1/E2 (peak 1) and E3 (peak 2) components were separated from each other by gel filtration on a Superose 12 column (100 ml bed volume) attached to an FPLC. The column running buffer was 50 mM Tris-HCl (pH 9), 1 M NaCl, 1 mM DTT and 0.01% (v/v) Triton X-100, with a 4 M NaCl pre-injection onto the column to maintain dissociating conditions ahead of the moving components.

M= M_r marker

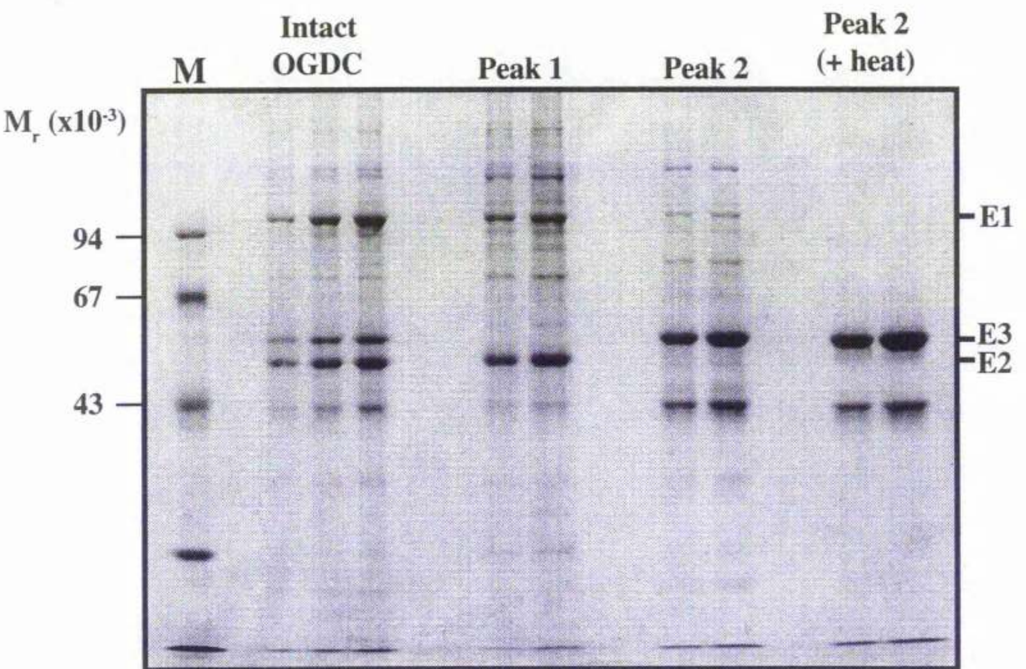


Fig. 4.3: Dissociation of bovine heart OGDC at 1 M NaCl at neutral pH

OGDC (20 mg) was incubated at 4°C for 20 min in 50mM potassium phosphate buffer (pH 7.6), 1M NaCl, 1 mM DTT, 0.01% (v/v) Triton X-100, 1 mM $MgCl_2$ and 0.2 mM TPP. The dissociated E1/E2 (peak 1) and E3 (peak 2) components were separated from each other on a Superose 12 gel filtration column (100 ml bed volume) attached to a FPLC. The column running buffer was as the dissociation buffer described above. After separation, the E3 fraction (peak 2) was heat treated at 65°C for 10 min and the denatured proteins pelleted by centrifugation at high speed in a microfuge. The resulting supernatant was also analysed on the gel.

M= M_r markers

4.2.2 Time Course for Reconstitution of OGDC Activity

Initially, a time course was carried out to determine the optimal conditions required for maximal reconstitution of bovine heart OGDC activity after mixing the E1/E2 core with the native OGDC E3 (purified as in section 2.2.9.7). Reconstitution assays were performed as described in Materials and Methods (section 2.2.10.4). The zero time point was accomplished by adding the two components directly to the assay cuvette without pre-incubation. The extent of reconstitution was calculated as a percentage of original OGDC specific activity, determined before dissociation. As can be seen in Fig. 4.4, 60% of intact activity is already observed at zero time. This spontaneous re-assembly is an inherent property of multienzyme complexes. However, it was unclear as to why it took an additional 3-5 min for maximal re-assembly (70-80%) to occur. Recently, Sanderson *et al.* (1996b) provided an explanation for this effect by performing a time course of reconstitution of PDC activity after pre-incubation of the E2/X and E1/E3 components in 50 mM potassium phosphate buffer containing 2.6 mM cysteine-HCl. They reported that maximal levels of reconstitution (70-80%) of PDC activity was attained immediately and this high level was maintained over 40 min (Fig. 4.5). The explanation offered for the elimination of the lag phase for maximal activity was that pre-incubation in buffer containing cysteine-HCl (a reducing agent) enabled a degree of re-activation of the separated components due to the reducing influence of sulphydryl agents. Therefore, the increase in activity over time shown in Fig. 4.4 does not represent further reassembly of the E1/E2 and the E3 components but a minor reactivation of the individual enzymes. A similar effect is often observed when purified OGDC is incubated in buffers containing DTT or β-mercaptoethanol.

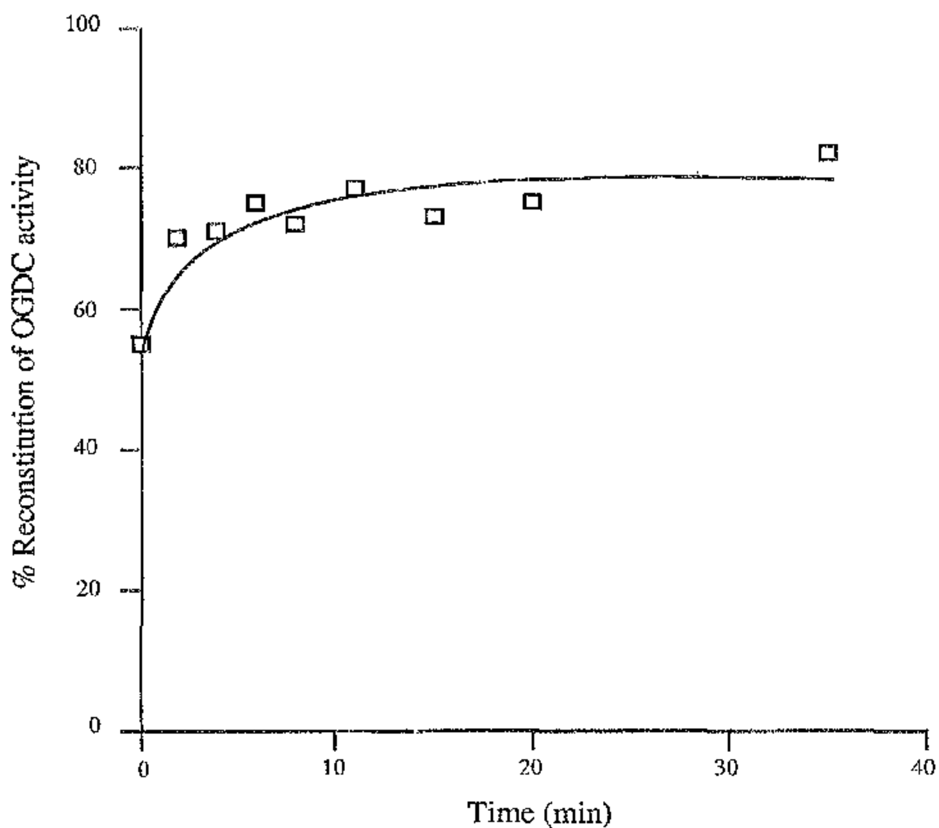


Fig. 4.4: Time course of reconstitution of bovine heart E1/E2 core with native E3

E1/E2 (60 μg) from bovine heart OGDC was incubated with an excess of E3 (30 μg) at 30°C. Aliquots (equivalent to 5 μg E1/E2) were removed and assayed for OGDC activity (section 2.2.10.1) at prescribed time points. For the zero time point, components were added directly to the assay cuvette. Results represent the average of duplicate estimates differing by less than 5% and are presented as percent of original OGDC activity, measured before dissociation.

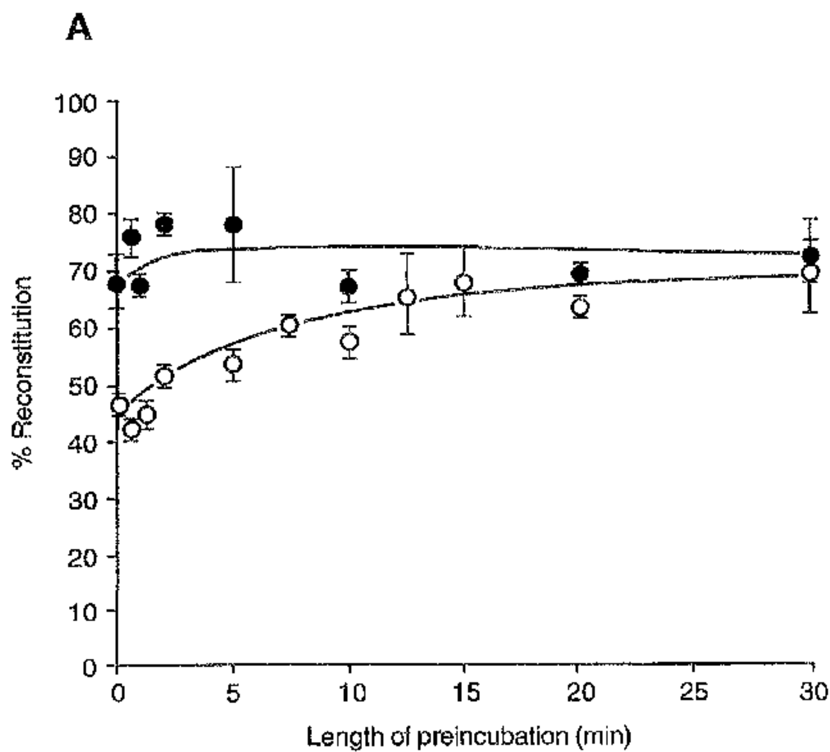


Fig. 4.5: Time course of reconstitution of bovine heart PDC activity by re-assembly of E2/X and E1/E3 subcomplexes

Equal volumes of E2/X and E1/E3, diluted 1:10 in assay buffers A (○--○) or A and B (●--●) (see Methods section 2.2.10.1 for buffers) to 0.1-0.2 mg.ml^{-1}, were preincubated together at 30°C in the cuvette containing buffers A and B minus substrate for the times indicated, prior to initiation of catalysis by the addition of pyruvate. The dissociated fractions exhibited negligible complex activity. Specific activity reconstituted was determined from duplicate assays and expressed as a % of the specific activity of the intact complex prior to dissociation.

Reproduced from Sanderson *et al.* (1996)

4.2.3 Reconstitution of OGDC Activity by Homologous and Heterologous E3 Enzymes

The aim of these experiments was to analyse the ability of the potato tuber E3 isoenzymes to reconstitute OGDC activity. The levels of reconstitution were compared with those obtained for native OGDC E3, bovine heart PDC E3, *Candida utilis* E3 (yeast) and porcine heart E3s. The E3 samples were all prepared as in section 2.2.9.7 and potato E3 isoforms as in 2.2.9.3.

The method of reconstitution was described in section 2.2.10.4 and results were presented as a percentage of original OGDC activity (Fig. 4.6). Native E3 and bovine heart PDC E3 consistently gave a high level of reconstitution of activity (60-80%) over several experiments, confirming that E3s from both these complexes were completely interchangeable. In comparison, porcine heart E3 was only able to sustain intermediate levels (20%) of activity. This was considered highly unusual since it has high sequence identity with human liver E3 (96%). The porcine heart and yeast E3s, being commercial preparations may have been contaminated with an inhibitory component. However, these E3s, after purification by Mono Q ion exchange chromatography, still provided the same results. An interesting situation was discovered when reconstitution was attempted using bovine mucosal E3. Activities of only about 15% were sustainable with bovine mucosal E3 (data not shown) indicating that there may be tissue specific isoforms of the E3 component in mammalian systems. In each case, reconstitution activity reached saturation with 1.5 μg of E3 to 5 μg of E1/E2 core. This correlated closely with the theoretical stoichiometry of E3 representing one fifth of the intact complex by weight and suggested high affinity of binding to the E1/E2 subcomplex.

Having demonstrated the ability to stimulate reconstitution of OGDC activity to varying levels employing E3s from various sources, the ability of the three individual potato isoforms to sustain OGDC activity was analysed. Unfortunately, the potato isoenzymes were unable to stimulate any intact activity of bovine heart OGDC, even with 100-fold excess enzyme (not shown in Fig. 4.6). The situation

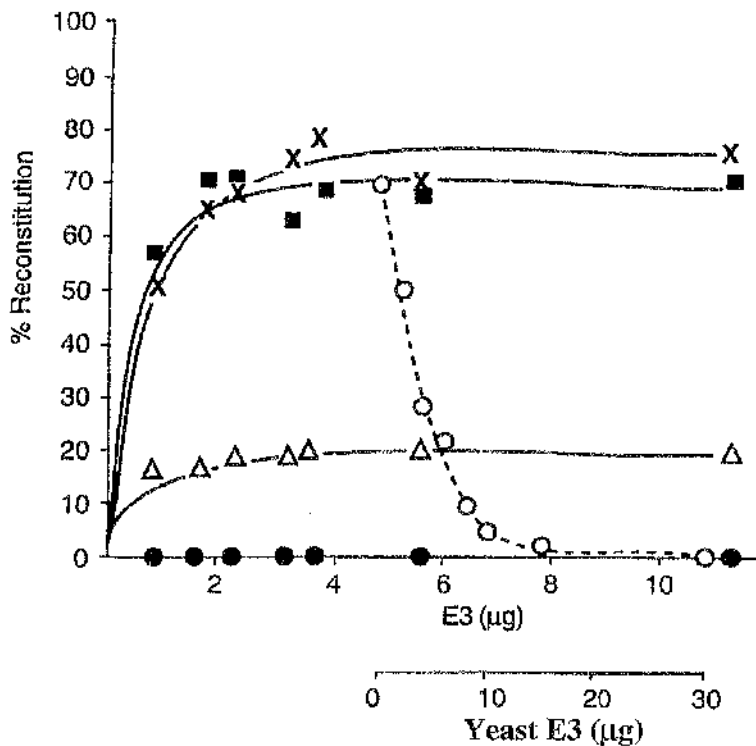


Fig. 4.6: Comparison of reconstitution of bovine heart OGDC activity by re-assembly of native and heterologous E3s with the E1/E2 subcomplex. Purified E1/E2 subcomplex of OGDC (Materials and Methods, section 2.2.9.2), was incubated with increasing amounts of E3 from bovine heart OGDC (**X–X**), bovine heart PDC (■–■), porcine heart E3 (△–△), yeast E3 and the three potato tuber E3 isoforms (●–●). Samples were incubated for 12 min at 30°C with a fixed amount of bovine heart OGDC E1/E2 core (7 μg). All samples were present in assay buffer A (minus NAD^+). Intact OGDC was assayed in duplicate (section 2.2.10.1) and reconstituted activity expressed as a % of original complex activity.

A competitive binding assay was performed with E1/E2 core (7 μg) and a fixed level of native bovine heart OGDC E3 (2 μg) in the presence of increasing amounts of yeast E3 as indicated (○–○) and incubated for 12 min at 30°C before assaying.

Reproduced from Sanderson *et al*. (1996b)

was similar to that of yeast E3, which failed to reconstitute any OGDC activity. The complete lack of reconstitution obtained with these E3s was unexpected so it was essential to determine if this was caused by an inability of these E3s to associate with the OGDC core in the reconstitution assay.

4.2.4 E3 Binding Analysis

As the potato and yeast E3 failed to stimulate reconstitution with the E1/E2 core, their binding capacity for the core was investigated. Binding to the mammalian core by the potato E3s had been expected as they had been purified initially in small amounts by binding to bovine heart OGDC E1/E2 subcomplex, immobilised to cyanogen bromide Sepharose 4B matrix. Assays were carried out, whereby increasing amounts of yeast E3 were incubated with a fixed amount of E1/E2 subcomplex and bovine heart E3. The maximum activity of reconstitution obtained with the core and the bovine heart E3 was rapidly decreased to zero activity with about 10 μg of yeast E3 (Fig. 4.6). At equimolar concentrations of the two E3s, activity was less than 40% of the original reconstituted activity suggesting a specific, high affinity interaction between yeast and bovine E3 for the E3 binding sites on the E1/E2 core. At a 4-fold excess of yeast E3 with respect to bovine E3, there was total inhibition of reconstituted OGDC activity.

Similar experiments with the potato E3 found that they competed poorly with the native E3 (Fig. 4.7), although there appeared to be a weak interaction of the α_2, $\alpha\beta$ and β_2 isoenzymes with the core since there was a 40% decrease in original activity in the presence of 40 to 50-fold excess of these isoforms. Unfortunately, limitations in amounts of each isoform precluded a more detailed investigation of their ability to interact with the mammalian E1/E2 OGDC subcomplex. The possibility of the contaminating bands in the potato E3 preparations interfering in binding interactions does exist, but it was not investigated further in this context. In summary, it appears that the potato mitochondrial α_2, $\alpha\beta$ and β_2 E3 isoforms interact weakly with the mammalian OGDC E1/E2 subcomplex. This is in contrast to yeast E3, which forms a tight

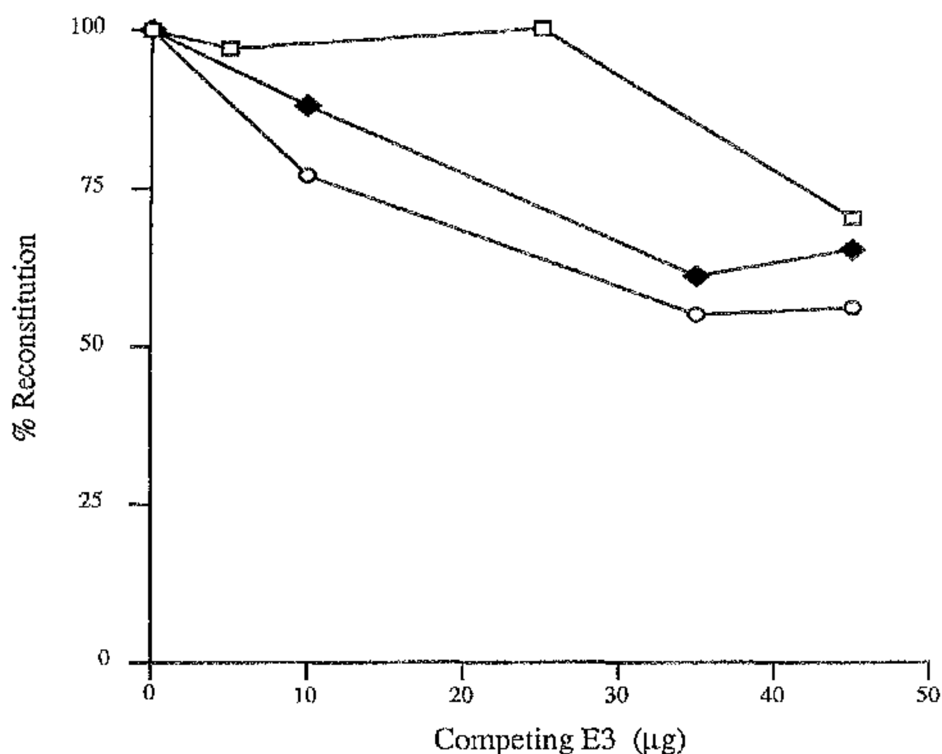


Fig. 4.7: Percent decrease in reconstituted OGDC activity resulting from competition between the native E3 and potato tuber E3 isoforms
Increasing amounts ot the E3 isoforms from potato tubers were added to a fixed amount of E1/E2 core (5 μg) from bovine heart OGDC and the lowest amount of native E3 required to give maximal, initial reconstitution (2 μg). Following a 12 min incubation at 30°C, OGDC activity was assayed as described in section 2.2.10.1. Results were presented as percent of original OGDC activity.

□ α_2 E3 (peak 1) ◆ $\alpha\beta$ E3 (peak 2) ○ β_2 E3 (peak 3)

association with this assembly as judged by their relative abilities to act as competitive inhibitors with the homologous E3 enzyme.

4.2.5 Partial Purification of OGDC from Potato Tubers

The next stage was to test the potato isoforms with a homologous, potato OGDC core complex as reconstitution had failed to stimulate mammalian OGDC activity. Intact OGDC has never been purified, mainly because it is present at low levels in plant tissues and is thought to dissociate in dilute solution. Consequently, all reported investigations have concentrated on PDC from various plant sources. Poulson and Wedding (1970) are credited with the only known report of partial OGDC purification from cauliflower florets where intact activity was achieved by reconstitution with porcine heart E3.

A crude isolation of OGDC was attempted from a potato tuber mitochondrial extract, following the work of A. Carmichael (Ph.D. Thesis, Glasgow University, 1994). Potato tuber mitochondria were isolated as described in Materials and Methods (section 2.2.9.5) and solubilised with 0.1% (v/v) Triton X-100. Total OGDC activities were very low compared with PDC activities (0.81 $\mu mol.min^{-1}$ and 20.6 $\mu mol.min^{-1}$, respectively). Poulson and Wedding (1970) had reported that NADH oxidase activity, which tended to mask OGDC activity in plant extracts, could be eliminated by freeze/thaw treatment of the mitochondria. This method of extraction was examined but found to have no effect on the final recorded activities.

The extract was applied to a Superdex 200 gel filtration column (bed volume-100 ml). As illustrated in Fig. 4.8, PDC and OGDC activities were eluted in a single peak in the void volume; however, OGDC was barely detectable. E3 activity was found in two peaks, at 46-50 ml (void volume fractions) and 70-80 ml. The void volume E3 is clearly associated with the PDC and OGDC and the second peak is thought to represent E3 dissociated from PDC, OGDC and GDC. This conclusion was reached because GDC is known to exist in a dissociated state on

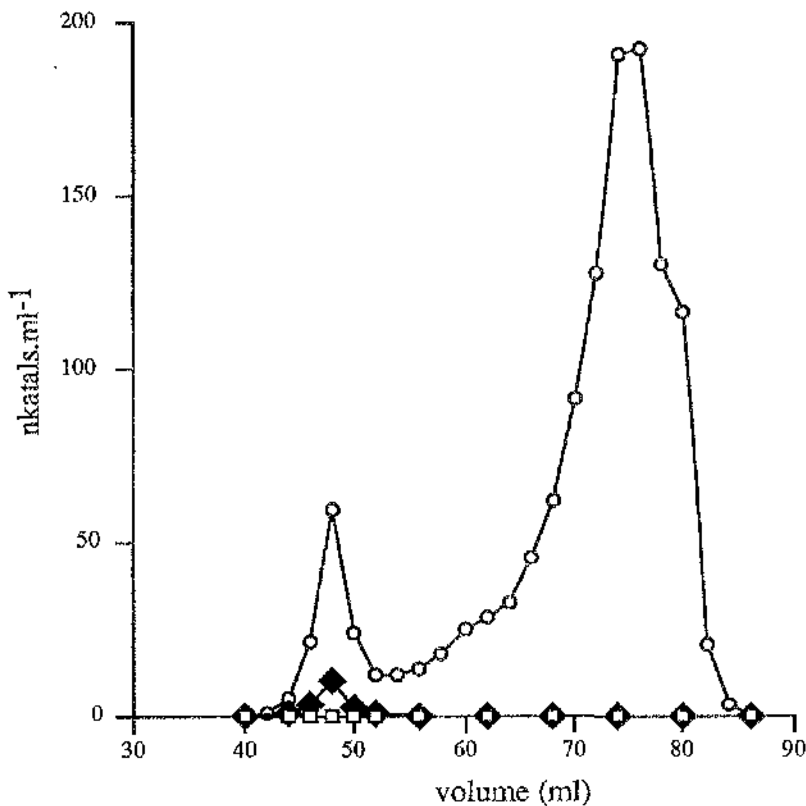


Fig. 4.8: Elution profile from a Superdex 200 gel filtration column of PDC, OGDC and E3 activities from a potato mitochondrial extract
Potato tuber mitochondrial extract (3ml) was applied to a Superdex 200 gel filtration column (100 ml bed volume) and protein eluted with 50 mM imidazole, pH 6.8 containing 1 mM DTT. E3 , OGDC and PDC activities were recorded as $nkatals.ml^{-1}$.

○ E3 activity ◆ PDC activity □ OGDC activity

release from pea mitochondria (Bourguignon *et al.*, 1988; Walker & Oliver, 1986a); therefore, the L-protein would be expected to elute as 'free' enzyme.

It was possible that the Superdex gel filtration step was promoting the instability of the complexes due to excessive dilution of the extract. Accordingly, a smaller (25 ml) column was used, which led to poor resolution of the E3 peaks. PDC and OGDC activities were still found in the same ratios as previously. Reconstitution assays were carried out in the hope that the poor recovery of OGDC activity by gel filtration was due to the complete dissociation of the E3 component. In this case the E1/E2 would still be present in the void volume fraction. Although PDC was also present in the sample, differential binding of the three E3 isoforms should be recognisable. However, all three potato E3 isoforms failed to stimulate OGDC and PDC activity when added to the void volume fraction.

4.2.6 Partial Purification of OGDC from Cauliflower Mitochondria

It was possible that potato tubers, being storage organs, naturally contained less OGDC than other plant tissues. Hence, as the isolation procedure from potato tuber mitochondria had been unsuccessful, it was decided to use a modified method of mitochondrial purification by Poulson and Wedding (1970) from cauliflower florets. This is the only known report of the partial purification of OGDC from plants and, importantly for these experiments, intact activity had been obtained by reconstitution with porcine heart E3. The method is described in Materials and Methods (section 2.2.9.6). The procedure by which Poulson and Wedding had isolated the E1/E2 core of OGDC, by a series of ammonium sulphate precipitation, was not repeated since optimisation of the method was impossible in the time available. Instead, the isolated mitochondria were extracted with a 30 min incubation at 4°C in the presence of 0.1% (v/v) Triton X-100. The crude extract was assayed for enzyme activity and then reconstitution assays were performed with increasing concentrations of porcine heart E3 to try to stimulate OGDC and PDC activity. Fig. 4.9 presents the results of reconstitution with porcine heart E3.

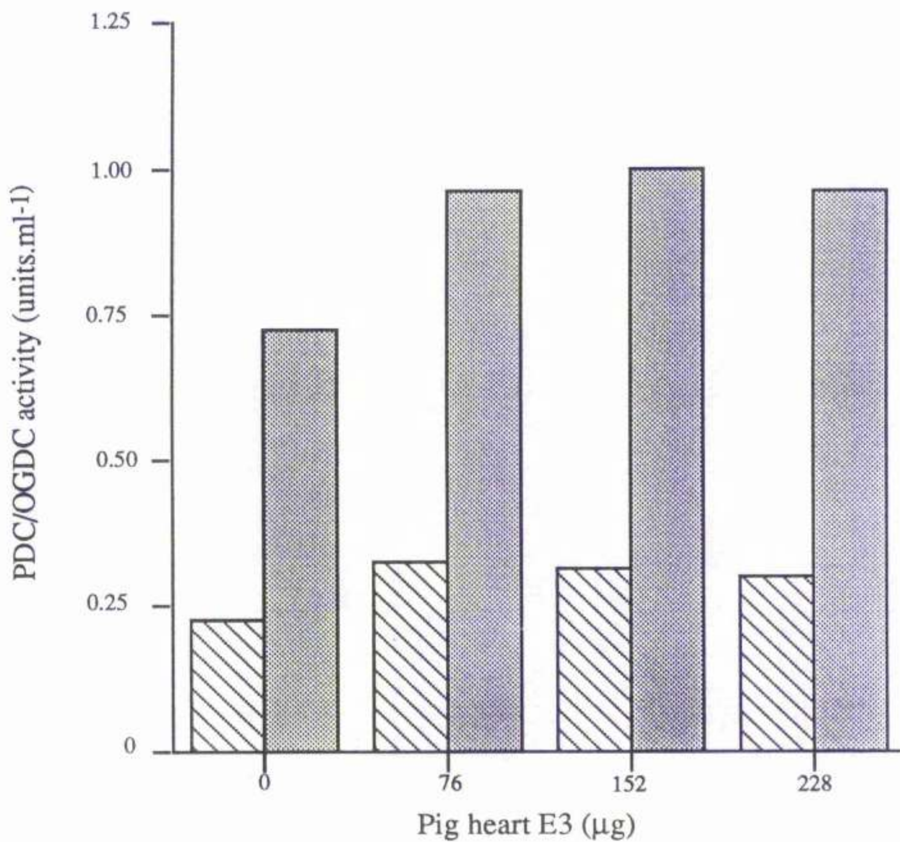


Fig. 4.9: Stimulation of cauliflower mitochondrial OGDC and PDC by reconstitution with increasing amounts of porcine heart E3

A fixed amount of mitochondrial extract was incubated at 30°C for 20min with increasing quantities of pig heart E3. Assays for PDC and OGDC were carried out as in section 2.2.10.1. Insufficient mitochondrial material precluded duplicate assays.

▩ PDC activity ⧅ OGDC activity

Levels of recordable OGDC activity were low but the ratio of OGDC to PDC activity (1:3) was readily measurable, in comparison with the potato tuber extract. Approx. 1.3-fold stimulation of OGDC and PDC activity is evident with excess porcine E3, indicating that the low levels of PDC and OGDC activity are unrelated to dissociation of the E3 component.

Confirmation of the presence of intact complexes was obtained by dissociating the E3 component from the core with high salt levels and then reconstituting intact activity with porcine heart E3. The extract was incubated in the presence of 1 M NaCl for 30 min at 4°C to remove the E3 components from the complexes. The extract was then centrifuged at 80 000 g for 16 h and the pellet, containing the E1/E2 core of OGDC and E2/X core of PDC, was resuspended in 20 mM potassium phosphate buffer, pH 7.6. As illustrated in Fig. 4.10, salt treatment led to a 75% decrease in OGDC activity. Significantly, E3, at a level of 15 μg, produced a 3-fold stimulation in activity to 70-80% of original activity but surprisingly, reconstitution with the potato E3 isoenzymes did not promote any stimulation in activity (data not presented). Similar results were found when PDC was assayed in the same sample.

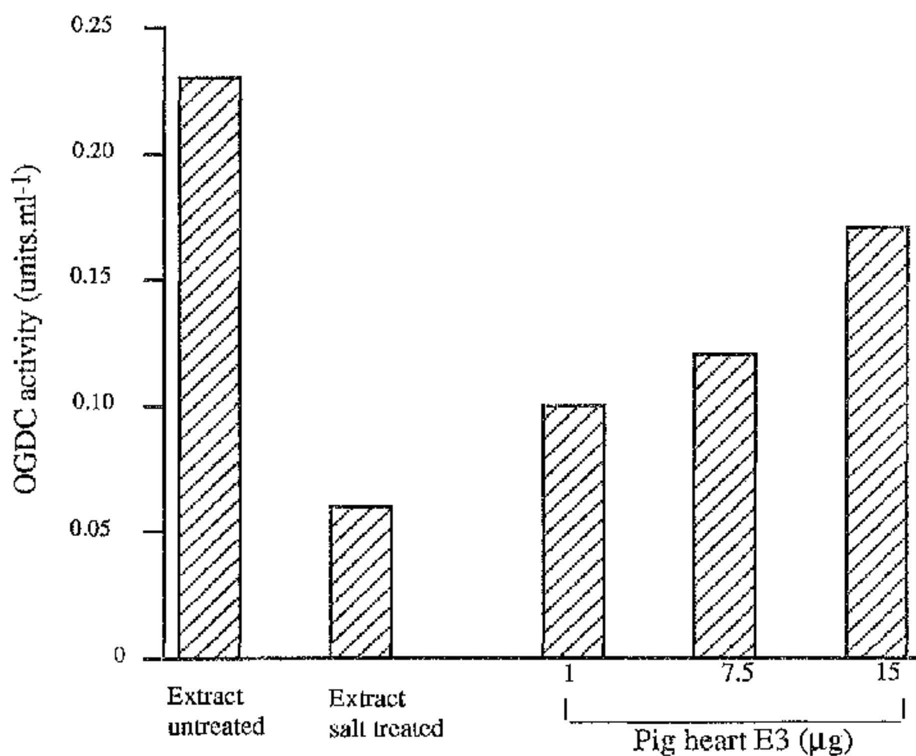


Fig. 4.10: Reconstitution of OGDC activity from cauliflower mitochondria with increasing amounts of pig heart E3

Salt treated (1 M NaCl) mitochondrial extract was centrifuged at 80 000 g to pellet the E1/E2 core leaving the E3 component in the supernatant. Following dialysis to remove the salt, a fixed amount of E1/E2 core was incubated at 30°C for 20 min with increasing amounts of pig heart E3. Reconstituted OGDC activity was plotted against increasing amounts of pig heart E3. Insufficient mitochondrial material precluded duplicate assays.

4.3 Discussion

The reconstitution experiments with the E1/E2 core of bovine heart OGDC highlighted the contrasting abilities of a number of different E3 to stimulate OGDC activities. In spite of the extensive homologies between bovine mucosal, porcine heart and yeast E3s, these E3 preparations gave poor levels of reconstitution compared with native bovine heart OGDC and PDC E3s. The affinity of binding was high as maximal activity was obtained with stoichiometric levels of E3 but only 15-20% restoration of OGDC activity was sustained by porcine heart and bovine mucosal E3. Surprisingly yeast E3 was unable to promote any reconstitution, although competition studies showed that there was a high affinity of association between this E3 and the bovine heart E1/E2 core. In each case, the specific activities of the individual E3s were similar and their purity was confirmed by SDS-PAGE. Subtle differences in the interaction of heterologous E3s to the bovine heart OGDC core may explain the low levels of OGDC activity. Thus, imprecise binding orientations with heterologous E3s may prevent the optimal interaction of the flexible lipoyl domains of E2 with the active sites of E1 and E3. The result with bovine mucosal E3 was unexpected because the view currently held by the majority of researchers is that a single E3 in mammalian species serves all the complexes and should, therefore, be interchangeable. A similar situation was later found with E3 binding to the bovine heart PDC core (Sanderson *et al.*, 1996b) indicating that the results obtained with OGDC were not artefactual.

Complex specificity of the E3 isoforms from potatoes was not demonstrated by the reconstitution experiments using the E1/E2 core of bovine heart OGDC. Reduction of OGDC activity through competition of the potato isoforms with the native E3 was low, even in the presence of a large excess of E3. These data suggest that binding of the plant E3 is quite different to that of the mammalian or yeast E3, with only low affinity interactions occurring with the bovine OGDC core. This is consistent with the fact that potato E3s were known to bind to the

immobilised E1/E2 subcomplex of OGDC, which could be employed in their purification. However, the orientation and low affinity of binding are clearly incompatible with an ability to support the activity of bovine heart OGDC. The yeast E3 behaved in a similar way to the potato E3 isoforms as it was unable to promote detectable reconstitution; however, unlike the plant E3s, it exhibited a high affinity binding to the bovine heart core as evidenced by its ability to compete with the parent E3 enzyme (bovine heart E3).

Reconstitution experiments with the OGDC core from potatoes were impossible to conduct owing to difficulties in obtaining sufficient amounts of active complex from this source. Measurable recoveries of OGDC were obtained from cauliflower but were still approx. one third of PDC activities. The presence of low OGDC activities has been noticed by several researchers working with the plant 2-oxoacid dehydrogenase complexes and the reason for this is unclear since both PDC and OGDC are present at similar levels in mammalian cells. Again reconstitution with the potato E3 isoforms failed to occur with the E3-depleted cauliflower core, although it was shown to occur with porcine heart E3. The ability of porcine heart E3 to stimulate cauliflower OGDC activity is surprising in view of the lack of success with the potato E3 isoforms. The discovery of distinct mitochondrial E3 isoforms in some plants (potato and barley) and the presence of single E3s in other plants (pea and turnip) seems to place plants in two groups. Cauliflower appears to possess only one E3 for GDC and all the 2-oxoacid dehydrogenase complexes, which may have evolved differently from the potato isoforms. These observations may offer a possible explanation for the failure of specific potato E3 isoforms to promote reconstitution of the cauliflower OGDC.

Further work on the purification of OGDC from plants was not feasible in the time available but continuing work in this laboratory recently suggested complex specificity of the isoforms using PDC isolated from potato mitochondrial extracts (R. Cook, unpublished observations). The modified purification method maintained the initial mitochondrial extract in a concentrated state by the use of a

small (10 ml) Sephadex G-200 gel filtration column which also separated the dissociated E3 component. Preliminary reconstitution studies showed that both the $\alpha\beta$ and the β_2 isoforms appeared to promote recovery of overall PDC activity to 40-50% of original levels, but much lower levels of the β_2 isoform were required, indicating that it bound to the PDC core with much higher affinity. The α_2 form, on the other hand failed to reconstitute PDC activity. The conclusion reached was that the β_2 form was most likely to be the complex specific form of PDC. The specificity of β_2 was extended to OGDC but levels of this complex, extracted by the modified method, were still found to be low. The roles of the α_2 and the $\alpha\beta$ form are less clear, but there is some evidence for tissue specific expression, with α_2 predominating in leaves. It still needs to be elucidated whether it represents a GDC specific form as this complex is present at high levels in leaves, representing 30-50% of mitochondrial matrix proteins.

The direction of the project was altered at this stage, in view of the difficulties, highlighted in these chapters, in working with the plant 2-oxoacid dehydrogenase complexes. The isolation of the gene(s) encoding the α and β polypeptides of the potato E3 isoenzymes was attempted since the N-terminal sequence of the two polypeptides was known (Fullerton *et al.*, 1996) and forms the subject of the following chapters. With regard to the complex specific nature of the E3 isoforms, the exact amino acid sequence would allow the determination of the relationship between the α and β polypeptides and the sequence comparison with other cloned E3s of known complex specificity. It would also permit their subsequent over-expression for further structure-function analysis.

Chapter 5

Cloning the Potato Mitochondrial Dihydrolipoamide Dehydrogenase (E3) Gene(s)

5.1 Introduction

Chapter 4 was concerned with reconstitution studies of the purified isoforms of potato mitochondrial dihydrolipoamide dehydrogenase (E3) with the E1/E2 cores of 2-oxoglutarate dehydrogenase (OGDC) from bovine heart and plant mitochondria. Attempts were made to assess possible complex specific roles of the mitochondrial α_2, $\alpha\beta$ and the β_2 isoforms in potatoes. However, as this was not possible owing to the difficulties in assay, purification and reconstitution of the complexes, it was decided to approach the question by sequence comparison of the gene(s) encoding the α and β subunits in order to determine their relationship with each other.

At the time of this work a number of prokaryotic and eukaryotic E3 genes had already been cloned and sequenced, yielding much information on domain organisation and amino acid identities with other pyridine nucleotide-disulphide oxidoreductases (Carothers *et al.*, 1989). Human erythrocyte glutathione reductase (GR) was the first enzyme of this family to be cloned and sequenced. Its domain structure and the role of specific amino acids in catalysis was elucidated from crystallographic data (Thieme *et al.*, 1981; Krauth-Siegel *et al.*, 1982) and has been extremely valuable in predicting the domain organisation of various E3s, since there is considerable homology at the primary amino acid sequence level.

In *E. coli* a single gene (*lpd*) has been cloned which encodes the E3 component for PDC and OGDC (Stephens *et al.*, 1983). It was found to be part of an operon containing the *aceE* and *aceF* genes for E1 and E2 of PDC. The OGDC E1 and E2 genes (*sucA* and *sucB* respectively) were subsequently found in a separate operon. All three genes of the PDC operon can be transcribed in a single transcriptional unit (Spencer & Guest, 1985) from a promoter located upstream from the first gene, *aceE*. However , the *lpd* gene can be transcribed from its own promoter allowing excess E3 production to complement with OGDC. The *lpd* gene of *A. vinelandii* and *Pseudomonas fluorescens* is located downstream of the

succinyltransferase gene (E2 of OGDC) and not in the PDC operon (Benen *et al.*, 1989; Westphal & Kok, 1988).

The situation is complicated in other bacterial species. For example, *Pseudomonas putida* contains three different E3 enzymes encoded by separate genes. The *lpd-val* gene is part of a tightly linked operon encoding all the components of the BCDC (Sykes *et al.*, 1987) and *lpd-glc*, providing E3 for PDC, OGDC and GDC, is part of the OGDC cluster. A third gene (*lpd-3*) was found not to be part of a 2-oxoacid dehydrogenase operon (Burns *et al.*, 1989a). A similar arrangement exists in *P. aeruginosa* where the presence of two genetically distinct E3 enzymes has been proved (McCully *et al.*, 1986). In the strict aerobe, *Alcaligenes eutrophus*, two distinct E3 enzymes were also detected; one was part of the PDC operon (*pdhL*) (Hein & Steinbuchel, 1994) and the other on the OGDC operon (*odhL*) (Hein & Steinbuchel, 1996). Remarkably, the *pdhL* product carries a lipoyl domain at the N-terminal region.

Less is known about the genetic aspects of E3 from eukaryotes. *S. cerevisiae* is similar to *E. coli* in having a single E3 structural gene (*lpd1*) for PDC and OGDC (Ross *et al.*, 1988; Roy & Dawes, 1987). Otulakowski & Robinson (1987) concluded that the human gene encodes a single gene product representing L-protein of GDC and the E3 of the 2-oxoacid dehydrogenase complexes.

The only E3 to be cloned and sequenced in plants is that of L-protein, the dihydrolipoamide dehydrogenase component of GDC, from pea leaf mitochondria (Bourguignon *et al.*, 1992; Turner *et al.*, 1992). The cDNA encoded a 501 amino acid polypeptide including a 31 amino acid presequence. Mature L-protein was predicted to consist of 470 amino acids, giving a protein of M_r 49 721 (or 50 441 if the FAD cofactor is included). The sequence reported by Turner *et al.* (1992), produced a C-terminal extension to the open reading frame (ORF) of 28 residues resulting from an extra nucleotide at position 1535, but otherwise was identical to the sequence of Bourguignon *et al.* (1992). Re-sequencing confirmed that the extra nucleotide was an error and the cDNA sequences obtained by the two groups were

derived from the same gene. There is strong evidence in pea for the presence of a single E3 gene involved in the 2-oxoacid dehydrogenase complexes and GDC (Bourguignon *et al.*, 1996). The overall sequence similarity between the L-protein and human E3 is very high (70%); therefore, it was expected that the secondary and tertiary structure of L-protein would be similar to the human enzyme.

The wealth of sequence information has permitted comparisons of the domain organisation and tertiary structures of E3s from various organisms as discussed in Chapter 1 (1.3.4.1). All E3 enzymes whose sequences have been reported have the same domain structure as GR, as exemplified by the E3 polypeptide from *E. coli* which was subdivided into an FAD-binding domain, an NAD^+- binding domain, a central domain and an interface domain. Characteristic sequence motifs have been identified which are highly conserved within the domains of the family of pyridine nucleotide-disulphide oxidoreductases (Carothers *et al.*, 1989). These are the FAD-binding fold (residues 7-37), the region around the two redox-active cysteine residues found in the FAD-binding domain (cys-44 and cys-49), the NAD(H) binding fold (residues 176-212) and the interface region (residues 430-462). The 'fingerprint' structural motif in the dinucleotide-binding fold of either FAD or NAD is manifested as a $\beta\alpha\beta$ fold.

In contrast to the single E3 purified from pea leaf mitochondria, three separate isoforms of E3 were isolated from potato tuber mitochondria by our group (Fullerton *et al.*, 1996). As discussed in chapter 3, analysis by SDS-PAGE showed that the first isoform to elute from a FPLC Mono Q column was an alpha homodimer (α_2), the second an alpha/beta heterodimer ($\alpha\beta$) and the third a beta homodimer (β_2). Preliminary data suggest that they may have distinct kinetic and biochemical properties and heat stabilities and consequently it has been tentatively proposed that they represent complex specific forms of E3 (R. Cook, unpublished observations). However, the two polypeptides, α and β are known to be very similar to each other because they cross-react strongly with yeast antisera, have identical N-terminal sequences and are very close in molecular mass (as estimated

by electrospray mass spectrometry) differing by only approx. 116 Da. Furthermore, the N-terminal sequence was identical to that reported for L-protein from pea mitochondria apart from one conservative substitution at residue 11. Therefore, it was unknown whether they represented distinct gene products or were a result of alternate splicing or covalent modification.

The results presented in the next section concern the attempted cloning and sequencing of cDNAs encoding E3 from potatoes. The starting material was total RNA from potato leaves, the sole reason being ease of preparation in comparison to potato tuber RNA. It was valid to employ RNA purified from potato leaves since the three isoforms were also expressed in the mitochondria of this tissue (A. Carmichael, Ph.D Thesis, Glasgow University, 1994), although it was known that the α_2 homodimer predominated in this tissue as opposed to the β_2 homodimer in tubers.

5.2 Strategy for Cloning the E3 Gene

A PCR-based strategy was employed for the amplification of the cDNA generated from the mRNA of potato leaves. The availability of the N-terminal amino acid sequence allowed for gene specific primer design. Firstly, the exact nucleotide sequence of the 5′ end of the gene was determined to permit design of E3 gene specific primers for future amplification of the entire E3 cDNA. This was accomplished by amplifying the region of cDNA using highly degenerate primers based on the N-terminal amino acid sequence and cloning the product.

Once the nucleotide sequence at the extreme 5′ end was known, specific nested primers were designed to amplify the entire gene in conjunction with a universal primer at the 3′ end of the cDNA. Initially *Pfu*, a proofreading DNA polymerase (Stratagene) was used to obtain faithful copies of the gene but problems were encountered with the cloning procedure using this product. Subsequently, a *Taq* amplified product was cloned and sequenced successfully. As a result of the cloning strategy, the *Taq* amplified cDNA lacked the 5′-sequence encoding the

ATG translation start site, the leader sequence and the first seventeen amino acids of the mature protein. Originally, it was intended to obtain this sequence by 5′ RACE (Rapid Amplification of 5′ cDNA ends) of the first strand cDNA; however, as described in chapter 7, it was obtained by the isolation and sequencing of full-length potato leaf cDNA library clones.

5.3 Results

5.3.1 RNA Preparation from Potato Leaves

The initial phase of this project involved preparation of good quality, undegraded total RNA for the synthesis of full length, double-stranded cDNA. An mRNA separation step from the total cellular RNA was not required for this purpose as an oligo-dT primer was used to prime the synthesis of the first-strand cDNA.

RNA was extracted from potato leaves (*Solanum tuberosum*, cv. Maris Piper) in the presence of guanidinium isothiocyanate, a powerful protein denaturant which served to denature endogenous nucleases and deproteinate the RNA. Denaturation was enhanced by the addition of 2-mercaptoethanol to break intramolecular protein disulphide bonds. Phenol/chloroform extraction was used to remove the degraded protein and RNA was precipitated with salt and isopropanol.

The quantity of the RNA was assessed by measuring its absorbance at 260 nm (3.87 $\mu g.\mu l^{-1}$) and its purity by measuring the A_{260}/A_{280} ratio. A value of 1.74 was obtained which indicated pure RNA (a ratio of 1.8-2.0 is generally appropriate for pure RNA). The preparation was judged as being relatively undegraded as discrete bands representing 28S and 18S rRNA (approx. 5 kb and 2 kb respectively) were observed (Fig. 5.1). As expected, a heterogeneous smear of mRNAs ranging in size from 0.5 kb to over 10 kb was present. Ribosomal RNA is the most abundant species, representing 80-90% of total cellular RNA. In contrast mRNA represents only 1-5% of total cellular RNA.

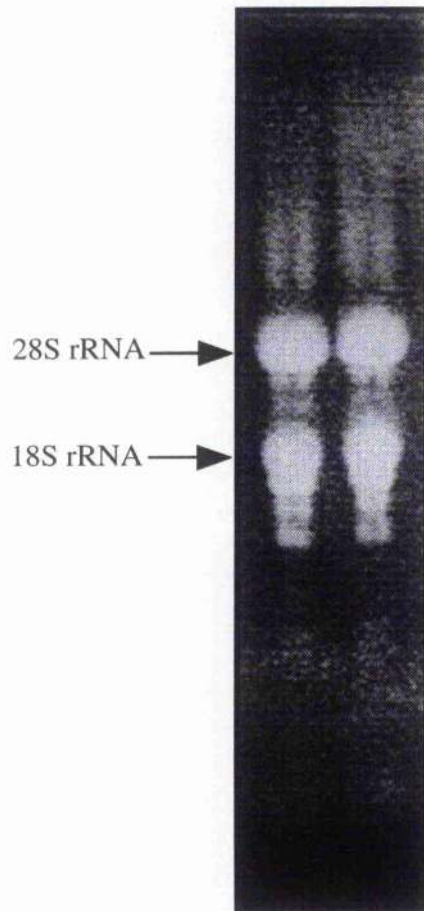
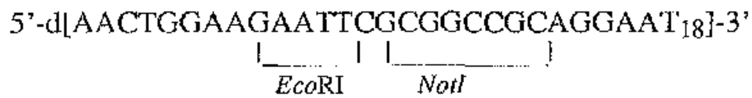


Fig. 5.1: Analysis of the RNA preparation from potato leaves

RNA was prepared from frozen potato leaves as described in Materials and Methods (section 2.2.28). Approx. 10 μg of RNA were electrophoresed in duplicate on a 1% (w/v) denaturing formaldehyde/agarose gel and stained with ethidium bromide following the method in 2.2.17.

5.3.2 First-Strand cDNA Synthesis

Synthesis of first-strand cDNA (as described in section 2.2.23.1) was primed using the *Not*I-d(T)$_{18}$ bifunctional primer corresponding to polyadenylated mRNA. It was termed 'bifunctional' because it was composed of a string of d(T) residues as well as an 'anchor' domain which contained a *Not*I and an *Eco*RI restriction site for cloning. The sequence of the primer was:



Another primer (*Not*I-G), lacking the oligo d(T) portion, was made because the *Not*I-d(T)$_{18}$ primer does not always function efficiently for subsequent PCR of the first-strand cDNA. The completed first-strand cDNA reaction of 33 µl was used directly for conversion to double-stranded cDNA and amplification by PCR as described in the next section.

5.3.3 PCR Amplification of the 5′-Region of the ORF of the Potato Mitochondrial E3 Gene

First-strand cDNA was amplified by PCR using degenerate primers designed from the known N-terminal sequence of potato mitochondrial E3 (Fig. 5.2). For cloning purposes an *Eco*RI site was added to the 5′ end of primer A and a *Bam*HI site to primer B. Primer sequences are shown in Table 5.1.

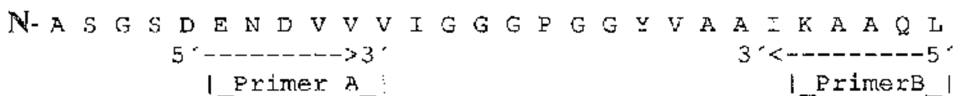


Fig. 5.2: N-terminal sequence obtained for the potato mitochondrial α and β polypeptides, showing positions of the degenerate primers A and B.

Table 5.1

Sequence of the forward and reverse primers used to amplify the 5´ region of the potato mitochondrial E3 gene

Primer	Sequence (5´--->3´)	Product size	Degeneracy
A (forward)	G GAATTC GAC GAA AAC GAC GTN GT T G T T (variant bases under GAC, GAA, AAC, GAC)	80bp	64-fold
B (reverse)	CG GGATCC AGT TGN GCN GCT TTT AT C C (variant bases under AGT, GCT)		64-fold

a)N= A, T, G and C

b)Restriction sites underlined

5.3.3.1 Amplification of First-Strand cDNA

The first-strand cDNA product was used as a template to amplify double-stranded cDNA corresponding to amino acids 5 to 28 of the N-terminal sequence. Primers A and B were present at 1000 pmol instead of 10 pmol in each 50 µl reaction as they were highly degenerate. The PCR mix is described in section 2.2.22 and the cycling parameters are described in Fig. 5.3.

The products of PCR amplification with primers A and B are illustrated in Fig. 5.3, lane 3. The product at approx. 80 bp was assumed to be the product of interest as it corresponded to the expected size. Lanes 1 and 2 show the products generated with primer A or primer B respectively. This control, carried out to eliminate non-E3 specific products, found that neither primer amplified a 80 bp product on its own. There was a contamination in the negative template control (Lane 4); however, it was not considered to be a problem as the 80 bp product was well separated from it for future gel purification.

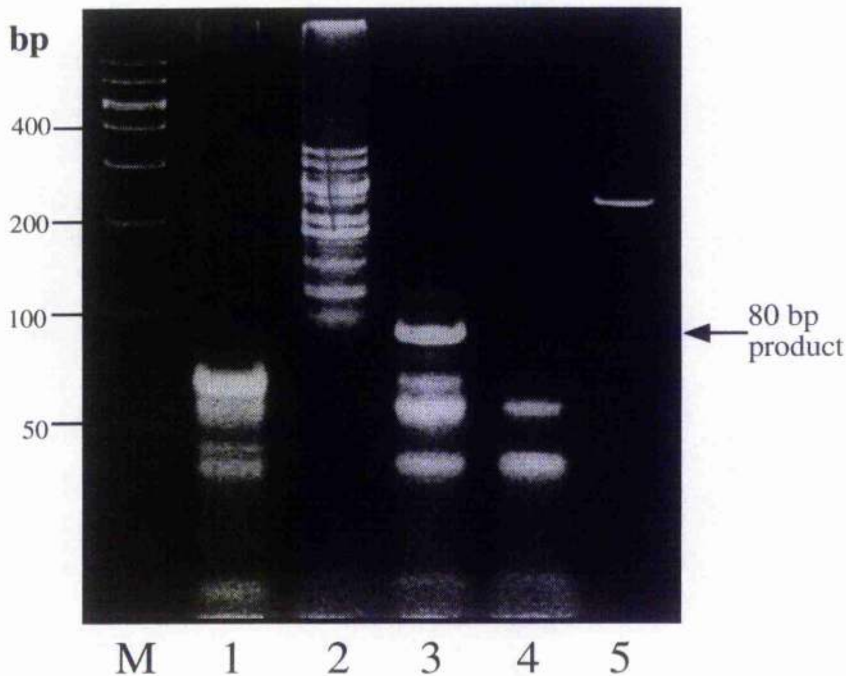


Fig. 5.3: Analysis of PCR products obtained by *Taq* polymerase amplification of first-strand cDNA

Potato leaf RNA served as a template for first-strand cDNA synthesis as described in Materials and Methods (section 2.2.23.1). The 5'-region of the first-strand cDNA (5 µl) was then converted to double-stranded cDNA and amplified by PCR (section 2.2.23.2) in the same reaction. DNA synthesis was primed using 1000 pmol of each degenerate primer, A and B, shown in Table 5.1, and in addition contained 1 mM $MgCl_2$. The cycling was as follows: 1 x 5 min at 95°C, 30 x [30 s at 55°C, 30 s at 65°C and 30 s at 95°C], 1 x 2 min at 65°C and 1 x 3 min at 72°C. Products were separated on a 12% (w/v) polyacrylamide gel

M: size marker 1: primer A only, 2: primer B only, 3: primers A+ B, 4: negative DNA control, 5: Positive DNA control (pBluescript DNA and M13 forward and reverse primers from Pharmacia).

5.3.3.2 Cloning and sequencing the 80 Base Pair Fragment

The purified product was cloned directionally into the *Eco*RI and *Bam*HI sites of pBluescript KS+ transformed into *E. coli* XLI Blue cells. Nine colonies out of fourteen colonies were found to contain an insert of the correct size (80 bp). Fig. 5.4 presents six of the recombinant plasmids (named pBkA, pBkB, pBkC, pBkD, pBkF and pBkG) after digestion with *Eco*RI and *Bam*HI.

Automated sequencing (section 2.2.33) was carried out on the double-stranded recombinant plasmid template to confirm that the correct PCR product had been cloned and to provide the exact nucleotide sequence for further primer design. Each of the four dyedeoxy terminators (ddNTPs) were tagged with a different fluorescent dye. Thus, the growing chain was simultaneously terminated and labelled with the dye corresponding to that base. Each dye emitted light at a different wavelength when excited with laser light. This had the advantage that four sequencing reactions could be performed in one tube and all four reactions could be detected and distinguished in a single gel lane.

The sequence of the clones corresponding to amino acids 5-28 of the mature E3 enzyme is shown in Fig. 5.5.

```
        ___Primer A______ _______SP1_____   ______NSP1_____
        gatgagaatgacgtggttgttatcggtggtggtcccggcggctatgtggcggcgatt
        ---------+---------+---------+---------+---------+-------
        ctactcttactgcaccaacaatagccaccaccagggccgccgatacaccgccgctaa
A S G S D  E  N  D  V  V  V  I  G  G  G  P  G  G  Y  V  A  A  I

         aaagccgcccaact
         --+---------+-
         tttcggcgggttga
         K  A  A  Q
```

Fig. 5.5: DNA sequence of clones pBkC, pBkD and pBkF representing amino acids 5-28 of the mature E3 protein

The first 4 amino acids from the N-terminal sequence, not used for primer design, are shown in italics. Sequences chosen for production of unique nested primers, SP1 and NSP1, are indicated.

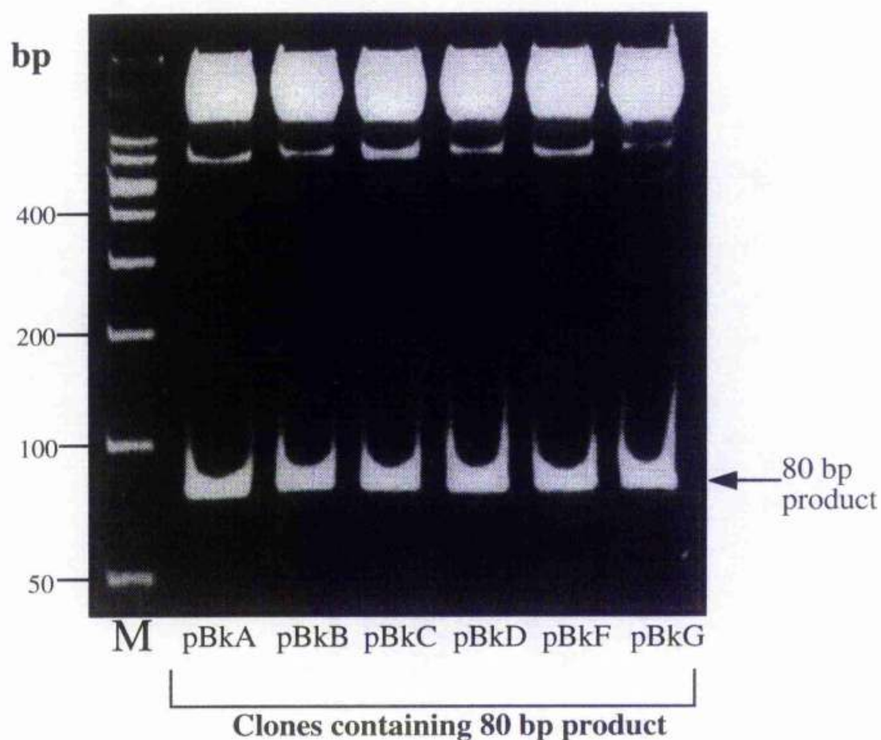


Fig. 5.4: Analysis of clones containing the 80 bp insert

Recombinant pBluescript (KS+) plasmids containing the 80 bp product, (ligated into the *Eco*RI and *Bam*HI site), were transformed into XL1 Blue *E. coli* cells as described in section 2.2.27.1. Cells were grown on LB (amp/tet) plates and positive colonies selected by inactivation of β-galactosidase activity. Purified recombinant plasmid DNA (1 μg) was digested with *Eco*RI and *Bam*HI and separated on a 12% (w/v) polyacrylamide gel.

M= size marker

5.3.4 Nested-PCR Amplification of the Full Length cDNA Encoding the Potato Mitochondrial E3

A nested-PCR approach was employed to permit the specific amplification of the E3 cDNA from the first-strand cDNA sample. This approach was necessary to identify the relevant cDNA from a number of PCR products. The method requires PCR amplification with one primer set followed by a subsequent PCR using a second primer set, internal to the amplified DNA. The primers were expected to be specific as they were designed from the sequenced 5′ end of the open reading frame of the gene, described in section 5.3.3.4.

5.3.4.1 Design of the Nested-PCR Primers

Two N-terminal sense primers, SP1 and NSP1, were designed downstream of the original sequence corresponding to primer A to avoid the ambiguity of the codons specified in this region owing to the degeneracy of primer A (Fig. 5.5 and Table 5.2). The first primer set, used to amplify the E3 gene from the first-strand cDNA reaction, was SP1 and the *Not*1-d(T)$_{18}$ reverse primer and the second primer set was NSP1 and the *Not*I-d(T)$_{18}$ antisense primer. Hence the PCR product obtained from the nested-PCR reaction was expected to start at the codon specifying amino acid 18 (glycine) of the mature E3 protein as indicated in Fig. 5.5.

Table 5.2

Sequence of the nested forward primers used to amplify the E3 cDNA from first-strand cDNA

Primer	Sequence (5′--->3′)	Approx. product size (kb)
SP1	CG GGATCC TGTTATCGGTGGTGGTCCC	1.8
NSP1	G GAATTC GC GGCTATGTGGCGGC	1.8

a) Restriction sites underlined

5.3.4.2 PCR Amplification of First-Strand cDNA

Amplification of DNA was catalysed with thermostable *Pfu* polymerase (Stratagene) as it had a 12-fold higher fidelity of DNA synthesis than *Taq* polymerase due to its 3′ to 5′ proof-reading exonuclease activity. High fidelity of synthesis was required as the product was to be cloned and the sequence of the gene obtained. Two sources of *Pfu* were available; one was the native enzyme from *Pyrococcus furiosus* and the other was recombinant *Pfu*. PCR was performed with both enzymes in order to compare the products.

The 'Hot Start' technique (detailed in section 2.2.22) was employed for PCR with *Pfu* to reduce the possibility of competing side reactions such as the amplification of non-target sequences. Mis-priming mainly occurs during pre-PCR setup when all reactants have been mixed at room temperature. Therefore, the reaction was designed so that all reactants do not mix until they have reached a high enough temperature to suppress primer annealing to non-target sequences. In addition, the 3′ to 5′ exonuclease activity associated with *Pfu* DNA polymerase, will begin to degrade the template DNA in the absence of nucleotides. Thus, it was necessary to add the polymerase last to the reaction mix.

The cycling conditions, stated in Fig. 5.6, began with an initial 90 min extension to ensure that the second strand of DNA was fully extended. Lanes 1 and 2 are the negative DNA controls with recombinant and native *Pfu* respectively. A smear of DNA, of unknown origin, was observed in these samples which was absent in the reactions amplifying the first-strand cDNA template (lanes 3 and 4). Surprisingly, a ladder of products was only observed with the native *Pfu* polymerase (lane 4). The size of cDNA product, estimated from the cloned pea mitochondrial E3, was expected to be above 1500 bp; therefore, the likely candidate was the 1.8 kb product. An increase in the annealing temperature from 60°C to 63°C, 65°C and 68°C resulted in a loss of amplification of the 1.8 kb product but the inexplicable smear of DNA in the negative controls was always present.

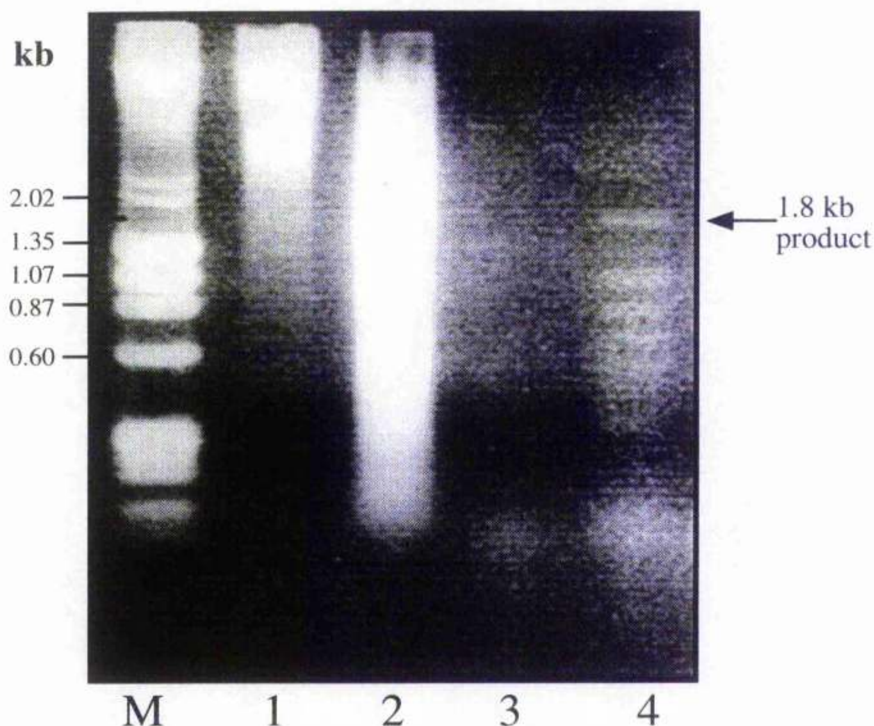


Fig. 5.6: PCR amplification of first-strand cDNA with *Pfu* (native and recombinant) DNA polymerase

First-strand cDNA, generated from potato leaf RNA as described in section 2.2.23.1, was primed with the E3 gene specific primer, SP1 (Table 5.2) and the *Not*I-dT$_{18}$ primer. The reactions contained 2.5 μl of first-strand cDNA and 10 pmol primers (in a 50μl reaction). Cycling conditions were as follows: 1 x [4 min at 95°C, 10 min at 55°C and 90 min at 75°C], 30 x [5 min at 60°C, 15 min at 75°C and 1 min at 95°C] and 1 x 90 min at 75°C. One fifth of the reaction was separated on a 1% (w/v) agarose gel.

M: size marker, 1: negative DNA control with recombinant *Pfu* polymerase, 2: negative DNA control with native *Pfu* polymerase, 3: PCR amplification with recombinant *Pfu* polymerase, 4: PCR amplification with native *Pfu* polymerase.

5.3.4.3 Nested-PCR Of The Amplified DNA Products

The sample previously amplified with the native *Pfu* polymerase (Fig. 5.6, lane 4) was re-amplified using the nested, forward primer, NSP1 (Table 5.2) and the *Not*I-G primer described in section 5.3.2. Once more, the activities of native and cloned *Pfu* were compared in case only one polymerase was able to amplify the target DNA. Only one product, at 1.8 kb was re-amplified (Fig. 5.7), confirming that it was likely to be specific for the E3 gene. Interestingly, this time the recombinant *Pfu* polymerase was effective, whereas the native *Pfu* polymerase had amplified the original product. Furthermore, the negative DNA control was completely clear.

Nested PCR on the previously amplified DNA (Fig. 5.6, lane 4) was also attempted with *Taq* polymerase as a comparison. Fig. 5.8 illustrates the separated products on a 1% (w/v) agarose gel. The PCR reaction, resolved in lane 3, re-amplified the 1.8 kb product. Lanes 1 and 2 are the controls containing only the NSP1 primer and the *Not*I-G primer respectively. As expected no product was amplified with the N*ot*I-G primer only; however, the NSP1 primer alone generated products of the same size as the 1.8 kb product. This was thought to be due to carry over of the *Not*I-d(T)$_{18}$ antisense primer from the DNA template which would have primed DNA synthesis in conjunction with the NSP1 primer. A few species of DNA with slightly different sizes could be distinguished on the gel because of the low concentration of product resulting from the presence of the limiting *Not*I-d(T)$_{18}$ contaminant. These were inconspicuous in lane 3 because of the high concentration of amplified DNA product. A possible explanation for this result was the occurrence of different polyadenylation sites for the E3 gene leading to heterogeneity in mRNA size. The d(T)$_{18}$ portion of the primer also had the potential to bind non-specifically to sites along the poly A tail producing products of varying lengths depending on where it primed DNA synthesis. On the other hand, if the α and β polypeptides were produced by different, but closely related genes then the two forms may have been simultaneously amplified.

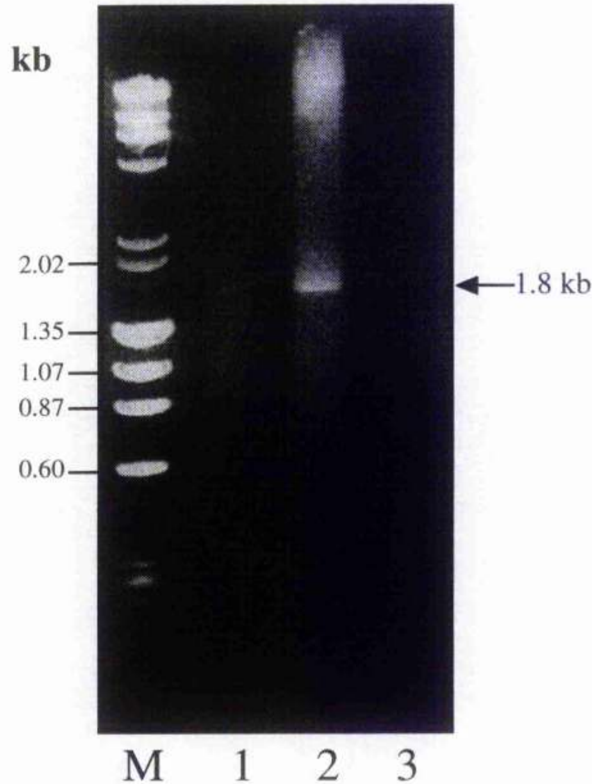


Fig. 5.7: Analysis of the potato leaf cDNA re-amplified by nested-PCR using native and recombinant *Pfu* polymerase

The DNA sample (1 µl), illustrated in Fig. 5.6 (lane 4), was re-amplified using the recombinant and native *Pfu* polymerases as described in section 2.2.22. The reaction was primed using 10 pmol of the E3, nested gene specific primer (in a 50 µl reaction), NSP1 (Table 5.2) and the *Not*I-G primer. Cycling conditions were as follows: 1 x 95°C for 3 min, 37 x [5 min at 55°C, 15 min at 75°C and 1 min at 95°C] and 1 x 90 min at 75°C. Products were separated on a 1% (w/v) agarose gel.

M: size marker, 1: PCR amplification with native *Pfu* polymerase, 2: PCR amplification with recombinant *Pfu* polymerase, 3: negative DNA control.

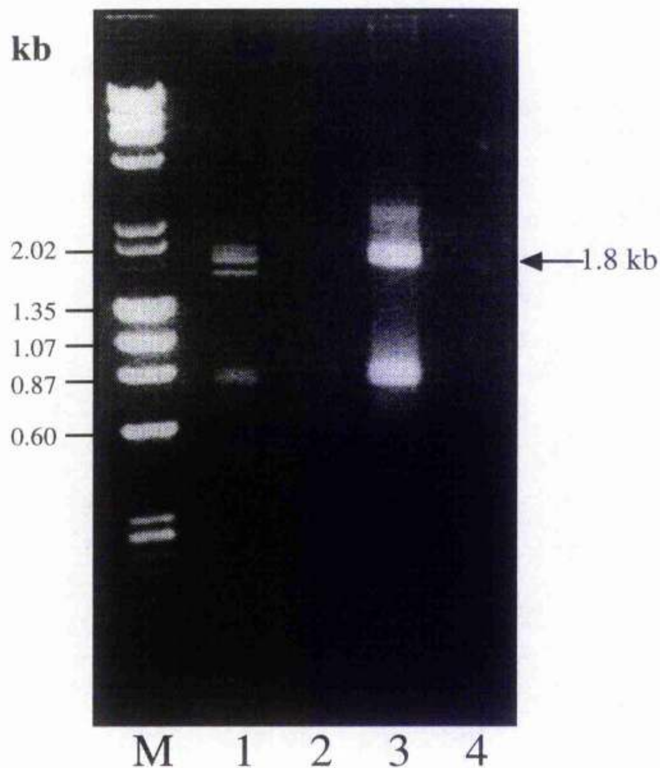


Fig. 5.8: Analysis of the potato leaf cDNA re-amplified by nested-PCR using *Taq* DNA polymerase

The DNA sample (1 µl), illustrated in Fig. 5.6 (lane 4), was re-amplified using *Taq* DNA polymerase. The reaction was primed with 10 pmol of the E3 nested gene specific primer, NSP1 (Table 5.2) and the *Not*I-G primer and carried out in the presence of 1 mM $MgCl_2$. Cycling conditions were as follows: 1 x [3 min at 95°C], 35 x [3 min at 55°C, 3 min at 65°C, 30 s at 95°C], and 1 x [6 min at 72°C]. One fifth of the products were separated on a 1% (w/v) agarose gel.

M: size marker, 1: NSP1 primer only, 2: *Not*I-G primer only, 3: both primers, 4: negative DNA control.

5.3.4.4 Cloning the Nested PCR Product

The purified 1.8 kb products, amplified by both *Pfu* and *Taq* polymerase, are visualised in Fig. 5.9. Significantly, the figure shows quite clearly that both the *Pfu* and *Taq* polymerase products are the same size. The products were ligated into the *Eco*RI site of pBluescript (KS+) following the protocol detailed in section 2.2.26.

E. coli DH5α cells were transformed with each ligation mix. Thirty positive colonies were isolated from the plates containing cells transformed with the *Taq* construct and in comparison only seven colonies were produced with the *Pfu* construct. Visualisation of the purified recombinant plasmids (Fig. 5.10), after digestion with *Eco*RI, revealed that colonies transformed with the *Taq* construct contained inserts of interest. However, the seven colonies transformed with the *Pfu* construct were found not to contain any inserts. The presence of four different sizes of insert was not surprising as the PCR with *Taq* and only the NSP1 primer (Fig. 5.8) had already suggested a number of species differing in size. Subsequent sequencing and expression studies were required to determine whether this was due to different polyadenylation sites, non-specific priming of the *Not*I-$d(T)_{18}$ primer or different cDNA products.

The cloning strategy with the *Pfu* amplified product was repeated using *Epicurian Coli* SURE supercompetent cells (Stratagene). This recombination deficient host strain carries mutations in DNA repair pathways directly involved in the rearrangement and deletion of DNA. Fifty colonies were isolated and analysis of the recombinant plasmid DNA, digested with *Eco*RI, revealed that part of the inserts had been deleted, leaving inserts ranging from 50 to 500 bp (Fig. 5.11). Owing to lack of time, further PCRs using Vent DNA polymerase (New England Biolabs, USA), another proof-reading enzyme, were not conducted. Analysis was continued with the clones containing the *Taq* amplified insert, although the fidelity of DNA polymerisation by *Taq* was a problem to be considered.

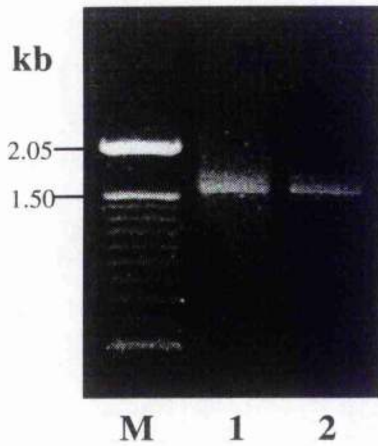


Fig. 5.9: Gel purified PCR products amplified with *Pfu* or *Taq* DNA polymerase for cloning purposes

M: size marker, 1: *Pfu* PCR product, 3: *Taq* PCR product.

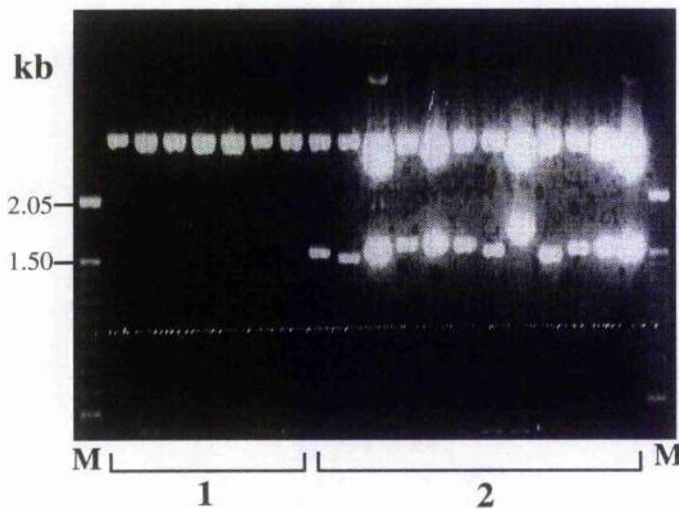


Fig. 5.10: Analysis of clones obtained by transformation with the recombinant *Taq* and the *Pfu* PCR products

The PCR products, illustrated in Fig. 5.9, were ligated to the *Eco*RI site of pBluescript (KS+) vector and transformed into *E. coli* DH5α cells as described in Materials and Methods (section 2.2.27.1). Cells were grown on LB (amp) plates and putative positive colonies were selected for a mini plasmid preparation (2.2.20). Plasmid DNA (1 μg) was digested with *Eco*RI to restrict out the inserts and separated on a 1% (w/v) agarose gel.

M: size marker, 1: putative positive *Pfu* clones, 2: putative positive *Taq* clones (showing only 12 of 30 obtained).

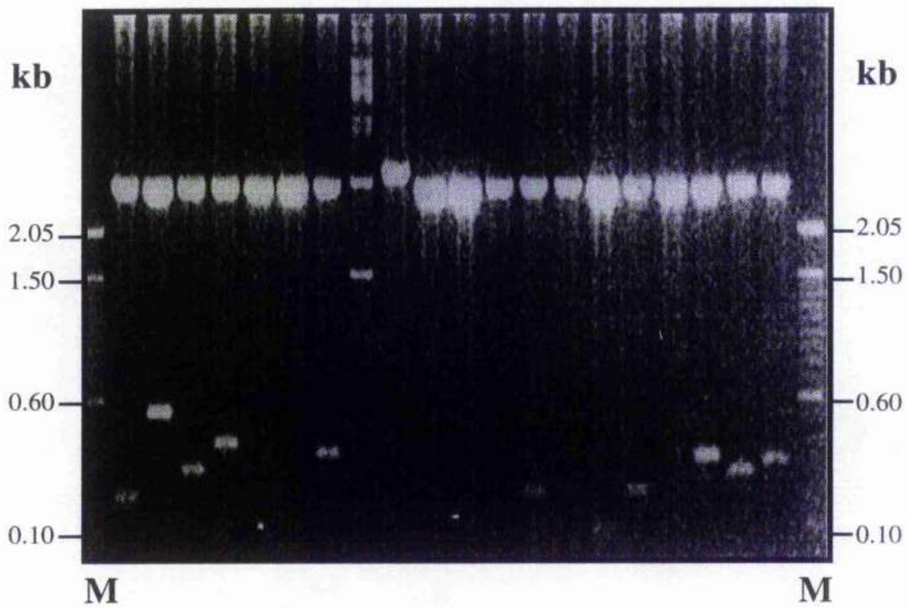


Fig. 5.11: Analysis of clones obtained by transformation of SURE *E. coli* cells with the recombinant *Pfu* PCR product

Twenty clones were isolated for a mini plasmid preparation of which 1μg was digested with *Eco*RI to restrict out the insert. DNA was analysed on a 1% (w/v) agarose gel.

M= size marker

5.3.4.5 Confirmation of Identity by Partial Sequencing and Restriction Mapping of the Clones Containing the *Taq* Amplified Insert

Eight of the clones containing the *Taq* amplified insert were chosen to represent the range of insert sizes. They were named pTaq13, pTaq14, pTaq15, pTaq16, pTaq22, pTaq23, pTaq25, and pTaq30. The inserts were sequenced from both ends using M13 universal and M13 reverse primers. This partial sequencing was required to confirm the identity of the clones and to determine their orientation in the plasmid since cloning was non-directional. Automated sequencing was performed as described in section 2.2.33, providing 350-400 bases of sequence from each end.

The first 30 bases of the sense strand were translated and the sequence was consistent with amino acids 18 to 27 of the N-terminal sequence of the E3 enzyme, shown in Fig. 5.5. The PCR with the M13 reverse primer provided sequence data from the 5′ end of the sense strand for clones pTaq15, 25, and 30, therefore these clones had been inserted in the sense orientation. In contrast pTaq13, 14, 16, 22 and 23 were found to be cloned in the antisense orientation as the M13 reverse primer allowed the sequencing of the antisense strand from the polyA tails.

Confirmation of the identity of the clones by analysis of their 5′ sequence was followed by their over-expression in bacteria, as will be described in chapter 6. This control was carried out to ensure that *Taq* DNA polymerase had faithfully amplified the cDNA and to determine if any of the clones encoded the α and β polypeptides.

The clones were restriction digested with *Eco*RI (the cloning site) and one of the following enzymes: *Bam*HI, *Pst*I, *Sal*I, *Hin*dIII and *Kpn*I in order to prepare a crude restriction map. The only enzymes found to cut within the insert were *Pst*I and *Kpn*I. Fig. 5.12 depicts an agarose gel of the pattern of digestion obtained for all eight clones with *Eco*R1 and *Pst*I or *Kpn*I. All the clones had two internal *Kpn*I sites generating three fragments. One fragment was relatively constant for all the

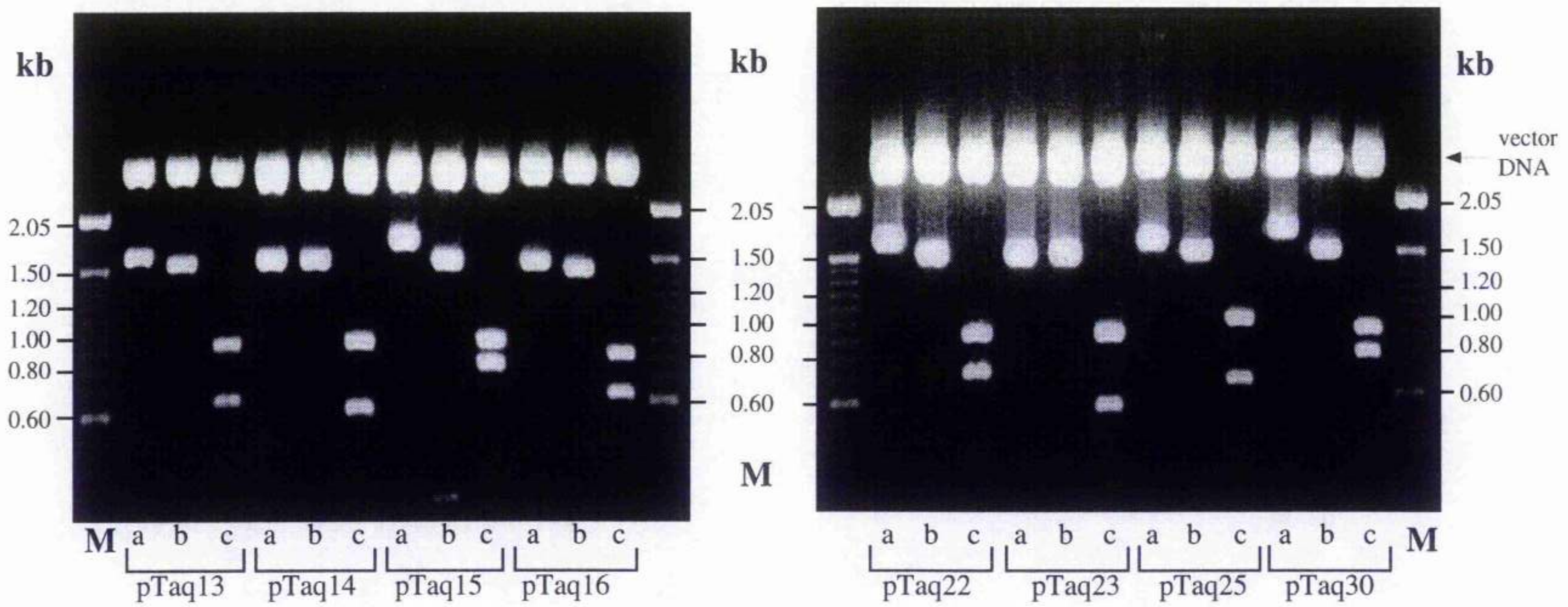


Fig. 5.12: Restriction digests of the recombinant plasmids containing the *Taq* PCR product

The recombinant plasmids (1 μg) were each digested with *Eco*RI (**a**), *Eco*RI and *Pst*I (**b**) and *Eco*RI and *Kpn*I (**c**) as described in Materials and Methods (section 2.2.15). Digested plasmids were separated on a 1% (w/v) agarose gel.

M= size marker

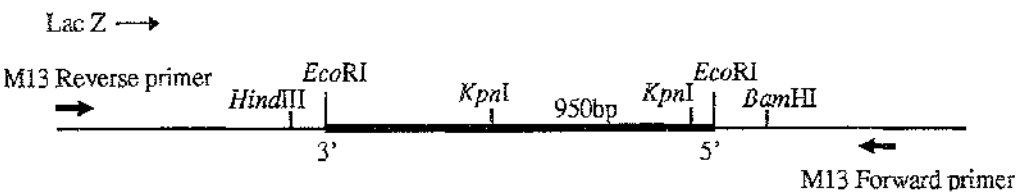


Fig. 5.13a: Restriction map of clone pTaq22 showing orientation of insert in pBluescript KS+

The map and orientation of the insert is representative of pTaq13, pTaq14, pTaq16 and pTaq23. The orientation of the plasmid is indicated by the direction of transcription of the Lac Z.

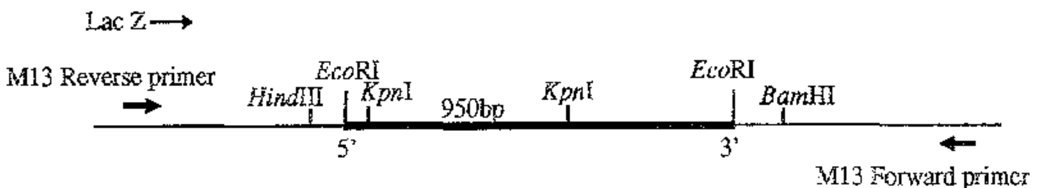


Fig. 5.13b: Restriction map of clone pTaq15 showing orientation of insert in pBluescript KS+

The map and orientation of the insert is representative of pTaq25 and pTaq30

clones and was roughly 950 bp, whereas the smaller fragment varied in size between the clones (600 to 800 bp). The last fragment was too small to be visualised but was known to exist by sequencing the ends of the clones described above. pTaq15 and 30 were thought to be the same clone, as were pTaq14 and 23 since they had the same digestion pattern. Therefore, there were six clones in total exhibiting obvious differences; they either had different polyadenylation start sites or differences at the *Kpn*I sites. A representative restriction map of pTaq22 with *Eco*RI and *Kpn*I is presented in Fig. 5.13A and that of pTaq15 in Fig. 5.13B.

5.3.5 Generation Of Clones For Sequencing By Directed Deletions

The DNA cloned into pBluescript (KS+) was too large to be sequenced conveniently from a single primer binding site. An efficient way to sequence large DNA inserts was to generate a nested set of deletions using Exonuclease III (Exo III) in the target DNA, thereby moving the priming site closer to the sequence of interest. Therefore, deletions are from the end of the insert, closest to the priming site chosen for sequencing. Expression studies (Chapter 6) had demonstrated that pTaq22 and pTaq16 encoded proteins with different M_r values, possibly corresponding to the α polypeptide and the β polypeptide respectively; thus, these two clones were selected for sequencing. In these cases, the M13 universal primer site was used as it allowed the sequencing of the sense strands. This protocol, described in section 2.2.32 and schematically in Fig. 5.14, was favoured instead of primer walking because only one primer was required for sequencing and all the reactions could be analysed on a single sequencing gel.

The enzyme producing the 3′ overhang, which is resistant to digestion by Exo III thereby protecting the M13 universal primer binding site, was *Sac*I. The 5′ overhang adjacent to the insert was generated by *Bam*HI. The uniform rate of digestion of Exo III from the 5′ overhang allowed deletions of pre-determined lengths to be made by removing timed aliquots from the reaction. S1 nuclease was used to remove the remaining single-stranded tails, the low pH and the zinc cations

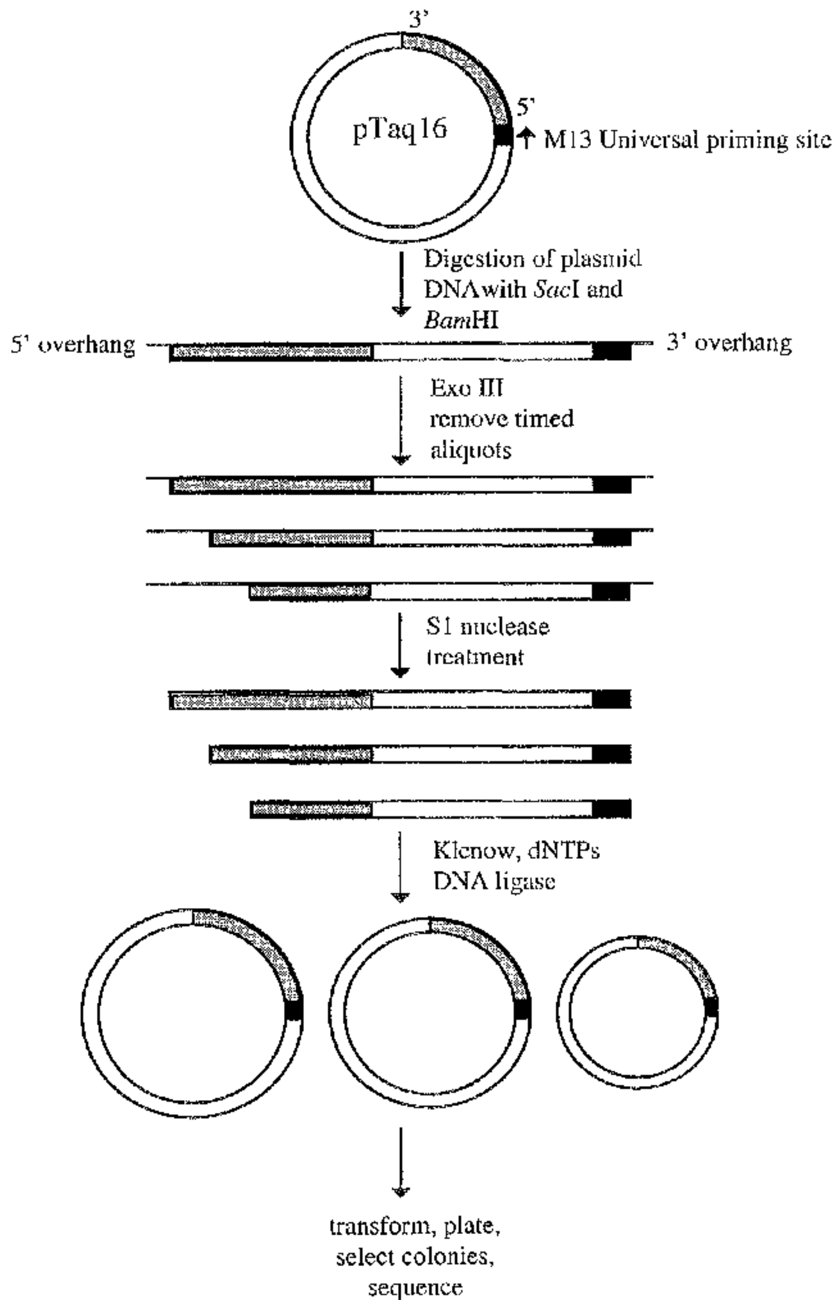


Fig. 5.14: Schematic for the generation of nested-deletions of pTaq16

The general method was followed as described in section 2.2.32. Plasmid DNA (5 µg) was digested with *Bam*HI (to generate 5' overhang) and *Sac*I (to generate 3' overhang) before being digested with Exo III. The outer strand of the insert represents the sense strand. Exo III deletions of pTaq22 were carried out in the same manner, using identical restriction enzymes.

in the S1 nuclease buffer inhibiting further digestion by Exo III. After neutralisation and heat inactivation of the nuclease, the 3´ to 5´ activity of Klenow DNA polymerase was used to remove any remaining protruding 3´ termini. Fig. 5.15a presents an example of the nested deletions produced at 30 s intervals for Taq16, performed in conjuction with Dr S. Pradhan. An excellent set of overlapping deletions were produced which were then ligated and transformed into DH5α cells. Each time point yielded a collection of subclones containing clustered deletions. Nine subclones from each time point were selected randomly for a small scale plasmid preparation and treated with a restriction enzyme to digest out the inserts. *Pvu*II and *Bss*HII contained two sites each within the pBluescript (KS+) plasmid: one site for each enzyme was placed on either side of the multiple cloning site (refer to Fig. 1.1, Materials and Methods). *Bss*HII was used to cut out the insert because there were no sites within the insert itself. The inserts were sized on a 1% agarose gel along side a 100 bp marker (data not shown). A set of subclones, containing a decreasing ladder of inserts with deletions of approx. 100 bp extending further into the original insert, were chosen for plasmid preparations (Fig. 5.15b).

5.3.5.1 DNA Sequence the pTaq22 and pTaq16 Subclones

Plasmids from all the subclones were subjected to automated sequencing (section 2.2.33) using the M13 universal primer. The sequences representing pTaq22 and pTaq16 were aligned using the Gel Assemble programme and were found to be overlapping for the entire length of the gene. A putative consensus sequence of the sense strand was constructed for each clone which was found to start at the codon specifying amino acid 18 (gly18) and continued to the polyA tail. The antisense strand was sequenced by primer walking using primers designed from the consensus sequence (Table 5.3). The primers (20 bases in length) had annealing temperatures between 50-55°C. The final sequence of both clones and the positions of the forward and reverse sequencing primers is presented in Fig. 5.16.

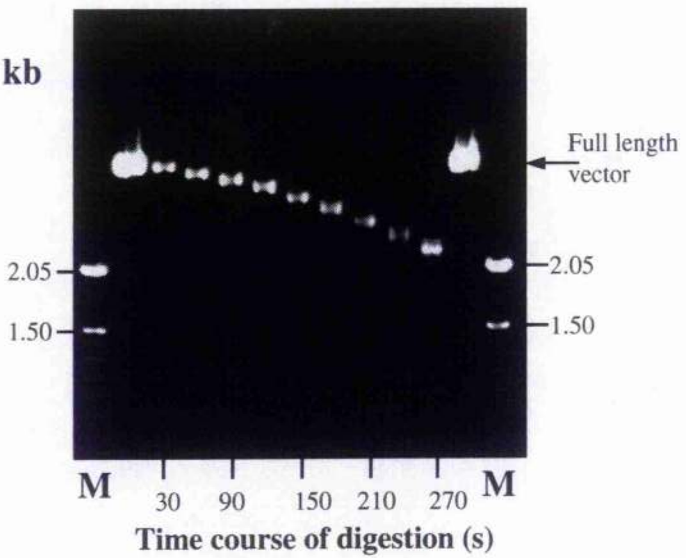


Fig. 5.15a: Time course of Exo III deletions of clone, pTaq16 produced at 30 s intervals

Plasmid, pTaq16, was digested with Exo III according to the method in section 2.2.32.1. Aliquots were removed at 30 s intervals and a fifth of the volume analysed on a 1% (w/v) agarose gel to follow the time course of deletions.

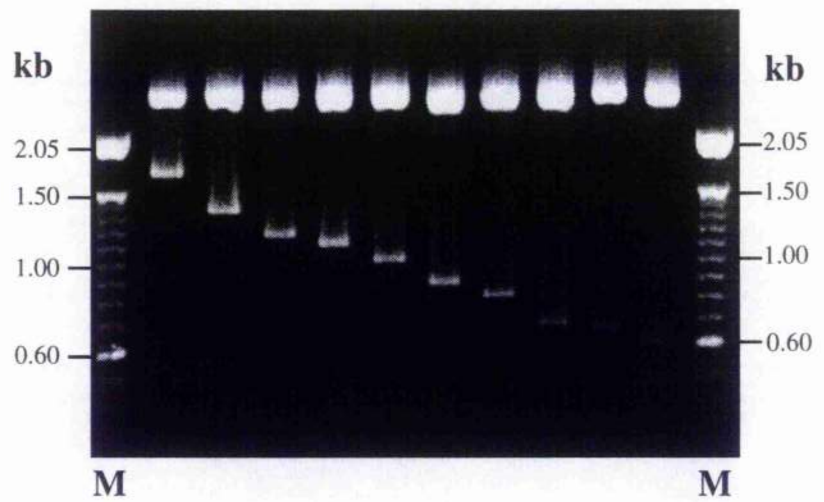


Fig. 5.15b: Subclones of clone, pTaq16 containing a decreasing ladder of inserts

The Exo III deleted plasmids were cloned as described in section 2.2.32.2. Subclones differing by 100 bp were chosen for sequencing by analysing the inserts, digested with *Bss*HII, on a 1% (w/v) agarose gel.

M= size marker

Table 5.3

Sequencing primers designed from the consensus sequence of the sense strand of pTaq22

Reverse Sequencing Primers (5´-3´)	Position along clone +	Forward Sequencing Primers (5´-3´)	Position along clone+
W) GGCTATCGAGGATATAACAC	1386	P) GATTCCTTTACAGCACTTCC	160*
X) CGAGTGACTTTACCTGTTCC	1072	1) CTTCTGTCGAAGTTGATCTC	182
Y) GGTACTCTTCCAGCAGAGAC	786	2) CTGTCGAAGGTGGATAACTC	338
Z) ATTCAACAACAGTCACCTC	571	3) CTCAGAGGTGACTGTTGTTG	548
		4) CTGACAAGGCTGGTAGAATC	827
		5) CTTCCTAGCTAACAGTAGAGCC	1101
		6) TTTTGTTGCTGCCGTTACTC	1434

+ Following numbering scheme of Fig. 5.16

* following numbering scheme of Fig. A, Appendix

The open reading frame of pTaq22 consisted of 1359 bp, encoding a polypeptide of 453 amino acid residues. The full length mature E3 protein was expected to be 470 amino acids with the inclusion of the first 17 amino acids from the N-terminus. The predicted M_r values from the deduced amino acid sequences (including the first 17 amino acids) were 49 558 (pTaq22) and 46 060 (pTaq16). The M_r of the polypeptide from pTaq16 was surprisingly low in comparison with the known M_r of either the native α or β polypeptides (49 446 $\pm$12.28 and 49 562 $\pm$6.83, respectively). This irregularity was resolved when the two putative consensus sequences were aligned and an in frame deletion of 96 bases (corresponding to 32 amino acids with an approx. M_r of 3500) was discovered in Taq16 at position 249 of the ORF (Fig. 5.16). To determine whether the deletion was a cloning artifact, the pTaq23 clone was sequenced by primer walking using the reverse primers designed above and forward primers also designed from the consensus sequence (Table 5.3). pTaq23 was chosen because it was known from expression studies, described in chapter 6, to encode an E3 polypeptide slightly larger than the pTaq16 product but smaller than the pTaq22 product. This had suggested that pTaq23 and pTaq16 may encode a protein similar in size to the β polypeptide and pTaq22 may encode a protein similar in size to the α polypeptide. Sequencing was initiated at the site of primer 1; therefore, the sequence of pTaq23 presented in Fig. 5.16 begins at base 223 of the ORF. Sequence comparison showed that pTaq23 did not have the deletion found in pTaq16 and that its sequence was identical with that of pTaq22 and pTaq16 apart from the base differences indicated in Fig. 16. In order to calculate its mass (49 517.44) from the deduced amino acid sequence, the assumption was made that the upstream sequence was identical with that of pTaq16. Additionally a deletion in this upstream sequence was unlikely in pTaq23 because the digest pattern with *Eco*RI/*Kpn*I presented in Fig. 5.12, showed that its 950 bp band was similar in size to that of pTaq22.

In relating the three clones with each other, it was discovered that there were eight separate base changes in total causing amino acid differences, at the positions

indicated in Fig. 5.16. However, these amino acid differences did not give rise to a mass difference between any of the clones of around 116 Da (found between the native potato α and β polypeptides) and they appeared to be random, leading to the belief that some may represent mutations caused by the use of *Taq* polymerase. It was noted that the clones differed slightly in size due to the polyA start site; the polyA tail of pTaq16 appeared 60 bases after that of pTaq23 and 67 bases before the polyA tail of pTaq22. Nevertheless, the stop codon was in the same place for all the clones. There appear to be no frame shifts, deletions or insertions of one or more amino acids in the sequences.

Fig. 5.16: DNA sequence comparison of the clones pTaq22, pTaq23 and pTaq16.

The sequence of pTaq23 starts at base number 223. A changed base leading to an amino acid change is indicated by '*' and a changed base causing no amino acid change is indicated by '+'. The forward and reverse sequencing primers found in Table 5.3 are shown by half arrows. The stop codon is in bold and the poly A tails are underlined.

```
M13 Universal primer          1
                              GGCTATGTGGCGGCGATCAAAGCAGCA pTaq22
                              GGCTATGTGGCGGCGATCAAAGCAGCA pTaq16
                              G  Y  V  A  A  I  K  A  A

28                *
CAGcTTGGGCTTAAAACTCCATGTATCGAGAAACGTGGTACCCTCGGTGG pTaq22
CAGCTTGGGCTTAAAACTACATGTATCGAGAAACGTGGTACCCTCGGTGG pTaq16
Q  L  G  L  K  T  P  C  I  E  K  R  G  T  L  G  G
                  T

78
TACTTGCCTTAACGTTGGTTGTATTCCTTCTAAGGCACTTCTTCATTCGT pTaq22
TACTTGCCTTAACGTTGGTTGTATTCCTTCTAAGGCACTTCTTCATTCGT pTaq16
 T  C  L  N  V  G  C  I  P  S  K  A  L  L  H  S  S

128
CCCACATGTTTCATGAAGCGCAACATTCATTTGCCAATCATGGTGTGAAG pTaq22
CCCACATGTTTCATGAAGCGCAACATTCATTTGCCAATCATGGTGTGAAG pTaq16
  H  M  F  H  E  A  Q  H  S  F  A  N  H  G  V  K

178         Primer 1                        *
TTTTCTTCTGTCGAAGTTGATCTCCCTGCCATGATGGGACAAAAAGATAA pTaq22
                                             GATAA pTaq23
TTTTCTTCTGTCGAAGTTGATCTCCCTGCCATGATGGGACGAAAAGATAA pTaq16
F  S  S  V  E  V  D  L  P  A  M  M  G  Q  K  D  K
                                       R
228
AGCCGTGTCTAACTTAACACGAGGTATAGAGGGTCTATTCAAGAAAAACA pTaq22
AGCCGTGTCTAACTTAACACGAGGTATAGAGGGTCTATTCAAGAAAAACA pTaq23
AGCCGTGTCTAACTTAACACGA............................ pTaq16
 A  V  S  N  L  T  R  G  I  E  G  L  F  K  K  N  K

278
AAGTAAACTATGTGAAGGGATATGGTAAATTCCTCTCCCCTTCTGAAATT pTaq22
AAGTAAACTATGTGAAGGGATATGGTAAATTCCTCTCCCCTTCTGAAATT pTaq23
.................................................. pTaq16
  V  N  Y  V  K  G  Y  G  K  F  L  S  P  S  E  I

328                 Primer 2
TCTGTTGACACTGTCGAAGGTGGTAACTCTGTTGTTAAAGGCAAAAATAT pTaq22
TCTGTTGACACTGTCGAAGGTGGTAACTCTGTTGTTAAAGGCAAAAATAT pTaq23
..................GGTGGTAACTCTGTTGTTAAAGGCAAAAATAT pTaq16
S  V  D  T  V  E  G  G  N  S  V  V  K  G  K  N  I

 *
378
TATTATTGCAACTGGCTCTGATGTCAAAGGTCTACCTGGCATAACCATTG pTaq22
TATTATTGCAACTGGCTCTGATGTCAAAGGTCTACCTGGCATAACCATTG pTaq23
TGTTATTGCAACTGGCTCTGATGTCAAAGGTCTACCTGGCATAACCATTG pTaq16
 I  I  A  T  G  S  D  V  K  G  L  P  G  I  T  I  D
 V

428
ATGAGAAGAAAATTGTGTCATCCACCGGAGCATTAGCTTTGACTGAAATT pTaq22
ATGAGAAGAAAATTGTGTCATCCACCGGAGCATTAGCTTTGACTGAAATT pTaq23
ATGAGAAGAAAATTGTGTCATCCACCGGAGCATTAGCTTTGACTGAAATT pTaq16
  E  K  K  I  V  S  S  T  G  A  L  A  L  T  E  I
```

```
478
CCAAAAAGATTAGTTGTTATTGGTGCTGGCTATATAGGCCTTGAGATGGG pTaq22
CCAAAAAGATTAGTTGTTATTGGTGCTGGCTATATAGGCCTTGAGATGGG pTaq23
CCAAAAAGATTAGTTGTTATTGGTGCTGGCTATATAGGCCTTGAGATGGG pTaq16
P  K  R  L  V  V  I  G  A  G  Y  I  G  L  E  M  G

528                        Primer 3
ATCTGTCTGGGGTCGCCTCGGCTCAGAGGTGACTGTTGTTGAATTTGCAT pTaq22
ATCTGTCTGGGGTCGCCTCGGCTCAGAGGTGACTGTTGTTGAATTTGCAT pTaq23
ATCTGTCTGGGGTCGCCTCGGCTCAGAGGTGACTGTTGTTGAATTTGCAT pTaq16
 S  V  W  G  R  L  G  S  E  V  T  V  V  E  F  A  S
                            Primer Z
578
CTGATATTGTTCCAACCATGGATGGTGAAGTTCGCAAGCAATTTCAACGT pTaq22
CTGATATTGTTCCAACCATGGATGGTGAAGTTCGCAAGCAATTTCAACGT pTaq23
CTGATATTGTTCCAACCATGGATGGTGAAGTTCGCAAGCAATTTCAACGT pTaq16
  D  I  V  P  T  M  D  G  E  V  R  K  Q  F  Q  R

628
GCTCTTGAGAAGCAAAAGATGAAATTCATGCTAAACACTAAGGTGGTGTC pTaq22
GCTCTTGAGAAGCAAAAGATGAAATTCATGCTAAACACTAAGGTGGTGTC pTaq23
GCTCTTGAGAAGCAAAAGATGAAATTCATGCTAAACACTAAGGTGGTGTC pTaq16
A  L  E  K  Q  K  M  K  F  M  L  N  T  K  V  V  S

678
AGTTGATGCTACAGGTGATGGTGTGAAATTGACCCTTGAACCTTCAGCTG pTaq22
AGTTGATGCTACAGGTGATGGTGTGAAATTGACCCTTGAACCTTCAGCTG pTaq23
AGTTGATGCTACAGGTGATGGTGTGAAATTGACCCTTGAACCTTCAGCTG pTaq16
 V  D  A  T  G  D  G  V  K  L  T  L  E  P  S  A  G

728
GTGGTGATCAAACTATTCTCGAGGCTGATGTTGTTCTCGTCTCTGCTGGA pTaq22
GTGGTGATCAAACTATTCTCGAGGCTGATGTTGTTCTCGTCTCTGCTGGA pTaq23
GTGGTGATCAAACTATTCTCGAGGCTGATGTTGTTCTCGTCTCTGCTGGA pTaq16
  G  D  Q  T  I  L  E  A  D  V  V  L  V  S  A  G
                                          Primer Y
778
AGAGTACCATTCACTTCAGGACTTGGATTGGACACGATAGGAGTTGAAAC pTaq22
AGAGTACCATTCACTTCAGGACTTGGATTGGACACGATAGGAGTTGAAAC pTaq23
AGAGTACCATTCACTTCAGGACTTGGATTGGACACGATAGGAGTTGAAAC pTaq16
R  V  P  F  T  S  G  L  G  L  D  T  I  G  V  E  T

828   Primer 4
TGACAAGGCTGGTAGAATCTTGGTCAATGAACGTTTTGCCACTAACGTCC pTaq22
TGACAAGGCTGGTAGAATCTTGGTCAATGAACGTTTTGCCACTAACGTCC pTaq23
TGACAAGGCTGGTAGAATCTTGGTCAATGAACGTTTTGCCACTAACGTCC pTaq16
 D  K  A  G  R  I  L  V  N  E  R  F  A  T  N  V  P

878                                       +
CAGGGGTACATGCAATTGGTGATGTCATTCCTGGACCAATGCTGGCTCAC pTaq22
CAGGGGTACATGCAATTGGTGATGTCATTCCTGGACCAATGTTGGCTCAC pTaq23
CAGGGGTACATGCAATTGGTGATGTCATTCCTGGACCAATGTTGGCTCAC pTaq16
  G  V  H  A  I  G  D  V  I  P  G  P  M  L  A  H
```

```
928
AAGGCAGAGGAGGATGGCGTTGCTTGCGTAGAGTTCATTGCAGGTAAGGA pTaq22
AAGGCAGAGGAGGATGGCGTTGCTTGCGTAGAGTTCATTGCAGGTAAGGA pTaq23
AAGGCAGAGGAGGATGGCGTTGCTTGCGTAGAGTTCATTGCAGGTAAGGA pTaq16
K  A  E  E  D  G  V  A  C  V  E  F  I  A  G  K  E

978
GGGTCATGTAGACTATGATATGGTACCTGGTGTTGTTTACACCCACCCGG pTaq22
GGGTCATGTAGACTATGATATGGTACCTGGTGTTGTTTACACCCACCCGG pTaq23
GGGTCATGTAGACTATGATATGGTACCTGGTGTTGTTTACACCCACCCGG pTaq16
 G  H  V  D  Y  D  M  V  P  G  V  V  Y  T  H  P  E

1028                               |           *
AGGTGGCTTATGTTGGGAAAACCGAGGAACAGGTAAAGTCACTCGGAGTT pTaq22
AAGTGGCTTATGTTGGGAAAACCGAGGAACAGGTAAAGTCACTCGGAATT pTaq23
AAGTGGCTTATGTTGGGAAAACCGAGGAACAGGTAAAGTCACTCGGAATT pTaq16
  V  A  Y  V  G  K  T  E  E  Q  V  K  S  L  G  V
                          <-------Primer X---    I

1078                            ----Primer 5----->
GATTATCGTGTTGGCAAATTCCCCTTCCTAGCTAACAGTAGAGCCAAGGC pTaq22
GATTATCGTGTTGGCAAATTCCCCTTCCTAGCTAACAGTAGAGCCAAGGC pTaq23
GATTATCGTGTTGGCAAATTCCCCTTCCTAGCTAACAGTAGAGCCAAGGC pTaq16
D  Y  R  V  G  K  F  P  F  L  A  N  S  R  A  K  A

1128             +                              +
AATTGATGATGCTGAGGAAATTGTCAAGGTAATTGCTGAGAAAGAGAGCG pTaq22
AATTGATGATGCTGAGGGAATTGTCAAGGTAATTGCTGAGAAAGAGAGTG pTaq23
AATTGATGATGCTGAGGGAATTGTCAAGGTAATTGCTGAGAAAGAGAGTG pTaq16
 I  D  D  A  E  G  I  V  K  V  I  A  E  K  E  S  D

1178         + *                  +
ACAAGATATTGGGTGTCCATATTATGTCACCTAACGCAGGGGAGCTTATT pTaq22
ACAAGATATTGGCCGCCCATATTATGTCACCTAATGCAGGGGAGCTTATT pTaq23
ACAAGATATTGGGCGTCCATATTATGTCACCTAATGCAGGGGAGCTTATT pTaq16
  K  I  L  G  V  H  I  M  S  P  N  A  G  E  L  I
              A
1228       +
CACGAAGCTGTCCTGGCATTGCAGTACGGAGCATCAAGTGAGGACATTGC pTaq22
CACGAAGCTGTACTGGCATTGCAGTACGGAGCATCAAGTGAGGACATTGC pTaq23
CACGAAGCTGTACTGGCATTGCAGTACGGAGCATCAAGTGAGGACATTGC pTaq16
H  E  A  V  L  A  L  Q  Y  G  A  S  S  E  D  I  A

1278                                    *
TCGTACTTGTCATGCACATCCAACAATGAGTGAGGCACTCGAAGAAGCAG pTaq22
TCGTACTTGTCATGCACATCCAACAATGAGTGAGGCACTCAAAGAAGCAG pTaq23
TCGTACTTGTCATGCACATCCAACAATGAGTGAGGCACTCAAAGAAGCAG pTaq16
 R  T  C  H  A  H  P  T  M  S  E  A  L  E  E  A  A
                                        K
1328                          *
CCATGGCCACTTATGACAAGCCCATCCACATGTAGGCTTGTGTTATATCC pTaq22
CCATGGCCACTTATGACAAGCCCATCCACACGTAGGCTTGTGTTATATCC pTaq23
CCATGGCCACTTATGACAAGCCCATCCACATGTAGGCTTGTGTTATATCC pTaq16
  M  A  T  Y  D  K  P  I  H  M            <-Primer W
                             T
1378
TCGATAGCCTCATTTTTCTTATTCCGTTGAGAATCTTGAACGCTTAGAAT pTaq22
TCAATAGCCTCATTTTTCATnTnCCGTTGAAAATCTTGAATGCTTAnAAT pTaq23
TCGATAGCCTCATTTtTCATATTCCGtTGAGAATCTTGAATGCTTAGAAT pTaq16
```

1428 Primer 6

AGGCTTTTTGTTGCTGCCGTTACTCAAGCAGCCCATATTCCTCTTnATTC pTaq22
AGGCTTTTTGTTGCTGCCGTTACTCAAACAGCCCATATTCCCCTTTATtC pTaq23
AGGCTTTTTGTTGCTGCCGTTACTCAAGCAGCCCATATTCCCCTTTATTC pTaq16

1478

AGGACTCCAAATACACTATATGAATGTTTATCTTCTGCATTTTTCCATAA pTaq22
aGGACTCCAAATACATTATATGAATGTTTATCATCTGCATTTTTCCATAA pTaq23
AGGACTCCAAATACATTATATGAATGTTtATCAtCTGCATttttCCAtAA pTaq16

1528

TAAAGATGAGATGTAACCTGCTTTAAACCCTGCAGAAGTTAACTGCCTG pTaq22
AAAAAAAAAAAAAAA **pTaq23**
TAAAGATGAGATGTAACCTGCTTTAAACCCTGCAGAAGTTAACTGCCTG pTaq16

1578

GCACTAAsATGGTGCTAGACATTACAAAACTTCCCATTTACTTTGCTTTT pTaq22
GCACTAAGATGGTGCTAGACATTACAAAAAAAAAAAAAAAAAAA **pTaq16**

1628

TTGGCATGTATCCCTAAATTCTATATTTATCTTGTTTGATTTGGTAGATC pTaq22

1678

GGAGTCGATTTGGTGAAAAAAAAAAAAAAAAAA **pTaq22**

M13 Reverse primer

5.4 Discussion

PCR amplification of cDNA derived from RNA was the method used for cloning the potato mitochondrial E3 gene for the reason that the possible presence of two transcripts of the gene, encoding the α and β polypeptides, may be isolated more efficiently than screening a library.

The clones, pTaq13, pTaq14, pTaq15, pTaq16, pTaq22, pTaq23, pTaq25 and pTaq 30 were identified as containing E3 cDNA by comparison of their deduced amino acid sequence with the N-terminal sequence information obtained from the purified potato E3 protein. However, as a result of the cloning strategy the cDNA sequence started at the bases encoding amino acid eighteen (glycine).

Clones pTaq22, pTaq16 and pTaq23 were selected for sequencing after expression studies (chapter 6) showed that they expressed E3 proteins which could possibly be related to the α and β polypeptides. In chapter 6, pTaq22 appeared to express a polypeptide with lower mobility as assessed by SDS-PAGE, indicating that it could represent the α polypeptide. However, the theoretical M_r, 49 558 (calculated from the deduced amino acid sequence), was closer to the native M_r of the β polypeptide, which appears as the higher mobility subunit on SDS-PAGE. pTaq23 and pTaq16 appeared to express a polypeptide with higher mobility than the pTaq22 product by SDS-PAGE in chapter 6, which could possibly correspond to the β polypeptide. In the case of pTaq16, the low M_r of the deduced amino acid sequence and the higher mobility of the polypeptide was explained by an internal deletion of 96 bp, encoding 32 amino acids. Although the deletion was in frame, it was probably artifactual since the mass of the polypeptide, estimated from the deduced amino acid sequence, was too distant from the native α and β polypeptides. The activity of *Taq* DNA polymerase would be expected to mis-incorporate or add a base and not to delete large sequences from a PCR product. Therefore, it is suggested that the sequence may have been deleted by a recombination event in DH5α during cloning. However, the existence of

secondary structures in the DNA, causing *Taq* polymerase to skip parts of the DNA can not be ruled out. Since only three of the clones were sequenced, it was not known if a deletion, similar to pTaq16, existed in any of the other clones. The theoretical M_r of the deduced amino acid sequence of pTaq23 was 41Da smaller than the M_r of the pTaq22 (being 49 517.44) protein due to four amino acid substitutions between them.

The double restriction digests with *Eco*RI/*Kpn*I generated two bands: the top one was approximately 950 bp and the lower one varied in size (Fig. 5.12). Full length sequencing of both strands of the cDNAs found that the variation in the lower band was owing to heterogeneity at the polyA site. The top band of pTaq16 was slightly smaller in size at 850 bp, coinciding with the deletion of 96 bp. Sequencing pTaq23, which had expressed a protein slightly larger that pTaq16, showed that it did not have this deletion. The reason for the slightly larger top band in pTaq25 (Fig. 5.12) remained a mystery as it was not sequenced.

A comparison of the deduced amino acid sequences presented in the Appendix, Fig. B, revealed that there were no gross differences in amino acids between the clones, the translation stop site was in an identical position and there were no frame shifts. With the data available, it was difficult to conclude that the clones represented the separate α and β cDNAs. It appeared that only the cDNA encoding the β polypeptide had been cloned. However, two amino acid residues per clone (three in pTaq16) were different from the consensus. These differences in amino acids were at dissimilar positions and provided no obvious distinction between pTaq22 and the clones, pTaq16 and23. Additionally, at five other positions in the sequence (919, 1145, 1176, 1191 and 1239, in Fig. 5.16) there were differences in bases which caused no change in the amino acid. It seemed that the differences were random and may be attributed to the use of *Taq* DNA polymerase due to the relatively high error rate of this enzyme. Base substitutions occur about one per 9000 bp and frameshifts occur about one per 40 000 bp. The mutation rates are highly template-specific and are caused by the lack of a 3´-5´

exonuclease (proofreading) activity. One error in 9000 bp seems insignificant but after 30 cycles of PCR it leads to an error in every 300 bp of product. As single products were cloned this error rate may be encountered in the sequencing. However, it was difficult to judge which amino acid differences were real and not due to *Taq* DNA polymerase.

An attempt had been made to reduce the chances of including mutations in the sequence by the use of the proof-reading *Pfu* DNA polymerase. It was unclear as to why the *Pfu* amplified product was "unclonable". In general, many failures are attributed to the fact that some foreign inserts form non-standard secondary and tertiary structures, which are then deleted by the host. Two frequently occurring unstable structures in eukaryotic DNA are cruciforms (caused by inverted repeats) and Z-DNA (formed within alternating purine-pyrimidine stretches). Homologous recombination between the genome of the host and the recombinant plasmid is also a possibility. However, these explanations are unlikely as the *Taq* amplified DNA products, which were expected to be the same as the *Pfu* amplified products, had been cloned successfully. The seven false positives, shown in Fig. 5.10, may have arisen by a small deletion in the Lac Z gene as a result of excision of the inserted DNA at some stage during colony growth (Brown, 1992).

Expression of the cloned gene in *E. coli*, as described in the following chapter, was carried out in an attempt to determine if the clones represented only one type of polypeptide, α or β. Moreover, active E3 protein would suggest that the PCR derived clones were faithful copies of the E3 gene. Since the 5′ upstream sequence (encoding the leader sequence and the first seventeen amino acids) was unavailable from the PCR derived cDNA, a potato leaf cDNA library was later screened to isolate full length cDNA clones (chapter 7). The added advantage of this was that the sequence of the PCR derived clones could be compared with that of the clones isolated from the library and they could be used for future expression studies if expression of the *Taq* clones proved to be unsuccessful.

Chapter 6

Heterologous Expression of the Potato Mitochondrial Dihydrolipoamide Dehydrogenase Gene(s) in *Escherichia coli*

6.1 Introduction

Owing to the error rate of *Taq* polymerase, it was decided to determine whether the eight *Taq* clones, isolated in chapter 5, encode a functional lipoamide dehydrogenase (E3) by attempting its expression in *E. coli*. However, it was recognised that the absence of the nucleotide sequence encoding the first seventeen amino acids of the mature protein from the cloned *Taq* inserts, as a result of the cloning strategy, may also have posed problems in the binding of the FAD cofactor or the folding of the E3 protein or both. The domain structure of E3 was elucidated when a tertiary structure for *E. coli* E3 was determined by fitting its primary amino acid sequence into the known tertiary structure of glutathione reductase (Rice *et al.*, 1984). The FAD-binding domain is composed of residues 1-145 and is characterised by a βαβ structural fold involving residues 8-35, of which 5 residues are known to make contact with the FAD. For this reason the successful binding of FAD to the truncated protein during over-expression was in some doubt. Although the consequences of the absence of this sequence were unknown, an additional advantage of the heterologous expression work was to determine whether polypeptides of differing mass were expressed by any of the clones which could be related to either the α or β polypeptides of E3. Furthermore, successful over-expression of the potato E3 in *E. coli* would provide a plentiful supply of the enzyme, which is not available from potato tubers.

The pET system (Novagen) was chosen to subclone and express the *Taq* inserts, encoding the E3 gene, in *E. coli*. Genes are under the control of strong bacteriophage T7 transcription and translation signals: expression is induced by providing a source of T7 RNA polymerase in the host cell. pET28a, b and c are translation vectors, for the cloning of genes without the ribosome binding site (rbs) and ATG start codon or for eukaryotic genes in general as they do not carry the compatible rbs for expression in prokaryotes. The vectors carry the efficient rbs from the phage T7 major capsid protein and sequences adjacent to the cloning site

for the histidine 'Tag' for purification of the protein on a nickel affinity column. This stretch of 6 histidine residues can be cleaved off at a later stage by thrombin. The three forms of the vector, 'a', 'b' and 'c' denote the reading frame relative to the *Bam*H1 cloning site (see Fig. 1.1, Materials and Methods).

At first, genes are cloned using DH5α, a host strain that does not contain the T7 RNA polymerase. As a result, they are transcriptionally silent, eliminating plasmid instability owing to the production of proteins potentially toxic to the host cell. Once established in a non-expression host, the plasmids are then transferred into expression hosts - *E. coli* BL21(DE3) cells - containing a chromosomal copy of the T7 RNA polymerase gene under the control of the lacUV5 promoter which is induced by the addition of IPTG (0.4mM). The added advantage of these strains is that they have been engineered to lack the *lon* protease and the *omp* T outer membrane protease. Thus, target proteins are expected be more stable than in host strains containing these proteases. An additional strain - BL21(DE3)pLysS - can be used if expression of the heterologous protein takes place at an accelerated rate and is detrimental to the cell or to correct folding of the protein. The pLysS plasmid encodes T7 lysozyme which forms a specific complex with T7 RNA polymerase, reducing the rate of transcription (Moffat & Studier, 1987).

6.2 Results

6.2.1 Subcloning the *Taq* Inserts into pET28c Expression Vectors

The general procedure for subcloning into the *Bam*HI and *Hind*III sites of the expression vector, pET28c from pBluescript was described in section 2.2.34.1. pET28c expresses from the ATC triplet of the GGATCC *Bam*HI recognition sequence and was chosen because expression of the E3 gene would be in the correct reading frame i.e. from the first glycine residue of the cloned gene as described in Fig. 6.1.

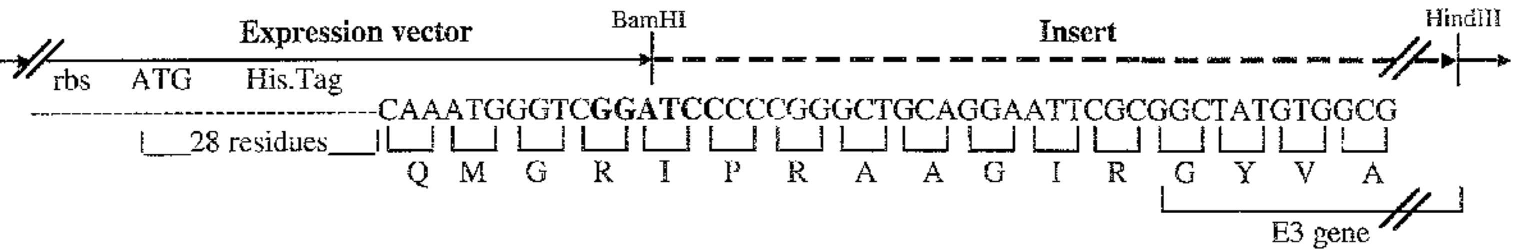


Fig 6.1: Reading frame for expression of the pET28c expression construct, containing the E3 cDNA

The *Taq* amplified inserts (cloned in chapter 5) were subcloned from pBluescript into the *Bam*H1 and *Hind*III sites of pET28c. The figure presents only the last four residues of the expression plasmid and the first four codons of the E3 gene, to describe the reading frame of the construct from the *Bam*HI site. The first glycine residue of the E3 cDNA is actually residue 18 of the mature E3 protein.

With reference to chapter 5, the *Taq* amplified inserts representing the E3 gene (clones pTaq13, 14, 16, 22 and 23) had been inserted into pBluescript KS+ in the antisense direction; therefore, excising the insert from the plasmid with *Bam*HI and *Hind*III would allow their insertion into pET28c in the sense orientation. However, the *Taq* amplified inserts in pTaq15, 25 and 30 were cloned in the sense orientation of pBluescript already and subcloning with *Bam*HI and *Hind*III orientated them in the antisense direction in the expression vector. As a result, expression of these genes should either be absent or should produce nonsense protein. These three constructs were used as a negative control of expression. An additional negative control for expression was the insert from pTaq22 subcloned into pET28a. As this insert would have been cloned in the wrong reading frame, protein expression should not occur or if it did, nonsense protein would be expressed.

Fig. 6.2 is an analytical gel showing the inserts subcloned into pET28c and propagated in DH5α *E. coli* cells. They were of the correct size expected for the respective cloned *Taq* insert. These new clones were named pTaqE16, 22, 23, 25 30, 13, 14 and 15

6.2.2 Time Course of Expression from the pET28c Clones

The transfer of the recombinant expression vectors from the cloning host to the expression host strain was described in section 2.2.34.2. The expression hosts used were BL21(DE3) and BL21(DE3)plysS. The latter strain possesses an additional plasmid, pLysS, which encodes for T7 RNA polymerase lysozyme. As described in the introduction to this chapter, T7 lysozyme provides a stringent level of control for rapid over-expression of proteins by complexing with a certain amount of the T7 RNA polymerase produced on induction.

The time course of protein expression, after induction of the lacUV5 promoter with 0.4 mM IPTG, was performed following the method detailed in section 2.2.34.3. It was important to induce protein expression in mid log phase (A_{600} of between 0.3-0.5) as this would ensure that there were still plenty of

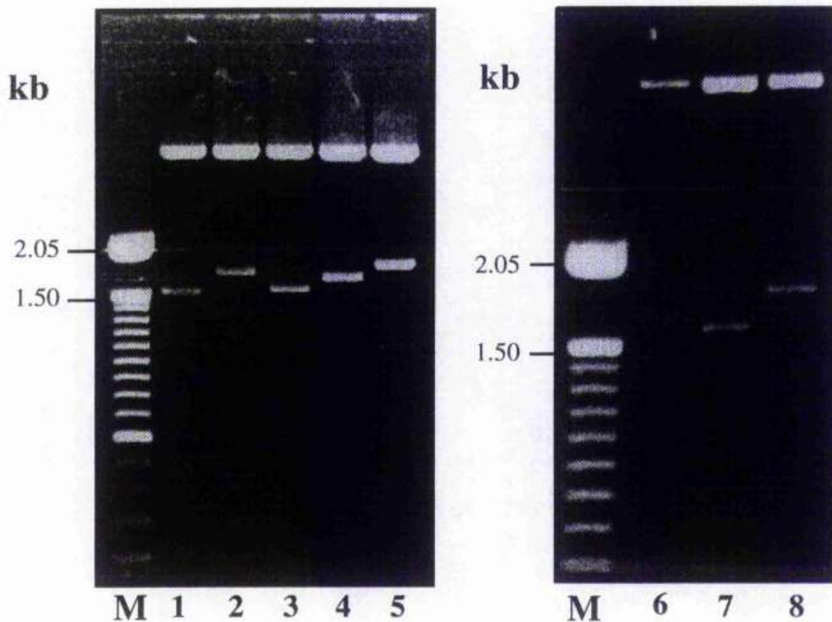


Fig. 6.2: Analysis of the *Taq* amplified inserts subcloned from pBluescript into the pET28c expression vector

The *Taq* amplified inserts were subcloned from pBluescript (KS+) clones (section2.2.34.1) into the *Bam*HI and *Hind*III sites of pET28c. Positive clones containing each of the inserts were isolated and purified plasmid DNA (1 µg) was digested with *Bam*HI and *Hind*III. Fragments were separated on a 1% (w/v) agarose gel.

M: size marker, 1: pTaqE16, 2: pTaqE22, 3: pTaqE23 4: pTaqE25, 5: pTaqE30, 6: pTaqE13, 7:pTaqE14, 8: pTaqE15

nutrients in the media to support continued growth of the cells. Samples (1 ml) were collected 0, 1, 2 and 3 h after induction and separated by 10% SDS-PAGE as described in section 2.2.5.

The results (Table 6.1) were very encouraging as it was found that the clones pTaqE13, 14, 16, 22 and 23 were all found to express protein at both 30°C and 37°C in one of the strains - BL21(DE3). The complete lack of expression in *E. coli* BL21(DE3)plysS was surprising because T7 lysozyme is only expected reduce the rate of expression and not inhibit it. As expected, clones pTaqE15, 30 and 25 did not express protein because, as mentioned earlier, they had been cloned in the wrong orientation. There were two additional negative controls for expression. The first was expression of insert cloned in the wrong reading frame i.e. in pET28a and the second was induction of untransformed *E. coli* BL21(DE3). Neither was found to express the heterologous protein (data not shown).

Table 6.1

Results from the induction of expression of the Taq clones

Clone	Expression in BL21(DE3)	Expression in BL21(DE3)plysS
pTaqE13	yes	no
pTaqE14	yes	no
pTaqE16	yes	no
pTaqE22	yes	no
pTaqE23	yes	no
pTaqE15	no	no
pTaqE25	no	no
pTaqE30	no	no

Fig. 6.3 shows the results of expression of only two of the clones at 37°C. High level of over-expression of clone pTaqE13 in BL21(DE3) can be seen over the 4 h; in contrast no expression was obtained in the BL21(DE3)plysS strain. The recombinant plasmid is slightly 'leaky' as some expression is evident before induction (at 0 h). The lack of expression in either of the host cells is shown for clone pTaqE15.

Fig. 6.4A and 6.4B present 10% (w/v) SDS-polyacrylamide gels analysing the expression at 37°C of clones pTaqE13, 14, 16, 23 and comparing the expression of pTaqE22 and 23. Interestingly, proteins having slightly differing mobilities on SDS-PAGE were expressed. In order to show the apparent differences in mobilities of the expressed proteins, four times less sample was loaded as compared with the gel in Fig. 6.3. The plasmids, pTaqE13, pTaqE14 and pTaqE22 were found to express a protein having a lower mobility than the protein expressed from pTaqE16 and pTaqE23 by SDS-PAGE. It was thought at first that they may represent the α and β forms; however, sequencing of pTaq22 and pTaq23 had shown that they had very similar cDNAs and the unusually small polypeptide expressed by pTaqE16 was explained by a deletion of 96 bp leading to the absence of 32 amino acids from the mature protein. The protein encoded by pTaqE14 does appear to be intermediate in size between the protein of pTaq13 and pTaq23. The M_r of the expressed proteins will be 2.7 kDa more than the native proteins purified from potatoes, contributed by the plasmid from the ATG initiating residue to the *Bam*HI cloning site (ref. Fig. 6.1). However, since the native E3 proteins were not analysed alongside the expressed proteins, this increase in mass would not be discernible on the gel.

The identity of the expressed proteins as E3 was confirmed by Western blotting the proteins expressed from all clones and immunological detection by ECL (section 2.2.7). Yeast antiserum was used to probe the blots as it was known to cross react with potato tuber mitochondrial E3. A one in five thousand dilution of the antibody was made and incubated with blots of gels showing the time course of

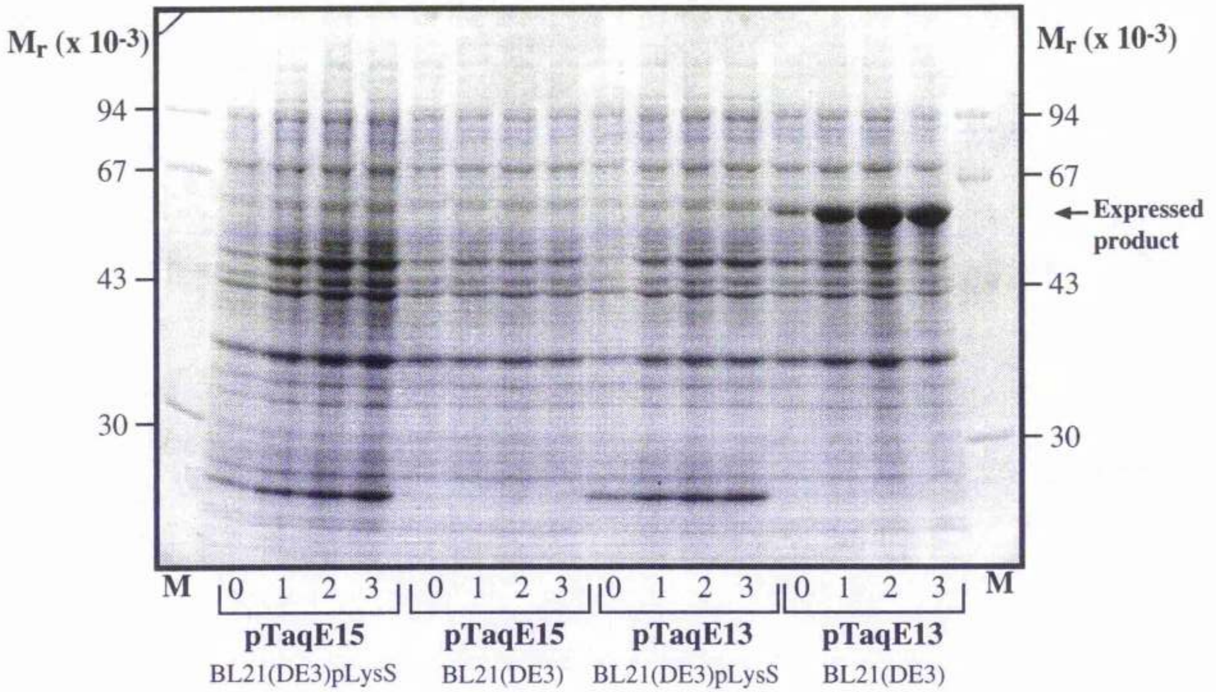


Fig. 6.3: Analysis of the time course of expression, at 37°C from clones pTaqE13 and pTaqE15 in *E. coli* BL21(DE3) or BL21(DE3)pLysS cells

The clones were propagated in either 1) BL21(DE3) or 2) BL21(DE3)pLysS *E. coli* cells in 100 ml LB media (supplemented with kanamycin or kanamycin/chloramphenicol, respectively). Expression was induced at an A_{600} of 0.5 with 0.4 mM IPTG. Samples (1 ml) were removed at hourly intervals (0, 1, 2, and 3 h) after induction and the pelleted cells resuspended in Laemmli buffer (10µl/0.1 absorbance unit). The samples (20 µl) were separated on 10% (w/v) SDS-PAGE and stained with Coomassie blue.

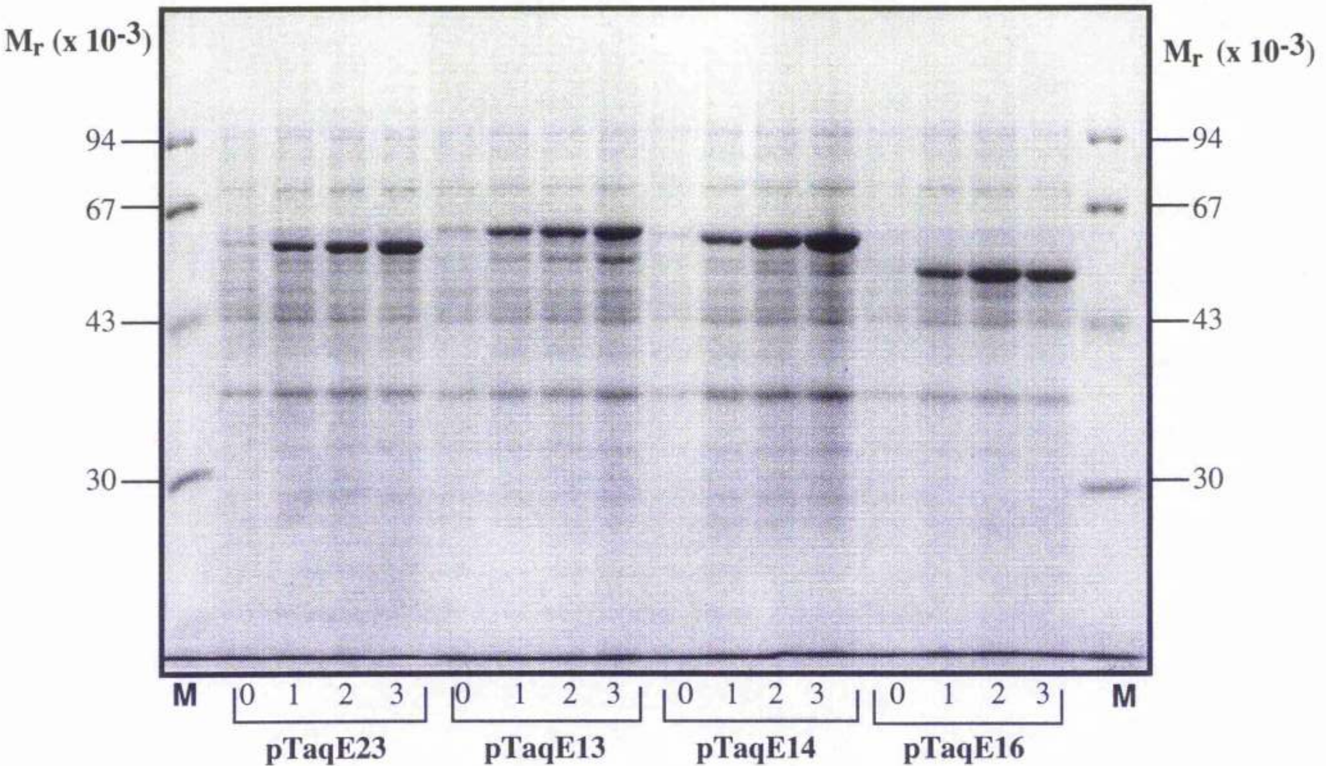


Fig. 6.4a: Time course of expression at 37°C, from clones pTaqE23, pTaqE13, pTaqE14 and pTaqE16 constructs in *E. coli* BL21(DE3) cells

The clones were propagated in BL21(DE3) *E. coli* cells in 100 ml LB (kanamycin) media. Expression was induced at an A_{600} of 0.5 with 0.4 mM IPTG. Samples (1 ml) were removed at hourly intervals after induction and the pelleted cells resuspended in Laemmli buffer (10μl/0.1 absorbance unit). The samples (5 μl) were separated on 10% (w/v) SDS-PAGE and stained with Coomassie blue.

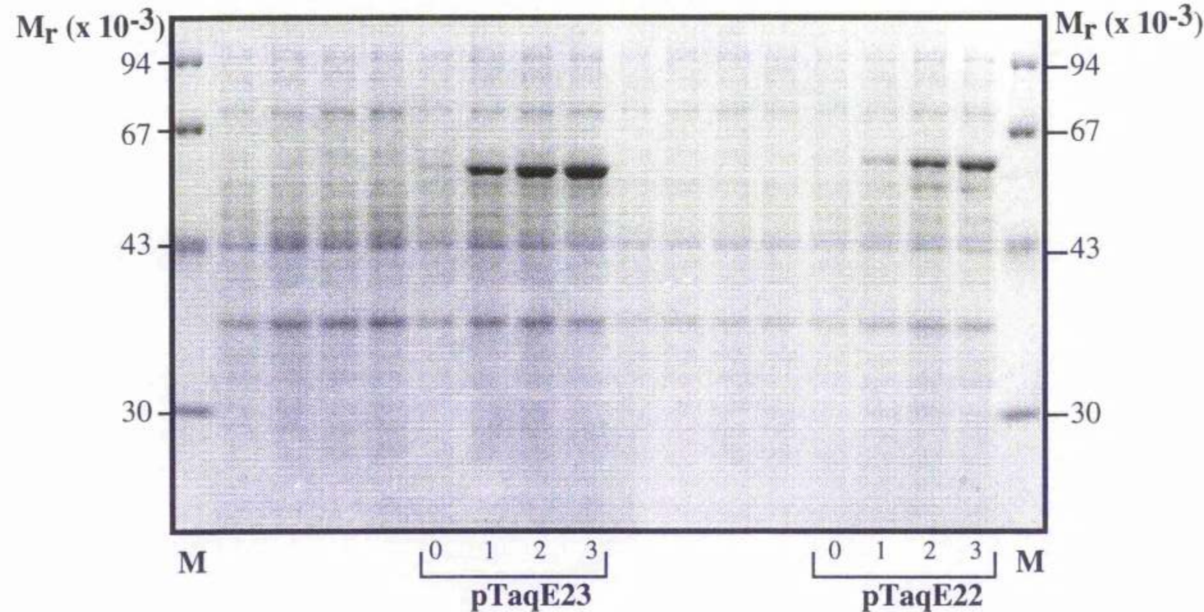


Fig. 6.4b: Time course of expression at 37 °C, from clones pTaqE23 and pTaqE22 constructs in *E. coli* BL21 (DE3) cells

The clones were propagated in BL21(DE3) *E. coli* cells as described in Fig. 6.4a. Samples (1ml) were removed at hourly intervals after induction and the pelleted cells resuspended in Laemmli buffer (10µl/0.1 absorbance unit). The samples were separated on 10% (w/v) SDS-PAGE and stained with Coomassie blue.

expression for all the clones. The results of the ECL detection of protein expression from plasmids pTaqE13, 14 and 16 are presented in Fig. 6.5. Both forms of the protein cross reacted with the antisera very specifically. The blot shows the 'leaky' nature of pTaqE13 since E3 is detectable before induction at zero time. Furthermore, no background was observed and there was no cross reaction with *E. coli* E3 as judged by probing the cell extracts of the untransformed host cells (data not shown).

6.2.3 Purification of the Expressed Dihydrolipoamide Dehydrogenase from BL21(DE3) Cells

Having established expression of E3 in *E. coli*, the next step was to purify the enzyme in an active form. Initial experiments concentrated on trying to assay over-expressed protein in a crude extract before attempting purification on nickel affinity columns using the six histidine tag on the N-terminal end of the enzyme which binds to the divalent nickel cations.

Lysate preparation from 100 ml cultures of the expressing BL21(DE3) was described in section 2.2.34.3 of Materials and Methods. The control sample was untransformed BL21(DE3) which was cultured and induced in the same way as the experimental samples. After three passages through a French press, the control sample emerged as a transparent, straw coloured liquid. In contrast, the lysates from the expressing samples were opaque containing fine white particulate matter. Not surprisingly, centrifugation at 12 000 g for 20 min resulted in the white insoluble matter pelleting down leaving a clear supernatant. E3 activity was assayed before and after centrifugation as in section 2.2.10.2 but no activity was found above background *E. coli* E3 activity. Evidently, E3 was being expressed in an insoluble form, accumulating in inclusion bodies. The pelleted samples were solubilised in Laemmli sample buffer and resolved on a 10% (w/v) SDS-polyacrylamide gel which showed, in lanes 2 and 3 of Fig. 6.6, that the E3 was indeed in the insoluble fraction. Furthermore, all concentrated preparations of E3 tend to be a bright yellow colour due to the presence of the FAD cofactor.

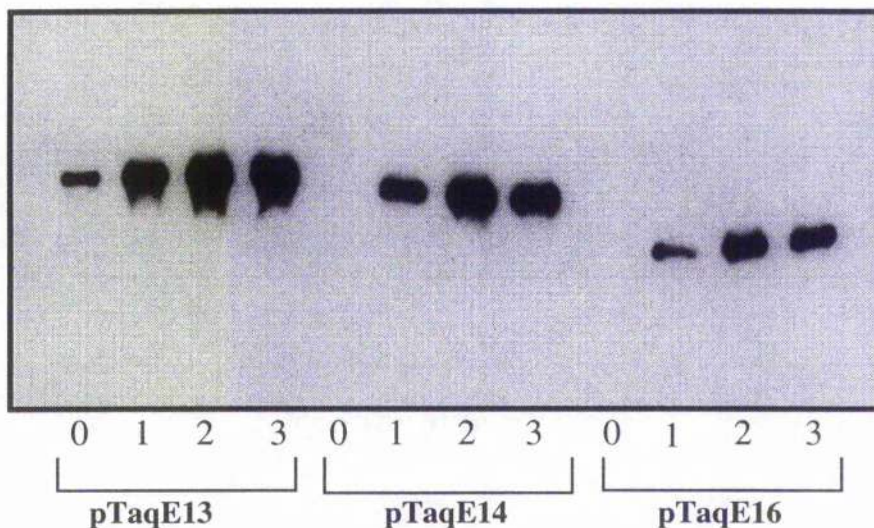


Fig. 6.5: Immunological identification of the protein expressed from clones pTaqE13, pTaqE14 and pTaqE16

A time course of expression was carried out as described in section 2.2.34.3. Samples (1 ml,) collected at the time points of 0, 1, 2 and 3 h after induction with 0.4 mM IPTG, were pelleted and resuspended in laemmli buffer (10μl/0.1 absorbance unit). The samples were separated by 10% (w/v) SDS-PAGE and transferred to Hybond C membrane. ECL detection of the protein was performed with yeast E3 anti-serum (1/5000 dilution) as described in Materials and Methods, section 2.2.7.

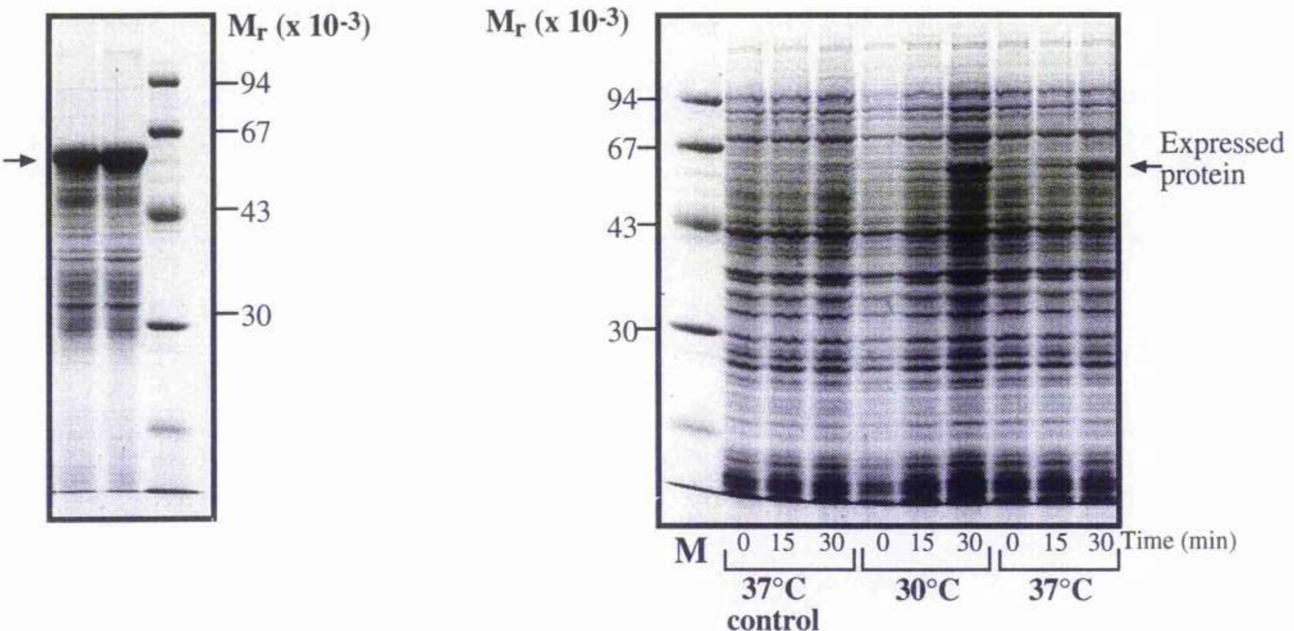


Fig. 6.6: Analysis of the insoluble, expressed protein

E. coli BL21(DE3), containing pTaqE22, was grown until an A_{600} of 0.5 was obtained. Protein expression was induced and cells lysed after 3 h, as described in section 2.2.34.3. The insoluble protein was pelleted at high speed in a microfuge and resuspended in Laemmli buffer. The protein was resolved by 10% (w/v) SDS-PAGE (in duplicate) and stained with Coomassie blue.

Fig. 6.7: Time course of expression from pTaqE22 over 30 min at 30°C and 37°C

pTaqE22 was propagated in *E. coli* BL21(DE3) in 100 ml LB media, supplemented with kanamycin. Samples (1 ml) were collected 15 and 30 min after induction of protein expression with 0.4 mM IPTG and the pelleted cells resuspended in Laemmli buffer (10μl/0.1 absorbance unit. The samples (20 μl) were resolved by 10% (w/v) SDS-PAGE and stained with Coomassie blue.

Therefore, since the precipitated material was uncharacteristically white, it was possible that there was no binding of the FAD cofactor.

A number of conditions for growth of the expressing strains were tested using the pTaqE22 clone, in an attempt to increase the solubility of the expressed E3. It is a well known phenomenon that reduction in the temperature of growth can lead to 30-90% of the heterologous protein accumulating in active, soluble form (Hockney, 1994). Expression of E3 at 30°C had already been shown (section 6.2.2) to occur at similar levels as 37°C, but again the protein was insoluble and inactive.

It was conceivable that E3 was being expressed too rapidly and was aggregating in the bacterial cells over the 4 h of induction period. In view of this, samples for analysis on a 10% (w/v) SDS-PAGE gel were collected 15 and 30 min after induction, after which time cells were harvested for protein purification. As can be seen in Fig. 6.7, a high level of expression had already occurred after 30 min at both 30°C and 37°C and this is certainly too rapid to permit correct protein folding. A five times lower concentration of IPTG (0.8 mM) caused no detectable reduction in the rate of expression. In both cases E3 was insoluble and therefore inactive. The protein was still insoluble even when the cells were extracted in the presence of 50% glycerol using the French press.

As the rate of expression was so fast it was possible that the cells could not produce enough FAD for the vast amounts of E3. Finally, expression was carried out in the presence of riboflavin, the precursor required for FAD production. Riboflavin was known to be quite insoluble in aqueous solution that it was necessary to add the solid riboflavin directly to 50ml of sterile LB media at a concentration of approx. 0.05-0.1 mM. The media was re-sterilised by passing it through a 0.22 μm filter and the culture flasks were foil wrapped as riboflavin is sensitive to light. Again, this attempt failed to produce active, soluble E3 (data not shown).

6.3 Discussion

Over-expression of the E3 enzyme in *E. coli* was successful in that it showed that cDNAs encoding E3 had been cloned by the specific amplification of RNA derived cDNA. It appeared that each clone was encoding products with different mobilities on SDS-PAGE which may have been be related to the α and β polypeptides. However, subsequent amino acid sequence information deduced from the DNA sequence for pTaq16, pTaq22 and pTaq23 (Chapter 5) found that there were no gross differences between the clones which could account for the expression of the separate polypeptides. Importantly, the open reading frame of pTaq22 and pTaq23 were the same length. Calculation of the M_r values from deduced amino acid sequence of the clones pTaq22, pTaq23 and pTaq16 in chapter 5 also dispelled the idea that the expressed proteins may represent both the α and β E3 polypeptides. The M_r of the protein encoded by pTaq22 was found to be closer to that of the β polypeptide which had been obtained by electrospray mass spectrometry. The 96 base deletion, discovered on sequencing pTaqE16 in chapter 5, made an observable difference to the subunit M_r of the expressed protein as the polypeptide expressed from this plasmid was slightly smaller than the polypeptide expressed by pTaqE23. The predicted M_r of this protein would have been 49 589 if the deletion was not present; a value close to the M_r of the native β polypeptide (49 562, determined by Es-Ms). Similarly the predicted M_r of the protein encoded by pTaq23 was closer to the M_r of the β polypeptide.

Sequence information of three of the *Taq* clones pointed to only one type of cDNA having been cloned from RNA of potato leaves and hence only one type of E3 being expressed (β_2). The question of why the expressed proteins displayed slightly different sizes still remained to be answered. Proteolysis within *E. coli* is a possibility depending on the conformation of the insoluble protein.

Unfortunately, this work was inconclusive in terms of confirming the authenticity of the *Taq* polymerase derived clones because the over-expressed proteins were inactive. In addition, the expressed products could not be identified

as specifically α or β subunits without a comparison of their kinetic profiles with those of the native isoforms. Reducing the temperature of growth of the cultures from 37°C to 30°C and lysing the cells in the presence of 50% glycerol did not improve the solubility of the expressed protein leading to the production of active enzyme. The complete lack of expression in BL21(DE3)pLysS was surprising, although it is possible that the T7 lysozyme was controlling T7 DNA polymerase too efficiently.

Other prokaryotic E3s have been successfully over-expressed in *E. coli* (Benen *et al.*, 1989; Westphal & Kok, 1988) but this is the first known attempt to express a eukaryotic E3 in a prokaryotic system. Many proteins are known to accumulate in the form of insoluble, biologically inactive inclusion bodies and in most cases the problem stems from the fact that the protein being expressed is eukaryotic in origin. Prokaryotes used as hosts for heterologous expression, like *E. coli* do not provide the appropriate environment for many eukaryotic proteins. They are unable to carry out many of the co- or post-translational modifications required by eukaryotic proteins e.g. glycosylation. In addition, eukaryotic proteins may not be able to fold correctly because the rate of protein synthesis is far more rapid in prokaryotes and they lack the specific chaperones.

It was suggested in the introduction to this chapter that the absence of the first 17 amino acids from the N-terminal end of the mature protein may be the main contributing reason for the resulting insoluble protein. As discussed in chapter 1, the E3 polypeptides from all sources are highly conserved and can be divided into four functional domains: the FAD-binding domain, the NAD(H)-binding domain, the central domain and the interface domain (Carothers *et al.*, 1989). Sequence analysis of the pea and potato mitochondrial E3 indicated that the important sites for enzyme activity are strongly conserved (Bourguignon *et al.*, 1992). It was expected therefore, that their secondary and tertiary structures would be the same as other E3s.

A structural comparison will be made with *E. coli* E3 which has 40% identity with the pea E3. The FAD cofactor is deeply buried within the FAD-binding domain and in *E. coli* E3 this encompasses residues 1 to 145. The conserved structural motif for FAD-binding is the βαβ fold which is found between residues 8 to 35 and within this group, the residues Leu11, Ala13, Gly14, Ala16, and Gly17 are known to make contacts with the FAD cofactor. There are an additional nineteen main chain residues with polar or non-polar side-chains, found right along the polypeptide, making contacts with the FAD cofactor (Rice *et al.*, 1984). Clearly, there are important contacts made with the FAD cofactor in the first twenty residues which would be missing in the over-expressed E3 proteins. As FAD is an integral part of E3, its inability to bind to the protein could lead to incorrect folding and structure formation. In support of this, recent work has shown that FAD is a necessary component in the re-folding of functionally active E3, previously unfolded in the presence of guanidine-HCl (N. Beaumont, Glasgow University: personal communication).

In conclusion, possible reasons for the expression of inactive E3 in the *E. coli* host strain are: 1) errors in the sequence of the clones containing a *Taq* amplified insert; 2) the absence of the N-terminal region of the E3 protein; 3) expression in a non-compatible system; 4) the presence of the positively charged histidine tag at the N-terminus of the proteins which could affect correct folding. Expression of the full length E3 clones isolated from the potato leaf cDNA library in the following chapter would provide a confirmation of any of these possibilities as they would be expected to be faithful copies of the E3 gene. A number of methods are available to overcome expression of insoluble protein. For example, 'leaky' plasmids can be left to express protein over a long period of time i.e 20 h without induction and a temperature of 30-35°C would help to reduce the log phase of growth so that nutrients are not depleted too quickly. It is also important to try different expression vectors with different host strains before resorting to use of a eukaryotic expression system.

Chapter 7

Screening a Potato Leaf cDNA Library for Mitochondrial Dihydrolipoamide Dehydrogenase Genes

7.1 Introduction

In chapter 5 the isolation of clones encoding dihydrolipoamide dehydrogenase (E3) by PCR amplification of cDNA, derived from potato leaf RNA was described. The clones expressed E3 subunits of slightly different sizes; however, there was no conclusive evidence to identify them as the separate α and β polypeptides. The disadvantage of using *Taq* DNA polymerase to amplify products for cloning was its lack of proof-reading ability. Therefore, the error frequency of *Taq* may have been responsible for introducing sufficient mutations in the cDNA to yield inactive protein products forming inclusion bodies. In addition, the *Taq* amplified clones were truncated at the 5′ end. Finally, a potato cDNA library was screened for full length E3 gene(s) in order to be able to make valid conclusions concerning the precise sequence. Unfortunately, the only library available was a potato leaf library whereas a tuber library would have been more appropriate. However, it was known that both α and β polypeptides were expressed in leaf tissue although it was found that the α_2 isoform was predominant in this tissue compared with tubers where the β_2 isoform predominated (A.Carmichael, Ph.D thesis, Glasgow University, 1994). Thus, the only difficulty envisaged was isolating the gene encoding the β isoform from the leaf library, if indeed there is a separate gene. It was hoped that the isolated clones would include the 5′ upstream sequence information not present in the *Taq* clones.

The potato leaf library was made in Lambda ZAP II by Dr. H. Hesse at the Max Planck Institute for Plant Molecular Physiology, Germany. Lambda ZAP II is an insertion type vector for uni-directional cloning into the *Eco*RI and *Xho*I sites (Chapter 2, Fig. 2.4). It harbours a complete copy of the phagemid vector, pBluescript SK-, which contains the multiple cloning site. pBluescript is flanked on one side by the initiator region and on the other by the terminator region of f1 filamentous phage origin of replication where DNA synthesis initiates and terminates, respectively. The presence of these regions allows for the *in vivo*

excision of the pBluescript (-)strand which re-circularises, eliminating the need for laborious subcloning procedures. The proteins necessary for initiation and termination of replication, re-circularisation of the (-)-strand phagemid and packaging as filamentous phage are provided by simultaneously infecting the *E. coli* host with lambda helper phage.

Screening of the library was performed in three stages. Nearly 500 000 phage were probed in the first screen to isolate areas on the plates (roughly 1 cm^2) corresponding to a positive signal. Each area would contain negative clones as well as the positive clone resulting from the presence of overlapping plaques as plating was at high density. Therefore, the subsequent rounds of screening were aimed at purifying the positive clones by plating the phage from the primary screen at far lower densities to obtain separate plaques.

7.2 Results

7.2.1 Titre of the Potato Leaf cDNA Library

The first step in the screening protocol was to obtain the titre of the library, i.e. the number of plaque forming units per ml of the library. This was achieved by counting the number of plaques produced by a known dilution of the library stock (section 2.2.30.1). The average titre was 5.8×10^8 $pfu.ml^{-1}$ of the cDNA library.

7.2.2 Probe Preparation

The E3 clones obtained in chapter 5 provided a homologous probe for screening the potato leaf cDNA library. Restriction digests of the pTaq22 clone presented in that chapter had shown that a double digest with *Eco*RI and *Kpn*I produced two bands, one of 750 bp and the other of 950 bp (Chapter 5, Fig. 5.12). The 950 bp band, found to be close to the 5′ end of the gene, was used as the probe for the library. The advantage of this was that any clones hybridising with it would be most likely to harbour inserts containing the initiating codon of the E3 gene and would therefore provide the leader and N-terminal sequence missing from the *Taq*

clones. The likelihood of false negatives was also reduced as the fragment was completely gene specific, containing no vector sequences.

The purified 950 bp product, shown in Fig. 7.1, was estimated to be 75 $ng.\mu l^{-1}$ using the Hoefer minifluorimeter (section 2.2.24.2). The probe was labelled using the Megaprime DNA Labelling kit from Amersham (2.2.29.1) based on the use of random sequence nonamers to prime DNA synthesis on denatured template DNA at numerous sites along its length. Unlabelled dCTP was substituted by a radiolabelled equivalent $[\alpha\text{-}^{32}P]$dCTP (3000 $Ci.mmol^{-1}$) which was incorporated into newly synthesised DNA. The Klenow fragment of DNA polymerase I was used to synthesise DNA and the absence of the 5´-3´ exonuclease activity associated with DNA polymerase I ensured that the labelled nucleotide was not removed as a monophosphate. The protocol was so efficient that only 25 ng of double-stranded template DNA were required to produce probes with a high specific activity.

7.2.3 Screening the Potato cDNA Library for the E3 Gene

For the primary screen approximately 500 000 phage were plated at a density of 60 000 pfu per plate. Plaque lifts were carried out exactly as described in section 2.2.30.2 using the positively charged nylon membrane, Hybond N+.

The membranes were incubated overnight with the heat-denatured probe at 65°C, followed by washing of the membranes at high stringency (section 2.2.29.2). Fig. 7.2A presents one autoradiographed membrane showing a single positive signal (and two putative positives) obtained from a plate containing 60 000 pfu. Probing with the 950 kb fragment appears to be very specific as there is no background hybridisation. A total of 25 positive signals were identified (spread over eight plates) from their position on the autoradiograph and cored from the plates as agar plugs for elution and storage of the phage in 1 ml of SM buffer and 50 μl of chloroform (section 2.2.30.3).

Ten of the positively hybridising plaques were re-screened in the second stage. The aim was to screen approximately 100-200 plaques per sample. This

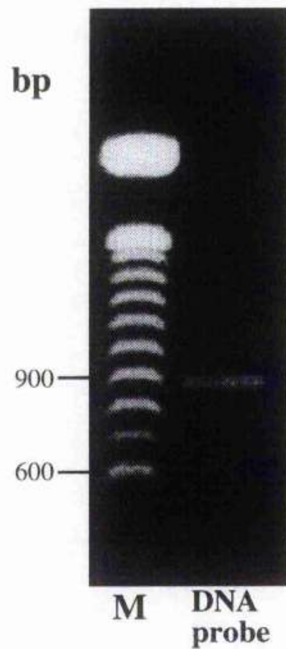


Fig. 7.1: DNA probe (900 bp) prepared from a *Eco*RI/*Kpn*I digest of pTaq22

The plasmid, pTaq22 (25 μg) was digested with the named restriction enzymes and the 900 bp digested fragment gel purified using the method described in section 2.2.18. The final purified product was analysed on a 1% (w/v) agarose gel.

M= size marker

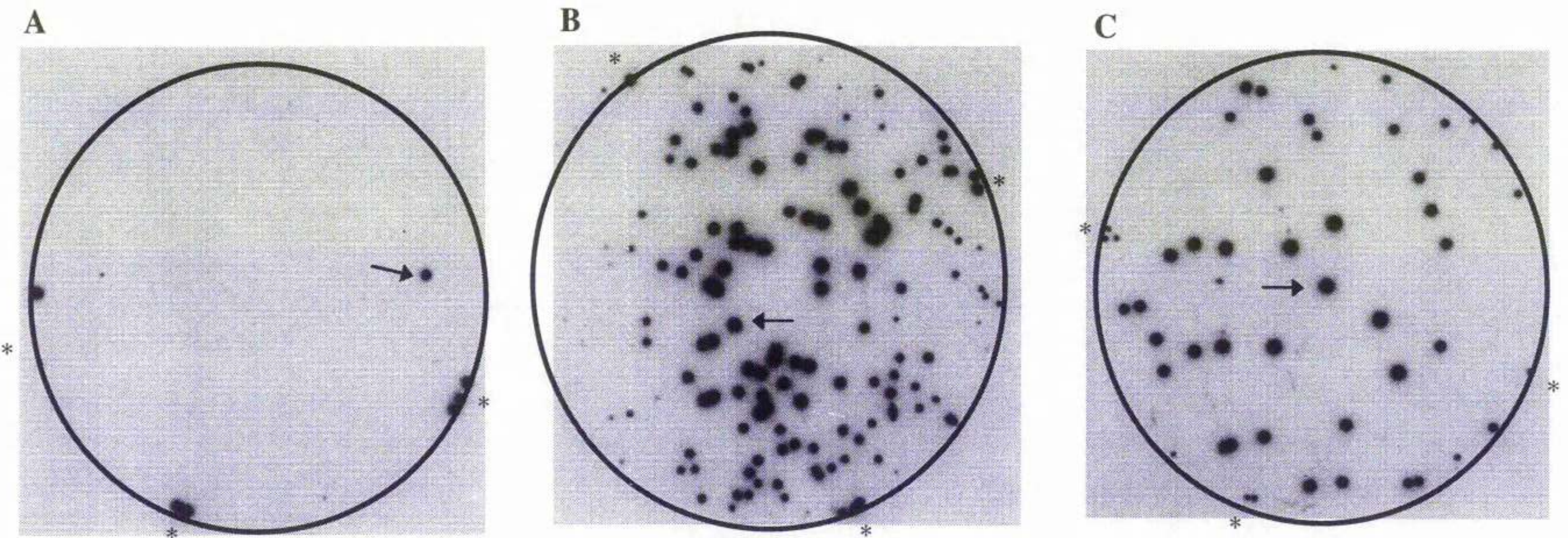


Fig. 7.2: Autoradiographs showing the progress of screening the potato leaf cDNA library for a clone containing the putative E3 gene

Plaque lifts were carried out as described in Materials and Methods (section 2.2.30.2) using Hybond N+ (positively charged nylon membrane). The probe used for hybridisation was the 900 bp fragment of pTaq22, digested with *Eco*RI and *Kpn*I (Fig. 7.1). Hybridisation and washing of the membranes was performed at 65°C. The figure presents the second and third screens of a single hybridising signal indicated on autoradiograph A (first-screen) with an arrow. The hybridising plaque isolated for the third-screen and the plaque isolated for *in vivo* excision of the recombinant pBluescript (SK-) are indicated on autoradiographs B and C respectively. **A**: first-screen (60 000 pfu/plate), **B**: second-screen (100-200 pfu/plate), **C**: third-screen (50-70 pfu/plate) * orientation markers

number is low enough for individual plaques to be distinguished. A large number of the plaques were expected to hybridise with the probe as each sample from the primary screen was enriched for its particular positive clone. The primary screen samples were titred as before and diluted thirty times to achieve the plating density required. Fig. 7.2B is an example of the positive signals obtained with one of the samples after hybridisation to the membranes and autoradiography. The positive signals are very clearly defined and the absence of smearing of the plate surface owing to condensation or the plaque lifts is also evident as there are no streak marks across the film. Two or three positively hybridising plaques from each plate were cored from the agar and stored in 1 ml of SM buffer and 50 μl of chloroform. Although the plaques were separate from each other it was difficult to remove a single hybridising plaque and the estimate is that two negative plaques were also removed with the one positive plaque.

One positive sample from each plate of the secondary screen was re-screened for a final time. This time it was important to ensure that the plaques were well separated so that single plaques could be isolated. This was achieved by plating the phage at a density of 50-70 plaques per plate, using the same procedures as above. Fig. 7.2C shows the positive signals obtained from one of the plates; about a half of the plaques on the plates were positive. A couple of positive plaques were cored from each plate for storage in 1 ml of SM buffer and 50 μl of chloroform. Each selected plaque was expected to contain a single type of clone.

Eight of the purified clones from the tertiary screen were chosen for PCR analysis as described in the next section. For this purpose these samples were amplified using the protocol described in section 2.2.30.4 to provide a phage stock. The average titre of the stocks was increased from 1×10^6 to 6×10^9 $pfu.ml^{-1}$ by this procedure.

7.2.4 PCR Analysis of the Isolated Library Clones

A quick test to ensure that the eight Lambda clones from above were carrying the E3 cDNA was to amplify the insert DNA of the clones with the

degenerate primers used in chapter 5 (Table 5.1). They were designed from N-terminal amino acid sequence of the E3 polypeptide from potato tuber mitochondria to amplify an 80 bp fragment from the 5′ end of the E3 gene for cloning purposes. The cycle parameters and the products obtained are presented in Fig. 7.3. The degenerate primers successfully enabled the amplification of an 80 bp product from all the clones apart from clone pB7-1. Priming was very specific as only one product was obtained. The experiment indicated that the E3 gene had been cloned and that the clones, most probably, contained the extreme 5′ end of the gene as the 80 bp fragment corresponds to amino acids 5 to 28 of the mature E3 protein.

PCR was used again to determine if the upstream sequence encoding the leader sequence and the first seventeen amino acids of E3 were present in the clones. This sequence was important for the full length sequence of the E3 gene as it had been absent from the *Taq* clones. Foreign DNA is cloned directionally into the multiple cloning site of pBluescript SK- which is part of the Lambda ZAP II vector. The 5′ end of the gene is inserted in the *Eco*RI site and the 3′ end into the *Xho*I site. The T3 primer site is adjacent to the *Eco*RI site; therefore it is possible to use the T3 sense primer in conjunction with an antisense, gene specific primer (AGSP) to amplify the region in between the two primers. The antisense, gene specific primer was designed from the sequence already available from the *Taq* amplified clones in chapter 5, (shown below in Table 7.2). The position of AGSP, is indicated in Fig. A (appendix) and was 120 bases downstream from the 5′ end of the *Taq* clones. A product of around 450 bases was obtained from five of the clones (Fig. 7.4), a size which agreed well with the estimate of 350-500 bases from the pea E3 sequence (Bourguignon *et al.*, 1992). It was unclear why other non-specific products were amplified from the clones pB6-2, 3-1 and 7-1 since a control reaction containing the T3 primer only was not carried out.

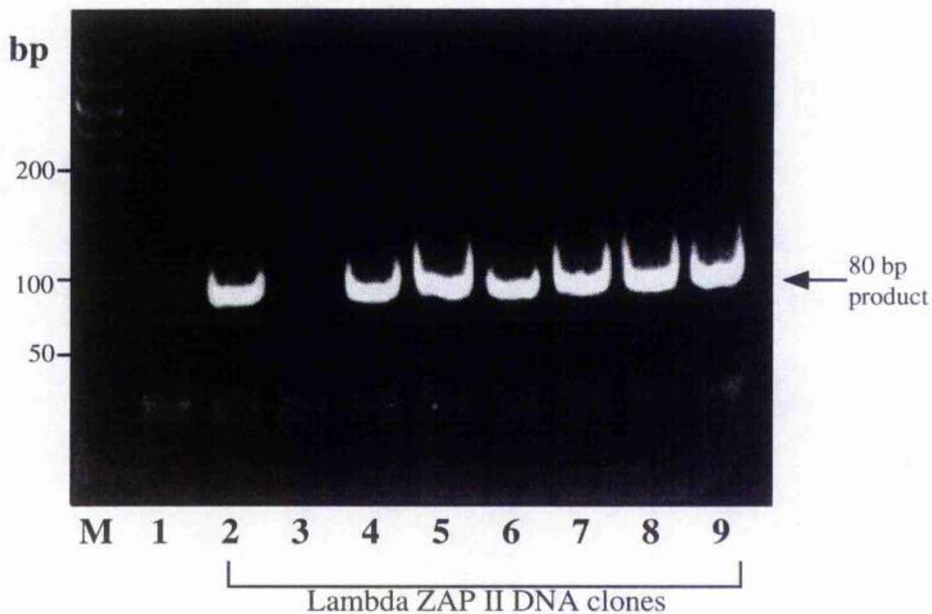


Fig. 7.3: PCR analysis of the 5' region of the library clones using the degenerate primers designed from the N-terminal region of the E3 enzyme
The degenerate primers (1000 pmol per 50 μl reaction) shown in Table 5.1 were used to prime DNA synthesis from Lambda ZAP II target DNA (0.05μg) purified using the method described in section 2.2.30.4. Amplification with *Taq* polymerase was carried out in the presence of 1 mM $MgCl_2$ as described in section 2.2.22. The cycles were as follows: 1 x 5 min at 95°C, 30 x [30 s at 55°C, 30 s at 65°C and 3 s at 95°C], 1 x 2 min at 65°C and 1 x 3 min at 72°C. One fifth of the sample was analysed on a 12% (w/v) polyacrylamide gel.

M: size marker, 1: negative DNA control, 2: pB6-1, 3: pB7-1, 4: pB9-2, 5: pB3-1, 6: pB9-3, 7: pB10-1, 8:pB8-1, 9:pB6-2

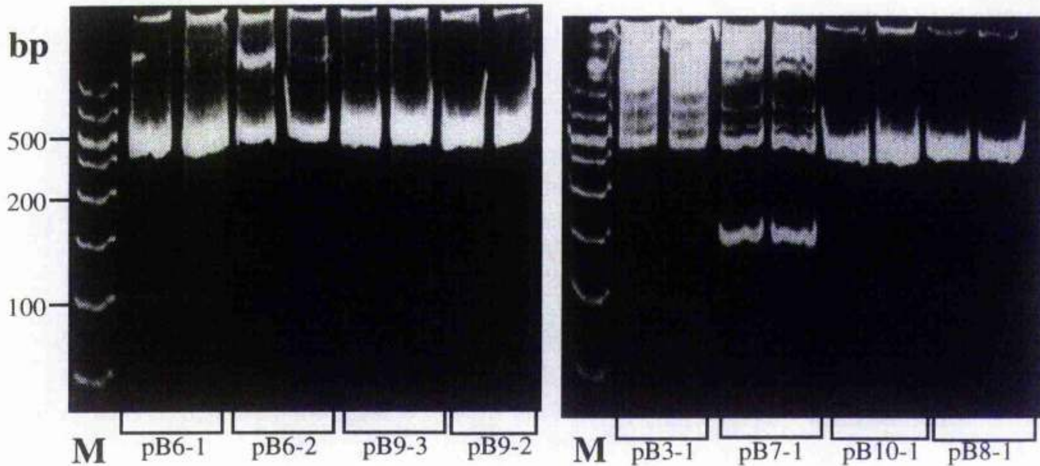


Fig. 7.4: PCR amplification of the 5' region of the library clones - priming with T3 and an E3 gene specific primer (AGSP)

The amplification of the Lambda ZAP II DNA (0.05 μg), purified as in section 2.2.30.4, with *Taq* polymerase was carried out in duplicate, in the presence of 1 mM $MgCl_2$. The cycling conditions were as follows: 1 x 5 min at 95°C, 35 x [1 min at 50°C, 2 min at 60°C, 30 s at 95°C] and 1 x 10 min at 72°C. The Lambda ZAP II clones were analysed on a 12% (w/v) polyacrylamide gel and are named below their respective lanes (not in the same order as Fig. 7.3).

M= size marker

Table 7.2

Primer sequences used for the amplification of the 5′ upstream sequences of the isolated library clones.

Primer	Sequence (5′-3′)	Number of Bases
AGSP	ATTCTAGA TGAAGAAGTGCCTTAGAAGG	28
T3	ATTAACCCTCACTAAAGGGA	20

7.2.5 *In Vivo* Excision of the pBluescript Phagemid from Lambda ZAP II

The use of Lambda ZAP II as a cloning vector eliminated the need for subcloning the inserts into plasmid vectors. As explained in the introduction, the vector contains the pBluescript SK- phagemid and in *E. coli* it can be excised by f1 or M13 based helper phage. The protocol was followed as described in section 2.2.31. The success of the *in vivo* excision was apparent as many colonies appeared on the plates which were expected to contain the double-stranded pBluescript plasmid with the cloned insert. pBluescript was able to be propagated as a plasmid, as well as a phagemid, as it carried the colE1 origin of replication. The helper phage were not ampicillin resistant and contained an amber mutation which meant that they were unable to co-infect SOLR cells (non-suppressing strain).

Two colonies from each plate were isolated and used directly for PCR analysis to confirm that they still contained the E3 gene. Each colony was suspended in 100 µl of sterile water and 5 µl used for PCR amplification with the degenerate primers designed in chapter 5 (Table 5.1). The reaction mix and cycling conditions were described in Fig. 7.3 for the amplification of the 80 bp fragment from the 5′ end of the E3 gene. The Triton X-100 in the *Taq* polymerase buffer

was sufficient to solubilise the SOLR cells. The same 80 bp product as in Fig. 7.3 was amplified from all the excised clones verifying the identity of the inserts.

7.2.6 Restriction Digest Analysis

Owing to the incomplete sequence information available so far it was impossible to know if the α and β polypeptides were encoded by one gene or two very similar genes. It was already known that the α form predominated in potato leaf mitochondria and hence the α gene was more likely to be isolated from the library. In order to narrow down the number of clones for sequencing, restriction mapping was necessary to ascertain whether all the cloned inserts represented the same gene. Of course, two very closely related genes would not necessarily be differentiated by this method.

Restriction maps of six clones were prepared using a series of restriction enzymes. All the clones were found to have the same restriction map apart from clone pB3-1. On closer inspection it was found that this clone also had the same map but had been inserted into the vector in the antisense orientation. This result was unexpected as the clones had been isolated from a unidirectional library. Fig 7.5 a and 7.5b present the final restriction maps.

The two fragments of 700 and 950 bp obtained upon *Kpn*I digestion of the clones were identical with those of the *Kpn*I digested *Taq* clones (Chapter 5, Fig. 5.12). In addition, the existence of different polyadenylation start sites was indicated because the *KpnI/Xho*I fragment near the 3′ end of the gene was slightly larger in clones pB9-3 and pB8-1. The restriction maps provided important reference points for the subsequent sequence analysis, but the existence of two different cDNAs was still unclear.

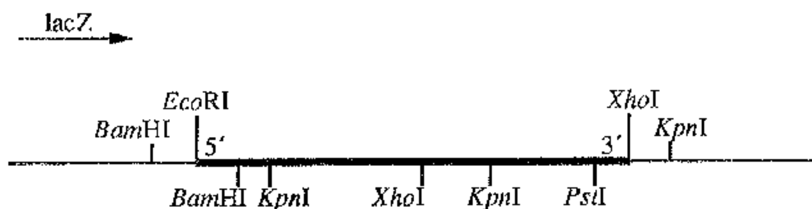


Fig. 7.5a: Restriction map of clones pB9-3, pB10-1, pB9-2, pB6-1 and pB8-1

Direction of transcription of the lacZ gene is indicated.

Fig. 7.5b: Restriction map of clone pB3-1

Direction of transcription of the lacZ gene is indicted.

7.2.7 Sequence Analysis of the Potato Leaf cDNA Clones

Three clones were selected for automated sequencing; pB3-1, pB9-3 and pB10-1. Fig 7.6 shows the plasmid preparations used for the sequencing reactions after *Kpn*I digestion. Primer walking was performed using the forward and reverse primers (Chapter 5, Table 5.3) designed for sequencing the *Taq* clones: pTaq16, pTaq22 and pTaq23. The primers had been designed from the sequence of the sense-strand of pTaq22 obtained by sequencing the Exo III deletion clones in chapter 5.

The nucleotide sequence of the cDNA from pB3-1 and the deduced amino acid sequence of the protein is shown in Fig. 7.7. The ATG triplet designated as the translation start site was confirmed by the similarity of the surrounding sequence (ACAA<u>ATG</u>GC) with the favoured sequence (AACA<u>ATG</u>GC) flanking the consensus functional plant-initiator codon (Lütcke *et al.*, 1987). This is the only upstream ATG codon in the sequence. The synthesis of potato mitochondrial E3 will therefore begin with methionyl-alanine as do most plant proteins. The open reading frame encodes a protein with 504 amino acid residues. The identity of the encoded protein was confirmed by comparison with the N-terminal sequence of the purified potato E3. The N-terminus corresponds to Ala35, indicating that the protein is synthesised with a 34-amino acid leader sequence. Thus, the mature E3 consists of 470 amino acids, giving a protein of M_r 49 561. This is exactly the same as the mass of the β polypeptide obtained by electrospray mass spectrometry. The proposed polyA signal (AATAAA) is indicated on Fig. 7.7. The positions of the restriction sites shown agrees with the restriction map shown in Fig. 7.5B. The leader sequence exhibits a typical enrichment for alanine (9%), leucine (15%), arginine (12%), threonine (15%) and serine (20%) and contains no negatively charged residues. The sequence of pB9-3 was exactly the same as pB3-1, apart from a two base deletion at position 766 (indicated by 'XX' on Fig. 7.7). This deletion was expected to be a cloning artefact because translation of the sequence

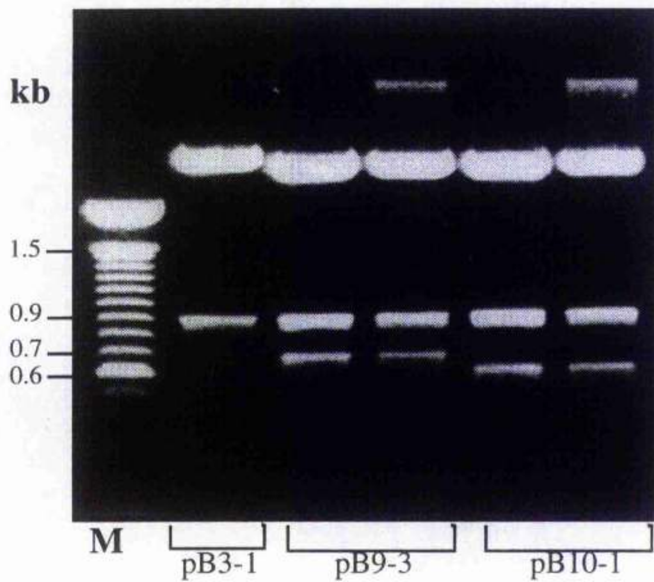


Fig. 7.6: Analysis of the purified plasmids pB3-1, pB9-3 and pB10-1 used for sequencing the E3 gene from potato

The purified plasmids (1 μg) were digested with *Eco*RI and *Kpn*I and the fragments analysed on a 1% (w/v) agarose gel.

M= size marker

Fig. 7.7: Nucleotide sequence of potato E3 and its deduced amino acid sequence

The bases and amino acids are numbered separately. The N-terminal sequence is underlined and vector sequences are in bold. The position of the two base deletion in clone pB9-3 is marked with an 'X' and the active-site cysteins are marked with '§'. The proposed poly A signal is indicated by '◆'.

```
                                                  GGCACGAGCAACAACTCA   18

ACAGTTCACATCGAATCACCAAAATTGACCAAAAAACCTACCGGAAAGTTCACCGGAATA   78
                                 ⌐→ Leader sequence
CACGATTCCGATAACTGACCTTTACAAATGGCGATTTCTACCTTAGCTAGACGAAAGGCT  138
                                 M  A  I  S  T  L  A  R  R  K  A     11

ACCACGTTTCTGTCATCCAGACTCCTTTACAGCACTTCCAAGTATTCGTTCTCTTTAACC  198
T  T  F  L  S  S  R  L  L  Y  S  T  S  K  Y  S  F  S  L  T     31

         ⌐→ E3
               BamHI
AGAGGTTTCGCTTCAGGATCCGATGAGAACGACGTCGTTGTTATCGGTGGTGGTCCCGGA  258
R  G  F  A  S  G  S  D  E  N  D  V  V  V  I  G  G  G  P  G     51

GGCTATGTGGCGGCGATCAAAGCAGCACAGCTTGGGCTTAAAACTACATGTATCGAGAAA  318
G  Y  V  A  A  I  K  A  A  Q  L  G  L  K  T  T  C  I  E  K     71

      KpnI           §            §
CGTGGTACCCTCGGTGGTACTTGCCTTAACGTTGGTTGTATTCCTTCTAAGGCACTTCTT  378
R  G  T  L  G  G  T  C  L  N  V  G  C  I  P  S  K  A  L  L     91

CATTCGTCCCACATGTTTCATGAAGCGCAACATTCATTTGCCAATCATGGTGTGAAGTTT  438
H  S  S  H  M  F  H  E  A  Q  H  S  F  A  N  H  G  V  K  F    111

TCTTCTGTCGAAGTwGATCTCCCTGCCATGATGGGACAAAAAGATAAAGCCGTGTCTAAC  498
S  S  V  E  V  D  L  P  A  M  M  G  Q  K  D  K  A  V  S  N    131

TTAACACGAGGTATAGAGGGTCTATTCAAGAAAAACAAAGTAAACTATGTGAAGGGATAT  558
L  T  R  G  I  E  G  L  F  K  K  N  K  V  N  Y  V  K  G  Y    151

GGTAAATTCCTCTCCCCTTCTGAAATTTCTGTTGACACTGTCGAAGGTGGTAACTCTGTT  618
G  K  F  L  S  P  S  E  I  S  V  D  T  V  E  G  G  N  S  V    171

GTTAAAGGCAAAAATATTATTATTGCAACTGGCTCTGATGTCAAAGGTCTACCTGGCATA  678
V  K  G  K  N  I  I  I  A  T  G  S  D  V  K  G  L  P  G  I    191

ACCATTGATGAGAAGAAAATTGTGTCATCCACCGGAGCATTAGCTTTGACTGAAATTCCA  738
T  I  D  E  K  K  I  V  S  S  T  G  A  L  A  L  T  E  I  P    211

                              XX
AAAAGATTAGTTGTTATTGGTGCTGGCTATATAGGCCTTGAGATGGGATCTGTCTGGGGT  798
K  R  L  V  V  I  G  A  G  Y  I  G  L  E  M  G  S  V  W  G    231

CGCCTCGGCTCAGAGGTGACTGTTGTTGAATTTGCATCTGATATTGTTCCAACCATGGAT  858
R  L  G  S  E  V  T  V  V  E  F  A  S  D  I  V  P  T  M  D    251

GGTGAAGTTCGCAAGCAATTTCAACGTGCTCTTGAGAAGCAAAAGATGAAATTCATGCTA  918
G  E  V  R  K  Q  F  Q  R  A  L  E  K  Q  K  M  K  F  M  L    271

AACACTAAGGTGGTGTCAGTTGATGCTACAGGTGATGGTGTGAAATTGACCCTTGAACCT  978
N  T  K  V  V  S  V  D  A  T  G  D  G  V  K  L  T  L  E  P    291

                             XhoI
TCAGCTGGTGGTGATCAAACTATTCTCGAGGCTGATGTTGTTCTCGTCTCTGCTGGAAGA 1038
S  A  G  G  D  Q  T  I  L  E  A  D  V  V  L  V  S  A  G  R    311

GTACCATTCACTTCAGGACTTGGATTGGACACGATAGGAGTTGAAACTGACAAGGCTGGT 1098
V  P  F  T  S  G  L  G  L  D  T  I  G  V  E  T  D  K  A  G    331
```

AGAATCTTGGTCAATGAACGTTTTGCCACTAACGTCCCAGGGCTACATGCAATTGGTGAT 1158
R I L V N E R F A T N V P G V H A I G D 351

GTCATTCCTGGACCAATGCTGGCTCACAAGGCAGAGGAGGATGGCCTTGCTTGTGTAGAG 1218
V T P G P M L A H K A E E D G V A C V E 371

KpnI
TTCATTGCAGGTAAGGAGGGTCATGTAGACTATGATATGGTACCTGGTGTCGTTTACACC 1278
F I A G K E G H V D Y D M V P G V V Y T 391

CACCCGGAAGTGGCTTATGTTGGGAAAACCGAGGAACAGGTAAAGTCACTCGGAGTTGAT 1338
H P E V A Y V G K T E E Q V K S L G V D 411

TATCGTGTTGGCAAATTCCCCTTCCTAGCAAACAGTAGAGCCAAGGCAATTGATGATGCT 1398
Y R V G K F P F L A N S R A K A I D D A 431

GAGGGAATTGTCAAGGTAATTGCTGAGAAAGAGAGTGACAAGATATTGGGCGTCCATATT 1458
E G I V K V I A E K E S D K I L G V H I 451

ATGTCACCTAATGCAGGGGAGCTTATTCACGAAGCTGTACTGGCATTGCAGTACGGAGCA 1518
M S P N A G E L I H E A V L A L Q Y G A 471

TCAAGTGAGGACATTGCTCGTACTTGTCATGCACATCCAACAATGAGTGAGGCACTCAAA 1578
S S E D I A R T C H A H P T M S E A L K 491

GAAGCAGCCATGCCCACTTATGACAAGCCCATCCACATGTAGGCCTGTGTTATATCCTCG 1638
E A A M A T Y D K P I H M * 504

ATAGCCTCATTTTTCTTATTCCGTTGAGAATCTTGAACGCTTAGAATAGGCTTTTTGTTG 1698

CTGCCGTTACTCAAGCAGCCCATATTCCCCTTAATTCAGGACTCCAAATACACTATATGA 1758

ATGTTTATCTTCTGCATTTTTCCATAATAAAGATGAGATGTAACCTGCTTTAAAAAAAAA 1818

PstI BamHI
AAAAAAAAAAACCTCGTGCCGAA**TCCTGCAGCCCGGGGGATCCACTAGTTCT**

following it produced many stop codons. Incomplete sequence information was obtained for pB10-1 and it was not analysed further.

A comparison of all the DNA sequences is presented in Fig. A (Appendix) which also includes the sequence of the PCR derived clones from chapter 5. The sequences of pTaq22, pTaq23 and pTaq16 were missing upstream sequence encoding the N-terminal and leader sequences due to the cloning procedures employed. Nevertheless, all the sequences are identical with each other, except at the bases indicated. Some base changes caused amino acid differences as indicated in bold in Fig.B (Appendix). A comparison of the clones reveals that all the differences, except one (at residue 19), are in the PCR derived clones. The deduced amino acid sequence of all the clones (Fig. B, Appendix) found that the stop codon is in exactly the same position (at base number 1618 in Fig. A, Appendix). The upstream sequences of pTaq16 and pTaq22 are expected to be the same as the two library clones because they encompass the N-terminal sequence derived from protein sequencing. However, as explained in chapter 5, the sequence of pTaq23 begins later than the other *Taq* amplified clones merely because the sequencing strategy for this clone was started further downstream. There could be some differences in the upstream sequence of pTaq23 but this was not investigated further.

7.3 Discussion

Confidence in the identity of the clones isolated from the potato leaf cDNA library comes from the fact that the predicted N-terminal sequence matches the sequence derived from protein sequencing. The cDNA clone for potato E3 revealed a 1512-nucleotide open reading frame encoding a 504 residue protein. The first 34 residues were found to bear some of the hallmarks of a mitochondrial leader sequence (a positive hydrophobic-polar design).

Surprisingly, the predicted M_r of the clone pB3-1, 49 561, was found to correspond exactly with the M_r of the β polypeptide as determind by electrospray

mass spectrometry. Preliminary protein studies had shown that the β_2 isoform was the minor species in pea leaf mitochondria, with the α_2 form predominating. Hence, there would have been a higher probability of cloning the α gene if there were two separate genes encoding the α and β polypeptides. The reasons for cloning the apparently minor species remains unclear at this stage. Preliminary copy number analysis had been inconclusive in determining whether the E3 gene was in single copy; therefore, it was impossible to conclude that another gene encoding the α polypeptide did not exist in the cDNA library. Furthermore, although the *Taq* amplified clones did not match the sequence of pB3-1 exactly they were thought to represent the gene for the β polypeptide. Initially, the assumption made in selecting *Taq* amplified clones for sequencing was that pTaq16 and pTaq23 may encode the β polypeptide and pTaq22 may encode the α polypeptide. However, there were no obvious differences in amino acid residues consistent only with pTaq22. The predicted M_r values (including the 5′ amino acid sequence absent from the clones) were 49 558 (pTaq22), 49 517 (pTaq23) indicating identity with the β polypeptide and 46 060 for the protein encoded by pTaq16; none of them showing a difference in M_r of 116 found between the native α and β polypeptides. Characteristically for mitochondrial E3s, the M_r of the β polypeptide obtained by electrospray mass spectrometry is lower than the value of 56 000 based on its mobility in SDS-PAGE.

The putative element signalling polyadenylation (AATAAA) in higher eukaryotes was found at the 3′ end of the clones (Fig. A, Appendix), twenty bases from the polyA tail of pB3-1. However, since the polyA tail of pTaq23 began immediately at this sequence and those of pB9-3, pTaq16, pTaq22 and pTaq23 began a various points later in the sequence it was not known if there were other related signals before and after the proposed sequence. Multiple polyA signals are quite common in plant genes as it appears that the standard polyadenylation signal is not as strong as it is in other eukaryotic genes.

There are various possibilities at the genetic level for the generation of the α and β polypeptides, having a mobility difference visible on SDS-PAGE and a M_r difference of 116. A deletion of three bases causing the loss of one amino acid or the cleavage of the first or terminal amino acid during post-translational processing could generate the α polypeptide since it is smaller by 116 Da. Neither of these were found to occur and cleavage of the first last amino acids would not have generated proteins with the small mass difference since their M_r values would be 71 (alanine) and 149 (methionine), respectively. A post-translational modifcation involving an addition would not be possible since the gene for the larger polypeptide had been cloned. Finally, the α polypeptide could be a physiologically insignificant form generated by an unusual type of allelic variation which allowed the separation of the two polypeptides.

It is feasible that there are two separate genes having a number of conservative substitutions in amino acids which coincidentally leads to polypeptides having a small difference in mass. The validity of this hypothesis is apparent when comparing the amino acid sequence of the potato and pea mitochondrial E3 which have 85% identity in sequence, giving rise to a difference in M_r of only 168. Copy number analysis is required to prove the presence of a second gene. Furthermore, the presence of a second transcript originating from a single gene by RNA splicing, could be investigated by analysis of mRNA by Northern blots.

The overall sequence similarities at the amino acid level between E3 from potato and E3s from eukaryotic sources is very high (Table 7.3). It has 92.8% similarity with pea E3 and 72-73% with yeast, human and pig E3s. Although potato E3 is found to be less similar to prokaryotic E3s, the values are still high (60-70%). Therefore, the potato enzyme in not expected to be very different from these E3s in secondary and tertiary structure. The primary amino acid sequence comparisons between the potato clone pB3-1 and several other E3s (Fig. 7.8) showed that important sites for enzyme activity are highly conserved. In fact, the amino acid sequence surrounding the two redox-active cysteines (residues 75-91

located in the FAD-binding domain) is identical in all the E3s of eukaryotic origin. Likewise, two regions in the mature potato E3 protein are very similar to the 'fingerprint' structural motif in the adenine-binding domain of either FAD (residues 42-63) or NAD (residues 214-241) of many enzymes, including the dihydrolipoamide dehydrogenases. The presence of several clustered regions of high sequence similarity suggests that all these proteins have evolved from a common ancestor.

Table 7.2

Percent identity and similarity of several E3s with potato E3

Source of E3	Percentage identity	Percentage similarity	Accession Number	Reference
Pea (*Pisum sativum*)	85.2	92.8	X63464	Bourguignon *et al.* (1992)
Trypanosoma brucei	56.2	75.2	X70646	Else *et al.* (1993)
Human	54.6	73.0	J03490	Otulakowski & Robinson. (1987)
Pig	54.6	73.6	J03489	Otulakowski & Robinson. (1987)
Yeast	55.7	72.2	J03645	Browning *et al.* (1988)
Pseudomonas putida (*lpd*3)	54.4	73.3	X55704 S74684	Palmer *et al.* (1991)
Azotobacter vinelandii	48.1	68.0	P18925	Westphal & de Kok. (1988)
Klebsiella pneumoniae (*AcoD**)	40.2	61.1	U30887	Peng *et al.* (1996)
Staphylococcus aureus	44.1	63.8	X58434	Hemila. (1991)
Bacillus stearothermophilus	47.9	64.9	X53560	Borges *et al.* (1990)
E. coli K12	43.0	62.6	P00391	Stephens *et al.* (1983)

* *Klebsiella pneumoniae* Acetoin dehydrogenase system (*acoD* gene for E3)

Fig. 7.8: Primary amino acid sequence comparison of potato E3 (clone pB3-1) and E3s from other sources

The comparison was created using the pileup program of the GCG sequence analysis package. The boxed regions show the conserved amino acid domains. Abbreviations for the source of the E3 sequence: pB3-1,potato; Tryp,*Trypanosoma brucei*; Pseud,*Pseudomonas putida*; Azo, *Azotobacter vinelandii*; Aco-D,*Klebsiella pneumoniae*; Staph, *Staphylococcus aureus*; Bac, *Bacillus stearothermophilus*; Coli, *Escherichia coli*.

§=redox active cysteines

*= amino acids conserved in higher Eukaryotes

+= amino acids conserved in all E3s

```
                                                            FAD-domain
        1                                                           50
pB3-1   MAISTLARRK ATTFLSSRLL YSTSKYSF.. ..SL.TRGFA SGSDENDVVV
Pea     MAMANLARRK GYSLLSSETL ....RYSF.. ..SLRSRAFA SGSDENDVVI
Tryp    .......... .......... .......... ....MFRRCF PIFNPYDVVV
Human   ....MQSWSR VYCSLAKRGH FNRISHGLQG LSAVPLRTYA DQPIDADVTV
Pig     ....MQSWSR VYCTLAKRGH FNRIAHGLQG VSAVPLRTYA DQPIDADVTV
Yeast   .......... ...MLRIRSL LNNKRAFSST VRTLTINK.. ....SHDVVI
Pseud   .......... .......... .......... .......... ..MKSYDVVI
Azo     .......... .......... .......... .......... .MSQKFDVIV
Aco-D   .......... .......... .......... .......... .MHDKYDVLI
Staph   .......... .......... .......... ......MVVG DFPIETDTIV
Bac     .......... .......... .......... ......MVVG DFAIETETLV
Coli    .......... .......... .......... .......... STEIKTQVVV
                                                         **

        51                                           §      §  100
pB3-1   IGGGPGGYVA AIKAAQLGLK TTCIEK.... ..RGTLGGTC LNVGCIPSKA
Pea     IGGGPGGYVA AIKAAQLGFK TTCIEK.... ..RGALGGTC LNVGCIPSKA
Tryp    VGGGPGGYVA AIKAAQLGLK TACVEK.... ..RGALGGTC LNVGCIPSKA
Human   IGSGPGGYVA AIKAAQLGFK TVCIEK.... ..NETLGGTC LNVGCIPSKA
Pig     IGSGPGGYVA AIKAAQLGFK TVCIEK.... ..NETLGGTC LNVGCIPSKA
Yeast   IGGGPAGYVA AIKAAQLGFN TACVEK.... ..RGKLGGTC LNVGCIPSKA
Pseud   IGGGPGGYNA AIRAGQLGLT VACVEG.... ..RSTLGGTC LNVGCMPSKA
Azo     IGAGPGGYVA AIKSAQLGLK TALIEKYKGK EGKTALGGTC LNVGCIPSKA
Aco-D   IGGGPGGYVA AIRAGQLGLR TVLVEK.... ...QHLGGIC LNWGCIPTKA
Staph   IGAGPGGYVA AIRAAQLGQK VTIVEKG... ...N.LGGVC LNVGCIPSKA
Bac     VGAGPGGYVA AIRAAQLGQK VTIVEKG... ...N.LGGVC LNVGCIPSKA
Coli    LGAGPAGYSA AFRCADLGLE TVIVERY... ...NTLGGVC LNVGCIPSKA
        *+ ++*++*+ +****+++ *  * * +*           +++*+ ++*++*+*++

        101                                                    150
pB3-1   LLHSSHMFHE AQ.HSFANHG VKF.SSVEVD LPAMMGQKDK AVSNLTRGIE
Pea     LLHSSHMYHE AK.HSFANHG VKV.SNVEID LAAMMGQKDK AVSNLTRGIE
Tryp    LLHATHMYHD AHA.NFERYG LMGGAGVTMD VAKMQQQKEK SVNGLTSGVE
Human   LLNNSHYYHM AHGTDFASRG IEM.SEVRLN LDKMMEQKST AVKALTGGIA
Pig     LLNNSHYYHM AHGKDFASRG IEM.SEVRLN LEKMMEQKSN AVKALTGGIA
Yeast   LLNNSHLFHQ MHTEA.QKRG IDVNGDIKIN VANFQKAKDD AVKQLTGGIE
Pseud   LLHASELYEA ASGDEFAHLG IEVKPT..LN LAQMMKQKDE SVTGLTKGIE
Azo     LLDSSYKFHE AH.ESFKLHG IS.TGEVAID VPTMIARKDQ IVRNLTGGVA
Aco-D   LLHGAEV... AHTITHASQL GISVGEVNVD LQKLVQFSRT VSQQLTAGVA
Staph   LLHASHRFVE A..QHSENLG V.IAESVSLN FQKVQEFKSS VVNKLTGGVE
Bac     LISASHRYEQ A..KHSEEMG I.KAENVTID FAKVQEWKAS VVKKLTGGVE
Coli    LLHVAKVIEE A..KALAEHG I.VFGEPKTD IDKIRTWKEK VINQLTGGLA
        +*   *   *   *          *                *  **    *  ++ +

        151                                                    200
pB3-1   GLFKKNKVNY VKGYGKFLSP SEISVDTVEG ......GNSV VKGKNIIIAT
Pea     GLFKKNKVTY VKGYGKFVSP SEISVDTIEG ......ENTV VKGKHIIIAT
Tryp    YLLKKNKVTY YKGEAGFVTP NTLNVKGIDG ......KDEA IEAKNTIIAT
Human   HLFKQNKVVH VNGYGKITGK NQVTATKADG ......GTQV IDTKNILIAT
Pig     HLFKQNKVVR VNGYGKITGK NQVTATKADG ......STEV INTKNILIAT
Yeast   LLFKKNKVTY YKGNGSFEDE TKIRVTPVDG LEGTVKEDHI LDVKNIIVAT
Pseud   YLFRKNKVDW IKGWGRLDGV GKVVVKAEDG ......SETA LQAKDIVIAT
Azo     SLIKANGVTL FEGHGKLLAG KKVEVTAADG ......SSQV LDTENVILAS
Aco-D   YLLKKNGVRV IDGTARLRGK GQITVEDARG ......EARD YRADHVILAT
Staph   GLLKGNKVNI VKGEAYFVDN NSLRVMDEKS .......AQT YNFKNAIIAT
Bac     GLLKGNKVEI VKGEAYFVDA NTVRVVNGDS .......AQT YTFKNAIIAT
Coli    GMAKGRKVKV VNGLGKFTGA NTLEVEGENG .......KTV INFDNAIIAA
          * * **+     +                   *                  *+*
```

```
FAD-domain  NAD-domain

        201                                                               250
pB3-1   GSDVKGLPGI TIDEKKIVSS TGALALTEIP KRLVVIGAGY IGLEMGSVWG
Pea     GSDVKSLPGV TIDEKKIVSS TGALALSEIP KKLVVIGAGY IGLEMGSVWG
Tryp    GSEPTALPFL PFDEKVVLSS TGALALQQVP KKMVVIGGGV IGLEIGSVWA
Human   GSEVTPFPGI TIDEDTIVSS TGALSLKKVP EKMVVIGAGV IGVEIGSVWQ
Pig     GSEVTPFPGI TIDEDTVVSS TGALSLKKVP EKMVVIGAGV IGVELGSVWQ
Yeast   GSEVTPFPGI EIDEEKIVSS TGALSLKEIP KRLTIIGGGI IGLEMGSVYS
Pseud   GSEPTPLPGV TIDNQRIIDS TGALSLPQVP KHLVVIGAGV IGLELGSVWR
Azo     GSKPVEIPPA PVDQDVIVDS TGALDFQNVP GKLGVIGAGV IGLELGSVWA
Aco-D   GARPRALPGI APDGEHIWTY FEALRPKLLP KSLLLIGSGA IGVEFASLYN
Staph   GSRPIEIPNF KFGK.RVIDS TGALNLQEVP GKLVVVGGGY IGSELGTAFA
Bac     GSRPIELPNF KFSN.RILDS TGALNLGEVP KSLVVIGGGY IGIELGTAYA
Coli    GSRPIQLPFI PHEDPRIWDS TDALELKEVP ERLLVMGGGI IGLEMGTVYH
        +*    +     **    ** **+  *    +    ***+*+    +* + ****

        251                                                               300
pB3-1   RLGSEVTVVE FASDIVP.TM DGEVRKQFQR ALEKQKMKFM LNTKVVSVDA
Pea     RLGSEVTVVE FASEIVP.TM DAEIRKQFQR SLEKQGMKFK LKTKVVGVDT
Tryp    RLGSDVTVVE FAPRCAP.TL DSDVTDALVG ALKRNGEDEV PMTGIEGVNG
Human   RLGADVTAVE FLGHVGGVGI DMEISKNFQR ILQKQGFKFK LNTKVTGATK
Pig     RLGADVTAVE LLGHVGGIGI DMEVSKNFQR ILQKQGFKFK LNTKVIGATK
Yeast   RLGSKVTVVE FQPQIGA.SM DGEVAKATQK FLKKQGLDFK LSTKVISAKR
Pseud   RLGSQVTVIE YLDRICP.GT DTETAKTLQK ALAKQGMVFK LGSKVTQATA
Azo     RLGAEVTVLE AMDKFLP.AV DEQVAKEAQK ILTKQGLKIL LGARVTGTEV
Aco-D   DLGCKVTLVE LASQILPVE. DAEVSAAVRK SFEKRGIQIH TQTLVTQVQL
Staph   NFGSEVTILE GAKDILG.GF EKQMTQPVKK GMKEKGVEIV TEAMAKSAEE
Bac     NFGTKVTILE GAGEILS.GF EKQMAAIIKK RLKKKGVEVV TNALAKGAEE
Coli    ALGSQIDVVE MFDQVIP.AA DKDIVKVFTK RISKK.FNLM LETKVTAVEA
        *  +   **  *+          *     * ***  * **       * ***

                                      NAD-domain  Central-domain
        301                                                               350
pB3-1   TGDG..VKLT LEPSAGGDQT ILEADVVLVS AGRVPFTSGL GLDTIGVETD
Pea     SGDG..VKLT VEPSAGGEQT IIEADVVLVS AGRTPFTSGL NLDKIGVETD
Tryp    TNNGSIALTL EVEQAGGQAE TLHCDALLVS VGRRPYTAGL GLEKNNVSLN
Human   KSDGK.IDVS IEAASGGKAE VITCDVLLVC IGRRPFTKNL GLEELGIELD
Pig     KSDGN.IDVS IEAASGGKAE VITCDVLLVC IGRRPFTQNL GLEELGIELD
Yeast   NDDKNVVEIV VEDTKTNKQE NLEAEVLLVA VGRRPYIAGL GAEKIGLEVD
Pseud   SADG..VSLV LEPAAGGTAE SLQADYVLVA IGRRPYTKGL NLESVGLETD
Azo     KNKQ..VTVK FVDAEGEKSQ AF..DKLIVA VGRRPVTTDL LAADSGVTLD
Aco-D   TDTGVRCTL. ...NNTGGEY SQDVERVLLA VGVQPNIEDL GLETLGVELD
Staph   TDNGVKVTY. ..EA.KGEEK TIEADYVLVT VGRRPNTDEL GLEELGVKFA
Bac     REDGVTVTY. ..EA.NGETK TIDADYVLVT VGRRPNTDEL GLEQIGIKMT
Coli    KEDGIYVTM. ..EGKKAPAE PQRYDAVLVA IGRVPNGKNL DAGKAGVEVD
          **          **        ** **   +*+**   +    *    * * *

        351                                                               400
pB3-1   KAGRILVNER FATNVPGVHA IGDVIP.GPM LAHKAEEDGV ACVEFIAG..
Pea     KLGRILVNER FSTNVSGVYA IGDVIP.GPM LAHKAEEDGV ACVEYLAG..
Tryp    ERGFVKIGSH FETNVAGVYA IGDVVDKGPM LAHKAEDEGV ACAEILAG..
Human   PRGRIPVNTR FQTKIPNIYA IGDVVA.GPM LAHKAEDEGI ICVEGMAG..
Pig     PRGRIPVNTR FQTKIPNIYA IGDVVA.GPM LAHKAEDEGI ICVEGMAG..
Yeast   KRGRLVIDDQ FNSKFPHIKV VGDVTF.GPM LAHKAEEEGI AAVEMLKT..
Pseud   KRGMLAQRTP .PTSVPGVWV IGDVTS.GPM LAHKAEDEAV ACIERIAG..
Azo     ERGFIYVDDY CATSVPGVYA IGDVV.RGAM LAHKASEEGV VVAERIAG..
Aco-D   .RGFIKTDAA CRTNVFGLYA IGDVAG.PPC LAHKASHEGV LCVETLAGVE
Staph   DRGLLEVDKQ SRTSISNIYA IGDIVP.GLP LAHKASYEAK VAAEAIDG..
Bac     NRGLIEVDQQ CRTSVPNIFA IGDIVP.GPA LAHKASYEGK VAAEAIAG..
Coli    DRGFIRVDKQ LRTNVPHIFA IGDIVG.QPM LAHKGVHEGH VAAEVIAG..
          +** ** * * *        * *++*  *** ++++**  *    * *  **
```

Central-domain Interface-domain

```
        401                                                  450
pB3-1   KEGHVDYDMV PGVVYTHPEV AYVGKTEEQV KSLGVDYRVG KFPFLANSRA
Pea     KVGHVDYDKV PGVVYTNPEV ASVGKTEEQV KETGVEYRVG KFPFMANSRA
Tryp    RPGHVNYDVI PGVIYTMPEV ASVGKTEEEL KKAGVAYKVG KFPFNANSRA
Human   GAVHIDYNCV PSVIYTHPEV AWVGKSEEQL KEEGIEYKVG KFPFAANSRA
Pig     GAVHIDYNCV PSVIYTHPEV AWVGKSEEQL KEEGIEYKVG KFPFAANSRA
Yeast   GHGHVNYNNI PSVMYSHPEV AWVGKTEEQL KEAGIDYKIG KFPFAANSRA
Pseud   KPHEVNYNLI PGVIYTRPEL ATVGKTEEQL KAEGRAYKVG KFPFTANSRA
Azo     HKAQMNYDLI PAVIYTHPEI AGVGKTEQAL KAEGVAINVG VFPFAASGRA
Aco-D   GAHPLDRDYV PGCTYARPQV ASLGLTESTA LARGRPIRIG KFSYQSNGKA
Staph   QAAEVDYIGM PAVCFTEPEL ATVGYSEAQA KEEGLAIKAS KFPYAANGRA
Bac     HPSAVDYVAI PAVVFSDPEC ASVGYFEQQA KDEGIDVIAA KFPFAANGRA
Coli    KKHYFDPKVI PSIAYTEPEV AWVGLTEKEA KEKGISYETA TFPWAASGRA
           * **   * + *  **  +** *  *+* ***   *  +    * ** *+** ****+

        451                                                  500
pB3-1   KAIDDAEGLV KVIAEKESDK ILGVHIMSPN AGELIHEAVL ALQYGASSED
Pea     KAIDNAEGLV KIIAEKETDK ILGVHIMAPN AGELIHEAAI ALQYDASSED
Tryp    KAVATEDGFV KVLTDKATDR ILGVHIVCSA AGELIAGALL AMEYGASSED
Human   KTNADTDGMV KILGQKSTDR VLGAHILGPG AGEMVNEAAL ALEYGASCED
Pig     KTNADTDGMV KILGQKSTDR VLGAHIIGPG AGEMINEAAL ALEYGASCED
Yeast   KTNQDTEGFV KILIDSKTER ILGAHIIGPN AGEMIAEAGL ALEYGASAED
Pseud   KINHETEGFA KVIADAETDE VLGVHLVGPS VSEMIGEFCV AMEFSASAED
Azo     MAANDTAGFV KVIADAKTDR VLGVHVIGPS AAELVQQGAI AMEFGTSAED
Aco-D   LVSGETEGFV KTIFDAETGE LLGAHMVGAQ VTEQIQGFGI ARHLEATDES
Staph   LSLDDTNGFV KLITLKEDDT LIGAQVVGTG ASDIISELGL AIEAGMNAED
Bac     LALNDTDGFL KLVVRKEDGV IIGAQIIGPN ASDMIAELGL AIEAGMTAED
Coli    IASDCADGMT KLIFDKESHR VICGAIVGTN GGELLGEIGL AIEMGCDAED
        *       + * +    *  *  *+ **   * ***   **  +* * ** +*

                      Interface-domain
        501                                 538
pB3-1   IARTCHAHPT MSEALKEAAM AT.YDKPIHM ........
Pea     IARVCHANPT MSEAIKEAAM AT.YDKPIHI ........
Tryp    VGRTCHAHPT MSEAVKEACM AC.FAKTINF ........
Human   IARVCHAHPT LSEAFREANL AASFGKSINF
Pig     IARVCHAHPT LSEAFREANL AASFGKAINF.
Yeast   VARVCHAHPT LSEAFKEANM AA.YDKAIHC ........
Pseud   IALTCHPHPT RSEALRQAAM NVD.GMAMQI ........
Azo     LGMMVFAHPA LSEALHEAAL AVS.GHAIHV ANRKK...
Aco-D   LLSMIFAHPT LSEAMHESIL AACDQPLHQ. ........
Staph   IALTIHAHPT LGEMTMEAA. EKAIGYPIHT M.......
Bac     IALTIHAHPT LGEIAMEAA. EVALGTPIHI ITK.....
Coli    IALTIHAHPT LHESVGLAA. EVFEGSITDL PNPKAKKK
        *** **+ +* *+*  ***  *    * *
```

Chapter 8

Discussion and Future Research

8.1 Dihydrolipoamide Dehydrogenase (E3) from Potato Tuber Mitochondria

8.1.1 Purification and M_r Analysis of Potato Mitochondrial E3

The novel isolation of multiple E3 isoforms from a plant source (*Solanum tuberosum* c.v. Maris piper potatoes) was first reported by this laboratory (Fullerton *et al*, 1996) where α_2, $\alpha\beta$ and β_2 activities (eluted from Mono Q ion exchange column as peaks 1, 2 and 3, respectively) were shown to arise from the various combinations of the 58 000 (α) and the 56 000 (β) subunits. The existence of the heterodimeric $\alpha\beta$ form was surprising since E3 has been reported to exist only as a homodimer in all organisms studied to date. Recently, similar isoforms were discovered in mitochondria from barley leaves, although several other plants such as horseradish (*Amoracia rusticana*), turnip (*Brassica campesteris*) and pea (*Pisum sativum*) have only a single form of mitochondrial E3 (M. Conner, Ph.D Thesis, Glasgow University, 1997). The expression of isoforms in potatoes was not a result of the tetraploid nature of the genes since a dihaploid strain of *Solanum tuberosum* and a diploid strain of *Solanum phoreta* also expressed three isoforms of E3.

Although the isoforms had been purified to homogeneity, their final yields were poor making further characterisation difficult. The original purification protocol was redesigned to scale up the procedure by removing an affinity chromatography step and replacing it with an ion exchange step (DEAE-52 chromatography) having greater capacity. The scale-up, was successful in obtaining approx. 20-fold more units of E3 activity (16 units by the original protocol to 320 units by the new protocol). However, since the extract was obtained from a crude pellet and not pure mitochondria the scale-up procedure increased the final yield of the three isoforms at the cost of slightly reduced purity.

Increased yield of the E3 isoforms permitted the determination of precise M_r values of the α and β polypeptides by electrospray mass spectrometry (Es-Ms). The M_r for the α polypeptide (49 446 ±12.28) and the β polypeptide (49 562 ±6.83) were found to be lower than the masses deduced by SDS-PAGE (58 000 and 56 000, respectively). This discrepancy is characteristic of mitochondrial E3s and the M_r values obtained by Es-Ms were in a similar range to other mitochondrial E3s reported to date. However, it was observed that the polypeptide having the larger mass by SDS-PAGE (α) is actually the smaller form by electrospray mass spectrometry.

8.1.2 Complex Specificity of the Mitochondrial E3 isoforms

The discovery of E3 isoforms in potato mitochondria had suggested that their physiological significance may be related to complex specificity. This would be an unusual situation since the single mitochondrial E3 from pea has been shown to be a common component for both PDC, OGDC and GDC (Bourguignon *et al.*, 1992; Turner *et al.*, 1992; Bourguignon *et al.*, 1996). In general, a single E3 is thought to complement the activities of the PDC, OGDC, BCDC and GDC in higher eukaryotes, although there are increasing examples of complex specific forms in prokaryotes.

In support of a complex specific role, tissue specific expression of the three potato E3 isoforms was reported by A. Carmichael (Ph.D Thesis, Glasgow University, 1994). The α_2 form was observed to be expressed at higher levels in potato leaf mitochondria (at a ratio of 9 α_2:2 $\alpha\beta$:2 β_2 and $\alpha\beta$ and β_2 were the major form in tubers (a ratio of 1 α_2:3 $\alpha\beta$:5 β_2). This result suggests that the α_2 isoform may be specific for GDC, which constitutes 30-50% of total leaf mitochondrial protein. Additionally, some preliminary kinetic characterisation carried out in the laboratory had shown that the three isoforms differ considerably in their enzymatic properties. For example, striking differences were found in the K_m for NAD^+ (120μM [α_2], 200μM [$\alpha\beta$] and 500μM [β_2]). They were also found to have varying heat stabilities, with the α form being more stable, and varying salt

sensitivities (R. Cook, unpublished observations, Glasgow University). It is not clear whether the enzymatic properties of the isoforms were dependent on the conditions of assay used i.e. different assay temperature and buffers, since previous results had suggested that the kinetic differences were not as pronounced (R. Fullerton, Ph.D Thesis, Glasgow University, 1996) as those found more recently.

Subsequently, complex specificity was investigated by reconstitution of potato OGDC and PDC activities (after being stripped of their E3 components) with the potato E3 isoforms. This approach proved to be more difficult than at first envisaged owing to the difficulties in purifying plant OGDC. Thus, there was no direct evidence for the preferential use of any of the isoforms by OGDC.

Although difficulties were encountered with detection and purification of the plant complexes, interesting differences in the selective binding of E3s from a variety of sources with the E1/E2 core of mammalian OGDC (bovine heart) were shown. Although the potato E3s were unable to promote any OGDC activity, weak binding was shown by competition experiments. Surprisingly, reconstitution by a variety of E3s highlighted subtle differences in catalytic and/or binding orientations between bovine heart E1/E2 and E3 from different species. Bovine mucosal and porcine heart were found to give poor reconstitution (15-20%) with bovine heart OGDC core and yeast E3 did not promote any complex activity. However, they were all found to have a high affinity of binding. E3 is a highly conserved enzyme with approx. 96% sequence identity between human liver and porcine heart E3; therefore, these E3s had been expected to promote the activity of OGDC to the same levels as native bovine heart E3. Similar experiments carried out by Sanderson *et al.* (1996b) had shown that these results were not artefactual since the same pattern of reconstitution was obtained with bovine heart PDC core complex.

Recent preliminary data on reconstitution of plant PDC has suggested some complex specificity of the β_2 isoform for PDC (R. Cook, unpublished observations, Glasgow University). This would seem to agree with the suggestion

that the α_2 form is putatively specific for GDC as it is expressed in high levels in leaf tissue. A recent development has been the discovery of a distinct chloroplastic E3 from pea leaves (Conner, *et al.* 1996). It is thought that this E3 is a component of the PDC found in chloroplasts which is distinct from mitochondrial PDC. The M_r of the plastidic E3 was found to be 52 614 by Es-Ms, approx. 3000 Da larger than the pea mitochondrial E3.

8.2 Cloning the Potato Mitochondrial E3 gene(s)

Since conclusions could not be made about the role of the three E3 by reconstitution studies, the relationship between the α and β polypeptides was investigated by the cloning their genes. This is only the second mitochondrial E3 to be cloned from a plant source.

8.2.1 Cloning by PCR Amplification of E3 cDNA

A PCR-based strategy was used for the amplification of cDNA generated from potato leaf RNA in the event of two separate transcripts encoding E3. Generating a PCR product for the purposes of sequencing requires the use of a proof-reading polymerase (*Pfu*) but the amplified product was unable to be cloned. The cloning of a *Taq* amplified product of the same size as the *Pfu* product led to the isolation of eight clones containing E3 cDNA inserts of varying size. The inserts were expected to encode E3 from amino acid 18 of the mature protein as a result of the cloning strategy. Expression of the cDNAs in *E. coli* showed that they also encoded proteins of different size. However, the proteins were inactive and insoluble either because they were truncated at the N-terminus causing FAD-binding sites to be lost or they had crucial mutations caused by the use of *Taq* DNA polymerase.

The sequence of pTaq22 consisted of an open reading frame of 1359 bp, encoding a polypeptide of 453 amino acids, which with the inclusion of the first 17 amino acids at the N-terminus would be 470 amino acids in total. The M_r was

found to correspond approx. to the β polypeptide (49 558). The amino acid sequence of pTaq23 was identical with pTaq22 apart from four amino acid substitutions. It was expected that the two clones represented the β polypeptide and expression of proteins of differing size was merely due to non-conservative amino acid substitutions. The sequence of pTaq16 was also identical with pTaq22 except for a deletion of 96 bp which was thought to be a cloning artefact.

8.2.2 Isolation of E3 clones from a Potato leaf cDNA Library

A 5′ restriction fragment of the pTaq22 cDNA was used as a probe for screening the cDNA library for full length E3 clones. A total of six clones were isolated and all were found to have the same restriction maps. Clones selected for sequencing, pB3-1 and pB9-3, had an open reading frame of 1512 bp, encoding a protein with 504 residues. The mature protein was 470 residues with a leader sequence of 34 residues. The theoretical M_r of the deduced protein was 49 561, being identical to the mass of the β polypeptide by Es-Ms. pB9-3 had a two base deletion, but this was expected to be a cloning artefact. Partial sequence of another clone, pB10-1 was obtained and was expected to be identical with pB3-1 and pB9-3. The sequence of these clones were identical to the sequence of the clones containing the *Taq* amplified insert, except at a few positions. The base differences in the *Taq* clones were attributed to mutations caused by *Taq* polymerase.

The sequence information and calculated M_r values of all the clones indicated that the β polypeptide had been cloned. If there was a second gene encoding the α polypeptide it was unclear why it had not been identified by cloning from purified mitochondrial RNA and from the library. Preliminary data presented by R. Cook suggested that the three isoforms had very distinct kinetic properties which would suggest that α and β are products of two different genes. The α polypeptide may posses a significant number of amino acid substitutions which coincidentally has a similar M_r to the β polypeptide. This is evident with pea mitochondrial E3 with the potato E3 which have 10% non-similar amino acids in comparison with each other and a M_r difference of only 168. However, other

criteria shows that the two polypeptides were very similar i.e. they both cross-reacted strongly to yeast antiserum, they were found to have identical N-terminal sequences and are very close in mass (a difference of only approx. 116 by Es-Ms). This would imply that alternate splicing of RNA or a deletion was responsible for a second, closely related polypeptide. The possibility of a post-translational modification for the generation of the α polypeptide was eliminated because it was shown that the gene encoding the polypeptide with the larger predicted M_r value had been cloned. Moreover, the deletion of the first or last amino acid of the mature protein would not generate an α polypeptide with the known M_r.

The existence of leaf specific mitochondrial isoforms can be discounted because the clone isolated from a leaf cDNA library had an identical M_r value to the β polypeptide of tuber mitochondria origin. The potato E3 sequence was found to have very high sequence similarities with other eukaryotic and prokaryotic E3s; 62% (*E. coli*), 73% (human) and 93% (pea). This reflects the highly conserved nature of the E3 enzyme. The highest identities were in the conserved functional domains of the polypeptide. An unrooted evolutionary tree of several E3s was constructed, based on amino acid sequence (Fig. 8.1). The open circle is the presumed origin of all the sequences and the numbers are 'PAM-distances', giving the % amino acid deviation from the last common root. The eukaryotic genes are clustered to the bottom left of the tree, with the plant E3s forming a separate branch. The tree shows that LPD-3 from *Pseudomonas putida* is found in between yeast and *Trypanosoma brucei* E3. Sequence and evolutionary analysis by Palmer *et al.* (1991) had shown that this E3 was more similar to eukaryotic E3s than the other E3s from *P. putida*, LPD-val and LPD-glc which were closely related to the prokaryotic E3s. The authors suggest that the *lpd*3 gene may be close to the ancestral gene for eukaryotic E3.

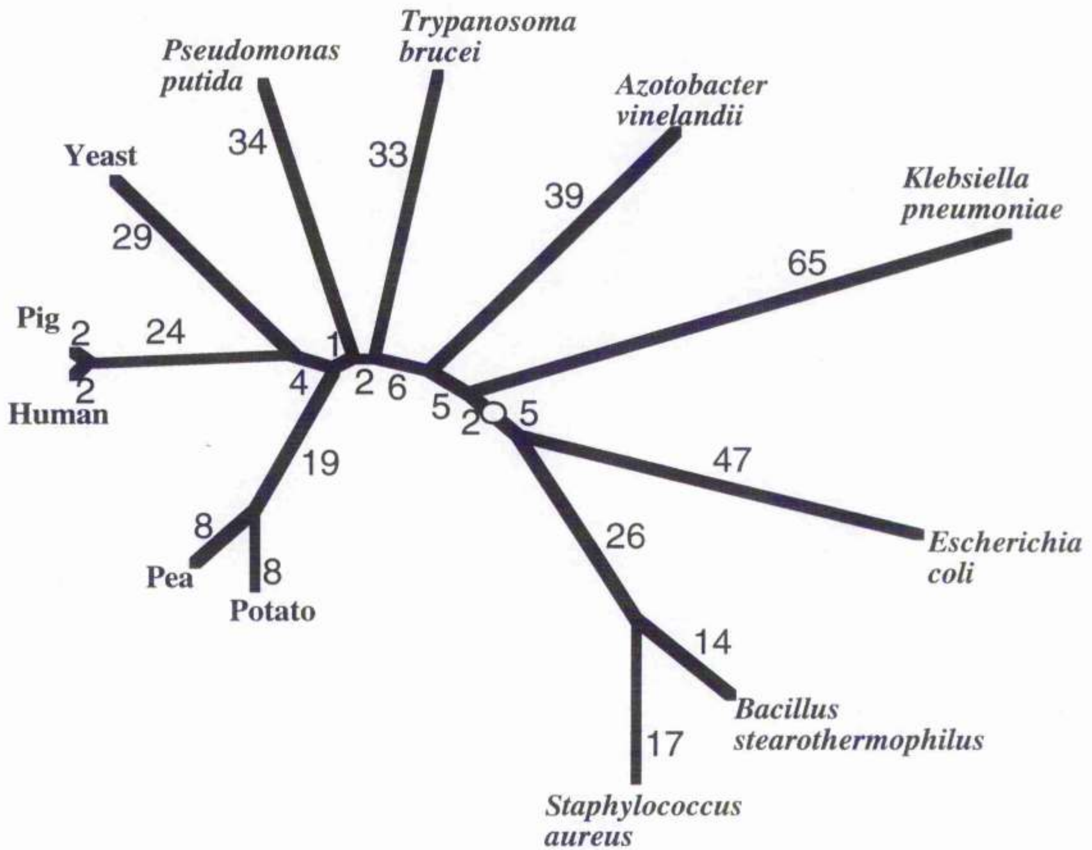


Fig. 8.1: Unrooted evolutionary tree of potato E3 in relation to other E3s, based on amino acid sequence.

The numbers indicate the relative evolutionary distances from the next common root (percent amino acid mutation) and were calculated using the Darwin system (AllAll: related peptide software) at the Computational Biochemistry Research Group server (ETHZ, Switzerland). The open circle indicates the theoretical origin of the tree.

8.3 Future Research

The exact relationship between the two polypeptides, α and β, forming the three isoforms of potato E3 has, as yet, not been established. Therefore, future work will concentrate on identifying the nature of a second possible gene encoding the α polypeptide. In this context, copy number analysis of the E3 gene in genomic potato DNA is a crucial first step. Extensive kinetic characterisation is required to clear the current confusion in this area. Additionally, comparisons of K_m, V_{max} and heat stabilities would help to determine if the α and β polypeptide are very closely related or not. This would then lead to the screening of a potato leaf or a potato tuber cDNA library if available.

Cloning of the newly discovered pea chloroplast E3 is underway and will provide an interesting comparison of structure-function relationships between the potato and pea mitochondrial E3s. For example, in characterising the selective binding of the chloroplast specific E3 to plastidic PDC. Cloning and sequencing of these plant genes is particularly important for the future production of transgenic plants. This would aid the assessment of the role played by plastidic PDC in fatty acid metabolism by its selective inhibition.

Northern blots would provide important information on the developmental regulation and tissue specific expression of the potato E3 isoforms. Of interest would be following the expression of the α_2 form with expression of GDC.

The over-expression of the clone, pB3-1, isolated from the potato leaf cDNA library is eagerly anticipated which would provide plentiful supply of the β_2 isoform. Initially this full length clone would be expressed in bacteria before attempting eukaryotic expression systems. This would be followed by crystallisation of the β_2 form for future structural determination at high resolution.

References

Allen, A. G., Perhan, R. N., Allison, N., Miles, J. S., & Guest, J. R. (1989) *J. Mol. Biol.*, **208**, 623-633.

Bates, D. L., Danson, M. J., Hale, G., Hooper, E. A., & Perham, R. N. (1977) *Nature*, **268**, 313-316.

Batteiger, B., Newhall, W.J. & Jones, R.B. (1982) *J. Immunol. Methods*, **55**, 297-307.

Behal, R. H., Browning, K. S., Hall, T. B., & Reed, L. J. (1989) *Proc.Natl. Acad. Sci. USA*, **86**, 8732-8736.

Benen, J. A. E., Van Berkel, W. J. H., Van Dongen, W. M. A. M., Muller, F., & De Kok, A. (1989) *J. Gen. Microbiol.*, **135**, 1787 - 1797.

Berg, P. A., & Klein, R. (1988) *Hepatology*, **8**, 200-201.

Bleile, D. M., Hackert, M. L., Pettit, F. H., & Reed, L. J. (1981) *J. Biol. Chem.*, **256**, 514-519.

Bleile, D. M., Munk, P., Oliver, R. M., & Reed, L. J. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 4385-4389.

Borges, A., Hawkins, C.F., Packman, L.C., & Perham, R.N. (1990) *Eur. J. Biochem.*, **194**, 95-102.

Bourguignon, J., Macherel, D., Neuberger, M., & Douce, R. (1992) *Eur. J. Biochem.*, **204**, 865-873.

Bourguignon, J., Merand, V., Rawsthorne, S., Forest, E., & Douce, R. (1996) *Biochem. J.*, **313**, 229-234.

Bourguignon, J., Neuberger, M., & Douce, R. (1988) *Biochem. J.*, **255**, 169-178.

Brookfield, D. E., Green, J., Ali, S. T., Machado, R. S., & Guest, J. R. (1991) *FEBS Lett.*, **295**, 13-16.

Brown, G. K., Brown, R.M., Scholem, D.M., Kirby, D.M. & Dahl, H-H.M. (1989) *Ann. N. Y. Acad. Sci.*, **573**, 360-368.

Brown, J.P., & Perham, R.N. (1976) *Biochem. J.*, **155**, 419-427.

Brown, T. A. (1992) in *The Practical Approach Series* (Rickwood, D., & Hames, B. D., Eds.), Oxford University Press, Oxford.

Browning, K.S., Uhlinger, D.J., & Reed, L.J. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 1831-1834.

Budde, R. J. A., Fang, T. K., & Randall, D. D. (1988a) *Plant Physiol.*, **88**, 1031-1036.

Budde, R. J. A., & Randall, D. D. (1987) *Arch. Biochem. Biophys.*, **258**, 600-606.

Budde, R. J. A., & Randall, D. D. (1988b) *Plant Physiol.*, **88**, 1026-1030.

Budde, R. J. A., & Randall, D. D. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 673-676.

Bullock, W. O., Fernandez, J. M., & Short, J. M. (1987) *Biotechniques*, **5**, 376-379.

Burns, G., Brown, T., Hatter, K., & Sokatch, J.R. (1989b) *Eur. J. Biochem.*, **179**, 61-69.

Burns, G., Sykes, P. J., Hatter, K., & Skokatch, J. R. (1989a) *J. Bacteriol.*, **171**, 665-668.

Camp, P. J., Miernyk, J. A., & Randall, D. D. (1988) *Biochim. Biophys. Acta*, **933**, 269-275.

Camp, P. J., & Randall, D. D. (1985) *Plant Physiol.*, **77**, 571-577.

Carmicheal, A. (1994) Ph.D. Thesis, Glasgow University.

Carothers, D. J., Pons, G., & Patel, M. S. (1989) *Arch. Biochem. Biophys.*, **268**, 409-425.

Carothers, D. J., Raefsky-Estrin, C., Pons, G., & Patel, M. S. (1987) *Arch. Biochem. Biophys.*, **256**, 597-605.

Chang, M., Naik, S., Johanning, G. L., Ho, L., & Patel, M. S. (1993) *Biochemistry*, **32**, 4263-4269.

Conner, M., Krell, T., & Lindsay, J. G. (1996) *Planta*, **200**, 195-202.

Cook, K. G., Bradford, A. P., & Yeaman, S. J. (1985) *Biochem. J.*, **225**, 731-735.

Cook, K. G., Lawson, R., Yeaman, S. J., & Aitken, A. (1983) *FEBS Lett.*, **164**, 47-50.

Coppel, R. L., McNeilage, L. J., Surh, C. D., Van De Water, J., Spithill, T. W., Whittingham, S., & Gershwin, M. E. (1988) *Proc. Acad. Natl. Sci. USA*, **85**, 7317-7321.

Craig, D.W. & Wedding, R.T. (1980) J. *Biol. Chem.*, **255**, 5763-5768.

Danson, M. J., Conroy, K., McQuattie, A., & Stevenson, K. J. (1987) *Biochem. J.*, **243**, 661-665.

Danson, M. J., Eisenthal, R., Hall, S., Kessell, S. R., & Williams, D. L. (1984) *Biochem. J.*, **218**, 811-818.

Danson, M. J., Fersht, A. R., & Perham, R. N. (1978) *Proc. Natl. Acad. Sci. USA*, **75**, 5386-5390.

Danson, M.J., Hale, G., Johnson, P., Perham, R.N., Smith, J., & Spragg, P. (1979) *J. Mol. Biol.*, **129**, 603-617.

Dardel, F., Davis, A. L., Laue, E. D., & Perham, R. N. (1993) *J. Mol. Biol.*, **229**, 1037-1048.

Dardel, F., Packman, L. C., & Perham, R. N. (1990) *FEBS Letters.*, **264**, 206-210.

Darlinson, M. G., Spencer, M. E., & Guest, J. R. (1984) *Eur. J. Biochem.*, **141**, 351-359.

Davis, P. F., Pettit, F. H., & Reed, L. J. (1977) *Biochem. Biophys. Res. Commun.*, **75**, 541-549.

Day, D.A., Moore, A.L., Dry, I.B., Wiskich, J.T. & Ascon-Bieto, J. (1988) *Plant Physiol.*, **86**, 1199-1204

Day, D.A., Neuberger, M., Douce, R. (1985) *Aust. J. Plant Physiol.*, **12**, 219-228.

De Marcucci, O. G., & Lindsay, J. G. (1985) *Eur. J. Biochem.*, **149**, 641-648.

Dennis, D. T., & Mierynk, J. A. (1982) *Annu. Rev. Plant Physiol.*, **33**, 27-50.

Denton, R. M., & McCormack, J. G. (1985) *Am. J. Physiol.*, **249**, E543-E554.

Diaz, F., & Komuniecki, R. (1995) *Mol. Biochem. Parasitol.*, **75**, 87-97.

Dickinson, J. R., Roy, D. J., & Dawes, I. W. (1986) *Mol. Gen. Genet.*, **204**, 103-107.

Douce, R., Bourguignon, J., Macherel, D., & Neuberger, M. (1994) *Biochem. Soc. Trans.*, **22**, 184-186.

Dry, I. B., & Wiskich, J. T. (1985) *Biochem. Biophys. Res. Commun.*, **133**, 397-403.

Else, A.J., Hough, D.W., & Danson, M.J. (1993) *Eur. J. Biochem.*, **212**, 423-429.

Ferrige, A. G., Seddon, M. J., Green, B. N., Jarvis, S. A., & Skilling, J. (1992) *Mass Spectrom.*, **6**, 707-711.

Fregeau, D. R., Roche, T. E., Davis, P. A., Coppel, R. L., & Gershwin, M. E. (1989) *J. Immunol.*, **144**, 1671-1676.

Freudenberg, W., Dietrichs, D., Lebertz, H., Andreesen, J.R. (1989), *J. Bacteriol.*, **171**, 3, 1346-1354.

Fuller, C. C., Reed, L. J., Oliver, R. M., & Hackert, M. L. (1979) *Biochem. Biophys. Res. Commun.*, **90**, 431-437.

Fullerton, R. (1995) Ph.D Thesis, Glasgow University.

Fullerton, R., Conner, M., Cook, R.M., Khan, S.S., Krell, T. & Lindsay, J.G. (1996) In preparation.

Fussey, S. P. M., Guest, J. R., James, O. F. W., Bassendine, M. F., & Yeaman, S. J. (1988) *Proc. Natl. Acad. Sci. USA.* **85**, 8654-8658.

Gardestrom, P., Bergman, A., & Ericson, I. (1980) *Plant Physiol.*, **65**, 389-391.

Gemel, J., & Randall, D. D. (1992) *Plant Physiol.*, **100**, 908-914.

Gerbling, H., & Gerhardt, B. (1987) *Planta*, **171**, 386-392.

Gerbling, H., & Gerhardt, B. (1988) *Plant Physiol.*, **88**, 13-15.

Gerbling, H., & Gerhardt, B. (1989) *Plant Physiol.*, **91**, 1387-1392.

Gershwin, M. E., & Mackay, I. R. (1991) *Gastroenterology*, **100**, 822-825.

Gopalkrishnan, S., Rahmatulla, M., Radke, G. A., Powers-Greenwood, S., & Roche, T. E. (1989) *Biochem. Biophys. Res. Commun.*, **160**, 715-721.

Graham, L. D., Packman, L. C., & Perham, R. N. (1989) *Biochemistry*, **28**, 1574-1581.

Green, J. D. F., Laue, E. D., Perham, R. N., Ali, S. T., & Guest, J.R. (1995) *J. Mol. Biol.*, **248**, 328-343.

Griffin, T.A., Lau, K.S., & Chuang, D.T. (1988) *J. Biol. Chem.*, **263**, 14 008-14 014.

Guan, Y., Rawsthorne, S., Scofield, G., Shaw, P., & Doonan, J. (1995) *J. Biol. Chem.*, **270**, 5412-5417.

Guest, J. R. (1987) *FEMS Microbiol. Lett.*, **44**, 417-422.

Guest, J.R., Angier, S.J., Russell, G.C. (1989) *Ann. N. Y. Acad. Sci.*, 573, 76-99.

Guest, J. R., Lewis, H. M., Graham, L. D., Packman, L. C., & Perham, R. N. (1985) *J. Mol. Biol.*, **185**, 743-754.

Guest, J. R., Quail, M. A., Dave, E., Cassey, B., & Attwood, M. M. (1996) in *Biochemistry and Physiology of Tiamine Diphosphate Enzymes*. (Bisswanger, H., & Schellenberger, A., Eds.) pp 326-333, A.u.C. Intemann, Prien.

Hanahan, D. (1983) *J. Mol. Biol.*, **166**, 557-580.

Hanemaaijer, R., Janssen, A., de Kok, A., & Veeger, C. (1988) *Eur. J. Biochem.*, **174**, 593-599.

Hawkins, C. F., Borges, A., & Perham, R. N. (1989) *FEBS Lett.*, **255**, 77-82.

Hay, B., & Short, J. M. (1992) *Strategies*, **5**, 16.

Hein, S., & Steinbuchel, A. (1994) *J. Bacteriol.*, **176**, 14, 4394 - 4408.

Hein, S., & Steinbuchel, A. (1996) *FEMS Microbiol. Lett.* **136**, 231-238.

Hemilä, H. (1991) *Biochim. Biophys. Acta.*, **1129**, 119-123.

Henderson, C. E., & Perham, R. N. (1980) *Biochem. J.*, **189**, 161-172.

Henderson, C. E., Perham, R. N., & Finch, J. T. (1979) *Cell*, **17**, 85-93.

Henikoff, S. (1984) *Gene*, **28**, 351-359.

Hind, G., Nakatani, H. Y., & Izawa, S. (1974) *Proc. Natl. Acad. Sci. USA*, **71**, 1484-1488.

Hipps, D. S., Packman, L. C., Allen, M. D., Fuller, C., Sakaguchi, K., Appella, E., & Perham, R. N. (1994) *Biochem. J.*, **297**, 137 - 143.

Hiraga, K., & Kikuchi, G. (1980) *J. Biol. Chem.*, **255**, 11671-11676.

Ho, L., Wexler, I. D., Kerr, D. S., & Patel, M. S. (1989) *Ann. N. Y. Acad. Sci.*, **573**, 347-359.

Hockney, R. C. (1994) *Trends Biochem. Sci.*, **12**, 456-463.

Hodgson, J. A., De Marcucci, O. G. L., & Lindsay, J. G. (1988) *Eur. J. Biochem.*, **171**, 609-614.

Hodgson, J. A., De Marcucci, O. G.L., & Lindsay, J. G. (1986) *Eur. J. Biochem.*, **158**, 595-600.

Hunter, A., & Lindsay, J. G. (1986) *Eur. J. Biochem.*, **155**, 103-109.

Husic, D. W., Husic, D. W., & Tolbert, N. E. (1987) *Crit. Rev. Plant Sci.*, 5, 45-100.

Jackman, S.A., Hough, D.W., Danson, M.J., Stevenson, K.J. & Opperdoes, F.R. (1990) *Eur. J. Biochem.*, **193**, 91-95.

James, A. G., Cook, R. M., West, S. M., & Lindsay, J. G. (1995) *FEBS Lett.*, **373**, 111-114.

Jeng, J., Huh, T. L., & Song, B. J. (1994) *Biochem. Biophys. Res. Commun.*, **203**, 225-230.

Jilka, J. M., Rahmatullah, M., Kazemi, M., & Roche, T. E. (1986) *J. Biol. Chem.*, **261**, 1858-1867.

Jones, S. M. A., & Yeaman, S. J. (1986) *Biochem. J.*, **237**, 621-623.

Kalia, Y. N., Brocklehurst, S. M., Hipps, D. S., Appella, E., Sakaguchi, K., & Perham, R. N. (1993) *J. Mol. Biol.*, **230**, 323-341.

Kang, F., & Rawsthorne, S. (1994) *Plant J.*, **6**, 795-805.

Kaplan, M. M. (1987) *N. Eng. J. Med.*, **316**, 521-528.

Keha, E.E., Ronft, H., & Kresze, G.-B. (1982) *FEBS Lett.*, **145**, 289-292.

Kim, Y., & Oliver, D. J. (1990) *J. Biol. Chem.*, **265**, 848-853.

Klein, S. M., & Sagers, R. D. (1966) *J. Biol. Chem.*, **241**, 197-205.

Kochi, H., Seino, H., & Ono, K. (1986) *Arch. Biochem. Biophys.*, **249**, 263-272.

Korotchkina, L.G., Ali, M.S. & Patel, M.S. (1996) in *Alpha-Keto Acid Dehydrogenase Complexes* (Patel, M. S., Roche, T. E., & Harris, R. A., Eds.) pp 17-31, Birkhauser Verlag, Basel.

Korotchkina, L. G., Tucker, M. M., Thekkumkara, T. J., Madhusudhan, K. T., Pons, G., Kim, H., & Patel, M. S. (1995) *Prot. Express. Purif.*, **6**, 79-90.

Krauth-Siegel, R. L., Blatterspiel, R., Saleh, M., Schlitz, E., Schirmer, R. H., & Untucht-Grau, R. (1982) *Eur. J. Biochem.*, **121**, 259-267.

Krömer, S., Lernmark, U., & Gardestrom, P. (1994) *J. Plant Physiol.*, **144**, 485-490.

Kuriyan, J., Kong, X. P., Krishna, T. S. R., Sweet, R. M., Murgolo, N. H., Field, H., Cerami, A., & Henderson, G. B. (1991b) *Proc. Natl. Acad. Sci. USA*, **88**, 8764-8768.

Kuriyan, J., Krishna, T. S. R., Wong, L., Guenther, B., Pahler, A., Williams, C. H., & Model, P. (1991a) *Nature*, **352**, 172-174.

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

Lau, K. S., Herring, W. J., Chuang, J. L., McKean, M., Danner, D. J., Cox, R. P., & Chuang, D. T. (1992) *J. Biol. Chem.*, **267**, 24090-24096.

Lawson, J. E., Behal, R. H., & Reed, L. J. (1991a) *Biochemistry*, **30**, 2834-2839.

Lawson, J. E., Niu, X. D., & Reed, L. J. (1991b) *Biochemistry*, **30**, 11 249-11 254.

Lernmark, U., & Gardestrom, P. (1994) *Plant Physiol.*, **106**, 1633-1638.

Lessard, I. A. D., & Perham, R. N. (1995) *Biochem. J.*, **306**, 727-733.

Linn, T. C., Pettit, F. H., Hucho, F., & Reed, L. J. (1969) *Proc. Natl. Acad. Sci. USA*, **64**, 234-241.

Liu, S., Baker, J. C., & Roche, T. E. (1995) *J. Biol. Chem.*, **270**, 793-800.

Lowe, P.N., Hodgson, J.A., & Perham, R.N. (1983) *Biochem. J.*, **215**, 133-140.

Lowry, O.H., Rosenbrough, N.J., Farr, A.L. & Randall, R.J. (1951) *J. Biol. Chem.*, **193**, 265-275.

Luethy, M. H., David, N. R., Elthon, T. E., Miernyk, J. A., & Randall, D. D. (1995) *J. Plant. Physiol.*, **145**, 443-449.

Luethy, M. H., Miernyk, J. A., David, N. R., & Randall, D. D. (1996) in *Alpha-Keto Acid Dehydrogenase Complexes* (Patel, M. S., Roche, T. E., & Harris, R. A., Eds.) pp 71-92, Birkhauser Verlag, Basel.

Luethy, M. H., Miernyk, J. A., & Randall, D. D. (1994) *Biochim. Biophys. Acta.*, **1187**, 95-98

Lütcke, H. A., Chow, K. C., Mickel, F. S., Moss, K. A., Kern, H. F., & Scheele, G. A. (1987) *EMBO J.*, **6**, 43-48.

Machado, R. S., Clark, D. P., & Guest, J. R. (1992) *FEMS Microbiol. Lett.*, **100**, 241-246.

Maeng, C.Y., Yazdi, M.A., Niu, X.-D., Lee, H.Y. & Reed, L.J. (1994) *Biochemistry*, **33**, 13 801-13 807.

Mande, S. S., Sarfaty, S., Allen, M. D., Perham, R. N., & Hol, W. G. J. (1996) *Structure*, **4**, 277-286.

Markwell, M.A., Haass, S., Bieber, L.L. & Tolbert, N.E. (1976) *Anal. Biochem.*, **87**, 206-210.

Marsac, C., Stanbie, D., Bonne, G., Cousin, J., Jehenson, P., Benelli, C., Leroux, J. P., & Lindsay, J. G. (1993) *J. Pediatr.*, **123**, 915-920.

Marshall, S.E., McCormack, J.G., Denton, R.M. (1984) *Biochem. J.*, **218**, 249-260.

Mattevi, A., Obmolova, G., Sokatch, J. R., Betzel, C., & Hol, W. G. J. (1992a) *Proteins: structure, function and genetics.*, **13**, 336-351.

Mattevi, A., Obmolova, G., Kalk, K. H., van Berkel, W. J. H., & Hol, W. G. J. (1993) *J. Mol. Biol.*, **230**, 1200-1215.

Mattevi, A., Obmolova, G., Schulze, E., Kalk, K. H., Westphal, A. H., de Kok, A., & Hol, W. G. J. (1992b) *Science*, **255**, 1544-1550.

Mattevi, A., Schierbeek, A. J., & Hol, W. G. J. (1991) *J. Mol. Biol.*, **220**, 975 - 994.

Matuda, S., & Saheki, T. (1985) *Biochem. Biophys. Res. Commun.*, **129**, 479-484.

McCully, V., Burns, G., & Sokatch, J. R. (1986) *Biochem. J.*, **233**, 737-742.

McCormack, J. G., & Denton, R. M. (1979) *Biochem. J.*, **180**, 533-544.

Miernyk, J. A., Camp, P., & Randall, D. D. (1985) *Curr. Top. Plant Biochem. Physiol.*, **4**, 175-189.

Miernyk, J. A., & Randall, D. D. (1987) *Plant Physiol.*, **83**, 311-315.

Moffat, B. A., & Studier, F. W. (1987) *Cell*, **49**, 221-227.

Nakano, K., Matuda, S., Yamanaka, T., Tsubouchi, H., Nakagawa, S., Titani, K., Ohta, S., & Miyata, T. (1991) *J. Biol. Chem.*, **266**, 19 013-19 017.

Nakano, K., Takase, C., Sakamoto, T., Nakagawa, S., Inazawa, J., Ohta, S., & Matuda, S. (1994) *Eur. J. Biochem.*, **224**, 179-189.

Neagle, J. C., & Lindsay, J. G. (1991) *Biochem. J.*, **278**, 423-427.

O'Brien, W. E. (1978) *Arch. Biochem. Biophys.*, **189**, 291-297.

Ohlrogge, J. B., Kuhn, D. N., & Stumpf, P. K. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 1194-1198.

Oliver, D. J., & Walker, G. H. (1984) *Plant Physiol.*, **76**, 409-413.

Oliver, R. M., & Reed, L. J. (1982) in *Electron Microscopy of Proteins* (Harris, J. R., Ed.) pp 1-48, Academic Press, London.

Oliver, D. J. (1994) *Annu. Rev.Plant Physiol. Plant Mol. Biol.*, **45**, 323-327.

Oliver, D. J., Neuberger, M., Bourguignon, J., & Douce, R. (1990) *Plant Physiol.*, **94**, 833-839.

Ono, K., Radke, G. A., Roche, T. E., & Rahmatullah, M. (1993) *J. Biol. Chem.*, **268**, 26 135-26 143.

Oppermann, F. B., & Steinbuchel, A. (1994) *J. Bacteriol.*, **176**, 469-485.

Otulakowski, G., & Robinson, B. H. (1987) *J. Biol. Chem.*, **262**, 17 313-17 318.

Packman, L. C., Borges, A., & Perham, R. N. (1988) *Biochem. J.*, **252**, 79-86.

Packman, L. C., & Perham, R. N. (1986) *FEBS Lett.*, **206**, 193-198.

Palmer, J.A., Madhusudhan, K.T., Hatter, K., Sokatch, J.R. (1991) *Eur. J. Biochem.*, **202**, 231-240.

Patel, M. S., & Harris, R. A. (1995) *FASEB J.*, **9**, 1164-1172.

Paxton, R., & Harris, R. A. (1984) *Arch. Biochem. Biophys.*, **231**, 48-57.

Peng, H.-L., Deng, W.-L., Yang, Y.-H., & Chang, H.-Y. (1996) *J. Biochem.*, **119**, 1118-1123.

Perham, R. N. (1991) *Biochemistry*, **30**, 8501-8512.

Perham, R. N. (1996) in *Alpha-Keto Acid Dehydrogenase Complexes* (Patel, M. S., Roche, T. E., & Harris, R. A., Eds.) pp 1-15, Birkhauser Verlag, Basel.

Perham, R. N., Duckworth, H. W., & Roberts, G. C. K. (1981) *Nature*, **292**, 474-477.

Pettit, F. H., Pelley, J. W., & Reed, L. J. (1975) *Biochem. Biophys. Res. Commun.*, **65**, 575-582.

Pettit, F. H., Roche, T. E., & Reed, L. J. (1972) *Biochem. Biophys. Res. Commun.*, **49**, 563-571.

Pons, G., Reafsky-Estrin, C., Carothers, D.J., Pepin, R.A., Javed, A.A., Jesse, B.W., Ganapathi, M.K., Samols, D. & Patel, M.S. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 1422-1426.

Popov, K. M., Kedishvili, N. Y., Zhao, Y., Gudi, R., & Harris, R. A. (1994) *J. Biol. Chem.*, **269**, 29 720-29 724.

Popov, K. M., Kedishvili, N. Y., Zhao, Y., Shimomura, Y., Crabb, D. W., & Harris, R. A. (1993) *J. Biol. Chem.*, **268**, 26 602-26 606.

Popov, K. M., Zhao, Y., Shimomura, Y., Kuntz, M. J., & Harris, R. A. (1992) *J. Biol. Chem.*, **267**, 13 127-13 130.

Poulsen, L. L., & Wedding, R. T. (1970) *J. Biol. Chem.*, **245**, 5709-5717.

Quail, M. A., Haydon, D. J., & Guest, J. R. (1994) *Mol. Microbiol.*, **12**, 95-104.

Rahmatullah, M., Gopalkrishnan, S., Andrews, P. C., Chang, C. L., Radke, G. A., & Roche, T. E. (1989) *J. Biol. Chem.*, **264**, 2221-2227.

Randall, P. J. (1986) *Biochem. Soc. Trans.*, **14**, 799-806.

Randall, D. D., & Miernyk, J. A. (1990a) in *Meth. Plant Biochem.*, 3, 175-192, Academic Press.

Randall, D. D., Miernyk, J. A., David, N. R., Budde, R. J. A., Schuller, K. A., Fang, T. K., & Gemel, J. (1990b) *Curr. Top. Plant Biochem. Physiol.*, **9**, 313-328.

Randall, D. D., Miernyk, J. A., Fang, T. K., Budde, R. J. A., & Schuller, K. A. (1989) *Ann. N. Y. Acad. Sci.*, **573**, 192-205.

Randall, D. D., & Rubin, P. M. (1977) *Plant Physiol.*, **59**, 1-4.

Randall, D. D., Williams, M., & Rapp, B. J. (1981) *Arch. Biochem. Biophys.*, **207**, 437-444.

Rao, K. P., & Randall, D. D. (1980) *Arch. Biochem. Biophys.*, **200**, 461-466.

Rapp, B. J., & Randall, D. D. (1980) *Plant Physiol.*, **65**, 314-318.

Reed, L. J. (1974) *Accounts. Chem. Res.*, **7**, 40-46.

Reed, L.J., Browning, K.S., Niu, X-D., Behal, R.H., & Uhlinger, D.J. (1989) *Ann. N. Y. Acad. Sci.*, **573**, 155-167.

Reed, L. J., & Hackert, M. L. (1990) *J. Biol. Chem.*, **265**, 8971-8974.

Reed, L.J., Koike, M., Levitch, M.E. & Leach, F.R. (1958) *J. Biol. Chem.*, **232**, 143-148.

Reed, L. J., & Yeaman, S. J. (1987) in *The Enzymes* (Boyer, P. D., & Krebs, E. G., Eds.) **18**, pp 77-95, Academic Press, New York.

Reid, E. E., Thompson, C. R., Lyttle, D. T., & Dennis, D. T. (1977) *Plant Physiol.*, **59**, 842-848.

Rice, D. W., Schulz, G. E., & Guest, J. R. (1984) *J. Mol. Biol.*, **174**, 483-496.

Rice, J.E., Dunbar, B. & Lindsay, J.L. (1992) *EMBO. J.*, **11**, 3229-3235.

Richarme, G. (1989) *J. Bacteriol.*, **171**, 12, 6580 - 6585.

Robien, M. A., Clore, G. M., Omichinski, J. G., Perham, R. N., Appella, E., Sakaguchi, K., & Gronenborn, A. M. (1992) *Biochemistry*, **31**, 3463-3471.

Robinson, B. H. & Chun, K. (1993) *FEBS Lett.*, **328**, 99-102.

Robinson, B. H., MacKay, N., Petrova-Benedict, R., Ozalp, I., Coskum, T., & Stacpoole, P. W. (1990) *J. Clin. Invest.*, **85**, 1821-1824.

Robinson, B. H., Taylor, J., & Sherwood, W. G. (1977) *Peadiatric Res*, **11**, 1198-1202.

Roche, T.E., Liu, S., Ravindran, S., Baker, J.C. & Wang, L. (1996) in *Alpha-Keto Acid Dehydrogenase Complexes* (Patel, M. S., Roche, T. E., & Harris, R. A., Eds.) pp 33-51, Birkhauser Verlag, Basel.

Roche, T. E., Rahmatullah, M., Li, L., Radke, G. A., Chang, C. L., & Powers-Greenwood, S. L. (1989) *Ann. N. Y. Acad. Sci.*, **573**, 168-174.

Ross, J., Reid, G. A., & Dawes, I. W. (1988) *J. Gen. Microbiol.*, **134**, 1131-1139.

Roy, D. J., & Dawes, I. W. (1987) *J. Gen. Microbiol.*, **133**, 925-933.

Rubin, P. M., Zahler, W. L., & Randall, D. D. (1978) *Arch. Biochem. Biophys.*, **188**, 70-77.

Rubin, P. M., & Randall, D. D. (1977) *Arch. Biochem. Biophys.*, **178**, 342-349.

Russell, G. C., & Guest, J. R. (1990) *Biochem. J.* **269**, 443-450.

Sakurai, Y., Fukuyoshi, Y., Hamada, M., Hayakawa, T., & Koike, M. (1970) J. *Biol. Chem.*, **245**, 4453-4462.

Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989) *Molecular cloning: A laboratory manual*, 2nd ed., Cold Spring Harbour Laboratory Press, Cold Spring Harbour, NY.

Sanderson, S. J., Miller, C., & Lindsay, J. G. (1996a) *Eur. J. Biochem.*, **236**, 68-77.

Sanderson, S. J., Khan, S.S., McCartney, R.G., Miller, C. & Lindsay, J.G. (1996b) *Biochem. J.*, **319**, 109-116.

Sarojini, G., & Oliver, D. J. (1983) *Plant Physiol.*, **72**, 194-199.

Schiering, N., Kabasch, W., Moore, M. J., Distefano, M. D., Walsh, C. T., & Pai, E. F. (1991) *Nature*, **352**, 168-172.

Schuller, K. A., Gemel, J., & Randall, D. D. (1993) *Plant Physiol.*, **102**, 139-143.

Schuller, K. A., & Randall, D. D. (1989) *Plant Physiol.*, **89**, 1207-1212.

Schulz, G. E., Schirmer, R. H., Sachsenheimer, W., & Pai, E. F. (1978) *Nature* , **273**, 120-124.

Schulze, E., Westphal, A. H., Obmolova, G., Mattevi, A., & Hol, W. G. J. (1991) *Eur. J. Biochem.*, **255**, 561-568.

Short , J.M., Fernandez, J.M., Sorge, J.A., & Huse, W.D. (1988) *Nuc. Acids. Res.*, **16**, 7583-7600.

Smith, R. G., Gauthier, D. A., Dennis, D. T., & Turpin, D. H. (1992) *Plant Physiol.*, **98**, 1233-1238.

Sokatch, J. R., & Burns, G. (1984) *Arch. Biochem. Biophys.*, **228**, 660-666.

Sokatch, J. R., McCully, V., Gebrosky, J., & Sokatch, D. J. (1981) *J. Bacteriol.*, **148**, 639-646.

Sokatch, J. R., McCully, V., Gebrosky-Sahm, J., & Maguire, M. (1983) *J. Bacteriol.*, **153**, 969-975.

Spencer, M.E., Darlison, M.G., Lewis, H.M. & Guest, J.R. (1984) *Eur. J. Biochem.*, **141**, 361-374.

Spencer, M. E., & Guest, J. R. (1985) *Mol. Gen. Gen.*, **200**, 145-154.

Stanley, C.J., & Perham, R.N. (1980) *Biochem. J.*, 191, 147-154.

Springer, J., & Heise, K.-P. (1989) *Planta*, **177**, 417-421.

Stansbie, D., Wallace, S. J., & Mori, M. (1986) *J. Inherited Metab.*, **9**, 105-119.

Stephens, P. E., Lewis, H. M., Darlison, M. G., & Guest, J. R. (1983) *Eur. J. Biochem.*, **135**, 519-527.

Stepp, L.R., Pettit, F.H., Yeaman, S.J. & Reed, L.J. (1983) *J. Biol. Chem.*, **258**, 9454-9458.

Studier, F. W., & Moffatt, B. A. (1986) *J. Mol. Biol.*, **189**, 113-130.

Sugden, M.C., & Holness, M.J. (1994) *FASEB J.*, **8**, 54-61.

Sugden, P. H., Hutson, N. J., Kerby, A. L., & Randle, P. J. (1978) *Biochem. J.*, **169**, 433-435.

Surh, C. D., Ansari, A., & Gershwin, M. E. (1990) *J. Immunol.*, **144**, 3367-3374.

Sykes, P. J., Burns, G., Menard, J., Hatter, K., & Sokatch, J. R. (1987) *J. Bacteriol.*, **169**, 1619-1625.

Taylor, A. E., Cogdell, R. J., & Lindsay, J. G. (1992) *Planta.*, **188**, 225-231.

Teague, W. M., Pettit, F. H., Wu, T.-L., Silberman, S. R., & Reed, L. J. (1982) *Biochemistry*, **21**, 5585-5592.

Thekkumkara, T.J., Pons, G., Mitroo, S., Jentoft, J.E., & Patel, M.S. (1989) *Ann. N. Y. Acad. Sci.*, **573**, 113-129.

Thieme, R., Pai, E. F., Schirmer, R. H., & Schulz, G. E. (1981) *J. Mol. Biol.*, **152**, 763-782.

Towbin, H., Staehelin, T. & Gordon, J. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 4350-4354.

Turner, S. R., Ireland, R., & Rawsthorne, S. (1992) *J. Biol. Chem.*, **267**, 7745-7750.

Walker, J. L., & Oliver, D. J. (1986a) *J. Biol. Chem.*, **261**, 2214-2221.

Walker, J. L., & Oliver, D. J. (1986b) *Arch. Biochem. Biophys.*, **248**, 626-638.

Wallis, N. G., & Perham, R. N. (1994) *J. Mol. Biol.*, **236**, 209 - 216.

Wedding, R. T., & Black, M. K. (1971) *J. Biol. Chem.*, **246**, 1638-1643.

Westphal, A. H., & Kok, A. (1988) *Eur. J. Biochem.*, **172**, 299-305.

Westphal, A. H., & Kok, A. (1990) *Eur. J. Biochem.*, **187**, 235-239.

Williams, C. H. J. (1976) in *The Enzymes* , 3rd ed. (Boyer, P. D., Ed.) pp 89-173, Academic Press Inc., New York.

Williams, M., & Randall, D. D. (1979) *Plant Physiol.*, **64**, 1099-1103.

Wray, W., Boulikas, T., Wray, V.P. & Hancock, R. (1981) *Anal. Biochem.*, **118**, 197-203.

Wu, T. L., & Reed, L. J. (1984) *Biochemistry*, **23**, 221-226.

Yeaman, S. J., Bassendine, M. F., Fittes, D., Hodgson, D. R., Heseltine, L., Brown, H. J., Mutimer, D. J., James, O. F. W., & Fussey, S. P. M. (1989) *Ann. N. Y. Acad. Sci.*, **573**, 183-191.

Yeaman, S. J., Fussey, S. P. M., Danner, D. J., James, O. F. W., Mutimer, D. J., & Bassendine, M. F. (1988) *Lancet*, **i**, 1067-1070.

Yeaman, S. J. (1989) *Biochem. J.*, **257**, 625-632.

Yeaman, S. J. (1986) *Trends Biochem. Sci.*, **11**, 293-296.

Yeaman, S. J., Hutcheson, E. T., Roche, T. E., Pettit, F. H., Brown, J. R., Reed, L. J., Watson, D. C., & Dixon, G. H. (1978) *Biochemistry*, **17**, 2364-2370.

Zeiher, C. A., & Randall, D. D. (1990) *Plant Phisiol.*, **94**, 20-27.

Zhao, Y., Hawes, J., Popov, K. M., Jaskiewicz, J., Shimomura, Y., Crabb, D. W., & Harris, R. A. (1994) *J. Biol. Chem.*, **269**, 18 583-18 587.

Appendix

Fig. A: DNA sequence comparison of the clones pB3-1, pB9-3 and the PCR derived clones pTaq22, pTaq23 and pTaq16.

The sequence information of pTaq23 starts at base number 482. A changed base leading to an amino acid change is indicated by '*' and a changed base causing no amino acid change is indicated by '+'. The forward and reverse sequencing primers found in Chapter 5, Table 5.3 are shown by half arrows. The position of primer, 'AGSP', designed to amplify the 5' upstream sequences of the clones is shown. The stop codon is in bold, poly A tails are underlined and the polyadenylation signal is shown by '◆'

```
1                    GGCACG AGCAACAACT CAACAGTTCA CATCGAATCA pB3-1
        CGGCAC GAnAGACTC. TTCAACAACT CAACAGTTCA CATCGAATCA pB9-3

37    CCAAAATTGA CC.AAAAAAC CTACCGGAAA GTTCACCGGA ATACACGATT pB3-1
      CCAAAATTGA CCAAAAAAAC CAACCGGAAA GTTCACCGGA ATACACGATT pB9-3
                          ┌→ Leader sequence
87    CCGATAACTG ACCTTTACAA ATGGCGATTT CTACCTTAGC TAGACGAAAG pB3-1
      CCGTTAACTG ACCTCTACAA ATGGCGATTT CAACCTTAGC TAGACGAAAG pB9-3
          +                    *     Primer P    ⟶
137   GCTACCACGT TTCTGTCATC CAGACTCCTT TACAGCACTT CCAAGTATTC pB3-1
      GCTACAACGT TTCTGTCATC CAGATTCCTT TACAGCACTT CCAAGTATTC pB9-3
                            ┌→ E3 sequence
187   GTTCTCTTTA ACCAGAGGTT TCGCTTCAGG ATCCGATGAG AACGACGTCG pB3-1
      GTTCTCTTTA ACCAGAgGTT TCGCTTCAGG ATCCGATGAG AACGACGTCG pB9-3

237   TTGTTATCGG TGGTGGTCCC GGAGGCTATG TGGCGGCGAT CAAAGCAGCA pB3-1
      TTGTTATCGG TGGTGGTCCC GGAGGCTATG TGGCGGCGAT CAAAGCAGCA pB9-3
                               GGCTATG TGGCGGCGAT CAAAGCAGCA pTaq22
                               GGCTATG TGGCGGCGAT CAAAGCAGCA pTaq16
                          *
287   CAGCTTGGGC TTAAAACTAC ATGTATCGAG AAACGTGGTA CCCTCGGTGG pB3-1
      CAGCTTGGGC TTAAAACTAC ATGTATCGAG AAACGTGGTA CCCTCGGTGG pB9-3
      CAGcTTGGGC TTAAAACTCC ATGTATCGAG AAACGTGGTA CCCTCGGTGG pTaq22
      CAGCTTGGGC TTAAAACTAC ATGTATCGAG AAACGTGGTA CCCTCGGTGG pTaq16

337   TACTTGCCTT AACGTTGGTT GTATTCCTTC TAAGGCACTT CTTCATTCGT pB3-1
      TACTTGCCTT AACGTTGGTT GTATTCCTTC TAAGGCACTT CTTCATTCGT pB9-3
      TACTTGCCTT AACGTTGGTT GTATTCCTTC TAAGGCACTT CTTCATTCGT pTaq22
      TACTTGCCTT AACGTTGGTT GTATTCCTTC TAAGGCACTT CTTCATTCGT pTaq16
                                 ⟵        AGSP
387   CCCACATGTT TCATGAAGCG CAACATTCAT TTGCCAATCA TGGTGTGAAG pB3-1
      CCCACATGTT TCATGAAGCG CAACATTCAT TTGCCAATCA TGGTGTGAAG pB9-3
      CCCACATGTT TCATGAAGCG CAACATTCAT TTGcCAATCA TGGTGTGAaG pTaq22
      CCCACATGTT TCATGAAGCG CAACATTCAT TTGCCAATCA TGGTGTGAAG pTaq16
          Primer 1      ⟶                         *
437   TTTTCTTCTG TCGAAGTwGA TCTCCCTGCC ATGATGGGAC AAAAAGATAA pB3-1
      TTTTCTTCTG TCGAAGTwGA TCTCCCTGCC ATGATGGGAC AAAAAGATAA pB9-3
      TTTTCTTCTG TCGAAGTTGA TCTCCCTGCC ATGATGGGAC AAAAAGATAA pTaq22
                                                     GATAA pTaq23
      TTTTCTTCTG TCGAAGTTGA TCTCCCTGCC ATGATGGGAC GAAAAGATAA pTaq16
                                          +
487   AGCCGTGTCT AACTTAACAC GAGGTATAGA GGGTCTATTC AAGAAAAACA pB3-1
      AGCCGTGTCT AACTTAACAC GAGGTATAGA GGGTTTATTC AAGAAAAACA pB9-3
      AGCCGTGTCT AACTTAACAC GAGGTATAGA GGGTCTATTC AAGAAAAACA pTaq22
      AGCCGTGTCT AACTTAACAC GAGGTATAGA GGGTCTATTC AAGAAAAACA pTaq23
      AGCCGTGTCT AACTTAACAC G......... .......... .......... pTaq16

537   AAGTAAACTA TGTGAAGGGA TATGGTAAAT TCCTCTCCCC TTCTGAAATT pB3-1
      AAGTAAACTA TGTGAAGGGA TATGGTAAAT TCCTCTCCCC TTCTGAAATT pB9-3
      AAGTAAACTA TGTGAAGGGA TATGGTAAAT TCCTCTCCCC TTCTGAAATT pTaq22
      AAGTAAACTA TGTGAAGGGA TATGGTAAAT TCCTCTCCCC TTCTGAAATT pTaq23
      .......... .......... .......... .......... .......... pTaq16
                    Primer 2      ⟶
587   TCTGTTGACA CTGTCGAAGG TGGTAACTCT GTTGTTAAAG GCAAAAATAT pB3-1
      TCTGTTGACA CTGTCGAAGG TGGTAACTCT GTTGTTAAAG GCAAAAATAT pB9-3
      TCTGTTGACA CTGTCGAAGG TGGTAACTCT GTTGTTAAAG GCAAAAATAT pTaq22
      TCTGTTGACA CTGTCGAAGG TGGTAACTCT GTTGTTAAAG GCAAAAATAT pTaq23
      .......... .......AGG TGGTAACTCT GTTGTTAAAG GCAAAAATAT pTaq16
```

```
     *             +
637  TATTATTGCA ACTGGCTCTG ATGTCAAAGG TCTACCTGGC ATAACCATTG pB3-1
     TATTATTGCA ACTGGTTCTG ATGTCAAAGG TCTACCTGGC ATAACCATTG pB9-3
     TATTATTGCA ACTGGCTCTG ATGTCAAAGG TCTACCTGGC ATAACCATTG pTaq22
     TATTATTGCA ACTGGCTCTG ATGTCAAAGG TCTACCTGGC ATAACCATTG pTaq23
     TGTTATTGCA ACTGGCTCTG ATGTCAAAGG TCTACCTGGC ATAACCATTG pTaq16

687  ATGAGAAGAA AATTGTGTCA TCCACCGGAG CATTAGCTTT GACTGAAATT pB3-1
     ATGAGAAGAA AATTGTGTCA TCCACCGGAG CATTAGCTTT GACTGAAATT pB9-3
     ATGAGAAGAA AATTGTGTCA TCCACCGGAG CATTAGCTTT GACTGAAATT pTaq22
     ATGAGAAGAA AATTGTGTCA TCCACCGGAG CATTAGCTTT GACTGAAATT pTaq23
     ATGAGAAGAA AATTGTGTCA TCCACCGGAG CATTAGCTTT GACTGAAATT pTaq16

737  CCAAAAAGAT TAGTTGTTAT TGGTGCTGGC TATATAGGCC TTGAGATGGG pB3-1
     CCAAAAAGAT TAGTTGTTAT TGGTGCTGGC xxTATAGGCC TTGAGATGGG pB9-3
     CCAAAAAGAT TAGTTGTTAT TGGTGCTGGC TATATAGGCC TTGAGATGGG pTaq22
     CCAAAAAGAT TAGTTGTTAT TGGTGCTGGC TATATAGGCC TTGAGATGGG pTaq23
     CCAAAAAGAT TAGTTGTTAT TGGTGCTGGC TATATAGGCC TTGAGATGGG pTaq16
                                 Primer 3
787  ATCTGTCTGG GGTCGCCTCG GCTCAGAGGT GACTGTTGTT GAATTTGCAT pB3-1
     ATCTGTCTGG GGTCGCCTCG GCTCAGAGGT GACTGTTGTT GAATTTGCAT pB9-3
     ATCTGTCTGG GGTCGCCTCG GCTCAGAGGT GACTGTTGTT GAATTTGCAT pTaq22
     ATCTGTCTGG GGTCGCCTCG GCTCAGAGGT GACTGTTGTT GAATTTGCAT pTaq23
     ATCTGTCTGG GGTCGCCTCG GCTCAGAGGT GACTGTTGTT GAATTTGCAT pTaq16
                                        Primer Z
837  CTGATATTGT TCCAACCATG GATGGTGAAG TTCGCAAGCA ATTTCAACGT pB3-1
     CTGATATTGT TCCAACCATG GATGGTGAAG TTCGCAAGCA ATTTCAACGT pB9-3
     CTGATATTGT TCCAACCATG GATGGTGAAG TTCGCAAGCA ATTTCAACGT pTaq22
     CTGATATTGT TCCAACCATG GATGGTGAAG TTCGCAAGCA ATTTCAACGT pTaq23
     CTGATATTGT TCCAACCATG GATGGTGAAG TTCGCAAGCA ATTTCAACGT pTaq16

887  GCTCTTGAGA AGCAAAAGAT GAAATTCATG CTAAACACTA AGGTGGTGTC pB3-1
     GCTCTTGAGA AGCAAAAGAT GAAATTCATG CTAAACACTA AGGTGGTGTC pB9-3
     GCTCTTGAGA AGCAAAAGAT GAAATTCATG CTAAACACTA AGGTGGTGTC pTaq22
     GCTCTTGAGA AGCAAAAGAT GAAATTCATG CTAAACACTA AGGTGGTGTC pTaq23
     GCTCTTGAGA AGCAAAAGAT GAAATTCATG CTAAACACTA AGGTGGTGTC pTaq16

937  AGTTGATGCT ACAGGTGATG GTGTGAAATT GACCCTTGAA CCTTCAGCTG pB3-1
     AGTTGATGCT ACAGGTGATG GTGTGAAATT GACCCTTGAA CCTTCAGCTG pB9-3
     AGTTGATGCT ACAGGTGATG GTGTGAAATT GACCCTTGAA CCTTCAGCTG pTaq22
     AGTTGATGCT ACAGGTGATG GTGTGAAATT GACCCTTGAA CCTTCAGCTG pTaq23
     AGTTGATGCT ACAGGTGATG GTGTGAAATT GACCCTTGAA CCTTCAGCTG pTaq16

987  GTGGTGATCA AACTATTCTC GAGGCTGATG TTGTTCTCGT CTCTGCTGGA pB3-1
     GTGGTGATCA AACTATTCTC GAGGCTGATG TTGTTCTCGT CTCTGCTGGA pB9-3
     GTGGTGATCA AACTATTCTC GAGGCTGATG TTGTTCTCGT CTCTGCTGGA pTaq22
     GTGGTGATCA AACTATTCTC GAGGCTGATG TTGTTCTCGT CTCTGCTGGA pTaq23
     GTGGTGATCA AACTATTCTC GAGGCTGATG TTGTTCTCGT CTCTGCTGGA pTaq16
                                                  Primer Y
1037 AGAGTACCAT TCACTTCAGG ACTTGGATTG GACACGATAG GAGTTGAAAC pB3-1
     AGAGTACCAT TCACTTCAGG ACTTGGATTG GACACGATAG GAGTTGAAAC pB9-3
     AGAGTACCAT TCACTTCAGG ACTTGGATTG GACACGATAG GAGTTGAAAC pTaq22
     AGAGTACCAT TCACTTCAGG ACTTGGATTG GACACGATAG GAGTTGAAAC pTaq23
     AGAGTACCAT TCACTTCAGG ACTTGGATTG GACACGATAG GAGTTGAAAC pTaq16
          Primer 4
1087 TGACAAGGCT GGTAGAATCT TGGTCAATGA ACGTTTTGCC ACTAACGTCC pB3-1
     TGACAAGGCT GGTAGAATCT TGGTCAATGA ACGTTTTGCC ACTAACGTCC pB9-3
     TGACAAGGCT GGTAGAATCT TGGTCAATGA ACGTTTTGCC ACTAACGTCC pTaq22
     TGACAAGGCT GGTAGAATCT TGGTCAATGA ACGTTTTGCC ACTAACGTCC pTaq23
     TGACAAGGCT GGTAGAATCT TGGTCAATGA ACGTTTTGCC ACTAACGTCC pTaq16
```

```
                                                  +
1137 CAGGGGTACA TGCAATTGGT GATGTCATTC CTGGACCAAT GCTGGCTCAC pB3-1
     CAGGGGTACA TGCAATTGGT GATGTCATTC CTGGACCAAT GTTGGCTCAC pB9-3
     CAGGGGTACA TGCAATTGGT GATGTCATTC CTGGACCAAT GCTGGCTCAC pTaq22
     CAGGGGTACA TGCAATTGGT GATGTCATTC CTGGACCAAT GTTGGCTCAC pTaq23
     CAGGGGTACA TGCAATTGGT GATGTCATTC CTGGACCAAT GTTGGCTCAC pTaq16
                                 +
1187 AAGGCAGAGG AGGATGGCGT TGCTTGTGTA GAGTTCATTG CAGGTAAGGA pB3-1
     AAGGCAGAGG AGGATGGCGT TGCTTGCGTA GAGTTCATTG CAGGTAAGGA pB9-3
     AAGGCAGAGG AGGATGGCGT TGCTTGCGTA GAGTTCATTG CAGGTAAGGA pTaq22
     AAGGCAGAGG AGGATGGCGT TGCTTGCGTA GAGTTCATTG CAGGTAAGGA pTaq23
     AAGGCAGAGG AGGATGGCGT TGCTTGCGTA GAGTTCATTG CAGGTAAGGA pTaq16
                                         +
1237 GGGTCATGTA GACTATGATA TGGTACCTGG TGTCGTTTAC ACCCACCCGG pB3-1
     GGGTCATGTA GACTATGATA TGGTACCTGG TGTTGTTTAC ACCCACCCGG pB9-3
     GGGTCATGTA GACTATGATA TGGTACCTGG TGTTGTTTAC ACCCACCCGG pTaq22
     GGGTCATGTA GACTATGATA TGGTACCTGG TGTTGTTTAC ACCCACCCGG pTaq23
     GGGTCATGTA GACTATGATA TGGTACCTGG TGTTGTTTAC ACCCACCCGG pTaq16
                                                        *
1287 AAGTGGCTTA TGTTGGGAAA ACCGAGGAAC AGGTAAAGTC ACTCGGAGTT pB3-1
     AAGTGGCTTA TGTTGGGAAA ACCGAGGAAC AGGTAAAGTC ACTCGGAGTT pB9-3
     AGGTGGCTTA TGTTGGGAAA ACCGAGGAAC AGGTAAAGTC ACTCGGAGTT pTaq22
     AAGTGGCTTA TGTTGGGAAA ACCGAGGAAC AGGTAAAGTC ACTCGGAATT pTaq23
     AAGTGGCTTA TGTTGGGAAA ACCGAGGAAC AGGTAAAGTC ACTCGGAATT pTaq16
                                 \______ Primer X ______
                                                      
                               _______ + Primer 5 __\
1337 GATTATCGTG TTGGCAAATT CCCCTTCCTA GCAAACAGTA GAGCCAAGGC pB3-1
     GATTATCGTG TTGGCAAATT CCCCTTCCTA GCTAACAGTA GAGCCAAGGC pB9-3
     GATTATCGTG TTGGCAAATT CCCCTTCCTA GCTAACAGTA GAGCCAAGGC pTaq22
     GATTATCGTG TTGGCAAATT CCCCTTCCTA GCTAACAGTA GAGCCAAGGC pTaq23
     GATTATCGTG TTGGCAAATT CCCCTTCCTA GCTAACAGTA GAGCCAAGGC pTaq16
                       +                                 +
1387 AATTGATGAT GCTGAGGGAA TTGTCAAGGT AATTGCTGAG AAAGAGAGTG pB3-1
     AATTGATGAT GCTGAGGGAA TTGTCAAGGT AATTGCTGAG AAAGAGAGCG pB9-3
     AATTGATGAT GCTGAGGAAA TTGTCAAGGT AATTGCTGAG AAAGAGAGCG pTaq22
     AATTGATGAT GCTGAGGGAA TTGTCAAGGT AATTGCTGAG AAAGAGAGTG pTaq23
     AATTGATGAT GCTGAGGGAA TTGTCAAGGT AATTGCTGAG AAAGAGAGTG pTaq16
                   + *                    +
1437 ACAAGATATT GGGCGTCCAT ATTATGTCAC CTAATGCAGG GGAGCTTATT pB3-1
     ACAAGATATT GGGTGTCCAT ATTATGTCAC CTAACGCAGG GGAGCTTATT pB9-3
     ACAAGATATT GGGTGTCCAT ATTATGTCAC CTAACGCAGG GGAGCTTATT pTaq22
     ACAAGATATT GGGCGCCCAT ATTATGTCAC CTAATGCAGG GGAGCTTATT pTaq23
     ACAAGATATT GGGCGTCCAT ATTATGTCAC CTAATGCAGG GGAGCTTATT pTaq16
                 +
1487 CACGAAGCTG TACTGGCATT GCAGTACGGA GCATCAAGTG AGGACATTGC pB3-1
     CACGAAGCTG TCCTGGCATT GCAGTACGGA GCATCAAGTG AGGACATTGC pB9-3
     CACGAAGCTG TCCTGGCATT GCAGTACGGA GCATCAAGTG AGGACATTGC pTaq22
     CACGAAGCTG TACTGGCATT GCAGTACGGA GCATCAAGTG AGGACATTGC pTaq23
     CACGAAGCTG TACTGGCATT GCAGTACGGA GCATCAAGTG AGGACATTGC pTaq16
                                                 *       +
1537 TCGTACTTGT CATGCACATC CAACAATGAG TGAGGCACTC AAAGAAGCAG pB3-1
     TCGTACTTGT CATGCACATC CAACAATGAG TGAGGCACTC AAAGAAGCTG pB9-3
     TCGTACTTGT CATGCACATC CAACAATGAG TGAGGCACTC GAAGAAGCAG pTaq22
     TCGTACTTGT CATGCACATC CAACAATGAG TGAGGCACTC AAAGAAGCAG pTaq23
     TCGTACTTGT CATGCACATC CAACAATGAG TGAGGCACTC AAAGAAGCAG pTaq16
                                      *
1587 CCATGGCCAC TTATGACAAG CCCATCCACA TGTAGGCCTG TGTTATATCC pB3-1
     CCATGGCCAC TTATGACAAG CCCATCCACA TGTAGGCCTG TGTTATATCC pB9-3
     CCATGGCCAC TTATGACAAG CCCATCCACA TGTAGGCTTG TGTTATATCC pTaq22
     CCATGGCCAC TTATGACAAG CCCATCCACA CGTAGGCTTG TGTTATATCC pTaq23
     CCATGGCCAC TTATGACAAG CCCATCCACA TGTAGGCTTG TGTTATATCC pTaq16
                                                 \_Primer W
```

```
1637 TCGATAGCCT CATTTTTCTT ATTCCGTTGA GAATCTTGAA CGCTTAGAAT pB3-1
     TCGATAGCCT CATTTTTCTT ATTCCGTTGA GAATCTTGAA CGCTTAGAAT pB9-3
     TCGATAgCCT CATTTTTCTT ATTCCGTTGA GAATCTTGAA CGCTTAGAAT pTaq22
     TCAATAGCCT CATTTTTCAT ATTCCGTTGA AAATCTTGAA TGCTTAGAAT pTaq23
     TCGATAGCCT CATTTTTCAT ATTCCGtTGA GAATCTTGAA TGCTTAGAAT pTaq16
                   Primer 6
1687 AGGCTTTTTG TTGCTGCCGT TACTCAAGCA GCCCATATTC CCCTTAATTC pB3-1
     AGGCTTTTTG TTGCTGCCGT TACTCAAGCA GCCCATATTC CCCTTAATTC pB9-3
     AGGCTTTTTG TTGCTGCCGT TACTCAAGCA GCCCATATTC CTCTTnATTC pTaq22
     AGGCTTTTTG TTGCTGCCGT TACTCAAACA GCCCATATTC CCCTTTATtC pTaq23
     AGGCTTTTTG TTGCTGCCGT TACTCAAGCA GCCCATATTC CCCTTTATTC pTaq16
                                                       ◆◆
1737 AGGACTCCAA ATACACTATA TGAATGTTTA TCTTCTGCAT TTTTCCATAA pB3-1
     AGGACTCCAA ATACACTATA TGAATGTTTA TCTTCTGCAT TTTTCCATAA pB9-3
     AGGACTCCAA ATACACTATA TGAATGTTTA TCTTCTGCAT TTTTCCATAA pTaq22
     AGGACTCCAA ATACATTATA TGAATGTTTA TCATCTGCAT TTTTCCATAA pTaq23
     AGGACTCCAA ATACAtTATA TGAATGTTtA TCATCTGCAT TTTTCCATAA pTaq16
     ◆◆◆
1787 TAAAGATGAG ATGTAACCTG CTTTAAAAAA AAAAAAAAAA pB3-1
     TAAAGATGAG ATGTAACCTG CTTTAAACCC TGCAGAAAGT TAACTGCCTA pB9-3
     TAAAGATGAG ATGTAACCTG CTTTAAACCC TGCAGAA GT TAACTGCCTG pTaq22
     AAAAAAAAAA AAAAA pTaq23
     TAAAGATGAG ATGTAACCTG CTTTAAACCC TGCAGAA GT TAACTGCCTG pTaq16

1837 GCACTAAGAT GGTGCTAGAC ATTACAAAAC TTCCCATTTA CTTTGCTTTT pB9-3
     GCACTAAGAT GGTGCTAGAC ATTACAAAAC TTCCCATTTA CTTTGCTTTT pTaq22
     GCACTAAGAT GGTGCTAGAC ATTACAAAAA AAAAAAAAAA AAAAtT     pTaq16

1887 TTGGCATGTA TCCCTAAATT CTATATTTAT CTTGTTTGAT TTGGTAAAAA pB9-3
     TTGGCATGTA TCCCTAAATT CTATATTTAT CTTGTTTGAT TTGGTAGATC pTaq22

1937 AAAAAAAAAA AAAAAAAAAA ACTC  pB9-3
     AGTCGATTTGG TGAAAAAAAA AAAAAAAAAA pTaq22
```

Fig. B: Amino acid comparison of the library clones, pB3-1 and pB9-3 with the *Taq* amplified clones pTaq22, pTaq23 and pTaq16

§§= redox active cysteines.

```
          1                                        ↱         50
 pB3-1    MAISTLARRK ATTFLSSRLL YSTSKYSFSL TRGFASGSDE NDVVVIGGGP
 pB9-3    MAISTLARRK ATTFLSSRFL YSTSKYSFSL TRGFASGSDE NDVVVIGGGP

          51                              §     §           100
 pB3-1    GGYVAAIKAA QLGLKTTCIE KRGTLGGTCL NVGCIPSKAL LHSSHMFHEA
 pB9-3    GGYVAAIKAA QLGLKTTCIE KRGTLGGTCL NVGCIPSKAL LHSSHMFHEA
pTaq22     GYVAAIKAA QLGLKTPCIE KRGTLGGTCL NVGCIPSKAL LHSSHMFHEA
pTaq23
pTaq16     GYVAAIKAA QLGLKTTCIE KRGTLGGTCL NVGCIPSKAL LHSSHMFHEA

          101                                               150
 pB3-1    QHSFANHGVK FSSVEVDLPA MMGQKDKAVS NLTRGIEGLF KKNKVNYVKG
 pB9-3    QHSFANHGVK FSSVEVDLPA MMGQKDKAVS NLTRGIEGLF KKNKVNYVKG
pTaq22    QHSFANHGVK FSSVEVDLPA MMGQKDKAVS NLTRGIEGLF KKNKVNYVKG
pTaq23                              KDKAVS NLTRGIEGLF KKNKVNYVKG
pTaq16    QHSFANHGVK FSSVEVDLPA MMGRKDKAVS NLTR...... ..........

          151                                               200
 pB3-1    YGKFLSPSEI SVDTVEGGNS VVKGKNIIIA TGSDVKGLPG ITIDEKKIVS
 pB9-3    YGKFLSPSEI SVDTVEGGNS VVKGKNIIIA TGSDVKGLPG ITIDEKKIVS
pTaq22    YGKFLSPSEI SVDTVEGGNS VVKGKNIIIA TGSDVKGLPG ITIDEKKIVS
pTaq23    YGKFLSPSEI SVDTVEGGNS VVKGKNIIIA TGSDVKGLPG ITIDEKKIVS
pTaq16    .......... ......GGNS VVKGKNIVIA TGSDVKGLPG ITIDEKKIVS

          201                                               250
 pB3-1    STGALALTEI PKRLVVIGAG YIGLEMGSVW GRLGSEVTVV EFASDIVPTM
 pB9-3    STGALALTEI PKRLVVIGAG XIGLEMGSVW GRLGSEVTVV EFASDIVPTM
pTaq22    STGALALTEI PKRLVVIGAG YIGLEMGSVW GRLGSEVTVV EFASDIVPTM
pTaq23    STGALALTEI PKRLVVIGAG YIGLEMGSVW GRLGSEVTVV EFASDIVPTM
pTaq16    STGALALTEI PKRLVVIGAG YIGLEMGSVW GRLGSEVTVV EFASDIVPTM

          251                                               300
 pB3-1    DGEVRKQFQR ALEKQKMKFM LNTKVVSVDA TGDGVKLTLE PSAGGDQTIL
 pB9-3    DGEVRKQFQR ALEKQKMKFM LNTKVVSVDA TGDGVKLTLE PSAGGDQTIL
pTaq22    DGEVRKQFQR ALEKQKMKFM LNTKVVSVDA TGDGVKLTLE PSAGGDQTIL
pTaq23    DGEVRKQFQR ALEKQKMKFM LNTKVVSVDA TGDGVKLTLE PSAGGDQTIL
pTaq16    DGEVRKQFQR ALEKQKMKFM LNTKVVSVDA TGDGVKLTLE PSAGGDQTIL

          301                                               350
 pB3-1    EADVVLVSAG RVPFTSGLGL DTIGVETDKA GRILVNERFA TNVPGVHAIG
 pB9-3    EADVVLVSAG RVPFTSGLGL DTIGVETDKA GRILVNERFA TNVPGVHAIG
pTaq22    EADVVLVSAG RVPFTSGLGL DTIGVETDKA GRILVNERFA TNVPGVHAIG
pTaq23    EADVVLVSAG RVPFTSGLGL DTIGVETDKA GRILVNERFA TNVPGVHAIG
pTaq16    EADVVLVSAG RVPFTSGLGL DTIGVETDKA GRILVNERFA TNVPGVHAIG

          351                                               400
 pB3-1    DVIPGPMLAH KAEEDGVACV EFIAGKEGHV DYDMVPGVVY THPEVAYVGK
 pB9-3    DVIPGPMLAH KAEEDGVACV EFIAGKEGHV DYDMVPGVVY THPEVAYVGK
pTaq22    DVIPGPMLAH KAEEDGVACV EFIAGKEGHV DYDMVPGVVY THPEVAYVGK
pTaq23    DVIPGPMLAH KAEEDGVACV EFIAGKEGHV DYDMVPGVVY THPEVAYVGK
pTaq16    DVIPGPMLAH KAEEDGVACV EFIAGKEGHV DYDMVPGVVY THPEVAYVGK
```

```
        401                                                450
pB3-1   TEEQVKSLGV DYRVGKFPFL ANSRAKAIDD AEGIVKVIAE KESDKILGVH
pB9-3   TEEQVKSLGV DYRVGKFPFL ANSRAKAIDD AEGIVKVIAE KESDKILGVH
pTaq22  TEEQVKSLGV DYRVGKFPFL ANSRAKAIDD AEGIVKVIAE KESDKILGVH
pTaq23  TEEQVKSLGI DYRVGKFPFL ANSRAKAIDD AEGIVKVIAE KESDKILGAH
pTaq16  TEEQVKSLGI DYRVGKFPFL ANSRAKAIDD AEGIVKVIAE KESDKILGVH

        451                                                500
pB3-1   IMSPNAGELI HEAVLALQYG ASSEDIARTC HAHPTMSEAL KEAAMATYDK
pB9-3   IMSPNAGELI HEAVLALQYG ASSEDIARTC HAHPTMSEAL KEAAMATYDK
pTaq22  IMSPNAGELI HEAVLALQYG ASSEDIARTC HAHPTMSEAL EEAAMATYDK
pTaq23  IMSPNAGELI HEAVLALQYG ASSEDIARTC HAHPTMSEAL KEAAMATYDK
pTaq16  IMSPNAGELI HEAVLALQYG ASSEDIARTC HAHPTMSEAL KEAAMATYDK

        501
pB3-1   PIHM
pB9-3   PIHM
pTaq22  PIHM
pTaq23  PIHT
pTaq16  PIHM
```

GLASGOW
UNIVERSITY
LIBRARY