CLONING AND CHARACTERISATION OF THE PLANT PYRUVATE DEHYDROGENASE COMPLEX COMPONENTS

Donna McGow

Thesis submitted to the University of Glasgow for the degree of Doctor of Philosophy.

Division of Biochemistry and Molecular Biology, IBLS

September, 2002

Declaration

I hereby declare that the work presented in this thesis is my own, except where otherwise cited or acknowledged. No part of this thesis has been presented for any other degree.

Donna McGow

September, 2002

In loving memory of Benji,
the best friend a girl could ever wish for,
and for Spock and Ginger,
there will always be a place in my heart for you all.

ACKNOWLEDGEMENTS

I would like to thank first and foremost Professor J. G. Lindsay for giving me the chance to do my PhD in his laboratory and for all the help, support and advice throughout my work and especially through the writing of this thesis. I am very grateful.

Thanks goes to the BBSRC for the funding throughout the past three years.

My thanks go to Dr. Sharon Kelly for the help and assistance in performing the CD and fluorescence measurements and for the help with MALDI-TOF analysis and finally for all the help in the interpretation of the results. I would also like to thank Dr. Olwyn Byron and Gordon Campbell for their time spent performing sedimentation equilibrium experiments and analysing the data. I am grateful to Dr. Olwyn Byron for her time spent on explaining the finer points of this technique.

I would like to thank all the past and present members of the laboratory, Heather, Susan, Audrey, Louise, Anu and Lorraine, they have made the laboratory a very happy place to work in and have made the past 3 years go very fast. My friends, Jane and Marie Anne for the long phone calls and longest tea breaks ever known in Whistler's Mother and my family, I thank them all for their support and encouragement throughout the PhD years and especially through the writing stage, it can't have been easy, I'm sorry.

I'd like to thank my sister Fiona, for the use of her computer, or the complete domination of her computer. Thank you so much.

Finally, I would like to thank the computer, who was always there for me.!!

ABSTRACT

Multienzyme complexes are non-covalent assemblies composed of one or more copies of their constituent enzymes that act cooperatively to catalyse a linked series of reactions. These multienzymes complexes are of great interest in the study of protein-protein interactions, substrate channelling and the catalytic and regulatory advantages conferred by their organisation into multimeric complexes. One such complex is pyruvate dehydrogenase complex (PDC), the principal member of the 2-oxo acid dehydrogenase complex family. PDC is located in the mitochondrial matrix in eukaryotes, where it catalyses the irreversible oxidative decarboxylation of pyruvate committing its acetyl-CoA product to enter into the citric acid cycle or ketone body formation. Plants are unique in having two distinct and spatially separate complexes; a mitochondrial complex similar to the mammalian and a chloroplastic complex with a unique anabolic function in providing acetyl-CoA and NADH for *de novo* fatty acid synthesis. Although much is known about the organisation and stoichiometry of the PDC from mammalian, yeast and prokaryotic sources little is known of the composition, subunit organisation and stoichiometry of the plant equivalent. Until recently, studies have been hindered by the low abundance of PDC in plant tissues. Nevertheless the plant PDC from both organelles are thought to be composed of the same three enzymes found in mammalian PDC designated E1, E2 and E3. In all PDCs studied to date E1 and E3 are arranged around an oligomeric E2 core. Two types of E2 core have been identified; the 24meric cube displaying octahedral (432) symmetry and the 60meric pentagonal dodecahedron displaying icosahedral (532) symmetry.

In this thesis, various enzymes from plant mitochondrial and chloroplastic PDCs have been cloned from either recombinant plasmids containing the genes encoding PDC components or from cDNA produced from mRNA. These enzymes were cloned into different expression vectors that incorporate expression tags (His-tag or GST-fusion) to the N-terminal region of the recombinant enzymes and successfully expressed in *E. coli*. Subsequently these enzymes were purified to near homogeneity by Zn^{2+} ion affinity chromatography, as judged by Coomassie blue staining. High yields of enzymatically active recombinant enzymes were achieved

using this purification method. The plant mitochondrial and chloroplastic E2 component was expressed as either the mature E2 enzyme or as a truncated form comprising the lipoyl domain(s) and adjacent subunit (E1 and/or E3) binding domain. Successful lipoylation of E2 indicates that these recombinant enzymes were capable of folding into their native 3D structures. A monoclonal antibody (mAb) specific for lipoylated human E3-BP and E2 was used to detect lipoylated plant E2s. However, this antibody only recognised the mitochondrial E2. The specificity of the antibody for mitochondrial E2s suggests that in addition to the presence of the lipoic acid cofactor the antibody also recognises regions of sequence specific to the mitochondrial E2 lipoyl domain. Sequence alignments of mitochondrial and chloroplastic/ plastidic lipoyl domains reveal differences that may be the potential epitope(s) for the mAb. Lipoylation of the chloroplastic E2 was verified by measuring the mass of the E2 didomain GST-fusion proteins by MALDI-TOF after expression in media with or without exogenous lipoic acid.

The probable subunit organisation of plant mitochondrial and chloroplastic E2 cores was investigated by analysing their M_r values using analytical ultracentrifugation in the sedimentation equilibrium mode. These preliminary studies show a distinct difference in M_r value between the two E2 complexes. The chloroplastic E2 was estimated to have a M_r value consistent with it assembling into a 24meric oligomer while the M_r value estimated for the mitochondrial E2 suggests assembly into a 60meric core.

Finally, the expression of recombinant potato E3 and pea L-protein in *E. coli* did not hinder these enzymes from forming their native dimeric structure. The stability of the dimeric structure was studied and was proved to be essential for E3 activity, with loss of secondary and tertiary structure correlating with the loss of activity. Fluorescence studies on both the potato E3 and pea L-protein showed the release of FAD with the unfolding of the E3 enzyme and is consistent with the loss of E3 activity. The effect of the 31 amino acid mitochondrial targeting sequence on the structure of the L-protein was also studied using circular dichroism and fluorescence. The *in vivo* function of a targeting peptide is to maintain the protein in a transient unfolded form until the protein reaches its final destination with the

cell. The L-protein precursor was expressed in *E. coli* in a partially unfolded form possibly similar to that found *in vivo.* This was further substantiated by its inactivity and insolublity and by the structural changes observed by CD and fluorescence after pre-treatment with urea.

CONTENTS

TABLE OF CONTENTS **Page**

LIST OF FIGURES **Page**

CHAPTER 1

CHAPTER 3

CHAPTER 4

CHAPTER 5

CHAPTER 6

CHAPTER 7

LIST OF TABLES **Page**

ABBREVIATIONS

BCDC	branched chain 2-oxoacid dehydrogenase complex
BSA	bovine serum albumin
CD	circular dichroism
cDNA	complementary deoxyribonucleic acid
CoA	coenzyme A
(k)Da	(kilo) Daltons
DCPIP	2, 6-dichlorophenol-indophenol
DEPC	diethyl pyrocarbonate
DHL	dihydrolipoamide, [Lip(SH_2)]
DTT	dithiothreitol
E1	pyruvate dehydrogenase/ decarboxylase
E2	dihydrolipoamide acetyltransferase
E3	lipoamide dehydrogenase
EDTA	ethylene diamine tetra-acetate
FAD	flavin adenine dinucleotide
FPLC	fast protein liquid chromatography
GDC	glycine decarboxylase complex
GdmCl	guanidinium chloride
GST	glutathione S-transferase
h	hour
His	histidine
IPTG	isopropyl β-thiogalactopyranoside
KPi	potassium phosphate
mA/min	milli-absorbance units/minute
MALDI-TOF	matrix assisted laser desorption ionisation-time of flight
min	minute
MOPS	3-[N-Morpholino] propanesulphonic acid
mM	millimolar
M_r	relative molecular mass
NaAC	sodium acetate

NAD^+	nicotinamide adenine dinucleotide (oxidised form)
NADH	nicotinamide adenine dinucleotide (reduced form)
NaPi	sodium phosphate
OGDC	2-oxoglutarate dehydrogenase complex
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PDC	pyruvate dehydrogenase complex
PEG	polyethlene glycol
PMSF	phenymethylsulphonyl fluoride
RNA	ribonucleic acid
RT-PCR	reverse transcriptase-polymerase chain reaction
s	seconds
SDS	sodium dodecyl sulphate
TCA	trichloroacetic acid
TEMED	N, N, N^1, N^1-tetramethlethylene diamine
ThDP	thaimine diphosphate
Tris	Tris(hydroxymethyl) aminomethane
vol	volumes
w	weight

CHAPTER 1

INTRODUCTION

1. INTRODUCTION

1. 1. THE 2-OXO ACID DEHYDROGENASE COMPLEXES

The family of 2-oxo acid dehydrogenase complexes are involved collectively in the oxidative decarboxylation of a variety of 2-oxo acids releasing CO_2 and generating the relevant acyl CoA and NADH (Perham and Packman, 1990). As a group they represent a unique example of stable multienzyme assemblies with molecular weights in the range 4-10 million. Multienzyme complexes are non-covalent aggregates of protein subunits which work together to enhance overall catalytic activity by successive substrate channelling in an intermolecular manner (Perham, 1991). These multienzyme complexes have been intensively studied as general models for analysis of macromolecular design and assembly, the importance of protein-protein interactions and the catalytic and regulatory advantages conferred by ordered arrangement into multi-subunit structures.

The three principal members of this family are involved in key positions in carbohydrate and amino acid metabolism (see Figure 1.1). The pyruvate dehydrogenase complex (PDC) catalyses the irreversible oxidative decarboxylation of pyruvate, regulating the entry of 2-carbon units into the citric acid cycle but also providing acetyl CoA for the formation of ketone bodies. The second member, the 2-oxoglutarate dehydrogenase complex (OGDC) is a key regulatory enzyme of the tricarboxylic acid cycle converting 2-oxoglutarate to succinyl CoA. The third member is the branched-chain 2-oxo acid dehydrogenase complex (BCDC) which catalyses an irreversible step in the catabolism of several branched-chain 2-oxo acids produced by transamination of leucine, isoleucine and valine (Reed and Hackert, 1990; Perham, 1991). A close relative of this multienzyme family is the glycine decarboxylase complex (GDC), found in the mitochondria of animals and plants where it is responsible for the decarboxylation of glycine into CO_2, NH_4^+ and N^5, N^{10}-methylene-tetrahydrofolate (THF) (Oliver, 1994).

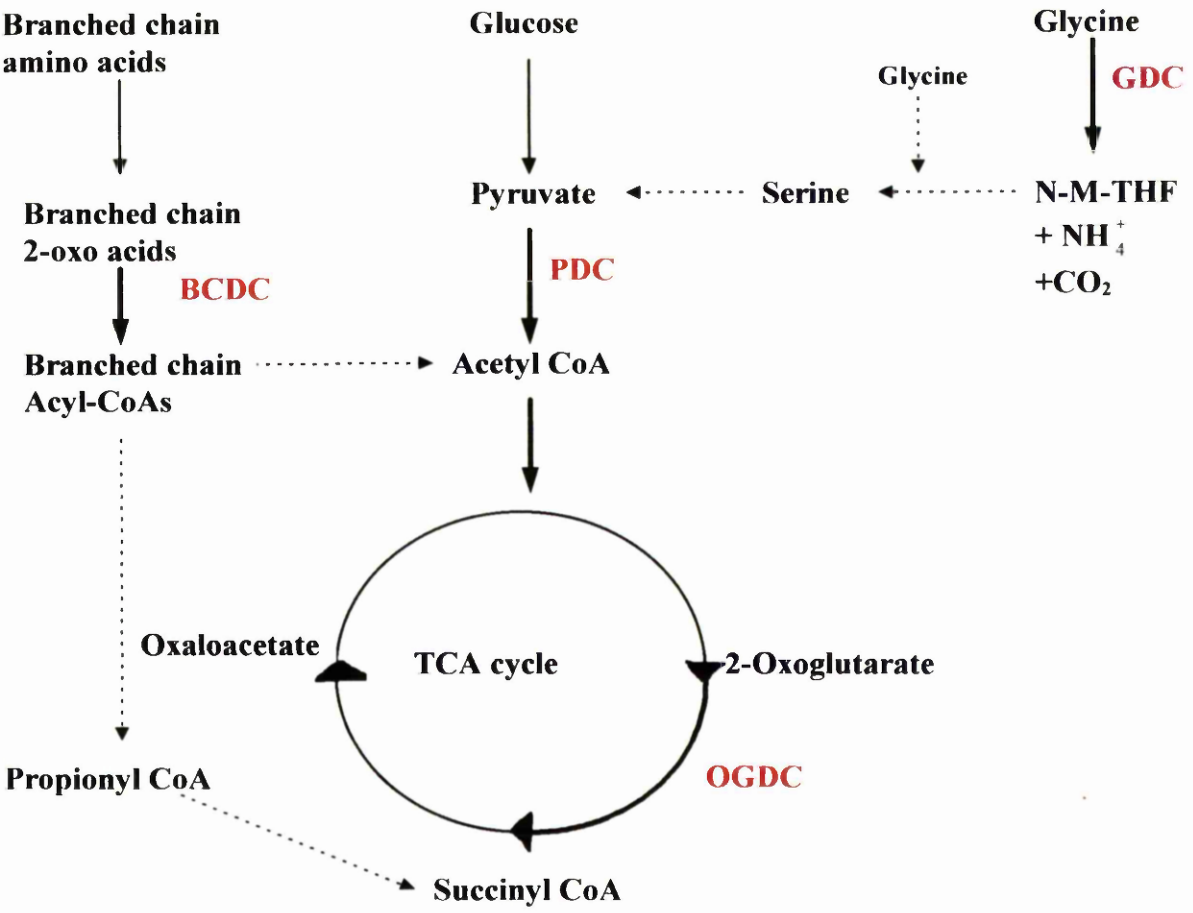


FIGURE 1.1: Metabolic Pathways Involving the 2-Oxo Acid Dehydrogenase Complexes from Mitochondria

PDC: pyruvate dehydrogenase complex, catalysing the conversion of pyruvate to acetyl CoA and committing the entry of carbon to the citric acid cycle.

OGDC: 2-oxoglutarate dehydrogenase complex, an integral enzyme in the citric acid cycle converting 2-oxoglutarate to succinyl CoA.

BCDC: branched-chain 2-oxo acid dehydrogenase complex oxidatively decarboxylates the branched-chain 2-oxo acids produced from the transamination of leucine, isoleucine and valine.

GDC: glycine decarboxylase complex, converts glycine to N^5, N^{10}-methylene tetrahydrofolate (N-M-THF), NH_4^+ and CO_2

Plants are unique in containing two distinct, spatially separate pyruvate dehydrogenase complexes. Most of the early work on crude mitochondrial and chloroplastic extracts or partially purified complexes showed the presence of distinct enzymatic activities within these compartments. One is located within the mitochondrial matrix similar to other eukaryotic systems, but a distinct PDC is located in the chloroplast stroma (Figure 1.2). With the apparent absence of the citric acid cycle in the chloroplastic compartment (Elias and Givan, 1979), this raised the question of the function of the plastidic PDC.

Chloroplastic PDC fulfils a unique anabolic role linking pyruvate metabolism and chloroplast fatty acid and isoprenoid biosynthesis by providing acetyl-CoA as a precursor (Williams and Randall, 1979). Plastids contain all the essential enzymes for fatty acid biosynthesis (Figure 1.3). Besides fatty acids, acetyl-CoA is required for synthesis of isoprenoids (carotenoids and the phytyl chain of chlorophylls). The incorporation of acetate, acetyl-CoA, malonate and malonyl-CoA into fatty acids was demonstrated using ruptured immature spinach chloroplasts; however, they were not effective precursors in fatty acid synthesis in intact chloroplasts where the membrane was impermeable to CoA derivatives (Williams and Randall, 1979). A pool of acetyl-CoA within the chloroplast is therefore required for *de novo* fatty acid production.

The source of acetyl-CoA in plastids was identified from the incorporation of ^{14}C-acetate into fatty acids. Acetate generated in the mitochondrial compartment from the hydrolysis of acetyl-CoA by acetyl-CoA hydrolase, penetrates cellular biomembranes. On entering the plastid it is then converted back to acetyl-CoA by acetyl-CoA synthetase (Figure 1.3), located exclusively in plastids (Golz *et al.*, 1994). However, with the discovery of plastid PDC which also provides acetyl-CoA and NADH for *de novo* fatty acid synthesis and polyisoprenoid biosynthesis (Camp and Randall, 1985) it is evident that there are at least two routes of providing acetyl-CoA as a substrate for these anabolic pathways.

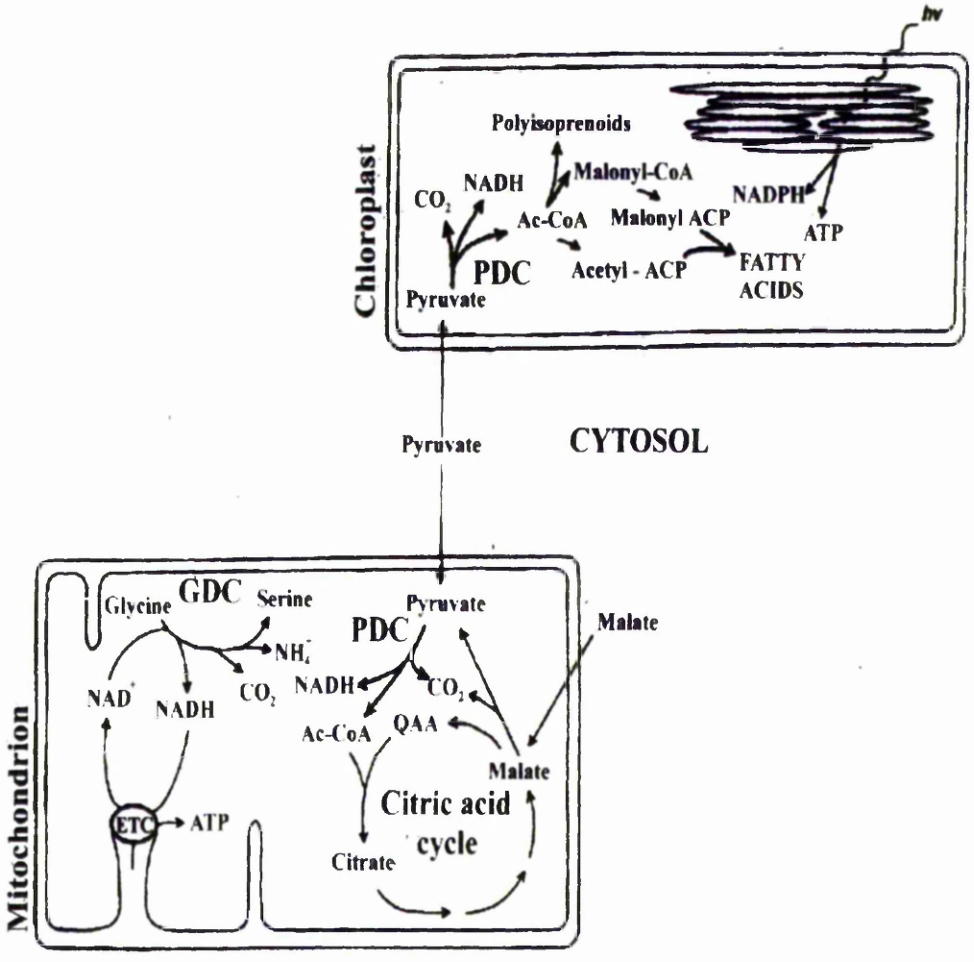


FIGURE 1.2: Location of the Plant Pyruvate Dehydrogenase Complexes

The dual location of the plant pyruvate dehydrogenase complexes and their role within their respective compartments. The chloroplast PDC provides acetyl-CoA (Ac-CoA) for the biosynthesis of fatty acids and polyisoprenoids and the mitochondrial PDC provides acetyl-CoA as the entry point of carbon into the citric acid cycle.

(Adapted from Luethy *et al.*, 1996)

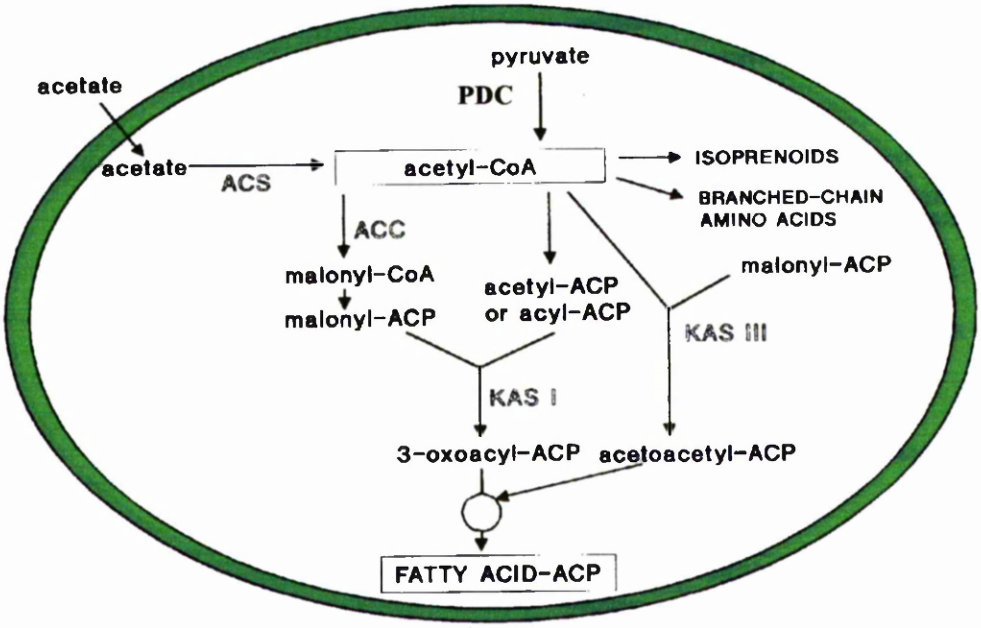


FIGURE 1.3: Scheme of the *De Novo* Fatty Acid Biosynthesis Pathway from Acetate and Pyruvate in Plastids of Higher Plants

ACS: acetyl-CoA synthetase; ACC: acetyl-CoA carboxylase; KAS I and III: β-keto-ACP synthases; PDC: plastidic pyruvate dehydrogenase complex.

(Taken from Golz *et al.*, 1994).

The plant pyruvate dehydrogenase complexes may share the same substrates, products and cofactors but they differ in their biochemical requirements. The optimal pH for pea chloroplastic PDC is 7.8-8.2 while its mitochondrial counterpart favours a slightly more acidic 7.3-7.7 environment (Williams and Randall, 1979). The organelle-specific complex requirement for divalent cations (Mg^{2+}) also differs with pea chloroplastic PDC having a higher optimum concentration than its mitochondrial counterpart (Williams and Randall, 1979). These differences are consistent with the idea that the chloroplastic PDC is the major provider of fatty acid precursors. *De novo* fatty acid synthesis occurs during periods of illumination (Rougham *et al.*, 1980). During photosynthesis the chloroplast stroma becomes more alkaline and the concentration of Mg^{2+} ions also increases producing the optimum conditions for chloroplastic PDC activity.

1. 1. 1. The Components of the 2-Oxo Acid Dehydrogenase Complexes

All three multienzyme assemblies are composed of multiple copies of three separate enzymes: a substrate specific decarboxylase-dehydrogenase (E1); a complex specific dihydrolipoamide acyltransferase (E2) and lipoamide dehydrogenase (E3), a common component of all three complexes in mammals. The overall reaction requires the action of all three enzymes working in concert plus the involvement of five cofactors; thiamine diphosphate (ThDP); lipoamide [Lip(S_2)]; nicotinamide adenine dinucleotide (NAD^+); flavin adenine dinucleotide (FAD) and CoA. Each of these cofactors associate with specific enzymes within the complex. A fourth component found in eukaryote PDCs is the E3 binding protein, which also carries a lipoic acid prosthetic group. The sole function of this polypeptide described to date is a role in binding E3 to the E2 core with a high affinity. The components for each 2-oxo acid dehydrogenase complex are listed in Table 1.1.

The related glycine decarboxylase complex is a loose association of four individual enzymes, which individually are similar to their 2-oxo acid dehydrogenase complex counterparts (see Table 1.2). The P-protein is a pyridoxal-5'-phosphate dependent amino acid decarboxylase; H-protein is a carrier protein that has a bound lipoic acid cofactor; T-protein is a tetrahydrofolate-linked methyl transferase and the fourth enzyme involved is the L-protein, a FAD-containing dihydrolipoamide dehydrogenase. Like the 2-oxo acid dehydrogenase complexes these four enzymes act in concert to convert one molecule of glycine to N^5, N^{10}-methylene tetrahydrofolate, releasing CO_2 and NH_4^+ with the concomitant production of NADH.

		Complex		
Enzyme	**Cofactor**	**PDC**	**OGDC**	**BCDC**
E1	ThDP	Pyruvate Dehydrogenase	2-Oxo acid Dehydrogenase	Branched-chain 2-Oxo acid Dehydrogenase
E2	Lipoic acid CoA	Dihydrolipoamide Acetyltransferase	Dihydrolipoamide Succinyltransferase	Dihydrolipoamide Acyltransferase
E3	FAD NAD^+	Dihydrolipoamide Dehydrogenase		
E3-BP (Protein X)	Lipoic Acid	Binds E3 to E2 core	NONE	NONE

TABLE 1.1: Enzyme Components of the Mammalian 2-Oxo Acid Dehydrogenase Complexes

PDC: pyruvate dehydrogenase complex

OGDC: 2-oxo acid dehydrogenase complex

BCDC: branched-chain 2-oxo acid dehydrogenase complex

In addition to these components, an E3-binding-protein (E3-BP) is also found in the mammalian, yeast and nematode PDCs. This subunit also houses a catalytically active lipoamide cofactor. In general the E3 component is common to all complexes although there are complex specific and organelle specific E3s found in plants and bacteria.

Subunit	Enzyme Activity	Cofactor	PDC Equivalent
P-protein	glycine decarboxylase	pyridoxal phosphate (PLP)	**E1**
H-protein	carrier protein	lipoic acid	**E2 (lipoyl domain)**
T-protein	THF transferase	tetrahydrofolate (THF)	**E2 (catalytic domain)**
L-protein	dihydrolipoamide dehydrogenase	FAD NAD^+	**E3**

TABLE 1.2: Component Enzymes that Form the Glycine Decarboxylase Complex

The constituent enzymes that make the glycine decarboxylase complex (glycine cleavage system) with their essential prosthetic group and their counterpart in the pyruvate dehydrogenase complex. The H-protein is in excess and, in the dissociated complex, it acts as a substrate for P, T and L-proteins.

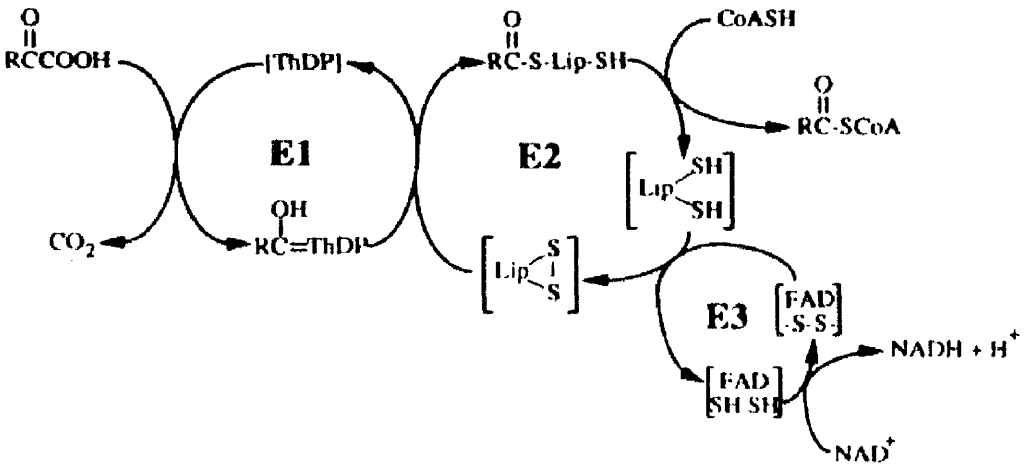
1. 1. 2. Enzyme Mechanism for the 2-Oxo Acid Dehydrogenase Complexes

The overall reaction mechanism of the 2-oxo acid dehydrogenase complexes involves the three non-covalently bound enzymes acting in five successive steps to catalyse the reaction shown in Figure 1.4.

The first step involves a substrate specific 2-oxo acid dehydrogenase (E1) which catalyses both the decarboxylation of the cognate 2-oxo acid (reaction 1) and reductively acylates the lipoyl moiety (reaction 2) covalently bound to E2. In the case of PDC, for example, decarboxylation is promoted by the E1-linked ThDP cofactor. ThDP is ionized to form a carbanion which reacts with the carbonyl group of pyruvate, forming a hydroxyethylidene-ThDP intermediate with the release of CO_2. This first step is the rate limiting step in the overall reaction scheme.

In the second step, also E1-catalysed, the cognate ThDP intermediate participates in the reductive acetylation of a lipoic acid prosthetic group covalently bound to the dihydrolipoamide acyltransferase (E2). In PDC, the addition of the acetyl group to E2 produces a S^6 or S^8 acetyl-dihydrolipoamide intermediate. The E2 enzyme catalyses the transfer of the acetyl group to CoA, (reaction 3) producing the relevant acyl-CoA and leaving the E2-linked lipoamide in its reduced state.

Dihydrolipoamide dehydrogenase (E3) is responsible for catalysing the reoxidation of the dihydrolipoyl moiety with NAD^+ as the ultimate electron acceptor. The penultimate step involves a charge transfer complex formed from a disulphide bridge and the FAD cofactor at the catalytic centre. The dihydrolipoamide cofactor of E2 is oxidised with the transfer of a hydride ion to the disulphide bridge and FAD allowing the lipoic acid to return to its oxidised state (reaction 4). The final step, is the oxidation of the reduced E3-bound FAD cofactor by NAD^+, the final electron acceptor (reaction 5).



Net Reaction

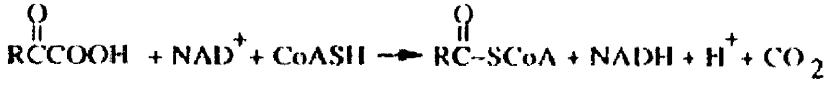


FIGURE 1.4: The Reaction Mechanism of Oxidative Decarboxylation by 2-Oxo Acid Dehydrogenase Complexes

For PDC R=CH_3; for OGDC R=CH_2CH_2COOH; for BCDC R=$(CH_3)_2CH$, $(CH_3)_2CHCH_2$, or $(CH_3)(C_2H_5)CH$.

ThDP, thiamine diphosphate; $Lip(S)_2$ and $Lip(SH)_2$, oxidised and reduced forms of the lipoic acid cofactor; FAD-$(S)_2$ and FAD-$(SH)_2$, oxidised and reduced forms of the charge transfer complex including flavin adenine dinucleotide and two cysteine residues (S); NAD^+ and NADH, oxidised and reduced forms of nicotinamide adenine nucleotide.

(Taken from Perham, 1991)

1. 1. 3. The Action of the Glycine Decarboxylase Complex

The glycine decarboxylase complex (GDC) catalyses the oxidation of glycine into CO_2, NH_4^+, N^5, N^{10}-methylene tetrahydrofolate and NADH within the mitochondrial matrix. The N^5, N^{10}-methylene tetrahydrofolate produced reacts with another molecule of glycine to form serine in a reaction catalysed by serine hydroxymethyltransferase (SHMT) as shown in Figure 1.1.

The reaction scheme of GDC mirrors that of PDC, requiring the action of multiple enzymes working in concert. Glycine oxidation occurs in step one of a five step reaction (Figure 1.5) where a Schiff base forms between the carbonyl group of pyridoxal phosphate (PLP) and α-amino group of the glycine (Walker and Oliver, 1986). This reaction is promoted by the glycine decarboxylase enzyme (P-protein).

The aminomethyl moiety is transferred (step 2) to one of the sulphydryl groups of the lipoamide cofactor bound to the H-protein via an amide linkage to a specific lysine residue. During this transfer the α-carboxyl group of glycine is released as CO_2 while electrons reduce the sulphydryl groups on lipoamide.

In step three the aminomethyl-lipoamide intermediate leaves the active site of the P-protein and moves to that of the T-protein where the α-amino group is released as NH_4^+ while the tetrahydrofolate (THF) cofactor acts as an acceptor for the remaining one carbon unit.

The final two steps of the overall GDC reaction are identical to the last two steps of PDC and involve the oxidation of the reduced lipoamide cofactor on the P-protein via the FAD cofactor of L-protein (step 4) with the concomitant reduction of NAD^+ to NADH + H^+ (step 5) (Walker and Oliver, 1996; Sarojini and Oliver, 1983).

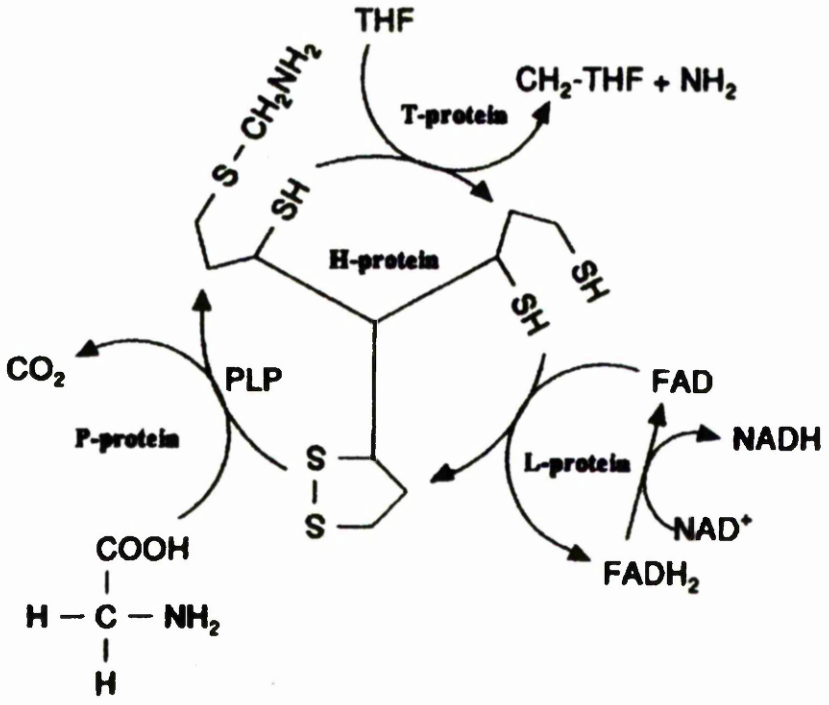


FIGURE 1.5: The Reaction Mechanism of the Glycine Decarboxylase Complex

PLP: pyridoxal phosphate cofactor

Lipoamide cofactor of H-protein in oxidised form (S—S); reduced form (SH_2); aminomethyl-lipoamide intermediate (S-CH_2NH_2)

THF: tetrahydrofolate cofactor

FAD, $FADH_2$: oxidised and reduced forms of flavin adenine dinucleotide

NAD^+, NADH: oxidised and reduced forms of nicotinamide adenine nucleotide

(Taken from Oliver, 1994)

1. 1. 4. Composition of the Organelle-Specific PDCs

In addition to being spatially and biochemically distinct, the composition of these assembles has proved to differ between organelles and various plant species. Table 1.3 shows the M_r of the PDC components from the purified PDC of three plant sources as judged by their migration on SDS-PAGE. Different techniques such as N-terminal sequencing of protein bands excised from SDS polyacrylamide gels and immuno-detection using subunit specific antibodies were employed to aid their identification.

The E2, putative E3-binding protein and the E3 components from pea chloroplast and mitochondrial PDC were initially identified by cross-reacting with mammalian subunit specific antisera. A M_r value of 50, 000 for both the chloroplast and mitochondrial E2 subunit of pea PDC (Taylor *et al.*, 1992a) was observed and was unusually low compared to the M_r for human (74, 000) but close to that of the yeast (58, 000). The validity of these low M_r values for E2 was strengthened by the detection of two E2 subunits from maize with a M_r values of 52, 000 and 53, 000 (Thelen *et al.*, 1998). The low M_r may be indicative of the presence of a single lipoyl domain while the potato E2 subunit, identified by N-terminal sequencing, has a M_r of 78, 000, similar to that of the human E2 suggesting that it contains two lipoyl domains (Millar *et al.*, 1998).

The first identification of multiple E3 components in plants was obtained from studies in pea and potato. Pea contains organelle specific E3 isoforms with distinct M_r values of 52, 000 for the chloroplastic form and 56, 000 for its mitochondrial counterpart (Taylor *et al.*, 1992a). Verification of two E3 species was obtained by the purification of each enzyme (Conner *et al.*, 1996a & b). Two E3 proteins were subsequently found to exist in maize with M_r values of 62, 000 and 63, 000 (Thelen *et al.*, 1998). Purification of E3 from potato tubers revealed the existence of three mitochondrial isoforms, designated α_2, β_2 and $\alpha\beta$ (Cook *et al.*, 1996) where the α polypeptide had a M_r of 58, 000 and 56, 000 for the β polypeptide. This observation was further substantiated by the detection of two E3 bands identified

by N-terminal sequencing, where these E3 proteins had distinct M_r values of 55, 000 and 58, 000 (Millar *et al.*, 1998).

The presence of an E3-binding protein (protein-X) has not been established unequivocally in plants although cross-reacting components (M_r of 48, 000 (chl) and 67, 000 (mt)) were identified in pea by employing mammalian protein-X antisera. A potential candidate for E3-binding protein may have been detected in potato mitochondrial PDC, where N-terminal sequencing of a single band from purified PDC preparations (M_r of 55, 000) produced two distinct sequences. One was an E3 isoform whereas the other had a sequence similar to but distinct from potato E2 (M_r of 78, 000). Both E2 proteins were shown to be acetylated (Millar *et al.*, 1998). Further investigation is required to confirm the function of this polypeptide. Currently, the existence of E3-binding proteins in plant mitochondrial and chloroplastic PDC isoforms has still to be established conclusively.

The E1 subunits are poorly characterised in the three plants that have been the subject of the most intensive study (pea, potato and maize). No cross-reaction was observed with mammalian E1 antisera; however, yeast E1 antisera highlighted a putative mitochondrial E1α subunit with a M_r of 41, 000 (Taylor *et al.*, 1992a) in pea. Anti-potato E1 antibodies were employed in the recognition of the corresponding maize E1α and E1β subunits (Thelen *et al.*, 1998). The results of N-terminal sequencing of purified potato mitochondrial PDC proteins resolved on SDS-PAGE produced two E1α subunits that were identical at the amino acid level and a putative E1β subunit (Millar *et al.*, 1998).

Although these studies have provided an overview of the general composition of the PDC from plant cells, there are still many inconsistencies with respect to M_r values, the exact function of individual polypeptides resolved on SDS-PAGE and the apparent presence of multiple isoforms of various subunits from purified PDC. Therefore improvements in the purification and analysis of organelle-specific complexes are still necessary in order to understand their structure, organisation and function precisely.

Subunit	Pea chl [1]	Pea mt [1]	Potato mt [2]	Maize mt [3]
E2	50	50	78/55	52/53
X	48	67	55 ?	
E3	52	56 (50)[4]	58/55	62/63
E1α		41	43/41	43
E1β			37	40

TABLE 1.3: Comparison of the Subunit M_r from Different Species of Plant PDC

Tentative identification of the M_r values ($x10^{-3}$) of PDC components from pea mitochondria and chloroplasts and from maize mitochondria. Identification relied on cross-reactivity with specific subunit antisera. Potato mitochondrial components were identified by N-terminal sequencing.

1, Taylor *et al.* (1992a); 2, Millar *et al.* (1998); 3, Thelen *et al.* (1998); 4, Conner *et al.* (1996a & b)

1. 2. ENZYMATIC COMPONENTS OF THE PYRUVATE DEHYDROGENASE COMPLEX

The pyruvate dehydrogenase complex is composed of multiple copies of three separate enzymes, each performing a specific function in the overall conversion of pyruvate to acetyl-CoA. Each enzyme is individually folded and assumes its own characteristic oligomeric structure before associating with its companions to form the multienzyme complex. *In vitro*, individual enzymes can be purified from the native mammalian and bacterial complexes and can spontaneously re-associate in high yield with the correct stoichiometry to form a fully functional PDC.

1. 2. 1. The Pyruvate Decarboxylase (E1) Enzyme

Pyruvate decarboxylase is the first enzyme to participate in the overall reaction mechanism; it also catalyses the rate determining step while ensuring the specificity of the complex for its substrate. Two types of E1 enzyme exist. In Gram negative bacteria (eg. *E. coli*) E1 is a homodimer with a subunit mass of approximately 100 kDa (de Kok *et al.*, 1998) where the E1 polypeptide chains are unsplit (Perham, 1991). In contrast, in Gram positive organisms (eg. *B. stearothermophilus*) and eukaryotes the E1 component consists of a heterotetramer, containing two α subunits and two β subunits, with subunit masses of approx. 41 kDa and 36 kDa, respectively (Patel and Roche, 1990).

Recent crystal structures are available for the *E. coli* PDC-E1 homodimer (Arjunan *et al.*, 2002), the PDC-E1β subunit from *Pyrobaculum aerophilum* (Kleiger *et al.*, 2001) and the heterotetrameric E1 components of BCDC from *P. putida* and human (Aevarsson *et al.*, 1999; 2000). The E1 homodimer is assembled from multidomain polypeptides. Each subunit has a N-terminal domain, a middle domain and a C-terminal domain. The N-terminal domain is the largest of the three, centred on a five stranded parallel β-sheet with connecting helices on either side and on top. The middle domain contains six helices surrounding a six stranded parallel β-sheet connected to the N-terminal by an extended loop. The C-terminal

domain is formed from a four stranded parallel and one antiparallel β-sheet, which is flanked by six helices with the last helix extending towards the C-terminus. Dimerisation into the homodimer involves interactions between residues from all three domains in each subunit.

No crystal structure for the human PDC-E1 tetramer has yet been determined but the structure of the E1 tetramer of BCDC has been resolved (Aevarsson *et al.*, 2000). The structure of the $\alpha_2\beta_2$ tetramer has four lobes each corresponding to one subunit that makes extensive contacts with the other three as seen in Figure 1.6. The N-terminal tails of the α subunit crossover and the C-terminal domains extend to hold onto the β subunit (Aevarsson *et al.*, 2000). The E1α subunit is composed of a large α/β domain with an extended N-terminal tail and a small helical C-terminal. The major domain contains a 14 stranded β-sheet surrounded by 13 α-helices (Aevarsson *et al.*, 2000). The E1β subunit is composed of two similar sized domains separated into the N- or C-terminal region. Both domains comprise a central β-sheet surrounded by several helices. The N-terminus has a six stranded parallel β-sheet flanked by seven helices while the C-terminus has a five stranded β-sheet surrounded by four helices. Both domains are connected by an extended loop (Aevarsson *et al.*, 1999).

A conserved motif was identified in all three E1β subunits (BCDC-E1β from *P. putida* and human and PDC-E1β from *Pyrobaculum aerophilium*) that promotes a helix-helix interaction between the β and α subunits (Kleiger *et al.*, 2001). Conserved interactions were also found between the two domains, involving hydrogen bonds between the C-terminal (Gln238) and the N-terminus (Asn157) and a salt bridge formed between Arg255 and Asp155. The residues refer to those in the *P. aerophilum* E1β subunit, although equivalent residues were found in all the E1β sequences aligned. The function of these interactions is thought to lock the domains into a specific orientation that will permit the N-terminus to interact with E1α and the C-terminus to interact with E2 (Kleiger *et al.*, 2001).

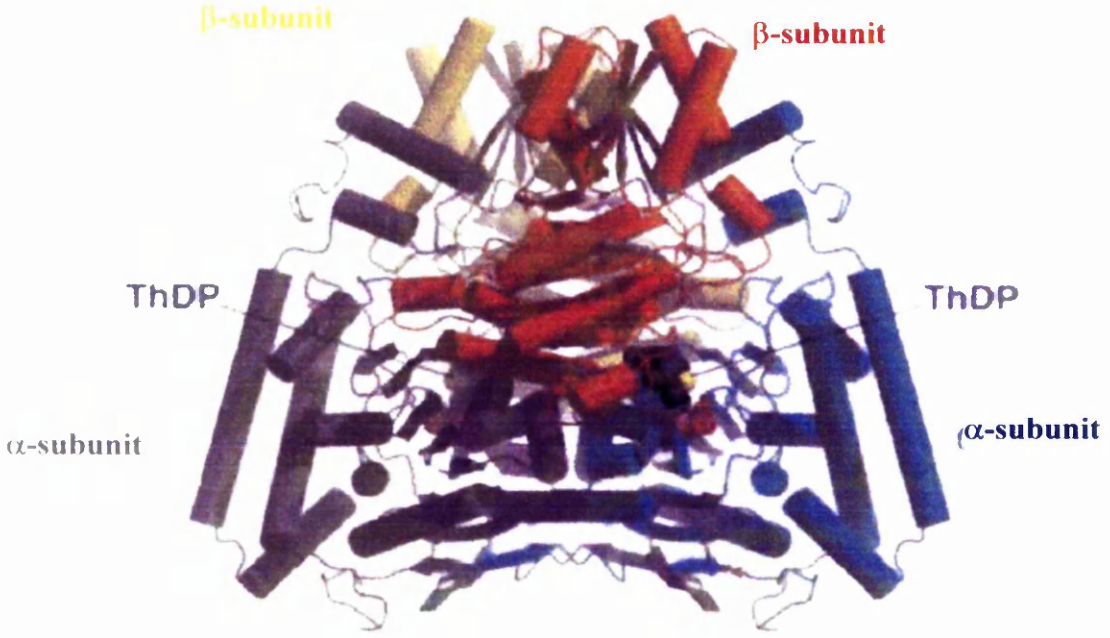


FIGURE 1.6: Ribbon Diagram for the Human BCDC-E1 Tetramer

The E1 tetramer is formed from the association of 2 α subunits and 2 β subunits. Both subunits are composed of α-helices and β-strands depicted as cylinders and arrows, respectively. The ThDP cofactors bind at the active sites located at the two interfaces.

(Taken from Aevarsson *et al.*, 2000)

The *E. coli* PDC-E1 homodimer appears similar to the *P. putida* BCDC-E1 tetramer with the equivalent domains as follows; PDC-E1 N-terminal corresponds to BCDC-E1α, the PDC-E1 middle domain with BCDC-E1β N-terminal domain and the PDC-E1 C-terminal domain with BCDC-E1β subunit C-terminal domain (Arjunan *et al.*, 2002).

Although the ThDP-dependent enzymes have low sequence homology they do share a conserved motif of approx. 30 residues (Dyda *et al.*, 1990). This is the ThDP binding site beginning with -GDG- and ending with -NN-, responsible for binding Mg^{2+} ions that anchor the ThDP via the pyrophosphate group. The ThDP cofactor is bound in the "V" conformation where the pyrimidine ring is close to the reactive C2 of the thiazolium ring (Fang *et al.*, 1998). The cofactor is found bound at the two active sites situated at the α-β interfaces of the tetramer and the interface of the homodimer. Comparison of ThDP enzymes has identified a catalytic glutamate responsible for promoting the formation of the active carbanion by interacting with the aminopyrimidine ring (Fang *et al.*, 1998). The catalytic glutamate has been identified in the *E. coli* PDC-E1 (Glu571) and human BCDC-E1 (Glu76-β').

The mammalian PDC-E1α subunit has three phosphorylation sites which are involved in the regulation of the complex. These sites are highly conserved; however site two in the *A. thaliana* mitochondrial E1α subunit is replaced by a threonine residue. The crystal structure of the human BCDC-E1 tetramer shows that the primary serine residue that is phosphorylated is close to the active site and suggests that the phosphate group prevents substrate and/ or cofactor binding through steric hindrance or electrostatic repulsion (Aevarsson *et al.*, 2000).

1. 2. 2. The Dihydrolipoamide Acetyltransferase (E2) Enzyme

The dihydrolipoamide acetyltransferase of the pyruvate dehydrogenase complex, and indeed the corresponding E2 enzymes of all the members of the 2-oxo acid dehydrogenase complex family, has three major roles in governing its overall

structure and function; (1) the E2 component forms the central 24 or 60meric symmetrical core to which multiple copies of E1 and E3 bind; (2) it houses the acyltransferase activity and (3) it is designed to allow the attached lipoic acid cofactor to interact sequentially with the active sites of all three constituent enzymes (Yeaman, 1989). The ability of the oligomeric E2 to fulfil these multiple roles is facilitated by its oligomeric structure and unique multidomain organisation (Figure 1.7). The modular domain structure of E2 was initially shown by proteolysis of the E2 core releasing the lipoyl domains and subunit binding domain but leaving the core structure intact, suggesting that these domains protrude from an inner region responsible for its self-assembly into higher M_r oligomers.

Further evidence for the multidomain nature of E2 came with the determination of the *E. coli* E2 protein sequence where three equivalent domains were found in the N-terminal region. These all housed a potential lipoate attachment site and were followed by another smaller domain involved in E1 and/ or E3 binding which preceded the largest domain at the C-terminal end of the polypeptide, the catalytic domain.

Elucidation of the multidomain structure of E2 was initially derived from the study of the E2 components of *E.coli* PDC and OGDC (Perham and Packman, 1989). *E. coli* E2 consists of three highly homologous domains, each about 80 amino acids long linked together by long peptide linker sequences of 20-30 residues although the number of lipoyl domains can vary from one to three depending on the species (Figure 1.7). The adjacent domain located C-terminal to the inner lipoyl domain is smaller (approx. 35-40 residues) and indeed is one of the smallest globular domains known (Mande *et al.*, 1996). The role of this region is to bind either E1 and E3 in the case of *B. stearothermophilus* PDC, or to bind E1 only (mammalian PDC) or E3 only (*E. coli* PDC). Linked to the subunit binding domain by another stretch of linker sequence is the large C-terminal region, which remains assembled after proteolysis and contains the catalytic site for the acetyltransferase activity in addition to being responsible for self-assembly. In the case of *E. coli* PDC it also contains an E1 binding site.

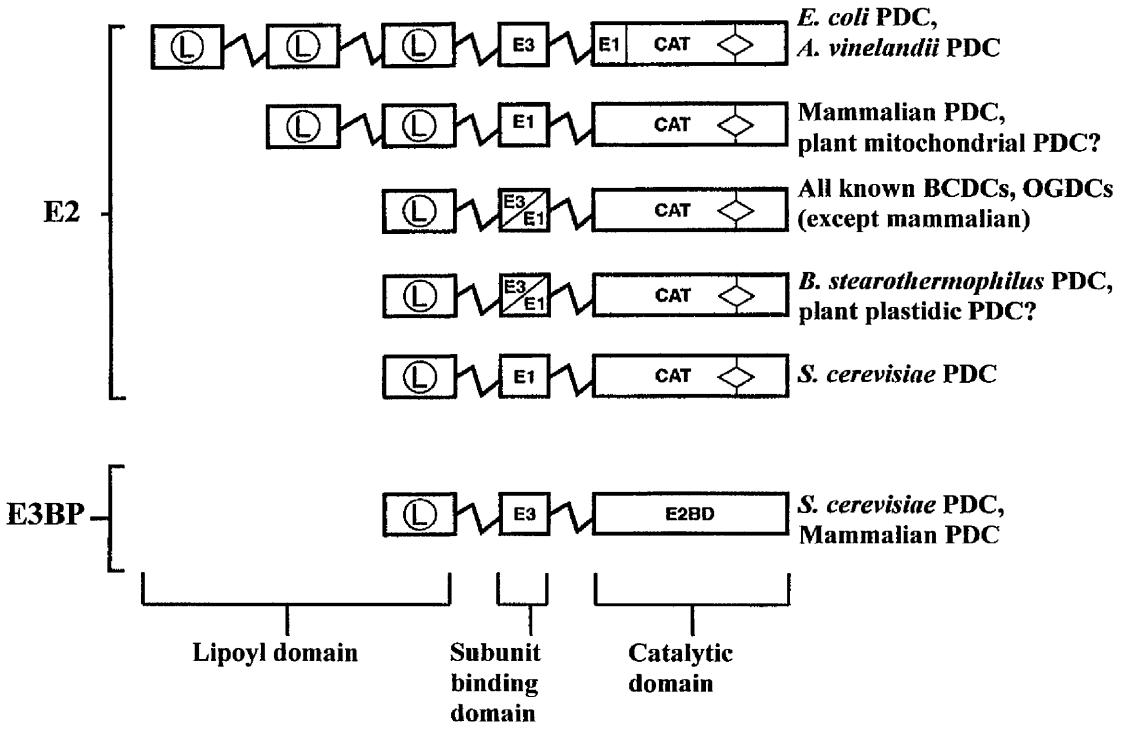


FIGURE 1.7: Domain Structure of the E2 Components from the 2-Oxo Acid Dehydrogenase Complexes From Various Sources

The E2 component from PDC, OGDC and BCDC has a multidomain structure comprising of independently folded domains linked together by small peptides sequences (zig-zag line). Located at the N-terminus are the lipoyl domains capable of covalent modification with lipoic acid. C-terminal to these domains is the subunit binding domain, which binds E1 or E3 or both depending on the source. Linked to the C-terminus via a linker sequence is the catalytic domain that houses the acyltransferase activity (active site, <>) and the oligomerisation information. A related PDC component (E3-binding protein, E3BP) found in yeast and mammals is also included. E3BP comprises a single lipoyl domain and subunit binding domain that only binds E3, both domains are linked by small peptide sequences. However, instead of a catalytic domain there is an E2-binding domain to allow association with the E2 oligomeric core.

1. 2. 2. 1. The Lipoyl Domain

As indicated previously, lipoyl domains are present singly or in tandem repeat at the N-terminus of the multidomain E2 enzyme with their number varying from one to three depending on the source (Figure 1.7). They each house a specific lipoylation site where the lipoic acid cofactor is covalently attached via an amide linkage to the N^6-amino group of a specific lysine residue.

The lipoyl group is reductively acylated in an efficient manner only when bound to the lipoyl domain. The lipoyl domain is recognised by the E1 component in the parent complex (Wallis and Perham, 1994). In *B. stearothermophilus* certain residues (11-15) form part of a surface loop close in space to the β-turn of the lipoyl-lysine that appear to be involved in domain recognition in that the lipoyl-lysine could no longer be reductively acylated by E1 in their absence. It was also reported that the N-terminal residues of this loop were involved in the recognition of E1 (Wallis *et al.*, 1996). Specificity is thus conferred by the residues on the lipoyl domain and the subunit binding domain. During catalysis by these complexes the protein-bound lipoamide acts as a "swinging arm" carrying activated acyl groups and reducing equivalents between the active sites of the E1, E2 and E3 components (Reed and Hackert, 1990).

All lipoyl domains have similar amino acid sequences and are of 80-100 amino acids in length suggesting that they have a common basic three dimensional structure. The overall structure of the lipoyl domain has been solved for *Bacillus stearothermophilus* PDC (Dardel *et al.*, 1993); *E. coli* PDC (Green *et al.*, 1995); *Azotobacter vinelandii* PDC (Berg *et al.*, 1997) and the inner domain of *Homo sapiens* PDC (*Howard et al.*, 1998) by the means of NMR. The structure of each lipoyl domain closely resembles the others and so the description that follows generally applies to every lipoyl domain. The 3D structure of the lipoyl domain is a flattened eight-stranded β-barrel, formed from two four stranded anti-parallel β-sheets (Wallis and Perham, 1994; Green *et al.*, 1995), folded around a core of hydrophobic residues (Dardel *et al.*, 1993).

The N and C-terminal amino acids are in close proximity on adjacent β-strands. Located at the opposite side of the domain is the β-turn which houses the lipoylatable lysine residue (Figure 1.8). The polypeptide backbone exhibits a 2-fold quasi-symmetry with certain strands almost superimposable (Dardel *et al.*, 1995). The location of the specific lysine residue that is subject to covalent modification is highly conserved and found at the tip of a type-I β-turn. Highly conserved aspartic acid and alanine residues flank the lysine residue (DKA). The precise position of the lysine residue is essential for recognition by lipoylating enzymes but the conserved aspartic acid and alanine are not (Wallis and Perham, 1994). However, they are important in the recognition of the lipoyl domain by the E1 subunit of the parental 2-oxo-acid dehydrogenase complex as mutations at these positions reduce the rate of reductive acylation (Wallis and Perham, 1994).

FIGURE 1.8: The N-terminal Lipoyl Domain from *A. vinelandii* PDC-E2

The lipoyl domain resembles a flattened β-barrel formed from 2 β-sheets constructed from four antiparallel β-strands (depicted as arrows). The cofactor lipoic acid (ball and stick) is attached to the specific lipoylatable lysine residue situated at the tip of the β-turn. The N and C-terminal ends (denoted N and C) are located at the opposite side of the barrel of the β-turn and in close proximity to each other.

(Taken from de Kok *et al.*, 1998)

1. 2. 2. 2. The Subunit Binding Domain

The next independently folded domain located adjacent to the lipoyl domains and attached via a linker region rich in alanine and proline is the subunit binding domain (SBD). It is one of the smallest functional domains known being 33 residues in length for the *B. stearothermophilus* PDC (Kalia *et al.*, 1993) or 51 amino acids in the *E. coli* OGDC (Robien *et al.*, 1992) and lacks disulphide bridges or a prosthetic group. They share a significant number of sequence similarities with comparable domains in other acyltransferase components. The function of this domain is essentially to bind the cognate E1 and/or E3 enzymes to the oligomeric core of the 2-oxo acid dehydrogenase complex. Their binding specificity is variable. Whereas the mammalian SBD binds only the E1 component, the bacterial subunit binding domain binds either the E3 component only (*E. coli*) or both the E1 and E3 components (*A. vinelandii*) (Patel and Roche, 1990).

The 3D-structures of the SBD from *B. stearothermophilus* PDC (Kalia *et al.*, 1993) and *E. coli* OGDC (Robien *et al.*, 1992) have been solved by NMR spectroscopy. Both consist of two short almost parallel α-helices, two short extended strands, a five residue helix-like turn and an irregular disordered loop (Robien *et al.*, 1992). These structures enclose a hydrophobic core that is highly conserved in all subunit binding domains of other E2 polypeptides suggesting that it plays a crucial role in determining and stabilising the polypeptide fold. The loop forms a finger-like structure consisting of two antiparallel extended regions (Robien *et al.*, 1992). Within this loop, glycine residues are conserved in nearly all E2-derived subunit binding sequences, for example, Gly37 *in E. coli* (Robien *et al.*, 1992) and a series of conserved glycine residues in *B. stearothermophilus* PDC-E2, Gly23, Gly25, Gly28 (Kalia *et al.*, 1993). The reason for the conservation of this residue is that the presence of a side chain would result in unfavourable steric hindrance with the rest of the protein structure or that they would interfere with the binding of E1 or E3 to the E2 core (Kalia *et al.*, 1993).

Binding studies involving the SBD of the E2 component of *B. stearothermophilus* PDC with the E3 component of the same complex showed that saturation occurred

at a molar ratio of 1:1. Thus the E3 homodimer binds only one subunit binding domain (Hipps *et al.,* 1994). This 1:1 stoichiometry may be explained by steric hindrance or an induced conformational change within the E3 homodimer preventing interaction of a second binding domain. Further structural evidence for the 1:1 molar ratio was obtained from the crystallisation of the *B. stearothermophilus* subunit binding domain with its E3 partner. The interaction of the subunit binding domain and the E3 component is close to the twofold axis of symmetry of the E3 homodimer (Figure 1.12). The entrance of the E3 active site remains unblocked with the binding to the subunit binding domain. The E3-subunit binding domain complex is stabilised by electrostatic interactions between the interface domain of both E3 monomers and the first helix of the subunit binding domain causing slight conformational changes in the loop structure of the subunit binding domain and in one monomer of the E3 dimer. A hypothetical model (Mande *et al.*, 1996) of two SBDs binding at the E3 dimer axis, showed a steric clash between the SBDs providing a structural basis for the observed 1:1 stoichiometry.

1. 2. 2. 3. The Catalytic Domain

The C-terminal domain in the multidomain E2 polypeptide is the catalytic domain consisting of about 250 residues (Mande *et al.*, 1996). It is responsible for the transacetylase activity in addition to the structural information for assembly into the large oligomeric cores observable by electron microscopy. The basic structure of the *A. vinelandii* trimeric subunit comprises three monomers that consist of two distinct parts: first, a main body with a globular shape folded into a six stranded β-sheet surrounded by five α-helices and a small two stranded β-sheet (Mattevi *et al.,* 1993a) as shown in Figure 1.9. The second remarkable feature of the E2 monomer is the N-terminal "elbow" shaped arm that does not interact with the rest of the monomer but extends over the neighbouring E2 monomer within the trimer structure contributing to the intertrimer interactions (Mattevi *et al.*,1992a). The topology of the E2 trimer is virtually identical to that of chloramphenicol acetyltransferase (CAT) with one noticeable feature being the significantly shorter

N-terminal helix in CAT as shown in the schematic diagram (Figure 1.10) of both the E2 trimer and CAT.

A common sequence element found in E2 polypeptides was the -H-X-X-X-D-G- motif known to be critically involved in the active site of CAT (Perham, 1991). Chloramphenicol acetyltransferase catalyses the O-acetylation of chloramphenicol in antibiotic resistant bacteria. Comparison of the *E. coli* E2 catalytic domain with that of the chloramphenicol acetyltransferase proved that they were well aligned homologs of each other. This region of high homology contained the CAT active site histidine residue that is well conserved in all E2 sequences. The histidine residue is though to act as a general base promoting the nucleophilic attack of CoA on S^8-acyldihydrolipoamide (Guest, 1987; Perham, 1991). The active site motif is situated near the C-terminal of the catalytic domain of all E2 polypeptides as it is also in the structurally related CAT.

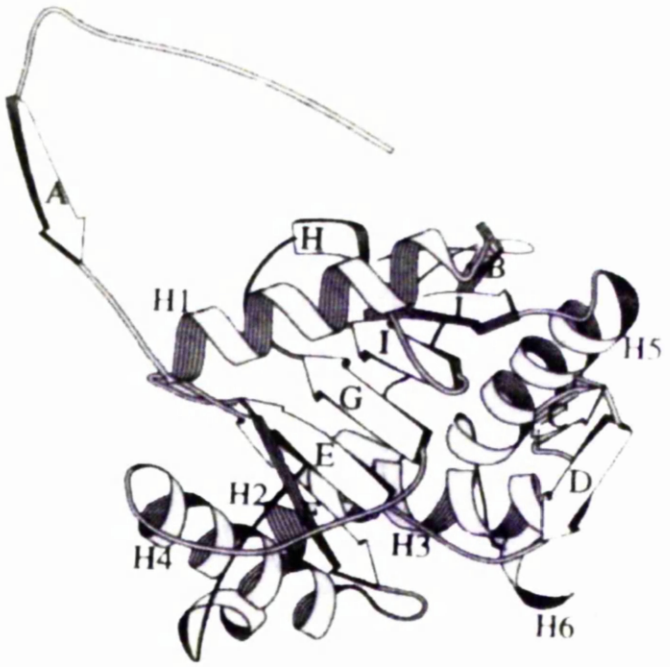


FIGURE 1.9: Ribbon Diagram of the Catalytic Domain of a Single Subunit of the PDC-E2 from *A. vinelandii*

The catalytic domain of the monomeric PDC-E2 component is comprised of β-sheets depicted as arrows, and α-helices. At the N-terminal region of the domain there is an extended "elbow" shaped arm, which extends away from the monomer body over one of the neighbouring E2 subunits of the trimer.

(Taken from Mattevi *et al.*, 1992a).

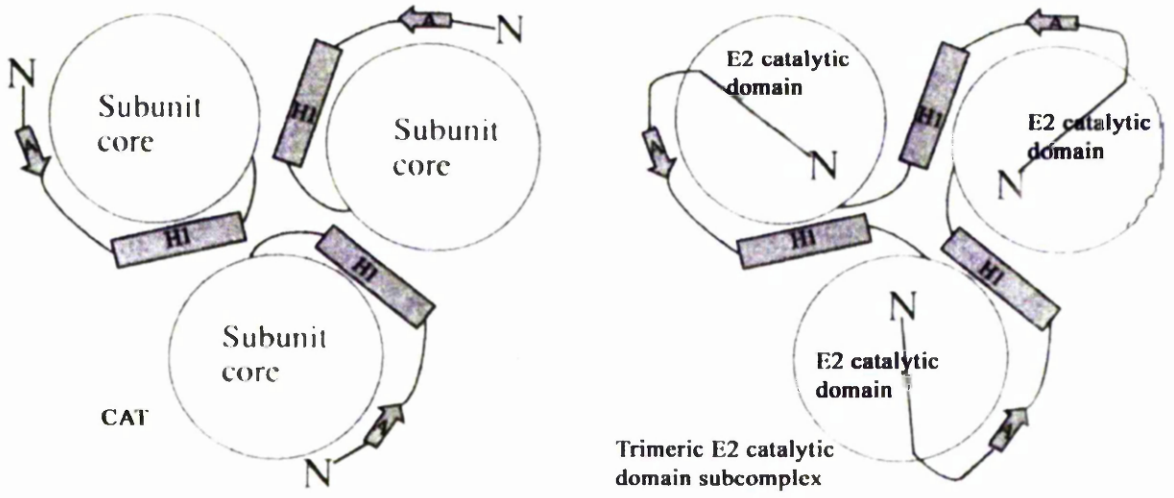


FIGURE 1.10: Schematic Representation of the Trimer Organisation of the Chloramphenicol Acetyltranseferase (CAT) and the Dihydrolipoamide Acetyltransferase (E2) Catalytic Domain

The C-terminal region of the E2 component (see Figure 1.7) not only houses the acetyltransferase activity but also contains information on the organisation of the multimeric cores. The PDC-E2 core is constructed from the catalytic domains of three E2 components. The topology of these trimers is virtually identical to that of the CAT enzymes. However, the N-terminal regions of the individual E2 catalytic domains interact with a second E2 within the trimer. Whereas the N-terminal region of each CAT subunit interacts with the same subunit.

(Taken from Mattevi *et al.*, 1992a)

Two distinctive polyhedral forms of E2 are detected by electron microscopy, the cube and the pentagonal dodecahedron. The cube is composed of 24 E2 monomers displaying overall octahedral symmetry while the pentagonal dodecahedron is composed of 60 E2 polypeptides organised into an oligomeric core with icosahedral symmetry. Schematic diagrams of both cores are presented in Figure 1.11. The underlying morphological unit of both core structures is the E2 trimer (Reed and Hackett, 1990). In the cubic core the E2 subunits are thought to assemble via trimeric intermediates around the eight vertices, while the E2 trimers in the 60meric structure constitute the twenty vertices of the core. The crystal structure of the *A. vinelandii* E2 cubic core has a hollow interior with large pores (3nm in diameter) in the faces of the cube. On trimer formation 25% of the monomer accessible surface becomes buried whereas in contrast the intertrimer interaction is weak and only 8% of the monomer accessible surface is buried (Mattevi *et al.*, 1992a). The trimers are more stable than the overall oligomeric core structure, explaining the tendency of the *A. vinelandii* E2 core to dissociate into its constituent trimers upon binding of the peripheral components (Hanemaaijer *et al.*, 1989).

The catalytic centre formed in the cubic core is 3nm long, located at the subunit interface and extends across the trimer. CoA enters the catalytic centre from inside the cube while the dihydrolipoamide prosthetic group enters from the outside. The distribution of these two cores in the 2-oxo acid dehydrogenase complex family is shown in Table 1.4.

24meric core	**60meric core**
OGDC (all sources)	Mammalian PDC
BCDC (all sources)	*S. cerevisiae* PDC
Gram negative bacterial PDC (*A. vinelandii, E. coli)*	Gram positive bacterial PDC (*B. stearothermophilus)*
	Fungal PDC
	Plant mitochondrial PDC (*Zea mays, S. tuberosum*)

TABLE 1.4: Distribution of the Multimeric E2 Core Structures of the 2-Oxo Acid Dehydrogenase Complexes

Yeaman, (1989); Reed and Hackett, (1990)

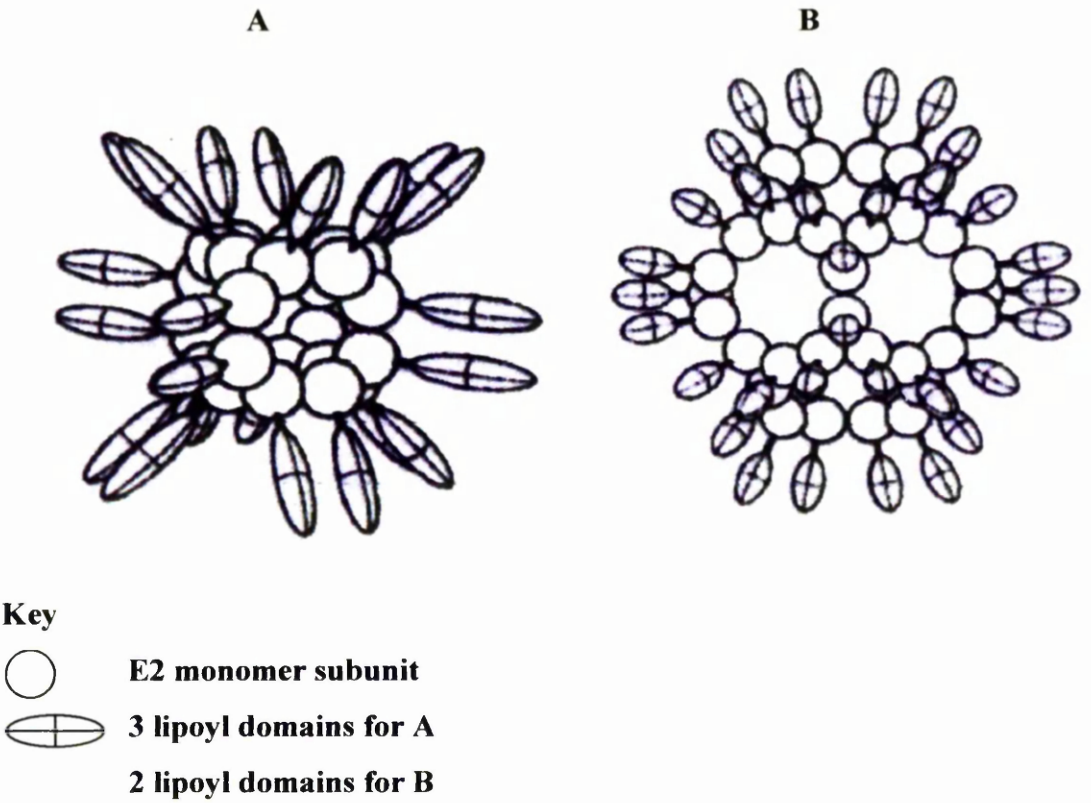


FIGURE 1.11: Quarternary Structure of the Multimeric E2 Core of PDC

Two oligomeric E2 cores have been identified; the 24meric structure of the E2 core form *E. coli* PDC (A) and the 60meric structure E2 core from mammalian PDC (B). The multimeric core is formed from trimers of E2 (as seen in Figure 1.10) assembled at the 8 vertices of the cubic core or the 20 vertices of the pentagonal dodecadral core.

(Taken from Reed *et al.*, 1989)

1. 2. 3. The Dihydrolipoamide Dehydrogenase (E3) Enzyme

Dihydrolipoamide dehydrogenase belongs to a family of flavoproteins known as pyridine nucleotide-disulphide oxidoreductases which includes structurally well characterised flavoproteins such as glutathione reductase (GR) (Carothers *et al.*, 1989). Human GR has four structurally distinct domains; FAD-binding, NADP-binding, central and interface domain. Structural comparison of GR to E3 of *E. coli* and many other E3s of various species reveals a high degree of homology in their core structures particularly in the residues around the FAD and NAD(P) binding sites. Conservation of the overall secondary structure of the human GR to human E3 were found to be very similar with the highest degree of sequence identity situated in the β-sheets, that form the internal core of the domains (Thekkumkara *et al.*, 1989). Crystal structures of E3 from *A. vinelandii,* (Mattevi *et al.*, 1991); *P. putida* (Mattevi *et al.*, 1992b); *P. fluorescens* (Mattevi *et al.*, 1993b); *B. stearothermophilus* (Mande *et al.*, 1996); *S. cerevisiae* (Toyoda *et al.*, 1998) and L-protein of GDC from *P. sativum* leaf mitochondria (Faure *et al.*, 2000) have been determined and show a high degree of structural conservation in their 3D organisation shown in Figure 1.12.

As suggested by their names each region has its specific function, the FAD-binding region has a conserved FAD binding motif (GxGxxG/AxxxG/A) but also found in this domain is the disulphide active site (CL/VNxGC). Similarly in the NAD^+-binding domain the conserved motif (GxGxIGxExxxVxxxxG) is found for the binding of the NAD^+ substrate (Lutziger and Oliver, 2001).

In its active state the E3 enzyme consists of two subunits which associate via hydrogen bonds between the amino acid side chains mainly from the interface domain and two long α-helices in the FAD-binding domain (Mattevi *et al.*, 1991). The catalytic site located at the interface domain contains the FAD prosthetic group and a redox-active disulphide bridge on monomer A and a histidine residue on monomer B (de Kok *et al.*, 1998).

FIGURE 1.12: The Complex Between the Subunit Binding Domain and the E3 Homodimer of *B. stearothermophilus*

The complex between the subunit binding domain (blue) and the E3 homodimer. Each monomer of the dimer consists of four domains, the FAD binding domain (yellow), the NAD binding domain (green), the central domain (red) and the interface domain (lilac). The FAD cofactor is indicated in black binding to its specific binding domain.

(Taken from de Kok *et al.*, 1998; Mande *et al.*, 1996)

In mammals dihydrolipoamide dehydrogenase is the common component of the 2-oxo acid dehydrogenase complexes (PDC, OGDC and BCDC as well the GDC, named L-protein). E3 is responsible for reoxidising the dihydrolipoamide group bound to E2 (or H-protein for GDC). Oxidation of dihydrolipoamide reduces the E3 enzyme, causing the opening of the disulphide bridge. A charge transfer complex is formed between the sulphydryl group of one cysteine residue and the FAD group (Mattevi *et al.*, 1991). In this reduced state the NAD^+ substrate can enter its binding site to receive the transfer of the hydride ion (Mattevi *et al.*, 1992b; 1993b).

Mammalian E3 exists as a homodimer comprising two subunits with a monomeric M_r of 51, 000. *E. coli* E3 also exists as a homodimer of two subunits with M_r of approx. 51, 000 and the *lpd* gene which encodes E3 is regulated by the PDC (*ace*E and *ace*F genes) operon and a secondary promotor of the *lpd* gene (Patel and Roche, 1990). Most of PDC E3 is encoded as a read-through transcript of the PDC operon while the transcription of the *lpd* gene supplies most of the E3 for OGDC (Carothers *et al.*, 1989).

Pseudomonas species were found to have three E3 genes (*lpd*-val, *lpd*-glc and *lpd*-3) encoding three distinct lipoamide dehydrogenases differing in M_r values. (Carothers *et al.*, 1989; Burns *et al.*, 1989b). LPD-Glc was the only E3 found in *P. putida* when cultured in glucose synthetic medium, while both LPD-Glc and LPD-Val were found in cells grown in valine rich medium. *P. putida* LPD-Val (M_r 49, 000) is specific for BCDC while the LPD-Glc (M_r 56, 000) is the E3 component of OGDC, and probably also PDC and the L-protein of the GDC (Burns *et al.*, 1989b). *P. aeruginosa* also contains these two forms of E3, LPD-Val (M_r 50, 000) and LPD-Glc (M_r 54, 000) (Carothers *et al.*, 1989). LPD-Val and LPD-Glc are specific for their functions and not interchangeable (Burns *et al.*, 1989a). Isolation of a third E3 enzyme from a mutant *P. putida* strain, affected in the *lpd*G gene encoding the LPD-Glc lead to the identification of LPD-3, a distinct E3 enzyme (M_r 53, 000) similar to LPD-Glc but not LPD-Val at the amino acid level. The function of this third E3 is unclear but purified LPD-3 restored PDC activity and

OGDC activity, albeit less efficiently than LPD-Glc with regard to OGDC activity (Burns *et al.*, 1989b).

The *Pseudomonas* species are not the only species to have multiple E3s. Recently plants have also been found to have multiple forms of this enzyme. Mitochondrial E3 from *Pisum sativum* is encoded by a single gene and was identified as the E3 component of PDC and OGDC as well as the L-protein of GDC (Bourguignon *et al.*, 1996). Comparative analysis of partially purified mitochondrial and chloroplastic PDC from *P. sativum* using antibodies to *Brassica oleracea* mitochondrial PDC, reported that both complexes have a common protein (M_r 58, 100) corresponding to the E3 component (Camp and Randall, 1985). However, using antibodies to mammalian and yeast PDC E3, two distinct organelle specific E3s were found in pea, suggesting the existence of separate chloroplastic and the mitochondrial isoform, the chloroplastic form exhibiting a lower molecular mass than the mitochondrial E3 (Taylor *et al.*, 1992). These specific forms were then subsequently purified as distinct enzymes with different N-terminal sequences (Conner *et al.*, 1996a) and shown to have distinct enzymatic properties.

Complex specific E3 isoforms have also been identified to exist in mitochondria from tubers of *Solanum tuberosum*. Three E3 activities were purified by Mono-Q ion exchange chromatography and gave distinct bands on SDS-PAGE composed of two subunits (α and β). These two polypeptides exhibit total identity in their N-terminal sequences but differ in their M_r values (α, 58, 000 and β, 56, 000). Dimerisation of these subunits produce the purified three isoforms of E3 (α_2, $\alpha\beta$, and β_2). Reconstitution of E3 deficient PDC and OGDC with the purified isoforms revealed that $\alpha\beta$ and β_2 could restore the PDC and OGDC activity but α_2 was inactive in this regard (Cook *et al.*, 1996). These results support the idea of complex specific E3s. Tissue specific expression was also observed with the α_2 isoform dominant in leaves, suggesting a possible role as the L-protein of GDC while the β_2 isoform was most abundant in tubers.

Recently, a surprising finding has been the identification of four E3 genes in *Arabidopsis thaliana* found due to the determination of the *A. thaliana* genome sequence. Two forms were identified as chloroplastic and two mitochondrial (Lutziger and Oliver, 2000; 2001). Where the plastid E3s are expressed in parallel in different tissues, the mitochondrial E3s are not. One gene, *mtLPD1* is strongly light induced while the *mtLPD2* gene is less light dependent. High levels of *mtLPD1* RNA are found in leaves while the highest levels of *mtLPD2* RNA are found in the roots. This difference of expression may shed light on the roles of these two E3s. Thus, the product of the gene *mtLPD1* may be involved in the glycine decarboxylase complex where the highest levels are found in the leaves, and the product of the gene *mtPLD2* may be the E3 component of the α-keto acid dehydrogenases (Lutziger and Oliver, 2001).

In addition to the above several types of dihydrolipoamide dehydrogenase have been identified in various organisms that do not contain PDC. In halophilic archaebacteria, a dihydrolipoamide dehydrogenase (M_r of 112, 000-120, 000) was discovered that was specific for dihydrolipoamide in the presence of NAD^+. No 2-oxo acid dehydrogenase complexes are present and pyruvate and 2-oxoglutarate are converted to acyl-CoA by ferredoxin oxidoreductases and so the role of this E3 is unclear (Danson *et al.*, 1984). A second E3 was found in the blood stream form of *Trypanosoma brucei*. *T. brucei* contains only a single mitochondrion but lacks all 2-oxo acid dehydrogenase complexes; however, an E3 activity was found associated with the cytosolic side of the plasma membrane. It was specific for NAD^+ and dihydrolipoamide (Danson *et al.*, 1987). Lipoic acid (either protein bound, free or modified was unclear) was discovered in the bloodstream form of *T. brucei*. This was the first time that lipoic acid was found in a eukaryote that lacks 2-oxo acid dehydrogenase complexes albeit at low concentrations (Jackman *et al.*, 1990).

The most unusual E3 to date was discovered in *Neisseria meningitidis* where it was named p64k (M_r 64, 000). It is present on the cell surface and is involved in the immunological response to this organism. It contains a FAD cofactor and is functionally active as a homodimer, exhibiting oxidoreductase activity on

dihydrolipoamide in the presence of NAD^+. A unique feature of this E3 is its N-terminal 111 residues, which are homologous to E2 lipoyl domains. Its crystal structure is similar to E3 with the FAD cofactor bound in a similar manner. The major difference between the FAD binding domain of this E3 and others is the extension of two strands at the surface of the protein with the insertion of a loop of approx. 10 residues. This contains the antigenic sites recognised by the host (Li de la Sierra *et al.*, 1997).

1. 2. 4. The E3-Binding Protein (Protein-X)

In PDC from mammals and *Saccharomyces cerevisiae* (Kreze and Ronft, 1981), an additional polypeptide has been found tightly associated with the E2 core. This E3-binding protein (E3-BP) or protein-X has a subunit M_r value of approx. 50, 000 and is antigenically distinct from other PDC subunits (de Marcucci *et al.*, 1986). N-terminal sequencing of the E3-binding protein reveals significant homology to E2 which suggested that this protein was a functional variant or an isoenzyme of E2 (Neagle *et al.*, 1989) but contained a single lipoyl domain. The lipoyl domain is also covalently modified by the addition of lipoic acid to a specific lysine residue. Similar to E2 E3-BP can be reductively acetylated in the presence of pyruvate or acetyl-CoA and NADH. As yet the possible role of this polypeptide in the enzymatic reaction is not clearly understood. The presence of a single lipoyl domain accounts for the small M_r in comparison to mammalian E2s (70, 000) and was identified by tryptic cleavage of ^{14}C-labelled acetylated lipoyl domains producing fragments with M_r values of 34, 000 and 15, 000 for E2 and E3-BP, respectively.

The role of the E3-BP, hence its name, was determined by selective proteolysis of E3-BP lipoyl domains from the E2/E3-BP subcomplex leaving a truncated core bound E3-binding protein. Removal of the lipoyl domain was accompanied by the loss of activity which could be restored adding an excess of the E3 component. Comparison of the truncated polypeptide to that of other subunit-binding domains of PDC from *B. stearothermophilus* indicated several conserved residues important in structural organisation and the binding of E3, thereby suggesting the presence of

an equivalent domain. It was found that the cleavage site in E3-BP was located at the boundary of a putative E3-SBD suggesting that minor disruption in this domain may explain the lower E3 binding affinity and loss of activity (Sanderson *et al.*, 1996). The multidomain structure (see Figure 1.6) of the E3-BP comprises a N-terminal lipoyl domain, an E3-binding domain and a C-terminal domain equivalent to that of E2. The C-terminal domain may include assembly information allowing association with E2 during core assembly.

1. 3. THE QUARTERNARY STRUCTURE OF PDC

PDC is a multienzyme complex formed from the non-covalent association of multiple copies of three enzymes. The centre of the complex is formed from the E2 oligomeric core as already discussed, which either forms a cubic or pentagonal dodecahedral structure. PDC from *E. coli* (M_r approx. 4.6 million) is arranged around a cubic core that can bind 12 E1 homodimers and 6 E3 dimers (Reed and Hackert, 1990; Patel and Roche, 1990). The PDC from *E. coli* shows that the E3 component only interacts with the subunit binding domain found on E2 and the E1 component interacts with the C-terminal region of the catalytic domain. The difference in binding sites may direct the assembly by positioning the E1 along the vertices, leaving the faces of the cube for the binding of E3. The E3 dimer has a potential of two binding sites for association with the subunit binding domain at its dimer interface. However, as mentioned before (see section 1. 2. 2. 2), these sites are close together so that occupancy of one binding domain excludes the binding to the second, therefore inducing asymmetry in the complex (de Kok *et al.*, 1998).

The *A. vinelandii* PDC is also built around a 24meric E2 core, where unlike other PDCs, the *A. vinelandii* E2 core dissociates into smaller suboligomers upon the binding of E1 and E3 components. Refined crystal structures have identification the most likely morphological subunit as being a trimer with 3-fold rotational axis (Mattevi *et al.*, 1992a). In view of the tight associations between the monomers in the trimer, the dissociated suboligomers are probably trimers (Mattevi *et al.*, 1992a).

Mammalian PDC is assembled around a 60meric pentagonal dodecahedral core that forms the framework for the binding of E1, E3 and E3-binding protein. Bovine heart PDC (M_r approx. 10 million) contains 20-30 E1 tetramers and 6 E3 homodimers associated with the E2 core with the E1 components located near the 30 edges of the core and the E3 components binding at the faces. The quarternary structure of the mammalian PDC is made more complicated by the involvement of the E3-binding protein which is responsible for the binding of the E3 dimers to the E2 core. Twelve E3-binding protein subunits are included with the complex overall structure (Sanderson *et al.*, 1996). This stoichiometry of 30 E1 tetramers and 6 E3 dimers is also seen in the PDC of *B. stearothermophilus* (M_r approx. 9 million), although no E3-binding protein is involved and E1 or E3 binds via the subunit binding domain of E2 in a mutually exclusive manner (Mande *et al.*, 1996).

The quarternary structures of the plant complexes are unknown, but the structure of mitochondrial PDC-E2 cores of at least two species are known. Both mitochondrial PDC-E2s from maize (*Z. mays*) and from potato (*S. tuberosum*) show the presence of a 60meric structure. As the plant mitochondrial core is similar to the mammalian core it may be assumed that the overall organisation of the complex is similar also. However, the organisation of the chloroplast complex still remains elusive.

1. 4. REGULATION OF THE PLANT PYRUVATE DEHYDROGENASE COMPLEX

1. 4. 1. Regulation of the Mitochondrial PDC by Product Inhibition

Plant mitochondrial PDC, like its mammalian counterpart, is inhibited by the products of the overall enzymic reaction i.e. NADH and acetyl-CoA which exhibit competitive inhibition with their respective substrate (NAD^+ and CoA). The plant mitochondrial complex is more sensitive to the NADH/NAD^+ ratio than the acetyl-CoA/CoA ratio as shown by their kinetic values. The K_i of NADH (34μM) is 3.5 fold lower than the K_m of NAD^+ (125μM) while the K_i of acetyl-CoA (13μM) is twice that of the K_m of CoA (7μM) (Randall *et al.*, 1977).

1. 4. 2. Regulation of the Mitochondrial PDC by Reversible Phosphorylation

Plant mitochondrial PDC also resembles mammalian PDC in that it can be reversibly phosphorylated, by a closely associated PDH-kinase and PDH-P-phosphatase. Investigation into these components has been hampered by the loss or inactivation of these enzymes during the purification of the PDC from plant sources. The PDH-kinase is tightly bound to the complexes while the PDH-P-phosphatase is loosely bound (Miernyk and Randall, 1987). Mammalian PDC is inactivated by the phosphorylation of serine residues located at three sites on the E1α component. These serine residues at sites one and two are conserved among all eukaryotic E1α subunits; however, in *A. thaliana*, the serine is replaced with a threonine at site two (Luethy *et al.*, 1996).

Plant mitochondrial PDC from pea and broccoli have been shown to be phosphorylated in an ATP-dependent (100-200μM) manner with concomitant loss of activity (Miernyk & Randall, 1987). This deactivation required Mg^{2+} as shown by EDTA inhibition of deactivation although only low concentrations of Mg^{2+} are required (Randall *et al.*, 1981). Activity of the phosphorylated PDC was restored by incubation with $MgCl_2$ (10-20mM) (Miernyk & Randall, 1987), Mn^{2+} and Co^{2+} were also good activators of the PDH-P-phosphatase but to a lower extent than Mg^{2+}. Mono- or polyvalent cations or Ca^{2+} were unable to activate PDH-P-phosphatase; neither did orthophosphate or polyamines (Miernyk and Randall, 1987). Ca^{2+} ions act as an activator of mammalian PDH-P-phosphatase whereas in plant mitochondria it appears to act as an inhibitor (Budde *et al.*, 1988).

In vitro effectors of the plant kinase activity include ADP, pyruvate, acetyl-CoA, NADH, citrate, 2-oxoglutarate and monovalent cations as shown in Figure 1.13. In addition K^+ and NH_4^+, both stimulate the PDH-kinase and are therefore responsible for the inactivation of PDC. K^+ (0.5-4.0mM) apparently lowers the K_m (ATP) of the PDH-kinase, while NH_4^+ are more effective at activating the PDH-kinase with levels of 20-80μM required (Schuller and Randall, 1989), and is discussed further in the next section.

1. 4. 3. Light Effects on the Mitochondrial PDC Activity

Light-dependent inactivation of the mitochondrial PDC *in vivo* can be reversed by placing the plant in dark conditions (Gemel and Randall, 1992). PDC is inactivated when mitochondria are oxidising other carbon sources like glycine or succinate. With the oxidation of succinate the inactivation of PDC may be attributed to the increase in mitochondrial ATP levels whereas with glycine conversion to serine, NADH is produced to drive oxidative phosphorylation (increasing ATP levels) along with an increase in NH_4^+. During photorespiration and photosynthesis, NH_4^+ levels increase to approx. 3mM (Schuller *et al.*, 1989) within the mitochondrial matrix. As previously mentioned, NH_4^+ can enhance inactivation of PDC by stimulating PDH-kinase, hence regulating carbon flux into the citric acid cycle during photosynthesis (Gemel and Randall, 1992). As a result pyruvate levels will rise consequently re-activating the PDC via increased PDH-P-phosphatase activity and so maintain a balance between substrate and product levels.

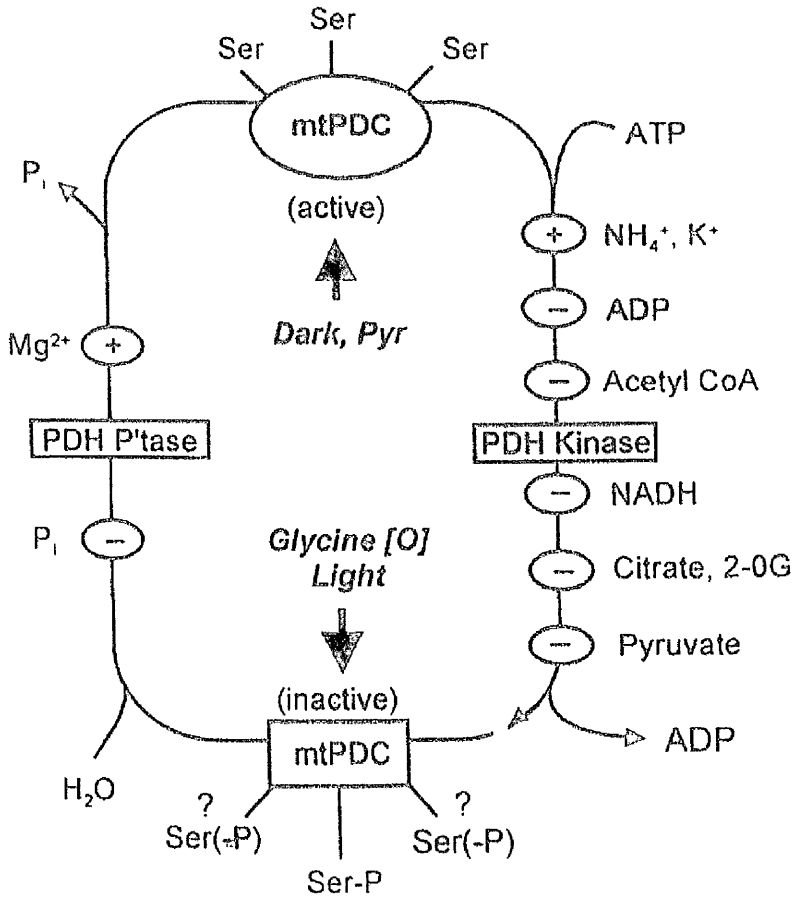


FIGURE 1.13: Model Illustrating Effectors of the Reversible Phosphorylation of Pea Leaf Mitochondrial PDC

Diagram showing the various products, substrates and intermediates that effect the PDH-kinase and PDH-P-phosphatase associated with pea mitochondrial PDC. (Taken from Luethy *et al.*, 1996)

1. 4. 4. Regulation of the Chloroplast PDC by Product Inhibition

Pea chloroplastic PDC shares similar kinetic properties with bacterial and mitochondrial PDC. Competitive inhibition is exhibited for acetyl-CoA versus CoA and NADH versus NAD^+. However uncompetitive inhibition is observed for pyruvate versus NADH or acetyl-CoA, with NAD^+ and CoA versus acetyl-CoA and NADH respectively, exhibiting noncompetitive inhibition. The main regulatory properties are enforced by NADH and acetyl-CoA against their respective substrates. Plastid PDC, like its mitochondrial equivalent is more sensitive to the $NADH/NAD^+$ ratio than the acetyl-CoA/CoA ratio (Camp *et al.*, 1988).

1. 4. 5. Regulation of the Chloroplast PDC by Phosphorylation and Metabolites

Plastid PDC resembles bacterial PDC and is not regulated by reversible phosphorylation (Randall et al., 1989). The activity of pea chloroplast *in vitro*, demonstrated insensitivity towards intermediates of glycolysis. Regulation of plastid PDC was influenced by the products of fatty acid synthesis with palmitate stimulating and oleate inhibiting overall PDC activity. The acyl-CoAs elicited less pronounced effects than their corresponding free fatty acids (Camp *et al.*, 1988).

1. 4. 6. Light Regulation of the Chloroplast PDC

Light regulation of plastid PDC appear to occur indirectly via changes in the pH and ionic composition of chloroplast stroma. Pea chloroplast PDC has a pH optimum of 7.8-8.2 and a higher Mg^{2+} concentration (1-5mM) requirement compared to its mitochondrial counterpart. An alkaline pH optimum is a key regulatory mechanism for PDC during photosynthesis when the chloroplast stroma becomes more alkaline (pH 7-8) and the Mg^{2+} concentration increases from 1-3mM (Williams and Randall, 1979). Changes in the physical environment of the chloroplast PDC correlate well with the determined optimum conditions.

Pea chloroplast PDC is relatively insensitive towards inhibition by NADPH and ATP, both of which are present at high levels within the chloroplast during photosynthesis. The requirement of an alkaline pH optimum and high Mg^{2+} concentrations are all consistent with active chloroplast PDC in times of illumination when demand for acetyl-CoA and NADH for fatty acid synthesis is high (Camp *et al.*, 1988).

1. 5. AIMS

The aim of the work described in this thesis was to clone various plant pyruvate dehydrogenase complex components and express them in *E. coli* with the intention of studying structural and functional aspects of the, as yet poorly characterised, plant complexes.

Cloning these plant enzymes into an appropriate expression vector has the potential of producing a readily available and sustainable source of recombinant protein with yields in the range of 5-30mg per litre of culture. Incorporation of expression tags, such as GST-fusion protein and His-tag, during the expression of these PDC components permits specific and efficient purification of these recombinant proteins from crude bacterial lysates.

Purification of these enzymes in high yields to near homogeneity will permit investigation into the degree of lipoylation of both the mitochondrial and chloroplastic E2 and allow initial studies into characterising the oligomeric core structure. Measuring the potentially large oligomeric core structures of these E2 will involve the use of an analytical ultracentrifugation to determine the M_r of the protein species present in the purified E2 sample.

The importance of the dimeric structure of E3 will also be investigated by first determining if the dimer structure is assembled by these recombinant E3s and what if any influence does the N-terminal mitochondrial targeting sequence has on the activity, folding and stability of the E3 structure. This will include the use of the spectroscopic properties of the proteins as measured by CD and fluorescence.

The long term aim is to reconstitute the complete complex producing a model mitochondrial and chloroplastic PDC *in vitro* for detailed structure function analysis and will also allow investigation of the regulatory properties of these organelle-specific complexes.

CHAPTER 2

MATERIALS AND METHODS

2. MATERIALS AND METHODS

2. 1. MATERIALS

2. 1. 1. Bacterial Strains

***Escherichia coli* DH5α:** a recombination deficient strain for the propagation of plasmid DNA. This is a high transformation efficiency strain producing good yields of plasmid.

Genotype: *sup*E44 Δ*lac*U169 (ϕ80*lac*ZΔM15) *hsd*R17 *rec*A1 *end*A1 *byr*A96 *thi*1 *rel*A1

***Escherichia coli* BL21 (DE3) pLysS:** employed for the high level expression of genes cloned into expression vectors under the control of the bacteriophage T7 promoter.

Genotype: F⁻*omp*T *hsd*SB ($r^-_B m^-_B$) *gal dcm* (DE3) pLysS.

2. 1. 2. Bacterial Media

All strains of *E. coli* were grown in Luria Broth (10g Bacto-tryptone, 5g Bacto yeast extract and 10g NaCl per litre, pH 7.5). LB plates were made by adding 7.5g Bacto Agar to 500ml LB. All media were autoclaved before use and where necessary, both the media and the plates were supplemented with ampicillin at 50μg/ml, or kanamycin at 30μg/ml, or both. Sambrook *et al.*, (1989).

2. 1. 3. Chemicals

All chemicals used were of the highest grade available commercially. Specialised chemicals like ampicillin, kanamycin, glutaraldehyde, NaF, glutathione (reduced form) and lipoic acid were all from Sigma. Glutathione Sepharose 4B was supplied from Amersham Pharmacia Biotech. Imidazole and $ZnCl_2$ were bought from MERCK, BDH. Isopropyl β-thiogalactopyranoside (IPTG) and dithiothreitol (DTT) were obtained from Melford Laboratories Ltd, Suffolk. Bradfords reagent

(BIO-RAD protein assay reagent) was purchased from BIO-RAD. Pre-made 29:1(v/v) acrylamide:bisacrylamide solution was supplied by Severn Biotech. Ltd.

2. 1. 4. Enzymes

DNA restriction enzymes and DNA T4 ligase were supplied by Boehringer Mannheim Ltd. *Taq* polymerase and calf intestinal alkaline phosphatase were supplied by Promega Ltd. Cloned *Pfu* polymerase was supplied by Stratagene and Promega.

2. 1. 5. Molecular Size and Weight Markers

For the determination of DNA product size, a 1Kb DNA Ladder (Promega) was used and for the quantification of DNA products, a Low DNA Mass Ladder (Life Technologies) was used. The relative molecular mass (subunit M_r) of proteins separated on SDS-PAGE was determined by comparison with low M_r markers (Amersham Pharmacia Biotech). Molecular mass protein standards for size exclusion chromatography were purchased from Sigma.

2. 1. 6. Oligonucleotides

Oligonucleotides were specifically designed as required and were synthesised by Genosys Biotechnologies Ltd, Operon Technologies Inc. and VH Bio Ltd.

The coloured sequences indicate the incorporated restriction sites for BamHI, XhoI and EcoRI.

a) *P. sativum* chloroplast PDC E2 forward primer

5'-CTTCAAGGATCCAGCCAAGATCCGCGAAATCTTC-3'

b) *P. sativum* chloroplast PDC E2 reverse primer

5'-TGTGAAGGATCCCTATAATGTTAGGATTTCTGG-3'

c) *P. sativum* chloroplast PDC E2 didomain forward primer

5'-CTTCAAAGGATCCGCCAAGATCCGCGAAATCTTC-3'

d) *P. sativum* chloroplast PDC E2 didomain reverse primer

5'-TTGTGAGGATCCCTATACTCGAACCAGGCAACG-3'

e) *S. tuberosum* mitochondrial PDC E3β isoform froward primer

5'-GCTTCAGGATCCCGATGAGAACGACG-3'

f) *S. tuberosum* mitochondrial PDC E3β isoform reverse primer

5'-TATAACGGATCCCTACATGTGGATGGGCTTG-3'

g) *A. thaliana* mitochondrial PDC E2 forward primer

5'-GCAAAACTTTCTCGAGACAGGACCCATATC-3'

h) *A. thaliana* mitochondrial PDC E2 reverse primer

5'-GTTGGACTCGAGTCACAGAAGAAGTCTTCG-3'

i) *A. thaliana* mitochondrial PDC E2 tridomain forward primer

5'-GCAAAACTTTGGATCCACAGGACCCATATC-3'

j) *A. thaliana* mitochondrial PDC E2 tridomain reverse primer

5'-CTTCATACTAGAATTCTGACTGTGTCACAGAAGGC-3'

k) *P. sativum* mitochondrial GDC L-protein precursor forward primer

5'-GATCGAATGGATCCAATGGCTATGGCGAACTTGG-3'

l) *P.sativum* mitochondrial GDC L-protein mature forward primer

5'-GGTCAAGGGATCCCGCCTCCGGATCTG-3'

m) *P. sativum* mitochondrial GDC L-protein reverse primer

5'-CTGAATGGATCCTTAACACACAAAGTATC-3'

2. 1. 7. Plasmid Vectors

pGEX-2T: for the expression of glutathione S-transferase fusion recombinant proteins; the GST-fusion protein is added to the N-terminal end of the cloned protein (ampicillin resistant). (Pharmacia, Appendix 1.)

pET-14b: for the expression of His-tagged recombinant proteins; the six histidine residues are added to the N-terminus of the cloned protein (ampicillin resistant). (Novagen, Appendix 2.)

pET-28b: for the expression of His-tagged recombinant proteins; the six histidine residues are linked to the N-terminus of the cloned protein (kanamycin resistant). (Novagen, Appendix 3.)

pCR®T7/ NT-TOPO®: for the expression of His-tagged recombinant proteins, adding six histidine residues to the N-terminal region of the cloned protein (ampicillin resistant). (Invitrogen, Appendix 4.)

pDCE2: (kind gift from Dr. Steve Rawsthorne, John Innes Centre). This vector contains the cDNA encoding the *Arabidopsis thaliana* mitochondrial PDC-E2.

pGDLE E3: (kind gift from Dr. Steve Rawsthorne, John Innes Centre). This vector contains the cDNA encoding the *Pisum sativum* mitochondrial GDC L-protein (PDC-E3).

pB3.1: this plasmid contains the cDNA sequence encoding the *Solanum tuberosum* mitochondrial PDC-E3β isoform.

pBE2: (kind gift from Prof. Jurgen Söll, University of Kiel, Germany) contains the cDNA encoding the *Pisum sativum* chloroplast E2 protein, cloned into pBluescript II SK^+

2. 2. METHODS

2. 2. 1. Polymerase Chain Reaction (PCR) of the PDC Components

PCR reactions were carried out in a total volume of 50µl containing 50ng of template DNA, 0.25mM dATP, dTTP, dCTP and dGTP (Boehringer Mannheim), Required primers (100ng each) were added to 1x cloned *Pfu* polymerase reaction buffer (supplied) containing 1.5mM $MgCl_2$. Each PCR reaction used 2.5U of cloned *Pfu* polymerase. PCR was carried out in a PTC-100™ thermocycler (Genetic Research Instrumentation Ltd) in 30 cycles of 45s at 95°C, 45s at various annealing temperatures (see TABLE 2. 1) and either 2 or 4min at 72°C. This was preceded by an initial denaturation step (45s at 95°C) and finished with a 10min extension step (72°C). Annealing temperatures were determined by trail and error. PCR products were viewed by agarose gel electrophoresis (2% (w/v) agarose) and stored at -20°C before and after agarose gel purification.

Target PCR product	**Annealing temperature (°C)**
P. sativum chloroplast PDC E2	60°C
P. sativum chloroplast PDC E2 Didomain	70°C
Potato Mitochondrial PDC E3	65°C
A. thaliana Mitochondrial PDC E2	60°C
A. thaliana Mitochondrial PDC E2 Tridomain	65°C
P. sativum Mitohchondria GDC L-protein	50°C

TABLE 2. 1: Annealing Temperatures for the PCR of PDC Components

2. 2. 2. Agarose Gel Electrophoresis

Plasmid DNA and PCR products were electrophoresed using 1%-2% (w/v) agarose gels with a 1Kb DNA Ladder. DNA was loaded with 6x loading buffer (0.25% (w/v) Bromophenol Blue, 0.25% (w/v) Xylene Cyanol FF, 15% (w/v)

Ficoll). Gels were run at 100 volts in 1x TAE buffer (50x, 2M Tris-HCl, pH 7.6, 50mM EDTA, 5.7% (v/v) acetic acid). Protocol was adapted from Sambrook *et al.*, (1989). Agarose gels were stained for 30min with ethidium bromide and DNA samples were viewed under UV light and photographed using E. A. S. Y imaging software.

2. 2. 3. DNA Isolation from Agarose

PCR products and digested plasmid DNA was subjected to electrophoresis as above, then extracted and purified from the agarose gel using the QIAquick Gel Extraction Kit (Qiagen) as instructed by the manufacturer. Digested PCR products were purified from the reaction mixture using this kit but without gel electrophoresis. The purity and yield of DNA (5µl) was then viewed by agarose gel electrophoresis.

2. 2. 4. Restriction Digestion of Plasmid DNA

A. Restriction Digestion of pET-14b

20µl of plasmid DNA was digested in a total volume of 40µl with BamHI (40U) or XhoI (40U) in the appropriate buffers. Digestion was performed at 37°C for 3h; then 2U of calf intestinal alkaline phosphatase was added to the total volume and then incubation continued at 37°C for a further 30min. The digested plasmid was purified from agarose gels as described in section 2. 2. 3.

B. Restriction Digestion of pGEX-2T

20µl of plasmid DNA was digested in a total volume of 40µl with BamHI (40U) or doubly digested with BamHI (40U) and EcoRI (40U) in the appropriate buffers. The DNA was digested at 37°C for 3h, to which 2U of calf intestinal alkaline phosphatase was then added and incubated at 37°C for a further 30min. The digested plasmid was purified from agarose gels as described in section 2. 2. 3.

2. 2. 5. Restriction Digestion of PCR Products

A. Restriction Digestion of Pea Chloroplastic Full Length E2/Didomain, Potato Mitochondrial E3β isoform and Pea Mitochondrial GDC L-protein

PCR products (30µl) were digested in a total volume of 60µl with BamHI (60U) in the appropriate buffer for 3h at 37°C. The digested PCR products were purified as described in section 2. 2. 3.

B. Restriction Digestion of *Arabidopsis thaliana* Mitochondrial E2

PCR products (30µl) were digested in a total volume of 60µl with XhoI (60U) in the appropriate buffer for 3h at 37°C. The digested PCR products were purified as stated in section 2. 2. 3.

C. Restriction Digestion of *Arabidopsis thaliana* Mitochondrial E2 Tridomain

PCR products (30µl) were digested in a total volume of 60µl with BamHI (60U) and EcoRI (60U) in the appropriate buffer for 3h at 37°C. The digested PCR products were purified as in section 2. 2. 3.

2. 2. 6. Ligations of Digested Plasmids and PCR Products

Purified digested plasmid DNA was ligated with purified digested PCR product DNA at various ratios of vector:insert (1:3, 1:6 and 1:9). The reaction was performed in a total volume of 20µl containing 3U T4 DNA ligase (Boehringer Mannheim). Routinely ligation reactions were incubated at room temperature (approx. 20°C) overnight. Ligation protocols were adapted from Sambrook *et al.*, (1989).

2. 2. 7. Making Competent *E. coli* Cells

Buffers

Buffer I: 30mM KAc, 100mM $RbCl_2$, 10mM $CaCl_2$, 75mM $CaCl_2$, 50mM $MnCl_2$, 15% (v/v) glycerol, pH 5.8

Buffer II: 10mM 3-[N-Morpholino] propanesulfonic (MOPS), 10mM $RbCl_2$, 15% (v/v) glycerol, pH 6.8

Both buffers were filtered sterilized and stored at 4°C.

Protocol for Making Competent Bacteria

E. coli was grown overnight at 37°C on minimal LB plates. A single colony was used to grow an overnight 5ml LB culture and used to inoculate 100ml LB and grown until an optical density at 550nm of 0.48 was obtained. The culture is chilled on ice for 5min, and then centrifuged at 3, 000rpm for 10min at 4°C. The pellet was resuspended in 40ml of Buffer I incubated on ice for 5min, and then centrifuged as before. The pellet was finally resuspended in 4ml Buffer II and chilled on ice for a further 15min. The cells were stored in 200µl aliquots at -80°C.

2. 2. 8. Transformation of Competent *E. coli*

E. coli DH5α cells were routinely used for the propagation and harvesting of recombinant plasmid DNA, while *E. coli* BL21 DE3 pLysS cells were used for the expression of the desired recombinant proteins. These bacterial strains were made competent and were transformed following the same method.

Buffers for Transformation

SOB medium (100ml): 2g Bacto tryptone, 0.5g Bacto yeast extract, 0.05g NaCl, pH 7.5
Autoclave and cool.
Add 10mM $MgCl_2$ and 10mM $MgSO_4$
both filtered sterilised

SOC medium (100ml): SOB supplemented with 0.4% (w/v) glucose (filter sterilised)

A. Transformation of Competent *E. coli* with Ligated Plasmid

Competent cells were thawed on ice, 50µl of cells was used for a single transformation. The entire ligation reaction mixture (20µl) was added to the cells and incubated on ice for 15min. Cells were heat shocked for 90s at 42°C, then returned to the ice for a further 2min. SOC medium (450µl) was added to the cells and incubated with shaking at 37°C for 45min. This mix (200µl) was plated out onto an appropriate antibiotic plate and incubated at 37°C overnight (16h).

B. Transformation of Competent *E. coli* with Plasmid DNA

Transformation of competent bacteria with plasmid DNA follows the same method as above (section 2. 2. 8A) but with one exception in that only 1µl of plasmid DNA is used per transformation.

2. 2. 9. Isolation of Wild Type or Recombinant Plasmid

Single colonies of bacteria were grown in 5ml LB with the appropriate antibiotic overnight (16h) with constant shaking at 37°C. The cells were pelleted by centrifugation at 3, 000rpm for 5min and the LB was discarded. Plasmid DNA was then prepared using the Wizard® Plus *SV* Minipreps DNA Purification System

(Promega) as instructed by the manufacturer. Isolated DNA (5µl) was viewed by agarose gel electrophoresis.

2. 2. 10. Small Scale Over-expression of Proteins

Single colonies of transformed BL21 (DE3) pLysS cells were grown overnight in 5ml LB, supplemented with the appropriate antibiotic, ampicillin (50µg/ml). LB (50ml) plus antibiotic was inoculated with 1ml of the overnight culture and grown at 37°C until an A_{600} of 0.5 was obtained. In the case of over-expression of PDC-E2 (full length and di/tridomains), the LB was supplemented with lipoic acid (final concentration 0.1mM) at the time of IPTG (final concentration 1mM) induction. *E. coli* cells were induced for 3h at 30°C with vigorous shaking. Samples were taken at hourly intervals throughout the induction and centrifuged at 13, 000rpm for 1min to pellet the cells. Pellets were resuspended in Laemmli sample buffer (section 2. 2. 14) using 10µl buffer per 0.1 A_{600} unit per 1ml of sample. Protein expression was observed by SDS-PAGE (section 2. 2. 14). Cells were harvested by centrifugation at 3, 000rpm for 15min and the supernatant discarded. The pellet was then resuspended in either buffer A (section 2. 2. 17) for His-tagged proteins or PBS (section 2. 2. 16) for GST-fusion proteins, before being stored at -20°C.

2. 2. 11. Large Scale Over-expression of His-Tagged Proteins

A single colony was used to inoculate 10ml of LB, supplemented with ampicillin (50µg/ml) and grown at 37°C overnight with shaking. LB (500ml) supplemented with ampicillin was inoculated with the 10ml overnight culture and grown at 37°C with constant shaking until an A_{600} of 0.5 was reached. The cultures were induced by the addition of IPTG (final concentration of 1mM). PDC-E2 proteins were expressed in the presence of lipoic acid (final concentration of 0.1mM). The cultures were incubated for 3h at 30°C (or 5h at room temperature for pea L-protein precursor) with constant vigorous shaking. Again 1ml samples were taken at hourly intervals, the A_{600} was measured, and samples centrifuged at 13, 000rpm for 1min to pellet the cells which were then resuspended in Laemmli buffer (section

2. 2. 14) using 10µl buffer per 0.1 A_{600} unit per 1ml of sample. Protein expression was viewed by SDS-PAGE (section 2. 2. 14). The *E. coli* cells were harvested by centrifugation at 8, 250rpm for 15min in a JA14 rotor in a Beckman J2-21 centrifuge to pellet the cells. The supernatant was discarded and the pellet was resuspended in 10ml buffer A (section 2. 2. 20) before being stored at -20°C.

2. 2. 12. Solubilisation of Expressed Fusion Proteins

A. French Pressure Cell

E. coli cells were disrupted by 3-4 passes through an automatic French pressure cell at 750psi (for the small French pressure cell, 5ml) or 950psi (for the large French pressure cell, 40ml) using an internal pressure of 14, 000psi. The French pressure cells were pre-cooled on ice before use.

B. Solubilization of Expressed Fusion Proteins

Over-expressed proteins in 50ml cultures were centrifuged at 3, 000rpm for 15min at 4°C. The pellets were resuspended in 3ml of 50mM KPi buffer, pH 7.6 prior to disruption by French pressure treatment (section 2. 2. 12). An aliquot (100µl) of the treated sample was retained before the remainder was centrifuged at 4°C at 10, 000rpm in a Beckman centrifuge using a JA17 rotor. An aliquot (100µl) of the supernatant was saved and the rest discarded. The pellet was resuspended in the original volume of 50mM KPi buffer, and centrifuged as before. This was repeated a total of three times. Each time the supernatant was discarded and replaced with fresh 50mM KPi buffer. Finally the pellet was resuspended in 50mM KPi buffer, pH 7.6 and an aliquot (100µl) of this pellet suspension was saved. An equal volume of Laemmli sample buffer (section 2. 2. 14) was mixed with the saved samples and the solubility of the recombinant protein was viewed by SDS-PAGE (see section 2. 2. 14).

2. 2. 13. Solubilisation of Insoluble Proteins from Inclusion Bodies

This protocol has been adapted from the Protein Folding Kit from Novagen.

Buffers

Inclusion body wash buffer: 20mM Tris-HCl buffer, pH 7.5, 10mM EDTA
1% (v/v) Triton X-100

Solubilization buffer: 50mM KPi buffer, pH 7.6

30% (w/v) N-lauroylsarcosine.

Preparation and Solubilisation of Inclusion Bodies

Following over-expression, the harvested cells were resuspended in 0.1x culture volume of wash buffer. Cells were disrupted by French pressure treatment as before, and the insoluble material was removed by centrifugation at 8, 250 rpm for 10min at 4°C. The pellet was resuspended in 0.1x culture volume of wash buffer and re-centrifuged. This was repeated 3 times. The inclusion body pellet was resuspended at a concentration of 10-20mg/ml in 50mM KPi buffer, supplemented with the appropriate volume of 30% (w/v) N-lauroylsarcosine to give a final concentration ranging from 0%-2%. Complete solubilisation was generally achieved with 0.5% (w/v) detergent. The suspension was incubated at room temperature for 15min with continuous gentle mixing and clarified by centrifugation at 8, 250 rpm for 10min at 4°C. N-lauroylsarcosine was partially removed by dialysis into two changes of 50mM KPi buffer using greater than 50x the volume of sample. Solubilisation of inclusion bodies was viewed after SDS-PAGE and staining with Coomassie blue.

2. 2. 14. SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE)

SDS-PAGE Buffers

Resolving gel: 8-15% (w/v) acrylamide,
0.5M Tris-HCl buffer, pH 8.8,
0.1% (w/v) sodiumdodecyl sulphate (SDS),
0.1% (w/v) ammonium persulphate,
0.1% (v/v) N, N, N^1, N^1-tetramethlethylene diamine (TEMED)

Stacking gels: 5.4% (w/v) acrylamide,
0.06M Tris-HCl buffer, pH 6.8,
0.1% (w/v) SDS,
0.1% (w/v) ammonium persulphate,
0.1% (v/v) TEMED

SDS buffer: 0.025mM Tris-HCl buffer, pH8.3,
0.2M glycine,
1% (w/v) SDS

Laemmli buffer: 62.5mM Tris-HCl buffer, pH 6.8,
2% (w/v) SDS,
10% (w/v) sucrose,
trace of pyronin Y

1M DTT : final concentration 150mM

SDS-PAGE

Proteins were resolved under denaturing conditions using the method of Laemmli (1970). Samples were resuspended in Laemmli buffer to which 1M DTT is added to a final concentration of 150mM, and boiled for 5min before loading onto the gel. Gels were run using vertical electrophoresis kits at a constant 50mA per gel.

Proteins were stained using 0.1% (w/v) Coomassie Brilliant Blue R-250 (Fluka) dissolved in 50% (v/v) methanol, 10% (v/v) acetic acid for 30min with shaking. Protein bands were visualised by destaining the gels in 10 % (v/v) methanol, 10% (v/v) acetic acid. Long term storage of these gels was achieved by soaking for 2h in 35% (v/v) ethanol and 2% (v/v) glycerol before being dried between two sheets of cellophane (Pharmacia) in an Easy Breeze apparatus (Hoefer Scientific Instruments, USA).

2. 2. 15. Enhanced Chemiluminescence (ECL™) for Detection of Specific Antigens (Western Blot)

A light emitting non-radioactive method for detecting immobilised specific antigens. Western blotting was adapted from Sambrook *et al.*, (1989).

Buffers

10x Transfer buffer: 25mM Tris-HCl buffer, pH 7.2, 192mM glycine, 0.02% (w/v) SDS, 20% (v/v) methanol

Blocking solution: 20mM Tris-HCl buffer, pH 7.2, 15mM NaCl, 5% (w/v) Non Fat Milk, 5% (v/v) Normal Donkey Serum, 0.2% (v/v) Tween 20

Primary Antibody: 20mM Tris-HCl buffer, pH 7.2, 1% (w/v) Non Fat Milk, 1% (v/v) Normal Donkey Serum, 1% (v/v) Tween 20, 1:2500 dilution of antibody

Wash Solution: 20mM Tris-HCl buffer, pH 7.2, 15mM NaCl, 1% (w/v) Non Fat Milk,1% (v/v) Normal Donkey Serum, 1% (v/v) Tween 20

Secondary Antibody : 20mM Tris-HCl buffer, pH 7.2, 150mM NaCl,
1% (w/v) Non Fat Milk, 1% (v/v) Normal Donkey Serum,
1:1000 dilution of HRP-antibody

High Salt Wash: 20mM Tris-HCl buffer, pH 7.2, 150mM NaCl

Method For ECL™ Western Blot

Following separation by SDS-PAGE, proteins were electrophoretically transferred onto nitrocellulose paper using a BioRad Trans-blot cell at 40mA overnight in the presence of 1x transfer buffer. The nitrocellulose was briefly stained with Ponceau S solution, allowing visualisation of transferred protein bands. This was removed by washing with distilled water and excess binding sites were blocked by immersing the nitrocellulose for 2h with agitation in blocking solution at room temperature. The nitrocellulose was briefly washed twice and then once for 15min followed by two 5min washes. The membrane was then incubated for 1h at room temperature with the primary antibody solution (100ml/membrane). Excess primary antibody was removed by a series of washes as described previously. The membrane was then incubated for a further 1h at room temperature with the secondary antibody conjugated with horeseradish peroxidase. After the secondary antibody incubation, the membrane was washed for 15min, then four 5min washes and finally for 30min in high salt solution. The detection step was carried out as described in the Amersham Protocol for ECL detection. Exposure times were typically 10s to 1min.

2. 2. 16. Purification of GST-fusion Proteins

Buffers

1x PBS : 170mM NaCl, 3mM KCl, 10mM Na_2HPO_4,
1mM KH_2PO_4, pH 7.2

Cation Exchange Purification Method

Protein purified from the POROS MC Zn^{2+}-imidoacetate column was pooled and dialysed for 2-4h at 4°C; the sample was then centrifuged to remove any precipitated salts or protein at 10, 000rpm, for 10min at 4°C.

The POROS® 20S column was washed with 1M NaCl at a flow rate of 10ml/min to remove any residual bound proteins before equilibration with 50mM KPi buffer, pH 7.0, 20mM NaCl for 10 CV. Dialysed sample was manually loaded and washed onto the column with 50mMKPi buffer, pH 7.0, 20mM NaCl for a further 10 CV. Multiple injection could be programmed into the method with small washes between each loading. Bound proteins were eluted with an increasing concentration gradient of NaCl, 20mM-1M, over 8 CV. Eluted material was automatically collected in 2ml fractions throughout the gradient, a final wash with 2M NaCl was used to clear the column of tightly bound proteins. The eluted protein was TCA precipitated (see section 2. 2. 21) and the purity was judged by SDS-PAGE (see section 2. 2. 14) and Coomassie blue staining.

B. Anion Exchange Purification

Buffers

Dialysis Buffer: 20mM Tris-HCl buffer, pH 7.0, 20mM NaCl,
Stock buffers: 100mM Tris-HCl buffer, pH 7.0
3M NaCl

All buffers are filtered and stored at 4°C, deionized water is also filtered.

Anion Exchange Purification Method

POROS® MC Zn^{2+}-imidoacetate purified protein samples were pooled and dialysed into buffer for 2-4h at 4°C with a change in buffer. After dialysis the sample was

centrifuged to remove any precipitated salts or protein at 10, 000rpm, for 10min at 4°C.

An initial column wash was performed with 1M NaCl to remove any bound proteins followed by equilibration with 20mM Tris-HCl buffer, pH 7.0, 20mM NaCl buffer for 10 CV. The sample was loaded manually in 5ml volumes in the presence of 20mM Tris-HCl buffer, pH 7.0, 20mM NaCl and washed for 10 CV. Elution of bound proteins was achieved by an increasing concentration gradient of NaCl, 20mM-1M, over 8 CV. Fractions of the eluted material (2ml) were collected throughout the gradient. Finally a wash with 2M NaCl buffer was used to clean the column. The eluted protein was TCA precipitated (see section 2. 2. 21) and the purity was judged by SDS-PAGE (see section 2. 2. 14) and Coomassie blue staining.

2. 2. 19. Production of Dihydrolipoamide

Lipoamide (1g) (DL-6,8-thioctic acid amide) was dissolved in 25ml 80% (v/v) methanol and cooled on ice. Sodium borohydride (1g) was dissolved in 5ml distilled H_2O cooled on ice. The sodium borohydride was added to the lipoamide solution while stirring on ice; once added the mixture was removed from the ice and stirred for a further 20min and the pH was adjusted to 2.0 with 0.25M HCl. Dihydrolipoamide (DHL) was extracted from the solution by mixing thoroughly with 50ml chloroform. The mixture was allowed to settle into two phases, the bottom organic phase was decanted into another flask and the chloroform extraction was repeated 4 times. Chloroform was removed by evaporation using a rotary evaporator. The resulting precipitate was dissolved in a 2.5:1.0 ratio of toluene/heptane and heated to ensure complete resuspension. The mixture was dried down until a small volume was left using a rotary evaporator. This was further dried until a white precipitate formed. DHL was stored at -20°C. The quality of the DHL was checked by assaying E3 activity with increasing amounts of DHL, dissolved in 100% ethanol at a concentration of 20mg/ml (100mM), until no further increase in activity was recorded.

2. 2. 20. Enzymatic Assays

A. Principle of Pyruvate Decarboxylase (E1) Assay

Pyruvate decarboxylase (E1) activity was assayed by monitoring the reduction of a dye, 2, 6-dichlorophenolindophenol (DCPIP), to its colourless form as a difference in absorbance at 600nm (Khailova *et al.*, 1977).

Pyruvate Decarboxylase Assay Buffers

Solution A: 50mM KPi buffer, pH 7.6, 3mM NAD^+, 2mM $MgCl_2$, 0.2mM ThDP

DCPIP: 1mg/ml

Solution C: 100mM pyruvate.

Pyruvate Decarboxylase Assay

To a cuvette containing 670µl of solution A, pre-warmed to 30°C, the E1 sample was added along with 14µl DCPIP and monitored until a constant absorbance was obtained. Pyruvate (14µl) was then added. The ability of the E1 protein to reduce DCPIP was measured using a Shimadzu UV-2101PC UV-VIS scanning spectrophotometer and the reduction was recorded as a negative milli-absorbance units /min due to the loss of colour.
E1 activity can be calculated by using the molar extinction coefficient for DCPIP of 22000 $M^{-1}.cm^{-1}$.

B. Principle of Dihydrolipoamide Acetyltransferase (E2) Assay

The assay measures acetyl dihydrolipoamide formation as a change in absorbance at 232nm over 45s. Acetyl CoA was generated from acetyl phosphate in the presence of CoA and phosphotransacetylase. Acetyl groups were transferred to

free dihydrolipoamide in the presence of the E2 source. Although a molar extinction coefficient of $4400M^{-1}.cm^{-1}$ was described for the thioester bond formation (Millar *et al.*, 1999) the precise molar extinction coefficient of the immediate S^8-acetyl-dihydrolipoamide product cannot be determined due to the exchange of acetyl forming a 3:1 mixture of S^8- or S^6-acetyl-dihydrolipoamide but also diacetylated products (Yang *et al.*, 1997). Hence the activity was measured in $\Delta A_{232}\ min^{-1}$.

Dihydrolipoamide Acetyltransferase Assay Buffers

Stock Buffers: 1M Tris-HCl, pH 7.4,
100mM dihydrolipoamide in 100% ethanol
100mM acetyl phosphate,
1mM CoA

Dihydrolipoamide Acetyltransferase Assay

The assay was performed at 30°C in a total reaction volume of 1ml in a quartz cuvette containing 30mM Tris-HCl, pH 7.4, 1mM DHL, 1mM acetyl phosphate, 20μM CoA and 2U of phosphotransacetylase. A baseline was established after the formation of acetyl-CoA, before the E2 sample was added. The change in absorbance was measured over 45s using a Shimadzu UV-2101PC UV-VIS scanning spectrophotometer and recorded as $\Delta A_{232}\ min^{-1}$.

C. Principle of Dihydrolipoamide Dehydrogenase (E3) Assay

The activity of E3 was measured by the formation of NADH as an increase in absorbance at 340nm. E3 oxidizes the free dihydrolipoamide reducing the protein bound FAD^+ cofactor, which is then oxidized by NAD^+ to form NADH.

Dihydrolipoamide Dehydrogenase Assay

The assay mixture contained 670µl of solution A (section 2. 2. 20A), pre-warmed at 30°C to which the E3 source was added. The reaction was initiated by the addition of 20µl of DHL (section 2. 2. 20B) (final concentration of 3mM). The increase in absorbance was measured over a period of 45s and recorded as mA min^{-1} (milli-absorbance units /min) using the kinetics software on a Shimadzu UV-2101 PC UV-VIS scanning spectrophotometer. The increase in absorbance was converted into µmol NADH min^{-1} ml^{-1} using the molar extinction coefficient of NADH (6220 $M^{-1}.cm^{-1}$).

2. 2. 21. Trichloroacetic Acid (TCA) Precipitation of Purified Proteins

This method was used to precipitate purified proteins in preparation for SDS-PAGE analysis. To 100µl of each fraction, 10µl of 100% (w/v) TCA was added and incubated on ice for 30min. Precipitated protein was collected by centrifugation at 14, 000rpm for 15min at 4°C. The pellet was washed with 500µl ice cold acetone to remove residual amounts of TCA. The centrifugation step was repeated and the acetone was discarded. Remaining acetone was removed by allowing the sample to air dry. Finally the pellet was resuspended in 20µl of Laemmli buffer and 1M DTT added to final concentration of 150mM. The samples were boiled for 5min before SDS-PAGE analysis with Coomassie blue staining to visualise the proteins.

2. 2. 22. Dialysis of Protein Samples

Dialysis tubing was prepared by autoclaving in distilled water. When cool, the tubing was repeatedly rinsed with distilled water before being stored in 20% ethanol. Before use the tubing was rinsed thoroughly with distilled water. Routine dialysis of protein samples was performed at 4°C with continual stirring on a magnetic stirrer against 50x the sample volume with several changes of buffer over a period of time.

2. 2. 23. Concentrating Protein Samples

Protein samples were concentrated by absorbing excess buffer from the protein sample contained in dialysis tubing using PEG-6000 at room temperature until the desired concentration was reached. The second method employed Centricon Plus-20 centrifugal concentration filters (Amicon) with a molecular weight cut-off of 30, 000 Da to concentrate the protein sample by filtration. Samples were centrifuged in a swinging bucket centrifuge for 20min at 3, 000 rpm and collected by centrifugation of the inverted filter at 1, 000rpm for 2min as instructed by the manufacturer. Centrifugation steps were repeated until the desired concentration was obtained.

2. 2. 24. Determination of Protein Concentration

Protein concentration was determined by the method of Bradford (1976). The dye Coomassie Brilliant Blue G-250 is converted from red to blue upon protein binding followed by an increase in absorption at 595nm. The A_{595} value for known protein concentrations can be used to produce a standard curve of A_{595} versus concentration. Thereafter the A_{595} of unknown protein samples can be determined from the standard curve. Routinely bovine IgG was used to construct a standard curve at 595nm on a Shimadzu UV-2101 PC UV-VIS scanning spectrophotometer.

2. 2. 25. Size Exclusion Chromatography

The native molecular weights of purified protein samples were determined using a Superose 6 column (HR10, Pharmacia, 25ml bed volume) in conjunction with a fast protein liquid chromatography system (FPLC; Amersham Pharmacia Biotech). Protein samples were dialysed against 50mM KPi buffer, pH 7.2, 150mM NaCl, 1mM DTT at 4°C for approximately 6h. The sample was applied to a pre-equilibrated column and eluted with the same buffer at a flow rate of 0.5ml/min. A calibration curve using high molecular weight standards was constructed to allow estimation of M_r values. Standards included Blue Dextran (2 MDa), Thyroglobulin

(669 kDa), Apoferritin (443 kDa), β-Amylase (200 kDa), Alcohol dehydrogenase (150 kDa), BSA (66 kDa), Carbonic anhydrase (29 kDa), Cytochrome c (12.4 kDa).

2. 2. 26. Glutaraldehyde Crosslinking of Purified Protein

Oligomeric proteins were subjected to chemical crosslinking with glutaraldehyde. This method was modified from that described by White *et al.* (1993).

Cross-linking Buffers

25%(w/v) Glutaraldehyde
2M $NaBH_4$ (in 0.1M NaOH)
10% (w/v) sodium deoxycholate
100% (w/v) TCA

Cross-linking Reaction

Protein samples (20-50μg) were denaturated by incubating on ice in 100μl 2M GdmCl before cross-linking with glutaraldehyde. Glutaraldehyde (20-40μl) was added and mixed thoroughly for 2min at room temperature. $NaBH_4$ (25μl) was added, mixed vigorously and then incubated at room temperature for 20min. 3μl sodium deoxycholate was added and again mixed. To precipitate the protein TCA (10μl) was added and incubated on ice for a minimum of 15min. The samples were centrifuged at 10, 000rpm at 4°C for 10min, the supernatant was discarded and the pellet washed with ice cold acetone (500μl) before centrifugation at 10, 000rpm at 4°C for 10min. Finally the pellet was resuspended in 30μl Laemmli buffer and resolved by gel electrophoresis on phosphate buffered SDS polyacrylamide gels.

2. 2. 27. SDS-Phosphate Gel Electrophoresis

Preparation of SDS-phosphate buffered acrylamide gels was adapted from the method described by Weber and Osborn (1969).

Gel Buffers

25mM NaPi: 17.2mM $NaH_2PO_4.H_2O$,
6.95mM $Na_2HPO_4.7H_2O$, pH 7.0

10% (w/v) SDS

10% (w/v) ammonium persulphate

Phosphate Gel

6% gel: 6ml acrylamide,
15ml NaPi buffer, pH 7.0
8.73ml H_2O,
600µl SDS,
225µl ammonium persulphate,
45µl TEMED

Tank buffer: 125ml NaPi buffer, pH 7.0
125ml H2O
0.2% SDS

Phosphate Gel Electrophoresis

The protein samples resuspended in Laemmli sample buffer were resolved under denaturing conditions; 1M DTT was added to a final concentration of 150mM, and

boiled for 5min before loading onto the gel. Gels were run at 50mA per gel for 2-3h and stained as described in section 2. 2. 14.

2. 2. 28. Matrix Assisted Laser Desorption Ionisation (MALDI) Mass Spectrometry

MALDI mass spectrometry measures the mass of a molecule as a ratio of mass/ charge (m/z). The molecule of interest was mixed with matrix before being ionised by an intense laser beam. Ionised molecules were detected and their mass/charge (m/z) ratios were recorded by a Time-of-Flight Mass Spectrometer Voyager System 6138 (PE Biosystems). Each spectrum was the product of accumulated results from three different runs (approx. 200 laser shots) across the analyte-matrix crystals.

Sample Preparation

The matrix (sinapinic acid or α-cyano-4-hydroxy-cinnamic acid) was prepared at a concentration of 10mg/ml in 50% acetonitrile containing 0.1% (v/v) trifluoroacetic acid. The protein samples (analyte) were mixed in a 1:1 (v/v) ratio with matrix before being spotted (1μl) onto a stainless steel plate and allowed to air dry prior to mass spectrometry analysis. External calibration was performed with myoglobin (m/z=16, 000). Mass spectra were recorded in collaboration with Dr. S. Kelly, IBLS, University of Glasgow.

2. 2. 29. Circular Dichroism

Polarised light is split into two circularly polarised components, left and right polarised light. Organic molecules appear to rotate the plane of polarisation of the polarised light due to their chiral properties as a result of their structure, or their position in an asymmetric environment. These molecules interact differently with the left and right circularly polarised light. Circular dichroism (CD) is a spectroscopic property that is sensitive to the differential absorption of left and right circularly polarised components of plane polarised radiation. It is widely used

to study protein structure. CD in the near-ultraviolet region (290-250nm) is principally due to tryptophan, tyrosine, phenylalanine and cystinyl groups, reflecting the asymmetric environment of these aromatic side chains characteristic of the tertiary structure. CD in the far-UV region (240-180nm) is due to the peptide bond and reflects the different regular forms of secondary structure of the protein (α-helix, β-sheet etc.). For most proteins the native structure produces a strong CD spectra over most of the far- and near-UV wavelengths. The loss of intensity that occurs upon treatment of the protein with a denaturant, such as guanidinium chloride (GdmCl), combined with fluorescence measurements, allow the stability of the tertiary and secondary structure to be compared (Kelly and Price, 1997; 2000).

CD spectra were recorded at 20°C on a JASCO J-600 spectropolarimeter. All spectra were recorded in collaboration with Dr. S. Kelly, IBLS, University of Glasgow.

Sample Preparation

Purified protein samples were dialysed against approx. 6l of 150mM NaF, 50mM KPi buffer, pH7.2 at 4°C for 6-8h with several buffer changes. The dialysed protein sample was concentrated using Centricon Plus-20 centrifugal filters as described in section 2. 2. 23.

2. 2. 30. Fluorescence

Most chromophores when excited by light return to their unexcited ground state by a series of non-radiative transitions, e.g. releasing heat. Chromophores that are quite rigid and inflexible however return to ground state by losing a portion of the absorbed energy as light (fluorescence). Fluorescence spectroscopy is a sensitive technique which can be employed continuously or discontinuously to obtain information on protein tertiary structure. Proteins have intrinsic chromophores, called fluors, in the form of tryptophan, tyrosine and phenylalanine (Trp gives stronger spectra). In the native state these side chains are buried in the folded hydrophobic cores, but unfolding the protein allows these side chains to freely

rotate around the C_α-C_β bond and they then can then be excited. The light emitted is lower in energy than absorbed light and so as the protein unfolds a change in environment of the residues to the polar solvent is detected as fluorescence at a longer wavelength.

2. 2. 31. Analytical Ultracentrifugation

Analytical ultracentrifugation is the only technique that can determine the native M_r accurately usually within 5% of the calculated mass based on the protein sequence. It is also a means of studying the hydrodynamic (conformation and shape) and thermodynamic (stoichiometry and association) properties of a protein or macromolecule. Analytical ultracentrifugation relies on sedimentation of the molecule with gravitational force under controlled conditions. For accurate calculations of mass, physical attributes such as solvent density(ρ), partial specific volume$(\bar{v})$ and solvent viscosity(η) are required.

Analytical ultracentrifugation experiments can be performed in two modes; sedimentation equilibrium and sedimentation velocity. Only sedimentation equilibrium studies were employed and these are described in detail in Chapter 6.

Sample Preparation

For analytical ultracentrifugation, protein samples must be as near to their native states as possible. Proteins are ionic in character so the pH of their environment is important. To maintain a constant pH, a buffer, normally at a concentration of 10-50mM is used. In addition NaCl is added to suppress non-ideality owing to non-specific interactions. It is important that the protein solution is dialysed against a large excess of buffer; it then can be treated as a simple two component system. All protein samples were dialysed into 50mM KPi, pH 7.2, 150mM NaCl, at 4°C for approximately two days with 4x 1l buffer changes. The dialysate was used as a reference buffer in analytical ultracentrifugation studies to enabling accurate background subtraction.

For sedimentation equilibrium (SE) experiments, protein samples were prepared from the stock protein solution by dilution with dialysate to give a series of 9 concentrations (usually A_{280} of 0.1-0.5au). SE experiments employed a six channel cell that allowed three concentrations to be analysed in one experiment with aliquots of dialysate pipetted into the opposing cells.

All analytical ultrcentrifugation experiments were carried out using a Beckman Optima XL-I analytical ultracentrifugation, at a constant temperature of 4°C. All measurements of proteins were at a wavelength of 280nm. Analytical ultracentrifugation experiments, sample preparation and data handling were performed by G. Campbell and Dr. O. Byron, Division of Infection and Immunity, University of Glasgow.

CHAPTER 3

CLONING AND OVEREXPRESSION OF THE PYRUVATE DEHYDROGENASE COMPLEX COMPONENTS

3. THE CLONING AND EXPRESSION OF THE PLANT PYRUVATE DEHYDROGENASE COMPLEX COMPONENTS

3. 1. INTRODUCTION

Studies on the plant pyruvate dehydrogenase complex (PDC) are still at an early stage relative to the comprehensive body of work which has been performed on the bacterial and, more recently, their mammalian and yeast equivalents. In plants the majority of the research conducted has focussed on native complexes from plant sources such as pea (*Pisum sativum*), broccoli (*Brassica oleracea*), barley (*Hordeum vulgare*), spinach (*Spinacia oleracea*), potato (*Solanum tuberosum*) and maize (*Zea mays*). Much of the research on plant PDCs involved analysis of the enzyme in crude or partially purified fractions and only recently have purified preparations with a defined subunit composition become available. However, there are still ambiguities with respect to the subunit composition and organisation of the plant complexes owing to the presence of multiple isoforms of all the constituent enzymes, as seen in SDS-PAGE analysis of maize and potato mitochondrial PDCs (Thelen *et al.*, 1998; Millar *et al.*, 1998). Moreover, the presence of the E3BP in plant mitochondrial and chloroplastic complexes has still to be established.

Investigation into the structure of the native complexes has been hindered by the low abundance of PDC obtained from plant sources. For example, Thelen and co-workers (1998) reported that only 0.4mg of mitochondrial PDC was purified from 955g of fresh weight maize shoots. To allow structural investigation into PDC or the separate enzyme components substantial amounts of purified complex would be required; hence the aim of producing recombinant enzymes with the long term goal of reconstitution of the complete complex.

Molecular Characterisation of Plant PDC Components

Within the last ten years several cDNA sequences encoding the mitochondrial and plastidic PDC components from a variety of plant sources have been cloned, characterised and published. To date the following clones have been identified:
The α and β subunits of the *Arabidopsis thaliana* pyruvate decarboxylase (E1) component of mitochondrial PDC have both been cloned (Luethy *et al.*, 1994; 1995). The E1α cDNA encodes a polypeptide of 389 amino acids. From the deduced protein sequence the N-terminal 32 residues were predicted to be the mitochondrial targeting peptide. The deduced protein sequence was 50% identical to the human E1α sequence (Luethy *et al.*, 1995). This same group characterised the E1β cDNA which encoded a polypeptide of 363 amino acids that contained a 29 amino acid N-terminal mitochondrial targeting sequence. A high degree of sequence identity (59%) was observed with the human E1β subunit (Luethy *et al.*, 1994). Both of these proteins have been successfully expressed as His-tagged proteins after subcloning into the pET-29b expression vector (Thelen *et al.*, 1995).

In addition to these mitochondrial subunits, the *A. thaliana* plastidic forms of E1α and E1β have also been identified and cloned (Johnston *et al.*, 1997). These cDNAs were analysed and the results showed that the E1α cDNA encodes a 428 amino acid polypeptide while the E1β cDNA produces a polypeptide of 406 amino acids. Both these proteins were identified as being chloroplastic proteins owing to the presence of a predicted N-terminal chloroplast targeting sequence comprising 68 residues (E1α) and 73 residues (E1β) respectively.

The pea mitochondrial PDC E1α and E1β subunits have also been cloned and both expressed in *Bacillus subtilis* with Staphylococcal protein A as a fusion partner to aid the secretion of the proteins (Moreno *et al.*, 2000a). Both subunits have also been reported to assemble into their heterotetrameric structure when they are co-expressed in the cytoplasm of the methylotrophic yeast, *Pichia pastoris* (Moreno *et al.*, 2000b).

Several other E1 subunit sequences have been cloned inculding the potato E1α subunit and the maize E1α and E1β subunits, both from mitochondrial PDC (Grof *et al.*, 1995; Thelen *et al.*, 1999a).

The dihydrolipoamide acetyltransferase (E2) component of the mitochondrial and plastidic PDC have been identified from *A. thaliana* (Guan *et al.*, 1995; Mooney *et al.*, 1999). The mitochondrial E2 cDNA contains an open reading frame encoding a 610 amino acid polypeptide. The predicted protein sequence contains two highly conserved regions situated within larger repeating units. A similar organisation has been seen in human E2 where they have been identified as lipoyl domains. The second (inner) lipoyl domain of the *A. thaliana* exhibits 79% sequence similarity to the equivalent domain of the human E2. The cDNA that encodes the plastid E2 produces a smaller peptide of 480 amino acids. Comparison of the deduced protein sequence to other eukaryotic E2 sequences revealed the presence of three distinct homologous regions, the lipoyl domain, subunit binding domain and the catalytic domain. However from the sequence it is evident that the plastidic E2 contains only a single lipoyl domain. The chloroplastic origins of this E2 was further substantiated by importing this E2 precursor into isolated pea chloroplasts (Mooney *et al.*, 1999).

Maize mitochondrial PDC-E2 has been cloned and expressed producing a 542 amino acid polypeptide with 30-37% sequence identity to other mitochondrial E2 sequences (yeast, human and *A. thaliana*) but only 29% identity to *A. thaliana* chloroplastic E2 (Thelen *et al.*, 1999b). In contrast to the *A. thaliana* mitochondrial E2 sequence which contains two lipoyl domains, this maize enzyme houses only a single domain but contains a large mitochondrial targeting sequence (107 residues).

The genes responsible for encoding the third enzyme of PDC has also been identified from various plants and revealed the existence of multiple isoforms. Two groups identified the gene responsible for the pea mitochondrial L-protein of the glycine decarboxylase complex (Turner *et al.*, 1992; Bourguignon *et al.*, 1992). The predicted amino acid sequences from both groups show high homology to

lipoamide dehydrogenase (E3) of PDC from human, yeast, porcine and *E. coli* (average 70%). Bourguignon *et al.* (1992) predicted that a 470 amino acid polypeptide would be the mature product while Turner and co-workers predicted a mature protein of 498 amino acids. Alignment of both cDNAs by Bourguignon *et al.* (1996) indicated that they differed at the 3' end of the cDNA by one base. Re-sequencing the cDNA of Turner *et al.* (1992) showed that the extra base was an error and after correction the cDNA encoded a polypeptide of 470 amino acids, although three conservative amino acid substitutions were present. This could be accounted by allelic variation as seen in other enzymes from the same species but different cultivars (Bourguignon *et al.*, 1996).

The first chloroplastic E3 was identified in this laboratory from pea plants. Purification of pea E3 on a Mono Q HR ion-exchange column produced two separate peaks of E3 activity. The two E3s were distinct in their relative molecular mass judged by SDS-PAGE and electrospray mass spectrometry (Conner *et al.*, 1996).

Multiple isoforms of E3 (α_2, $\alpha\beta$, β_2) were also found in potato tubers (*Solanum tuberosum,* Maris Piper) and leaves. These may represent tissue and complex-specific isoforms of E3 (Cook *et al.*, 1996). The β isoform has been isolated and the cDNA sequence determined along with the predicted amino acid sequence, (Khan, S. S., 1996).

With the completion of the *A. thaliana* genome project, four genes were found that encode for the dihydrolipoamide dehydrogenase (E3) component, two of chloroplastic and two of mitochondrial origin. The chloroplastic E3 isoforms showed high sequence identity, 81% at the DNA level and 88% at the protein level (Lutziger and Oliver, 2000). These cDNAs have been named *ptlpd1*, encoding a 570 amino acid polypeptide and *ptlpd2*, a 567 amino acid polypeptide. The precursor protein of the *ptlpd1* is directed towards the stroma of chloroplasts where the mature protein was seen to accumulate (Lutziger and Oliver, 2000), providing evidence for a chloroplastic location. The nucleotide sequences of the mitochondrial E3 isoforms were 83% identical and both encoded 507 amino acid

polypeptides that were 92% identical (Lutziger and Oliver. 2001). Both display 85% identity with pea mitochondrial E3 and 40-53% identity to *E. coli*, yeast and human E3, but only 33% identity to the chloroplastic E3 isoforms of *A. thaliana*. The N-terminal 36 amino acids of these proteins exhibited the expected characteristics of a mitochondrial targeting sequence suggesting a mitochondrial location for both proteins.

The aim of this chapter was to clone the chloroplastic and the mitochondrial pyruvate dehydrogenase complex components and subsequently express them as either His-tagged proteins or as GST fusion proteins in *E. coli*. Cloning and overexpressing these components has the potential to provide a readily available source of recombinant enzymes for future enzymatic, protein-chemical and biophysical studies.

3. 2. CLONING THE PYRUVATE DEHYDROGENASE COMPLEX COMPONENTS

DNA sequences encoding the enzyme components of the pyruvate dehydrogenase complex were amplified specifically by PCR-based approaches using a high fidelity polymerase. The PCR sequences were then subcloned into either: pGEX-2T (Appendix 1) for expression as glutathione S-transferase (GST) fusion proteins, pET-14b or pET-28b (Appendices 2 and 3, respectively) for expression as His-tagged (six histidine residues) proteins. In each of these vectors the PCR sequence was ligated into the multiple cloning sites situated C-terminal to the GST sequence or the His-tag. Thus full-length recombinant proteins were expressed with a His-tag at their N-terminus while the truncated forms i.e. the didomain and the tridomain of the E2 component were expressed as GST-fusion proteins. Expressing an integral part of a protein rather than the mature sequence may produce a partially unfolded and insoluble protein. However, expressing these truncated proteins as GST-fusion proteins aids their folding and provides a convenient method for their purification. Addition of the His-tag or GST-fusion protein can be to either end of the protein. The N-terminus was chosen because it is not involved in the catalytic action or the assembly of these proteins.

His-tagged polypeptides were expressed in *E. coli* under the control of the T7 RNA promoter, specific for the bacteriophage T7 RNA polymerase, which is itself chemically induced. GST-fusion proteins are under the control of another chemically inducible promoter, tac promoter. The recombinant proteins are expressed in the cytoplasm of *E. coli* and can be purified specifically via the His-tag or GST-fusion partner.

3. 3. PCR OF PEA CHLOROPLASTIC DIHYDROLIPOAMIDE ACETYLTRANSFERASE (E2)

Pea (*Pisum sativum)* chloroplast dihydrolipoamide acetyltransferase (E2) was subcloned as the mature E2 and a truncated protein consisting of the N-terminal lipoyl domain and adjacent subunit binding domain. The unpublished gene sequence of the pea chloroplast E2 component (Figure 3.1) was used to design primers that encompassed both the didomain (lipoyl domain and subunit binding domain) and the mature E2 (for primers see Materials section 2. 2. 6). Incorporation of BamHI restriction sites into these primers produced DNA sequences for both E2 and didomain with BamHI sites at either end. The mature E2 was subsequently cloned into the BamHI site of pET-14b vector. The mature E2 would be expressed as a His-tagged protein while the E2 didomain sequence was subcloned into the BamHI site of pGEX-2T vector for expression as a GST-fusion protein.

PCR was performed using the recombinant pBluescript II SK^+ plasmid containing the E2 gene sequence as a template following the protocol described in Methods section 2. 2. 1. PCR was successful as shown in Figure 3.2 with PCR products of the expected size being present in each case. The didomain DNA sequence (Figure 3.2, panel A) resolves at the 0.7 kb marker (actual length is 688 bp) while mature E2 sequence resolves at the 1.3 kb marker (Figure 3.2, panel B) as expected (actual length is 1327 bp). Both PCR products were gel purified as described in Methods section 2. 2. 4, and digested with BamHI (see Methods section 2. 2. 6A). The digested products were purified and then ligated (see Methods section 2. 2. 7) into previously BamHI digested plasmid (see Methods sections 2. 2. 5A and B). Recombinant plasmids were propagated in competent *E. coli* DH5α cells as described in Methods section 2. 2. 9A. In both cases, positive clones were grown overnight and the recombinant plasmid DNA was isolated as described in Methods section 2. 2. 10. The recombinant plasmids were then used to transform *E. coli* BL21 (DE3) pLysS cells in order to express these E2 proteins (see section 3.6).

FIGURE 3.1: Nucleotide and the Deduced Amino Acid Sequence of the Pea Chloroplast PDC Dihydrolipoamide Acetyltransferase Enzyme

The cDNA sequence encoding the dihydrolipoamide acetyltransferase component of the pea chloroplast pyruvate dehydrogenase complex was cloned into pBuescript II SK^+. The arrows denote the start of the presequence and of the mature sequence. The * denotes the relevant lysine residue that is likely to be post-translationally modified by the addition of the lipoic acid cofactor.

```
                                                GGCACGAGCGGCACGAGTCCAAA          23

ATGAACGCACATTCCTCTTCGATAATCTCAATGGCTTCTCCTTCTCCTTCC                              74
 M  N  A  H  S  S  S  I  I  S  M  A  S  P  S  P  S                               17

CAATCTCTCCGCAACTCAACATCCCTCCCTTTCTCAACTCTCATCCTCCCC                              125
 Q  S  L  R  N  S  T  S  L  P  F  S  T  L  I  L  P                               34

                                                                      E2
CGCAACTCCTCCCATCGCCGCCGTCACTCTTCCTTCAAAATCCAAGCCAAG                              176
 R  N  S  S  H  R  R  R  H  S  S  F  K  I  Q  A  K                               51

ATCCGCGAAATCTTCATGCCTGCACTCAGTTCCACTATGACCGAGGGTAAA                              227
 I  R  E  I  F  M  P  A  L  S  S  T  M  T  E  G  K                               68

ATCGTCTCGTGGATCAAATCCGAAGGCGACACTCTCTCCAAAGGTGATAGT                              278
 I  V  S  W  I  K  S  E  G  D  T  L  S  K  G  D  S                               85

GTTGTTGTTGTTGAATCGGATAAAGCTGATATGGATGTTGAGACTTTCTAT                              329
 V  V  V  V  E  S  D  K  A  D  M  D  V  E  T  F  Y                               102
                      *
GATGGAATTCTTGCTGCTATTGTTGTTTCTGAAGGTGAAACTGCTCCGGTT                              380
 D  G  I  L  A  A  I  V  V  S  E  G  E  T  A  P  V                               119

GGTGCTCCGATTGGTTTGCTCGCTGAGACTGCTGAGGATATCGCGGAAGCT                              431
 G  A  P  I  G  L  L  A  E  T  A  E  D  I  A  E  A                              136

CAAGCGAAAGCCAAATCGGTCAAATCCGGTTCCGGTTCTTCTTCTCCTCCT                              482
 Q  A  K  A  K  S  V  K  S  G  S  G  S  S  S  P  P                               153

CCTGAAATACCTGACACTCCTCCGGTGACTTCTCAGTCTCCTTCTCCTCCT                              533
 P  E  I  P  D  T  P  P  V  T  S  Q  S  P  S  P  P                               170

CCGCCACCACCGGTTAAGGCTGTTTCTGATGGTCCGAAGAAAATCACGGCG                              584
 P  P  P  P  V  K  A  V  S  D  G  P  K  K  I  T  A                               187

ACACCTCAGGCTAAGAAACTTGCGAAGCAGCACAAGGTTGACATTGCTTCC                              635
 T  P  Q  A  K  K  L  A  K  Q  H  K  V  D  I  A  S                               204

GTCAATGGGACTGGTCCATTTGGTAGGATTACTCCGGCTGACGTGGAGGCT                              686
 V  N  G  T  G  P  F  G  R  I  T  P  A  D  V  E  A                               221

GCTGCTGGGATTGTGCCGTCGAAGAGTAATGCACCTTCTGTAGTAGCTTCT                              737
 A  A  G  I  V  P  S  K  S  N  A  P  S  V  V  A  S                               238

GCTCCGGTGGCGTCGGCTCCACCAAAAGCTGCTGCTTCTGCTGCTGCACCA                              788
 A  P  V  A  S  A  P  P  K  A  A  A  S  A  A  A  P                               255

CCTGCGTTGCCTGGTTCGACTAATGTTGCTTTCACAACTATGCAGTCTGCT                              839
 P  A  L  P  G  S  S  N  V  A  F  T  T  M  Q  S  A                               272

GTTGCTAAGAACATGGTGGAGAGTCTTTCTGTTCCTACGTTCCGTGTTGGA                              890
 V  A  K  N  M  V  E  S  L  S  V  P  T  F  R  V  G                               289

TATCCGGTTACCACGGATGCTCTTGATGCTCTATATGAGAAGGTGAAACCT                              941
 Y  P  V  T  T  D  A  L  D  A  L  Y  E  K  V  K  P                               306
```

```
AAGGGTGTTACTATGACGGCTATATTGGCCAAGGCTGCTGCCATGGCACTT      992
 K   G   V   T   M   T   A   I   L   A   K   A   A   A   M   A   L      323

GTTCAGCATCCAGTTGTCAATGCCACCTGCAAAGACGGCAAGAACTTTCAC     1043
 V   Q   H   P   V   V   N   A   T   C   K   D   G   K   N   F   H      340

TATAACAGTAACATTAATGTTGCTGTTGCCGTGGCAATCAATGGTGGTTTG     1094
 Y   N   S   N   I   N   V   A   V   A   V   A   I   N   G   G   L      357

ATAACCCCAGTTCTTCAAGATGCTGATAAGTTGGACTTGTATCTGTTGTCC     1145
 I   T   P   V   L   Q   D   A   D   K   L   D   L   Y   L   L   S      374

CAAAAATGGAAAGAACTAGTTGGAAAGGCTCGTTCAAAGCAATTGCAACCT     1196
 Q   K   W   K   E   L   V   G   K   A   R   S   K   Q   L   Q   P      391

CATGAGTATAATTCAGGAACTTTCACGCTTTCAAACTTGGGTATGTTTGGA     1247
 H   E   Y   N   S   G   T   F   T   L   S   N   L   G   M   F   G      408

GTTGACAGGTTTGATGCCATACTTCCACCAGGCCAGGGGGCTATCATGGCT     1298
 V   D   R   F   D   A   I   L   P   P   G   Q   G   A   I   M   A      425

GTCGGAGCATCAAAGCCTACTGTTCTGGCTGATGCAGATGGATTCTTCAGG     1349
 V   G   A   S   K   P   T   V   L   A   D   A   D   G   F   F   R      442

GTGAAAAATAAAATGCTGGTGAATGTAACAGCTGATCACCGAATAATTTAT     1400
 V   K   N   K   M   L   V   N   V   T   A   D   H   R   I   I   Y      459

GGGGCTGACTTGGCTGCCTTCCTTCAGACATTCTCTAAAATCATTGAAAAC     1451
 G   A   D   L   A   A   F   L   Q   T   F   S   K   I   I   E   N      476

CCAGAAAGCCTAACATTATAGTCATTCTTCACAGAGATATATATTTAGTTC     1502
 P   E   S   L   T   L   *                                              482

CACCGGAAGGTGTTTCTTGAGAAGCTGTTTGGCTCAATTCTTAAACTTTTT     1553

GCACTTTCTCTTTACACTTTCATTTTGTTATACTTTACTAGATCAATCACC     1604

CGGGAATTCTTAAACTTTGTACTTTTCTTTGCACTTTCTTTTTGTTATATT     1655

TTACTAGCTCGATCACATAGAGATGCACATGCATAAGATTGTAATAATATT     1706

TGTGAGAGACCATAATTTAATTGTCTACTTTTTTTTTTTTTTT             1757
```

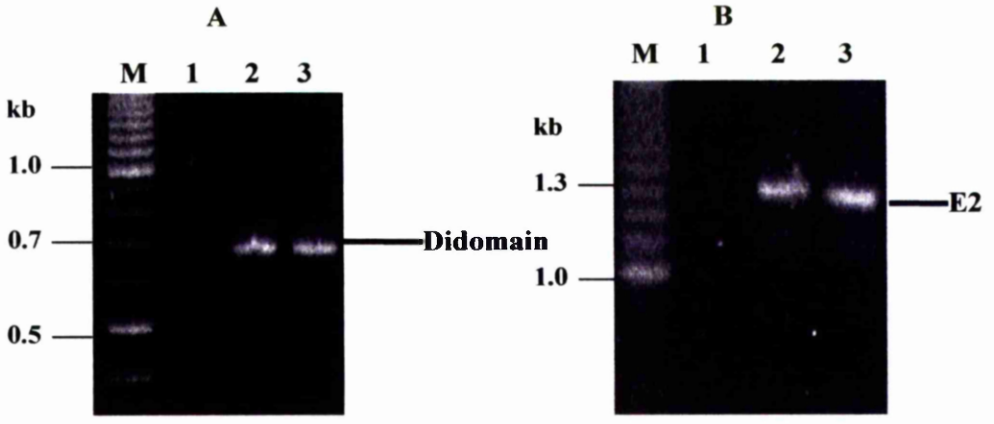


FIGURE 3.2: Polymerase Chain Reaction Products of the Pea Chloroplast E2 Didomain (A) and Full Length E2 (B) Sequences

5μl of each PCR reaction was resolved on 1.5% (w/v) agarose gels and stained with ethidium bromide.

Panel A: amplification of the E2 didomain; lane 1, control PCR with no DNA template. Lanes 2 and 3, positive PCR reactions producing products of 0.7 kb (actual size is 668 bp).

Panel B: PCR products of the amplified full length E2 sequence; lane 1, control PCR, no DNA template added. Lanes 2 and 3; PCR products of approx. 1.3 kb visible, actual size of sequence 1327 bp.

M: 1kb DNA size ladder.

3. 4. PCR OF THE *A. THALIANA* MITOCHONDRIAL DIHYDROLIPOAMIDE ACETYLTRANSFERASE (E2)

The existence of two types of oligomeric E2 cores in PDCs from various sources prompts the question as to the nature of the E2 organisation in the plant organelle-specific PDC complexes. It was essential to clone a mitochondrial E2 to allow comparison with the chloroplastic E2. As the pea mitochondrial E2 sequence was unavailable, the *A. thaliana* mitochondrial E2 sequence (Guan *et al.*, 1995) was employed. Primers were designed specifically for the tridomain (two N-terminal lipoyl domains and the adjacent subunit binding domain) and the mature protein using the published sequence (for primers see Materials section 2. 1. 6). The primers for the tridomain region of E2 were designed to incorporate a BamHI site at the 5' end and an EcoRI site at the 3' end, allowing directional subcloning into the pGEX-2T vector for expression of the tridomain as a GST-fusion protein. The primers designed to encompass the full-length mature protein incorporate a XhoI restriction site at either end to allow subcloning into pET-14b for the expression of E2 as a His-tag protein. XhoI was used instead of BamHI on this occasion as the full-length cDNA sequence contained an internal BamHI site.

PCR amplification of both sequences employed the PDC-E2 clone (a kind gift from Dr. Steve Rawsthorne, John Innes Centre, see section 2. 1. 7) as the template DNA. PCR was performed as described in the Methods section 2. 2. 1. and the amplified products were viewed by agarose gel electrophoresis as before. Figure 3.3 presents successful PCR results for the mitochondrial E2 tridomain with a PCR product resolving at 1 kb (predicted size is 1001 bp). The higher band observed is probably template DNA. The PCR product of full-length sequence resolved between 1.5 kb and 2 kb. The expected size of the DNA sequence is 1705 bp.

PCR products were gel purified and subjected to restriction digestion with the relevant restriction enzymes (see sections 2. 2. 6B and C) followed by ligation of full length E2 into pET-14b (digested with XhoI, see Methods section 2. 2. 5A) and the E2 tridomain into pGEX-2T (digested with BamHI and EcoRI, see Methods section 2. 2. 5B). Propagation of the recombinant plasmids in competent

E. coli DH5α cells completed the molecular cloning of *A. thaliana* mitochondrial E2 proteins. Isolated recombinant plasmids of both clones produced in *E. coli* DH5α were used to transform *E. coli* BL21 (DE3) pLysS cells for the expression of recombinant E2 proteins (see section 3.6).

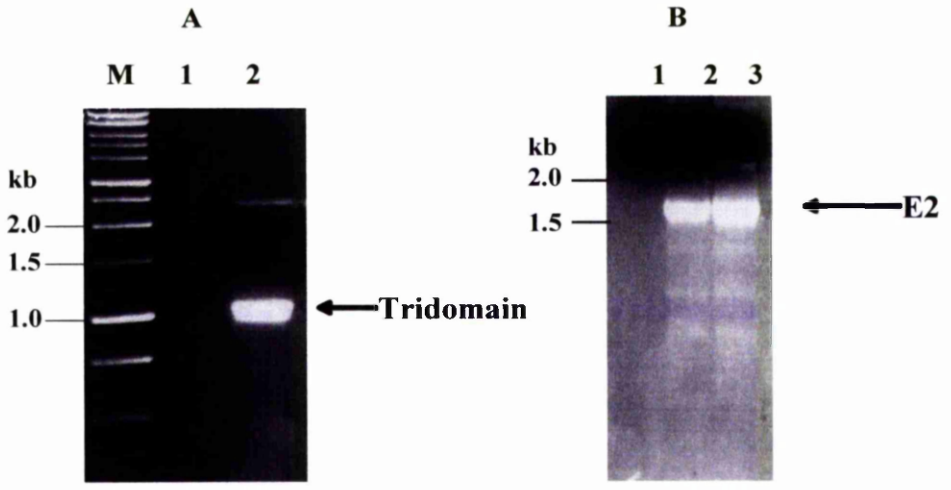


FIGURE 3.3: Amplification of the *A. thaliana* Mitochondrial PDC Tridomain and Full Length E2

PCR mixtures (5µl) were resolved on 1.5% agarose gels and the DNA viewed after staining with ethidium bromide under UV light.

Panel A: PCR of the E2 tridomain, Lane 1: negative control PCR mixture without template DNA. Lane 2: a positive result was seen by the presence of a single band of 1.0 kb (expected size 1001 bp).

Panel B: PCR results of the amplification of the full length E2 sequence. Lane 1: negative control PCR set up without template. Lanes 2 and 3: positive PCR with a major amplified DNA product between the 1.5 kb and 2.0 kb markers. Actual size of the sequence is 1705 bp.

M: 1 kb DNA ladder markers.

3. 5. POLYMERASE CHAIN REACTION OF THE MITOCHONDRIAL DIHYDROLIPOAMIDE DEHYDROGENASE (E3)

3. 5. 1. PCR of the Potato Mitochondrial Dihydrolipoamide Dehydrogenase

Previous work in our laboratory has demonstrated the existence of two mitochondrial E3 isoforms (α, β) from potato (*Solanum tuberosum*) that purified as homodimers and as a heterodimer. The gene encoding the E3β isoform has been identified by Khan, S. S. (1996) and available in the GenBank™/EMBL Data Bank. Primers were designed specifically for the mature E3 (see Materials section 2. 1. 6) with BamHI restriction sites included to aid with subcloning into the pET-14b vector. PCR was performed as described in Methods section 2. 2. 1, using the E3β cDNA as the template. The results of the PCR can be seen in Figure 3.4, panel A. Amplification of the E3 sequence produces a DNA band of approx. 1.5 kb as observed. Its predicted size was 1412 bp determined from the cDNA sequence. The PCR product was gel purified and digested with BamHI. After purification from the digest reaction, the E3 sequence was ligated into BamHI digested pET-14b vector and subsequently incorporated into competent *E. coli* DH5α cells for the propagation of recombinant plasmids.

3. 5. 2. PCR of the Pea Mitochondrial GDC L-protein (PDC-E3)

Two E3 enzymes were shown to exist in the pea plant (Taylor *et al.*, 1992 a,b) identified as the mitochondrial E3 and the chloroplastic E3 (Conner *et al.*, 1996a). The gene sequence encoding the pea mitochondrial glycine decarboxylase complex L-protein was published by two groups (Turner *et al.*, 1992; Bourguignon *et al.*, 1996). Analysis of this L-protein revealed that it was identical to PDC-E3 suggesting that the glycine decarboxylase complex and the pyruvate dehydrogenase complex from pea mitochondria share the same E3 (Bourguignon *et al.*, 1996). Primers were designed to encompass the mature protein and the precursor

(including the mitochondrial targeting sequence). Both sets of primers were designed to include BamHI restriction sites (see Materials section 2. 1. 6).

The sequences were amplified by PCR using the cDNA clone pGDLE (a kind gift from Dr. Steve Rawsthorne, John Innes Centre, see Materials section 2. 1. 7). The results of the PCR for each sequence are presented in Figure 3.4, panel B and 3.4, panel C, showing the amplification of the precursor and the mature L-protein (E3) DNA sequence, respectively. Both reactions produce amplified products of approx.1.5 kb. The predicted sizes of the precursor and mature sequences were 1583 bp and 1549 bp, respectively. PCR products were then gel purified (see Methods section 2. 2. 4), digested with BamHI and then ligated into the BamHI site of digested pET-14b vector. As before *E. coli* DH5α cells were used to propagate the recombinant plasmids.

Recombinant plasmids containing the potato E3 sequences and pea L-protein were incorporated into competent *E. coli* BL21 (DE3) pLysS cells which then permitted individual E3s to be expressed as His-tagged proteins (see section 3.6).

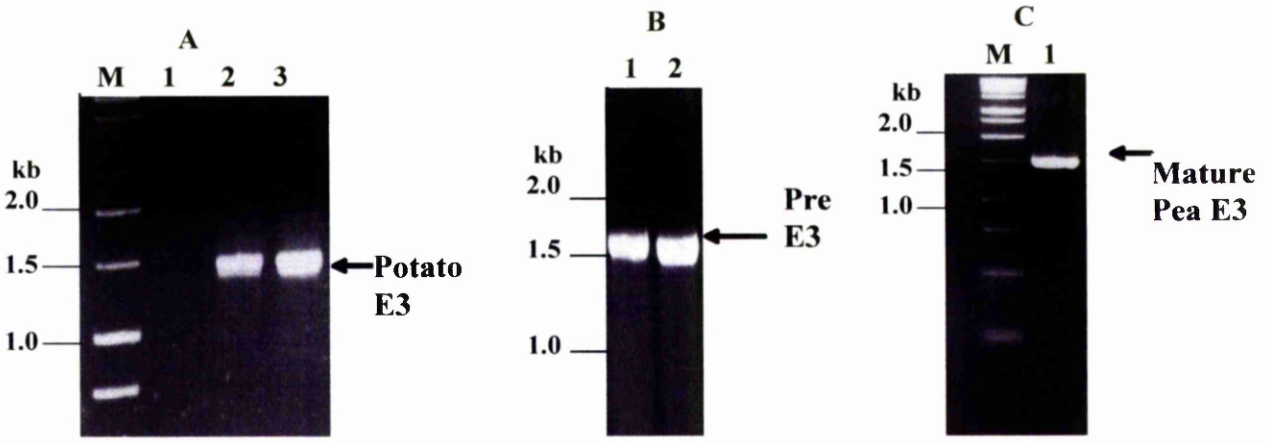


FIGURE 3.4: Polymerase Chain Amplification of the Mitochondrial E3 from Potato (A) and Pea (B and C)

Each PCR mixture (5µl) was viewed by gel electrophoresis on 1.5% agarose gels followed by staining with ethidium bromide and visualised under UV light.

Panel A: amplification of potato E3 sequence. Lane 1: negative control PCR with no template DNA. Lanes 2 and 3: positive PCR using clone pB3.1 as the template DNA, shows a product at approx. 1.5 kb (expected size of 1412 bp).

Panel B: amplification of pea L-protein precursor. Lanes 1 and 2: positive PCR producing a product of 1.5 kb (expected size of 1583 bp).

Panel C: amplification of mature pea L-protein sequence. Lane 1: positive PCR with a product of 1.5 kb in length (actual size is 1549bp).

M: 1 kb DNA marker ladder

3. 6. OVEREXPRESSION OF THE RECOMBINANT PROTEINS

The four expression vectors, pGEX-2T, pET-14b, pET-28b and pCR®T7/NT-TOPO®, employed to produce GST-fusion or His-tagged proteins, are chemically induced by the addition of 1mM isopropyl β-thiogalactopyranoside (IPTG). The cloned cDNAs are under the control of the promoters of the vectors, which are themselves controlled by the host cells. The expression host *E. coli* BL21 (DE3) pLysS has a copy of the T7 RNA polymerase gene integrated into its genome under the control of the lac UV5 promoter. In uninduced cells, low levels of transcription are often observed. It is therefore advantageous to use expression hosts carrying the pLysS plasmid that allows more stringent control of basal levels of transcription by providing T7 lysozyme, a natural repressor of T7 RNA polymerase. This minimises the possible toxic effects of heterologous protein expression prior to rapid induction of recombinant protein on addition of IPTG.

In the presence of IPTG, the T7 lysozyme is still present to inhibit the T7 polymerase; however, due to the dramatic increase in the T7 polymerase levels, rapid transcription of the target gene is initiated.

T7 RNA polymerase is selective and extremely active so that almost all of the host cell resources are converted to the expression of the target gene and, in some cases, the desired protein can be expressed to a level of 30-50% of total cellular protein. Overexpression of these proteins was performed as described in the Method sections 2. 2. 13 and 2. 2. 14.

3. 6. 1. Overexpression of the Chloroplastic and the Mitochondrial PDC Dihydrolipoamide Acetyltransferases

The N-terminal didomain of pea chloroplast PDC-E2 and the equivalent *A. thaliana* N-terminal mitochondrial E2 tridomain were both cloned into the pGEX-2T vector and expressed as GST-fusion proteins. Induction was performed routinely on a small scale in 50ml of LB supplemented with ampicillin, at 30°C. In contrast, mature proteins were expressed as His-tagged proteins in 500ml of LB cultures (supplemented with ampicillin) at 30°C.

The E2 component of PDC is post-translationally modified by the addition of lipoic acid to a specific lysine residue within the lipoyl domain. Over-expressing the E2 proteins in *E. coli* cells required the addition of exogenous lipoic acid (0.1mM) to the growth medium providing a pool of cofactor for the *E. coli* lipoyl ligase. Samples were normally taken at time zero and after 3h, prior to checking the expression of these proteins by SDS-polyacrylamide gel electrophoresis and Coomassie blue staining (see Methods section 2. 2. 17).

SDS-PAGE analysis of all the PDC-E2 gene products, either as GST-fusion or His-tagged proteins is shown in Figure 3.5. The expected M_r value of each recombinant polypeptide was calculated from their amino acid sequence. Strong bands can be seen in the T3 sample corresponding to chloroplastic E2 gene products in lanes A and B, respectively. Both pea chloroplast E2 constructs resolved with a M_r approx. 50, 000 while their predicted M_r values are 48, 219 for the E2 didomain and for the mature E2 it is 47, 703.

Expression of *A. thaliana* mitochondrial E2 tridomain and the mature E2 is shown in Figure 3.6 (lanes C and D, respectively) both proteins migrate with a M_r of approx. 70, 000. The expected M_r values for these proteins were 61, 368 and 62, 674 for the tridomain and mature E2, respectively.

The M_r for each PDC-E2 component observed by SDS-PAGE is higher than predicted. This aberrant migration has been observed for other E2 components and has been attributed to the lipoyl domains and the linker regions found in the E2 protein (Guest *et al.*, 1985; Allen and Perham, 1991; Thelen *et al.*, 1999). The high proportion of alanine and proline and the charged amino acids in the linker regions is thought to alter their migration during SDS-PAGE as so an aberrantly high M_r value is estimated (Miles *et al.*, 1988; Guan *et al.*, 1995).

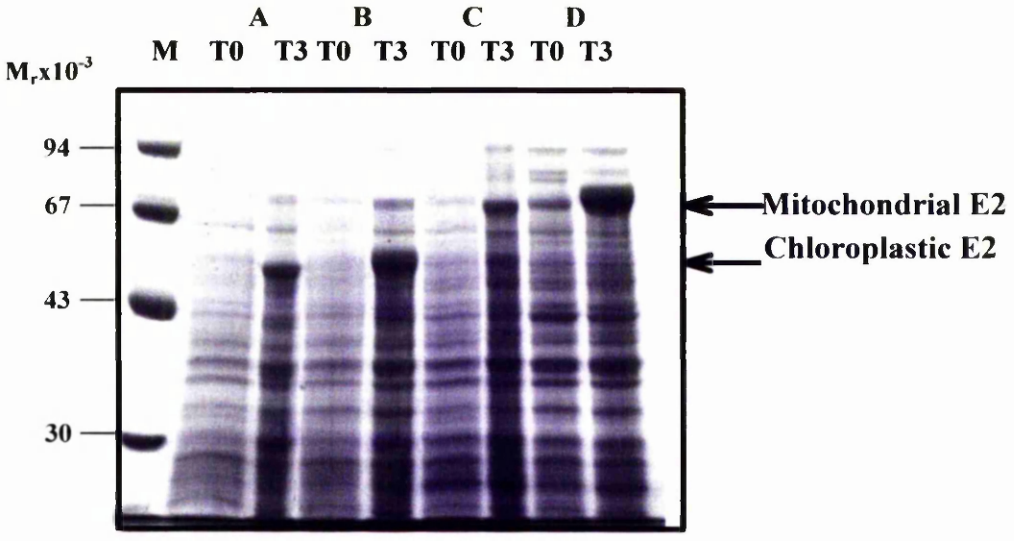


FIGURE 3.5: Expression of the Mature and N-terminal Domains of the Chloroplastic and Mitochondrial Dihydrolipoamide Acetyltransferase

Bacterial extracts taken at time 0 (T0) and 3h after IPTG (T3) were separated by SDS-PAGE on 10% gels and stained with Coomassie blue.

Lanes A and B: E2 didomain-GST-fusion with a M_r of 48, 219 (A) and mature E2 as a His-tagged protein with a M_r of 47, 703 (B).

Lanes C and D: *A. thaliana* mitochondrial E2 as a tridomain-GST-fusion with a M_r of 61, 368 (C) and the mature E2 as a His-tagged protein with a M_r of 62, 674 (D).

T0 indicates the time of induction with IPTG and T3 refers to 3h of growth after induction, usually at 30°C.

M: relative molecular mass markers.

3. 6. 2. Overexpression of the Mitochondrial Dihydrolipoamide Dehydrogenase

Both the potato E3 and pea L-protein (E3) were expressed as described in section 2. 2. 14. The potato E3 was expressed at 30°C for 3h, while the pea mitochondrial L-protein was expressed at room temperature (22-25°C) for 5h. The lower temperature was found to be beneficial in increasing the amount of soluble protein obtained with no significant effect on expression levels. The L-protein precursor was insoluble (see section 3. 6. 3) at all temperatures investigated and so it was expressed at 37°C for 3h. Samples were taken before (T0) and after IPTG induction (T3) and used to analyse expression by SDS-PAGE. The expected M_r values of the proteins were calculated from their amino acid sequences. The potato mitochondrial E3 has an expected M_r of 52, 015, the M_r value of the pea mitochondrial mature L-protein is 55, 897 and the precursor is 59, 463.

Expression of these proteins occurred in high yield as indicated by the strongly stained band on the SDS-polyacrylamide gel with a M_r of approx. 55, 000 shown in Figure 3.6. Low levels of E3 expression were visible in the T0 samples of potato E3 and mature L-protein. This is most likely due to a leaky T7 promoter allowing basal levels of transcription of the target protein in uninduced cells. There is no distinguishable difference in mass between the potato E3 and pea mature L-protein, while the L-protein precursor displays a lower mobility as expected.

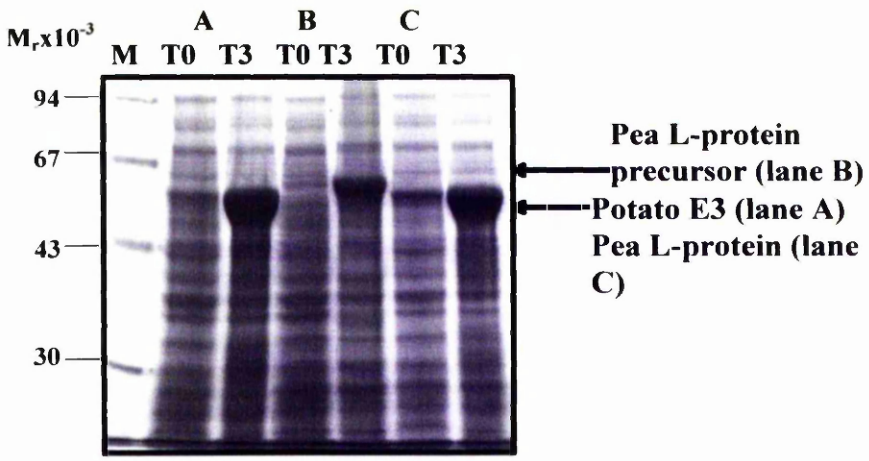


FIGURE 3.6: Expression of Mitochondrial Dihydrolipoamide Dehydrogenases

Soluble bacterial extract was resolved by SDS-PAGE on a 10% gel and the proteins were visualised by staining with Coomassie blue.

Lane A: expression of potato mitochondrial E3 with a predicted M_r of 52, 015,

Lane B: expression of pea mitochondrial L-protein precursor, with a predicted M_r of 59, 463.

Lane C: expression of mature L-protein, with a predicted M_r of 55, 897.

M: relative molecular mass markers.

3. 6. 3. Solubility of Recombinant Proteins

After successful cloning and expression of the recombinant polypeptides, the next essential step is to check the solubility of the proteins. This can easily be achieved by disrupting bacterial cells to release the soluble protein into the supernatant fraction and removal of insoluble cellular debris by centrifugation. Routinely two methods were used to disrupt bacterial host cells, these being French pressure treatment and sonication. The former method was chosen to aid the solubilisation of His-tagged proteins as it was a reproducible and more convenient method when dealing with large volumes of bacterial cultures.

The solubilities of all the recombinant proteins described in this chapter were checked using the same protocol described in Methods section 2. 2. 15. Aliquots of the lysate, supernatant and pellet fractions were then subjected to SDS-polyacrylamide gel electrophoresis and the solubility of the recombinant proteins was viewed by staining with Coomassie blue.

Figure 3.7 shows two SDS-polyacrylamide gels as examples of a solubility study involving the pea L-protein (panel A), a soluble protein and the pea L-protein precursor (panel B) an insoluble protein. The pea mitochondrial L-protein was shown to be entirely soluble by the presence of protein in the supernatant fraction (panel A, lane S), suggesting that the protein is completely folded (native state). In contrast its precursor form containing a N-terminal mitochondrial targeting sequence was shown to be insoluble and appeared entirely in the pellet fraction (panel B, lane P).

Expression of recombinant proteins at lower temperatures is a proven method for increasing the solubility in some cases; however, decreasing the temperature to 30°C or 22-25°C for the expression the L-protein precursor failed to produce soluble protein. Lowering IPTG concentrations may also slow the expression of protein, thereby allowing time for the correct folding of the expressed protein to occur in the cytoplasm. The expression of molecular chaperones (GroEL and GroES) within the host cells are known to aid protein folding. Chaperones bind to

newly translated proteins as either unfolded or partially-folded intermediates while present transiently in the cytoplasm. When they reach their final destination they finally fold into their native state generally after the cleavage of the targeting peptide. Preliminary expression experiments using GroEL and GroES failed due to unsuccessful expression of the chaperone proteins. However, it is unclear whether the presence of chaperones would increase L-protein precursor solubility as it contains the mitochondrial targeting sequence at the N-terminus which is thought to be responsible for maintaining the protein in a 'loosely-folded' translocation-competent state prior to mitochondrial import.

Parallel analysis of full-length mitochondrial and chloroplastic E2 and N-terminal didomain/tridomains expressed as GST-fusions indicated that overexpressed proteins were extremely soluble (not shown) with only minimal amounts detectable in the pellet fraction. Thus, the precursor form of pea mitochondrial L-protein was the sole exception in this regard.

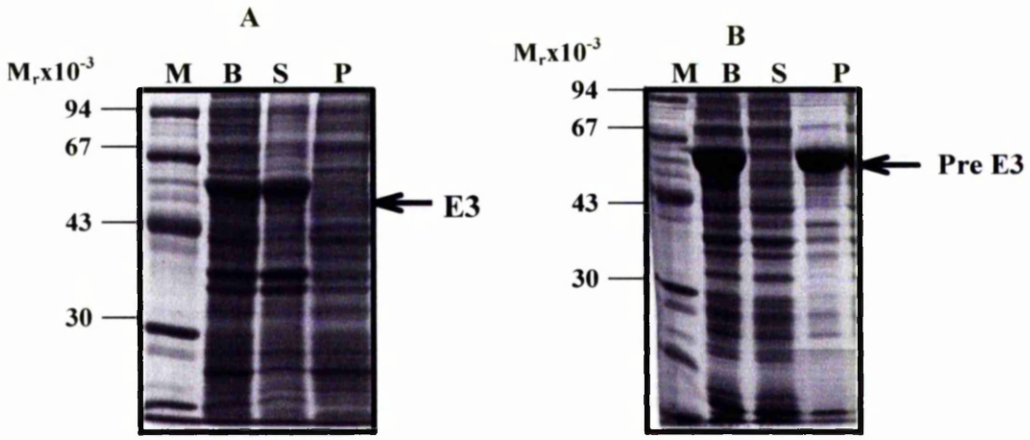


FIGURE 3.7: Solubility of Recombinant Pea Mitochondrial L-protein Expressed in *E. coli*

Bacterial cell extracts (B) were prepared by French pressure treatment and then centrifuged to yield supernatant (S) and pellet (P) fractions. Samples were analysed by SDS-PAGE with Coomassie blue staining. Equal volumes of each sample (5µl) was mixed with Laemmli sample buffer before being loaded onto the gel.

Panel A: Pea mature L-protein (M_r approx. 55, 000) present in the bacterial suspension and in the supernatant fraction, but not in the pellet fraction, indicating that the enzyme is soluble.

Panel B: Pea L-protein precursor (M_r approx. 59, 000) is present in the bacterial suspension but not in the supernatant, while it appears in the pellet fraction, indicating that the protein is insoluble.

M: relative molecular mass markers.

3. 7. DISCUSSION

The aim of this chapter was to clone a number of plant pyruvate dehydrogenase complex components and express them in *E. coli*. The availability of gene sequences coding for several enzymatic components of the mitochondrial and chloroplastic pyruvate dehydrogenase complexes made this goal achievable.

Cloning the Dihydrolipoamide Acetyltransferase

To permit detailed comparisons of the structures, organisation and functional properties of the chloroplastic and mitochondrial E2, to be performed, the PDC-E2 components from both organelles were cloned and overexpressed. The chloroplastic E2 (pea) was cloned as the mature and as a truncated form containing the N-terminal lipoyl domain and adjacent subunit binding domain (E2 didomain) into pET-14b and pGEX-2T, respectively. Successful high level expression of the E2 didomain as a GST-fusion protein and the mature E2 as a His-tagged protein was achieved as seen in Figure 3.5, lanes A and B.

Its mitochondrial E2 counterpart was obtained from an *A. thaliana* cDNA clone. Mitochondrial E2 was subcloned as the mature E2 and a truncated protein, consisting of the two N-terminal lipoyl domains and the subunit binding domain (tridomain). The DNA sequence encoding the tridomain was subcloned into pGEX-2T for the expression of the tridomain as a GST-fusion protein while the mature E2 was subcloned into pET-14b and expressed as a His-tagged protein as seen in Figure 3.5, lanes C and D, respectively. Availability of N-terminal regions containing catalytically active lipoyl domains and adjacent E1 and or E3 binding domains are useful tools for studying the E1-lipoyl domain recognition, lipoyl ligase specificity and subunit interactions in the assembly of the complex.

Cloning Dihydrolipoamide Dehydrogenase

Previously in this laboratory two dihydrolipoamide dehydrogenase (E3) isoforms were identified in potato mitochondria (Cook *et al.*, 1996). Work on these two

isoforms resulted in the isolation of cDNA for the E3β isoform; however, the gene encoding the E3α isoform remained elusive. The E3β protein had not been expressed previously but was achieved successfully, following its subcloning into pET-14b and subsequent production in *E. coli* as a His-tagged protein (Figure 3.6, lane A).

Pea E3 was found to be present as distinct organelle specific isoforms. At present only the pea mitochondrial E3 sequence has been identified (Turner *et al.*, 1992; Bourguinon *et al.*, 1992). Pea mitochondrial L-protein (E3) was cloned into pET-14b leading to the precursor and mature protein being expressed as His-tagged proteins in *E. coli* (Figure 3.6, lane B and C). The precursor was found to be insoluble (Figure 3.7, panel B) suggesting incomplete folding and the possible formation of non-specific aggregates or inclusion bodies. In contrast the mature L-protein was highly soluble (Figure 3.7, panel A).

Expression of Recombinant Proteins

Successful expression was observed for all recombinant proteins cloned in this chapter. Expressed proteins were visualised on SDS-PAGE. The PDC-E2 components displayed anomalous migration mobility behaviour with all the E2s resolving with a higher M_r value than predicted. This unusual migration pattern has been previously reported for PDC-E2 (Guan *et al.*, 1995; Thelen *et al.*, 1999b) and has been attributed to the unusual amino acid composition of the flexible linker regions which are rich in proline and alanine.

Expression of soluble recombinant proteins suggests that they are capable of folding into their native or near native state in the bacterial host. Their soluble nature also indicates that the presence of the N-terminal His-tag does not interfere with their folding or expression. The L-protein precursor has an additional mitochondrial targeting peptide (31 amino acids) and is insoluble, suggesting that the targeting peptide may prevent protein folding into its native state.

Recent Molecular Cloning of the *A. thaliana* PDC Components

Cloning the Pyruvate Decarboxylase

Access to *A. thaliana* genome sequence during this project permitted the cloning of the mitochondrial E1α and E1β subunits. This was achieved by producing genomic cDNA from freshly isolated mRNA from *A. thaliana* leaves with the use of reverse transcriptase-PCR (RT-PCR). The E1α subunit was cloned into pET-28b (see Appendix 3) the E1β subunit was cloned into pCR®T7/NT-TOPO® (see Appendix 4) permitting the expression of these subunits as His-tagged proteins.

Isoforms of Dihydrolipoamide Dehydrogenase

With the completion of the *A. thaliana* genome project, an unusual feature of the *A. thaliana* pyruvate dehydrogenase complex from mitochondria and chloroplast came to light. Four isoforms of E3 were identified (Figure 3.8); two predicted to be targeted to mitochondria and two towards the chloroplast (Lutziger and Oliver, 2000; 2001). All four genes were amplified from cDNA prepared from mRNA isolated from *A. thaliana* leaves and then subsequently subcloned into two vectors for the expression of each isoform as a His-tagged enzyme. The chloroplastic and mitochondrial *lpd1* genes were cloned into pET-14b and the chloroplastic and mitochondrial *lpd2* genes were cloned into the pCR® T7/NT®TOPO. Competent *E. coli* BL21 (DE3) pLysS cells were transformed with these recombinant plasmids and all four E3 isoforms were expressed. However, variable expression has been seen with the E3 encoded by chloroplastic *lpd1* gene expressing at high levels while the E3 encoded by chloroplastic *lpd2* gene and both mitochondrial E3s expressing at lower levels (not shown).

Successful expression has been observed for both E1 subunits and for all four E3 isoforms, however, due to time constraints no characterisation of these subunits or isoforms was possible.

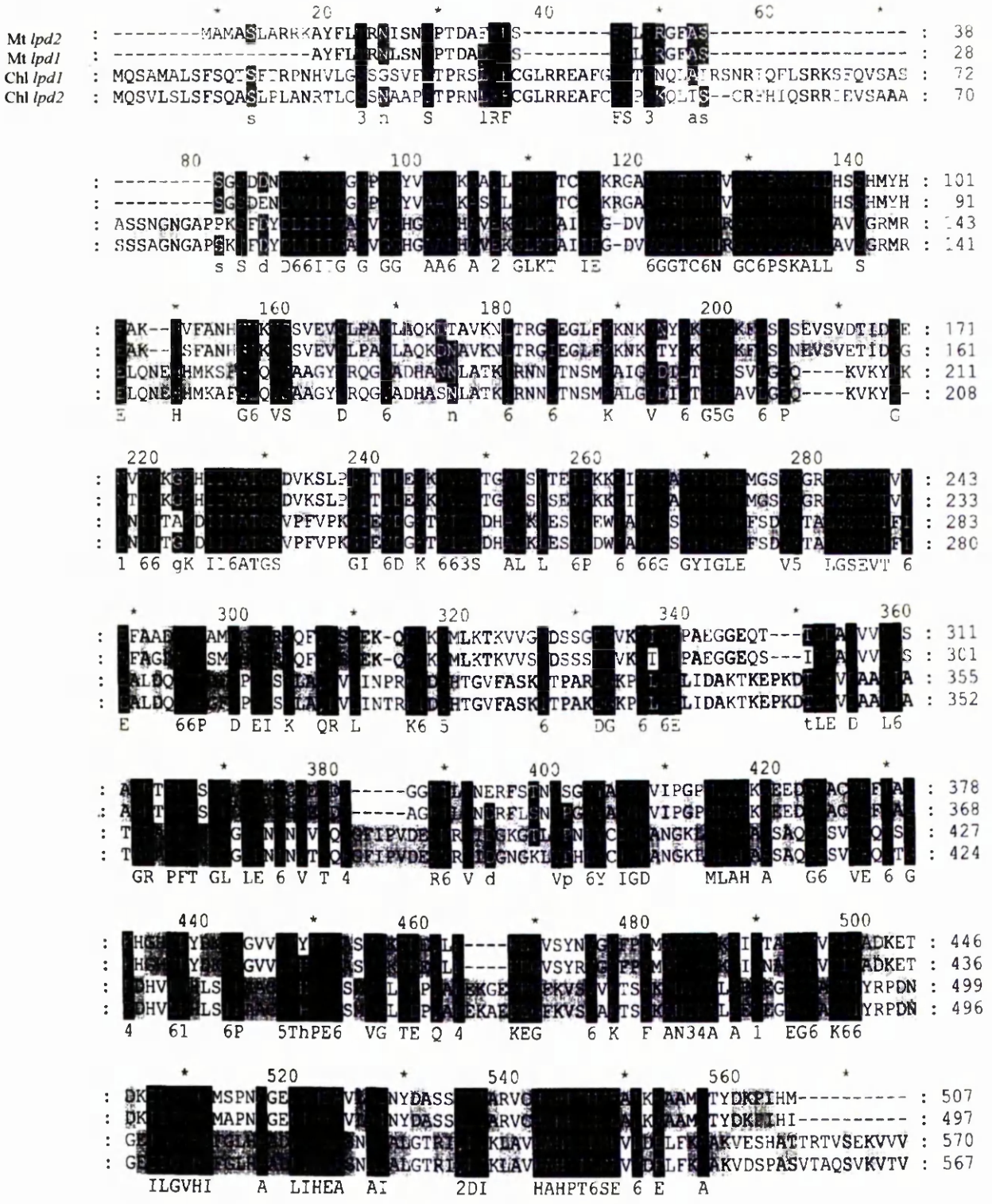


FIGURE 3.8: Comparison of the *A. thaliana* E3 Isoforms

With NCBI GeneBank accession numbers, AF228637 (Chl *lpd1*), AF228638 (Chl *lpd2*), AF228639 (Mt *lpd1*) and AF228640 (Mt *lpd2*).

Mt: mitochondrial, Chl: Chloroplastic

CHAPTER 4

PURIFICATION OF HIS-TAGGED ENZYMES

4. PURIFICATION OF HIS-TAGGED PROTEINS

4. 1. INTRODUCTION

Purification of the Plant Pyruvate Dehydrogenase Complex Components

In the past the pyruvate dehydrogenase was painstakingly purified by a multi-step method involving the homogenisation of large quantities of plant tissue, filtration and a series of differential centrifugation steps to produce intact mitochondria and chloroplasts (Williams and Randall, 1979). The ability to produce purified mitochondria and chloroplasts facilitated the detection of a distinctive PDC activity within the chloroplastic compartment for the first time. Subsequently plastid PDC was shown to be present in the chloroplast stroma and could be released in soluble form by osmotic rupture.

Early work on the plant PDC produced a series of publications reporting analyse of the enzymatic and regulatory properties of PDCs in crude or semi-pure preparations. More recently improved purification methods have been reported involving glycerol gradients to purify mitochondrial PDC from potato and maize (Millar *et al.*, 1998; Thelen *et al.*, 1998, respectively). These represent the highest purity preparations of plant PDCs reported to date; however the accurate composition of plant complexes still remains unclear as the purified maize PDC contained unindentified contaminant proteins visualised by Coomassie blue staining (Thelen *et al.*, 1998) whereas the purification of the potato PDC included the purification of a heat shock protein (Millar *et al.*, 1998). In addition to the contaminants, both reports provided evidence for the existence of multiple isoforms of PDC components preventing the definitive elucidation of PDC composition.

With the availability of genome sequences, EST databases and cDNA libraries from a variety of plant species cloning the PDC genes should allow the constituent enzymes to be individually expressed in *E. coli*. Cloning into vectors that contain appropriate tags such as glutathione S-transferase or His-tag permits the

expression of these proteins in a form that enables routine purification using affinity chromatography. Over-expression of these enzymes has the potential to produce protein in yields far greater than that possible from native plant sources. In future reconstitution of the purified components into fully assembled active complexes should be feasible permitting detailed analysis of their composition, subunit organisation, enzymatic properties and regulation.

As discussed in chapter three, recombinant PDC enzymes and specific domain combinations subcloned into the pET-14b or pGEX-2T vectors were expressed as His-tagged or GST-fusion proteins. The pET-14b vector permits the incorporation of six histidine residues unto the N-terminus of the polypeptide. In addition to the His-tag, a thrombin cleavage site is introduced C-terminal to the His-tag for removal if required. In this chapter, particular attention will be paid to the purification of these His-tag proteins, while similar studies on the GST-fusion proteins will be discussed in the following chapter.

Purification of His-tagged Proteins Using the BIOCAD SPRINT Perfusion Chromatography System

His-tagged enzymes were purified to near homogeneity on a medium to large scale (5-30mg) via specific binding of histidine residues to metal ions in an affinity chromatography step. A fast and effective protocol for purifying these proteins from the bacterial lysate employs a POROS® metal chelated resin column attached to a BIOCAD® SPRINT™ Perfusion Chromatography® System (PE Biosystems). This is a fully automated purification system, employing a pre-programmed purification scheme and plots the UV elution profile at 280nm against a number of parameters (e.g. % of elution buffer and pH). The column matrix consists of crosslinked poly(styrene-divinyl benzene) flow through particles that are surface coated with imidodiacetate groups ($CH_2N(CH_2CO_2^-)_2$) capable of binding a wide range of transition metals. The type of metal ion loaded onto the column affects the binding strength and selectivity towards the protein under investigation. Metal ions such as Cu^{2+}, Ni^{2+}, Zn^{2+} and Co^{2+} can be used for metal affinity chromatography with Cu^{2+} having the highest affinity and Co^{2+} displaying the weakest binding

property. Fe^{3+} and Al^{3+} ions can also be used but can be eluted from the column under the conditions required to elute protein. Metal chelate columns exhibit a high binding capacity (20mg/ml myoglobin using Cu^{2+} ions) for protein via binding to surface amino acids, notably the His-tag, although other significant interactions can involve cysteine and tryptophan side chains. Optimal purification was achieved by loading the column resin with metal ions before every run followed by saturation with buffer B (see Methods section 2. 2. 17). This replaces water molecules on the immobilised metal ions and increases the selectivity of purification. Prior to loading the sample the column was equilibrated with buffer A.

Initial trial purifications of the His-tagged enzymes utilised a POROS® MC column pre-loaded with Ni^{2+} ions; however, it was found that these E2s and E3s bound too tightly under these conditions. For this reason Zn^{2+} ions were routinely substituted as they display a lower binding affinity allowing these E2s and E3s to be recovered more readily in high yield.

4. 2. PURIFICATION OF MITOCHONDRIAL DIHYDROLIPOAMIDE DEHYDROGENASES

Sample Preparation

Recombinant potato mitochondrial E3 and L-protein of the pea mitochondria GDC were expressed as described in Methods section 2. 2. 11 in *E. coli* (DE3) pLysS. Proteins were released by disrupting bacterial cells by French pressure treatment (see Methods section 2. 2. 12A). An initial purification step was employed, exploiting the known heat stability of the E3 enzymes. The lysate was then incubated at 60°C for 30min and denatured proteins were removed by centrifugation at 10, 000 rpm at 4°C for 20min. The cleared lysate was applied manually onto the POROS® MC Zn^{2+}-imidoacetate column in the presence of buffer A. Elution of bound E3 was achieved by an increasing concentration gradient of imidazole as described in Methods section 2. 2. 17. The eluted enzyme was automatically collected in 2ml fractions. Peak fractions were then TCA

precipitated (see Methods section 2. 2. 21), and the purification was viewed by SDS-PAGE (see Methods section 2. 2. 14).

Results

4. 2. 1. Purification of the Potato Dihydrolipoamide Dehydrogenase β-isoform

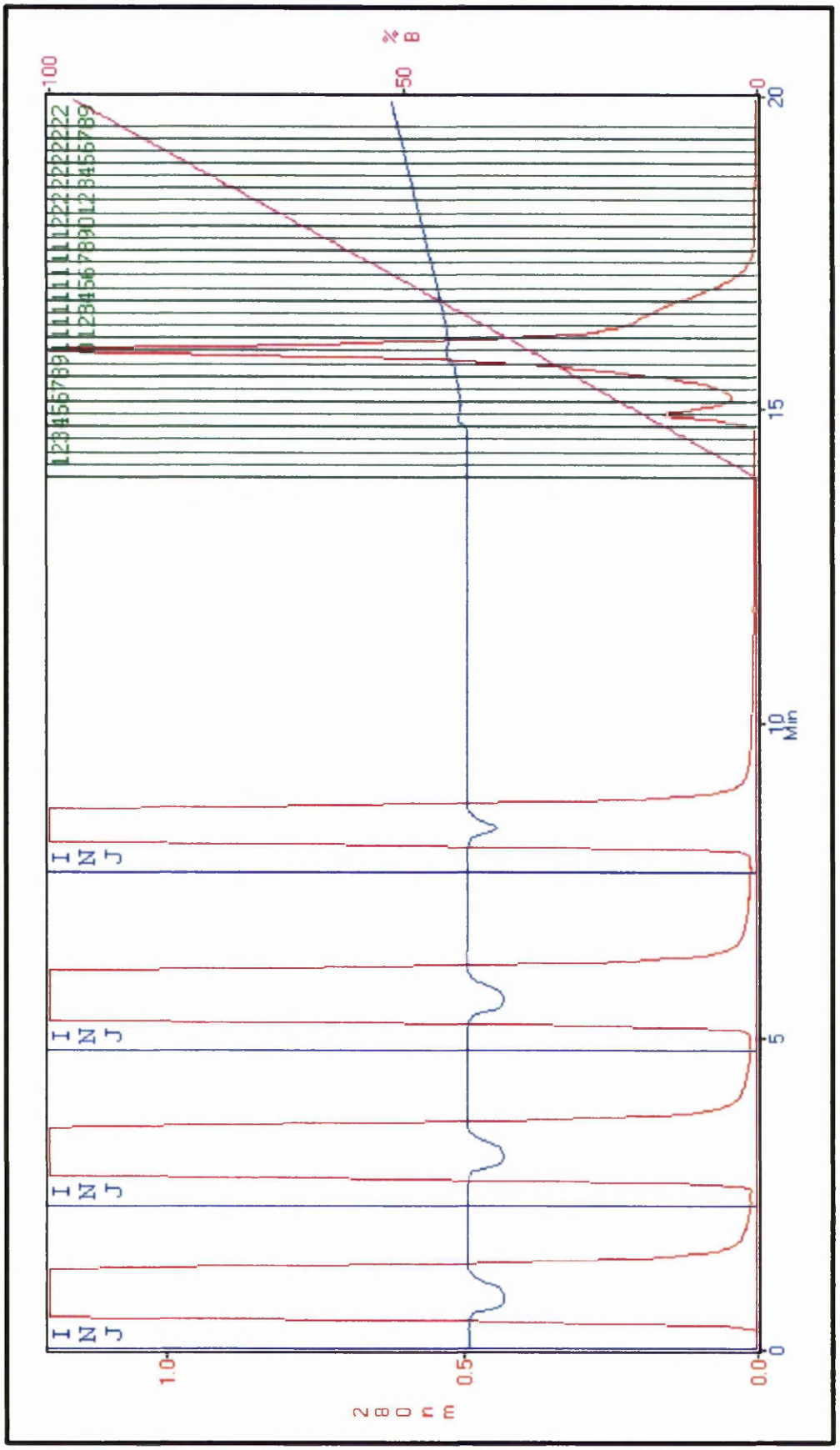
A purification run for the potato E3β isoform as recorded on the BIOCAD® SPRINT™ automated chromatography system is presented in Figure 4.1. Elution of protein washing off the column after each injection (bacterial proteins) and the subsequent specific elution of E3β during the imidazole gradient is depicted in red. The gradient (0.5mM-500mM imidazole) is shown in magenta while the fractions collected are shown in green. Peak fractions 4-15 were analysed by SDS-PAGE for E3 purity. As illustrated in Figure 4.2, the eluted enzyme has a M_r approx. 55, 000 corresponding to the potato E3β isoform. E3β was purified to near homogeneity as judged by Coomassie Blue staining although some minor contaminants were also present.

Enzymatic activity was assayed as described in the Methods section 2. 2. 20C. The activity and recovery of dihydrolipoamide dehydrogenase was measured at each step in the purification. The efficiency of the Zn^{2+}-affinity chromatography at purifying the His-tagged E3 from bacterial lysate is presented in the purification table, Table 4.1. A high recovery (82%) of expressed E3 from the bacterial lysate is achieved. The specific activity of the recombinant potato dihydrolipoamide dehydrogenase was calculated to be 293.8 μmol NADH min^{-1} mg^{-1} of E3 protein.

From three separate purification runs the potato E3 had an average specific activity of 315 ± 25 μmol NADH min^{-1} mg^{-1} of E3 with 80-95% of total expressed E3 being recovered. Routinely 20-30mg of E3 protein was purified from 1l of culture.

FIGURE 4.1: BIOCAD® Chromatogram of the Purification of the Potato Dihydrolipoamide Dehydrogenase

Purification of the potato mitochondrial E3β protein eluting from the Zn^{2+} ion affinity column shown in red as absorbance units at 280nm. The first four peaks correspond to the elution of non-interacting proteins from each sample injection. The linear gradient in magenta of 0%-100% buffer B represents the gradient of increasing imidazole concentration (0.5mM-500mM), where the E3 protein elutes in the range of 65-275mM imidazole.


% B
100
50
0
20
15
10
Min
5
0
I N J
I N J
I N J
I N J
1.0
0.5
0.0
280nm

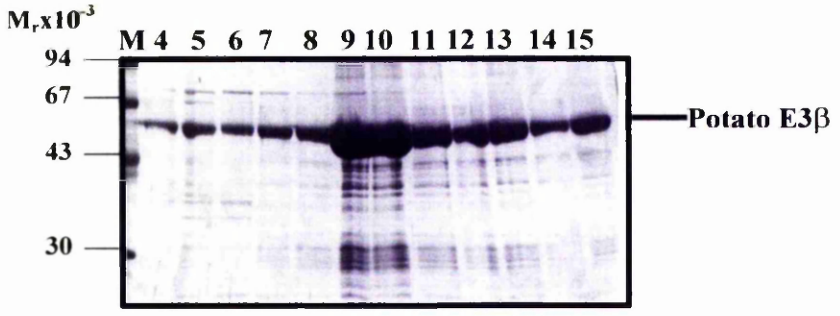


FIGURE 4.2: Purification of the Potato Mitochondrial E3β Isoform

SDS-PAGE analysis of the protein eluted in fractions 4-15. Lane numbers refer to the eluted fraction on the chromatogram (Figure 4.1). The eluted protein with a M_r approx. 55, 000 in fractions 4-17 corresponds to the E3 enzyme with a M_r of 52, 000 and was purified to near homogeneity as judged by Coomassie Blue staining. M: relative molecular mass markers

4. 2. 2. Purification of the Pea L-protein of GDC (PDC-E3)

The L-protein (E3) was purified from crude bacterial lysate following the two-step purification previously described for potato E3β. SDS-gel analysis of the pea L-protein purification is shown in Figure 4.3 A prominent protein band was observed in these fractions with a M_r approx. 55, 000 corresponding to the pea L-protein (M_r of 55, 897).

The dihydrolipoamide dehydrogenase activity of L-protein was measured using the same assay as potato E3 for each stage in the purification (see Table 4.2). An overall 56% of L-protein expressed was recovered and the specific activity was calculated as 246.7 μmol NADH min^{-1} mg^{-1} of L-protein. A yield of 5-10mg/1l culture was routinely obtained with an average specific activity of 200 ± 50 μmol NADH min^{-1} mg^{-1} of L-protein, close to the specific activity for native pea mitochondrial E3 of 269 ± 41 μmol NADH min^{-1} mg^{-1} of E3 (Conner *et al.*, 1996b). However this value is much greater than that recorded for the recombinant L-protein at 15.6μmol NADH min^{-1} mg^{-1} of L-protein (Bourguignon *et al.*, 1996).

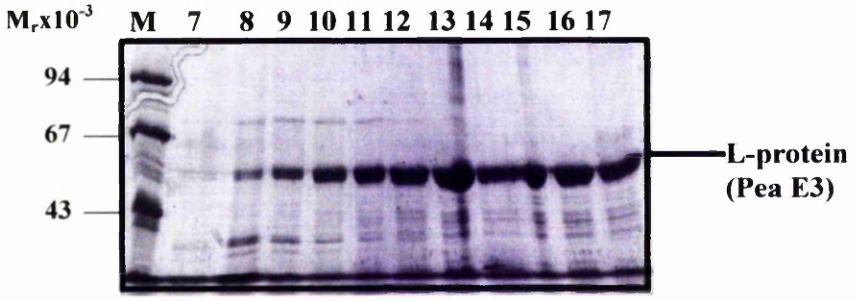


FIGURE 4.3. Purification of the Pea Mitochondrial L-protein of GDC

SDS-PAGE analysis of fractions 7-17 revealed the presence of a protein with a M_r approx. 55, 000 in fractions 8-17 corresponding to the L-protein with a M_r of 55, 897. L-protein (E3) was purified to near homogeneity as judged by Coomassie blue staining.

M: relative molecular mass markers

	Total Volume (ml)	Protein (mg/ml)	Total Protein (mg)	Activity (U/ml)	Total Activity (U)	Recovery (%)	Specific Activity (U/mg)	Purification Factor
Crude lysate	20	29.8	596	555.5	11110		18.6	
Post heat treatment	17.2	12.5	215	567.8	9766.2	87	45.5	7.8
Zn^{2+} column pool	24	1.3	31.2	382	9168	82	293.8	15

TABLE 4.1: Purification Table of the Potato Mitochondrial E3β Enzyme

Activity is measured as μmol NADH min^{-1} ml^{-1} of E3 and specific activity as μmol NADH min^{-1} mg^{-1} of E3. So 1 U = 1 μmol NADH min^{-1}

	Total Volume (ml)	Protein (mg/ml)	Total Protein (mg)	Activity (U/ml)	Total Activity (U)	Recovery (%)	Specific Activity (U/mg)	Purification Factor
Crude lysate	20	24.7	494	108.8	2176		4.4	
Post heat treatment	17.6	6.9	121.4	101.5	1786.4	82	14.7	3.3
Zn^{2+} column pool	18	0.3	5.4	74	1332	61	246.7	56

TABLE 4.2:Purification Table of the Pea Mitochondrial L-protein (E3)

Activity is measured as μmol NADH min^{-1} ml^{-1} of L-protein and specific activity as μmol NADH min^{-1} mg^{-1} of L-protein. So 1 U = 1 μmol NADH min^{-1}.

4. 3. PURIFICATION OF CHLOROPLASTIC & MITOCHONDRIAL DIHYDROLIPOAMIDE ACETYLTRANSFERASES (E2)

Sample Preparation for Affinity Chromatography

Recombinant pea chloroplastic E2 and *A. thaliana* mitochondrial E2 were routinely expressed in *E. coli* (DE3) LysS as described in Methods section 2. 2. 11. Recombinant proteins were released from the resuspended cells by disruption of bacteria employing French pressure treatment (see Methods section 2. 2. 12A). The lysate was cleared of insoluble material by centrifugation at 10, 000 rpm at 4°C for 20min, the supernatant was decanted and re-centrifuged as before. The cleared lysate was stored on ice until applied manually onto the POROS® MC Zn^{2+}-imidoacetate column in the presence of buffer A. Elution of bound E2 was achieved by an increasing concentration gradient of imidazole (0.5mM-500mM) as described in Methods section 2. 2. 17. Eluted material was collected and TCA precipitated, (see Methods section 2. 2. 21), before being resolved on SDS-PAGE (see Methods section 2. 2. 14).

Results

4. 3. 1. Purification of Pea Chloroplastic Dihydrolipoamide Acetyltransferase

His-tagged chloroplastic E2 eluted from the Zn^{2+}-imidoacetate column as a distinct sample peak. Eluted fractions were checked for the presence and purity of the His-tagged E2 by SDS-PAGE. Figure 4.4 shows the eluted material contains a dominant protein band with M_r approx. 50, 000 corresponding to the E2 (expected M_r of 47, 703). A doublet band pattern can be seen in fraction 8 and through 14-17 in this gel a possible consequence of partial lipoylation of plant E2 in the *E. coli* host (see section 5. 2. 2). Co-elution of contaminating proteins with the E2 can be seen and may occur either due to the bacterial proteins binding to the column (via regions rich in histidine, cysteine and tryptophan residues) or directly to the E2.
For enzymatic, biological and particularly structural investigation it is necessary that the E2 is purified to as near homogeneity as possible. As a way of removing

these contaminants, ion exchange chromatography was employed. Fractions containing chloroplastic E2 were pooled and the enzymatic activity was accessed by measuring S-acetyl dihydrolipoamide formation at 232nm, as described in the Methods section 2. 2. 20B. Assessing dihydrolipoamide acetyltransferase activity in the bacterial lysate proved difficult and unreliable possibly as a consequence of high background absorbance from bacterial proteins or low M_r UV-absorbing compounds. The enzyme activity is presented in Table 4.3. Specific activities of 2.5 -2.7 ΔA_{232} min^{-1} mg^{-1} were obtained, these values are considerably lower than those published for the maize mitochondrial E2, 191 ΔA_{232} min^{-1} mg^{-1} (Thelen *et al.*, 1999b) and recombinant human E2 of 19.4 ΔA_{232} min^{-1} mg^{-1} (Yang *et al.*, 1997). The difference in specific activities may be a consequence of differing methods of quantifying the E2 activity although similar assays were employed. Reported specific activities for this enzyme, from a variety of sources are extremely variable partly accounted for by the fact that differing assays and assay conditions are also often employed. Thus it was unclear whether the values obtained in this study were lower than expected. In addition, time constraints prevented further optimisation of the assay conditions for these plant E2s and it is also possible that the presence of the His-tag had a detrimental effect on the E2 activity.

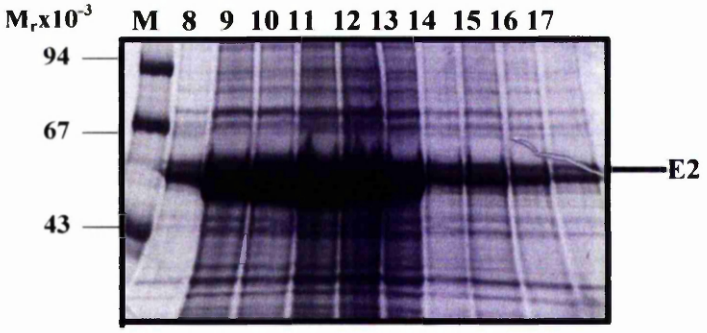


FIGURE: 4.4: Purification of the Pea Chloroplastic E2 By Zn^{2+} Affinity Chromatography

The major eluted protein species present in these eluted fractions has a subunit M_r approx. 50, 000 corresponding to the predicted value for E2 (M_r of 47, 703). A number of minor high and low M_r contaminant bacterial proteins co-elute with the His-tagged enzyme.

M: relative molecular mass markers

4. 3. 1. 1. Cation Exchange Chromatography of Pea Chloroplastic E2

Cation exchange chromatography was performed using a 4.6mm/100mm POROS® 20S column attached to a BIOCAD® SPRINT™ Perfusion Chromatography® System. The POROS® 20S column media is functionalized with sulphoethyl groups ($CH_2CH_2SO_3^-$), a strong cation exchanger. In cation exchange chromatography, the increase in ionic strength of the mobile phase is employed to promote the displacement of bound proteins depending on the affinity of the individual protein for the column matrix.

Sample Preparation

Purified His-tagged E2 was pooled, dialysed and loaded onto the cation exchange column in the presence of 20mMNaCl, 50mM KPi, pH 7.0 and eluted with NaCl (20mM-1.0M) in the same buffer as described in Methods section 2. 2. 18A. Eluted enzyme was TCA precipitated (see Methods section 2. 2. 21) and viewed by SDS-PAGE (see Methods section 2. 2. 14).

Results

The majority of co-eluting bacterial proteins from the Zn^{2+}-imidoacetate column were not retained on the cation exchange column, thus were effectively removed from the E2 sample. Pea chloroplastic E2 elutes midway through the NaCl gradient as a distinct protein peak and the individual fractions were analysed for their E2 purity (Figure 4.5). The major protein with a M_r approx. 50, 000 corresponding to E2 (M_r of 47, 703) was purified to near homogeneity as judged by Coomassie Blue. Some minor contaminants or possibly degradation products were apparent at high loading. E2 activity was measured as described in the Methods section 2. 2. 20B, no improvement was seen in the specific activity of the E2 after this second chromatographic step. Although the ion exchange chromatography step removed most of the contaminating proteins, the majority of the remaining E2 is enzymatically inactive (see Table 4.3).

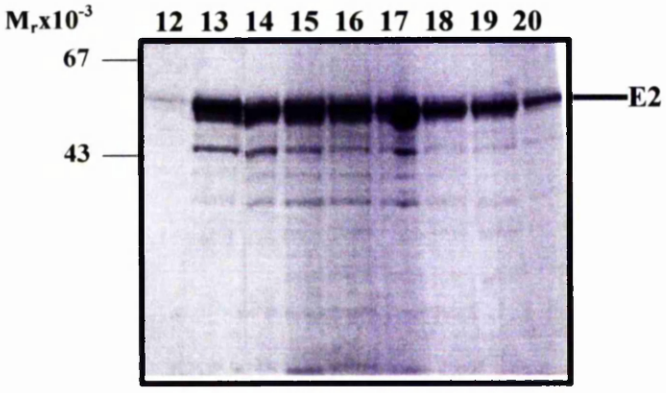


FIGURE 4.5: Cation Exchange Purification of the Pea Chloroplastic E2

The purity of E2 eluted from the cation exchange column was judged by Coomassie blue staining after SDS-PAGE. A major protein band with a M_r approx. 50, 000 was visible corresponding to E2 (M_r of 47, 703).

M: relative molecular mass markers

4. 3. 2. Purification of *A. thaliana* Mitochondrial Dihydrolipoamide Acetyltransferase

Results

A. thaliana E2 was purified from crude bacterial lysate following a similar purification regime to that previously described for chloroplastic E2. *A. thaliana* E2 eluted from the POROS® MC Zn^{2+}-imidoacetate column as a broad peak; this may indicate heterogeneity of binding of large oligomeric E2 to the column. Peak fractions were resolved on SDS-PAGE to identify the E2 rich fractions and check their purity (Figure 4.6). The appearance of a protein with a M_r approx. 70, 000 in the peak fractions corresponded to E2 (expected M_r of 62, 674). As seen in Figure 4.6, several minor bacterial contaminants co-eluted with E2 but were partly removed from the E2 sample by anion exchange chromatography.

Dihydrolipoamide acetyltransferase activity was assessed by measuring S-acetyl dihydrolipoamide formation at 232nm following the assay protocol described in the Methods section 2. 2. 20B and used to construct a purification table (Table 4.4). Enzyme activity detected in the lysate sample was difficult to measure and varied from assay to assay. However, from the activities measured the specific activity of the E2 was calculated to be $0.8\,\Delta\,A_{232}$ min^{-1} mg^{-1} which, similar to the chloroplastic E2 specific activity, was appreciably lower than published values. As mentioned before, the specific activities reported are variable and may be a consequence of differing assay methods and lack of optimisation of assay conditions. However, the differences may also be attributable to the difficulty of the E2 assay as it relies on the formation of the substrate immediately before the E2 source is added. Thus, variations in K_m of each individual E2 for acetyl-CoA may have a profound effect on enzymatic activity.

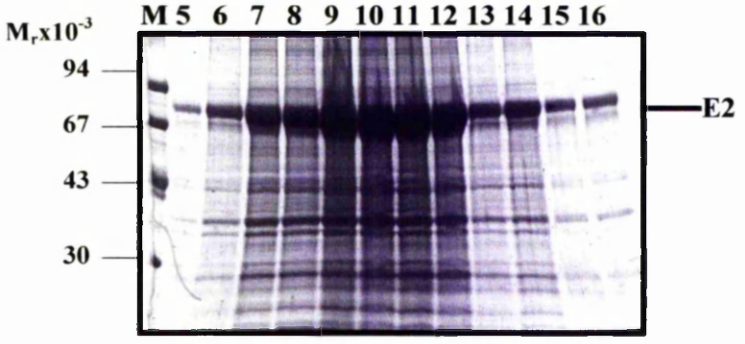


FIGURE 4.6: Purification of the *A. thaliana* Mitochondrial E2 By Zn^{2+} Affinity Chromatography

The major protein band on the SDS-PAGE of the precipitated fractions has a M_r approx. 70, 000 corresponding to *A. thaliana* E2 (M_r of 62, 674). Contaminating bacterial proteins were seen co-eluting in these fractions.

M: relative molecular mass markers

4. 3. 2. 1. Anion Exchange Chromatography of *A. thaliana* Mitochondrial E2

Anion exchange chromatography was performed using a POROS® HQ column attached to the BIOCAD® SPRINT™, Perfusion Chromatography® system. The HQ matrix is composed of crosslinked poly(styrene-divinylbenzene) resin particles that allow rapid movement of molecules. These particles are surface coated with polyethlyeneimine. Proteins are selectively eluted by increasing the salt concentration of the mobile phase.

Sample Preparation

A. thaliana E2 samples purified from the Zn^{2+}-imidoacetate step were prepared, loaded onto the anion exchange column and eluted with an increasing NaCl gradient (20mM-1M) as described in Methods section 2. 2.18B. Eluted protein was TCA precipitated (see Methods section 2. 2. 21) and viewed by SDS-PAGE (see Methods section 2. 2. 14).

Results

Bacterial proteins contaminating the mitochondrial E2 sample were effectively removed by anion exchange chromatography. Figure 4.8 shows the purity of E2 eluted from the anion exchange column as judged by Coomassie blue staining. Although some contaminants remain the overall purity of the E2 sample had improved markedly. The major band with an apparent M_r of 70, 000 corresponded to the mitochondrial E2, with a predicted M_r of 62, 674.

Purified fractions from anion exchange purification were pooled and the dihydrolipoamide acetyltransferase activity was measured as described in the E2 assay protocol in the Methods section 2. 2. 20B and recorded in Table 4.4.

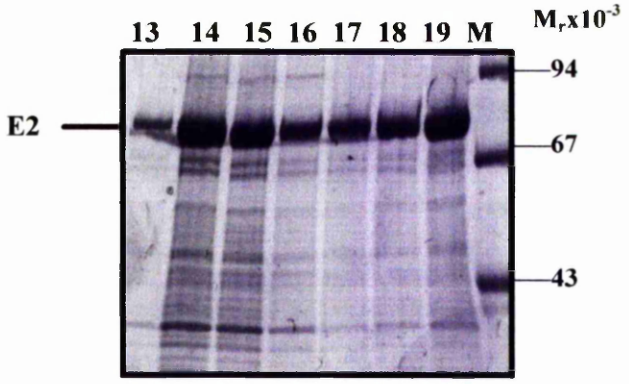


FIGURE 4.7: Anion Exchange Purification of the *A. thaliana* Mitochondrial E2

The protein present in peak eluted fractions were resolved on SDS-PAGE and the purity of E2 was judged by Coomassie blue staining. E2 was purified to near homogeneity especially in the later fractions as indicated by the protein band with a M_r approx. 70, 000 referring to E2 (predicted M_r value is 61, 368).

M: relative molecular mass markers

	Total Volume (ml)	Protein (mg/ml)	Total Protein (mg)	Activity (U/ml)	Total Activity (U)	Recovery (%)	Specific Activity (U/mg)	Purification Factor
Crude lysate	17.8	38.8	690.6	nd	nd	nd	nd	nd
Zn^{2+}-affinity pool	24	0.8	19.2	2.0	48.9		2.5	
Ion-exchange pool	18	0.6	10.8	1.6	28.8	58	2.7	1

TABLE 4.3: Purification Table of the Pea Chloroplast E2

Activity is measured as ΔA_{232} ml^{-1} of E2 so the specific activity is measured as ΔA_{232} min^{-1} mg^{-1} of E2.

1 U= 1 ΔA_{232} min^{-1}, nd, no data obtained for these values.

	Total Volume (ml)	Protein (mg/ml)	Total Protein (mg)	Activity (U/ml)	Total Activity (U)	Recovery (%)	Specific Activity (U/mg)	Purification Factor
Crude lysate	17.4	36	626.4	nd	nd	nd	nd	nd
Zn^{2+}-affinity pool	24	0.9	21.6	0.7	16.8		0.8	
Ion-exchange pool	9	0.8	7.2	0.9	8.1	48	1.1	1.4

TABLE 4.4: Purification Table of the *A. thaliana* Mitochondrial E2

Activity is measured as ΔA_{232} ml^{-1} of E2 so the specific activity is measured as ΔA_{232} min^{-1} mg^{-1} of E2.

1 U = 1 ΔA_{232} min^{-1}, nd, no data obtained for these values.

4. 4. PURIFICATION OF INSOLUBLE PROTEIN FROM INCLUSION BODIES

As shown in the previous chapter expression of the pea L-protein precursor resulted in the formation of insoluble aggregates. Expression of heterologous polypeptides in *E. coli* can often result in the accumulation of inclusion bodies. In many instances the appropriate environment to promote correct protein folding may not be present and so proteins may become misfolded or adopt a number of intermediate folding conformations (Thatcher *et al.*, 1996) with an increased tendency to form aggregates. Exposure of hydrophobic residues, not normally accessible in the native conformation can allow non-productive interactions between nascent intermediates giving rise to insoluble products. These misfolded intermediates/aggregates could result from the lack of post translational modification to complete the folding/assembly pathway, or the lack of suitable chaperones to aid the correct folding pathway.

Another possibility is the over-production of these recombinant proteins within bacterial cells. The protein levels may exceed the threshold concentration to which the protein can accumulate before aggregating or precipitating (Hartley *et al.*,1998). Thus the association of misfolded/ folded intermediates or the precipitation of protein can all lead to the formation of inclusion bodies.

Inclusion bodies formed within the cytoplasm of the bacterial cells are visible by light microscopy and are typically 0.2-1.5μm in size and usually present as a single structure per cell. They consist almost entirely of the recombinant protein, which is assumed to be densely packed aggregates of misfolded protein although some aggregates have been found to possess biological activity, suggesting that they may have a significant level of native secondary or tertiary structure. This can be advantageous in the purifying of these proteins by protecting them from proteases and permitting the purified protein to be solubilised in its active state or subjected to *in vitro* refolding in attempts to recover high levels of functionally mature protein.

Centrifugation and selective washing of insoluble aggregates achieved the recovery of inclusion bodies from *E. coli* cells. Inclusion bodies can then be solubilised by using various reagents such as guanidinium chloride (GdmCl) or urea, anionic/ cationic detergents or exposure to extremes of pH and then subjected to *in vitro* refolding. The mild detergent N-lauroylsarcosine, (Frangioni *et al.*, 1993) does not solubilise all inclusion bodies but, in some cases, can select those containing proteins in their native or near native state without interfering with their biological activity.

Sample Preparation

Inclusion bodies containing the pea L-protein precursor were used to investigate the effect of detergent on solubilisation of the protein. L-protein precursor was expressed in *E. coli* (DE3) pLysS as described in Methods section 2. 2. 11. Harvested cells were resupended in 20mM Tris-HCl buffer pH 7.5, 10mM EDTA, 1% (v/v) Triton X-100, and disrupted by French pressure treatment (see Methods section 2. 2. 13) prior to centrifugation. Inclusion bodies were purified and suspended in 50mM KPi buffer pH 7.5, (final concentration of 20mg/ml) as described in the Methods section 2. 2. 13. The inclusion body suspension was mixed with N-lauroylsarcosine for 15min at room temperature and then centrifuged (10, 000rpm for 10min) to separate the soluble supernatant from the insoluble pellet as described in Methods section 2. 2. 13.

In addition to N-lauroylsarcosine, a variety of different detergents including lauryl maltoside, β-octylglucoside (BOG) and lauryldimethylamine oxide (LDAO) were also investigated for their ability to solubilize inclusion bodies.

Results

Solubilisation of inclusion bodies containing the L-protein precursor was achieved only with N-lauroylsarcosine indicated by the presence of a protein band with a M_r approx. 55, 000 corresponding to the L-protein precursor appearing in the supernatant fraction produced from the solubilisation of inclusion bodies. No

effect on inclusion body solubility (results not shown) was observed with the other three detergents over the same range of concentrations.

Optimisation of the detergent concentration was investigated for N-lauroylsarcosine over the range, 0%-2% (w/v), (results not shown). It appeared that 0.5% (w/v) N-lauroylsarcosine was sufficient to solubilise the L-protein precursor as judged by its release from the pellet (lane P) into the detergent supernatant (lane D) as detected by SDS-PAGE and Coomasie blue staining, shown in Figure 4.8. Future studies of the L-protein precursor employed 0.5% (w/v) N-lauroylsarcosine.

The L-protein precursor was assayed for E3 activity; however, no activity was detected. The presence of the targeting sequence, in addition, to preventing complete folding seems to abolish E3 activity. The lack of E3 activity was inconsistent with results obtained for the human truncated E3 precursor where the E3 activity measured was half that of the human mature E3 activity (Kim *et al.*, 1991). The difference may be a result of the pea L-protein precursor containing the complete targeting sequence while the human truncated precursor contains only the C-terminal region of the targeting sequence.

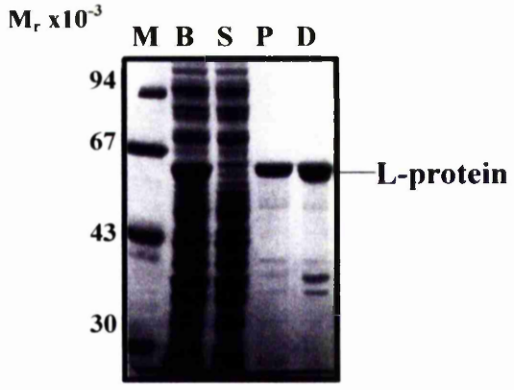


FIGURE 4.8: SDS-PAGE Analysis of the Solubilisation of the Pea L-protein Precursor

Bacterial lysate (B) was subjected to French pressure treatment and centrifuged yielding a soluble supernatant (S), no L-protein present, but found in the insoluble pellet (P) fraction indicating the insoluble nature of the protein. Inclusion bodies (pellet) were subsequently treated with 0.5% (w/v) N-lauroylsarcosine, thereby releasing the L-protein into the supernatant (D). Equal volumes of each fraction (5μl) were loaded on a 10% SDS gel and visualised by Coomassie blue staining.
M: relative molecular mass markers

4. 5. DISCUSSION

The aim of the work described in this chapter was to purify individual constituent enzymes of plant mitochondrial and chloroplastic PDC as recombinant His-tagged proteins from *E. coli*. As demonstrated the E2 and E3 components have been expressed in soluble form and were subsequently purified to near homogeneity by Zn^{2+}-affinity chromatography. Affinity chromatography employs a Zn^{2+}-imidodiacetate (POROS® MC resin) column connected to an automated BIOCAD® SPRINT™ Perfusion Chromatography® System enabling rapid purification of large quantities of recombinant protein from 1l of bacterial culture. The availability of sufficient quantities of highly purified enzymes will permit basic biochemical and biological studies.

Purification of the Dihydrolipoamide Dehydrogenases

Potato mitochondrial PDC-E3 and pea mitochondrial GDC L-protein were purified from crude bacterial lysates in a single chromatographic step facilitated by an initial heat treatment (60°C for 30min). High levels of purified E3 were obtained with recoveries of active enzyme ranging from 10-30mg of potato E3β isoform and 5-10mg of pea L-protein per litre of culture. Variations in yield were probably a consequence of variable expression rather than loss of protein during the heat treatment. Specific activities obtained for the recombinant E3 enzymes were 315 ± 25 μmol NADH min^{-1} mg^{-1} for potato E3 and 200 ± 50 μmol NADH min^{-1} mg^{-1} for pea L-protein consistent with the published activities of 280-350 μmol NADH min^{-1} mg^{-1} of E3 for the native pea mitochondrial PDC E3 enzyme (Conner *et al.*, 1996c).

However, these specific activities differed markedly from the value determined by Bourguignon and co-workers (1996) as 15.6μmol NADH min^{-1} mg^{-1}. A similar discrepancy was seen for the specific activity for porcine E3, where Bourguignon measured 15.2 μmol NADH min^{-1} mg^{-1} and 324 ± 39 μmol NADH min^{-1} mg^{-1} was

determined in our laboratory (Conner, 1996c). This inconsistency between values may result from differing enzymatic assay conditions or possible calculation errors.

Purification of Dihydrolipoamide Acetyltransferases

E2 from pea chloroplastic PDC and *A. thaliana* mitochondrial PDC were purified in two chromatographic steps involving affinity chromatography and then ion exchange chromatography. Zn^{2+}-affinty chromatography was used to purify both E2s from crude bacterial lysates although a number of contaminating bacterial proteins co-eluted with the E2s. These bacterial contaminants may bind directly to the Zn^{2+}-imidodiacetate functional groups on the column matrix via regions rich in histidine, cysteine and tryptophan residues or to the bound E2s. Removal of the majority of these contaminants was achieved by ion exchange chromatography. Chloroplastic E2 appeared to be markedly purer than the original affinity purified enzyme; moreover, after a final dialysis step and concentration step. The removal of low M_r contaminants was aided by the concentration step using centrifugal filters with M_r cut off of 30, 000, leaving a near homogeneous E2 sample as judged by Coomassie blue staining. A few contaminants were present on a heavily loaded SDS-gel (Figure 4.9, lane 1) possibly minor degradation products generated during storage.

A. thaliana mitochondrial E2 was also purified further from the contaminating proteins that co-eluted during affinity chromatography by ion exchange chromatography. This time an anion exchange matrix was employed, as the E2 did not appear to bind to the cation exchange column. An enriched E2 sample was obtained although a few bacterial proteins remainined as judged by Coomassie blue staining. After extensive dialysis and concentration the E2 enzyme displayed a high degree of purity although some low M_r proteins were visible on a heavily loaded SDS-gel in Figure 4.9 (lane 2).

Measuring E2 activity in crude bacterial lysate proved to be difficult and unreliable. This could be attributable to the high background absorbance from bacterial

proteins or UV absorbing compounds in the extract masking small changes in absorbance at 232nm.

A specific activity of 2.7 ΔA_{232} min^{-1} mg^{-1} was calculated for the pea chloroplast E2 enzyme while a figure of 1.2 $\Delta A_{232} min^{-1}$ mg^{-1} was calculated for the *A. thaliana* mitochondrial E2 enzyme. These values are much lower than the single published specific activity for recombinant maize E2 which was 191 ΔA_{232} min^{-1} mg^{-1} (Thelen *et al.*, 1999). However, this value is extremely high as recombinant human E2 had a reported specific activity of 19.4 ΔA_{232} min^{-1} mg^{-1} (Yang *et al.*, 1997). Comparison of the specific activities calculated for the pea and *A. thaliana* E2s to those published might suggest that these recombinant E2s, although soluble, are relatively inactive. However, both E2s are soluble, are assembled into high M_r oligomers (see sections 6. 3. 3 and 6. 3. 5) and are fully lipoylated (see sections 5. 3. 1 and 5. 3. 2) consistent with their correct folding and assembly. The difference may be due low substrate levels or assay conditions not optimised. Calculation error has to be considered as there is a massive difference in specific activities between the plant and human enzymes measured by two separate groups. In retrospect it would have been beneficial to devote more time and effort into optimising E2 assay conditions for the plant enzymes. It should be noted in this particular assay relies on rapid generation of acetyl-CoA via acetyl-phosphate and CoA. Thus the respective K_m of individual E2s for their acetyl-CoA substrate will have a profound influence on the level of E2 activity.

The four His-tagged mature mitochondrial and chloroplastic E2 and E3 enzymes were purified to near homogeneity as seen in Figure 4.9. The aim of this chapter was to achieve a method for obtaining large quantities of purified recombinant protein for each PDC component (see Table 4.5). This has been successfully achieved with purification yield in the range of 5-30mg of recombinant protein per litre of culture. In addition to the large yields, the recombinant enzymes show biological activity although further work is required to optimise the E2 assay conditions.

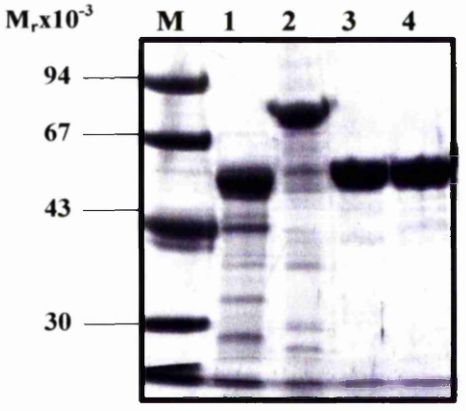


FIGURE 4.9: Purified His-tagged Enzymes of Plant Mitochondrial and Chloroplastic PDCs

Purified His-tagged proteins (15μg) were resolved on a 15% SDS polyacrylamide gel and the proteins were viewed by Coomassie blue staining.

Lane1: pea chloroplastic E2 with an M_r of 47, 703 but resolves at a higher M_r approx. 50, 000

Lane 2: *A. thaliana* mitochondrial E2 with a M_r of 62, 674 resolving at approx. 70, 000

Lane 3: potato mitochondrial E3β isoform with a M_r of 52, 015

Lane 4: pea mitochondrial L-protein (E3) with a M_r of 55, 897

M: relative molecular mass markers

Expressed Protein	Expression vector	Amount/ litre of culture	Specific Activity (U/mg)
Potato Mt E3	pET-14b	20-30mg	315 ± 25
Pea Mt E3	pET-14b	5-10mg	200 ± 50
Pea Mt E3 precursor	pET-14b	*	
Pea Chl E2	pET-14b	5-10mg	2.7
A. thaliana Mt E2	pET-14b	5-10mg	1.2

TABLE 4.5: Recombinant His-tagged Enzymes of Plant Mitochondrial and Chloroplastic PDCs

1U = 1 μmol NADH min^{-1} mg^{-1} E3

1U =1 ΔA_{232} units min^{-1} mg^{-1} E2

*, undetermined

Purification of Insoluble Proteins

Insoluble pea mitochondrial L-protein precursor was purified from inclusion bodies formed on the expression of the recombinant protein The insolubility of the L-protein precursor is not surprising as a major role of the 31 amino acid targeting sequence is to delay folding of the precursor. The L-protein may possibly be completely unfolded or in a loosely-folded tranlocation competent state until it reaches the mitochondria, where the target sequence is cleaved off during or shortly after entry allowing the mature protein to fold in the matrix compartment. Solubilising inclusion bodies with detergents is an inexpensive but effective method for producing soluble protein or aiding *in vitro* refolding and reconstitution to achieve active protein in large yields. N-lauroylsarcosine was successful in releasing the L-protein precursor from the inclusion bodies as determined by the presence of the L-protein in the detergent soluble supernatant by SDS-PAGE. Optimisation of detergent concentration found that 0.5% (v/v) was sufficient to promote complete solubilisation of inclusion bodies.

As to why N-lauroylsarcosine (Figure 4.10) is effective in solubilising inclusion bodies is not clearly understood, but its action may be attributable to the long acyl chain or perhaps the ionic head group having some protein stabilisation properties. These properties may aid the disruption of hydrophobic interactions between partially folded intermediates and on occasion permit assembly to continue along the native folding pathway.

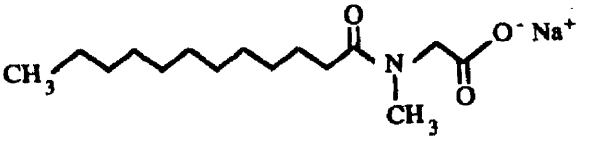


FIGURE 4.10: Chemical Structure of N-lauroylsarcosine

An alkyl anionic detergent (sodium salt) M_r 293.4

(Frangioni and Neel, 1993)

CHAPTER 5

INVESTIGATING THE LIPOYLATION OF PLANT E2s

5. INVESTIGATING THE LIPOYLATION OF PLANT E2

5. 1. INTRODUCTION

The oligomeric dihydrolipoamide acyltransferase (E2) components of the 2-oxo-acid dehydrogenase complex family play a crucial role, not only in providing the framework for assembly but also serving as the central enzyme of the overall reaction. The E2 component is post-translationally modified by the covalent attachment of its cognate prosthetic group (lipoic acid) to a specific lysine residue. Modification occurs within the N-terminal independently folded lipoyl domains. These domains have a characteristic flattened β-barrel structure housing the lysine residue at the tip of a type 1β-turn. Lipoic acid is linked via its carboxylic acid moiety to the ε-amino group of the lysine in an amide linkage. This lipoyl-lysine arm protrudes from the domain and acts as an acyl carrier visiting the catalytic sites on E1 and E3 within the pyruvate dehydrogenase complex.

Lipoic acid (6, 8-thioctic acid) is a sulphur containing cofactor found in many prokaryotic and eukaryotic organisms although its biosynthetic pathway is poorly understood (Millar *et al.*, 2000). *E. coli*, yeasts, and plants can synthesise lipoic acid endogenously or scavenge it from the environment. Several lines of evidence suggest that octanoic acid is a precursor for lipoic acid (Figure 5.1) and indeed octanoylation of recombinant lipoyl domains is possible under lipoic acid deficient conditions.

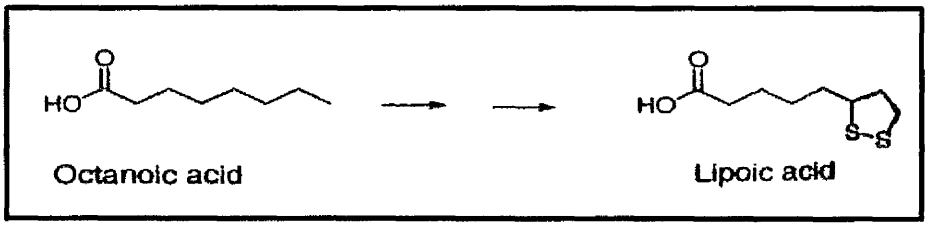


FIGURE 5.1: Structures of Octanoic Acid and Lipoic Acid
(Taken from Miller *et al.*, 2000)

In higher eukaryotes the transfer of lipoate to the apo-proteins requires the action of two enzymes: a lipoate-activating enzyme, that catalyses the formation of lipoyl-AMP and a lipoyl-AMP lipoyl ligase, which transfers the activated prosthetic group to the specific lysine residue on the apo-protein. Two distinct lipoate-protein activities were identified in *E. coli*, LPL-A and LPL-B, requiring ATP and Mg^{2+} and able to utilise L-lipoate and octanoyl adenylate (Brookfield *et al.*, 1991). However, the LPL-B gene product is the only one that can employ octanoic acid as substrate and was therefore thought to be responsible for the octanoylation of recombinant lipoyl domains expressed in *E. coli*. Exogenous lipoic acid is transferred to the unlipoylated apo-protein in an ATP-dependent process by lipoate-protein ligase A, LplA (LPL-A). In contrast, endogenous lipoate is transferred to the apo-protein by the *lipB* gene product lipB or LPL-B, a lipoyl-(acyl-carrier-protein)-protein-N-lipoyl transferase using the lipoyl-acyl carrier protein as the source of the lipoyl groups as shown in Figure 5.2 (Miller *et al.*, 2000).

In plants where lipoic acid containing complexes are located in both mitochondria and plastids, it would be reasonable to assume that both organelles would have a source of lipoyltransferase activity and lipoic acid. This assumption was substantiated with the identification of a mitochondrial lipoyltransferase (LIP2) that catalyses the transfer of lipoyl groups from lipoyl-acyl carrier protein to the apo-enzyme in *A. thaliana* (Wada *et al.*, 2001a). The plastid counterpart was subsequently identified as a lipoyl-ACP dependent lipoyltransferase (LIP2p) with homology to *E. coli* LPL-B (Wada *et al.*, 2001b).

As *E. coli* contain two lipoylating enzymes, potentially they should be capable of post-translationally modifying recombinant E2 proteins expressed within them. However, the expression of the recombinant *Bacillus stearothermophilus* N-terminal lipoyl domain in *E. coli* resulted in 80% of the expressed domains remaining unlipoylated while the remainder was apparently lipoylated and capable of being reductively acetylated by E1. However, 4% of the 'lipoylated' protein carried an aberrant substituent, identified by electrospray mass spectrometry as octanoic acid (Dardel *et al.*, 1990).

The expression of the inner lipoyl domain of the human pyruvate dehydrogenase complex in *E. coli* (Quinn *et al.*, 1993) also produced unlipoylated and lipoylated forms of the domain. This partial lipoylation of recombinant lipoyl domains in *E. coli* was observed not only for heterologous domains but was also observed for the expression of a hybrid *E. coli* lipoyl domain in *E. coli* (Ali and Guest, 1990). When lipoyl domains were expressed in mutant cells defective in lipoic acid biosynthesis and grown in lipoic acid deficient conditions, the lipoyl domains were also found to be modified by octanoylation on the lipoylatable lysine residue (Ali *et al.*, 1990).

The over-expression of various lipoyl domains in *E. coli* produced a mixture of unlipoylated and lipoylated products which suggested that over-expressing these domains can exceed the capacity of the cell for lipoylation (Brookfield *et al.*, 1991). Octanoylation represents a novel protein modification that occurs when lipoic acid is limiting and supports the idea that octanoic acid is a precursor of lipoic acid. However, when exogenous lipoic acid is added to the culture media the lipoyl domains were completely lipoylated indicating that the synthesis of lipoic acid is the limiting factor in lipoylation and that a salvage pathway exists to utilise exogenous cofactor.

Plant chloroplastic and mitochondrial E2 have already been shown to be expressed in *E. coli*. Taking into account the expression of other heterologous lipoyl domains in *E. coli*, it was important to determine if the expression of the equivalent plant domains would also result in the production of unlipoylated and lipoylated products. In this chapter, the lipoylation of chloroplastic E2 and mitochondrial E2 is investigated using the recombinant pea chloroplastic E2 didomain and the *A. thaliana* mitochondrial E2 tridomain as substrates for the *E. coli* lipoylation machinery.

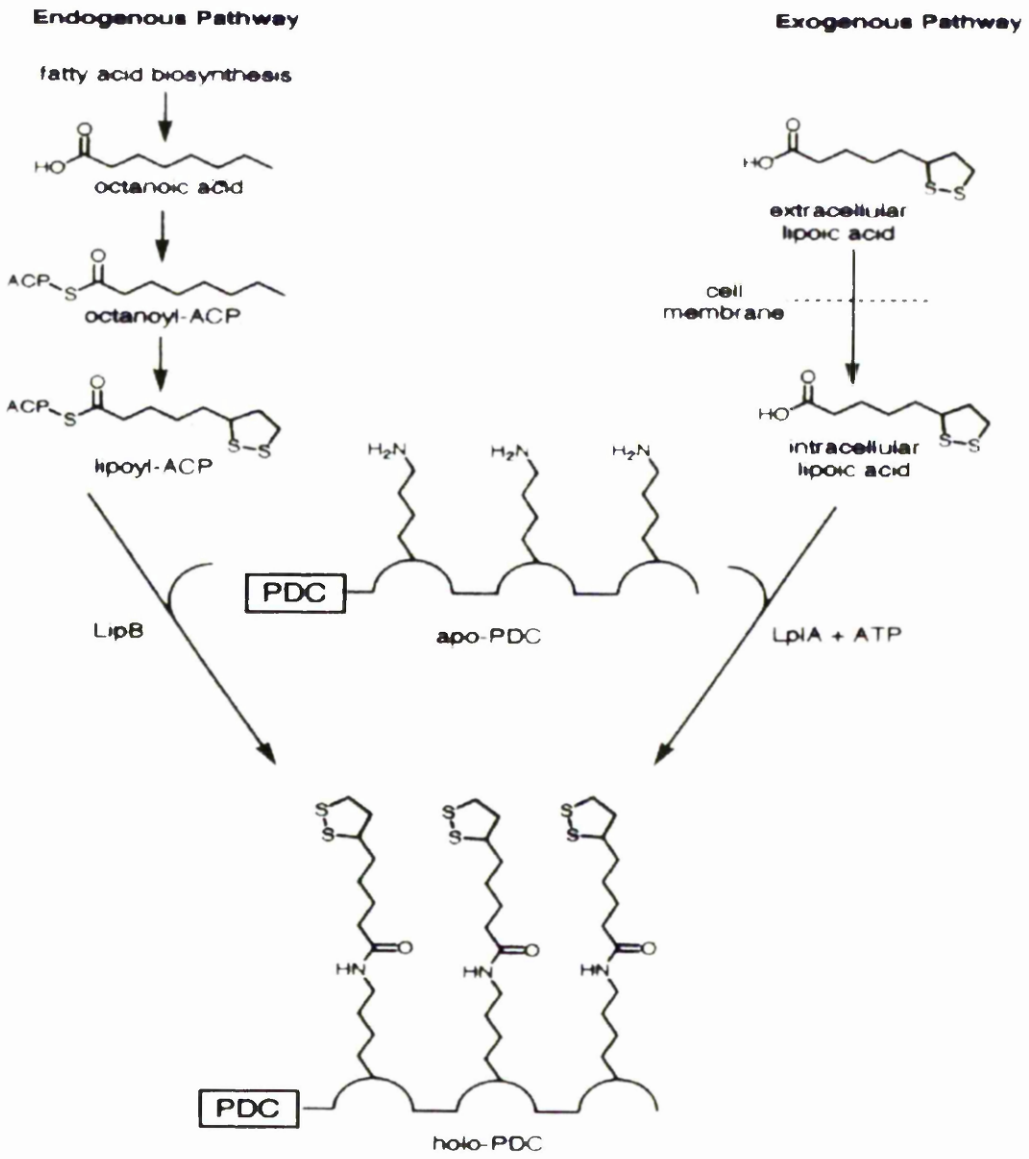


FIGURE 5.2: Schematic Representation of the Lipoylation Pathways in *E. coli.*

(Taken from Miller *et al.*, 2000)

5. 2. INVESTIGATING PLANT E2 LIPOYL DOMAINS

5. 2. 1. The 3D Structure of the Pea Chloroplastic Lipoyl Domain

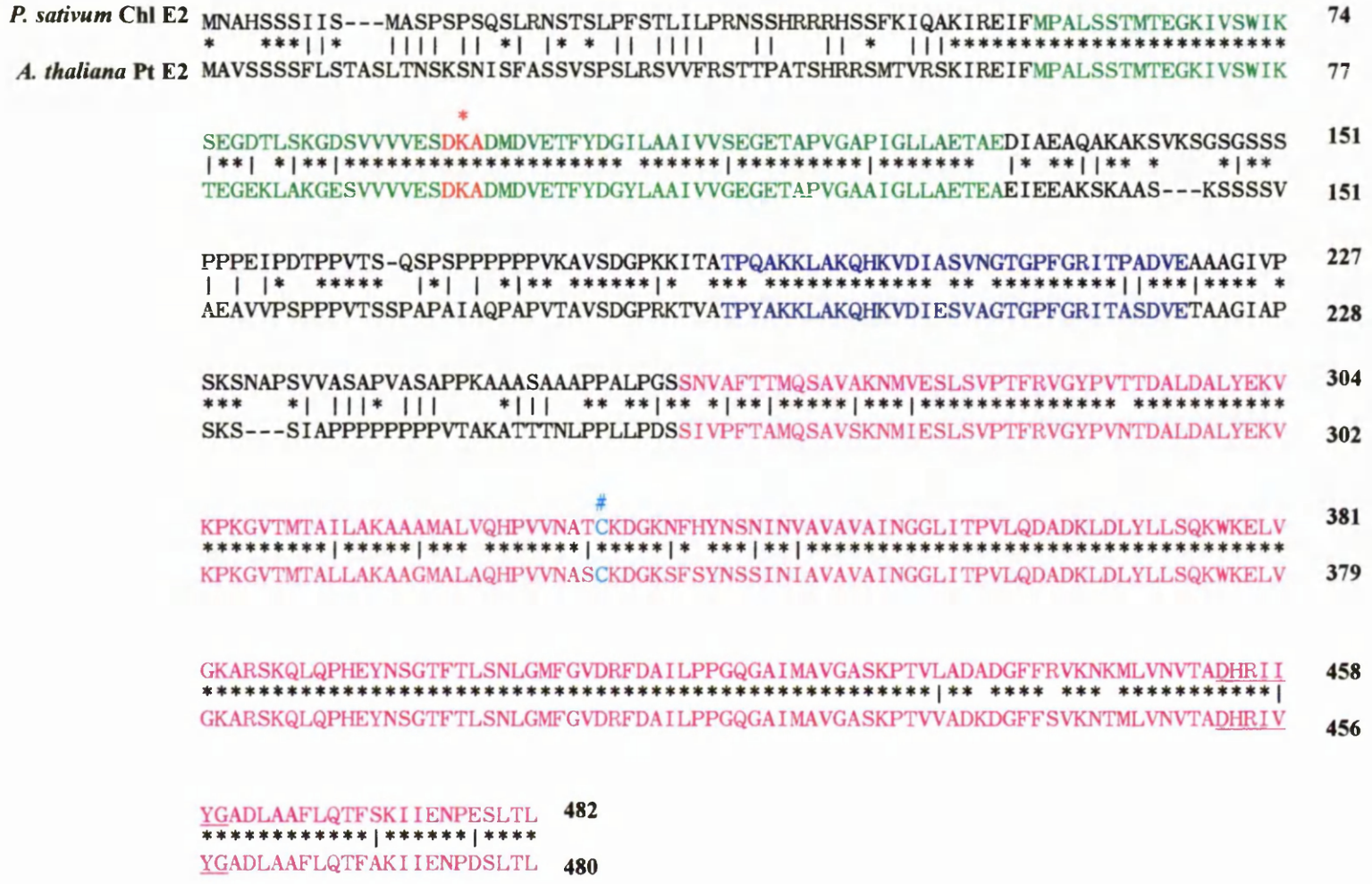
The pea chloroplast dihydrolipoamide acetyltransferase gene/protein sequence has not yet been fully characterised as a plastidic isoform. In order to identify the origin of the E2 sequence generously provided by Jurgen Söll, the deduced protein sequence was entered into the automated search engine, BLAST (Basic Local Alignment Search Tool). The deduced pea E2 protein sequence was employed to search databases of known proteins to find possible matches. The BLAST search aligned the pea E2 sequence with other plant E2 sequences. Alignment of pea E2 with the *A. thaliana* plastidic E2 sequence was found to exhibit the highest identity (Figure 5.3).

The 3D structure of its single lipoyl domain was predicted by submitting the protein sequence to SWISS MODEL an Automated Protein Modelling Server, which aligns protein sequences with similar/identical proteins with known crystal structures. The N-terminal 90 residues of the pea chloroplast E2 were aligned with three other sequences including human inner lipoyl domain of PDC-E2; lipoyl domain of OGDC-E2 from *A. vinelandii* and lipoyl domain of OGDC-E2 from *E. coli*. Figure 5.4 shows the predicted 3D structure for the lipoyl domain of the pea lipoyl domain (pink) aligned with the 3D structure for the human inner lipoyl domain (green) which had the highest sequence identity (37%).

Both protein sequences are practically superimposed with the position of the lysine residue at the tip of the type 1β-turn between the two sets of β-sheet. The relevant lipoyl-lysine residue for the pea E2 is Lys 44 while in the human lipoyl domain the lipoyl-lysine residue is Lys 50. Each lysine residue has an aspartic acid residue to the N-terminal side and an alanine residue to the C-terminal side of the lysine giving the highly conserved lipoylation motif -DKA-.

FIGURE 5.3: Sequence Alignment of the Deduced *P. sativum* Chloroplast E2 with the *A. thaliana* Plastid E2

The sequences presented are of *P. sativum* chloroplast E2 and *A. thaliana* plastid E2. The * denotes the lysine residue likely to be post-translationally modified by the addition of the lipoic acid cofactor. Three conserved domains of the *A. thaliana* E2, the lipoyl domain, the subunit binding domain, and the catalytic domain, * denotes the identical residues. # denoted the only cysteine residues in the sequence and the conserved motif of the active site of acetyltransferases is underlined.


P. sativum Chl E2
A. thaliana Pt E2
MNAHSSSIIS---MASPSPSQSLRNSTSLPFSTLILPRNSSHRRRHSSFKIQAKIREIFMPALSSTMTEGKIVSWIK 74
MAVSSSSFLSTASLTNSKSNISFASSVSPSLRSVVFRSTTPATSHRRSMTVRSKIREIFMPALSSTMTEGKIVSWIK 77
SEGDTLSKGDSVVVVESDKADMDVETFYDGILAAIVVSEGETAPVGAPIGLLAETAEDIAEAQAKAKSVKSGSGSSS 151
TEGEKLAKGESVVVVESDKADMDVETFYDGYLAAIVVGEGETAPVGAAIGLLAETEAEIEEAKSKAAS---KSSSSV 151
PPPEIPDTPPVTS-QSPSPPPPPPVKAVSDGPKKITATPQAKKLAKQHKVDIASVNGTGPFGRITPADVEAAAGIVP 227
AEAVVPSPPPVTSSPAPAIAQPAPVTAVSDGPRKTVATPYAKKLAKQHKVDIESVAGTGPFGRITASDVETAAGIAP 228
SKSNAPSVVASAPVASAPPKAAASAAAPPALPGSSNVAFTTMQSAVAKNMVESLSVPTFRVGYPVTTDALDALYEKV 304
SKS---SIAPPPPPPPPVTAKATTTNLPPLLPDSSIVPFTAMQSAVSKNMIESLSVPTFRVGYPVNTDALDALYEKV 302
KPKGVTMTAILAKAAAMALVQHPVVNATCKDGKNFHYNSNINVAVAVAINGGLITPVLQDADKLDLYLLSQKWKELV 381
KPKGVTMTALLAKAAGMALAQHPVVNASCKDGKSFSYNSSINIAVAVAINGGLITPVLQDADKLDLYLLSQKWKELV 379
GKARSKQLQPHEYNSGTFTLSNLGMFGVDRFDAILPPGQGAIMAVGASKPTVLADADGFFRVKNKMLVNVTADHRII 458
GKARSKQLQPHEYNSGTFTLSNLGMFGVDRFDAILPPGQGAIMAVGASKPTVVADKDGFFSVKNTMLVNVTADHRIV 456
YGADLAAFLQTFSKIIENPESLTL 482
YGADLAAFLQTFAKIIENPDSLTL 480

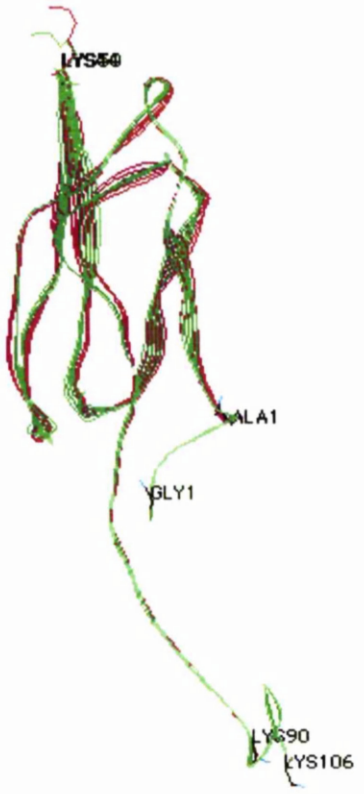


FIGURE 5.4: SWISS MODEL Structure of the *P. sativum* Chloroplast Lipoyl Domain

The predicted 3D structure of the pea chloroplast lipoyl domain (pink) was achieved by SWISS MODEL. The inner lipoyl domain of the human PDC-E2 (green) was found to have the highest sequence identity to the pea lipoyl domain. Prediction of the 3D structure of the pea chloroplast E2 shows the characteristic β-barrel structure of a lipoyl domain. The position of the lipoyl-lysine residue for each protein is indicated along with the N-terminal and C-terminal residues.

SWISS MODEL web site: w.w.w.expasy.org/swissmod/SWISS-MODEL.html and viewed using the Swiss-Pdb Viewer.

This similarity in protein sequence and structure together provide evidence to suggest that the pea E2 sequence contains a single lipoyl domain and that the cDNA encodes for the appropriate plastidic dihydrolipoamide acetyltransferase (E2) component.

E. coli was capable of expressing the chloroplastic E2; however, it was unknown if the *E. coli* lipoylation machinery would recognise plant E2 enzymes leading to their efficient post-translational modification as would be the case in their native environment. The first step into investigating the lipoylation potential of these proteins was to express them in growth media in the presence or absence of exogenous lipoic acid.

5. 2. 2. Overexpression of E2 Lipoyl Domains With or Without Exogenous Lipoate

The lipoyl domain(s) and subunit binding domain of pea chloroplastic E2 (didomain) and *A. thaliana* mitochondrial E2 (tridomain) were expressed in the *E. coli* BL21 (DE3) pLysS strain as glutathione S-transferase fusion proteins as described in Methods section 2. 2. 10.

The expression of the pea chloroplastic GST-E2 didomain grown in the presence (0.1mM) and the absence of lipoic acid is shown in Figure 5.5. After 1h no difference was observed between the expression of GST-E2 didomain in the two cultures. By 2h a distinct doublet band pattern was present in the minus lipoic acid sample while in the plus lipoic acid sample, a dominant upper band was visible. As the only difference between the two cultures was the presence of lipoic acid, it was surmised that the prominent upper band observed in lane 2h (+) and 3h (+) on the SDS-gel corresponds to the lipoylated form of E2. The results of the expression without lipoic acid suggest a mixture of lipoylated and unlipoylated E2 products was produced, with the upper band being the lipoylated form seen in lanes 2h (+) and 3h (+) and the lower band being the unlipoylated form seen in lanes 2h (-) and 3h (-) of Figure 5.5. Visual inspection of the gels suggest that approx. 50% lipoylation is acheived in the absence of exogenous cofactor consistent with partial

lipoylation levels reported for the human E2 inner lipoyl domain (Quinn *et al.*, 1993).

Clear separation of bands in SDS-PAGE corresponding to the unlipoylated and lipoylated forms of the pea chloroplast GST-E2 didomain was variable and not always evident. However, expression of the *A. thaliana* mitochondrial GST-E2 tridomain routinely produced only one band in both cultures as shown in Figure 5.6. Mobility shifts and the presence of two separate distinct bands such as those observed for the E2 didomain have also been observed for human E3-binding protein (E3-BP) didomain (Richards, 1999) but not for the equivalent E2-didomain comprising the inner lipoyl domain and adjacent subunit binding domain.

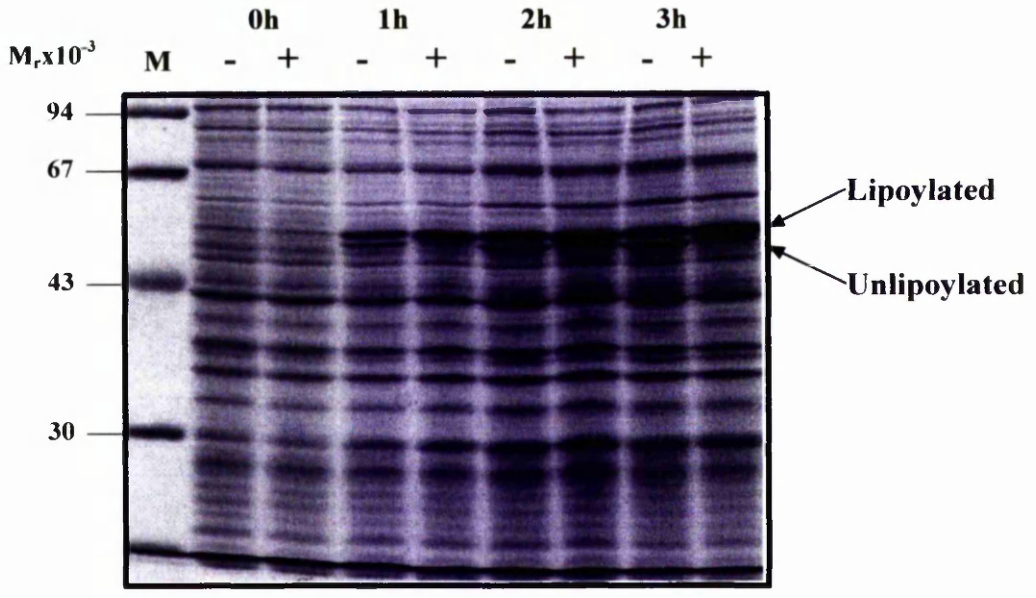


FIGURE 5.5: Expression of the Pea Chloroplastic GST-E2 Didomain in the Presence and Absence of Lipoic Acid

SDS-PAGE analysis of the expression of the pea chloroplast GST-E2 didomain grown in LB media plus/minus exogenous lipoic acid. Samples taken 0h, 1h, 2h and 3h after the addition of IPTG. Expression was observed after 1h, and by 2h it had reached its maximum judged by the appearance of a protein band with a M_r 50, 000 (GST-E2 didomain, 48, 219). The doublet band pattern seen in the 2h (-) and 3h (-) samples represent the lipoylated and unlipoylated forms, while in the 2h(+) and 3h (+) lanes a predominant single band is observed, probably representing the lipoylated form.

M: relative molecular mass markers.

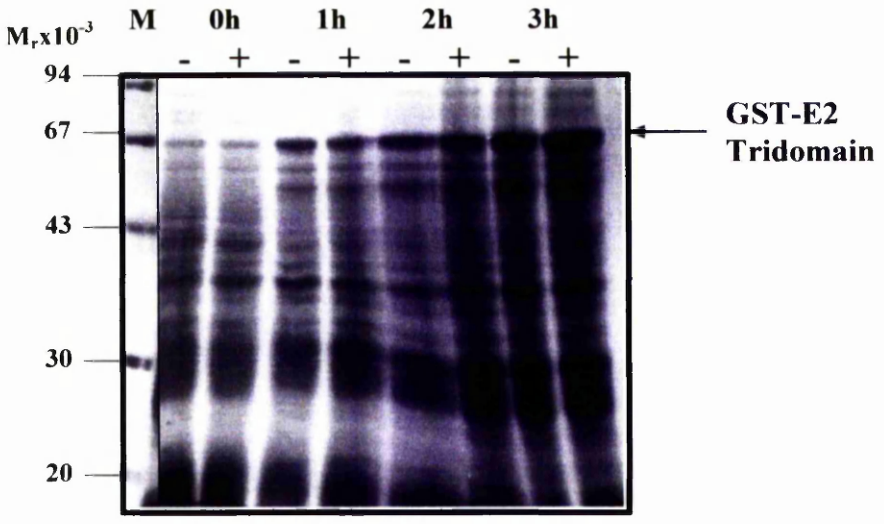


FIGURE 5.6: Expression of the *A. thaliana* Mitochondrial GST-E2 Tridomain Recombinant Protein in the Presence and Absence of Lipoate

SDS-PAGE analysis of the expression of the *A. thaliana* mitochondrial GST-E2 tridomain protein in media in the presence (+) or absence (-) of exogenous lipoic acid. Samples were taken at hourly intervals after the addition of IPTG. Expression was observed after 1h, and by 3h maximum levels was reached as judged by the appearance of a protein band with a M_r of 67, 000 (GST-E2 tridomain, M_r of 62, 674).

M: relative molecular mass markers.

5. 2. 3. Purification of the GST-E2 Didomain and the GST-E2 Tridomain

GST-fusion proteins of the chloroplastic E2 didomain and the mitochondrial E2 tridomain were expressed in media with and without exogenous lipoic acid as described in Method section 2. 2. 10.

GST-fusion proteins were purified by affinity chromatography using a glutathione Sepharose-4B (2ml x 1cm) column. Purification was carried out at room temperature although the samples and buffers were kept at 4°C. Bacterial lysate was prepared and clarified as described in Methods section 2. 2. 16. GST-E2 didomains and GST-E2 tridomains were purified from this clarified bacterial lysate following the purification procedure described in the Methods section 2. 2. 16. The GST-E2 didomain and GST-E2 tridomain proteins were specifically eluted and collected in 2ml fractions prior to SDS-PAGE analysis to view their purification.

Purification of the pea chloroplastic GST-E2 didomain expressed in the presence of exogenous lipoate is shown in Figure 5.7, panel A. GST- E2 didomain was completely soluble as shown by its presence in the supernatant (lane S) and absence from the pellet (P). This was also true for the pea GST-E2 didomain (-) lipoic acid and both *A. thaliana* GST-E2 tridomain proteins (+ and - lipoate). GST-E2 didomain was eluted over eight fractions and was purified to near homogeneity as judged by Coomassie blue staining. In Figure 5.7, panel B the purification of the *A. thaliana* mitochondrial GST-E2 tridomain plus lipoate is presented. GST-E2 tridomain with an expected M_r of 61, 368 eluted from the column in the first four fractions and migrates with a M_r value of approx. 70, 000 as observed previously.

The GST-E2 didomain and the GST-E2 tridomain expressed without lipoic acid were also purified by affinity chromatography. All GST-fusion proteins were purified to near homogeneity as judged by Coomassie blue staining, pooled and dialysed into Millipore water followed by PEG 6000 concentration (see Methods section 2. 2. 23).

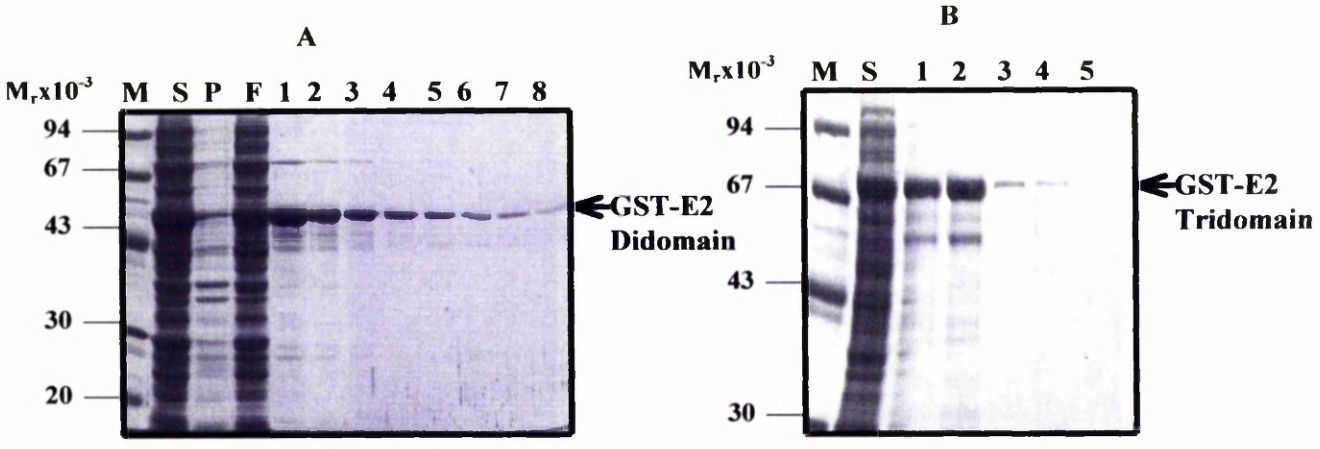


FIGURE 5.7: Purification of the Pea Chloroplastic GST-E2 Didomain and the *A. thaliana* Mitochondrial GST-E2 Tridomain

SDS-PAGE analysis shows the purification of both proteins grown in media supplemented with lipoic acid. The purity of the protein samples were judged by Coomassie blue staining. The supernatant (lane S) incubated on the GST-affinity column, followed by the eluted fractions (numbered).

Panel A: a significant amount of pea chloroplast GST-E2 didomain approx. M_r 50, 000 was present in the flow through buffer (lane F) probably as a result of column saturation. The GST-E2 didomain was seen eluting specifically in the following 8 fractions.

Panel B: the *A. thaliana* GST-E2 tridomain was also soluble, as evidenced by its presence in supernatant (S). The majority of the protein (M_r of approx. 70, 000) was eluted in the first two fractions.

M: relative molecular mass markers

5. 3. IDENTIFYING LIPOYLATED E2 LIPOYL DOMAINS

5. 3. 1. Immuno-detection of E2 Lipoylation

Two distinct bands were visible from the expression of the chloroplastic E2 didomain suggesting possible lipoylation of the domain. To establish if this was the case, lipoylation was investigated using immuno-detection with enhanced chemiluminesence (ECL™) to specifically identify lipoylated domains.

Access to monoclonal antibodies (mAb PD1/PD2) that exclusively recognise the holodomain form of the human inner E2 lipoyl domain, (generously provided by Prof. F. Stevenson, University of Southampton) permitted similar analysis of the plant E2 holodomains. Previous investigation into the ability of these antibodies to recognise lipoylated domains using purified human unlipoylated and lipoylated domains, showed that the antibody displayed an absolute specificity for the lipoylated form (Richards, 1999). However, it was not known if the antibody recognised E2s from diverse eukaryotic sources, as it does not cross-react with human OGDC-E2 or BCDC-E2 indicating that the antibody recognises other characteristics found specifically in PDC-E2 and E3-BP lipoyl domains.

For this reason, mAb (PD2) was used in conjunction with immobilised plant E2s, to examine its potential for assessing their lipoylation state. Cross-reaction was visualised using an anti-human IgG antibody (secondary antibody) conjugated with horseradish peroxidase followed by ECL™ (Amersham) detection as described in the Methods section 2. 2. 15.

Purified chloroplastic GST-E2 didomain and mitochondrial GST-E2 tridomain (expressed in growth media with or without lipoate) were resolved on SDS-polyacrylamide gels along with the purified mature His-tagged E2 equivalents (plus lipoate). Recombinant human PDC-E2 and E3-binding protein were used as positive controls. Samples (approx. 1μg) were resolved on duplicate gels. One gel was stained with Coomassie blue seen in Figure 5.8, (panel A), while the resolved

proteins on the duplicate gel were electrophoretically transferred onto nitrocellulose paper overnight (see Methods section 2. 2. 15).

In Figure 5.8, (panel A), the migration of the chloroplastic E2 GST-didomains and mature E2 (lanes1, 2 and 3, respectively) is shown as well as the mitochondrial GST-E2 tridomains and mature E2 (lanes 4, 5 and 6, respectively). Mature plant E2s were grown in the presence of lipoic acid, as was the human E2 and human E3-binding protein (lanes 7 and 8, respectively). The presence of two bands in the E3-BP sample (lane 8) is due to spill-over from the adjacent lane containing human E2. From the immuno-blot (Figure 5.8, panel B)it is clear that the mAb does not recognise the chloroplastic E2 either as the truncated GST-E2 didomain or in the mature form. However, the mitochondrial E2 was recognised by the monoclonal antibody (lanes 4, 5 and 6) to the same extent as the cross-reaction of the antibody to human E2 and E3-binding protein (E3-BP) indicating that the plant mitochondrial E2 lipoyl domains were lipoylated. Two possible reasons for the non-recognition of the chloroplastic E2 lipoyl domains are; either the chloroplastic lipoyl domains were not lipoylated or that the mAb is incapable of recognising the plant proteins.

Chloroplastic mature E2 and GST-didomains (lanes 1, 2 and 3) were not recognised by the mAb but a faint band was visible in these lanes, also present in the marker lane (M). This band could be an artefact produced from non-specific cross-reaction with a component of Laemmli buffer. A similar artefact was seen previously with 2-mercaptoethanol as the reducing agent but when DTT was substituted the artefact disappeared (Hunter, 1985).

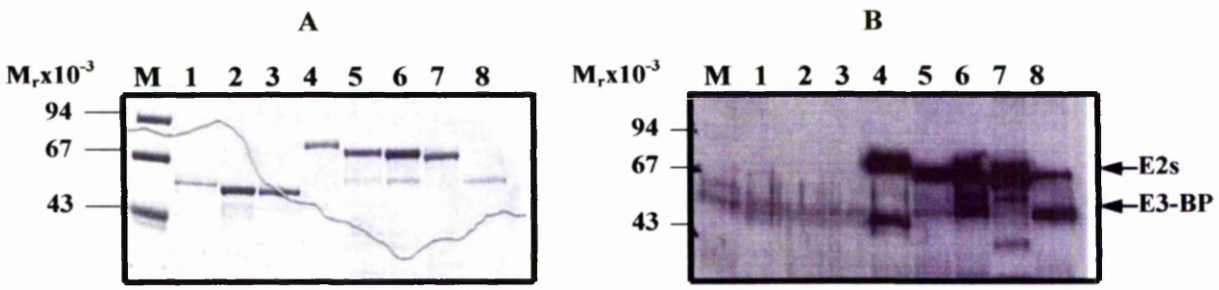


FIGURE 5.8: Immuno-detection of the Chloroplastic and Mitochondrial E2 Components with a mAb Specific for Lipoylated Mammalian E2

Panel A shows the Coomassie blue stained SDS-PAGE of plant E2s (1µg each)

Panel B shows the immuno-blot of plant E2s protein, using antibodies specific for the lipoylated lipoyl domains.

Lane 1: pea chloroplastic mature E2 (His-tagged), grown in media plus lipoic acid (M_r approx. 50, 000)

Lane 2: pea chloroplastic GST-E2 didomain, minus lipoic acid (M_r approx. 50, 000)

Lane 3: pea chloroplastic GST-E2 didomain, plus lipoic acid

Lane 4: *A. thaliana* mitochondrial mature E2 (His-tagged) grown in media plus lipoic acid (M_r approx. 67, 000)

Lane 5: *A. thaliana* mitochondrial GST-E2 tridomain, minus lipoic acid (M_r approx. 70, 000)

Lane 6: *A. thaliana* mitochondrial GST-E2 tridomain, plus lipoic acid

Lane 7: Purified His-tagged human E2, grown in media plus lipoic acid (M_r approx. 67, 000)

Lane 8: Purified His-tagged human E3-BP, grown in media plus lipoic acid (M_r approx. 50, 000)

5. 3. 2. Detection of Lipoylation by Mass Spectrometry

Lipoylation of chloroplastic E2 was undetermined with the use of the monoclonal antibody. Interestingly, however, the mAb (PD2) specifically recognised lipoylated plant mitochondrial E2 in addition to lipoylated human E2 and E3-BP. In order to determine the lipoylation status of the chloroplastic E2 the technique of mass spectrometry was employed with the aim of detecting a change in mass corresponding to the addition of a lipoic acid prosthetic group (188 Da).

Matrix Assisted Laser Desorption/ Ionisation Mass Spectrometry (MALDI)

MALDI has the power to generate ions from biomolecules and to analyse the mass-to-charge ratio of a particular ionised molecule, measured by the time-of-flight (TOF) analyser. A schematic diagram representing the general principles of MALDI-TOF detection is shown in Figure 5.9. The approach utilises a crystalline solution of matrix and analyte (sample) which is bombarded with a pulsed laser beam (typically nitrogen laser). The matrix absorbs this energy and, in doing so, becomes excited causing localised disintegration of the crystalline solution. Generation of a plume of single analyte molecules surrounded by excited matrix is followed by an energy transfer from the matrix to analyte. The ionised analyte is propelled towards the time-of-flight analyser. The TOF analyser measures for time of an ionised molecule to travel a set distance with the time taken directly related to the mass/charge (m/z) ratio. The TOF analyser records this information in the form of a mass spectrum, plotting the m/z ratio against intensity (indication of the abundance of each component in the sample). MALDI-TOF analysis is a "soft" ionisation technique generating singly charged molecular related ions regardless of M_r value.

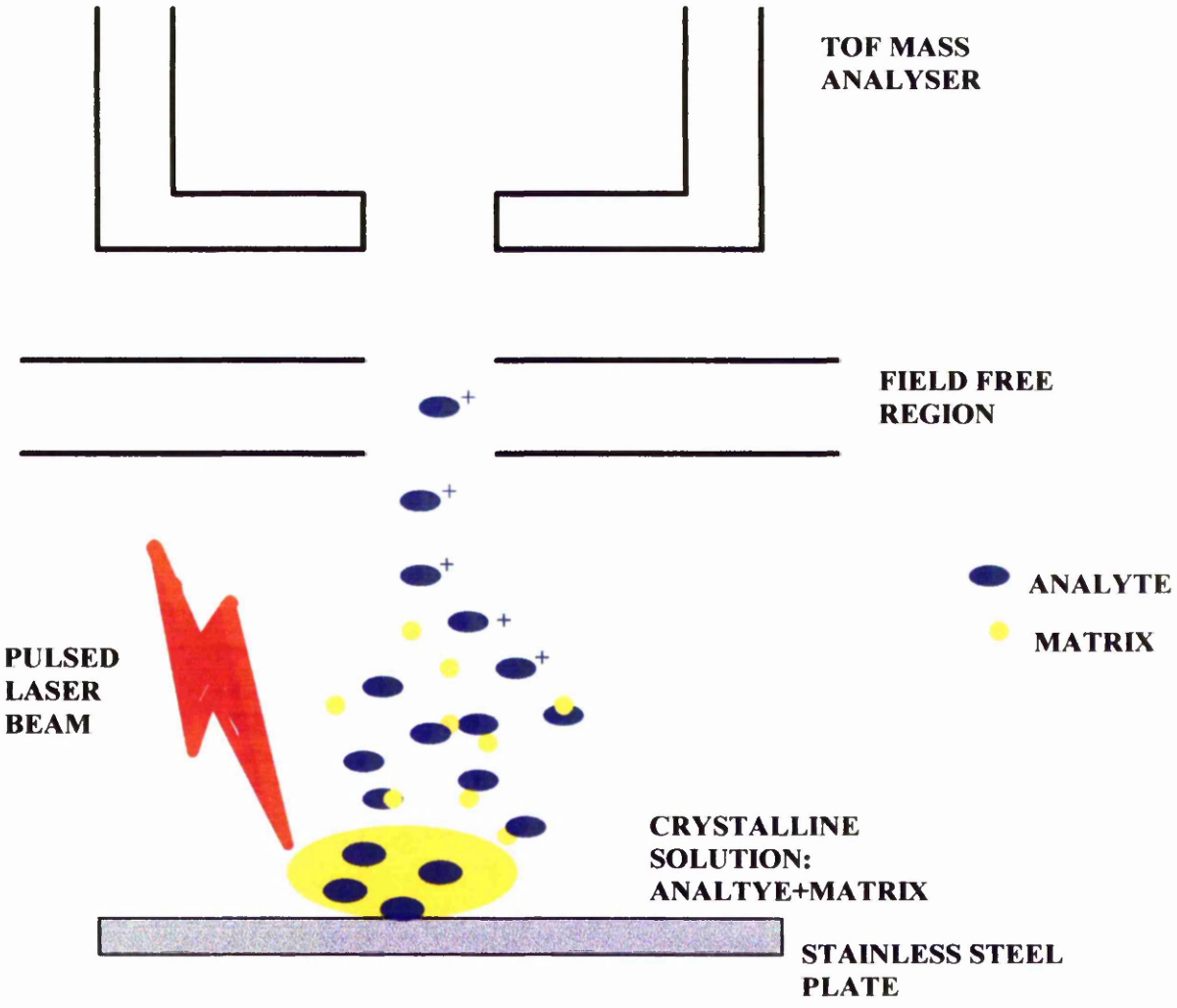


FIGURE 5.9: Schematic Diagram of MALDI (TOF) Mass Spectrometry

The laser bombards the analyte and matrix mixture spotted on the plate and disintegrates the crystals. The analyte becomes ionised (+) and is propelled towards the field free drift region and the time-of-flight analyser.

Sample Preparation

Purified chloroplastic GST-E2 didomains and mitochondrial GST-E2 tridomains were prepared as described in Methods section 2. 2. 28. Mixing with matrix dissolved in a volatile solvent with a trace of trifluoroacetic acid permits positive ionisation to occur. Spotted crystalline solutions on the sample plate are allowed to air dry before being introduced to the mass spectrometer. MALDI-TOF analyses were calibrated using an external calibration standard such as myoglobin (see section 2. 2. 28).

Results

5. 3. 2. 1. Detection of Lipoylated GST-E2 Didomains by MALDI-TOF

Modification by the attachment of a lipoate molecule will increase the proteins mass 190 Da; it is therefore logical to assume that when the lipoic acid prosthetic group is covalently attached to the E2 lipoyl domain the M_r of the lipoyl domain will increase by an equivalent amount. Detection of a mass increase of the correct magnitude would indicate lipoylation by the *E. coli* host. In the absence of lipoylation, measuring the mass could also permit the detection of octanoylated domains, described in the expression of lipoyl domains under lipoate deficient conditions (Ali *et al.*, 1990).

Using this approach, the mass/charge ratios of the chloroplastic GST-E2 didomains, grown in media supplemented with or without lipoate, were recorded as shown in Figure 5.10. The mass spectrum (panel A) of the GST-E2 didomain expressed in the absence of lipoic acid illustrates the presence of a major peak with a mass/ charge ratio or M_r of 48, 226. Assuming that the GST-E2 didomain was unlipoylated, the measured mass was consistent with the predicted M_r of 48, 219. In contrast a peak with a M_r of 48, 455 was measured for the GST-E2 didomain expressed in presence of exogenous lipoate (panel B). This higher mass suggests that the GST-E2 didomain was lipoylated under these conditions. However, the measured mass was higher than expected resulting in a difference of

229 Da instead of 190 Da. This may be accounted for by further modification of the lipoate moiety such as acetylation (addition of 42 Da) of the reactive sulphydryl groups of lipoic acid. Another consideration is biotinylation of the lipoyl domain (addition of 226 Da) but was dismissed owing to the requirement of a specific MKM motif. Other modifications such as the oxidation of the SH groups of lipoic acid to SO, SO_2H and SO_3H may all account for the additional difference in M_r values.

From the MALDI results lipoylation of the chloroplastic E2 didomain was suggested to occur when exogenous lipoate was added to the growth media. While the lipoyl domain remained in its apo-form when no exogenous lipoate was present. From the results no evidence was found to support covalent modification such as octanoylation where an increase in mass of 127 Da would be expected.

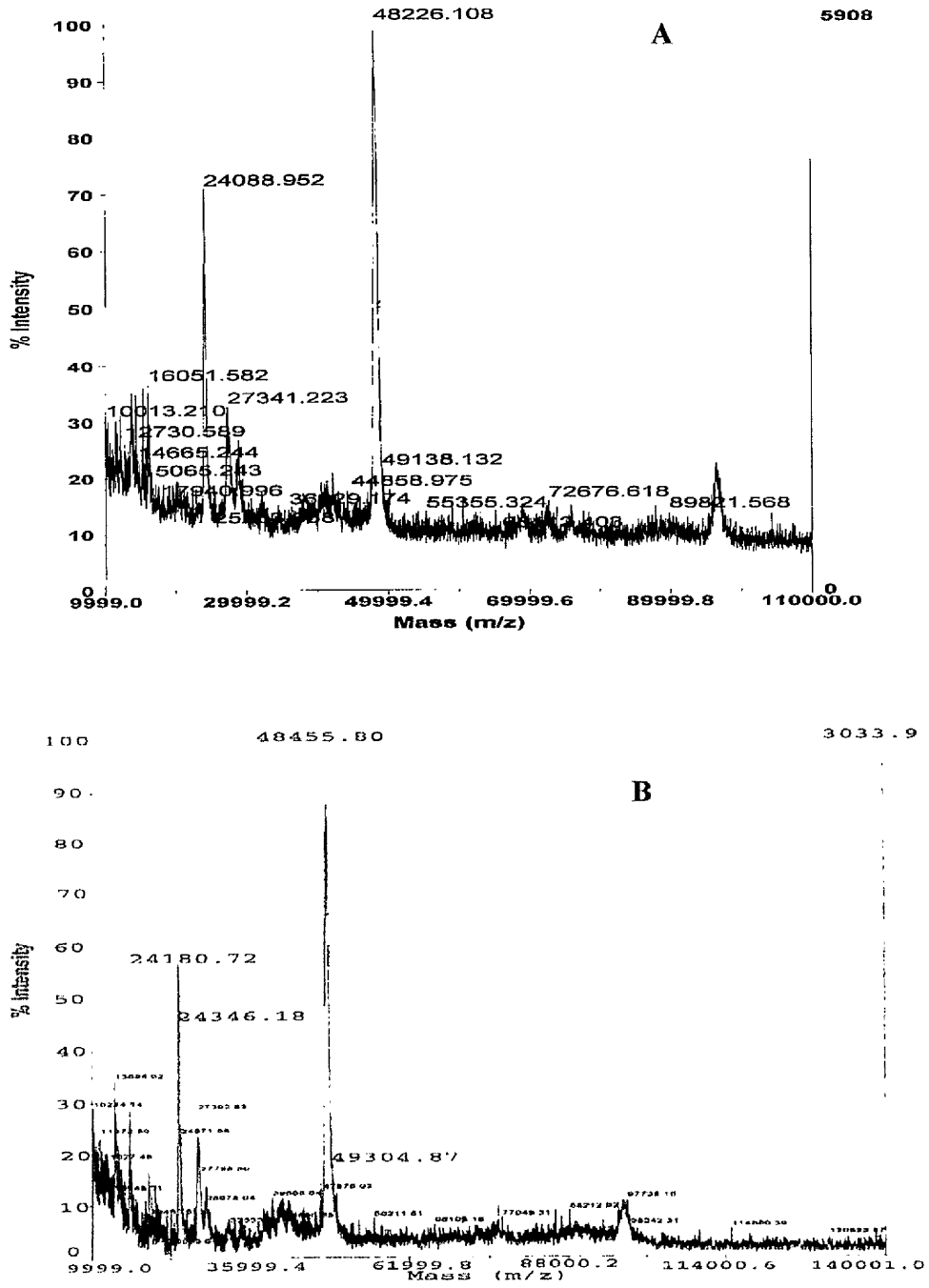


FIGURE 5.10: Mass Spectra of GST-E2 Didomain Minus/Plus Lipoate

Measured in a time of flight mass spectrometer with matrix assisted laser desorption ionisation, calibrated with myoglobin. Panel A, shows the spectrum for GST-E3 didomain minus lipoate. Panel B, is the spectrum of the GST-E2 didomain plus exogenous lipoate

5. 3. 2. 2. Detection of Lipoylated GST-E2 Tridomains by MALDI-TOF

Lipoylation of the mitochondrial E2 has previously been determined immunologically (see section 5. 3. 1) suggesting that the *E. coli* lipoylation apperatus recognised the plant mitochondrial E2. These findings were further substantiated by measuring the mass of the GST-E2 tridomains, with or without lipoate supplementation. Samples were prepared as described in Methods section 2. 2. 28, except that α-cyano-4-hydroxy-cinammic acid (CHCA) was used as the matrix since results were more reproducible although lower intensities were obtained. The M_r of the apo-protein of the GST-E2 tridomain was calculated to be 61, 368 while the fully lipoylated protein M_r was expected to be 61, 744 (two lipoyl domains present).

In Figure 5.11, (panel A), the mass spectrum of GST-E2 tridomain expressed without lipoic acid shows a small peak of M_r of 61, 744 exactly the M_r value expected for the lipoylated holo-protein. While in Figure 5.11, panel B, the mass spectrum of the GST-E2 tridomain expressed in the presence of lipoic acid is presented. A small peak with a M_r of 61, 751 was observed extremely close to the predicted M_r for lipoylated GST-E2 tridomains. The results suggest that the *A. thaliana* mitochondrial GST-E2 tridomain was fully lipoylated whether it was expressed with or without exogenous lipoate. However, owing to a decreased sensitivity in the mass analysis it is possible that non-lipoylated domains were not detected in this study. Both mass spectra contained several peaks of smaller M_r value, possibly due to multiply charged protein species of GST-E2 tridomain. The mass measured by the TOF mass spectrometer is the ratio of mass to charge (m/z); therefore the mass of a singly charged species is equal to the actual mass; however, if the protein species is doubly or trebly charged then the measured mass is correspondingly reduced.

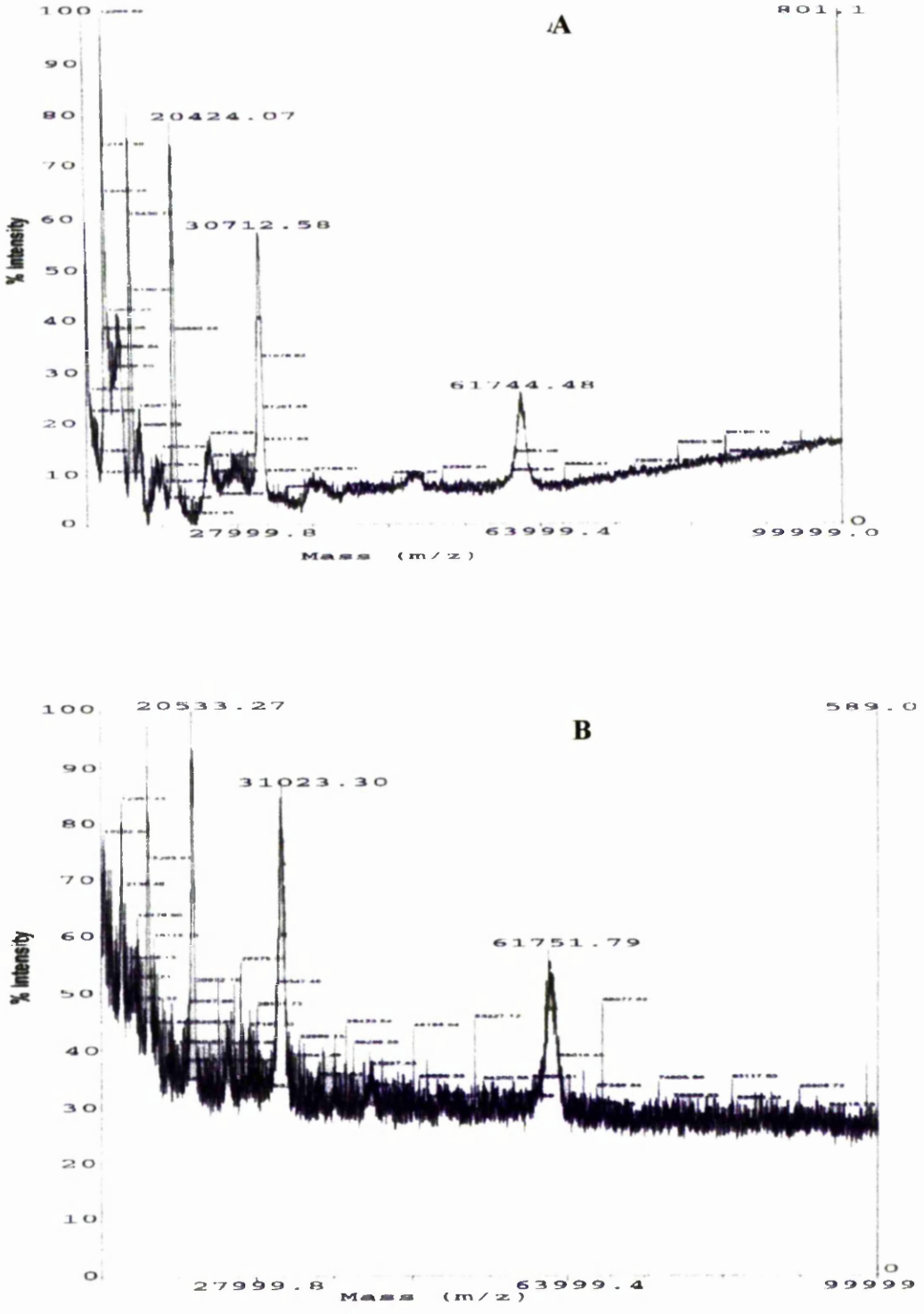


FIGURE 5.11: Mass Spectra of GST-E2 Tridomain Minus/Plus Lipoate

Measured in time of flight mass spectrometer with matrix assisted laser desorption ionisation, calibrated with myoglobim. Panel A, shows the spectrum of GST-E2 tridomain minus lipoate while panel B, shows the results for GST-E2 tridomain plus lipoate.

5. 4. DISCUSSION

The gene sequence encoding pea chloroplastic E2 was conclusively identified as the E2 component of the pyruvate dehydrogenase complex with 86% sequence identity to the *A. thaliana* plastid PDC-E2. The 3D structure of the single lipoyl domain was predicted by employing SWISS MODEL which aligned the first 90 N-terminal residues of chloroplastic E2 to the human inner lipoyl domain. Alignment of both sequences allows the chloroplastic E2 to be modelled on the 3D structure of the human lipoyl domain. The human and the predicted chloroplastic E2 lipoyl domains are almost superimposable with the location of the conserved lipoyl-lysine residue in the chloroplastic E2 in the identical position to that in the human lipoyl domain.

Native E2s are post-translationally modified by the covalent attachment of a lipoic acid group to a specific lysine residing in the lipoyl domain which is essential for the overall PDC reaction. Therefore the lipoylation status of the plant E2s expression was investigated.

Expression of chloroplastic GST-E2 didomains and mitochondrial GST-E2 tridomains in media supplemented with or without lipoic acid were visualised by SDS-PAGE. A band shift was observed in the chloroplastic GST-E2 didomain expression on the addition of lipoic acid. The appearance of a dominant upper band on the SDS-gel was considered to be the lipoylated form while the lower band the unlipoylated form. Both bands were visible in the minus lipoic acid sample leading to the assumption that a mixture of apo- and holo-domains were present. Band shifts were not always observed for the GST-E2 didomain and may be a consequence of variable expression of the protein resulting in changeable lipoylation. SDS-PAGE analysis of the expression of mitochondrial GST-E2 tridomain did not show any band shift pattern hinting that the protein was either in an unlipoylated or fully lipoylated state but, more likely the low resolution power of the SDS-PAGE within this high M_r range prevents the separation of these forms.

Lipoylation of heterologous human E2 and E3-BP expressed in *E. coli* has previously been detected with the use of a monoclonal antibody (PD2) specific for lipoylated domains. This method was used to identify the occurrence of lipoylation in the plant E2s. Mitochondrial E2 and GST-E2 tridomains, expressed in medium plus or minus lipoate, were recognised by the monoclonal antibody. However, the chloroplastic E2 and GST-E2 didomains (expressed +/- lipoic acid in medium) were not recognised suggesting that either they are not lipoylated or that the monoclonal antibody (PD2) is unable to recognise chloroplastic E2.

Sequence comparison of the lipoyl domains from human, *A. thaliana* mitochondria and plastid and pea chloroplast PDC-E2 and human E3-binding protein exhibit regions of homology especially surrounding the lipoyl-lysine residue as shown in Figure 5.12. The chloroplastic E2 sequence is 34-37% identical to human and *A. thaliana* mitochondrial PDC E2 lipoyl domains, but shows a higher identity to *A. thaliana* plastid lipoyl domain (86%). However, *A. thaliana* E2 lipoyl domains are more homologous to human E2 lipoyl domains having 49-61% sequence identity. This difference in sequence homology between the chloroplastic lipoyl domains and the mitochondrial domains of either *A. thaliana* or human may explain the preference of the mAb PD2 for mitochondrial E2 lipoyl domains. The specificity the monoclonal antibody PD2 for eukaryotic PDCs may be the result of a contiguous epitope present in the mitochondrial lipoyl domains in addition to the absolute dependence of the lipoate prosthetic group. A striking feature is shown in Figure 5.12 of the alignments of the lipoyl domains, there is a common motif of -EI/VET(DKA)V/T- found N-terminal to the DKA motif seen in all the lipoyl domains except the chloroplastic E2s and may account for the complex specific recognition of the mAb PD2.

In order to describe the lipoylation status of chloroplastic E2, mass spectrometry was employed using matrix assisted laser desorption ionisation equipped with a time-of-flight mass spectrometer. The measured mass was recorded as the ratio of mass to charge (m/z) with the dominant ionised species being singly charged; therefore the measured mass was directly proportional to the M_r of the protein. The GST-E2 didomain minus lipoic acid produced a peak of m/z equal to 48, 226,

indicative of an unlipoylated GST-E2 didomain (expected M_r of 48, 219). A mixture of lipoylated or unlipoylated domains as indicated by SDS-PAGE analysis was not found in this sample by MALDI-TOF analysis, but this may be the effect of variable lipoylation.

In contrast, the GST-E2 didomain plus lipoic acid shows a peak of m/z equal to 48, 455, suggesting the presence of a lipoylated lipoyl domain although the increase in M_r was greater than the addition of a lipoate moiety (188). Assuming lipoylation had occurred possible explanations for the additional mass of 48Da could be the acetylation (42Da) of the lipoic acid moiety or oxidation of the reactive S^6 and S^8 sulphydryl groups of the cofactor to either sulphenic acid (SO), suphinic acid (SO_2) or sulphoxide (SO_3). Biotinylation (biotin group, 266Da) of the lipoyl domain was also considered, however, site directed mutagenesis has shown that when the biotinyl motif, MKM, is substituted into the lipoyl domain for the DKA motif, the lysine was not recognised by the biotinyl protein ligase. This was also true when the DKA motif was incorporated into the biotinyl domain however it then becomes a poor substrate for the lipoyl ligase (Perham and Reche, 1998). It was therefore concluded that the mass increase detected by MALDI-TOF was likely to indicate the lipoylation of chloroplastic E2 and that the lipoate had been modified by acetylation or oxidation.

The mitochondrial GST-E2 tridomain, however, was found to be lipoylated in both samples. The GST-E2 tridomain minus lipoate produced a peak of m/z equal to 61, 744 expected for the lipoylated tridomain while the GST-E2 tridomain plus lipoate produced a peak of m/z equal to 61, 751. These MALDI-TOF results suggest that the mitochondrial E2 was lipoylated independently of exogenous lipoate.

In conclusion, the combined results from Western blotting and MALDI-TOF presented in this chapter provide evidence for the lipoylation of both recombinant pea chloroplastic E2 and *A. thaliana* mitochondrial E2 expressed either in their mature form or in their truncated di/tri domains in *E. coli*.

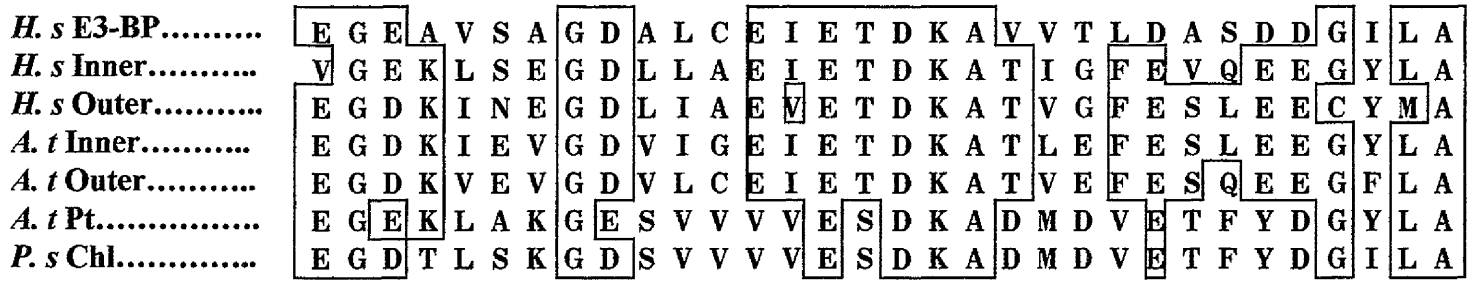


Figure 5.12: Alignment of PDC-E2 Lipoyl Domains from Various Species

Boxes show homologous regions of sequence in all of the lipoyl domain sequences. Many smaller regions of homology are not shown but these involve the species specific pairs e.g. the sequence "SVVVVESDKADMDVETFYDG*LA" is seen in the *A. thaliana* plastid lipoyl domain and also in the *P. sativum* chloroplastic lipoyl domain.

H. s: Homo sapien, E3 binding protein lipoyl domain
H. s: *Homo sapien*, where outer and inner refer to the mitochondrial outer N-terminal or inner lipoyl domain
A. t: *Arabidopsis thaliana*, where outer and inner refer to the mitochondrial lipoyl domains as in *Homo sapien*
A. t: *A. thaliana* plastid E2 lipoyl domain, contains a single lipoyl domain
P. s: *Pisum sativum* chloroplastic E2 lipoyl domain (single domain)

CHAPTER 6

STUDYING THE OLIGOMERIC STRUCTURE OF E2

6. STUDYING THE OLIGOMERIC STRUCTURE OF E2

6. 1. INTRODUCTION

The eukaryotic pyruvate dehydrogenase complex is an example of a multienzyme complex consisting of non-covalently bound enzymes arranged at precise locations around an oligomeric E2 core. Two forms of the oligomeric core have been identified by electron microscopy: a 60meric pentagonal dodecahedron as seen in eukaryotic mitochondrial PDC and Gram positive bacterial PDC and a 24meric cube seen in Gram negative bacterial PDC as shown in Figure 1.11.

The oligomeric core structures of PDC-E2 from plants are not well characterised; however, a 60meric E2 core from plant mitochondrial PDC was identified in maize and potato. The M_r value of recombinant maize E2 was determined to be 2.7 MDa (predicted 2.82 MDa) by size exclusion chromatography and the assembled core was visualised by electron microscopy (Thelen *et al.*, 1999). Its morphology was deduced to be a 29nm diameter pentagonal dodecahedron similar to the human E2 core. Analytical ultracentrifugation studies on the potato mitochondrial E2 core resulted in an estimated M_r value of 2.3-2.6 MDa consistent with a 60meric structure with respect to the monomer mass (Millar *et al.*, 1999).

There have been no reports on the morphology of the chloroplastic E2 cores as yet but elucidation of the oligomeric structure may aid the understanding of the evolutionary origin of mitochondria and chloroplasts. The most widely accepted theory is the "endosymbiont theory" which claims that mitochondria and chloroplasts originated from prokaryotic bacteria, which invaded host cells (proto-eukaryotic cell) and evolved a symbiotic relationship with the host cell. It is suggested that the mitochondria evolved from free-living aerobic bacteria while the chloroplast evolved from primitive blue-green algae (e.g. cyanobacteria). Tracing the evolution of both organelles has involved the comparison of protein sequences such as ferredoxins and ribosomal RNAs. In addition to these the morphology of the E2 core may also be used as a comparison.

In this chapter, the organisation of the plant mitochondrial and chloroplastic E2 will be investigated. Preliminary studies on both of the E2 cores by size exclusion chromatography using a Superose 6 column proved to be inadequate in ascertaining the precise M_r values of each E2 complex. Both E2s migrated as distinct protein species; however, both eluted at or near the void volume of the column and so no discrimination between either a 24meric or 60meric structure was possible. Therefore, the technique of analytical ultracentrifugation was employed as a more accurate means of measuring M_r values of large complexes with the potential of resolving the stoichiometry of the purified recombinant chloroplastic and mitochondrial E2 cores. Analytical ultracentrifugation studies can be performed in two modes: sedimentation velocity and sedimentation equilibrium. In this chapter only sedimentation equilibrium will be discussed.

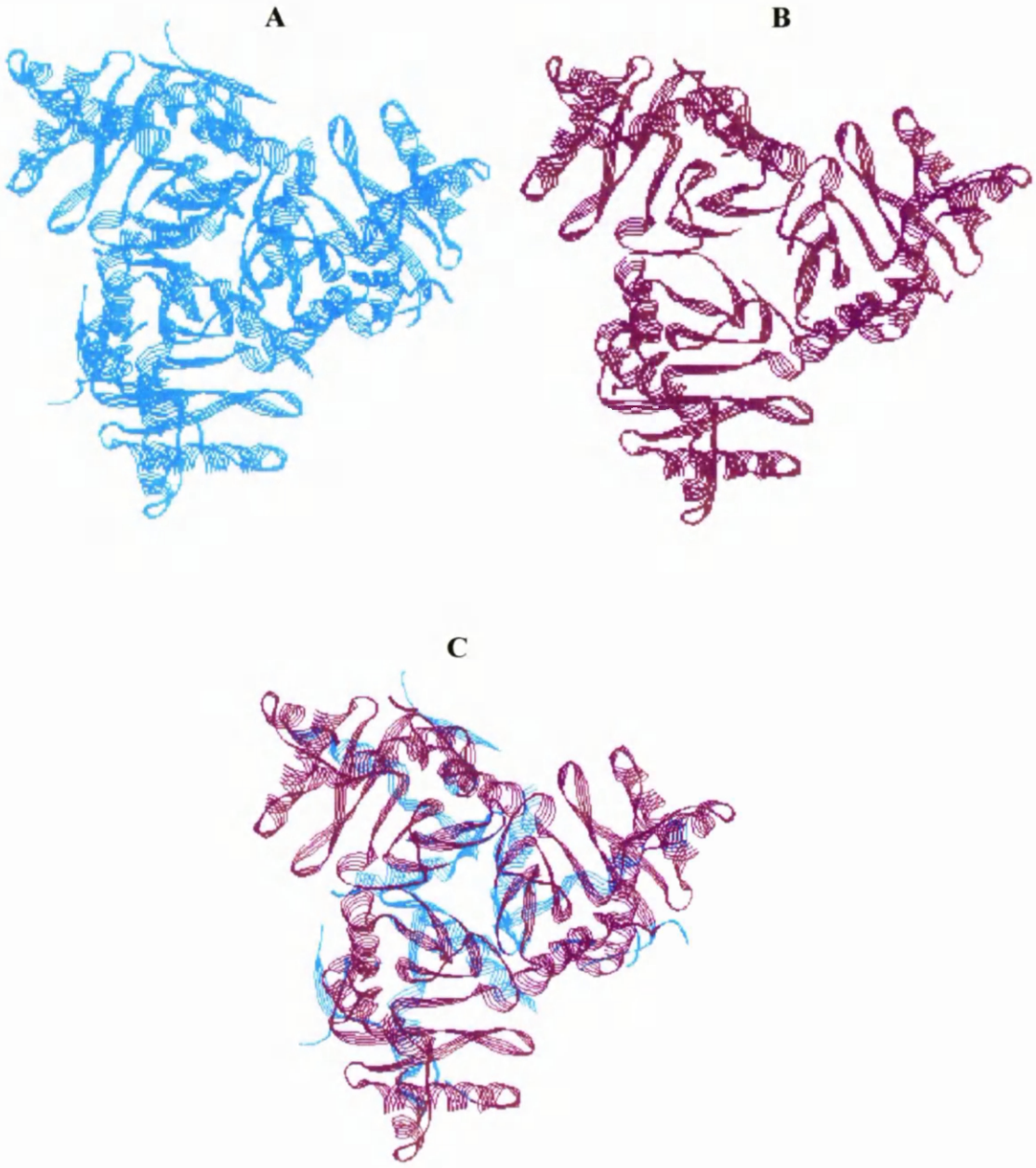
6. 2. Oligomeric Structure of the Catalytic Domain of E2

The oligomeric structure of the chloroplastic E2 has not yet been reported but preliminary sequence alignments with other E2s of known structure has allowed a predicted structure to be produced. The possible oligomeric core of chloroplastic E2 sequence was achieved by using the internet based search tool SWISS MODEL. The C-terminal 186 residues, identified as the catalytic domain, could be aligned with a subunit of the trimeric dihydrolipoamide succinyltransferase from *E. coli* and the dihydrolipoamide transacetylase catalytic domain of *B. stearothermophilus*. However, when three catalytic domain sequences were joined in succession and then aligned, the complete trimeric structure was predicted (Figure 6.1).

Although the pea chloroplastic E2 has a relatively low degree of sequence identity (32%) with this E2 from *E. coli* OGDC, the predicted structure of the chloroplastic trimer was similar to that of the OGDC-E2 trimeric catalytic domain. Modelling the chloroplast E2 into a trimer showed that it was capable of forming the morphological unit important in the assembly of either the 24meric cube or the 60meric pentagonal dodecahedron. However, the precise oligomeric form assembled from the chloroplast E2 trimer could be not be determined by SWISS MODEL

FIGURE 6.1: Trimeric Structure of the Catalytic Domain of Pea Chloroplastic E2 Aligned with *E. coli* OGDC-E2 Catalytic Trimer

The structure of the *E. coli* OGDC-E2 catalytic trimer (A), is shown along side the predicted pea chloroplastic PDC-E2 trimer (B). The chloroplastic structure is almost identical to the structure of the *E. coli* OGDC-E2 catalytic trimer, so much so that they can be superimposed (C). The SWISS MODEL search tool performed identification and alignments automatically.


A
B
C

6. 3. ANALYTICAL ULTRACENTRIFUGATION

6. 3. 1. The Ultracentrifuge Instrument

The technique of analytical ultracentrifugation (AUC) is widely employed for obtaining molecular masses and information about molecular size, conformation and protein-protein interactions. An easy-to-use, computer based Beckman Optima XL-I analytical ultracentrifuge (Palo Alto, USA) available with both scanning absorption and Rayleigh interference optics was employed for these measurements.

Optics

The optical system is used to record the concentration distribution of a sample (at different radial positions at a given time). The Beckman XL-I model (Figure 6.2) is equipped with scanning absorption optics. A high intensity xenon flash lamp means that wavelengths between 190-800nm can be utilised; it is fired briefly as the selected sector passes the detector during each revolution. A slit below the sample moves to allow sampling at different radial positions. Multiple readings at a single position are collected and averaged to reduce noise.

Sample Cells

The XL-I instrument measures the difference in absorbance between the sample and the reference buffer (ideally the dialysate) to account for the absorbing components of the solvent. Six channel cells can be used to examine up to three samples per cell for equilibrium experiments.

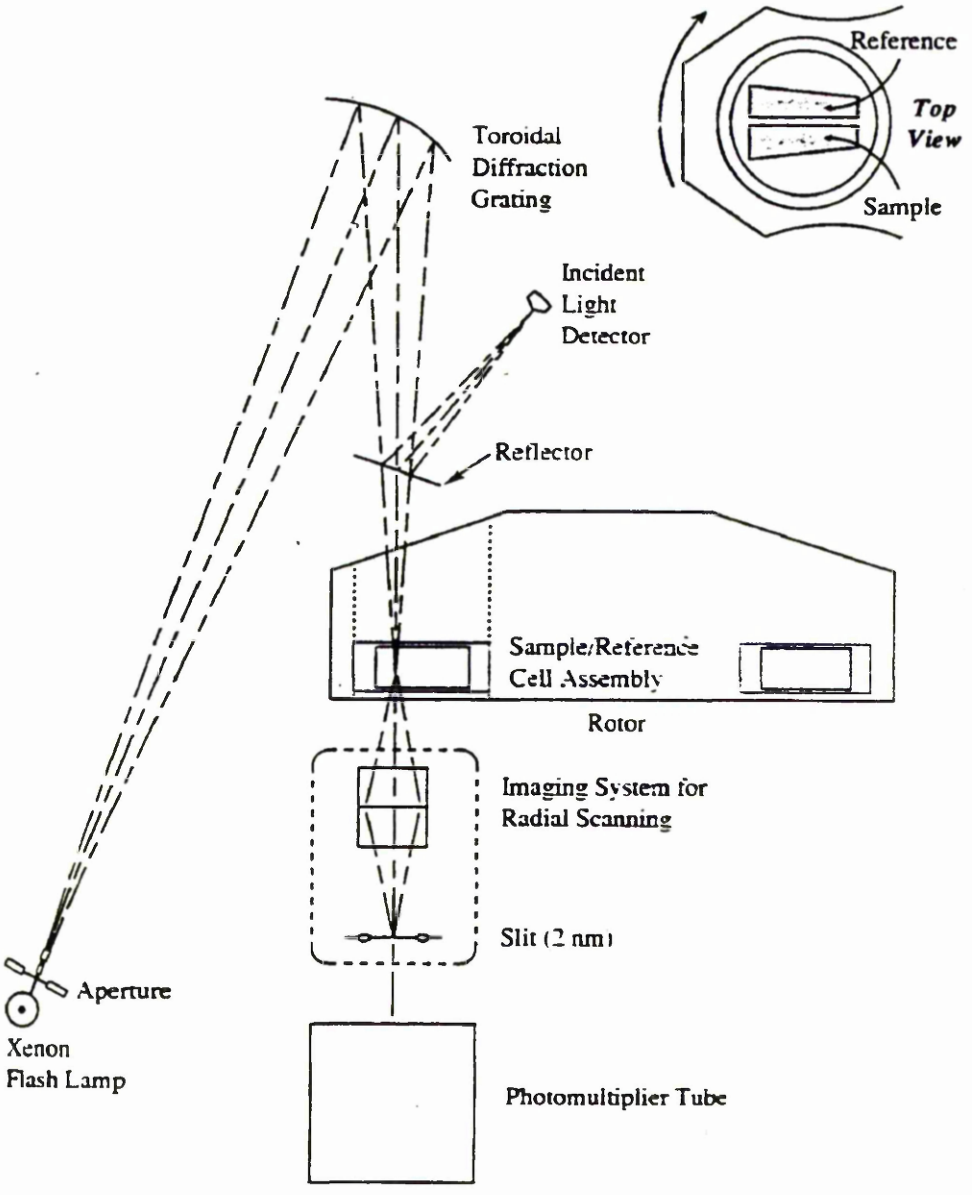


FIGURE 6.2: A Schematic Diagram of the Optical System of the Beckman Optima XL-A /XL-I Analytical Ultracentrifuge

(Taken from the Introduction to Analytical Ultracentrifugation, Beckman Handbook, Beckman Instruments Inc. California)

6. 3. 2. THE THEORY BEHIND SEDIMENTATION EQUILIBRIUM (SE)

In sedimentation equilibrium, small quantities of an initially uniform solution are centrifuged until no net movement of molecules occurs. At this point diffusion exactly balances sedimentation everywhere in the cell (Figure 6.3). Sedimentation is proportional to $\omega^2 r$, with r (radius) increasing towards the cell base so the tendency to sediment is greater nearer to the cell base. Diffusion is driven by the concentration gradient and so solute distribution will no longer vary with time. Thermodynamically, at equilibrium the total potential of a solute is the same at all points in the cell.

The molecular weight for a homogeneous, non-interacting solute species can be determined by its distribution in the cell whereas, for a heterogeneous solution at equilibrium the various species distribute at different radial positions depending on their mass. High molecular weight molecules distribute near the bottom (furthest radial position) of the cell while lower weight molecules dominate at the top (closest to the centre) see Figure 6.3. The measured molecular weight is based on the proportion of various species present so in a system of different solute species the weight average mass (M_w) is determined.

The raw data from SE experiments are fitted to defined models to produce a good fit that describes the sample system. As well as M_r value, homogeneity of the protein sample with respect to M_r and aggregation status can be determined along with stoichiometry and equilibrium constants for associative processes.

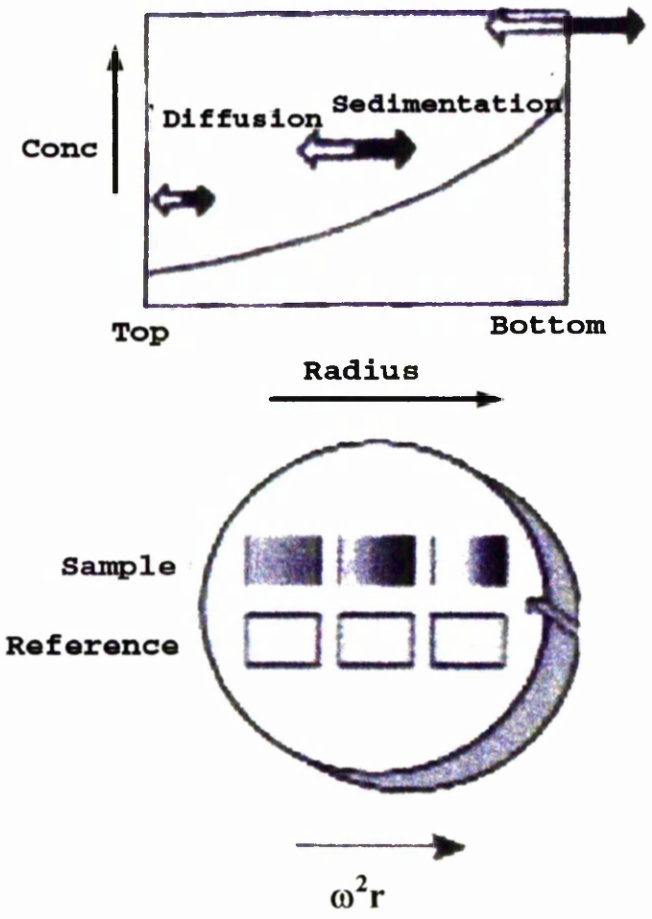


FIGURE 6.3: Schematic Diagram of Sedimentation Equilibrium

The top diagram illustrates the two forces on a molecule in solution within the cell in sedimentation equilibrium. The white arrows represent diffusion while the black arrows show sedimentation. After time both processes balance and equilibrium is reached at all points in the cell. The concentration increases at the bottom of the cell due to the increase in gravitational force sedimenting large molecules.

Below, a six-cell centrepiece that holds three samples varying in concentration at different radial distances and three volumes of reference buffer (dialysate) in the opposing cells.

Factors Affecting Sedimentation Equilibrium

For adequate distribution of molecules in a solution the speed required is dependent on the mass. For example, equilibrium achieved at high speeds is sensitive for detecting low molecular mass species while low speeds are more useful in the detection of higher mass species and for concentrated samples.

Non-ideality effects arise from the finite size of the molecule and the charge it carries, both of which reduce the apparent molecular mass. In homogeneous solutions the measured molecular mass is an apparent molecular mass, M_{app}, or apparent weight-average molecular mass, $M_{w,app}$, for heterogeneous solutions.

Self-association increases with increasing concentration (law of mass action). The apparent weight-average molecular mass ($M_{w,app}$) will also increase with increasing concentration, reflecting the increasing proportion of oligomers.

Analysis of the Sedimentation Equilibrium Data

All sedimentation equilibrium data were analysed using the Beckman Optima XL-I data analysis package and all data handling was performed by G. Campbell, Department of Infection and Immunity, University of Glasgow. Two models were used to analyse and determine the apparent weight-average molecular mass, $M_{w,app,}$ of the PDC-E2 oligomeric cores.

IDEAL 1

This model was initially used with SE data to estimate the apparent molecular mass (M_{app}) and to indicate any non-ideality or associative/ dissociative behaviour.
The model fits the raw data obtained in terms of a single, monodisperse thermodynamically ideal species according to the distribution in the cell.

ASSOC4

When the mass exceeds that of the monomer and self-association is suspected the equilibrium data can be analysed with the ASSOC4 model that takes all protein species present into account. The ASSOC4 model is similar to that for the ideal model, except that the total absorbance at a given radius is the sum of absorbances of all the species at that radius. Therefore at equilibrium the absorbance measurement by the analytical ultracentrifuge is recorded as the sum of all the species present. The model describes total macromolecule distribution for monomer $\rightarrow$ n-mer.

Sample Preparation

Analytical ultracentrifugation (AUC) experiments were performed on purified His-tagged E2s (see Chapter 4) dialysed as described in Methods section (2. 2. 22) and concentrated until an optical density at 280nm of 0.5 was reached using centrifugal concentration filters (see Methods section 2. 2. 31). The AUC experiments employed several protein concentrations created by diluting the stock protein solution.

Results

6. 3. 3. Sedimentation Equilibrium of Pea Chloroplastic E2

Sedimentation equilibrium studies on chloroplastic E2 were performed at 3, 000rpm during which absorption data were recorded. The data obtained were fitted to a single ideal model (IDEAL 1, assumes that only one species was present). An example of the IDEAL 1 model to data fit is shown in Figure 6.4. These fits were relatively poor as indicated by the residuals of the data to model fit. The residuals displayed a characteristic 'tick' shape suggesting that the solution was heterogeneous and was indicative of association reactions or aggregation.

The apparent weight-average molecular mass ($M_{w,app}$) of chloroplastic E2 as a function of concentration is plotted Figure 6.5 from the linear regression fit through the data. An apparent weight-average molar mass of 1.0-1.1 MDa at infinite dilution was obtained. This is in good agreement with the predicted M_r of a 24meric core with a M_r value of 1.14 MDa, determined using the monomer M_r value of 47, 703.

From Figure 6.5, no significant trend is observed with relation to mass with increasing concentration. A tendency for the mass to decrease with increasing protein concentration is associated with non-ideality; however, an increase in mass followed by a decrease, observed in this sample, may be due to heterogeneity or association (both lead to mass increase with concentration). Further fitting of the data to multimeric models using ASSOC4 incorporating large polymeric complexes allowed improvements to be made, therefore substantiating the theory of aggregates being present in solution.

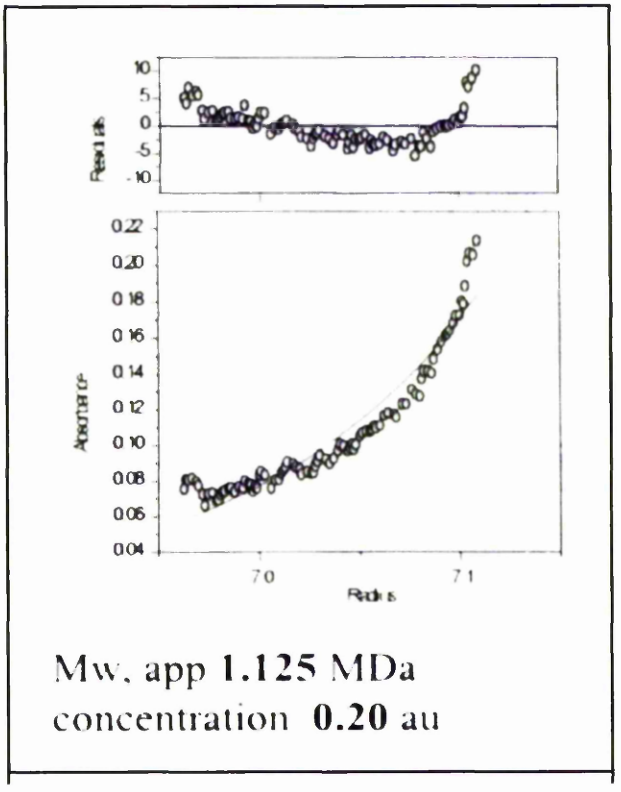


FIGURE 6.4: IDEAL 1 Model Fitted to Raw Sedimentation Equilibrium Data for Pea Chloroplastic E2

Rotor speed 3, 000 rpm at 4°C, concentration measured in absorbance units (au) at 280nm.

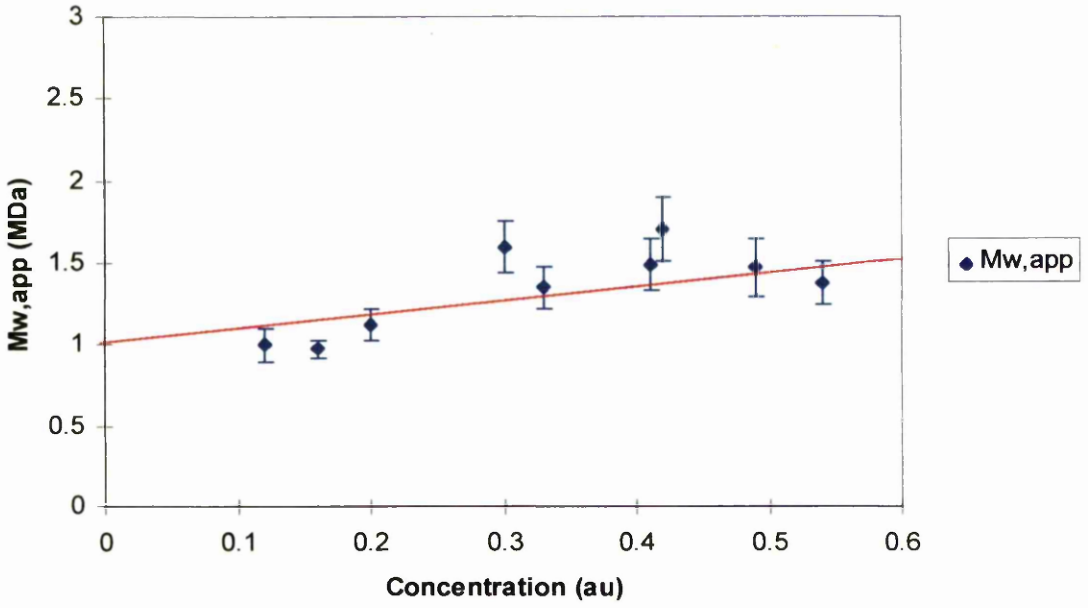


FIGURE 6.5: Apparent Weight-Average Molecular Mass of Pea Chloroplastic E2 Determined by Sedimentation Equilibrium

Rotor speed 3, 000 rpm at 4°C

The red line shows the line of best fit to the SE data.

6. 3. 4. Sedimentation Equilibrium of the Pea Chloroplastic E2 at Different Speeds

Sedimentation equilibrium experiments on chloroplastic E2 were repeated (using a second E2 sample) at three speeds (3, 000, 7, 000 and 15, 000rpm). It was thought, from the first 3, 000rpm SE run, that aggregates or associating species were present in the protein sample. Repeating the SE experiment at higher speeds permits detection of any smaller M_r protein species (potential associating species). The M_r value ranges detected by these speeds are as follows; 0.7-10 MDa, 150-600 kDa and 40-150 kDa respectively. As before the raw data were fitted to a single ideal model (IDEAL 1) for each speed to obtain a $M_{w,app}$ for the protein species present.

Results from the high speed centrifuge runs (7, 000 and 15, 000rpm) were plotted as a function of concentration (results not shown), a linear regression fit through the data points was used to obtain the M_r value ($M_{w,app}$) at infinite dilution. For the 7, 000rpm SE experiment, a $M_{w,app}$ of 150, 000-160, 000 Da was obtained while for the 15, 000rpm SE experiment a $M_{w,app}$ of 50, 000-80, 000 Da was estimated. Both sets of $M_{w,app}$ values are slightly larger than expected but are suggestive of trimeric E2 (approx. 143, 109 Da) and monomeric E2 (approx. 47, 703 Da) present in solution.

At low speed (3, 000rpm) the results were consistent with the $M_{w,app}$ value obtained for the first SE experiment. Indicating the presence of a protein species of 1.0-1.1 MDa, or a 20-23meric polymer.

The IDEAL 1 model fits were relatively poor as before. However, improvements were observed (judged by the evenly distributed residuals) when the ASSOC 4 multimeric model was utilised, an example is shown in Figure 6.6. The purpose of using multimeric models was to fully describe the associating species in solution. Unfortunately the E2 samples proved quite difficult to characterise precisely but the preliminary results from the ASSOC 4 model suggest possible trimer-12mer-24mer associating species.

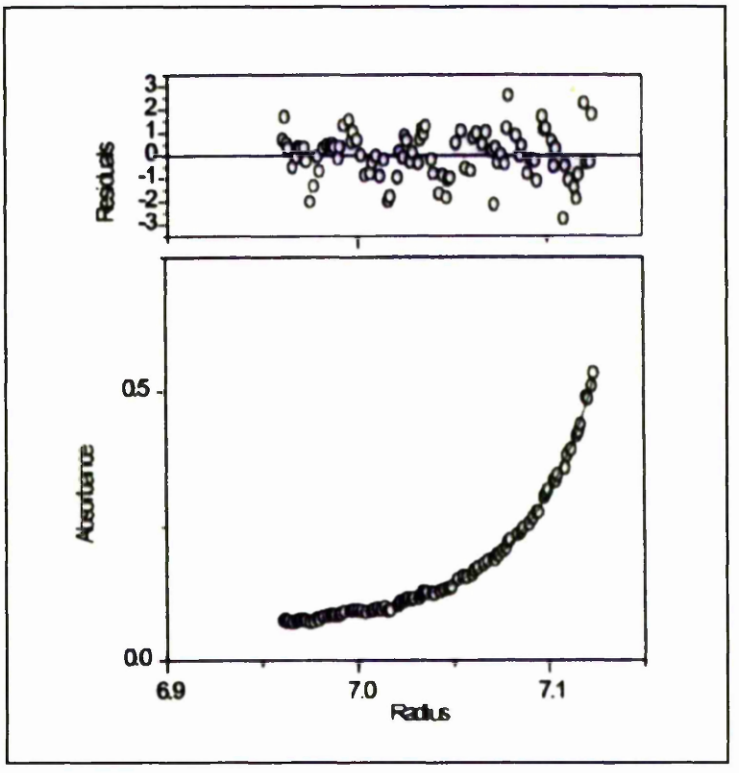


FIGURE 6.6: ASSOC4 trimer/12mer/24mer/72mer Model Fitted to Raw Sedimentation Equilibrium Data of Pea Chloroplastic E2

Rotor speed 7, 000 rpm at 4°C

6. 3. 5. Sedimentation Equilibrium of *A. thaliana* Mitochondrial E2

Plant mitochondrial E2 has been shown to exist as a 60meric species with respect to its elution from size exclusion chromatography but also in its appearance in negatively stained electron micrographs which was similar to E2 complexes from mammals and yeast (Millar *et al.*, 1999; Thelen *et al.*, 1999b). It could be assumed that the *A. thaliana* mitochondrial E2 complex will also form a 60meric core. To test this theory the purified *A. thaliana* E2 was subjected to similar hydrodynamic analysis as for chloroplastic E2.

Raw sedimentation equilibrium data were fitted using the IDEAL 1 model, an example result of which is presented in Figure 6.7. The data to model fits, like those for the chloroplastic E2, were poor with the residuals being unevenly distributed and displaying the characteristic 'tick' shape indicating association or aggregation. The $M_{w,app}$ determined for E2 as a function of concentration is plotted in Figure 6.8. A broad range of M_r values were determined, however, a linear regression fit through the data points yielded an estimated $M_{w,app}$ of 4.0-5.0 MDa. This result suggests a possible 71mer, higher than the expected 60meric structure with a M_r approx. 3.76 MDa, calculated using the monomer M_r of 62, 674. Although higher, the results indicate the presence of a much larger complex compared to that found for the chloroplastic E2. However, the estimated $M_{w,app}$ are within the range of the M_r values for 60meric and larger complexes.

Improvements to the data-model fit were obtained by using the ASSOC 4 model assuming that monomer-60mer-120mer-240mer associations were present. The data revealed that there was monomeric E2 and large aggregates present in the E2 sample but the SE data could not be described precisely by this particular model. Further association models require to be tested but due to time constraints this was not possible.

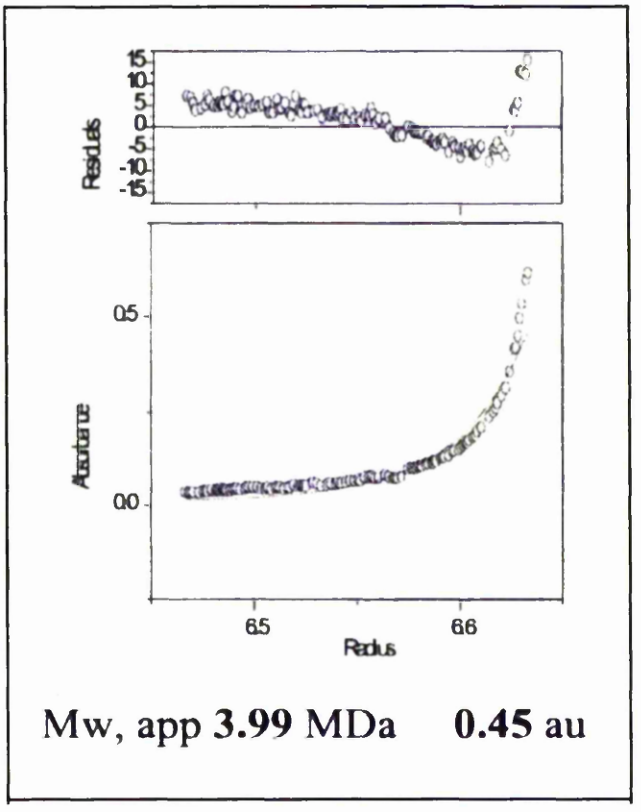


FIGURE 6.7: IDEAL 1 Model Fitted to Raw Sedimentation Equilibrium Data for *A. thaliana* Mitochondrial E2

Rotor speed 3, 000 rpm at 4°C, concentration is measured in absorbance units (au) at 280nm.

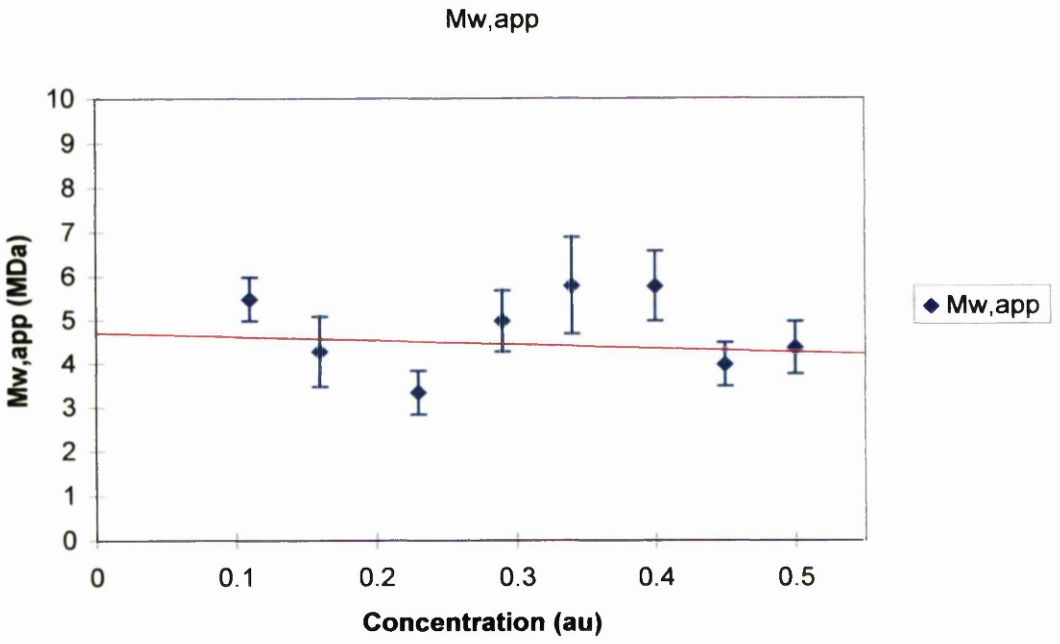


FIGURE 6.8: Apparent Weight-Average Molecular Mass of *A. thaliana* Mitochondrial E2 Determined by Sedimentation Equilibrium

Rotor speed 3, 000 rpm at 4°C

6. 4. DISCUSSION

Pea chloroplastic E2 had previously been shown to contain a single N-terminal lipoyl domain. In a similar manner SWISS MODEL identified the catalytic domain of E2 consisting of the C-terminal 189 residues. Sequence alignment with the trimeric E2 from *E. coli* OGDC, permitted the unknown structure of chloroplastic E2 to be deduced. The chloroplastic E2 monomer displayed a significant similarity to the monomeric subunits of the OGDC-E2 trimer from *E. coli*. Aligning chloroplast monomers in succession suggested the formation of a potential trimeric intermediate. Modelling the plastidic E2 with other E2 structures indicates that the chloroplastic E2 is capable of assembling into trimers generally thought to be the functional building block for assembly into either 24mers or 60mers.

Size exclusion chromatography of the recombinant chloroplastic and mitochondrial E2s showed that both formed large complexes that eluted from a Superose 6 column in the void volume. However, differentiating between the sizes of both complexes was not achieved as the resolution power of the matrix is low at high M_r ranges. From size exclusion studies, it was assumed that the E2s had successfully assembled into their large oligomeric cores; however the existence of higher M_r aggregates cannot be dismissed.

Analytical ultracentrifugation was employed as a means to measuring the M_r values of both E2 cores in solution. Sedimentation equilibrium experiments on chloroplastic E2 suggested a $M_{w,app}$ of 1.0-1.1 MDa indicating an oligomeric core in the range of 20-24 subunits close to the possible 24 meric cube structure. The data-model (IDEAL 1) fits suggested that aggregates of E2 were present in solution. Sedimentation equilibrium studies performed at higher speeds suggested the presence of monomeric and trimeric E2 species, however, the exact quantities of each could not be calculated.

Owing to time constraints only a 3, 000rpm centrifugation run was performed on the *A. thaliana* mitochondrial E2. The data were fitted to the IDEAL 1 model and

a $M_{w,app}$ of 4.0-5.0 MDa was estimated suggesting a possible range of 60meric to 80meric E2 core. Although a wide oligomeric range is suggested, the lower $M_{w,app}$ obtained are consistent with the published observations of a 60meric core from plant mitochondrial PDC. Again heterogeneity was suspected from the SE data. The protein species present in the E2 sample were not fully described but fitting the raw SE data to multimeric models indicated the presence of low level amounts of monomeric E2. It was thought that this monomeric E2 might be misfolded and unable to assemble into the native core.

Both sets of analytical ultracentrifugation experiments reveal the presence of small M_r protein species these were not detected during size exclusion chromatography analysis of these E2s. Dissociation of the E2 cores into smaller polymers during extensive dialysis required for AUC may explain the appearance of these monomeric and trimeric species. Assuming these aggregates of E2 or the low M_r E2 species are inactive the presence of which may explain the low specific activities obtained in chapter 4.

In conclusion, although both sets of experiments do not fully describe the associating protein systems present in these E2 solutions or provide accurate M_r values. What can be inferred from the SE results is that the E2 core complexes from both chloroplastic and mitochondrial PDC are distinctively different in size. The $M_{w,app}$ values obtained suggest that the chloroplastic PDC has a subunit organisation similar to that of Gram negative bacteria while the plant mitochondrial PDC is similar to eukaryotic PDC and Gram positive bacterial PDC.

CHAPTER 7

THE DIHYDROLIPOAMIDE DEHYDROGENASE DIMER

7. THE DIHYDROLIPOAMIDE DEHYDROGENASE DIMER

7. 1. INTRODUCTION

The dihydrolipoamide dehydrogenase (E3) component of the 2-oxo acid dehydrogenase complexes has been well characterised and crystallised from many sources including *A. vinelandii* (Mattevi *et al.*, 1991); *S. cerevisiae* (Toyoda *et al.*, 1998); *P. putida* (Mattevi *et al.*, 1992b) and *P. fluorescens* (Mattevi *et al.*, 1993b). Recently the L-protein of GDC from pea became the first reported crystal structure from a plant source (Faure *et al.,* 2000). From the crystal structure the functional molecule was revealed to be a dimer, as expected and shows similar domain organisation to enzymes of this family.

The L-protein monomer is composed of four domains found in all E3s: the FAD binding domain (residues 1-147), NAD^+ binding domain (148-279), central domain (280-348) and the interface domain (349-470). In terms of secondary structure, each monomer has 12α-helices and 23β-sheets (Faure *et al.*, 2000). This is consistent with the secondary structure identified for the *A. vinelandii* E3 monomer which contains approx. 30% of its residues in α-helical conformation and 27% forming β-strands, giving 11α-helices, two 3_{10}-helices and 23β-sheets (Mattevi *et al.*, 1991).

The crystal structure of the E3 dimer from *P. fluorescens* revealed that dimerisation 16% of monomer accessible surface area is buried (Mattevi *et al.*, 1993b) with the main interactions involving residues from the interface domain. Two regions of subunit-subunit interaction exist at the 2-fold axis of the dimer; these are separated into the lower and upper region. The lower region involves the two long α-helices that form the FAD binding domain and the upper region is constructed from the interface domain, including the catalytic centre and the flavin ring, as shown in Figure 7.1 (Mattevi *et al.*, 1991; 1993b). The C-terminal residues (5-10 residues) also contribute in subunit-subunit interactions as observed in E3 from *P. fluorescens* and *P. putida.* Although these are not distinguishable in the *A.*

vinelandii E3 structure, they appear to be required for enzymic activity (Mattevi *et al.*, 1992b). These residues form a tail that folds into the lipoamide binding site of the catalytic centre.

The FAD cofactor is present in an extended conformation, forming hydrogen bonds with the FAD binding domain and the central domain via its pyrophosphate moiety. The flavin ring of FAD separates the channel forming the dihydrolipoamide binding site and the NAD^+ binding site.

The lipoamide binding site of L-protein forms an 11-Å deep canal with a narrow aperture of several hydrophobic residues exposed to solvent. These residues have the potential to interact with the long aliphatic chain of dihydrolipoamide (Faure *et al.*, 2000). While the adenine ring of NAD^+ fits into a cleft surrounded by hydrophobic residues in a β-α-β region as seen in *P. putida* E3, the positive charge on NAD^+ prevents the substrate from binding completely to its putative binding domain.

The catalytic centre located at the dimer interface (see Figure 7.1) is formed from the flavin ring, a disulphide bridge in monomer A and a His residue in monomer B. Interaction of residues from each monomer is essential for the formation of the catalytic centre and explains the requirement of the dimeric form for activity. In the L-protein the major interactions involve Tyr19, Cys45, Cys50 and Glu329 of monomer A (Tyr16, Cys48, Cys53 in *P. fluorescens*) and His449, Glu454, His469 of monomer B (His450, Glu455, His470 in *P. fluorescens*).

The catalytic mechanism of E3 involves a redox ping-pong reaction where in the oxidised state the disulphide bridge is closed. During catalysis, the dihydrolipoamide is oxidised and the bridge is opened. A charge transfer complex is formed between Cys50 and the FAD cofactor. When FAD is reduced, NAD^+ can enter its binding site in order to be reduced. The adenine ring of NAD^+ binds independently from the redox state of FAD but the nicotinamide component only enters the binding site when the FAD cofactor is reduced (Mattevi *et al.*, 1992b;1993b). The NAD^+ substrate is reduced in turn oxidising the charge transfer

complex and the FAD cofactor is returned to its oxidised state ready to oxidise the next dihydrolipoamide substrate.

The crystal structure of the L-protein has already shown that the recombinant protein can be expressed in *E. coli* as a dimer and its activity depends on dimerisation. In this chapter the dimeric structure of recombinant potato E3 and pea L-protein (mature form) will be studied and their stability will be determined by using chemical denaturants. The insoluble precursor form of the L-protein from pea GDC which includes the N-terminal targeting sequence was also expressed. The targeting sequence is thought to hold the protein in a 'loosely' folded translocation-competent state until it reaches its final destination within the cell where it is cleaved allowing the mature protein to fold completely. Determination of secondary / tertiary structure of pre-E3 will be attempted and its stability to chemical denaturant will be compared to the mature form.

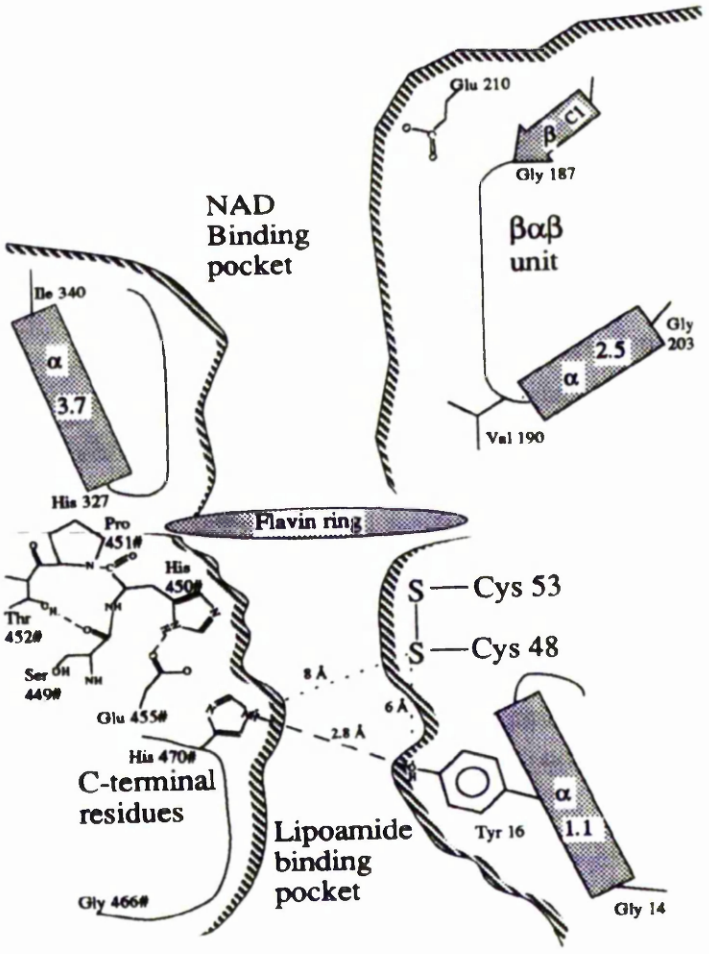


FIGURE 7.1: Schematic Representation of the Catalytic Centre of *P. fluorescens* E3

The diagram outlines the two channels of the catalytic centre separated by the flavin ring into the two substrate binding sites. The side chains of Tyr16 and His470 are involved in hydrogen bonds to each other and are more than 6Å away from the reactive disulphide bridge. However, they have important influences on the redox properties of the enzyme.

(Taken from Mattevi *et al.*, 1993b)

7. 2. IMMUNO-DETECTION OF THE RECOMBINANT E3s

Recombinant forms of the potato E3 and pea L-protein (E3 counterpart in GDC) were expressed and purified from *E. coli* as discussed in chapters 3 and 4 respectively. In addition to the specific activity being consistent with that expected for recombinant pea L-protein (E3), its identity was verified immunologically using antisera raised against purified native pea mitochondrial E3. Recombinant potato E3 was also probed using the same antibodies to check for the expected degree of cross-reactivity.

Sample Preparation

Potato E3 and pea L-protein (E3) were purified from crude *E. coli* extracts as described in Chapter 4 using the POROS MC Zn^{2+}-imidoacetate column (see Methods section 2. 2. 20). E3 samples were resolved on duplicate SDS polyacrylamide gels (see Methods section 2. 2.17). One gel was stained with Coomassie Blue while the other was electrophoretically transferred onto nitrocellulose for Western blotting (see Methods section 2. 2. 18). A positive control was provided by native pea mitochondrial E3 purified from crude pea extract (gift of Dr. A Kochhar, University of Glasgow).

Results

The mature forms of both the pea L-protein and potato E3 have a similar subunit M_r of approx. 55, 000 as determined by SDS-PAGE as seen in Figure 7.2 panel A, lanes 2 and 3 respectively. The native pea E3 band is present with the same M_r value (Figure 7.2 panel A, lane 4). As expected the pea precursor form migrates more slowly with a M_r of approx. 60, 000 (Figure 7.2 panel A, lane 1).

Figure 7.2 (panel B) show the results of the immuno-detection using the specific pea mitochondrial E3 antibodies. All E3 samples cross-react with the anti-pea E3 serum indicating that the recombinant pea L-protein (lanes 1 and 2) is immunologically similar to the E3 component of PDC.

Cross-reaction of potato E3 with the pea E3 antisera suggests that the E3 is similar to that of pea E3 to such a degree that it also it recognised by the antibody (lane 3).

The native E3 is used as a positive control, but unequal loading explains the varied intensity of the ECL signal obtained for the recombinant L-protein and E3 to the native E3. This is due to unreliable measurement of the native E3 concentration due to the presence of BSA in the purification procedure (Figure 7.2, panel A, lane 4).

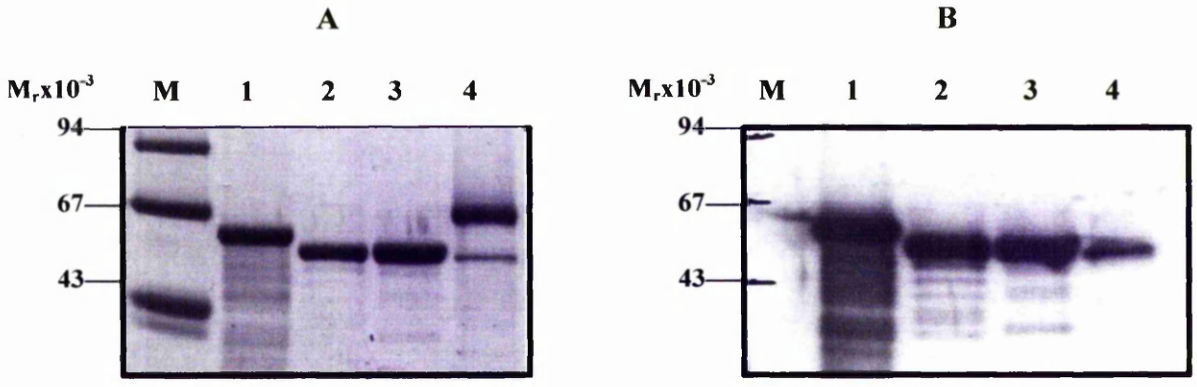


FIGURE 7.2: Immunological Analysis of Recombinant L-protein and E3

Duplicate gels were resolved with E3 samples for Coomassie Blue staining (panel A) and immunoblotting (panel B).

Lane 1: pea L-protein precursor, purified from inclusion bodies, approx. M_r 60, 000

Lane 2: pea mitochondrial L-protein, approx. M_r 55, 000

Lane 3: potato mitochondrial E3, approx. M_r 55, 000, cross-reacts to a similar extent as pea E3

Lane 4: purified native pea mitochondrial E3. The lower band corresponds to E3 as shown in panel B, detected by the anti-E3 antisera. The higher band is BSA included in the pea E3 preparation.

M: relative molecular mass markers

7. 3. INVESTIGATING THE DIMERISATION OF RECOMBINANT E3

As shown in the previous chapter assembly of recombinant E2s into high M_r oligomeric cores was achieved after expression in *E. coli*. The E3 component of PDC exists as a dimer formed from either two identical subunits or uniquely in some plants from two non-identical subunits. It was therefore important to show whether the recombinant E3s expressed in *E. coli* were also capable of forming their native dimeric states.

7. 3. 1. Size Exclusion Chromatography

Size exclusion chromatography on a Superose 6 column (25ml bed volume) was utilised initially to estimate the native M_r values of recombinant E3s. A standard curve was constructed by plotting the elution volumes of a series of marker proteins of known M_r values on a Superose 6 column as described if 2. 2. 28.

Sample Preparation

Purified recombinant potato E3 and pea L-protein were dialysed into the column running buffer (see Methods section 2. 2. 28) and then resolved in a similar manner to the marker proteins.

Results

Similar results for E3 and L-protein were obtained. Both eluted from the Superose 6 column as a distinct single species. From the standard curve the molecular mass of this species was estimated to be 100, 000 kDa. With respect to the monomer M_r of 55, 463, it can be determined that the pea L-protein eluted from the column in its dimeric form. The elution of the potato E3 enzyme also reveals its existence as a dimer with an estimated M_r of 100, 000, with respect to its monomer subunit M_r of 52, 015.

Precursor pea L-protein was also subjected to size exclusion chromatography; however it eluted as a broad peak (a series of aggregated species) over a wide range of M_r values. No identification of dimeric or other discrete oligomeric species could be identified. It is apparent that the precursor elutes as a range of various sizes consistent with the formation of non-specific aggregates. This is not surprising as this protein is insoluble when over-expressed in *E. coli*.

These results suggest that the mature forms of E3 and L-protein either potato or pea mitochondrial, both assembled in *E. coli* into their native dimeric structures while the precursor has a tendency to form large aggregates.

7. 3. 2. Cross-linking Studies on E3: Confirmation of Dimeric State

Resolution of E3 on SDS-polyacrylamide gels results in the enzyme being denatured and only the monomeric subunit M_r can be measured and visualised. However after chemically cross-linking the E3 enzyme with glutaraldehyde, the enzyme will migrate with a molecular mass equivalent to that of the crosslinked protein, in this case the expected dimeric state of E3.

Sample Preparation

E3 was subjected to cross-linking treatment with glutaraldehyde in the presence and absence of GdmCl which promotes dissociation and unfolding of E3 subunits (see Methods section 2. 2. 30). Subsequently samples were resolved on SDS-phosphate buffered gels as described in Methods section 2. 2. 31.

Results

Resolution of cross-linked E3 polypeptides is presented in Figure 7.3 showing potato E3 and pea L-protein (mature and precursor). The potato E3 (panel A) SDS denatured monomer (M_r 50, 000) was visible as a single band (lane 1). The following lanes in the gel correspond to cross-linked E3 monomer (lanes 2 and 3) denatured with GdmCl (2M). Cross-linking of non-denatured E3 (lanes 5 and 6) shows the presence of a putative dimeric species with an apparent M_r of 80, 000. This was smaller than expected given that the monomer mass was 50, 000. However the cross-linked monomer also displays an increased mobility (M_r 40, 000) resulting in the M_r of dimeric E3 equalling that of the cross-linked dimer present. The altered migration of cross-linked monomers and dimers reflects the occurrence of intramolecular cross-linking which prevents complete unfolding on addition of Laemmli sample buffer.

Similar results were observed for pea L-protein (panel B) with the denatured cross-linked E3 resolving at lower mass than the monomer (M_r 40, 000). Cross-linked

dimer resolves with an apparent M_r approx. 80, 000 again smaller than expected but consistent with dimer formation.

Cross-linking of the pea L-protein precursor (panel C) indicates the presence of aggregates as no protein bands are visible within the resolving range of the gel. The monomer resolves as normal with a M_r value of approx. 60, 000 while the denatured monomer subunit M_r was lower at approx. 42, 000. Cross-linked species are detected at the top of the gel suggesting their presence as high M_r aggregates.

In summary cross-linking of these recombinant E3s indicate that they are dimeric in structure while pre-E3 from pea is present as high M_r aggregates.

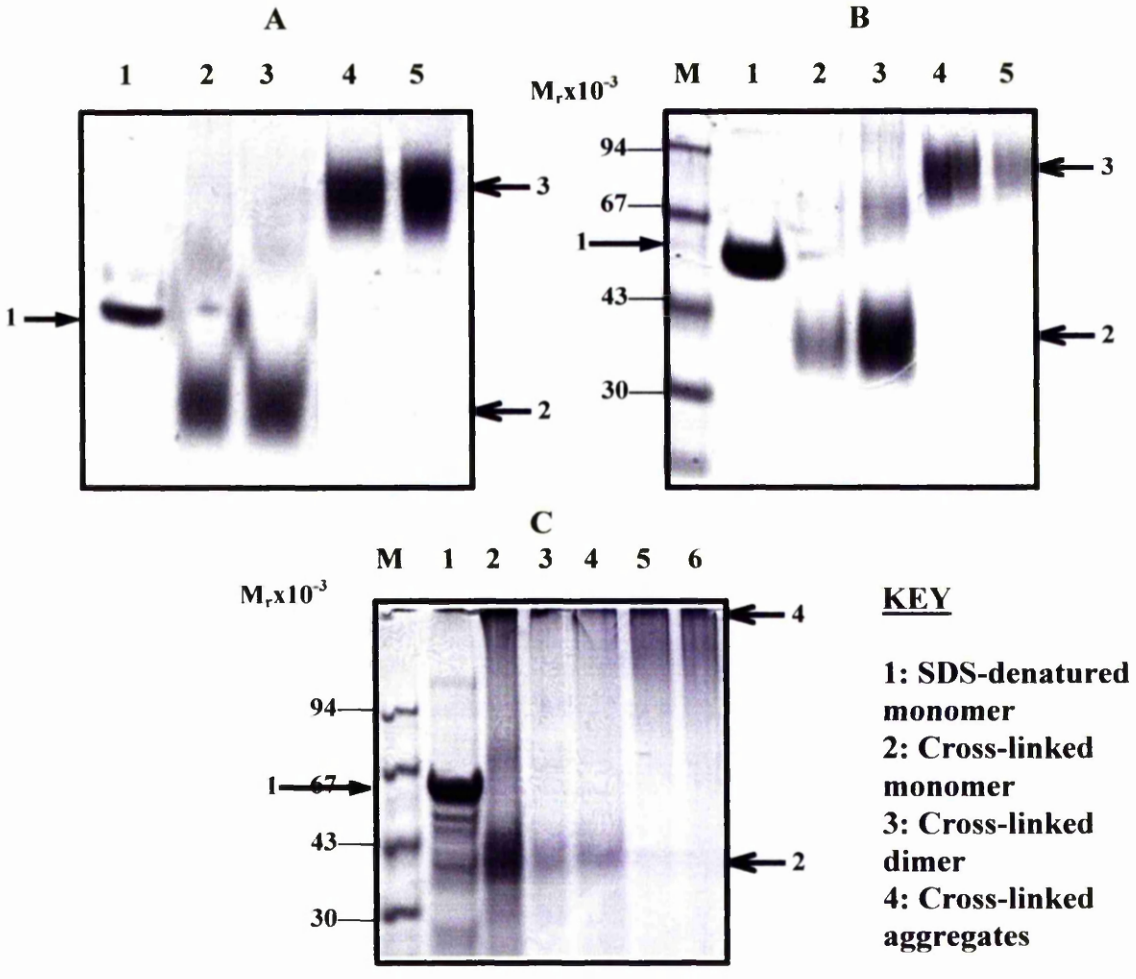


FIGURE 7.3: PAGE Analysis of Cross-linked E3 Enzymes

Panel A, potato E3, panel B, pea L-protein and panel C is pea L-protein precursor. Lane 1: SDS-denatured monomer subunit as seen in previous gels. E3 (50μg) was denatured in 2M GdmCl prior to cross-linking with glutaraldehyde (lanes 2 and 3). Whereas lanes 4 and 5 show native E3s after cross-linking.

M: relative molecular mass markers

7. 4. UNFOLDING STUDIES OF E3 AS MONITORED BY CIRCULAR DICHROISM AND FLUORESCENCE

Circular dichroism (CD) refers to a spectroscopic property that is sensitive to molecular conformation. CD measurements can be performed in two modes. In the near-ultraviolet (UV) CD (340-250nm) region the environment of aromatic amino acids such as tyrosine, phenylalanine and cystinyl groups, but principally tryptophan can be detected reflecting the tertiary and quaternary structure of the protein. In the far-UV CD (260-178nm) the amide linkage is the major chromophore that is detected, thereby the secondary structure of the protein (α-helix, β-sheet, β-turn) can be determined. Fluorescence changes in the protein reflect the loss of tertiary structure as a result of the exposure of internally located tryptophan residues to the external solvent. Chromophores such as tryptophan residues are quite rigid and inflexible. When they are excited by light they do not return to ground state by losing heat instead they lose a portion of energy as light (fluorescence).

Unfolding potato E3 and pea E3 samples with GdmCl, allows complementary measurement of the stability of the native secondary and tertiary structure by far-UV CD and fluorescence.

Sample Preparation

Purified E3 samples were dialysed against at least 6l of 50mM KPi, pH7.2, 150mM NaF over 6h with several buffer changes. Samples were concentrated using centrifugal concentrators (see Methods section 2. 2. 26). Concentrated E3 samples were incubated with a range of GdmCl concentrations prior to CD and fluorescence measurements (see Methods section 2. 2.32 and 2. 2. 33). E3s were also incubated with GdmCl for 15min prior to the assay of E3 activity (see Methods section 2. 2. 23c).

Results

7. 4. 1. Unfolding Potato E3 and Pea L-protein (E3) With GdmCl

The unfolding of E3 by GdmCl was monitored by changes in intrinsic protein fluorescence and far-UV CD and expressed in terms of the changes in these parameters relative to the total changes occurring between 0M and 6M GdmCl. The size of the ellipticity at 222nm and fluorescence emission maximum indicated that the E3 proteins were completely unfolded in 6M GdmCl.

Unfolding of potato E3 as monitored by CD and fluorescence is presented in Figure 7.4, along with the loss of activity due to GdmCl denaturation. Complete loss of E3 activity occurred with 2M GdmCl corresponding to the loss of approx. 80% of secondary structure. E3 was pre-treated with GdmCl before the activity was assayed in the absence of the chemical denaturant. The activity of the pre-treated E3 was performed in duplicate with values differing by a maximum of 10%. From the CD results the loss of secondary structure was 50% complete by 1.85M GdmCl, complemented by 50% loss of tertiary structure. However there was a noticeable lag in unfolding between 1-1.5M GdmCl, seen in both CD and fluorescence whereas a dramatic loss in E3 activity is observed over this range. Local structural perturbations in key regions of E3 essential for enzymatic activity or dissociation of E3 dimers prior to unfolding could account for this discrepancy.

Fluorescence of E3 in the absence of GdmCl was low, presumably reflecting energy transfer to the tightly bound FAD cofactor (West *et al.*, 1995). When E3 was incubated in 6M GdmCl and the fluorescence measured there was a red shift in emission maximum from 340nm-358nm due to the change in the tryptophan residues environment and an enhancement in fluorescence intensity as a result of cofactor dissociation or exposure (Figure 7.5).

Similar results were observed for pea L-protein (E3). Where 50% loss of secondary structure was found to occur between 1.5-1.6M GdmCl and 50% loss of tertiary structure observed between 1.4-1.5M GdmCl, (Figure 7.6). The activity of

pea L-protein (E3) declined to less than 5% of its original value by 2M GdmCl. In this case loss of E3 activity closely paralleled the loss of secondary and tertiary structure as monitored by CD and fluorescence properties.

An enhancement in fluorescence intensity of L-protein was obtained when incubated with increasing concentrations of GdmCl, along with the characteristic red shift in emission maximum (340-358nm), resulting from the change in the tryptophan environment while the increase in intensity reflects the loss of the FAD cofactor. The raw fluorescence data are not shown but was similar in appearance to that for potato E3 (Figure 7.5).

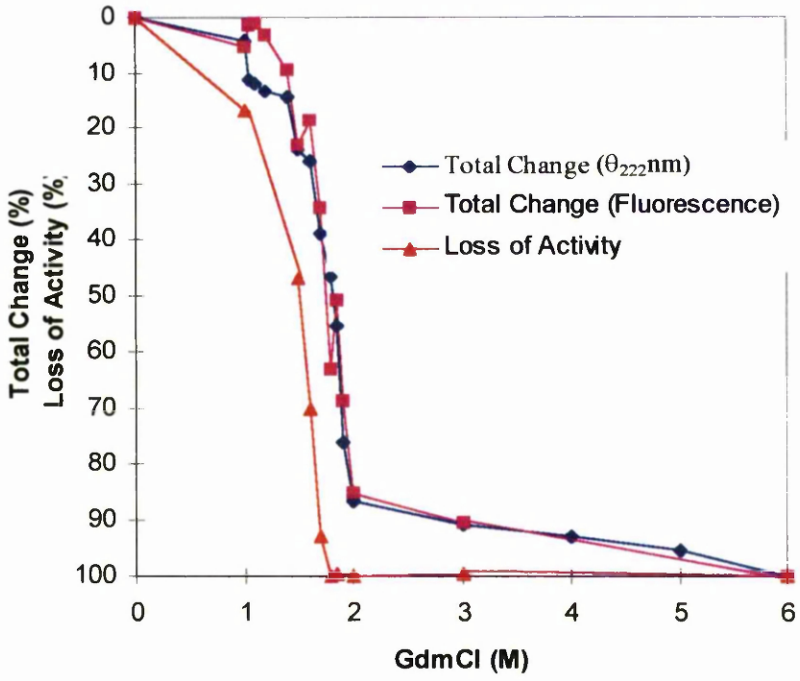
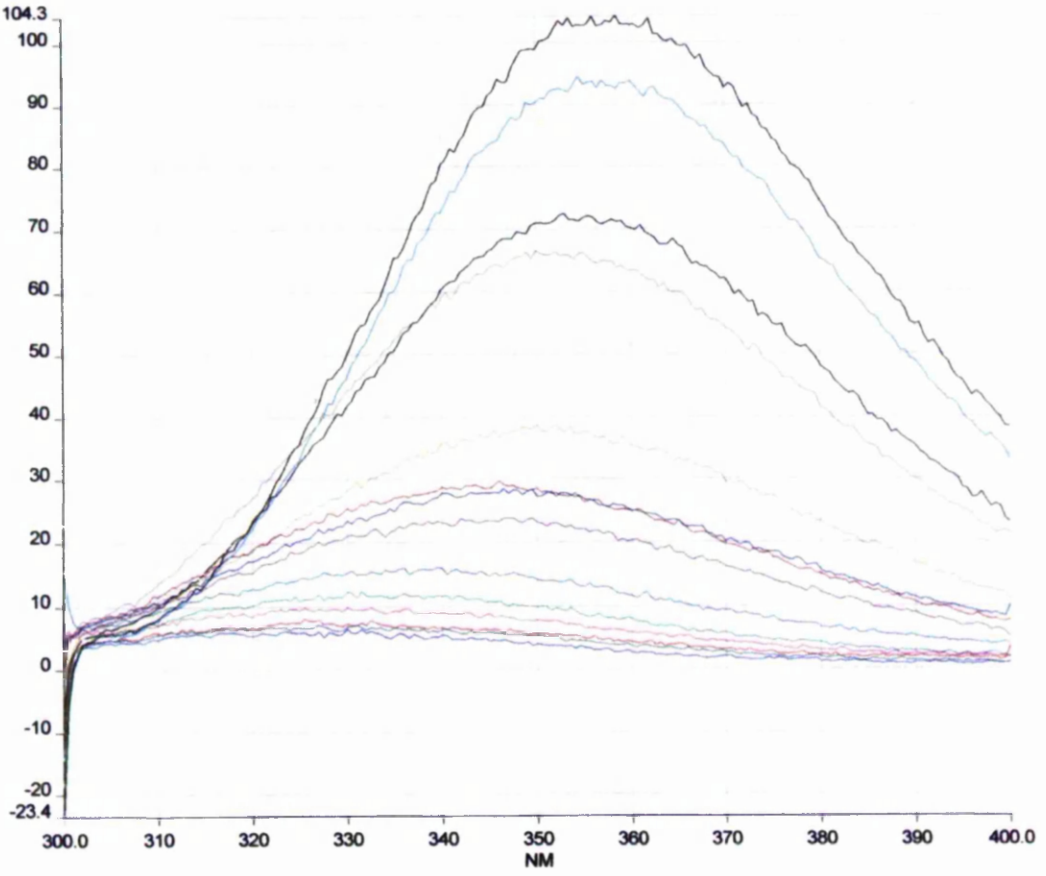


FIGURE 7.4: Unfolding of Potato E3 by GdmCl as Monitored by Loss of Activity and Changes in CD and Fluorescence Properties

E3 (500μg) was incubated in 50mMKPi pH7.2, 150mMNaF containing GdmCl for 15min at 20°C prior to measurement. In CD and fluorescence the changes were expressed relative to the total change observed between 0M and 6M GdmCl. CD was measured as change in ellipticity (θ) at 222nm. Fluorescence was measured with an excitation wavelength of 290nm and an emission wavelength of 355nm was chosen to show maximal fluorescence changes. E3 (70μg) was incubated as above, E3 activity was measured and expressed relative to a control sample from which GdmCl was omitted. E3 assays were performed in duplicate and differed by less than 10%.

FIGURE 7.5: Fluorescence Data for the Unfolding of Potato E3 with GdmCl

E3 (500μg) was incubated in 50mM KPi pH7.2, 150mM NaF containing GdmCl for 15min at 20°C prior to fluorescence measurements. An excitation wavelength of 290nm was used and the fluorescence of the unfolding protein was measured between 300-400nm.


104.3
100
90
80
70
60
50
40
30
20
10
0
-10
-20
-23.4
300.0
310
320
330
340
350
360
370
380
390
400.0
NM
PotE3 0M
PotE3 1.05M
PotE3 1.1M
PotE3 1.2M
PotE3 1.4M
PotE3 1.5M
PotE3 1.6M
PotE3 1.7M
PotE3 1.85M
PotE3 1.8M
PotE3 1.9M
PotE3 1M
PotE3 2M
PotE3 3M
PotE3 6M

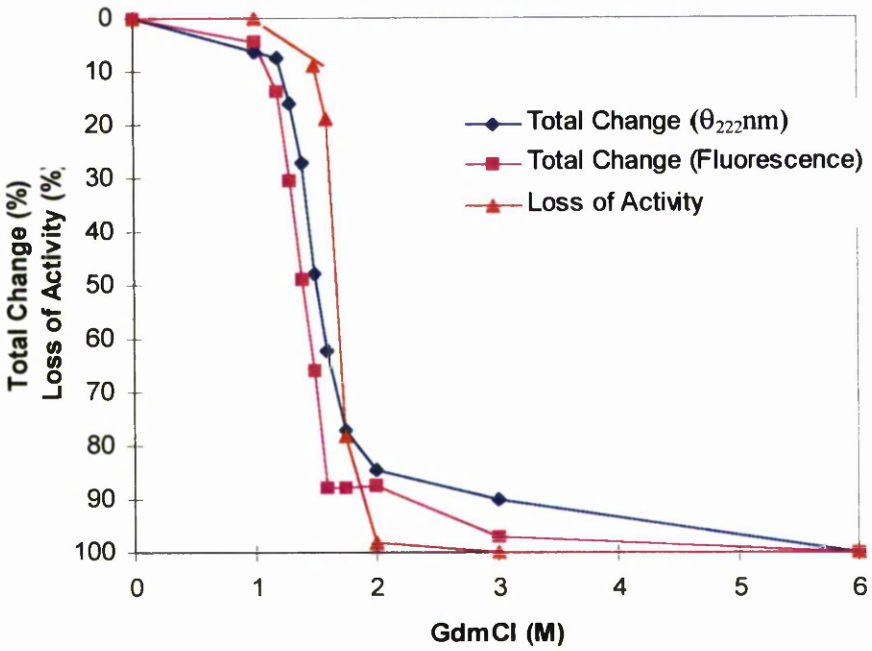


FIGURE 7.6: Unfolding of Pea L-protein by GdmCl as Monitored by Loss of Activity and Changes in CD and Fluorescence Properties

Pea L-protein (425µg) was incubated for 15min in 50mM KPi pH7.2, 150mM NaF containing GdmCl at 20°C prior to measurements. Changes in structure were expressed relative to the total change observed between 0M and 6M GdmCl. CD data were expressed as change in ellipticity (θ) at 222nm. Fluorescence results were recorded using an excitation wavelength of 290nm. An emission wavelength of 360nm was chosen to show maximal fluorescence changes.

E3 (70µg) was incubated in GdmCl as above before activity was measured and expressed relative to a control sample to which GdmCl was omitted. E3 assays in duplicate differed by less than 10%.

7. 4. 2. Unfolding Studies Pea L-protein Mature and Precursor Forms

The unfolding of pea L-protein mature and precursor forms by urea was monitored by changes in protein fluorescence and far-UV CD and expressed in terms of the changes in these parameters relative to the total changes occurring between 0M and 7.5M urea.

GdmCl was not used in this case, as precipitation occurred in the L-protein precursor sample on incubation with low GdmCl concentrations. Therefore, urea was utilised as the denaturant and the unfolding of the mature L-protein repeated for comparison. One point to note is that the precursor L-protein sample contains N-lauroylsarcosine (0.5% v/v), the sample is diluted in order to perform both the CD and fluorescence measurements, thereby decreasing the content of detergent in the sample approx. 7.5 fold.

Ellipticity measurements at 222nm and fluorescence emission data indicated the E3 precursor was completely unfolded in 7.5M urea. However unfolding was not observed for the mature E3. In Figure 7.7, mature pea E3 was stable in the presence of 7.5M urea and retained the majority of its secondary structure with no significant loss in activity occurring on the dilution of the chemical denaturant. Comparing the 0M and 7.5M urea samples in fluorescence and CD the results suggest that the tertiary structure was more sensitive to urea than the secondary structure. The greater change in structure as detected by fluorescence measurements may indicate slight perturbations of the tertiary structure corresponding to the tryptophan residue environment.

From the fluorescence data (Figure 7.8) a small red shift was observed between the 0M to 7.5M urea, the intensity was slightly enhanced (1.3 fold) indicating a minor local structural changes surrounding the FAD cofactor.

The precursor form was solubilised with detergent and so its state of folding was unknown; however, the precursor form was sensitive to urea and did exhibit unfolding as monitored by CD and fluorescence. In Figure 7.9, the loss of

secondary structure was gradual over a wide range of urea concentrations, with 50% loss observed between 4.0-4.2M. However, the fluorescence results show a dramatic loss of tertiary structure in the range of 4-5M urea, this may reflect the dissociation of aggregates. These results suggest that some degree of secondary or tertiary structure is present in the precursor form. Whether this is indicative of aggregates dissociating or of the detergent solubilised precursor unfolding is unclear. What can be taken from these results is that the additional 31 amino acids at the N-terminal end of the L-protein protein influences the correct folding and assembly of the precursor protein.

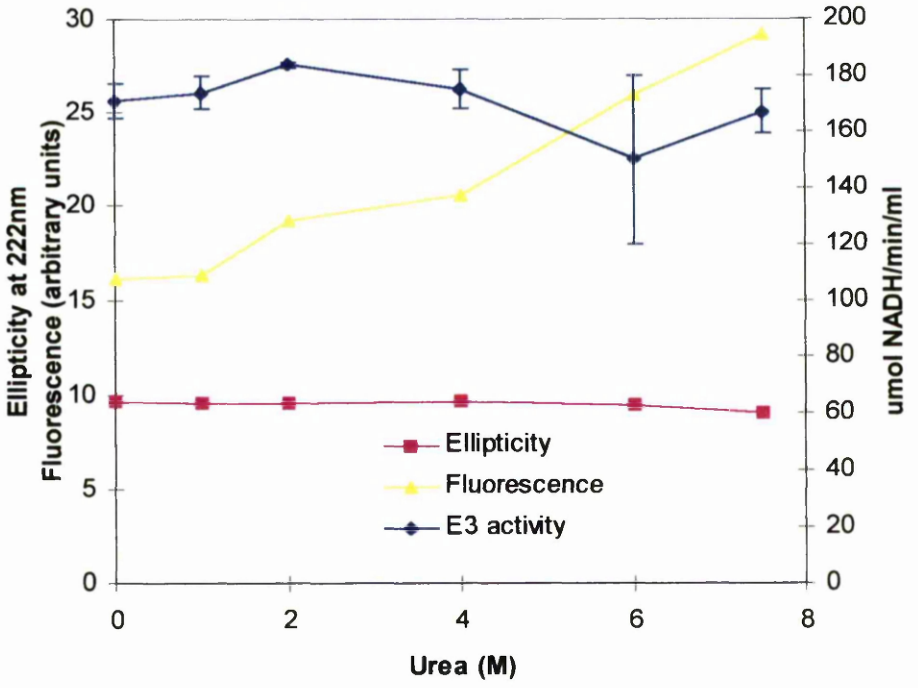
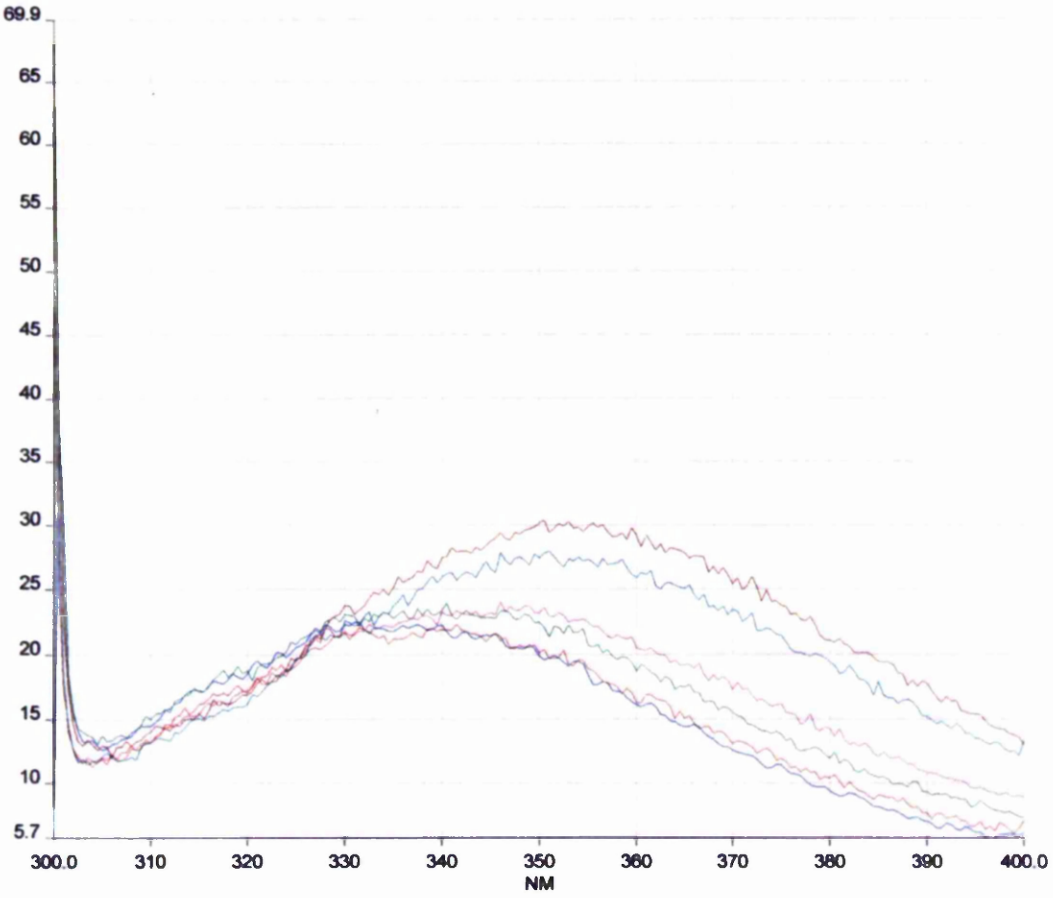


FIGURE 7.7: Unfolding of Pea L-protein by Urea as Monitored by Loss of Activity and Changes in CD and Fluorescence Properties

E3 (200μg) was incubated for 15min in 50mM KPi pH7.2, containing urea at 20°C before measurements. Changes in structure monitored by CD was expressed as the measured ellipicity (θ) at 222nm. Fluorescence results were recorded as the change in fluorescence arbitrary units at 360nm (emission wavelength) while the excitation wavelength was 290nm. E3 activity was measured and expressed as μmol NADH min^{-1} ml^{-1}.

FIGURE 7.8: Fluorescence Data of the Unfolding of Pea L-protein with Urea

L-protein (200μg) was incubated for 15min at 20°C in 50mM KPi pH7.2, containing urea prior to fluorescence measurements. Fluorescence emitted by the unfolding pea L-protein with urea were recorded from 300-400nm, using an excitation wavelength of 290nm.


69.9
65
60
55
50
45
40
35
30
25
20
15
10
5.7
300.0
310
320
330
340
350
360
370
380
390
400.0
NM
E30MU
E31MU
E32MU
E34MU
E36MU
E375MU

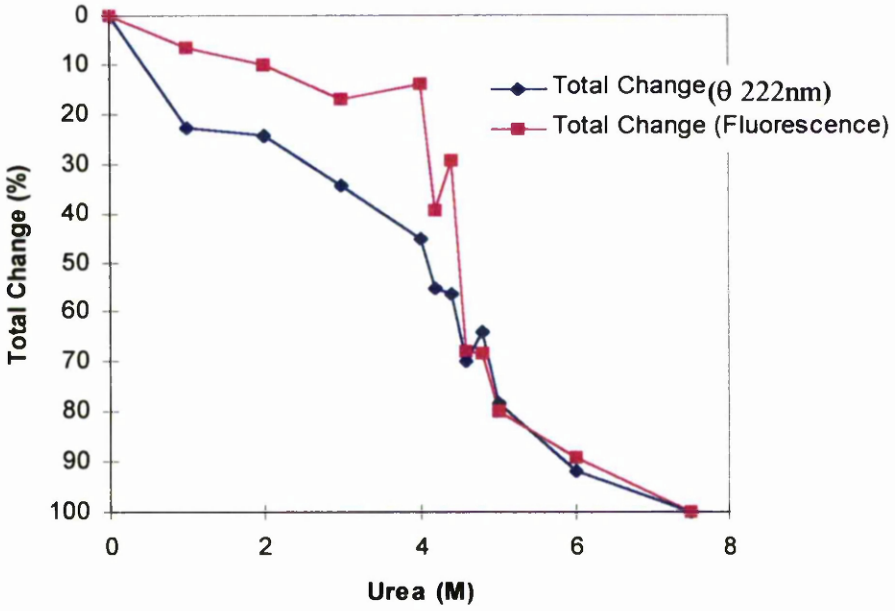


FIGURE 7.9: Unfolding of L-protein Precursor by Urea as Monitored by Changes in CD and Fluorescence Properties

L-protein precursor (200μg) was incubated in 50mM KPi pH7.2 containing urea for 15min at 20°C before measurements were recorded. Changes in structure were expressed relative to the total change observed between 0M and 7.5M urea. CD data were expressed as changes in ellipticity (θ) at 222nm and fluorescence data were obtained using an excitation wavelength of 290nm and an emission wavelength of 332nm was chosen to show fluorescence change.

7. 5. DISCUSSION

Expression of the three E3s (potato E3, pea L-protein, mature and precursor) produced 10-30mg enzyme per litre of culture. Both the potato E3 and pea L-protein were soluble and fully active suggesting that they were capable of folding correctly within their *E. coli* host. However the precursor form was found to be insoluble forming inclusion bodies, indicating that the N-terminal targeting sequence of 31 amino acids prevents folding into the native structure. In contrast the N-terminal His-tag sequence of 25 amino acids does not interfere with the folding and assembly of these recombinant mature enzymes. The work in this chapter involved initial structural analysis of the folding state and stability of recombinant E3s without resorting to detailed X-ray structural determination.

Recombinant potato E3 and pea L-protein were immunologically similar as both enzymes were recognised by antibodies raised against native pea mitochondrial E3. This result was expected for the pea L-protein as it has been previously reported that the L-protein and E3 from pea mitochondria are the same gene product (Bourguignon *et al.*, 1996).

Activity measurements of the purified enzymes show that both were fully active, with specific activities similar to those published, suggesting that they have folded correctly into their native dimeric structures. From previous reports the dimeric structure of E3 is essential for the activity of the enzyme. Verification of the dimeric structure of these recombinant E3s was determined using size exclusion chromatography giving an approx. M_r value of 100, 000 for both enzymes. Glutaraldehyde crosslinking studies revealed E3s in solution are present as dimers while the insoluble precursor was present as high M_r aggregated species.

Treatment of the E3s with chemical denaturants such as GdmCl and urea followed by measuring changes in spectroscopic properties by CD and fluorescence allowed the loss of native structure to be studied in relation to overall loss of enzyme activity. The unfolding of potato E3 and pea L-protein were consistent with conditions reported for unfolding native bovine E3 (West *et al.*, 1995). The

majority of potato E3 was unfolded at 2M GdmCl with no residual secondary or tertiary structure detectable by this stage indicating a 'random coil' conformation. In agreement with this result, 99% loss of E3 activity was detected on rapid dilution after pre-treatment with GdmCl. There was no significant difference in the unfolding of secondary or tertiary structure as indicated from both sets of results. However the fluorescence results revealed the release of the FAD cofactor after GdmCl incubation with an enhancement in the fluorescence intensity. This enhancement is thought to occur as a result of alleviating the quenching effect of the folded protein on the FAD moiety. As the protein unfolds the quenching effect on the FAD decreases until the FAD cofactor is solvent accessible and can dissociate from the protein. Previous work on the role of FAD has shown that it is essential for the correct folding and restoration of activity of the denatured E3 enzymes (Lindsay *et al.*, 2000).

Pea L-protein unfolds in a similar manner to potato E3 when incubated with GdmCl. Unfolding was complete at 2M GdmCl with 98% loss of activity being detected by this stage.

Comparing the stability of the mature and precursor structure to urea treatment revealed that mature L-protein was apparently stable while the precursor was not. The mature L-protein was not denatured by treatment of urea and no loss of activity was detected even after pre-treatment in 7.5M urea, although slight structural perturbations were detected in the tertiary structure. Fluorescence results showed a red shift in emission wavelength and a small enhancement in intensity. This small spectroscopic change may result from local minor structural changes allowing the tryptophan residue or the FAD cofactor access to the external environment.

The precursor was more sensitive to denaturation by urea showing a dramatic change in secondary and tertiary structure in the range 4-5M urea. These results suggest that some degree of secondary/tertairy structure ('molten globule') is present although pre-E3 forms aggregates within the cytoplasm of the *E. coli*. The presence 31 amino acids N-terminal targeting sequence clearly prevents the

precursor from folding into its native structure. Mitochondrial targeting sequences show no conserved homology blocks but are thought to adopt a common amphiphilic α-helical conformation. Mitochondrial targeting sequences are enriched in alanine, leucine, arginine and serine residues but sparse in aspartic acid, glutamic acid, isoleucine and lysine (Heijne *et al.*, 1989). In the pea L-protein, its N-terminal targeting sequence is consistent with the expected amino acid distribution and contains only one glutamic acid and one lysine. Comparison of several targeting sequences of non-homologous proteins found that the C-terminal 10-15 residues did not conform to the classical amphiphical α-helix. This is true for the pea L-protein targeting sequence, an amphiphic α-helix is formed when the C-terminal 13 residues are removed from the helical-wheel plot, presented in Figure 7.10. The majority of hydrophobic residues are found on one side of the helix, although two hydrophilic residues (basic and acidic) are also present.

The purpose of such an α-helix allows the precursor protein to be targeted to the mitochondria and promotes specific interactions with receptors on the surface of the organelle. The targeting sequence may hold the enzyme in a loosely / partially folded form. *In vivo*, the 'unfolded' state would be aided by the association of molecular chaperones that ensure that the precursor is maintained in an appropriate import-competent form prior to uptake into its final destination within the cell.

Overexpression of heterologous proteins in *E. coli* tend to form inclusion bodies rather than remain in the cytoplasm. Proteins in these inclusion bodies are thought to be loosely folded and may be classed as "molten globule" intermediates which can either fold correctly into the native protein or can form aggregates of loosely folded proteins. Dissociation of both the "molten globule" and aggregates may be observed in the CD and fluorescence results, respectively.

In summary, the results indicate that both the potato E3 and mature pea L-protein are expressed and can be purified in their active, dimeric states. As enzyme activity depends on its native dimeric structure and the loss of activity mirrors the

loss of native structure. The precursor L-protein is sensitive to mild chemical denaturants such as urea while the mature L-protein is quite insensitive, suggesting the that loosely folded form of the precursor may form aggregates in solution. Incubation with urea induces the 'loosely-folded' intermediates to unfold completely with increasing urea concentrations.

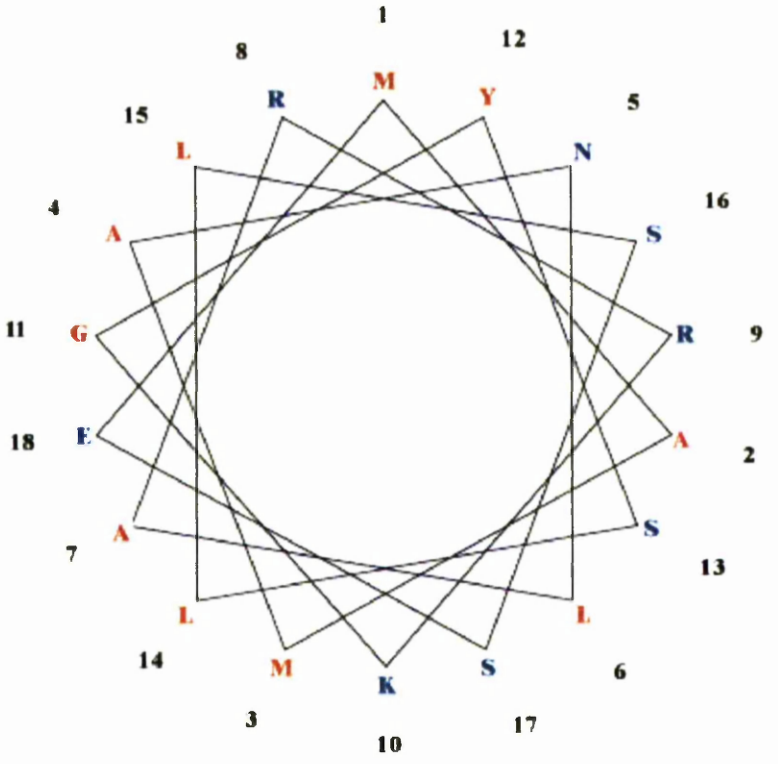


FIGURE 7.10: Helical-wheel Plot of the Pea Mitochondrial L-protein Mitochondrial Targeting Sequence.

The mitochondrial targeting sequence predicted to be 31 amino acids in length did not produce a distinct amphiphic α-helix when plotted on a helical wheel plot, but when the C-terminal 13 residues were removed (seen here) a distinct hydrophobic side was obtained. This plot presents the probable amphiphic α-helix expected for mitochondrial targeting.

Red residues are hydrophobic and blue residues are hydrophilic, the numbers denote the residue order in the sequence starting the initial methionine (M1).

CHAPTER 8

CONCLUSIONS AND FUTURE WORK

8. CONCLUSIONS AND FUTURE WORK

CONCLUSIONS

The majority of research conducted on the pyruvate dehydrogenase complex (PDC) has employed complexes purified from *E. coli*, *A. vinelandii*, *B. stearothermophilus*, yeast (*S. cerevisiae*) and mammals (bovine, human). However, studies on plant PDCs to date, while representing a substantial body of work, have been hindered by the low abundance of enzymes available from plant tissue. Consequently, difficulties have arisen in determining the precise subunit composition, organisation and stoichiometry of these complexes. Moreover, the pyruvate dehydrogenase complex has a dual location in plants adding to the complexity of the characterisation of the organelle specific complexes. Mitochondrial PDC appears similar to mammalian PDC, while a unique chloroplastic PDC is also present that plays an anabolic role in providing acetyl-CoA for *de novo* fatty acid synthesis. Both plant complexes have been poorly characterised with most work performed on crude preparations or partially purified extracts. This work has yielded valuable information on the differing modes of regulation and general kinetic parameters of the plant complexes. Both are regulated by product inhibition, but only the mitochondrial PDC is controlled by reversible phosphorylation. During periods of illumination, when ATP levels are high, the chloroplastic PDC activity is stimulated whereas the mitochondrial PDC is inactivated under these conditions.

Recent access to genome sequence information from a variety of sources combined with the increasing availability of plant cDNAs and ESTs has provided new avenues for producing large quantities of plant enzymes in recombinant form for detailed investigation. This expression in bacteria and the subsequent purification followed by *in vitro* reconstitution of the functional multienzyme complex will greatly facilitate future work in this area. Regulation of acetyl-CoA production by chloroplastic PDC is of particular interest since increasing oil yields

or manipulating oil production in plants such as soya bean and rape seed is of great commercial interest.

Cloning and Purifying Active PDC Components

Molecular cloning of plant PDC components and their subsequent overexpression in *E. coli* has the potential to produce a readily available and reliable source of recombinant protein. Expression of these recombinant proteins in *E. coli* allows high yields (in the range of 5-30mg per 1l of culture) of purified enzymes to be obtained in contrast to their low abundance purified from native plant sources.

Specific purification of these expressed PDC components was achieved by affinity chromatography using the incorporated expression tag (His-tag or GST-fusion protein) attached to the N-terminal region of the recombinant protein. The recombinant enzymes were purified to a high state of purity, as judged by SDS-PAGE and Coomassie blue staining. Potato and pea E3 were purified in their active states exhibiting specific activities similar to that of native pea mitochondrial E3 (Conner *et al.*, 1996a, b). However, purified E2s displayed much lower specific activities than that reported for the recombinant maize mitochondrial E2 (Thelen *et al.*, 1999b). These lower activities may be partly accounted for by the presence of aggregated E2 and small M_r species (monomeric or trimeric) detected by sedimentation equilibrium experiments. Gel filtration studies suggest that these small M_r proteins do not make up a major proportion of the overall E2 population. It should also be noted that the values reported by Thelen and co-workers are 10-fold higher than those published for human E2 (Yang *et al.*, 1997). A comprehensive investigation of the optimal assay conditions for plant E2s is also required to elucidate the basis for these apparent discrepancies.

The L-protein precursor expressed in *E. coli* included the 31 amino acid mitochondrial targeting sequence in addition to the His-tag. As the recombinant precursor protein was insoluble, it was subsequently purified from inclusion bodies using N-lauroylsarcosine. Solubilised L-protein was assayed for E3

activity. However, no activity was detected suggesting that the precursor was not present as a dimer but as either unfolded or 'loosely' folded forms which can aggregate even in the presence of detergent.

In addition to these components, the *A. thaliana* mitochondrial E1α and E1β subunits and four E3 isoforms have all been cloned and expressed in *E. coli* as His-tagged proteins. Characterisation of these proteins or *in vitro* reconstitution of the mitochondrial complex was not carried out due to insufficient time.

Lipoylation and Proposed Core Structure of PDC-E2

Post-translational modification involving lipoic acid attachment to the recombinant E2 lipoyl domains from *E. coli, B. stearothermophilus* and human sources expressed in *E. coli* has been shown to occur in previous work (Ali and Guest, 1990; Dardel *et al.*, 1990; Quinn *et al.*, 1993). However, there is no evidence reported on whether plant E2s will also be recognised by the lipoylating machinery of *E. coli*. It was therefore necessary to investigate the lipoylation state of plant E2 lipoyl domains when expressed in this host. Pea chloroplastic E2 lipoyl domains seem to be lipoylated with the appearance of a band shift on SDS-PAGE. The appearance of a dominant upper band in the sample grown in media plus lipoate was proposed to be the lipoylated form, while two bands were observed in the minus lipoate sample suggesting a mixture of lipoylated and unlipoylated products. These band shifts were not seen for the *A. thaliana* mitochondrial E2 lipoyl domains, but it is possible that due to the lower resolving power of the SDS-gel at the higher M_r range these two bands were not resolved. However, a monoclonal antibody (mAb) specific for the lipoylated human PDC-E2 specifically recognised the mitochondrial E2 lipoyl domains and did not cross-react with the chloroplastic E2 lipoyl domain. Aligning the protein sequences for human, plant mitochondrial and chloroplastic lipoyl domains revealed the poor sequence identity between the chloroplastic E2 lipoyl domains and the human lipoyl domains (35%) while a higher degree of sequence identity (60%) was seen between the human lipoyl domains and the mitochondrial E2 lipoyl domains. The lack of sequence homology between chloroplastic and mitochondrial domains may

explain why the mAb only recognises mitochondrial lipoyl domains. These combined observations suggest that in addition to the presence of lipoic acid the mAb also detects a specific sequence motif exclusively found in lipoyl domains of mitochondrial PDC-E2.

An enhancement was observed in the mAb recognition of the mitochondrial tridomain and full-length E2 expressed in medium plus lipoate suggesting that the lipoyl domains were fully lipoylated while only a proportion of lipoyl domains expressed without lipoate were co-translationally modified. In contrast, the MALDI-TOF mass spectrometry analysis of the tridomains plus or minus lipoate showed no significant difference in mass (61, 744 for minus lipoate and 61, 751 for plus lipoate) indicating that both E2 tridomains were fully lipoylated. It is not clear at present why a mixture of lipoylated and partially lipoylated species was not detected by mass spectrometry.

Lipoylation of the chloroplastic lipoyl domain was determined by MALDI-TOF mass spectrometry. The chloroplastic lipoyl domain expressed in the presence of lipoate had a measured M_r of 48, 455 which is 229 mass units bigger than the measured unlipoylated lipoyl domain (M_r of 48, 226) and can explained by the addition of a lipoyl group (expected value of 190 mass units). This increase in mass may also indicate the presence of an acetyl group (42 mass units) bound to the lipoic acid moiety. Biotinylation (226 mass units) was considered as a possibility owing to the similarity in mass difference but was rejected on the grounds that the specific sequence motif M-K-M is required for recognition by biotinyl protein ligase. MALDI-TOF analysis of the chloroplastic lipoyl domain expressed in growth medium without lipoate showed no mixture of lipoylated or unlipoylated domains as observed in SDS-PAGE. This may be a consequence of variable expression and/ or lipoylation of the different samples employed.

In conclusion, the *E. coli* lipoylating enzymes were capable of recognising both the chloroplastic and mitochondrial lipoylatable lysine residue residing at the tip of the β-turn of the lipoyl domains. However, the expression of chloroplastic lipoyl domains seem to require exogenous lipoate in order for the single lipoyl

domain to become fully lipoylated while the mitochondrial lipoyl domains are either partially or completely lipoylated without lipoic acid supplementation. The increase in the level of lipoylated products expressed in *E. coli* with the addition of exogenous lipoic acid has been previously reported (Quinn *et al.*, 1993), suggesting that the rate of lipoic acid synthesis is the limiting factor in lipoylation.

The pyruvate dehydrogenase complexes from prokaryotes and mammalian sources are composed of multiple copies of E1 and E3 arranged around a multimeric core of E2. Two types of oligomeric E2 core have been found to exist; the 24meric cubic core found in *E. coli* and *A. vinelandii* and the 60meric core found in mammalian PDC and *B. stearothermophilus* PDC. Determination of the oligomeric core structure may add further evidence to the evolutionary origin of the chloroplast or the mitochondria. The "endosymbiont theory" states that both organelles originated from free-living bacteria possibly an aerobic bacteria and phototrophic bacterium which invaded proto-eukaryotic cells and entered into a successful mutualistic union with the host. The PDC of the Gram positive bacteria contains a 60meric E2 core which is also found in mammalian PDC, consistent with the theory that mitochondria originated from a primitive Gram positive aerobic organism.

Purified recombinant mitochondrial and chloroplastic E2 permitted preliminary studies into the oligomeric structure assembled in *E. coli*. Their core stoichiometry was determined by measuring their apparent weight-average molecular mass ($M_{w,app}$) by sedimentation equilibrium (SE). The $M_{w,app}$ for chloroplastic E2 was estimated to be 1.0-1.1 MDa suggesting a 20-24meric core agreeing with a 24meric core seen in PDC of Gram negative bacteria. In contrast, the estimated $M_{w,app}$ for mitochondrial E2 was 4.0-5.0 MDa suggesting a 60-80meric E2 core higher than the expected 60meric structure previously observed for plant mitochondrial E2 (Millar *et al.*, 1999; Thelen *et al.*, 1999). PDC of the Gram positive *B. stearothermophilus* and mammals have a 60meric E2 core consistent with the theory that mitochondria originated from a common Gram positive aerobic bacterium. The SE results suggest that plant mitochondria also originate from a similar ancestor while the chloroplastic PDC has evolved from a Gram

negative ancestral bacterium. Determination of accurate M_r values for each E2 core was hindered by the presence of aggregates and small M_r species. However, the sedimentation equilibrium results present a clear distinction between the organisation of the E2 cores from chloroplastic PDC and mitochondrial PDC.

Analysis of the Structure and Stability of Recombinant Plant E3s

Structure function studies on the E3 component focused mainly on the stability of the dimer structure to chemical denaturation and the corresponding loss of enzyme activity. Previous reports state that the dimeric structure of E3 is essential for the enzyme activity owing to the location of the FAD cofactor and catalytic centre at the dimer interface. Chemical cross-linking of the recombinant E3s followed by resolution on SDS-phosphate gels allowed visualisation of their dimeric states with M_r values of approx. 100, 000. The presence of dimeric E3 was not surprising as both purified recombinant E3s were fully active with specific activities consistent with that reported for native pea E3 (Conner *et al.*, 1996).

The relationship between dimeric structure and E3 activity was investigated by treating E3 with GdmCl and following the loss of structure by circular dichroism. Complete loss of potato E3 activity at 2M GdmCl correlated with the major loss of secondary structure. Loss of activity seemed to precede the loss of structure between 1-1.5M GdmCl; this may reflect local structural perturbations possibly at the dimer interface concerning regions essential for activity. Similar results were obtained for the pea L-protein (E3) where the loss of activity paralleled the loss of structure.

Loss of E3 tertiary structure after treatment with GdmCl was detected by changes in intrinsic protein fluorescence. Denaturation of the potato E3 and pea L-protein produced a red shift in the emission wavelength (340 to 358nm) reflecting the change in environment of the tryptophan residue caused by the chemically induced unfolding. An enhancement in the fluorescence intensity was also observed probably mirroring alterations in the fluorescence of the FAD cofactor as it

becomes solvent accessible. In the fully folded state FAD fluorescence is quenched by the surrounding protein.

To investigate the folding status of the insoluble L-protein precursor urea was substituted for GdmCl, which induced precipitation of the polypeptide. The mature L-protein showed little evidence of effects on secondary and tertiary structure with no significant loss of activity being detected after pre-treatment with 7.5M urea. Fluorescence measurements revealed a slight red shift indicating a minor structural change reflecting an alteration in the environment of the single tryptophan.

In contrast to the mature L-protein, the precursor showed distinct structural changes. CD results suggested that there was a gradual unfolding of secondary structure over a wide range of urea concentrations (1M-6M), while the fluorescence results presented a more drastic loss of structure observed in the range of 4-5M urea this may reflect dissociation of aggregates. This difference in structural stability may reflect the occurrence of "molten globules". This term is used to describe partially-folded proteins that can either follow the correct folding pathway to fully folded native structures or form non-specific aggregates. The gradual change in CD signal may correspond to the unfolding of these "molten globule" structures, while the fluorescence results may be accounted for by the dissociation of aggregates previously observed in glutaraldehyde cross-linking experiments. The 31 amino acid targeting peptide permits the formation of elements of secondary structure, suggested by changes in the structure detected by CD and fluorescence, this partially-folded structure may be responsible for the insolubility and sensitivity to the mild denaturant urea. These results fit with the *in vivo* function of targeting sequences, by maintaining a level of partial-folding until the protein reaches its destination where the targeting sequence is cleaved allowing the mature protein to assemble correctly.

FUTURE WORK

The aim of this work was to produce a stable and reliable source of recombinant PDC components that would allow high yields of purified enzymes to be produced. The homogeneous protein purified by the method of affinity chromatography will permit future investigation into the structure and function of these components. Further study of the E2 core assemblies by electron microscopy is required to provide conclusive evidence about the subunit composition and morphologies of the recombinant cores as suggested by sedimentation equilibrium experiments.

Future crystallographic studies, requiring access to large quantities of enzymes on a routine basis, will now be possible as a result of the successful over-expression of these recombinant proteins in *E. coli* and their subsequent purification. This also permits structural work on the E2 cores from chloroplast and mitochondrial PDC to be achieved. To date structure of plant E2s have not yet been determined. However sequence alignments and SWISS modelling suggest that the predicted 3D-structure of the lipoyl domains and core complexes are similar to the mammalian and prokaryotic equivalents.

In vitro reconstitution of each organelle specific complex will permit each complex to be studied as a whole *in vitro*, and detailed comparisons between the two complexes can be performed. Unique features of plant PDCs can also be studied. For example, the presence of four E3 isoforms found in *A. thaliana*; two mitochondrial and two chloroplastic, all four have been cloned and have been expressed *in E. coli.*

The successful cloning of the constituent enzymes from the *A. thaliana* mitochondrial and plastidic PDC will allow future studies on the binding specificity of the E2 for E1 and E3 to be performed. Isothermal titration calorimetry (ITC) and surface plasmon resonance will be used to measure the binding specificity of the GST constructs (mitochondrial E2 tridomain and plastidic E2 didomain) for their complex specific E1 and E3 components. The

purpose of these studies will be to elucidate on the need for or the presence of an E3-binding protein, as seen in human and yeast (*S. cerevisiae*) PDC. These studies will also clarify the need for two organelle specific E3 isoforms as identified uniquely in *A. thaliana*.

REFERENCES

REFERENCES

Aevarsson, A., Seger, K., Turley, S., Sokatch, J. R., Hol, W. G. J. (1999) Crystal Structure of 2-Oxoisovalerate and Dehydrogenase and the Architecture of 2-Oxo Acid Dehydrogenase Multienzyme Complexes. *Nat. Struct. Biol.* **6**, 785-792

Aevarsson, A., Chuang, J. L., Wynn, R. M., Turley, S., Chuang, D. T., Hol, W. G. J. (2000) Crystal Structure of Human Branched-Chain α-Ketoacid Dehydrogenase and the Molecular Basis of Multienzyme Complex Deficiency in Maple Syrup Urine Disease. *Structure.* **8**, 277-291

Allen, A. G. and Perham, R. N. (1991) Two Lipoyl Domains in the Dihydrolipoamide Acetyltransferase Chain of the Pyruvate Dehydrogenase Complex of *Streptococcus faecalis. FEBS Lett.* **287**, 206-210

Ali, S. T. and Guest, J. R. (1990) Isolation and Characterization of Lipoylated and unlipoylated Domains of the E2p Subunit of the Pyruvate Dehydrogenase Complex of *Escherichia coli. Biochem. J.* **271**, 139-145

Ali, S. T., Moir, A. J. G., Ashton, P. R., Engel, P. C., Guest, J. R. (1990) Octanoylation of the Lipoyl Domains of the Pyruvate Dehydrogenase Complex in a Lipoyl-deficient Strain of *Escherichia coli. Mol. Microbiol.* **4**, 943-950

Arjunan P., Nemeria, N., Brunskill, A., Chandrasekhar, K., Sax, M., Yan, Y., Jordan, F., Guest, J. R., Furey, W. (2002) Structure of the Pyruvate Dehydrogenase Multienzyme Complex E1 Component from *Escherichia coli* at 1.85□ Resolution. *Biochemistry.* **41**, 5213-5221

Berg, A., Vervoort, J., De Kok, A. (1997) Three-dimensional Structure in Solution of the N-terminal Lipoyl Domain of the Pyruvate Dehydrogenase Complex from *Azotobacter vinelandii. Eur. J. Biochem.* **244**, 352-360

Bourguignon, J., Macherel, D., Neuburger, M., Douce, R. (1992) Isolation, Characterisation, and Sequence Analysis of a cDNA Clone Encoding L-protein, the Dihydrolipoamide Dehydrogenase Component of the Glycine Cleavage System from Pea leaf Mitochondria. *Eur. J. Biochem.*, **204**, 865-873

Bourguignon, J., Merand, V., Rawsthorne, S., Forest, E., Douce, R. (1996) Glycine Decarboxylase and Pyruvate Dehydrogenase Complexes Share the Same Dihydrolipoamide Dehydrogenase in Pea Leaf Mitochondria: Evidence from Mass Spectrometry and Primary Structure Analysis. *Biochem. J.*, **313**, 213-218

Bradford, M. M. (1976) A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding. *Anal. Biochem.* **72**, 248-254

Brookfield, D. E., Green, J., Ali, S. T., Machado, R. S., Guest, J. R. (1991) Evidence for Two Protein-Lipoylation Activities in *Escherichia coli*. *FEBS.* **295**,13-16

Burns, G., Brown, T., Hatter, K., Sokatch, J. R. (1989a) Sequence Analysis of the *lpdV* Gene for Lipoamide Dehydrogenase of Branched-Chain-Oxoacid Dehydrogenase of *Pseudomonas putida*. *Eur. J. Biochem.* **179**, 61-69

Burns, G., Sykes, P. J., Hatter, K., Sokatch, J. R. (1989b) Isolation of a Third Lipoamide Dehydrogenase from *Pseudomonas putida*. *J. Bacteriol.* **171**, 665-668

Camp, P. J., Randall, D. D. (1985) Purification and Characterization of the Pea Chloroplast Pyruvate Dehydrogenase Complex. *Plant Physiol.* **77**, 571-577

Camp, P. J., Miernyk, J. A., Randall, D. D. (1988) Some Kinetic and Regulatory Properties of the Pea Chloroplast Pyruvate Dehydrogenase Complex. *Biochim. Biophy. Acta.* **933**, 269-275

Carothers, D. J., Pons, G., Patel, M. S. (1989) Dihydrolipoamide Dehydrogenase: Functional Similarities and Divergent Evolution of the Pyridine Nucleotide-Disulfide Oxidoreductases. *Arch. Biochem. Biophys.* **268**, 409-425

Conner, M., Krell, T., Lindsay, J. G. (1996a) Identification and Purification of a Distinct Dihydrolipoamide Dehydrogenase from Pea Chloroplasts. *Planta.* **200**, 195-202

Conner, M., Lindsay, J. G. (1996b) Biochemical Characterisation of a Distinct Dihydrolipoamide Dehydrogenase from Pea Chloroplasts. *Flavins and Flavoproteins.* pp655-658., University of Calgary Press.

Conner, M. J. A. (1996c) Isolation and Characterisation of a Novel Dihydrolipoamide Dehydrogenase from Pea Chloroplast., PhD Thesis, University of Glasgow.

Cook, R. M., Munro, A., Lindsay, J. G. (1996) Distinct Forms of the FAD Containing Enzyme Dihydrolipoamide Dehydrogenase (E3) are Found in Tubers and Leaves of Potato (*Solanum tuberosum*). *Flavins and Flavoproteins.* pp659-661., University of Calgary Press.

Danson, M. J., Eisenthal, R., Hall, S., Kessell, S. R., Williams, D. L. (1984) Dihydrolipoamide Dehydrogenase from Halophilic archaebacteria. *Biochem. J.* **218**, 811-818

Danson, M. J., Conroy, K., McQuattie, A., Stevenson, K. J. (1987) Dihydrolipoamide Dehydrogenase from *Trypanosoma brucei*. Characterization and Cellular Location. *Biochem. J.* **243**, 661-665

Dardel, F., Davis, A. L., Laune, E. D., Perham, R. N. (1993) Three Dimensional Structure of the Lipoyl Domain from *Bacillus stearothermophilus* Pyruvate Dehydrogenase Multienzyme Complex. *J. Mol. Biol.* **229**, 1037-1048

Dardel, F., Packman, L. C., Perham, R. N. (1990) Expression in *Escherichia coli* of a Sub-gene encoding the Lipoyl Domain of the Pyruvate Dehydrogenase Complex of *Bacillus stearothermophilus*. *FEBS.* **264**, 206-210

Dyda, F., Furey, W., Swaminathan S., sax, M., Farrenkopf, B., Jordan, F. (1990) Preliminary Crystallographic Data for the Thiamin Diphosphate-Dependent Enzyme Pyruvate Decarboxylase from Brewer's Yeast. *J. Biol. Chem.* **265**, 17413-17415

Elias, B. A., Givan, C. V. (1979) Localisation of Pyruvate Dehydrogenase Complex in *Pisum sativum* Chloroplasts. *Plant Sci. Lett.* **17**, 115-122

Fang, R., Nixon, P. F., Duggleby, R. G. (1998) Identification of the Catalytic Glutamate in the E1 Component of Human Pyruvate Dehydrogenase. *FEBS. Letts.* **437**, 273-277

Faure, M., Bourguignon, J., Neuburger, M., Macherel, D., Sieker, L., Ober, R., Khan, R., Cohen-Addad, C., Douce, R., (2000) Interaction Between the Lipoamide-Containing H-protein and the Lipoamide Dehydrogenase (L-protein) of the Glycine Decarboxylase Multienzyme System. 2. Crystal Structures of H- and L-proteins. *Eur. J. Biochem.* **267**, 2890-2898

Frangioni, J. V. and Neel, B. G. (1993) Solubilization and Purification of Enzymatically Active Glutathione S-Transferase (pGEX) Fusion Proteins. *Anal. Biochem.* **210**, 179-187

Gemel, J., Randall, D. D. (1992) Light Regulation of Leaf Mitochondrial Pyruvate Dehydrogenase Complex. *Plant. Physiol.* **100**, 908-914

Golz, A., Focke, M., Lichyenthaler, H. K. (1994) Inhibitors of *de novo* Fatty Acid Biosynthesis in Higher Plants. *J. Plant. Physiol.* **143**, 426-433

Gorf, P. L., Winning, B. M., Scaybrook, T. P., Hill, S. A., Leaver, C. L. (1995) Mitochondrial Pyruvate Dehydrogenase: Molecular Cloning of the E1α Subunit and Expression Analysis. *Plant. Physiol.*, **108**, 1623-1629

Green, J. D. F., Laune, E. D., Perham, R. N., Ali, S. T., Guest, J. R. (1995) Three-dimensional Structure of a Lipoyl Domain from the Dihydrolipoamide Dehydrogenase Component of the Pyruvate Dehydrogenase Multienzyme Complex of *Escherichia coli*. *J. Mol. Biol.* **248**, 328-343

Guan, Y., Rawsthorne, S., Scofield, G., Shaw, P., Doonan, J. (1995) Cloning and Characterization of a Dihydrolipoamide Acetyltransferase (E2) Subunit of the Pyruvate Dehydrogenase Complex from *Arabidopsis thaliana*. *J. Biol. Chem.*, **270**, 5412-5417

Guest, J. R., Lewis, H. M., Graham, L. D., Packman, L. C., Perham, R. N. (1985) Genetic Reconstruction and Functional Analysis of the Repeating Lipoyl Domains in the Pyruvate Dehydrogenase Multienzyme Complex of *Escherichia coli*. *J. Mol. Biol.* **185**, 743-754

Guest, J. R. (1987) Functional Implications of Structural Homologies Between Chloramphenicol Acetyltransferase and Dihydrolipoamide Acetyltransferase. *FEMS Microbiol. Lett.*, **44**, 417-422

Hanemaaijer, R., Westphal, A. H., Van Der Heiden, T., De Kok, A., Veeger, C. (1989) The Quarternary Structure of the Dihydrolipoamide Transacetylase Component of the Pyruvate Dehydrogenase Complex from *Azobacter vinelandii*. *Eur. J. Biochem.* **179**, 287-292

Hartley, D. L. and Kane, J. F. (1988) Properties of Inclusion Bodies from Recombinant *Escherichia coli*. *Biochem. Soc.Trans.* **16**, 101-102

Heijne, G., Steppuhn, J., Herrmann, R. G. (1989) Domain Structure of Mitochondrial and Chloroplast Targeting Peptides. *Eur. J. Biochem.* **180**, 535-545

Hipps, D. S., Packman, L.C., Allen, M. D., Fuller, C., Sakaguchi, K., Appella, E. (1994) The Peripheral Subunit-Binding Domain of the Dihydrolipoamide Acetyltransferase Component of the Pyruvate Dehydrogenase Complex of *Bacillus stearothermophilus*: Preparation and Characterization of its Binding to the Dihydrolipoamide Dehydrogenase Component. *Biochem. J.* **297**, 137-143

Howard, M. J., Fuller, C., Broadhurst, R. W., Perham, R. N., Tang, J. G., Quinn, J., Diamond, A. G., Yeaman, S. J. Three-Dimensional Structure of the Major Autoantigen in Primary Biliary Cirrhosis (1998) *Gastroenterology*. **115**, 139-146

Hunter, A. (1985) Immunological and Biosynthetic Studies on the Mammalian 2-Oxoglutarate Dehydrogenase Complex. PhD Thesis. University of Glasgow.

Jackman, S. A., Hough, D. W., Danson, M. J., Stevenson, K. J., Opperdoes, F. R. (1990) Subcellular Localisation of Dihydrolipoamide Dehydrogenase and Detection of Lipoic Acid in Bloodstream forms of *Trypanosoma brucei*. *Eur. J. Biochem.* **193**, 91-95

Johnston, M. L., Luethy, M. H., Miernyk, J. A., Randall, D. D. (1997) Cloning and Molecular Analyses of the *Arabidopsis thaliana* Plastid Pyruvate Dehydrogenase Subunits. *Biochim. Biophys. Acta.*, **1321**, 200-206

Kalia, Y. N., Brocklehurst, S. M., Hipps, D. S., Appella, E., Sakaguchi, K., Perham, R. N. (1993) The High-Resolution Structure of the Peripheral Subunit-Binding Domain of Dihydrolipoamde Acetyltransferase from the Pyruvate Dehydrogenase Multienzyme Complex of *Bacillus stearothermophilus*. *J. Mol. Biol.* **230**, 323-341

Kelly, S. M. and Price, N. C. (1997) The Application of Circular Dichroism to Studies of Protein Folding and Unfolding. *Biochim. Biophys. Acta.* **1338**, 161-185

Kelly, S. M. and Price, N. C. (2000) The Use of Circular Dichroism in the Investigation of Protein Structure and Function. *Curr. Prot. Pep. Sci.* **1**, 349-384

Khailova, L. S., Bernkhardt, R., Khiubner, G. (1977) Study of the Kinetic Mechanism of the Pyruvate-2, 6-dichlorophenolinodphenol Reductase Activity of Muscle Pyruvate Dehydrogenase. *Biokhimiia.* **42**, 113-111 (Abstract only)

Khan, S. S. (1996) Biochemical and Molecular Characterisation of Dihydrolipoamide Dehydrogenase from Potato Mitochondria., PhD thesis, University of Glasgow

Kim H., Liu, T., Patel, M. S. (1991) Expression of cDNA Sequences Encoding Mature and Precursor Forms of Human Dihydrolipoamide Dehydrogenase in *Escherichia coli*. *J. Biol. Chem.* **266**, 9367-9373

Kleiger, G., Perry, J., Eisenberg, D. (2001) 3D Structure and Significance of the GIXXG Helix Packing Motif in Tetramers of the E1β Subunit of Pyruvate Dehydrogenase from the Archeon *Pyrobaculum aerophilum*. *Biochemistry.*, **40**, 14484-14492

de Kok, A., Hengeveld, A. F., Martin, A., Westphal, A. H. (1998) The Pyruvate Dehydrogenase Multi-enzyme Complex from Gram- negative Bacteria. *Biochim. Biopys. Acta.* **1385**, 353-366

Kresze, G. B., Ronft, H. (1981) Pyruvate Dehydrognase Complex from Bakers Yeast: Molecular Structure, Dissociation and Implications for the Origins of Mitochondria. *Eur. J. Biochem.* **119**, 581-587

Laemmli, U. K. (1970) Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature*, **227**, 680-685

Li de la Sierra, I., Pernot, L., Prange, T., Saludjian, P., Schiltz, M., Fourme, R., Padron, G. (1997) Molecular Structure of the Lipoamide Dehydrogenase Domain of a Surface Antigen from *Neisseria meningitidis. J. Mol. Biol.* **269**, 129-141

Lindsay, H., Beaumont, E., Richards, S. D., Kelly, S. M., Sanderson, S. J., Price, N. C., Lindsay, J. G. (2000) FAD Insertion is Essential for Attaining the Assembly-Competence of the Dihydrolipoamide Dehydrogenase (E3) Monomer from *Escherichia coli. J. Biol. Chem.* **275**, 36665-36670

Luethy, M. H., Miernyk, J. A., Randall, D. D. (1994) The Nucleotide and Deduced Amino Acid Sequences of a cDNA Encoding the E1β-Subunit of the *Arabidopsis thaliana* Mitochondrial Pyruvate Dehydrogenase Complex. *Biochim. Biophys. Acta.*, **1187**, 95-98

Luethy, M. H., Miernyk, J. A., Randall, D. D. (1995) The Mitochondrial Pyruvate Dehydrogenase Complex: Nucleotide and Deduced Amino Acid Sequences of a cDNA Encoding the *Arabidopsis thaliana* E1α-Subunit. *Gene*, **164**, 251-254

Luethy, M. H., Miernyk, J. A., David, N. R., Randall, D. D. (1996) Plant Pyruvate Dehydrogenase Complexes in Alpha-Keto Dehydrogenase Complexes. Patel, M. S., Roche, T. E., Harris, R. A. (Editors), Birkhauser Verlang Basel, Switzerland, pp71-92

Lutziger, I. Oliver, D. J. (2000) Molecular Evidence of a Unique Lipoamide Dehydrogenase in Plastids: Analysis of Plastidic Lipoamide Dehydrogenase from *Arabidopaia thaliana. FEBS Letters*, **484**, 12-16

Lutziger, I. Oliver, D. J. (2001) Characterization of Two cDNAs Encoding Mitochondrial Lipoamide Dehydrogenase from *Arabidopsis. Plant. Physiol.*, **127**, 615-623

Mande, S. S., Sarfaty, S., Allen, M. D., Perham, R. N., Hol, W. G. J. (1996) Protein-Protein Interactions in the Pyruvate Dehydrogenase Multienzyme Complex: Dihydrolipoamide Dehydrogenase Complexed with the Binding Domain of Dihydrolipoamide Acetyltransferase. *Structure*, **4**, 277-286

de Marcucci, G. G. L., Hodgson, J. A., Lindsay, J. G. (1986) The M_r-50 000 Polypeptide of Mammalian Pyruvate Dehydrogenase Complex Participates in the Acetylation Reactions. *Eur. J. Biochem*, **158**, 587-594

Mattevi, A., Schierbeek, A. J., Hol, W. G. J. (1991) Refined Crystal Structure of Lipoamide Dehydrogenase from *Azotobacter vinelandii* at 2.2A Resolution. *J. Mol. Biol.* **220**, 975-994

Mattevi, A., Obmolova, G., Schulze, E., Kalk, K. H., Westphal, A. H., de Kok, A. Hol, W. G. J. (1992a) Atomic Structure of the Cubic Core of the Pyruvate Dehydrogenase Multienzyme Complex. *Science*, **255**, 1544-1550

Mattevi, A., Obmolova, G., Sokatch, J. R., Betzel, C., Hol, W. G. J. (1992b) The Refined Crystal Structure of *Pseudomonas putida* Lipoamide Dehydrogenase Complexed With NAD^+ at 2.45□ Resolution. *Proteins*. **13**, 336-351

Mattevi, A., Obmolova, G., Kalk, K. H., Westphal, A. H., de Kok, A., Hol, W. G. J. (1993a) Refined Crystal Structure of the Catalytic Domain of Dihydrolipoamide Transacetylase (E2p) from *Azobacter vinelandii* at 2.6A Resolution. *J. Mol. Biol.* **230**, 1183-1199

Mattevi, A., Obmolova, G., Kalk, K. H., van Berkel, W. J. H., Hol, W. G. J. (1993b) Three-Dimensional Structure of Lipoamide Dehydrogenase from *Pseudomonas fluorescens* at 2.8□ Resolution: Analysis of Redox and Thermostability Properties. *J. Mol. Biol.* **230**, 1200-1215

Miernyk, J. A., Randall, D. D. (1987) Some Properties of Pea Mitochondrial Phospho-Pyruvate Dehydrogenase -Phosphatase. *Plant. Physiol.* **83**, 311-315

Miles, J. S., Guest, J. R., Radford, S. E., Perham, R. N. (1988) Investigation of the Mechanism of Active Site Coupling in the Pyruvate Dehydrogenase Multienzyme Complex of *Escherichia coli* by Protein Engineering. *J. Mol. Biol.* **202**, 97-106

Millar, A. H., Knopp, C., Leaver, C. J., Hill, S. A. (1998) Plant Mitochondrial Pyruvate Dehydrogenase Complex: Purification and Identification of Catalytic Components in Potato. *Biochem. J.* **334**, 571-576

Millar, A. H., Leaver, C. J., Hill, S. A. (1999) Characterization of the Dihydrolipoamide Acetyltransferase of the Mitochondrial Pyruvate Dehydrogenase Complex from Potato and Comparisons with Similar Enzymes in Diverse Plant Species. *Eur. J. Biochem.* **264**, 973-981

Miller, J. R., Busby, R. W., Jordan, S. W., Cheek, J., Henshaw, T. F., Ashley, G. W., Broderick, J. B., Cronan, J. E., Marletta, M. A. (2000) *Escherichia coli* LipA is a Lipoyl Synthase: *In Vitro* Biosynthesis of Lipoylated Pyruvate Dehydrogenase Complex from Octanoyl-Acyl Carrier Protein. *Biochemistry*, **39**, 15166-15178

Mooney, B. P., Miernyk, J. A., Randall, D. D. (1999) Cloning and Characterization of the Dihydrolipoamide S-Acetyltransferase Subunit of the Plastid Pyruvate Dehydrogenase Complex (E2) from *Arabidopsis*. *Plant Physiol.*, **120**, 443-451

Moreno, J. I., David, N. R., Miernyk, J. A., Randall, D. D. (2000a) Staphylococcal Protein A as a Fusion Partner Directs Secretion of the E1α and E1β Subunits of Pea Mitochondrial Pyruvate Dehydrogenase by *Bacillus subtilis*. *Protein Expression and Purification*, **18**, 242-248

Moreno, J. I., David, N. R., Miernyk, J. A., Randall, D. D. (2000b) *Pisum sativum* Mitochondrial Pyruvate Dehydrogenase can be Assembled as a Functional $\alpha_2\beta_2$ Heterotetramer in the Cytoplasm of *Pichia pastoris*. *Protein Expression and Purification*, **19**, 276-283

Neagle, J., De Marcucci, O., Dunbar, B., Lindsay, J. G. (1989) Component X of Mammalian Pyruvate Dehydrogenase Complex: Structural and Functional Relationship to the Lipoate Acetyltransferase (E2) Component. *FEBS Letts*. **253**, 11-15

Oliver, D. J. (1994) The Glycine Decarboxylase Complex From Plant Mitochondria. *Annu. Rev. Plant. Phsiol. Plant. Mol. Biol.* **45**, 323-337

Patel, M. S., Roche, T. E. (1990) Molecular Biology and Biochemistry of Pyruvate Dehydrogenase Complexes. *FASEB J.* **4**, 3224-3233

Perham, R. N., Packman, L. C. (1989) 2-Oxo Acid Dehydrogenase Multienzyme Complexes: Domains, Dynamics, and Design. *Ann. N. Y. Acad. Sci.* **573**, 1-20

Perham, R. N. (1991) Domains, Motifs and Linkers in 2-Oxo Acid Dehydrogenase Multienzyme Complexes: A Paradigm in the Design of a Multifunctional Protein. *Biochemistry* **30**, 8501-8512

Perham, R. N., Reche, P. A. (1998) Swinging Arms in Multifunctional Enzymes and the Specificity of Post-Translational Modification. *Biochem. Soc. Trans.* **26**, 299-303

Quinn, J., Diamond, A. G., Masters, A. K., Brookfield, D. E., Wallis, N. G. (1993) Expression and Lipoylation in *Escherichia coli* of the Inner Lipoyl Domain of the E2 Component of the Human Pyruvate Dehydrogenase Complex. *Biochem. J.* **289**, 81-85

Ralston, G. (1993) Introduction to Analytical Ultracentrifugation. Beckman Handbook, Beckman Instruments Inc., California.

Randall, D. D., Rubin, P. M., Fenko, M. (1977) Plant Pyruvate Dehydrogenase Complex: Purification, Characterisation and Regulation by Metabolites and Phosphorylation. *Biochim. Biophys. Acta.* **485**, 336-349

Randall, D. D., Williams, M. Rapp, B. J. (1981) Phosphorylation-Dephosphorylation of Pyruvate Dehydrogenase Complex from Pea Leaf Mitochondria. *Arch. Biochem. Biophys.* **207**, 437-444

Randall, D. D., Miernyk, J. A., Fang, T. K., Budde, R. J. A., Schuller, K. A. (1989) Regulation of the Pyruvate Dehydrogenase Complexes in Plants. *Ann. NY. Acad. Sci.* **573**, 192-205

Reed, L. J., Browning, K.S., Niu, X., Behal, R. H., Uhlinger, D. J. (1989) Biochemical and Molecular Genetic Aspects of Pyruvate Dehydrogenase Complex from *Saccharomyces cerevisiae. Ann. N. Y. Acad. Sci.* **573**, 155-167

Reed, L. J., Hackert, M. L. (1990) Structure-Function Relationships in Dihydrolipoamide Acyltransferases. *J. Biol. Chem.* **265**, 8971-8974

Richards, S. D. (1999) Protein-Protein Interactions within the 2-Oxoacid Dehydrogenase Complexes., PhD Thesis, University of Glasgow.

Robien, M. A., Clore, G. M., Omichinski, J. G., Perham, R. N., Appella, E., Sakaguchi, K., Gronenborn, A., M. (1992) Three-Dimensional Solution Structure of the E3-Binding Domain of the Dihydrolipoamdie Succinyltransferase Core from the 2-Oxoglutarate Dehydrogenase Multienzyme Complex of *Escherichia coli. Biochemistry.* **31**, 3463-3471

Roughan, P. G., Kagawa, T., Beevers, H. (1980) On the Light Dependency of Fatty Acid Synthesis by Isolated Spinach Chloroplasts. *Plant. Sci. Letts.* **18**, 221-228

Sambrook, J., Fritsch, E. F., Maniatis, T. (1989) Molcular Cloning. A Laboratory Manual. Second Edition, Editor, Chris Nolan. Cold Spring Harbour Laboratory Press.

Sanderson, S. J., Khan, S. S., McCartney, R. G., Miller, C., Lindsay, J. G. (1996) Reconstitution of Mammalian Pyruvate Dehydrogenase and 2-Oxoglutarate Dehydrogenase Complexes: Analysis of Protein X Involvement and Interaction of Homologous and Heterologous Dihydrolipoamide Dehydrogenases. *Biochem. J.* **319**, 109-116

Sarojini, G., Oliver, D. J. (1983) Extraction and Partial Characterization of the Glycine Decarboxylase Multienzyme Complex from Pea Leaf Mitochondria. *Plant. Physiol.* **72**, 194-199

Schuller, K. A., Randall, D. D. (1989) Regulation of Pea Mitochondrial Pyruvate Dehydrogenase Complex. *Plant. Physiol.* **89**, 1207-1212

Taylor, A.E., Cogdell, R. J., Lindsay, J. G. (1992a) Immunological Comparison of the Pyruvate Dehydrogenase Complexes from Pea Mitochondria and Chloroplasts. *Planta*, **188**, 225-231

Taylor A.E., Millar, R. E., Carmichael, A., Cogdell, R. J., Lindsay, J. G. (1992b) Dihydrolipoamide Dehydrogenase in plants: Differences in the Mitochondrial and Chloroplastic Forms. *Biochem. Soc. Trans.*, **21**, 38S

Thatcher, D. R., Wilks, P., Chaudhuri, J. (1996) Inclusion Bodies and Refolding. Proteins LabFAX, (Price, N. C. eds.) βios Scientific Publishers, Academic Press, pp119-130

Thekkumkara, T., Pons, G., Mitroo, S., Jentoft, J. E., Patel, M. S., (1989) Molecular Biology of the Human Pyruvate Dehydrogenase Complex: Structural Aspects of the E2 and E3 Components. *Ann. N. Y. Acad. Sci.* **573**, 113-129

Thelen, J. J., Luethy, M. H., Miernyk, J. A., Randall, D. D. (1995) Bacterial Expression and Mitochondrial Import of His-Tagged E1α and E1β Subunits of the *Arabidopsis thaliana* Pyruvate Dehydrogenase. *Plant. Physiol.*, **108**, 769

Thelen, J. J., Luethy, M. H., Miernyk, J. A., Randall, D. D. (1998) Partial Purification and Characterization of the Maize Mitochondrial Pyruvate Dehydrogenase Complex. *Plant. Physiol.*, **116**, 1443-1450

Thelen, J. J., Miernyk, J. A., Randall, D. D. (1999a) Molecular Cloning and Expression Analysis of the Mitochondrial Pyruvate Dehydrogenase from Maize. *Plant Physiol.* **119**, 635-643

Thelen, J. J., Muszynski, M. G., David, N. R., Luethy, M. H., Elthon, T. E., Miernk, J. A., Randall, D. D. (1999b) The Dihydrolipoamide S- Acetyltransferase Subunit of the Mitochondrial Pyruvate Dehydrogenase Complex from Maize Contains a Single Lipoyl Domain. *J. Biol. Chem.*, **274**, 21769-21775

Toyoda, T., Suzuki, K., Sekiguchi, T., Reed, L. J., Takenaka, A. (1998) Crystal Structure of Eucaryotic E3, Lipoamide Dehydrogenase from Yeast. *J. Biochem.* **123**, 668-674

Turner, S. R., Ireland, R., Rawsthorne, S. (1992) Purification and Primary Amino Acid Sequence of the L Subunit of Glycine Decarboxylase. Evidence for a Single Lipoamide Dehydrogenase in Plant Mitochondria. *J. Biol. Chem.*, **267**, 7745-7750

Wada, M., Yasumo, R., Jordan, S. W., Cronan, J. E., Wada, H. (2001a) Lipoic Acid Metabolism in *Arabidopsis thaliana*: Cloning and Characterization of a cDNA Encoding Lipoyltransferase. *Plant. Cell Physiol.* **42**, 650-656

Wada, M., Yasuno, R., Wada, H. (2001b) Identification of an *Arabidopsis* cDNA Encoding a Lipoyltransferase Located in Plastids. *FEBS.* **506**, 286-290

Walker J. L., Oliver, D. J. (1986) Glycine Decarboxylase Multienzyme Complex: Purification and Partial Characterization from Pea Leaf Mitochondria. *J. Biol. Chem.* **261**, 2214-2221

Wallis, N. G., Allen, M. D., Broadhurst, W., Lessard, I. A. D., Perham, R. N. (1996) Recognition of a Surface Loop of the Lipoyl Domain Underlies Substrate Channelling in the Pyruvate Dehydrogenase Multienzyme Complex. *J. Mol. Biol.* **263**, 463-474

Wallis, N. G. and Perham, R. N. (1994) Structural Dependence of Post-Translational Modification and Reductive Acetylation of the Lipoyl Domain of the Pyruvate Dehydrogenase Mulitenzyme Complex. *J. Mol. Biol.* **236**, 209-216

Weber, K., Osborn, M. (1969) The Reliability of Molecular Weight Determinations by Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis. *J. Biol. Chem.* **244**, 4406-412

West, S. M., Rice, J. E., Beaumont, E. S., Kelley, S. M., Price, N. C., Lindsay, J. G. (1995) Dissociation and Unfolding of the Pyruvate Dehydrogenase Complex by Guanidinium Chloride. *Biochem. J.* **308**, 1025-1029

White, M. F., Fothergill-Gilmore, L. A., Kelly, S. M., Price, N. C. (1993) Substitution of His-181 by Alanine in Yeast Phophoglycerate Mutase Leads to Cofactor-induced Dissociation of the Tetrameric Structure. *Biochem. J.* **291**, 479-483

Williams, M., Randall, D. D. (1979) Pyruvate Dehydrogenase Complex from Chloroplasts of *Pisum sativum* L. *Plant Phsyiol.* **64**, 1099-1103

Yang, D., Song, J., Wagenknecht, T., Roche, T. E. (1997) Assembly and Full Functionality of Recombinantly Expressed Dihydrolipoyl Acetyltransferase Component of the Human Pyruvate Dehydrogenase Complex. *J. Biol. Chem.* **272**, 6361-6369

Yeaman, S. J. (1989) The 2-Oxo Acid Dehydrogenase Complexes: Recent Advances. *Biochem. J.* **257**, 625-632