TARGETING, ASSEMBLY AND REGULATION OF THE PYRUVATE DEHYDROGENASE COMPLEX

Jason David Ward

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

Division of Biochemistry and Molecular Biology, IBLS

July 1998

Abstract

The majority of polypeptides in the mitochondrion are nuclear encoded and post-translationally transported to the mitochondrion from the cytosol. These include the polypeptides which make up the pyruvate dehydrogenase complex (PDC), the largest multienzyme complex of the 2-oxoacid dehydrogenase complex family. This transportation is initiated by an N-terminal mitochondrial targeting signal. However, whereas standard mitochondrial matrix targeting signals are approx. 15-30 amino acids long in mammals, the N-terminal presequence of the PDC-E2 component is 73 amino acids in length. Such long presequences usually carry additional information, such as a second targeting signal; however the presequence of PDC-E2 appears to have no such additional function.

Furthermore, mammalian PDC-E2 has the capacity to spontaneously self assemble into 60-meric dodecahedral cores, which form the superstructure for the complex itself. An integral part of mitochondrial targeting is prevention of such self assembly to allow translocation of the polypeptide from the cytosol to the matrix inside the mitochondrion.

This research details studies undertaken to examine the roles of the extended presequence of PDC-E2. A novel *in vivo* model was developed to monitor the natural behaviour of mitochondrial presequences with a reporter gene, utilising the reporter to measure both targeting and assembly. Sections of the PDC-E2 presequence were tested with this model system to discover their role within the translocation and folding of PDC-E2.

The regulation of PDC by PDC kinase was also investigated. PDC activity is strictly controlled by a variety of effector molecules, through the inactivation of PDC-E1 by PDC kinase. The kinase, by exerting control over the first, irreversible and rate limiting step, can block the entire reaction sequence.

Interestingly, phosphorylation of PDC-E1 can occur at 3 different sites, but previous research has demonstrated that not all sites must be phosphorylated to inactivate the complex.

To understand the behaviour of the kinase, and the circumstances under which the 3 phosphorylations take place, electrospray mass spectroscopy was employed to measure the number of phosphorylations on PDC-E1 made by PDC kinase, over time, under different circumstances.

Acknowledgements

For helping me at work, thanks go to

Professor Lindsay, for having the idea for this research, for helping me complete my Ph.D., and for personally supplying genomic DNA.

Past and present students and staff of the Davidson Building, including many from a time long ago in C35, namely Sanya Sanderson, Ruth Fullerton, Mark Conner, Saiqa Khan, Loïc Briand (for continuing the work), Graham McCartney, Susan Richardson, Clare Miller, Heather Lindsay, Lynn Anderson, and all in C30, including Margo Murphy; my thanks for some valuable help.

Thanks also to Tino Krell (for lots of help on the HPLC and ESMS, which all took a lot of time) and Wai-Yan Wan, for help on the Silicon Graphics Indy workstation.

For being my wife, love and thanks go to

Sheila Ward. Thanks for doing the dishes when I was playing computer games instead of writing my thesis. Thanks also to both our families for the support we've had, and that I've had during this research.

Many thanks go to

Jesus Christ, whom I confess to be my Lord and Saviour, and strength in times of trouble (of which there were plenty during this research and subsequent write up).

Finally, for funding this research, thanks go to

The Wellcome Trust. All Ph.Ds should be funded this well.

Table of Contents

page

List of Figures

page

List of Tables

Abbreviations

The following abbreviations used in this thesis are in addition to those recommended in "Instructions to authors" (*Biochem. J.* (1992) **281**, 1-19).

BCDC	branched chain amino acid 2-oxoacid dehydrogenase complex
CAT	chloramphenicol acetyltransferase
Ch.	chapter
CoA	coenzyme A
DMSO	dimethlysulphoxide
DTT	dithiothreitol
E3 (LDH)	dihydrolipoamide dehydrogenase
GDC	glycine dehydrogenase
GdnHCl	guanidine hydrochloride
Leupeptin	acetyl-L-leucyl-L-arginal
Lipoamide	6,8-thioctic acid amide
MOPS	3-[N-morpholino]propane-sulphonic acid
M_r	relative molecular mass
MSUD	maple syrup urine disease
NEM	N-ethylmaleimide
OGDC	2-oxoglutarate dehydrogenase complex
PBC	primary biliary cirrhosis
PDC	pyruvate dehydrogenase complex
PEG	polyethylene glycol
PMSF	phenylmethylsulphonylfluoride
TCA cycle	tricarboxylic acid cycle
TEMED	N,N,N',N'-tetramethylethylenediamine
ThDP	thiamine diphosphate
Tween 20	polyoxyethylenesorbitan monolaurate

TX-100	Triton X-100

Chapter 1

Introduction

1.1 Targeting, assembly and regulation of pyruvate dehydrogenase complex

Pyruvate dehydrogenase complex (PDC) is the largest multienzyme complex of the 2-oxoacid dehydrogenase complex family, located within the mitochondrion of eukaryotic organisms and regulated by a wide range of metabolites and hormones. The polypeptides which make up PDC are all nuclear encoded and post-translationally transported to the mitochondrion from the cytosol. The targeting of these polypeptides to the mitochondrion, the assembly of this large multimeric array from its component polypeptides and the regulation of the mature multienzyme complex are all investigated within this thesis. This chapter firstly describes the basic structural and functional characteristics of the 2-oxoacid dehydrogenase complexes before exploring the current understanding in targeting, assembly and regulation of PDC.

1.2 The 2-oxoacid dehydrogenase complexes

The 2-oxoacid dehydrogenase complexes are a family of 3 related high molecular mass polypeptide assemblies; the pyruvate dehydrogenase complex (PDC), the 2-oxoglutarate dehydrogenase complex (OGDC) and the branched-chain 2-oxoacid dehydrogenase complex (BCDC). They are all located within the mitochondrial matrix of eukaryotes, in association with the inner mitochondrial membrane. PDC can also be found in the plastids of plant cells.

The complexes have been intensively studied from a variety of organisms since they represent classic examples of multi-enzyme complexes; they have the capacity for spontaneous self assembly, the complexes are held together by non-covalent interactions between polypeptides and they catalyse a series of consecutive chemical reaction steps in metabolism.

Organising consecutive steps in a reaction pathway within multi-enzyme

complexes confers a number of advantages to the cell. Complexes can use 'substrate channelling', where metabolic intermediates are passed from one active site to another, enhancing catalytic efficiency by increasing the concentration of intermediates within a defined microenvironment. Intermediates can also be protected from a harsh external environment whilst in the complex; this is described as the 'hot potato' hypothesis, where a particularly labile intermediate can be quickly passed between successive enzyme sites (Perham, 1991).

1.2.1 Role in metabolism

The 2-oxoacid dehydrogenase complexes play key roles in intermediary metabolism, catalysing the irreversible oxidative decarboxylation of 2-oxoacids with the production of NADH, CO_2 and an acyl-CoA species (Fig. 1.1). PDC decarboxylates pyruvate to give acetyl-CoA, regulating the committed entry of 2 carbon units (acetate) from glycolysis into the TCA cycle. OGDC forms an important part of the TCA cycle, governing the flux of carbon units around the cycle, where it is involved in converting 2-oxoglutarate to succinyl-CoA. OGDC is also vital for providing succinyl-CoA for the synthesis of porphyrins, lysine and methionine. BCDC has a broad range of substrates, promoting the irreversible step in the degradation of branched-chain amino acids (leucine, isoleucine, valine), and as such regulates the breakdown of these compounds for entry of their carbon units into the TCA cycle. BCDC is also involved in the degradation of methionine and threonine, and is therefore integral to the control of branched-chain amino acid

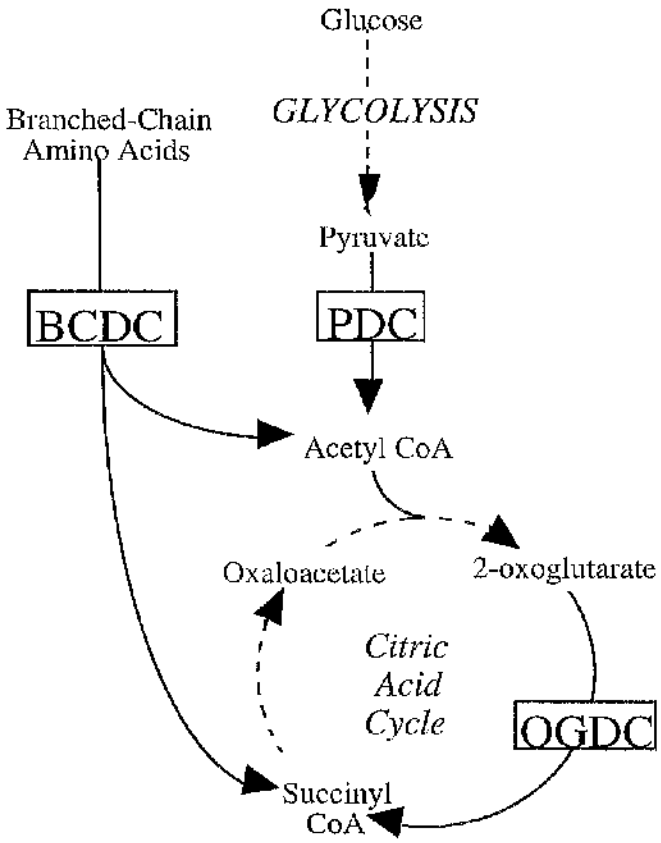


Figure 1.1 Diagram showing position of the 2-oxoacid dehydrogenase complexes in metabolism

As seen in the diagram, key positions in metabolism are occupied by these complexes, each of which catalyses a committed and irreversible step in carbohydrate or amino acid breakdown.

levels in mammals, converting excesses into acyl-CoA derivatives for recycling back into catabolism (see Perham, 1991, for a review).

All 3 complexes are regulated by the end-products of their reactions; PDC and BCDC are additionally regulated by complex-specific kinases and phosphatases. These kinases and phosphatases are controlled by various metabolites, such as end products, a variety of hormones and other effector molecules (see section 1.4).

1.2.2 Reaction sequence

The complexes are composed of multiple copies of 3 distinct enzymes: E1, a substrate-specific 2-oxoacid dehydrogenase, E2, a complex-specific dihydrolipoamide acyltransferase and E3, a dihydrolipoamide dehydrogenase.

Although the E1 and E2 components are complex-specific, the E3 component was previously thought to be common to all the complexes. However, evidence exists for E3 isoforms in certain cases. In rat liver mitochondria (Carothers *et al.*, 1987), bacteria (McCully *et al.*, 1986; Sokatch *et al.*, 1983), potato tuber mitochondria (R. Fullerton, Ph.D. thesis, Glasgow University, 1995) and pea leaf chloroplasts (Conner *et al.*, 1996), distinct E3 isoforms have been detected.

E1 catalyses the first, rate limiting step in the overall reaction sequence of each complex (Walsh *et al.*, 1986), the irreversible decarboxylation of its specific 2-oxoacid substrate. During the reaction, the cofactor of E1, thiamine diphosphate (ThDP), forms a 2-(1-hydroxyacyl)-ThDP intermediate, with the concomitant production of CO_2 (Fig. 1.2). E1 then catalyses the reductive acylation of the lipoyl prosthetic group of E2. This lipoamide moiety is found within a conserved motif located in the N-terminal region, covalently linked to a specific lysine via its ε-amino group. Acylation of the lipoamide moiety

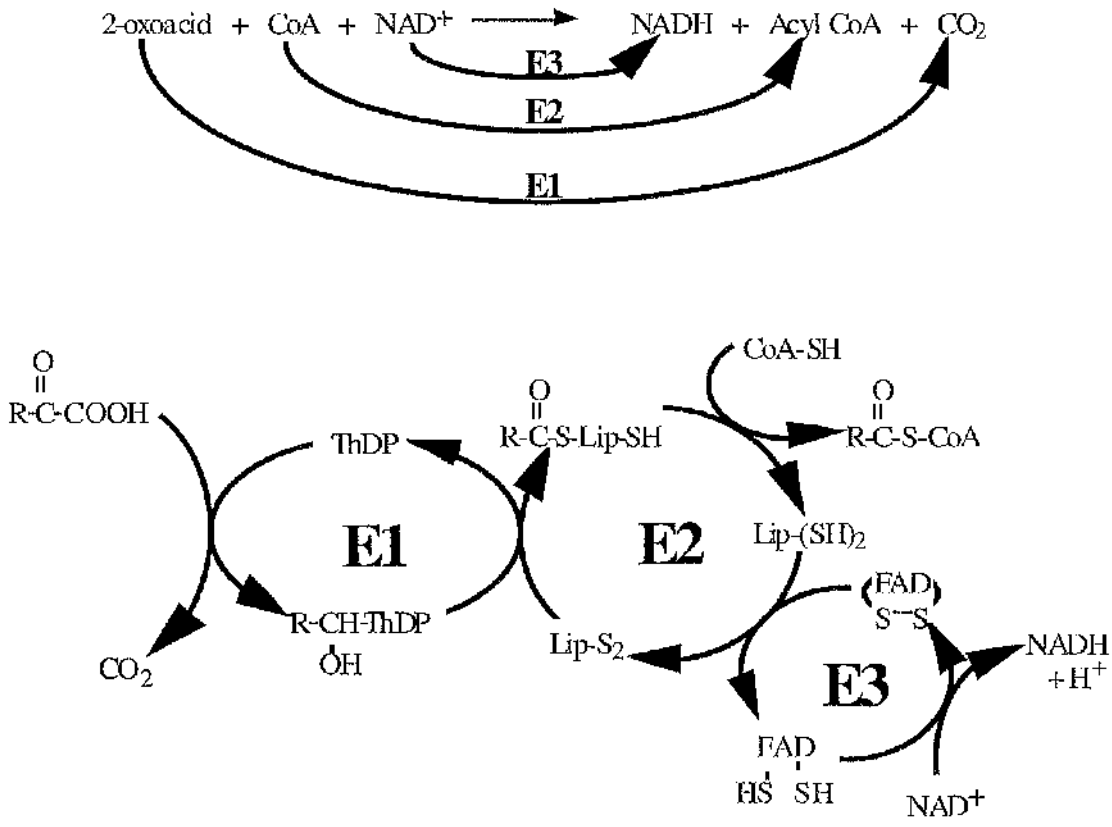


Figure 1.2 Reaction mechanism of the 2-oxoacid dehydrogenase complexes highlighting the specific reactions catalysed by the component enzymes

E1, 2-oxoacid dehydrogenase; E2, dihydrolipoamide acyltransferase; E3, dihydrolipoamide dehydrogenase; ThDP, thiamine diphosphate; Lip, lipoic acid; R=CH_3 for PDC; $COOHCH_2CH_2$ for OGDC; $(CH_3)_2CH$, $(CH_3)_2CHCH_2$ and $(CH_3)_2(C_2H_5)CH$ for BCDC.

forms an S^8-acyldihydrolipoamide intermediate; E2 then transfers the acyl group to its coenzyme A (CoA) acceptor, leaving the lipoamide prosthetic group in a reduced state. E3, using an FAD cofactor, reoxidises the reduced lipoamide prosthetic group. The FAD component is in turn reoxidised by NAD^+, which is the final electron acceptor. This leaves the complex in its original state, ready for the next reaction cycle. Each part of the complex will be discussed in more detail shortly.

1.2.3 Structure of the 2-oxoacid dehydrogenase complexes

Electron microscopy studies have shown the 2-oxoacid dehydrogenase complexes to be 30-50 nm in diameter, comparable in size to ribosomes (Henderson *et al.*, 1979). Negative stain electron microscopy of PDC (Oliver and Reed, 1982) and low resolution X-ray crystallography of E2/E3 subcomplexes (Fuller *et al.*, 1979) led to a model where E2 formed a core around which E1 and E3 bound on the vertices and faces, respectively. The core of the complex, an oligomer of E2 subunits, has been determined at a higher resolution by X-ray crystallography of recombinant E2 C-terminal catalytic domains (the N-terminal lipoyl domains, being highly flexible [section 1.3.2] appear to prevent such determinations). This, along with various reconstitution studies, have demonstrated that the E1 and E3 components are arranged around the surface of an E2 core. In OGDC and BCDC, this core forms a cube of 24 E2 subunits (octahedral 432 symmetry). PDC-E2 also exists as an octahedron in Gram negative bacteria, but in Gram positive bacteria, yeast, mammals, birds and fungi, PDC-E2 is present as a pentagonal dodecahedron of 60 subunits exhibiting icosahedral (532) symmetry (Fig. 1.3, Wagenknecht *et al.*, 1991).

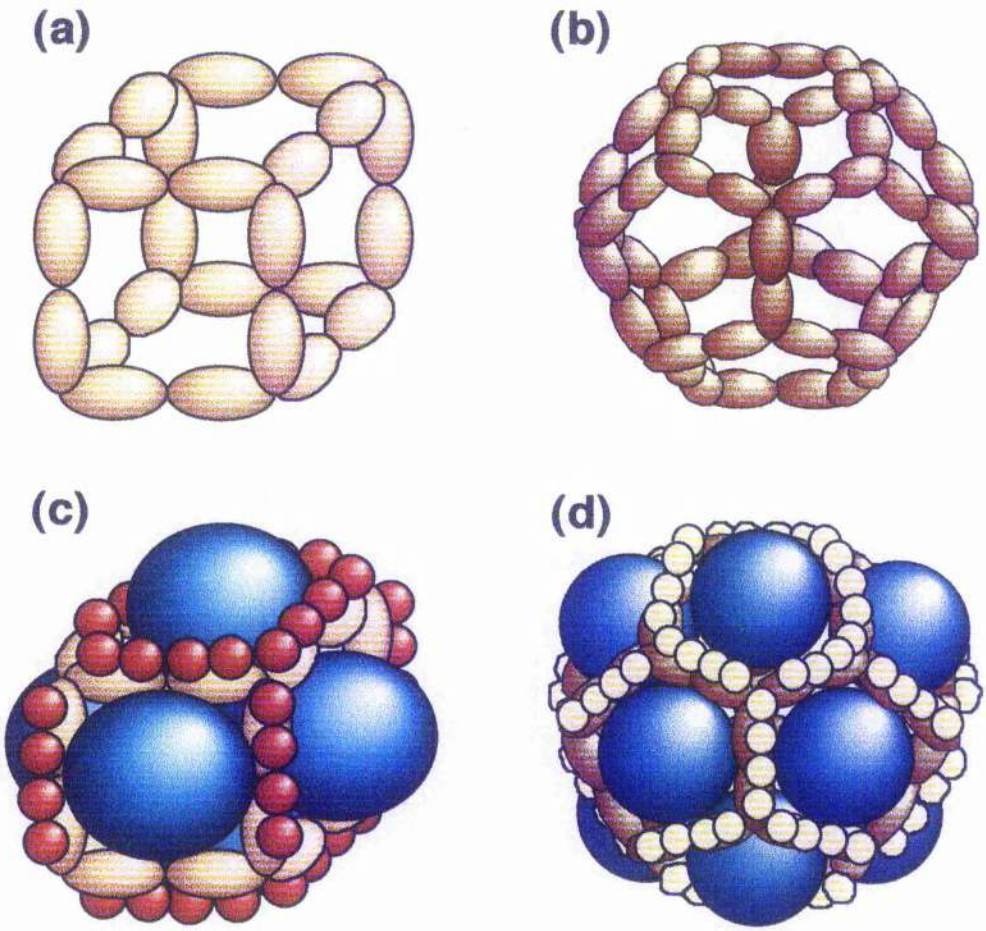


Figure 1.3 Proposed 3-D structure of PDC-E2 and native PDC

(a) Model of the cubic core of 24 E2 subunits arranged in groups of three as the eight vertices of a cube. E2 is represented by orange eliptoids.

(b) Proposed model of mammalian PDC core with 60 E2 subunits in groups of three forming the 20 vertices of an icosahedron. E2 is represented by orange elipses.

(c) Model of the cubic core form of PDC with 12 E1 dimers bound along the edges of the E2 core and 6 E3 dimers on the faces of the cube. The E3 dimers are represented by blue spheres, whilst the E1 dimers are represented as 4 smaller red spheres, to highlight the similarities to the model in (d).

(d) Model of the pentagonal dodecahedral PDC core with 20-30 E1 tetramers bound to the 30 edges and 12 E3 dimers bound to the 12 faces of the pentagonal dodecahedral core. The E3 dimers are represented by blue spheres, whilst the E1 tetramers are represented as 4 smaller yellow spheres

It has long been observed that PDC from mammalian sources, fungi, birds and *Bacillus* species have 60meric, icosahedral E2 cores. One reservation had previously been that these complexes could not be visualised using EM techniques such as negative staining, the complex dissociating during the various procedures of staining. Other EM methods have since been used to detect the complexes, and this 60meric structure has been confirmed in a number of more recent studies (Roche *et al.*, 1993; Wagenknecht *et al.*, 1991).

E1 is thought to bind the edges of such cores (20-30 in icosahedral PDC, 12 in cubic forms of the complexes, as seen in Table 1.1). E1 in OGDC is a homodimer, but in PDC and BCDC, it exists as a heterotetramer ($\alpha_2\beta_2$). E3 slots into the faces of cubic E2 (6 dimers, 1 dimer per face). In PDC it was previously thought that only 6 E3 dimers bound to the PDC dodecahedral E2 core (1 monomer per face), but recent evidence suggests that 12 dimers bind instead, with 1 dimer per face (Wu and Reed, 1984; Sanderson *et al.*, 1996). PDC from mammalian and yeast sources also possess protein X, a tightly associated subunit which exhibits strong similarities to E2. The polypeptide appears to have a function in binding E3 to the E2 core (Gopalkrishnan *et al.*, 1989). Recent evidence has increased the number of protein X subunits per complex from 6 to 12 monomers, possibly as 6 dimers (Maeng *et al.*, 1994; Sanderson *et al.*, 1996), which corresponds well with the stoichiometry expected for E3.

PDC and BCDC also posses specific kinase and phosphatase enzymes. In PDC, there are 1-3 tightly associated kinase and 5 loosely associated phosphatase molecules (see Table 1.1). These will be discussed in detail in a later section.

The subunit composition and mass of PDC is summarised in Table 1.1 below.

Table 1.1

	Number of Subunits			Mass of Subunit (Da)		
Subunit	OGDC	BCDC	PDC	OGDC	BCDC	PDC
E1(α)	12	12	60	96,000	46,000	42,000
E1(β)	-	12	60	-	37,000	36,000
E2	24	24	60	48,000	50,000	70,000
E3	6	6	12/24	55,000	55,000	55,000
Protein X	-	-	12	-	-	51,000
Kinase 1	-	N/D	1-3	-	43,280*	49,244*
Kinase 2	-	-	1-3	-	-	46,181*
Kinase 3	-	-	1-3	-	-	46,938*
Kinase 4	-	-	1-3	-	-	46,466*
Phosphatase a	-	N/D	N/D	-	38,000	50,000
Phosphatase b	-	-	N/D	-	-	97,000

(Values taken from bovine hearts or pig kidney, Lindsay, 1989, *except: Kinase 1-3 masses from human cDNA, Popov *et al.*, 1995; Kinase 4 mass from cDNA clone, Rowles *et al.*, 1996; BCDC Kinase from rat cDNA clone, Popov *et al.*, 1992. N/D- Not determined.)

1.2.4 Diseases associated with the 2-oxoacid dehydrogenase complexes

A number of genetic defects, causing a variety of metabolic acidoses where substrates accumulate to toxic levels, have been characterised which involve deficiencies in the activities of the 2-oxoacid dehydrogenase complexes. The majority of these diseases are in the E1 components of the complexes. One such example is 'Maple Syrup Urine Disease', an autosomal

recessive inborn condition where BCDC has deficient activity, arising from various genetic errors in the E1 and E2 subunits. The loss of activity can cause severe mental retardation, and leads to accumulation of branched-chain amino acids and derived metabolites in the urine of sufferers.

Another interesting case was a genetic error present in the protein X subunit of two brothers (Marsac *et al.*, 1993). These patients rapidly developed neurologic disorders in conjunction with chronic lactic acidaemia. PDC activity was reduced to 12% and 22%, demonstrating the necessity for this component (see McConnell *et al.*, 1997, for a recent study).

Of importance to the research described in this thesis is another disease associated with the 2-oxoacid dehydrogenase complexes. Primary biliary cirrhosis (PBC), a chronic autoimmune disease found predominantly in middle-aged women (average age at diagnosis is 51), is Britain's most common non-alcohol related cause of liver cirrhosis (O'Donohue and Williams, 1996). The disease is characterised by inflammation of septal and interlobular bile ducts arising from the immune reaction against bile duct epithelia. This progresses to cirrhosis of the liver, and often liver failure. Patients with PBC usually have antibodies to mitochondrial antigens present within their sera; the major autoantigens associated with this disease include the E2 components of PDC, OGDC and BCDC, component X of PDC and the PDC E1α and E1β subunits (Yeaman *et al.*, 1988; Fussey *et al.*, 1988; Surh *et al.*, 1989a, 1989b; Fregeau *et al.*, 1989). Approximately 95% of patients with PBC have antibodies directed at the inner lipoyl domain of PDC-E2: the most immunoreactive part of the lipoyl domain is the site of attachment of the lipoic acid. Previous research demonstrated that the lipoic acid itself plays a major part in the immunoreactivity of PDC-E2 (Fussey *et al.*, 1990). Recent evidence, however, suggests that it may be the lipoic acid of protein X which is implicated in the disease (Joplin *et al.*, 1997).

The major autoantigen involved in this condition is found on the surface

of the plasma membrane in the bile duct epithelial cells; additionally, antibodies are unable to enter the cells to react to mitochondrial antigens (Joplin *et al.*, 1995). It follows that a mistargeting or processing defect causing relocation of PDC-E2 or protein X epitopes to the cell surface could lead to an autoimmune reaction, and may be involved in promoting the onset of this disease.

1.3 Composition of the 2-oxoacid dehydrogenase complexes

1.3.1 2-oxoacid dehydrogenase

The 2-oxoacid dehydrogenases are ThDP-requiring enzymes, and catalyse both the initial irreversible decarboxylation of the 2-oxoacid and the reductive transfer of the resultant acyl group to the lipoyl prosthetic group on the lipoyl domains of the appropriate E2. As the rate limiting stage within each complex, it is these enzyme components which are regulated by phosphorylation in PDC and BCDC (Linn *et al.*, 1972).

E1s from many sources have been characterised both at the protein and cDNA level, and 3 families have been described based on the gene sequences. These groups were E1s from OGDC, E1s from PDC found in Gram negative bacteria, and E1s from all other PDC sources and BCDC. Between groups of E1s there is minimal homology, suggesting early evolutionary divergence (Darlinson *et al.*, 1984). In OGDC and PDC from Gram negative bacteria, E1 is a homodimer, with each subunit approximately 100,000 Da. In eukaryotic PDC and BCDC, E1 exists as a hetero-tetrameric enzyme ($\alpha_2\beta_2$), with α subunits of 41-46,000 Da and β subunits of 35-38,000 Da.

The cofactor, thiamine diphosphate (ThDP) is bound by the two α subunits. ThDP is bound by 2 domains: one which binds the pyrimidine ring, and one which binds the diphosphate. Although each α subunit in the tetramer has both domains to bind ThDP, the cofactor is located in the cleft between the

α and β subunits, so that one subunit binds the diphosphate whilst the other holds the pyrimidine ring (Stepp *et al.*, 1985).

Within the α subunit is also a β subunit binding domain, whilst the β subunit is responsible for binding the tetramer to the E2 core. It has been proposed that E1β is involved in the reductive acylation of the lipoyl group of E2, but this has yet to be proven (Yeaman, 1989; Patel and Roche, 1990).

1.3.2 Dihydrolipoamide acyltransferase

The E2 component forms the structural and functional framework of all the complexes. Not only do the other components bind to this oligomer in a non-covalent fashion to form the complex, but E2 also links the components functionally, coupling the activities of the various enzymes by providing flexible lipoyl domains which are able to visit the various active sites. E2, an acyltransferase, also catalyses the second step in the reaction sequence, by transferring the acyl group from its lipoyl domain(s) to CoA.

1.3.2.1 The domains of the dihydrolipoamide acyltransferases

Proteolytic cleavage of E2 has revealed a highly segmented structure consisting of a number of functional domains, between which are proteolytically sensitive linker regions (Fig. 1.4). All E2s have 3 main regions; the N-terminal lipoyl domain or domains, the E1 and/or E3 binding domain, and the C-terminal core forming domain, which also contains the catalytic part of the polypeptide. The E2s from all the complexes also share structural similarities with the 'H' protein of the glycine decarboxylase complex (GDC) (Fujiwara *et al.*, 1986; Fujiwara *et al.*, 1991).

At the N-terminal region the E2 polypeptide has 1-3 highly homologous lipoyl domains (Fig. 1.4). Each domain is approx. 80 amino acids long, joined

by linker regions of 25-30 amino acids which are rich in alanine, proline and various charged residues to allow conformational flexibility (Dardel *et al.*, 1991). These provide the flexibility required for interaction of the lipoyl domain(s) with the active sites of E1 and E3. Each of these domains is post-translationally modified by a ligase which covalently attaches the lipoic acid to a specific lysine on the domain via an amide linkage (Brookfield *et al.*, 1991). No conformational change is caused by attachment of the lipoic acid, which supports the concept that the domain is a free-swinging arm in solution (Perham, 1991). The flexibility of these domains is essential to the proper functioning of E2, providing a means of linking the various enzymatic activities.

The peripheral subunit binding domain is C-terminal to the lipoyl domain. This domain, 30-50 amino acids long, has the function of binding E1 and/or E3 to the core. In BCDC and *Bacillus* species, E1 and E3 appear to compete for binding to this domain which is unable to bind both simultaneously. In OGDC and octahedral PDC (e.g. in *E. coli*), the domain is largely associated with E3 binding, however the octahedral core of PDC in *A. vinelandi* binds E1, as does PDC-E2 in mammalian cells. E3 in mammalian cells is instead bound by protein X in PDC, and by E1 in OGDC, where the binding domain appears to be located in the gene for OGDC-E1 (Rice *et al.*, 1992).

The C-terminal domain of E2 acts as both the catalytic and the core-forming domain, with a compact structure. This domain also possesses E2 binding sites, allowing it to interact with other E2 polypeptides to form the high M_r core assembly. It is believed that the PDC-E2 core is generated via a trimeric intermediate (McCartney *et al.*, 1997), which then polymerises to make either the cubic or icosahedral structures.

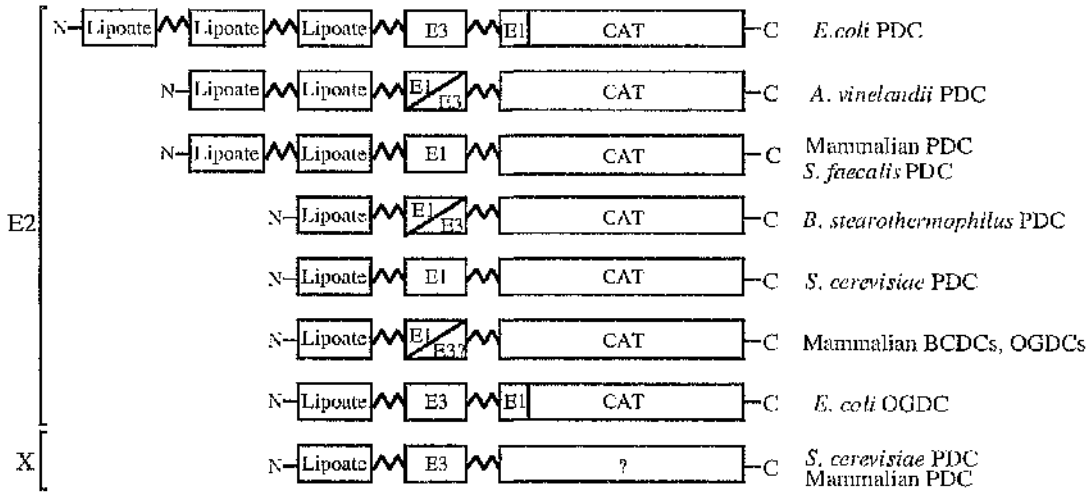


Figure 1.4 Diagram of the domain structure of the E2 and protein X components

Legend:

E3/E1- E3 and/or E1 binding domain

CAT- catalytic core domain

Zigzags- flexible linker regions

The domain structure of protein X is shown for comparison. Note that mammalian OGDC has been shown to lack the E3 binding domain. E3 binding sequences are found on E1 and the E1 binding sites on E2 are unknown as yet (Rice *et al.*, 1992).

1.3.2.2 Structure of dihydrolipoamide acetyltransferase

The entire structure of E2 from PDC has yet to be determined at the molecular level since the N-terminal domains of E2 are highly flexible and prevent crystal formation for X-ray crystallography. Whilst NMR is able to resolve the structure of the lipoyl domain, it cannot determine the structure of the whole molecule due to the size of PDC-E2. However, the structure of individual domains and the structure of the assembled core is known in detail, at least from prokaryotic sources.

In terms of the entire E2 polypeptide it is the C-terminal domain, forming the inner core of the complex, which has the most conserved amino acids between species. The active site histidine is especially conserved, surrounded by the motif HXXXDG, a region which is shared with the active site of chloramphenicol acetyltransferase (CAT). Indeed, throughout the C-terminal domain of E2 there is remote but significant sequence similarity between the two polypeptides. X-ray crystallography of this domain of PDC-E2 from *Azotobacter vinelandii* also reveals a structural similarity to CAT (Fig. 1.5). These comparisons may be a reflection of the similarities in catalytic mechanisms between E2 and CAT; E2 transfers acetyl groups from its lipoyl domain to CoA, whereas CAT transfers acetyl groups from acetyl-CoA to chloramphenicol to inactivate it. There is considerable variety between species in the active site residues of E2; however, a mechanism based on that demonstrated for CAT has been made and substantiated for a number of PDC-E2s (reviewed in Russell and Guest, 1991). In brief, protons are abstracted from the thiol of CoA, which then undergoes nucleophilic attack on the 8-acyl lipoamide thioester (Fig. 1.6).

The structure of the E3 binding domain has been solved by NMR studies on *E. coli* OGDC (Robein *et al.*, 1992) and PDC from *B. stearothermophilus* (Kalia *et al.*, 1993). The structure of the lipoyl domains has also been

determined by NMR studies on the purified segments from *B. stearothermophilus* (Dardel *et al.*, 1993) and *E. coli* (Green *et al.*, 1995). NMR has revealed that the lipoyl domain folds into a flattened barrel formed from two β-sheets, with the lipoic acid group attached to the tip of a β-turn.

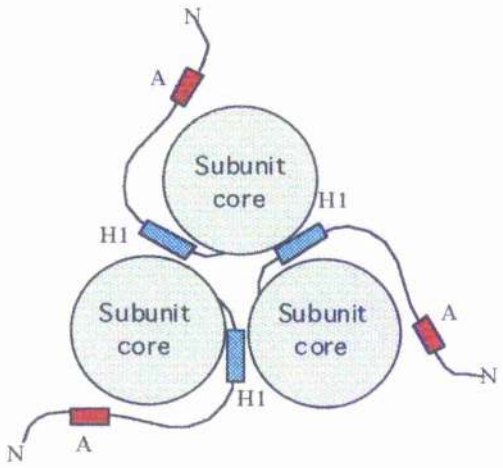
1.3.3 Dihydrolipoamide dehydrogenase

Dihydrolipoamide dehydrogenase is a member of the flavoprotein disulphide oxidoreductases (Carothers *et al.*, 1989), found in the 2-oxoacid dehydrogenase complexes, the distantly related GDC (where it is called L-protein) and the bacterial acetoin dehydrogenase system. In all these complexes, it is responsible for the NAD^+ dependent reoxidation of the dihydrolipoamide prosthetic groups of the E2 components in 2-oxoacid dehydrogenase complexes, and hydrogen carrier protein (H-protein) in GDC. This allows the acyltransferase component to re-enter the catalytic cycle (Kikuchi and Hiraga, 1982; Carothers *et al.*, 1989).

A single E3 was previously thought to be common to all three 2-oxoacid dehydrogenase complexes. However, evidence for complex specific isoforms of E3 is accumulating for a number of species, namely *Pseudomonas putida* (Palmer *et al.*, 1991a, b), *Pseudomonas aeruginosa* (McCully *et al.*, 1986), *E. coli* (Richarme, 1989) and rat liver (Carothers *et al.*, 1987).

E3 exists as a dimer of identical subunits of approx. 50,000 Da containing a molecule of non-covalently bound FAD. The sequence of E3 is highly conserved between species (e.g. porcine E3 has 96% identity with human E3, and 44% identity with *E. coli* E3). Structural similarity, as shown by X-ray crystallography, has been demonstrated between E3 and other pyridine nucleotide-disulphide oxidoreductases, such as glutathione reductase (Karplus and Schulz, 1987), mercuric reductase (Schiering *et al.*, 1991), NADH peroxidase (Stehle *et al.*, 1991), trypanothione reductase (Kuriyan *et*

Chloramphenicol Acetyltransferase- CAT



Dihydrolipoamide Acetyltransferase- PDC E2

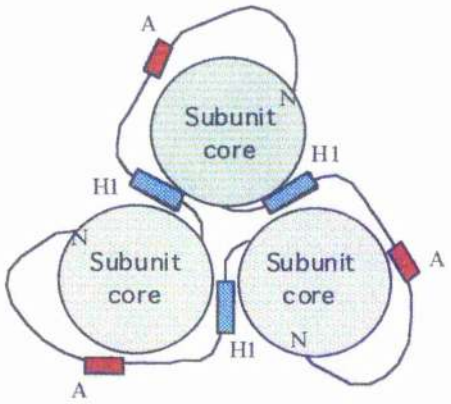


Figure 1.5 Comparison of CAT and PDC-E2 (Core domain)

This diagram from Mattevi *et al.* (1992) shows in two dimensions similarities between the domain structures of CAT and dihydrolipoamide dehydrogenase (PDC-E2), including their abilities to form trimers through similar interactions between the subunits. The shaded boxes represent important regions of

Figure 1.5 cont'd

secondary structure common to both polypeptides. The N-terminal region is the 'A' strand, which is involved in formation of an important β-sheet, the other is the H1 α-helix, which is involved in the 3-fold symmetry.

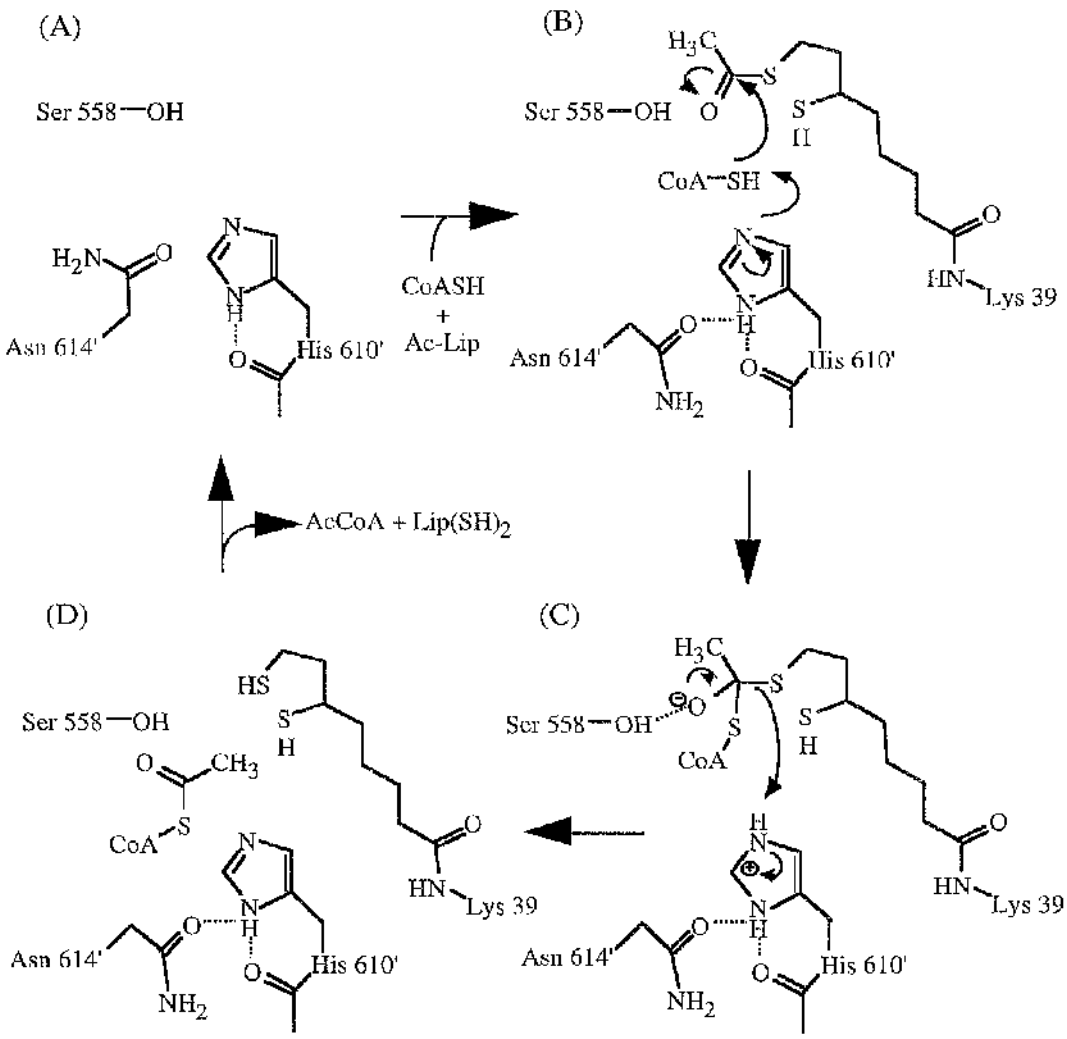


Figure 1.6 Catalytic mechanism of the acetyltransferase reaction

Two subunits of E2 participate in the reaction; residues provided by one subunit of E2 are represented by a prime ('), the other E2 residues are unlabelled. (A) In the ground state, the side chain of Asn 614' forms a hydrogen bond with Asp 508 (not shown), blocking the CoA-binding channel at the pantethiene moiety of CoA. (B) During binding of reduced CoA (CoASH), a reorientation of the Asn 614' side chain allows the formation of a new hydrogen bond with the reactive sulphydryl group of CoA. The sulphur atom subsequently attacks the carbonyl in the acetyl

Figure 1.6 cont'd

group of acetyl-dihydrolipoamide. (C) This intermediate is stabilised by a hydrogen bond to Ser 558. Rearrangement of the acetyl carbonyl group results in the simultaneous release of dihydrolipoamide and acetyl-CoA. (D) The reaction is completed when acetyl CoA and dihydrolipoamide are released from the active site, and Asn 614' has rearranged to go back to the original state (reproduced from Hendle *et al., 1995*).

al., 1991a) and thioredoxin reductase (Kuriyan *et al.*, 1991b).

1.3.4 Protein X

Protein X was discovered as a polypeptide tightly associated with the complex in yeast and mammalian PDC (De Marcucci and Lindsay, 1985). Protein X has an M_r of 50,000, and was originally thought to be a proteolytic fragment of E2. Protein X has since been cloned (Harris *et al.*, 1997); it possesses a lipoyl domain and a structure similar to E2 (Fig. 1.4). Protein X, or E3 binding protein as it has been recently called (E3BP), binds strongly to the complex through an unknown interaction with the E2 core. It was originally suggested that the 12 subunits integrated into the E2 core; however, protein X appears to have a post E2-assembly interaction in mammals, and is not a requirement for assembly of the core in yeast. Additionally, the presence of 60 PDC-E2 monomers per core would prevent any integration of protein X (Sanderson *et al.*, 1996; McCartney *et al.*, 1997).

Although no catalytic activity has been ascribed to protein X, a number of functions have already become apparent. The lipoyl domains of protein X are able to participate in the overall reaction since they are capable of being reductively acetylated, they also form a network of lipoyl domains over the PDC-E2 core which can sustain approx 15% of overall PDC activity after loss of all E2 lipoyl domains (Sanderson *et al.*, 1996). It has also been established that protein X is involved in positioning and binding E3 around the core to facilitate its role in electron transfer. Recent work has demonstrated that protein X binds the E3 domain to the PDC-E2 core; PDC without protein X has only 10-20% residual activity, which can be enhanced by the presence of excess exogenous E3 (McCartney *et al.*, 1997).

1.4 Regulation of complexes

The complexes are strictly controlled by short and long term regulation. In both bacteria and mammals the complexes are regulated in the short term by end-product inhibition (i.e. acyl CoA and NADH) when product concentrations increase in relation to the substrates. Such inhibition is competitively reversed by the substrates CoA and NAD^+ respectively.

Both short and long term regulation is achieved by phosphorylation from kinases and dephosphorylation by phosphatases in PDC from mammals, birds and *Neurospora crassa*, as well as mammalian BCDC. The kinases and phosphatases are modulated by a variety of hormones and other effector molecules. It is interesting to note that the complexes which have such regulation also have a $\alpha_2\beta_2$ tetrameric form of E1, whereas other complexes have only an α_2 form.

1.4.1 Kinases of PDC and BCDC

Inactivation of E1 is caused by phosphorylation of 1 of 3 possible sites in PDC E1, or 1 of 2 in BCDC E1. It is known from empirical studies that 1 phosphorylation per tetramer is sufficient for complete inhibition of an E1 (Korotchkina *et al.*, 1996). Both the kinase and phosphatase require Mg^{2+}. The kinase is inhibited by CoA, NAD^+, pyruvate, ThDP and is activated by acetyl CoA and NADH. The phosphatase is activated by Ca^{2+} and inhibited by NADH. Ca^{2+} binds at 2 sites, one site involved in binding to the E2 core, and the other site involved in catalysis (Reed and Yeaman, 1987). The kinase and phosphatase, by exerting control over the E1 component (which catalyses the first, irreversible, rate limiting step) can block the entire reaction sequence.

In rat heart PDC two kinases, p45 and p48, have been identified by N-terminal protein sequencing and subsequent cDNA cloning, whilst four

distinct isoforms have recently been cloned from the human genome (Rowles *et al.*, 1996). These all have sequences related to the prokaryotic protein histidine kinase family, quite dissimilar to regular cytoplasmic serine/threonine kinases (Popov *et al.*, 1993). The functions of each kinase would appear to be important to the metabolism of different tissues, since the isoenzymes are found in varying quantities in each tissue and are subject to differential allosteric regulation by metabolic effectors.

Liu *et al.* (1995b) proposed that the kinase dimer is able to move over the complex by swinging 'hand over hand', alternating between states where the kinase is bound specifically to 1 and 2 of the inner E2 lipoyl domains. This concept arose after demonstrating that the kinase repeatedly releases and binds these domains. An active kinase molecule would then be able to move over the surface of the E2 core, deactivating all the E1 components.

1.4.2 Phosphatases of PDC and BCDC

The phosphatases of PDC and BCDC are more loosely associated with their respective complexes, and are often lost during the purification of PDC and BCDC. In PDC the phosphatase is composed of two subunits, a catalytic subunit with an M_r of 50,000 and a subunit with an M_r of 97,000 which possesses an FAD cofactor. The function of this larger subunit is unknown to date (Teague *et al.*, 1982; Pratt *et al.*, 1982). PDC phosphatase requires Mg^{2+} for activity and Ca^{2+} promotes binding of the phosphatase to the E2 core of PDC, stimulating activity. Ca^{2+} is thought to elicit this behaviour by lowering the K_m for Mg^{2+}. This can also be achieved by spermine; by contrast, NADH inhibits the enzyme.

The corresponding phosphatase from BCDC is not as well characterised. The native enzyme has an M_r of 460,000 but its subunit composition and general properties remain unclear; only one catalytic subunit has been

identified with an M_r of 33,000 (Daumuni *et al.*, 1984; Daumuni and Reed, 1987).

1.5 Mitochondrial targeting

1.5.1 The requirement for protein import

The mitochondrion is an organelle containing several hundred polypeptides at a high concentration. Since only 13 of the polypeptides required by human mitochondria are encoded 'on site' by the genome present within the organelle, the rest must be imported from their site of synthesis in the cytoplasm (reviewed in Glover and Lindsay, 1992). All the other mitochondrial polypeptides, including those of the 2-oxoacid dehydrogenase complexes, are encoded by nuclear genes. These genes are transcribed into mRNA, which are then translated on free cytosolic ribosomes. Subsequently, the polypeptide which is synthesised is also cytosolic, but is required inside the mitochondrion. Such polypeptides must therefore be imported from the cytosol into the mitochondrion. Although co-translational transport of proteins into the mitochondrion has been observed (see Lithgow *et al.*, 1997, for a review), the majority of translocation occurs post-translationally; this is achieved through a variety of routes. One post-translation route appears to be used by more mitochondrial proteins than any other, and it is this route by which the polypeptides of the 2-oxoacid dehydrogenase complexes probably enter the mitochondrial matrix.

1.5.2 The method of protein import

Opinions vary on the details of how translocation of newly synthesised polypeptides into the mitochondrion is achieved, but most researchers agree that the following events take place: there is first an antifolding-unfolding step

where the emerging polypeptide or newly emerged polypeptide is kept from folding, i.e. is maintained in a translocation competent state. Evidence exists to suggest this is initiated by recruitment of chaperones as a direct result of the polypeptide possessing a specific N-terminal presequence. There is then a targeting event, which is facilitated by one or more presequence binding proteins, aiding the interaction of the polypeptide with the surface of the mitochondrion. This in turn leads to localisation of the polypeptide to the site of translocation. Following this is the translocation event itself, where the polypeptide is transported across the outer membrane, and then most frequently across the inner membrane, at a point where the 2 membranes of the mitochondrion are in close juxtaposition, called the 'translocation contact site'. Once across, the polypeptide is usually processed to remove the presequence, and is folded into its mature form, or retargeted to a sub-mitochondrial location.

In the mitochondrion, there are 4 sub-locations; the outer membrane, the inner membrane, the intermembrane space and the matrix. Polypeptides bound for the mitochondrion carry the 'default' matrix targeting signal, but often have additional information to divert the polypeptide to its final destination, sometimes, as in the case of outer membrane proteins, before reaching the matrix.

To complicate the picture further, there are a number of exceptions to the various steps; some polypeptides are able to insert themselves directly into the mitochondrion in a passive manner, some enter the intermembrane space yet do not pass through the inner membrane, and yet others possess a non-cleavable presequence. Such exceptions will be mentioned occasionally, but since they form the minority of cases, the focus will be on the route adopted by the majority of polypeptides, the 'default pathway' for mitochondrial targeting.

1.5.2.1 Antifolding

Early research in yeast noted that mitochondrial import *in vitro* was enhanced by the presence of some cytosolic fractions. Two proteins were purified subsequently which enhanced activity, one which was termed presequence binding factor (PBF), and one which was later shown to be a chaperone, Hsp70 (Fig. 1.7). Both proteins are critical in yeast for uptake of mitochondrial precursors.

Chaperones such as the cytosolic Hsp70 family appear to be involved at several steps in the import process in keeping the polypeptide unfolded. Chaperones also have a pivotal role in the unidirectional translocation of the polypeptide through the membranes of the mitochondrion, and are then responsible for folding the polypeptide correctly in the protein-dense mitochondrial matrix (Stuart *et al.*, 1994).

The polypeptide must usually be unfolded before entry into the mitochondrion, as shown by precursors which could not be imported when bound to antibody fragments, or presequence-dihydrofolate reductase (DHFR) constructs which would not import in the presence of its tightly-bound inhibitor, methotrexate (reviewed in Stuart *et al.*, 1986). In support of this view are the observations that preproteins are held in an open conformation before import, since precursors are often sensitive to digestion. Furthermore, import of precursors is accelerated by their unfolding; in experiments where precursors could not be imported into mitochondria due to mutations affecting the antifolding activity, polypeptides could still be imported if they were unfolded in 8 M urea beforehand (reviewed in Ryan and Jensen, 1995).

It has been proposed that cytosolic Hsp70 is one of the chaperones helping to keep the precursor in an unfolded state and prevent non-specific aggregation of these 'loosely folded' intermediates. Yeast mutant studies have demonstrated that when cytosolic Hsp70 is depleted, import of polypeptides

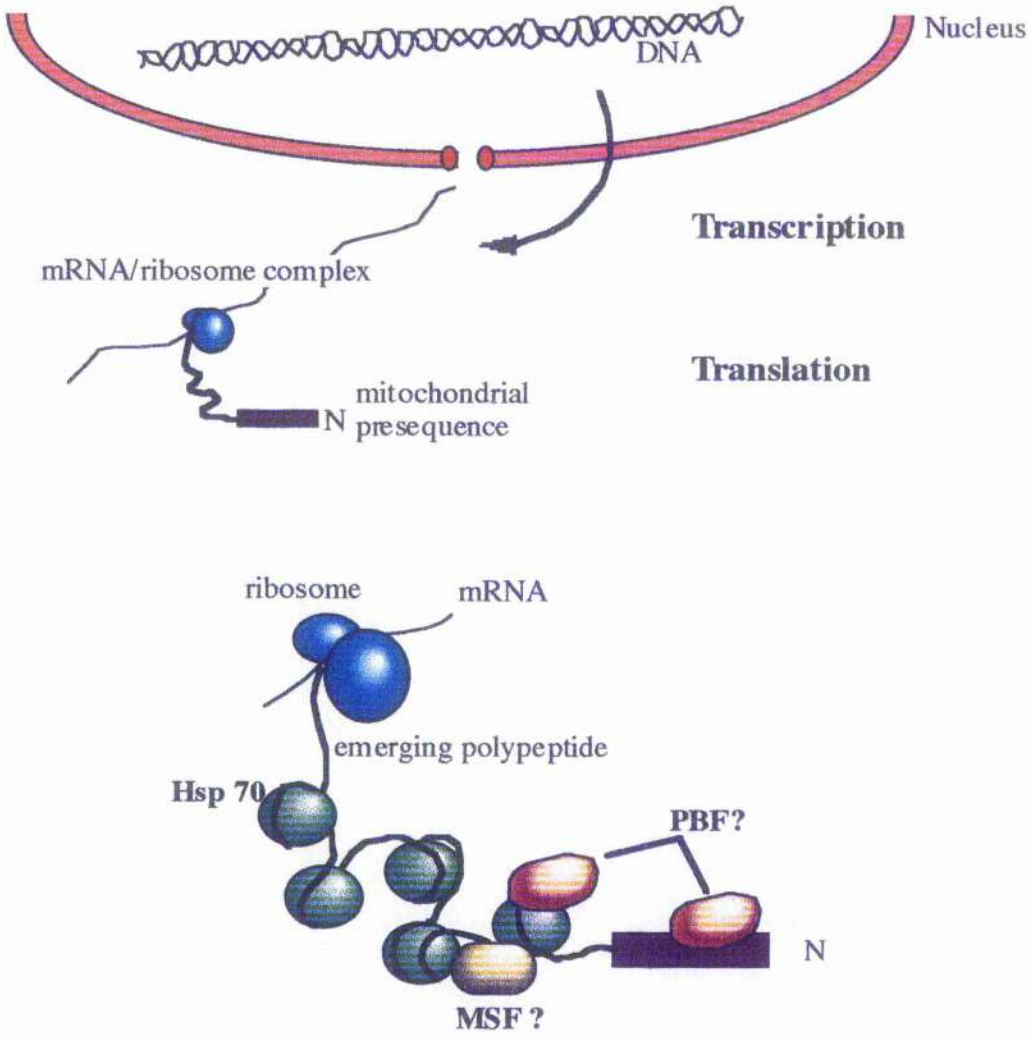


Figure 1.7. Illustration of the steps in antifolding

After translation of the mRNA to the polypeptide, the factors in antifolding bind the precursor as a result of a recognition event. Cytosolic Hsp70 is known to be involved in keeping many such polypeptides in a translocation-competent state. PBF (presequence binding factor) may bind along with Hsp70, or just bind the presequence. Mitochondrial import stimulating factor (MSF) may bind precursors without Hsp70 present.

into the mitochondrion is inhibited (reviewed in Glick and Schatz, 1991).

Since active unfolding by an 'unfoldase' has not been demonstrated thus far, it is believed that the chaperones are recruited to the emerging peptide as it exits the ribosome, preventing rather than reversing folding. The experiments of Endo *et al.* (1996) are especially relevant in this context, demonstrating that the chaperone Hsp70 in yeast is bound by mitochondrial presequences depending on the degree of amphiphilicity. With amphiphilicity being a defining characteristic of mitochondrial presequences, it follows that an important function of these sequences is to recruit chaperones to the emerging polypeptide.

PBF was characterised as a 50 kDa protein which could bind the immature form (pOTH) but not the mature form (OTH) of rat ornithine carbamoyltransferase. In experiments by Murakami and Mori (1990), the role of PBF was elucidated. When PBF was used in a mitochondrial import experiment with pOTH, it was found that it could stimulate import, and furthermore, could be further enhanced by the addition of Hsp70. The authors of this work also noted that Hsp70 alone could not stimulate uptake of pOTH.

Mitochondrial import stimulating factor (MSF) is another such antifolder/import enhancer. MSF is a chaperone, binding precursors in a partially folded conformation (especially aggregated mitochondrial precursors) and hydrolysing ATP. It differs from other chaperones in that it recognises mitochondrial signals, can direct precursors to mitochondrial receptors and has shown the ability to actively disperse protein aggregates (Ryan and Jensen, 1995). The mechanism by which MSF aids the fidelity of mitochondrial targeting is only partially understood; it is believed to work independently from Hsp70, possibly recognising presequences in conjunction with PBF.

1.5.2.2 Targeting

The unfolded polypeptide is then taken from the cytosol into the mitochondrion. This requires that the unfolded precursor be recognised by the mitochondrion, and bound by a receptor on the surface of the organelle, before being guided toward the translocation contact site (Fig. 1.8). Several receptors were first discovered during protease and antibody experiments, where the action of various mitochondrial membrane proteins was affected, causing different import behaviours in a number of precursors. On this basis, mitochondria were predicted to have a range of precursor receptors, and a number of putative receptors have since been identified, although not all have been characterised. These receptors have been found using a variety of methods, from simply using antibodies to mitochondrial membranes to the use of an anti-idiotypic antibody against a mitochondrial presequence peptide. Inhibition of a receptor can prevent uptake of several precursors into isolated mitochondria, but can be non-lethal; entire deletion of a receptor can also be non-lethal. Many of these receptors are therefore multiply redundant, with some precursors showing specificity for only one receptor, and others demonstrating a less particular nature, able to use one of a number of receptors (Glick and Schatz, 1991). A more complete analysis of the receptors MOM 20, 22, 37 and 70 can be found in Kubrich *et al.* (1995). The putative receptors MOM 19 ad MOM 40 have less defined roles, and are discussed in Seki *et al* (1995) and Rapaport *et al* (1997), respectively.

Such receptors are usually membrane bound with the receptor segment on the cytosolic side, recognising either the signal sequence (e.g. MOM 20 or MOM 22) or a cytosolic factor bound to the presequence, such as PBF. Differences between the receptors are manifold. For example, one receptor in *S. cerevisiae* binds the amphiphilic presequence directly without the requirement for any accessory proteins; by contrast, others, such as MOM 70

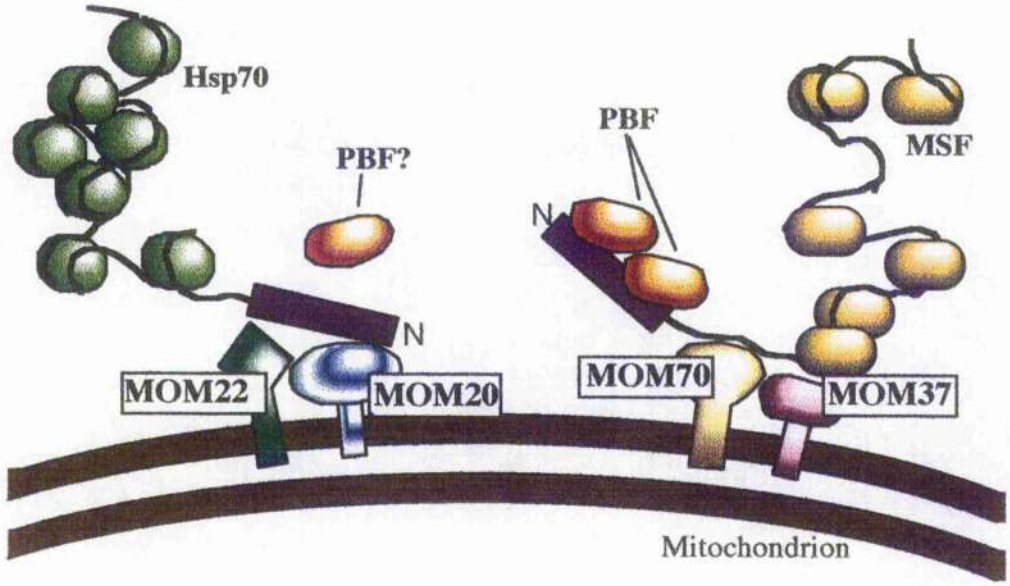


Figure 1.8 Schematic diagram of targeting by mitochondrial receptors

Mitochondrion outer membrane proteins (MOMs) and their possible mechanisms for targeting are shown. Receptors such as MOM 20 and 22 are expected to bind the presequences of precursor polypeptides directly. In the diagram, the peptide is held in an open conformation by Hsp70, and aided in targeting by PBF; this may not always be the case.

Other receptors, such as MOM 37 and 70 bind the mature part of the precursor, with bound chaperone, such as MSF.

and MOM 37, are proposed to bind only mitochondrial precursors which are held in an open conformation by chaperones such as MSF (Haucke *et al.*, 1995). It is also thought that receptors may work co-operatively, some binding the presequence, others binding mature regions of the same polypeptide; this may have important consequences for the differential kinetics of protein import (Shore *et al.*, 1995).

During recognition of the peptide and subsequent translocation, chaperones such as Hsp70 are removed or lost from the unfolded precursor. This may well involve another chaperone, known in *S. cerevisiae* as Ydjp1, a member of the DnaJ family and sometimes known in eukaryotes as Hsp40 (Fig. 1.9). This is required for efficient import although its precise role is not fully defined; it is postulated that it accelerates the ATPase activity of Hsp70 and subsequent dissociation of the chaperone from the precursor (reviewed in Martinus *et al.*, 1995). In yeast, Hsp40 is thought to bind the newly synthesised polypeptide along with Hsp70 to form a ternary complex; this then stimulates the ATPase activity of Hsp70 (Hendrick and Hartl, 1995). Thus, for efficient translocation, it is obvious that the chaperones, receptors and translocation apparatus must be able to function in a co-ordinated fashion.

Undoubtedly one function of these receptor proteins is to anchor the presequence to the membrane of the mitochondrion, limiting diffusion towards the translocation contact site to two-dimensions. A further implied role of the receptor is in aiding and co-ordinating insertion of the peptide into the mitochondrion and corresponding release of the peptide by the various chaperones.

It is of some interest that targeting may not be solely achieved through signals from the presequence. In yeast mutants incapable of targeting a non-mitochondrial protein with a mitochondrial presequence, one mutant was found which could distinguish between genuine and fusion mitochondrial polypeptides, transporting normal mitochondrial proteins to their correct

location, and leaving the 'synthetic' mitochondrial polypeptides in the cytosol (reviewed in Schatz and Dobberstein, 1996). It has been suggested therefore that further targeting information, at least in yeast, may be held within the mature part of the protein. Further investigation has also revealed that a large proportion of the mRNA coding for mitochondrial proteins is localised to the mitochondrial surface, and may be involved in co-translational translocation (Lithgow *et al.,* 1997). A number of mRNA binding proteins have been found which aid import of mitochondrial precursors in yeast, and microscopy studies have demonstrated a bias for mRNA encoding mitochondrial precursors to be in the vicinity of mitochondria, possibly mediated by an RNA signal in the 3' untranslated region. It is concluded therefore that both co-translational and post translational import of precursors can take place *in vivo.*

1.5.2.3 Translocation

Translocation is achieved by the combined efforts of a large number of proteins, some which form a pore in the membrane through which the polypeptide can travel, and some of which aid unidirectional movement of the amino acid chain through this proteinaceous pore (Fig. 1.9).

Despite the tremendous amount of research that has been invested into understanding the translocation systems of mitochondria, chloroplasts, endoplasmic reticula, peroxisomes and bacteria, the precise mechanisms remain the subject of considerable debate. Partly due to the difficulty in obtaining sufficient structural data on membrane proteins, a number of questions remain unanswered with regards to the translocation contact site, and indeed, the translocation apparatus of all membranes. A simple pore is not sufficient, since the translocation contact site is able to accommodate all amino acids, from glycine to tryptophan, without letting through any ions. Whilst the pore may not allow protons through from the intermembrane space

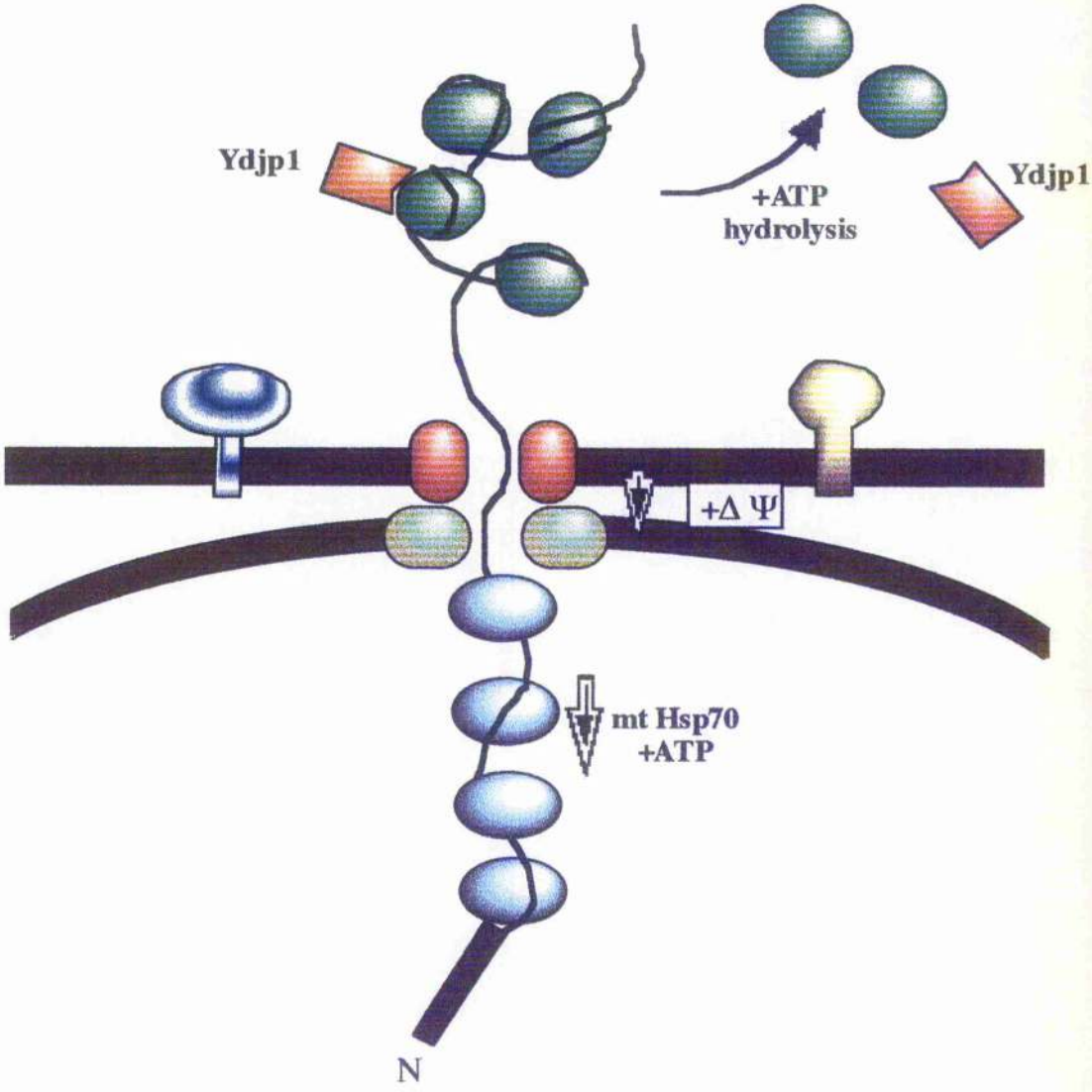


Figure 1.9 Diagram demonstrating the various contributing factors required for translocation into the mitochondrion

As shown, a number of factors are known to be required for efficient import in the majority of cases. Initial insertion of the presequence into the outer membrane translocation complex may be co-ordinated by one of the receptors, but the precise mechanism for insertion is poorly understood. It is thought, however, that insertion

Figure 1.9 cont'd

of the presequence into the inner mitochondrial membrane translocation pore is through an electrostatic interaction with the membrane potential ($\Delta\Psi$).

For release of the cytosolic chaperones during subsequent translocation, ATP is required, probably aided or coordinated by Ydjp1 (Hsp40). This allows movement of further sections of the polypeptide to be taken through the mitochondrial membranes. It is likely that the polypeptide is then pulled through the pore as proposed in the Brownian ratchet model, requiring ATP for the mitochondrial Hsp70.

to the matrix, it must, however, permit the transfer of bulky charged amino acids. Of interest is a protein called BSC1, a mitochondrial inner membrane protein which possesses an internal, unprocessed targeting signal. Since this polypeptide's signal is deep within the mature part of BSC1, it is expected to form a tight hairpin with a neighbouring transmembrane domain, and would be translocated through the pore as a loop. Import of such a loop would be analogous to the import of membrane spanning polypeptides into the endoplasmic reticulum (Fölsch *et al.*, 1996)

Many proteins of the translocation contact site have been identified by biochemical and genetic methods in yeast. Without structural data, there is only anecdotal evidence for the function of individual polypeptides within the complexes of the inner and outer translocation machinery (Kiebler *et al.*, 1993).

After being bound by a receptor, the polypeptide is then inserted into the outer membrane pore, termed the 'general insertion pore' since it is noted that a number of polypeptides are able to use the same site. Precursors for the matrix, inner membrane and some for the intermembrane space then cross the inner membrane through a second, separate pore. To do this, the mitochondrion must have a membrane potential ($\Delta\psi$) and a supply of ATP.

The proton gradient is required only for initial insertion of the presequence into the inner membrane translocation apparatus. It has been proposed that the positive charge on the presequence interacts with the gradient to allow it to overcome any kinetic barrier in inserting into the pore. Once inserted into the pore, the potential difference is no longer required, only the presence of ATP is necessary to complete translocation, as demonstrated by uncoupling and ATP depletion experiments (reviewed in Kiebler *et al.*, 1993).

Matrix ATP has been shown to be required for the reversible binding of mitochondrial Hsp70 to the incoming polypeptide. Studies by Stuart *et al.*

(1994) used ATP-depleted mitochondria to reveal the ability of mt Hsp70 to 'unfold' a polypeptide located on the cytosolic side of the mitochondrion and initiate its translocation from the mitochondrial matrix. This is in support of the 'Brownian ratchet' theory, where chaperones on the matrix side of the translocation contact site bind the polypeptide as it enters this compartment. As the polypeptide comes though by Brownian motion- either as a result of the folded polypeptide spontaneously unfolding a short section, or as the unfolded polypeptide is released from chaperones- the mt Hsp70 binds the precursor and makes any such translocation one way.

1.5.2.4 Processing and folding

Subsequent to this translocation, the polypeptide usually has the presequence removed and is then folded into the mature form (Fig. 1.10). The presequence is removed by a mitochondrial processing peptidase (called MPP), an unusual peptidase requiring zinc and cleaving peptides based not on primary sequence, but on a 3-dimensional motif. This peptidase has been classified with the pitrilysin family on the basis of the motifs it uses to bind its zinc cofactor. MPP is a heterodimer; the α subunit recognises the cleavage site according to an uncharacterised motif which may not be wholly contained within the presequence. Once recognised, the α subunit presents the polypeptide to the β subunit, which then cleaves the presequence from the rest of the polypeptide. Recognition of the peptide is tolerant of a large number of mutations (Luciano and Geli, 1996). As ever, there are exceptions to the rule. There are a few polypeptides in which the presequence is not processed (e.g. the N-terminal sequence of rhodanese) or is held within the mature part of the protein, such as BCS1, or many of the mitochondrial 'carrier proteins' which appear to translocate through the outer membrane by the default pathway, but then integrate into the inner membrane by a separate mechanism (Sirrenberg *et*

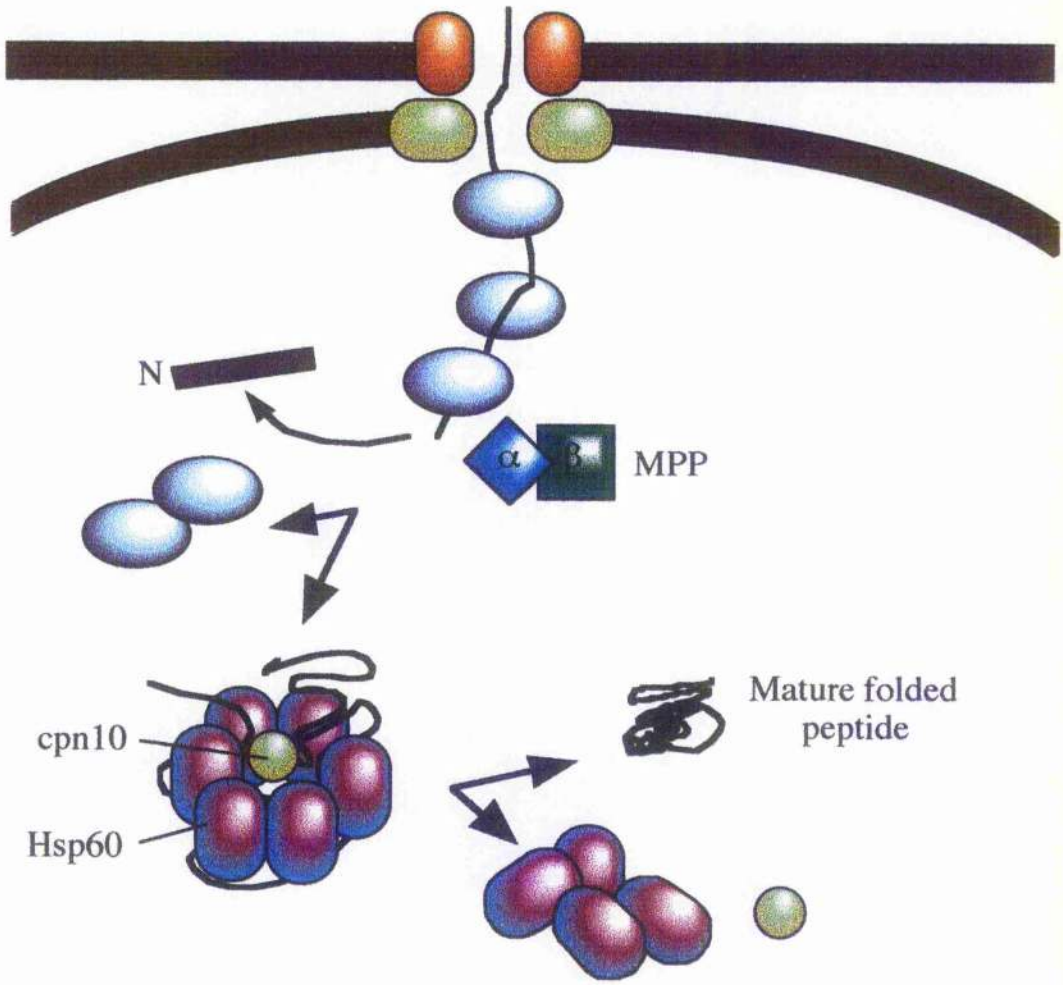


Figure 1.10 Diagram of the processing and folding of precursors in the mitochondrion

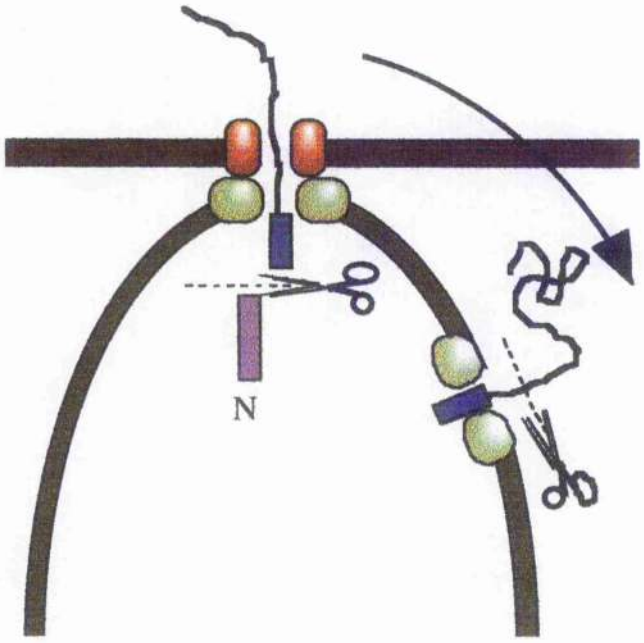
During translocation the presequence is cleaved by MPP. In the case of matrix bound polypeptides, mt Hsp70 is removed with hydrolysis of ATP, possibly during translocation. The mature-sized, partially folded peptide is then bound by Hsp60, along with Cpn10 (Hsp10). After release of these chaperones (with ATP hydrolysis), the peptide is fully folded.

al., 1998).

After processing, the polypeptide may be either re-targeted or folded. A matrix protein starts being folded shortly after entering the matrix; it is currently held that folding takes place after processing. Previously, however, opinion has been divided over when processing took place, it is possible that the two events may take place simultaneously. Those proteins which are bound for other compartments may be re-targeted. It is often the case that a polypeptide bound for either the intermembrane space or the inner membrane will posses a 'bipartite signal', that is, the presequence will be made up from one N-terminal signal that targets the polypeptide to the matrix, then a second signal is revealed when the first is cleaved off. This second N-terminal signal will redirect the polypeptide to its final destination; this re-addressing is called 'conservative sorting'. The second part of the signal shares similarities to bacterial secretory presequences, which supports the endosymbiotic theory of mitochondrial origins.

The re-addressing of polypeptides has been another area of some controversy, with the focus of debate being on the 'stop-transfer' theory or the contrasting 'conservative sorting' hypothesis. Both concepts are illustrated in Fig. 1.11. The question has often been one of whether a re-targeted protein uses matrix proteins for translocation or not. Experiments by Gruhler *et al.* (1996) provide some insight; in this research cytochrome b_2 presequences attached to DHFR were imported into mitochondria, but could not be fully imported due to the presence of methotrexate. Methotrexate binds DHFR, and can hold it in a closed, translocation-incompetent state, and if added before translocation, can halt import of the DHFR portion of the peptide. Such arrested polypeptides could be isolated which had already been re-addressed by their second targeting signal whilst still bound to mtHsp70, having lost their original mitochondrial targeting sequence. This demonstrates the polypeptide's contact with the matrix before re-targeting. Polypeptides

a)



b)

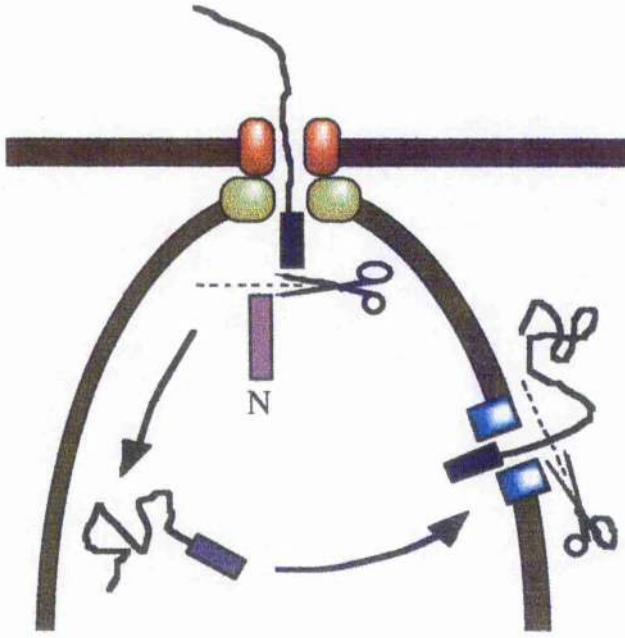


Figure 1.11 Stop-transfer and conservative sorting models

Bipartite signal sequences are thought to be sorted using either a 'stop-transfer' (a) or a 'conservative sorting' (b) method. In stop-transfer, the second part of the signal prevents further translocation; the inner membrane translocation machinery then dissociates from the outer membrane, allowing the polypeptide to enter the intermembrane space without passing through the matrix. The signal is then cleaved off by an inner membrane space processing peptidase.

In conservative sorting, the second part of the presequence acts as an export signal would in a bacterial cell, re-addressing the polypeptide to the inner mitochondrial membrane. The polypeptide is therefore fed through a new translocation pore as a hairpin loop, leaving the targeting signal in the matrix.

targeted to the intermembrane space are subsequently folded by chaperones present within this compartment.

Matrix polypeptides are folded in the presence of an unusually high concentration of proteins. It is not surprising that chaperones are present to prevent non-productive interactions and misfolding. Mitochondrial Hsp70 was identified from ATP-depletion studies, and as discussed, appears to have a role in translocation. It is likely that Hsp70 keeps incoming polypeptides in a loosely folded state before being folded by other chaperones. Another such chaperone is Hsp60; mitochondrial Hsp60 was identified from screens of temperature-sensitive yeast mutants able to import human ornithine transcarbamoylase into the mitochondrial matrix, but unable to assemble them at the non-permissive temperature (reviewed in Glick and Schatz, 1991). Other chaperones have also been implicated in the folding of mitochondrial proteins; these include mitchondrial GrpE and Cpn10 (also called Hsp10). These 2 chaperones have been discovered to be essential to the cell for survival and are not only implicated in protein import but also in regular protein maintenance; refolding damaged polypeptides (Martinus *et al.*, 1995).

1.5.3 Mitochondrial targeting sequences

Mitochondrially targeted polypeptides most usually posses a cleavable presequence of between 19 and 40 amino acids, although shorter and longer sequences have been found. These sequences exhibit no sequence homology, but have a high content of serine, threonine and basic amino acids. When the sequence is displayed as an α-helix, it will frequently reveal an amphiphilic nature, hydrophobic on one side, basic and hydrophilic on the other. Many studies have been carried out to detect patterns in all known presequences; no observed motif appears to be obligatory though some motifs are prominent, such as the amphiphilic helix, a lack of acidic residues and absence of proline

(Heijne *et al.*, 1989). Of all randomly generated synthetic presequences, 25% are able to target polypeptides to either the translocation apparatus of the endoplasmic reticulum, bacterial plasma membrane or the mitochondrion. However, these synthesised targeting sequences are much less effective than genuine presequences (Schatz and Dobberstein, 1996). It is therefore clear that whilst no consensus exists between the primary sequences of the mitochondrial targeting presequences, there are distinct signals located within these sequences which have yet to be completely deciphered.

1.6 The targeting and position of the components of 2-oxoacid dehydrogenase complexes

The components of the 2-oxoacid dehydrogenase complexes are all found in the matrix, in association with the inner membrane. Targeting to the mitochondrion is expected to proceed via the normal route through the translocation contact site.

1.6.1 The extended presequences of E2

The presequences of the components of the 2-oxoacid dehydrogenase complexes are all typical of mitochondrial targeting signals with the exception of the E2 components of each complex. These polypeptides possess highly extended presequences, PDC-E2 having the longest at 73 amino acids. Extended presequences often contain additional information, such as the bipartite signals of conservatively sorted inner membrane and intermembrane space proteins. However, they could also aid in maintaining the solubility of highly hydrophobic polypeptides, possibly through the recruitment of extra chaperones, or by acting directly as an intramolecular chaperone (Hajek *et al.*, 1997). Since E2s are known to be highly hydrophobic and easily form

aggregates, this is one possible reason for them having extended presequences. Antibodies raised to the mature form of E2 are unable to recognise any epitopes in precursor E2 (Hunter and Lindsay, 1986), consistent with the theory that PDC-E2 is kept in an especially unfolded state to avoid aggregation of the precursor.

1.7 Aims of research

Having discussed what is known about the targeting of PDC-E2 to the mitochondrion and the behaviour of PDC kinase, we are able to pin-point two areas which require some investigation.

1.7.1 The role of the extended presequence of PDC-E2

The first aim of this research was to uncover the properties held by the extended presequence of PDC-E2. Why is it so long- are there a number of signals held within the signal, and what are their functions? Is there a shorter portion of the sequence which is sufficient for targeting? If so, what is the function of the rest of the presequence? Finally, one could ask if this presequence was capable of mistargeting to the rough endoplasmic reticulum, and therefore a possible cause of the abnormal expression of E2 epitopes on the surface of bile duct epithelia in PBC.

Some answers to these questions may be obtained by testing parts of the presequence for their ability to target polypeptides to the mitochondrion. The main part of the research detailed in this thesis describes the strategy used to answer these questions, which involved initial construction of a new plasmid vector designed for such experiments, made not only for this research, but for more general studies on mitochondrial targeting *in vivo*. Secondly, this thesis describes the methods used to clone the presequence and, based on a

computer-aided analysis of the polypeptide sequence, the selection and synthesis of various presequence segments for testing in the new plasmid. Finally, this work reports on the *in-vivo* expression of the plasmid-presequence constructs and analysis of the results.

1.7.2 The behaviour of PDC kinase

It is known that PDC kinase phosphorylates three sites on E1 when both the E1 and the kinase are attached to the E2 core, but only acts upon one site when the E1 is freed from the core. However, it is unknown what takes place when E1 is bound to the core but the kinase is freed from the complex. If this was determined, it would help in understanding which interactions permit the phosphorylation of all three sites.

As an adjunct to the main research described in this thesis, further experiments were performed which were designed to answer this question, utilising electrospray mass spectroscopy to measure by mass the number of phosphorylations on E1 made by bound and free kinase.

Chapter 2

Materials and Methods

2.1 Materials

2.1.1 Kits

Sequenase™ Version 2.0 DNA Sequencing Kit and ECL Western blotting detection reagents were purchased from USB through Amersham International plc., Aylesbury, Bucks, UK.

Taq DyeDeoxy™ Terminator Cycle Sequencing kits and sequencing-grade phenol/chloroform/H_2O (68:18:14) were obtained from Applied Biosystems Inc. (ABI), Warrington, Cheshire, UK.

The MERmaid® kit was bought from BIO-101, 1070 Joshua Way, Vista, CA 92083, USA.

Micro BCA Protein Assay Reagent Kit was obtained from Pierce Laboratories, Rockford, Illinois, U.S.A.

Wizard™ *Plus* Mini Preps and Maxi Preps DNA purification system and the Wizard™ PCR preps were purchased from Promega, Southampton, UK.

Qiaex gel extraction kit and Qiagen midi and maxi kits were bought from Qiagen Ltd., Boundary Court, Gatwick Rd., Crawley, East Sussex, England.

PCR-Script™ Amp SK(+) Cloning Kit was obtained from Stratagene, Cambridge, UK.

2.1.2 Enzymes

Chloramphenicol acetyltransferase was a kind gift from Dr. Ann Lewendon and Prof. William Shaw, University of Leicester.

T4 DNA ligase with 10x buffer, and *Taq* DNA polymerase with 10x buffer and 25 mM $MgCl_2$ was supplied by Promega, Southampton, UK.

Vent$_R$® (exo-) DNA polymerase and 10x buffer was supplied by

New England Biolabs, Hitchen, Herts, UK.

'Native' and 'Cloned' *Pfu* DNA polymerase and supplied 10x buffers were obtained from Stratagene, Cambridge, UK.

*Eco*RI, *Hin*dIII, *Kpn*I, *Nhe*I, *Pst*I, calf intestinal alkaline phosphatase (CIP), and T4 DNA ligase, all with supplied 10x buffers, were obtained from Boehringer Mannheim Gmbh., Lewes, East Sussex, UK.

*Cpo*I and *Bsi*MI were purchased from Amersham International plc., Aylesbury, Bucks, UK.

2.1.3 Chemicals

Agarose (electrophoresis grade) was supplied by Appligene®-Oncor®, Parc d'Innovation, 67402, Illkirch, France.

Isopropylthio-β-D-galactoside, X-Gal and DTT were from Boehringer Mannheim Gmbh., Lewes, East Sussex, UK.

Acrylamide, *N,N'*-Methylenebisacrylamide, Poly(ethylene glycol) 6000 grade and Triton X-100 were obtained from Fisons, Loughborough, England.

Tris and caesium chloride were purchased from Gibco BRL Life Technologies, Inchinnan Business Park, Paisley.

Leupeptin was from the Marketing Association, Herts, England

Substrates and coenzymes for enzyme assays, PMSF, benzamidine-HCl, MOPS, Tween-20, antifoam A concentrate, Coomassie brilliant blue, ampicillin, tetracycline and collagenase were supplied by Sigma Chemical Company, Poole, Dorset, UK.

HPLC grade water and acetonitrile were from Rathburn Chemicals Ltd. Walkerburn, Scotland.

Methanol, propan-1-ol, propanol-2-ol and glycerol were from Fisher Scientific UK Ltd., Bishop Meadow Rd, Loughborough, Leicestershire,

UK.

Tris equilibrated phenol (Tris-phenol) was purchased from Fisons Scientific, Bishop Meadow Rd, Loughborough, Leicestershire, UK.

All other chemicals were bought from BDH Chemicals Ltd., Poole, U.K.

2.1.4 Electrophoresis markers

Low molecular weight ('Low M_r') marker proteins were from Pharmacia Biotech UK, St. Albans, Herts.

'100 bp ladder' DNA markers were from Gibco BRL Life Technologies, Inchinnan Business Park, Paisley.

'ϕX-174 RF DNA-*Hae*III 'digest DNA markers were purchased from Promega, Southampton, UK.

'DRIgest III' (λ DNA-*Hind*III/ϕX-174 RF DNA-*Hae*III digest) DNA markers were supplied by Pharmacia Biotech UK, St. Albans, Herts.

2.1.5 Media

Bacto®-tryptone, Bacto®-agar and Bacto®-yeast extract were supplied by Difco, East Molesey, Surrey, UK.

All cell culture materials were supplied by Gibco BRL Life Technologies, Inchinnan Business Park, Paisley.

2.1.6 Bacterial Strains

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

Genotype: Δ(*mcrA*)183, Δ(*mcrCB-hsdSMR-mrr*) 173, *endA*1, *supE*44, *thi*-1 *recA*1, *gyrA*96, *relA*1, Lac [F' *proAB*, *lacI*^{q}ZΔM15, Tn10(tet^{r})] (Bullock, 1987)

Escherichia coli DH5α (Stratagene Ltd., Cambridge): a recombination deficient, suppressing strain for the propagation of plasmid DNA.

Genotype: *supE*44 Δ*lacU*169 (f80 *lacZ*ΔM15) *hsdR*17 *recA*1 *endA*1 *gyrA*96 *thi*-1 *relA*1 (Hanahan, 1983)

2.1.7 Plastics

Centricon 3, 10 and 30 concentrators were bought from Amicon Ltd., Stonehouse, Gloucestershire.

Quickseal polyallomer ultracentrifuge tubes (3.9 ml) were from Beckman Instruments (UK) Ltd., Sands Industrial Estate, High Wycombe, Bucks., England.

2.1.8 Miscellaneous

Alconox^{TM} detergent was supplied by Aldrich Chemical Company Ltd., Gillingham, Dorset, UK.

Sterile Acrodisc® 0.2 μm filters were supplied by Gelman Sciences Ltd., Northampton, UK.

Gel-Mix® 6 was from Gibco BRL Life Technologies, Innchinan

Business Park, Paisley.

Permacel™ tape was obtained from Genetic Research Instruments, Dunmow, Essex, UK.

Dialysis tubing (Visking size 1-8/32') was supplied by Medicell International Ltd, London, UK.

X-Omat S film was obtained from Kodak Ltd., Dallimore Road, Manchester, England.

Hybond-C extra nitrocellulose sheets, High performance Hyperfilm-ECL and the ECL Western blotting kit were supplied by Amersham International plc., Aylesbury, Bucks, UK.

Non-immune donkey serum was supplied by the Scottish Antibody Production Unit (SAPU), Lanarkshire, Scotland.

Bovine hearts were obtained from Paisley Abbatoir, Sandyford Road, Paisley, Scotland.

2.1.9 Equipment

The automated sequencer was from Applied Biosystems Inc. (ABI), Warrington, Cheshire, UK.

The X-Omat-100 processor was purchased from Kodak Ltd., which was manufactured for Kodak Ltd. by L'accessori radio grafico, Burago, Molgora, Italy.

The Fast Protein Liquid Chromatograph (FPLC) , FPLC columns and all gel chromatography materials were purchased from Pharmacia Biotech UK, St. Albans, Herts.

The High pressure liquid chromatograph (HPLC) was from Hewlett Packard, Eskdale Rd., Wokingham, Berks.

The C4 reverse phase column was a Delta-Pak™ HPI C4, 2.0 x 150 mm, from Waters, Watford, Herts.

The Electrospray mass spectrometer (ESMS) was produced by VG Biotech Ltd., Altricham, Cheshire.

The centrifuges and the refrigerated benchtop centrifuge were bought from Beckman Instruments (UK) Ltd., Sands Industrial Estate, High Wycombe, Bucks., England.

The benchtop centrifuges (MSE Microcentaurs) were from Scotlab, Coatbridge, Scotland.

The PCR machines were purchased from Techne (Cambridge) Ltd., Duxford, Cambridge and from Genetic Research Instruments Ltd., Gene House, Dunmow Rd., Felsted, Dunmow, Essex, England.

2.2 General cell culture

Except where described, all following cell culture methods were carried out under sterile conditions. Mammalian cell culture procedures were performed inside sterile flow hoods with bottle necks flamed before sterile disposable pipettes were introduced to add or withdraw fluids. All media were preincubated to test for the presence of bacteria or yeast. Prokaryotic cell culture procedures could be carried out on the open bench. Bottle necks were still flamed before taking or adding fluid, again, using sterile pipettes. Platinum wire loops, used for taking isolated colonies of cells, or taking up small amounts of broth, were held in the flame of a bunsen until glowing red to sterilise between uses. Similarly, glass spreaders were dipped in 70% (v/v) ethanol and lit in a bunsen to obtain sterility. Platinum wires and glass spreaders were allowed to cool before use.

Most plasticware, such as flasks, graduated pipettes, syringes, filters, universals and polypropylene centrifuge tubes were purchased in pre-sterilised packs. Pipette tips for the auto-pipettors were placed in sealable racks and autoclaved (126°C, 15 psi, 15 min); eppendorfs were bottled in jars and

autoclaved.

Media were supplied as sterile for eukaryotic cell culture. For bacterial cell culture, media were either autoclaved or filter sterilised into sterile containers.

2.3 Eukaryotic cell culture

2.3.1 Starting cultures from frozen stocks

Cells which had been previously frozen (as described below) were revived by removing an aliquot of cells from liquid nitrogen and placing in a water bath at 37°C. Once defrosted, medium containing the cells was transferred to a medium sized flask (75 cm^2). Dulbecco's minimum essential medium ('DMEM', 450 ml dH_2O, 50 ml 10 x Dulbecco's minimum essential medium, 50 ml new born calf serum, 5 ml 200 mM '100x' L-glutamate, 5 ml '100x' MEM non essential amino acids, 35 ml 7.5% (w/v) sodium bicarbonate), preincubated at 37°C, was added (20 ml), and the flask was incubated at 37°C in 10%(v/v) CO_2, with the lid screwed on loosely to allow transfer of gases.

2.3.2 Passaging

Once cells had grown to confluence, they could be split into 3-7 new flasks of the same size, or put into new vessels, such as transfection dishes. To do this, the medium was taken off, and replaced with 5-20 ml of Versene (140 mM NaCl, 2.5 mM KCl, 8 mM Na_2HPO_4, 1.5 mM KH_2PO_4, 0.5 mM EDTA, 3 µM Phenol Red) depending on the size of the flask. The cell monolayer was gently washed with this, then the Versene was replaced with 1-2 ml of 1:4 trypsin/Versene (0.25% (v/v) trypsin). The flask was incubated with the trypsin until the cells started to appear round or detached from the flask, by

checking under a microscope. A gentle tap on the side of the dish could also help to dislodge the cells from the flask. Once the cells had all become dislodged, a small amount of DMEM was added to stop the tryptic digestion, and the medium was distributed among the new flasks. A volume of medium, appropriate to the size of flask, was added to each flask, and each flask incubated at 37°C in 10% (v/v) CO_2, with the lid screwed on loosely.

2.3.3 Freezing down cells

As before, cells were removed from the flask by use of a wash with Versene, then trypsin/Versene. For a large (162 mm^2) flask, 10-20 ml medium was added, and the cells spun down at 4°C at 2000 rpm (in a Beckman GS-15R benchtop centrifuge) for 5 min. The cells were gently taken back up in 9 ml medium, and 1 ml DMSO. The solution was then aliquoted into 1.8 ml sterile Nunc™ Cryotubes, and kept on ice until freezing. The aliquots (1.5 ml) were frozen in a -80°C freezer and placed in liquid nitrogen for long term storage 16-24 h later.

2.3.4 Transfection of DNA

2.3.4.1 Transfection by Lipofection

Cells were grown to 50% confluence in 60 mm transfection dishes, by seeding at low densities and culturing cells until they reached the desired numbers. The original growth medium was removed, and the cells were washed with OPTIMEM 1 medium. Meanwhile, 2-5 µg of DNA solution was added to 200 µl OPTIMEM 1 medium, and 4-40 µl of lipofectin reagent was diluted into another 200 µl OPTIMEM 1. The two solutions were then gently mixed together and allowed to stand for 15 min at room temperature. After this incubation, the medium was removed from the cells, and the

lipofectin/DNA mixture put in place. The cells were incubated in the presence of the DNA for 5-24 h (usually overnight) under normal growth conditions. The lipofectin/DNA mixture was removed, and the cells again grown in normal growth medium for 48 h before being harvested.

2.3.4.2 Transfection by DEAE dextran

The protocol employed was based on that devised by Sambrook *et al.* (1989). Cells were grown to 50% confluence in 60 mm transfection dishes, by seeding at low densities and culturing cells until they reached the desired numbers. The growth medium was removed, and the cells washed with OPTIMEM 1 medium. A DNA solution (250 µl 1 M Tris (pH 7.3), 200 µl 0.5 M HEPES (pH 7.2), 4.475 ml OPTIMEM 1, 75 µl DEAE-dextran solution (50 mg/ml) and 5 µg DNA) was put over the cells and incubated in normal growth conditions for 4 h. This solution was removed from the cells, and the cells were incubated for 1 min at room temperature in OPTIMEM 1 containing 25% (v/v) glycerol. After removal of the glycerol solution, the cells were again washed with OPTIMEM 1, and 5 ml normal growth medium was added to the cells. The cells were then allowed to grow in normal conditions for 48 h before harvesting.

2.3.4.3 Transfection by calcium chloride

The protocol employed was based on that devised by Sambrook *et al.* (1989). Cells were grown to 50% confluence in 60 mm transfection dishes, by seeding at low densities and culturing until they reached the desired numbers. A DNA solution was prepared to add to the growth mediumas follows; 5 µg DNA was added to 437.5 µl water and 62.5 µl 2 M $CaCl_2$. This was added to 500 µl HBS buffer (Hepes buffered saline; 280 mM NaCl, 1.5 mM Na_2HPO_4,

10 mM KCl, 12 mM dextrose, 50 mM HEPES (pH 7.05 with NaOH), all filter sterilised), and left to stand for 1 h, before being added to the cells in the growth medium. The cells were incubated with the $CaCl_2$/DNA precipitate for 20 h, under normal growth conditions, before the medium was removed from the cells and replaced with fresh growth medium. The cells were then grown for a further 48 h before being harvested.

2.4 Prokaryotic cell culture

Bacteria were grown either on agar plates or in broth cultures. Agar plates were made with a broth medium containing 1.5% (w/v) agar. Molten, sterilised agar medium could be poured out into Petri dishes in approx. 30 ml measurements. Once set and dried in a 37°C incubator for 1 h, bacteria could be streaked or spread onto the plate, and incubated upside down. This would produce isolated colonies of bacteria derived from a single cell, with any condensation from metabolism collected on the lid. Bacteria could also be grown as a suspension, or liquid culture, where the cells were grown in a suitable volume of medium, and incubated. Such cultures were usually grown aerobically, which entailed vigorous shaking of the culture at approx. 150-200 rpm. For this reason, a defined volume of medium was shaken in a vessel, often a conical flask, which had a greater capacity (at least 5 times the volume of the culture), as this allowed efficient aeration of the contents.

Bacteria were usually grown in Luria broth (10 g Bacto-tryptone, 10 g NaCl, 5 g yeast extract per litre H_2O, pH 7.4) which had been sterilised by autoclaving, or on Luria broth-agar (10 g Bacto-tryptone, 10 g NaCl, 5 g yeast extract, 15 g agar per litre H_2O, pH 7.4) which had been sterilised by autoclaving.

Bacteria were supplied in stabs, on plates, and in frozen cultures.

2.4.1 Starting cultures from frozen stocks

When the starting culture was a broth or a frozen liquid culture a small volume could be taken and spread out on a fresh agar plate using a sterile glass spreader. Alternatively, a sterile wire loop could be used to take a small volume of culture to be streaked out on a plate. The plate would then be incubated overnight at 37°C upside down.

2.4.2 Starting cultures from agar plates

When starting a culture from bacteria on a streaked-out plate, a single colony could be taken from the plate using a sterile wire loop, re-streaked on a fresh plate of a similar medium agar and incubated overnight at 37°C.

2.4.3 Growing and subculturing bacterial cultures

Subculturing bacterial cells could be performed by picking a single colony off an agar plate and re-streaking as above.

To grow larger numbers of cells, a single colony could be taken from the plate, using either a sterile pipette tip or a sterile wire loop, and dispersing the colony into 5 ml broth in a sterile universal. After an 8 h or overnight incubation, at 37°C with shaking, a small amount of the broth could be taken out with a sterile pipette, and diluted 100 fold into fresh broth. This broth could then be grown overnight again at 37°C with shaking.

2.4.4 Freezing down cells

Cells could be stored for 1-2 weeks as colonies on agar plates at 4°C. For longer term storage, up to a number of months, an 8 h or overnight broth

culture could be mixed with an equal volume of glycerol and kept at -20°C in sterile eppendorfs. The same culture/glycerol stock, stored at -80°C, could be kept for even longer term storage (a number of years).

2.4.5 Transformation with DNA

2.4.5.1 Competent cell preparation

Bacteria were regularly used for cloning and growing plasmids by transforming them with a small amount of plasmid DNA. A suitable strain was selected (e.g. *E. coli* XL1 Blue MRF') and streaked out onto a sterile LB agar plate and grown overnight at 37°C. Isolated colonies (4-5) were then taken from the plate with a sterile wire loop and dispersed into 15 ml LB broth (supplemented with 20 mM sterile $MgSO_4$) in a 50 ml polypropylene tube. Another 4-5 colonies were treated likewise so that, in total, a 30 ml culture was made. The two tubes were incubated with shaking at approx. 200 rpm at 37°C for 4.5 h, then spun at 4000 rpm in a JA17 rotor in a Beckman J2-21 centrifuge at 4°C for 10 min. The supernate was poured off and the pelleted cells gently resuspended in 10 ml modified 'TFB' (10 mM MES pH 6.3 with KOH, 50 mM $CaCl_2$, 100 mM KCl, filter sterilised and stored at 4°C). The cells were left on ice for 30 min with occasional swirling. At this stage, the cells were pelleted as previously, and resuspended in 4 ml modified 'TFB', and left for 15 min on ice. 'DnD' solution (140 µl) was added (1.53 g DTT, 9 ml DMSO, 100 µl 1 M potassium acetate (pH 7.5), made to 10 ml with H_2O, filter sterilised). These cells, competent for efficient transformation, remained viable for up to two weeks at 4°C using this method.

2.4.5.2 Transforming competent cells

Competent cells (400 µl) were pipetted into a sterile eppendorf for each transformation. Plasmid (10-50 ng), in a maximum of 20 µl, was added to the cells and gently mixed by pipetting up and down. This was incubated for 45 min on ice, then warmed to 42°C in a water bath for 90 sec. The transformation mix was then rapidly cooled on ice, and left for 2 min, before adding 800 µl SOC medium (20 g bacto-tryptone, 5 g yeast extract, 0.5 g NaCl, 10 ml 250 mM KCl, pH 7.0 with NaOH, volume made to 1 L, and sterilised by autoclaving; before use, 5 ml of filter sterilised 2 M $MgCl_2$ and 20 ml filter sterilised 1 M glucose were added), pre-warmed to 37°C, to the transformation mix. This was then incubated, with shaking, at 37°C, for 1 h.

After the 1 h incubation, the cells were pelleted by centrifugation for 1 min at 6500 rpm in a bench top centrifuge. The supernate was removed, and the cells gently resuspended in 100-200 µl of SOC medium. This was pipetted onto a suitable selective media agar plate, spread out on the plate using a sterile glass spreader and incubated overnight at 37°C.

2.4.5.3 Selective media

Plasmids most often contain a gene which when expressed in the host bacterium, confers a selectable advantage over other bacteria not carrying the plasmid. The 'marker' gene used throughout this work was the ampicillin resistance gene. Media which were selective for the bacteria carrying this plasmid contained 25 µg ampicillin per ml of broth, or 50 µg/ml of agar. The ampicillin was added as a filter sterilised 50 mg/ml solution after media had been autoclaved and cooled to below 50°C. Ampicillin was added to agar before it set, at approx. 45°C.

Agar plates could also have blue/white selection, to test for the

insertional inactivation of the β-galactosidase gene. Plasmids which had been used for cloning short pieces of DNA (see section 2.10.4.2) sometimes utilised this system. In addition to selecting for bacteria carrying the plasmid ampicillin resistance gene, it was possible to detect if the DNA had been inserted into the plasmid depending on whether or not the β-galactosidase gene had been inactivated by the presence of a DNA insert. To test for an active β-galactosidase gene, expression of the gene was induced by the presence of IPTG, and the gene product hydrolysed the X-Gal substrate, which then turned blue. LB-ampicillin plates were first coated with X-Gal (10% (w/v) X-Gal in DMSO) by spreading 20 µl solution into a plate until the solution had dried into the plate, followed by 20 µl IPTG (0.2 M IPTG solution, filter sterilised), spread on in the same way.

2.5 Protein methods

No protein methods required sterility. However, care was taken to ensure all plasticware and glassware was clean.

2.6 Protein purification

2.6.1 Protein extraction from cells

2.6.1.1 Obtaining cell suspensions

Cells grown in 60 mm culture dishes were washed by first discarding the culture medium, and then gently rinsing the cells with ice-cold PBS (140 mM NaCl, 2.5 mM KCl, 8 mM Na_2HPO_4, 1.5 mM KH_2PO_4, sterilised by autoclaving). Rinsing with PBS was performed twice, and the cells were scraped off the surface of the dish with a sterile cell scraper, resuspended in 1-1.5 ml ice-cold PBS and pipetted into a sterile eppendorf. The suspension was centrifuged at 2000 rpm for 5 min at 4°C, and the supernate removed with a

pipette before resuspending in 100 µl ice cold PBS. Cells were stored at 4°C until fractionated.

2.6.1.2 Fractionating cells

Cells could be broken open to release the contents of both the mitochondrial matrix and the cytoplasm by simply adding 1% (v/v) Triton X-100 to the cell suspension in PBS, and then vortexing momentarily. Debris, such as membranes and nuclei, could be pelleted by centrifugation at high speed in a benchtop centrifuge at 4°C and the supernate retained for further analysis. Alternatively, a freeze-thaw process, where the cells were frozen to -70°C, thawed and vortexed, could also achieve cell lysis in the absence of detergents.

Fractionation of the cells into cytoplasmic and cytoplasm + mitochondrial matrix fractions was carried out using digitonin. Digitonin is a detergent which crystallises cholesterol, causing lipid membranes to rupture. Since more cholesterol is found in plasma membranes than in mitochondrial membranes, less digitonin is required to permeabilise the plasma membrane. A cellular suspension in PBS was fractionated with 0.05 mg digitonin per million cells to 1mg digitonin per million cells. Digitonin was made as 10 and 1mg/ml standard solutions in Hepes-buffered saline (0.8 g NaCl, 0.037 g KCl, 0.027 g $Na_2HPO_4.2H_2O$, 0.2 g dextrose, 1 g HEPES in 100 ml H_2O, pH 7.05 with NaOH, and filter sterilised). Cells were incubated on ice for 2 min, and membranes and unbroken organelles pelleted by centrifugation at high speed in a benchtop centrifuge at 4°C, with the supernate retained for further experimental work.

2.6.2 Pyruvate dehydrogenase complex (PDC) purification

This method was adapted from Stanley and Perham (1980). PDC was isolated from bovine hearts which had been stored at -80°C and were subsequently defrosted overnight at 4°C. Prior to freezing, the fat and connective tissue was removed, and the rest of the heart diced into cubes approx. 1 cm x 2 cm. Bovine heart (300 g) was used in a single PDC preparation. The heart muscle was homogenised in a blender for 5 min, maximum speed, in an equal volume of 3% (v/v) Triton X-100 buffer (2.7 mM EDTA, 0.1 mM DTT, 3% (v/v) Triton X-100, 1 mM PMSF, 1 mM benzamidine-HCl, silicone antifoam (0.5 ml/l), 50 mM MOPS, pH 7.0). More 3% (v/v) Triton X-100 buffer was added to bring the total volume to approx. 1.5l. The homogenate was then spun at 10000 rpm in a JA-14 rotor in a Beckman J2-21 centrifuge for 20 min, and the pellets discarded.

The pH of the supernate was adjusted to pH 6.45 with 10% (v/v) acetic acid, and 0.12 vol of 35% (w/v) PEG 6000 added with stirring which was continued, on ice, for a further 30 min. The preparation was centrifuged at 14000 rpm in a JA-14 rotor in a Beckman J2-21 centrifuge for 15 min. The pellet was retained, resuspended and homogenised using a Teflon/glass homogeniser into a total volume of 200 ml using 1% (v/v) Triton X-100 buffer (2.7 mM EDTA, 0.1 mM DTT, 1% (v/v)Triton X-100, 1 mM PMSF, 0.15 µM leupeptin, 1 mM benzamidine-HCl, 50 mM MOPS, pH 6.8). The homogenate was centrifuged in a JA-17 rotor at 17000 rpm in a Beckman J2-21 centrifuge for 40 min, and the supernate retained. The solution was filtered through muslin to remove fat particles; 0.013 vol 1 M $MgCl_2$ was added to the solution, followed by 0.05 vol 1 M sodium phosphate buffer (pH 6.3), which was added whilst the solution was stirring. During the addition of the phosphate buffer, 0.5 M NaOH was used to prevent the pH of the solution falling below pH 6.5.

After adjusting the pH of the solution to 6.45 with 10% (v/v) acetic acid, 0.12 vol 35% (v/v) PEG 6000 was added with stirring, and retained for 30 min on ice. The solution was centrifuged as previously, in a JA-17 rotor at 16000 rpm for 10 min, and the supernate discarded. This pellet was then left overnight in 80 ml 1% (v/v) Triton X-100 buffer at 4°C, and allowed to solubilise in the buffer. Resuspension of the pellet was performed using the homogeniser, and then centrifuged at 17000 rpm in a JA-17 rotor for 60 min.

All subsequent steps were monitored for PDC and OGDC activity (enzyme assays 2.7.1.4 and 2.7.1.5). To the supernate was added 0.04 vol 35% (v/v) PEG 6000; this was stirred for 30 min on ice. An aliquot (1 ml) of the solution was then taken and centrifuged at high speed in a benchtop centrifuge to ascertain if the OGDC had been precipitated and pelleted whilst the PDC was left in solution. This was analysed by assaying the supernate for PDC and OGDC activity. If the OGDC had not been reduced to 5% of the initial activity, then another 0.01 vol 35% (v/v) PEG 6000 was added to the solution, and stirred for 30 min on ice. Another aliquot was taken to assay for PDC and OGDC activity as before, to determine if the OGDC was pelleting whilst the PDC remained in solution. Additional 0.01 vol 35% (w/v) PEG 6000 were added until this was achieved, after which the solution was centrifuged at 17000 rpm in a JA-17 rotor for 10 min.

The supernate was centrifuged in a Ti70 rotor at 45000 rpm for 2.5 h at 4°C prior to leaving the pellet overnight at 4°C in a minimal amount of 1% (v/v) Triton X-100 buffer. The softened pellet was subsequently resuspended in more 1% (v/v) Triton X-100 buffer, to obtain a concentration of 10-50 mg/ml. The purity of the PDC could then be checked by assays for OGDC and PDC activity, or by SDS-PAGE (section 2.7.3). A Lowry or Micro BCA protein assay was employed to measure PDC concentration (section 2.7.5.1. and 2.7.5.2., respectively).

2.6.3 E1/E3 purification

PDC solutions were mixed 1:1 with ice cold dissociation buffer (4 M NaCl, 50 mM imidazole, pH 7.0) and incubated on ice for 1 h. The PDC was subsequently spun in a benchtop microfuge at 13500 rpm for 15 min at 4°C to remove any insoluble material. Gel permeation chromatography of the supernate was performed on a Pharmacia FPLC Superose 12 column (16 mm x 500 mm, 100 ml bed volume). The column was pre-equilibrated with 200 ml running buffer (filtered and degassed 1 M NaCl, 50 mM imidazole, pH 7.0) at a flow rate of 0.1 ml min^{-1}. Dissociation buffer (2 ml) was injected onto the column at a flow rate of 0.5 ml min^{-1}, then 2 ml dissociated PDC sample was injected onto the column at the same flow rate. All the injected material was then eluted from the column in running buffer flowing at 1 ml min^{-1}. An absorbance trace at 280 nm recorded the elution profile, and 2 ml fractions were collected using a Pharmacia Frac-100 fraction collector.

The fractions containing the E1/E3 polypeptides were diluted in HPLC grade H_2O, and concentrated by spinning in centricon 10 tubes at 7000 rpm for 4 h or until concentrated. Again, HPLC grade H_2O (10 ml) was added to each sample to dilute out the salt, and reconcentrated as before. This was repeated 4 times in total to remove the majority of the salts, and leave the E1/E3 in HPLC grade H_2O.

2.6.4 HPLC of E1 and E3

The E1/E3 fraction was subsequently spun in a benchtop microfuge at 13500 rpm for 5 min at 4°C. The supernate was carefully removed and placed on ice in preparation for HPLC isolation of the polypeptides. HPLC was carried out on a Hewlett-Packard HPLC using a Delta-pak HPI C4 reverse phase column (2 ml). The column was first equilibrated in solution A (filtered

and degassed 98:2 H_2O:acetonitrile with 0.1% (v/v) trifluoroacetic acid) at a flow rate of 0.3 ml min^{-1}. The sample was injected onto the column at the same flow rate. The various proteins in the sample were eluted from the column by changing the running solution over a period of 60 min from solution A to solution B (filtered and degassed 10:90 H_2O:acetonitrile with 0.1% (v/v) trifluoroacetic acid). Fractions were collected from the column with on-line monitoring at 214 nm to detect eluted polypeptides.

2.6.5 Ultrafiltration

Amicon Centricon-10 concentrators were used, since these had a filter cut-off point at 10,000 M_r, therefore retaining E1α, E1β and E3. The solution was pipetted into the concentrator, and was centrifuged in a JA-17 rotor in a Beckman J2-21 centrifuge at 5,000 xg for 1 h at a time, until the desired volume was reached.

To dialyse into a different buffer, the same procedure was followed to concentrate the proteins into a minimal amount of buffer. The new buffer was then added, and the proteins reconcentrated. This concentration/wash cycle was repeated 3 or 4 times to ensure complete replacement of the original buffer.

2.7 Protein detection and characterisation

2.7.1 Enzyme assays

2.7.1.1 Citrate synthase

This method was based on that of Srere (1969). Into a 1 ml glass or plastic cuvette was pipetted 200 µl 0.5 M Tris/HCl pH 8.0, 200 µl 0.5 mM DTNB (which was freshly prepared), 100 µl 3 mM acetyl CoA (also freshly prepared) and sample added, with dH_2O to give a final volume of 0.9 ml. Once oxaloacetate was added, the solution had final concentrations of approx.

0.1 M Tris/HCl, 0.1 mM DTNB, 0.3 mM acetyl CoA, and 0.5 mM oxaloacetate. The assay mix was pre-warmed to 25°C, and the reaction started by adding 100 µl 5 mM oxaloacetate. Enzyme activity was measured by monitoring the increasing absorbance at 412 nm as the product, CoASH, generated a mercaptide with DTNB.

2.7.1.2. Lactate dehydrogenase

This method was adapted from Pesce *et al.* (1964). Into a 1 ml glass or plastic cuvette was pipetted 943 µl phosphate buffer (0.1 M potassium phosphate buffer, pH 7.5), 330 µl sodium pyruvate solution (23 mM sodium pyruvate in phosphate buffer), and 16.7 µl NADH solution (12 mM sodium salt of NADH in phosphate buffer). The solution had final concentrations of approx. 94 mM phosphate buffer, 7.6 mM sodium pyruvate and 0.2 mM NADH. The assay solution was mixed and pre-warmed to 25°C, sample was added, and the enzyme activity measured by loss of absorbance at 340 nm, on conversion of NADH to NAD^+.

2.7.1.3 Malate dehydrogenase

This method was adapted from Kun *et al.* (1967). Into a 1 ml glass or plastic cuvette was pipetted 943 µl phosphate buffer (0.1 M potassium phosphate buffer, pH 7.5), 330 µl oxaloactetate solution (15 mM oxaloacetate in phosphate buffer), and 16.7 µl NADH solution (12 mM sodium salt of NADH in phosphate buffer). The solution had final concentrations of approx. 94 mM phosphate buffer, 0.5 mM oxaloactetate and 0.2 mM NADH. This was mixed and pre-warmed to 25°C, sample was added, and the enzyme activity measured by loss of absorbance at 340 nm, on conversion of NADH to NAD^+.

2.7.1.4 Pyruvate dehydrogenase complex

This method is one based on the procedure of Brown and Perham (1978). Into a 1 ml glass or plastic cuvette was added 670 µl solution A (50 mM potassium phosphate buffer, pH 7.6, containing 3 mM NAD^+, 2 mM $MgCl_2$, 0.2 mM ThDP), 14 µl solution B (0.13 M cysteine HCl, 0.13 mM Li_2CoASH), and 14 µl solution C (100 mM pyruvic acid). This was mixed and preincubated at 30°C, then sample was added, and the overall complex activity measured by increasing absorbance at 340 nm, as the reaction generates NADH.

2.7.1.5 2-Oxoglutarate dehydrogenase complex

This method is one based on the method of Brown and Perham (1978). Into a 1 ml glass or plastic cuvette was added 670 µl solution A (50 mM potassium phosphate buffer, pH 7.6, containing 3 mM NAD^+, 2 mM $MgCl_2$, 0.2 mM ThDP), 14 µl solution B (0.13 M cysteine HCl, 0.13 mM Li_2CoASH), and 14 µl solution C (100 mM 2-oxoglutarate). This was mixed and preincubated at 30°C, sample was added, and the overall complex activity measured by increasing absorbance at 340 nm, as the reaction generates NADH.

2.7.1.6 Chloramphenicol acetyltransferase

This method was adapted from that of Seed and Sheen (1988). Cell extracts could be measured for CAT activity using a radioactive assay where radiolabeled chloramphenicol was linked by CAT to butyrate, causing the chloramphenicol to partition into an organic phase.

To a 1.5 ml eppendorf was added 10 µl 1 M Tris/HCl, pH 7.8, 32 µl

dH_2O, 40 µl cell extract, 8 µl 0.446 mM D-threo-[dichloroacetyl-1-^{14}C] chloramphenicol and 10 µl 2.5 mM butyryl CoA. This was mixed and incubated for 1 h at 37°C, and the reaction stopped by adding 210 µl TMPD:Xylene (2:1). The eppendorf was vortexed for 15 sec, and then spun in a bench top centrifuge at 13000 rpm for 2 min. The upper, organic, phase (180 µl) was removed with a pipette, and put in a scintillation vial. To this was added 3 ml of Ecoscint scintillation fluid, the vial was vortexed to mix the solutions, and the amount of ^{14}C incorporation into product was measured in a scintillation counter.

2.7.2 E1α and phosphorylated E1α detection by ESMS

Fractions collected from the HPLC were injected onto the electrospray mass spectrometer (ESMS) at 0.02 ml min^{-1}. The ESMS was operated by Dr. Tino Krell. Data were collected by a computer and processed using the VG Mass Lynx software.

2.7.3 SDS-Polyacrylamide gel electrophoresis (SDS-PAGE)

Proteins could be resolved by the use of SDS-PAGE, with the discontinous Tris/glycine buffered system of Laemmli (1970). Prior to electrophoresis, an equal volume of Laemmli sample buffer (62.5 mM Tris/HCl, pH 6.8, containing 2% (w/v) SDS, 10% (w/v) sucrose and 2% (w/v) pyronin Y) was added to the samples. DTT was included to give a final concentration of 10 mM. Samples were then boiled for 3-5 min.

Gels with dimensions of 170 mm x 145 mm x 1.5 mm were made from the following stock solutions to give a stacking gel of 4.4% acrylamide (with 0.12 M Tris/HCl, pH 6.8, 0.1% (w/v) SDS, 0.1% (w/v) ammonium persulphate and 0.08% (v/v) TEMED), and a resolving gel beneath of 7-15%

acrylamide (with 0.375 M Tris/HCl pH 8.8, 0.1% (w/v) SDS, 0.1% (w/v) ammonium persulphate and 0.08% (v/v) TEMED). The stock solutions were: acrylamide (29.2% (w/v) acrylamide, 0.8% (w/v) N,N'-methlene bisacrylamide), stacking gel buffer (0.17M Tris/HCl, pH 6.8, 0.14% (w/v) SDS), resolving gel buffer (0.75M Tris/HCl, pH 8.8, 0.2% (w/v) SDS) and ammonium persulphate solution (10% (w/v) ammonium persulphate). Polymerisation of the acrylamide solution was initiated by the addition of TEMED.

Gels were placed in vertical gel electrophoresis kits (Gibco BRL), and electrode buffer (24 mM Tris/192 mM glycine buffer, pH 8.3, containing 0.1% (w/v) SDS) added. The samples could then be applied to wells formed in the stacking gel, and a current of 40-70 mA applied. Electrophoresis was stopped as the Pyronin Y reached the end of the gel and entered the lower buffer reservoir.

Peptide subunits could either then be visualised immunologically by Western blotting, or by staining the gel for proteins. Gels were stained by using 0.04% (w/v) Coomassie Brilliant Blue R in 10% (v/v) acetic acid and 25% (v/v) methanol, and gently shaking overnight, followed by destaining with 20% (v/v) methanol and 10% (v/v) acetic acid.

The sizes of particular proteins could be estimated by using low M_r protein markers to construct a calibration curve relating relative mobility (R_f, as described below), to log(M_r) for each protein standard. In this way, the relative molecular mass of sample proteins on the same gel could be determined. The standards were: phosphorylase b (94000); bovine serum albumin (67000); ovalbumin (43000); carbonic anhydrase (30000); soybean trypsin inhibitor (20000); and α-lactalbumin (14000). Relative mobility was calculated by the equation:-

$$R_f = \frac{\text{distance migrated by protein}}{\text{distance migrated by tracking dye}}$$

Where the determination of relative molecular mass of a protein was required from a Western blot analysis, prestained M_r protein standards with extra protein, myosin (H-chain; 200000), were used for easy identification on nitrocellulose.

2.7.4 Western blotting

Proteins which had been resolved by SDS-PAGE were transferred to nitrocellulose membranes for subsequent immunological detection following the method of Towbin *et al.* (1979).

Gels were overlaid with pre-wetted Hybond-C extra transfer membrane and placed into cassettes, between pre-wetted Whatman 3 MM chromatography paper. Cassettes were then placed into a Bio-Rad Transblot™ tank filled with transfer buffer (25 mM Tris, 0.19 M glycine, 0.02% (w/v) SDS, 20% (v/v) methanol) and a current of 40 mA applied overnight. A non-fixative stain, Ponceau S, was used to check that proteins had transferred successfully onto the nitrocellulose. This was then rinsed off the membrane with wash buffer or dH_2O.

Enhanced chemiluminescence (ECL) was employed to detect polypeptides which were of interest. This utilises anti-rabbit antibody linked horseradish peroxidase (HRP), which catalyses a chemiluminescent reaction (the oxidation of luminol in alkaline conditions) whilst bound to the rabbit antisera, which in turn is bound to the antigen of interest. The protocol followed is a modified version of the protocol supplied with the reagents.

After blotting of the proteins onto the nitrocellulose, the nitrocellulose sheet was placed in a plastic tray, and gently shaken with blocking solution

(20 mM Tris/HCl pH 7.2, 15 mM NaCl, 5% (w/v) non fat dried milk, 5% heat inactivated donkey serum, 0.2% (v/v) Tween 20) for 2 h at 37°C or 4 h at room temperature. The nitrocellulose was then incubated overnight, with gentle shaking, at 4°C with primary antibody solution (20 mM Tris/HCl pH 7.2, 1% (w/v) non fat dried milk, 1% (v/v) heat inactivated donkey serum, 0.1% (v/v) Tween 20, 1:500 - 1:5000 dilution of primary antibody). The primary antibody solution was removed, and replaced with wash solution (20 mM Tris/HCl pH 7.2, 15 mM NaCl, 1% (w/v) non fat dried milk, 1% (v/v) heat inactivated donkey serum) and incubated at room temperature with agitation for 30 min. This washing procedure was repeated four times, before replacing the solution with a secondary antibody solution (20 mM Tris/HCl pH 7.2, 150 mM NaCl, 1% (w/v) non fat dried milk, 1% (v/v) heat inactivated donkey serum plus 1:1000 dilution of anti-rabbit HRP antibody) and incubated at room temperature, with shaking, for 2 h. The nitrocellulose sheet was then washed as previously, for 3 washes. It was finally washed in a solution of 20 mM Tris/HCl pH 7.2 with 150 mM NaCl for 30 min at room temperature, with shaking.

To detect the HRP, equal volumes of Amersham's ECL detection reagents 1 and 2 were mixed, so that there was 0.125 ml of the detection mixture for every 1 cm^2 of nitrocellulose sheet. The detection mix was poured onto the nitrocellulose, on the side where the proteins were bound, and incubated for 1 min. The blot was drained and dabbed dry with tissue, wrapped in SaranWrap™, and placed, protein side up, in a film cassette. A sheet of autoradiographic film (ECL Hyperfilm) was placed on top of the nitrocellulose in the dark, and the cassette closed. The film could be exposed for a few seconds to half an hour, depending on the strength of signal. The film was then developed in a Kodak X-Omat automatic film processor.

2.7.5 Protein quantitation

2.7.5.1 Lowry protein assay

This method was based upon that of Maxwell *et al.* (1978) derived from Lowry (1951). Solution A (2% (w/v) Na_2CO_3, 0.4% (w/v) NaOH, 0.16% (w/v) sodium tartate and 1% (w/v) SDS in water), solution B (4% (w/v) cupric sulphate pentahydrate in water), solution D (equal volumes of Folin-Ciocalteau reagent and dH_2O) were required. A standard curve was constructed, using BSA at 1 mg/ml, from 0 mg to 150 mg, with each standard made up to 1 ml with water. The unknown protein samples to be assayed were also made to a final volume of 1 ml. To each assay in the standard curve and unknown samples 3 ml solution C (1 ml solution B added to 100 ml solution A, prepared freshly) was added, mixed and left for 15 min. Solution D (0.3 ml) was then added and mixed, and left for a further 30 min. All the assays were then measured for absorbance at 660 nm, the unknown samples being compared to the standards for quantitation of protein content.

2.7.5.2 Micro BCA protein assay reagent kit

To measure protein concentrations between 20 µg/ml and 1.0 µg/ml, the Micro BCA protein assay reagent kit was used. The kit utilises bicinchoninic acid, a molecule which reacts with Cu^{1+} to form a purple reaction product with a peak absorbance at 562 nm. This allows the quantitation of Cu^{1+} which is generated by the reaction of peptide bonds and individual amino acid side chains with Cu^{2+}.

Three reagents were supplied; Micro reagent A (MA), containing sodium bicarbonate and sodium tartrate in 0.2 M NaOH, Micro reagent B (MB), containing 4% (w/v) BCA in water, and Micro reagent C (MC), containing 4% (w/v) cupric sulphate pentahydrate in water.

To assay protein, a working reagent was made. This consisted of 2 vol solution MC and 48 vol solution MB mixed together before 50 vol solution MA were added and mixed.

A set of protein standards, ranging from 20 µg/ml to 1.0 µg/ml of BSA were made, and 1 ml added to 1 ml of the working reagent. Likewise, the unknown protein samples were diluted to give a range of concentrations which would lie within the standard curve, and 1 ml each dilution added to 1 ml working reagent. All the reactions were incubated at 60°C for 1 h, and then the absorbance at 562 nm measured. The unknown protein concentrations could then be deduced from the standard curve.

2.7.6 Treating PDC with ATP

This method was adapted from that used by Hucho *et al.* (1972). PDC was diluted with 2 vol kinase buffer (20 mM potassium phosphate buffer, pH 7.5, 2 mM $MgCl_2$, 2 mM DTT). To start the treatment, 4 mM ATP was added. Loss of activity was followed using consecutive assays for PDC activity, noting the percentage loss of activity with time. Activity was expected to drop to below 5% after 20 min.

2.7.7 Treating PDC with collagenase

This method was adapted from that used by Rahmatullah *et al.* (1990). PDC was diluted with 2 vol of collagenase buffer (50 mM Tris/HCl pH 7.5, 2 mM $CaCl_2$, 3 mM NAD^+, 0.2 mM ThDP). Collagenase was added, 1mg collagenase for every 10mg PDC. Loss of activity was followed using consecutive assays for PDC activity, and noting the percentage loss of activity. Activity was expected to drop to 10-20% after 40 min.

2.8 DNA methods

Except where described, all the following methods were carried out under sterile conditions to reduce contamination with DNA from micro-organisms. Procedures could be carried out on the open bench using sterile plasticware, glassware and media. Most plasticware, such as flasks, graduated pipettes, syringes, filters, universals and polypropylene centrifuge tubes were purchased in pre-sterilised packs. Pipette tips for the auto-pipettors were placed in sealable racks and autoclaved; eppendorfs were bottled in jars and also autoclaved.

2.9 DNA purification

2.9.1 Plasmid purification

Plasmids were purified in a number of different ways over the course of this research. It was found that some methods were more reliable than others, and it was advantageous to vary the preparation method depending on how much DNA, what purity of DNA and how quickly the DNA was needed.

For small amounts of DNA (up to 5 μg) required relatively quickly, but at a higher expense, Wizard™ Mini Preps were routinely employed; for small amounts of DNA, but without using expensive kits (e.g. for screening a large number of clones by restriction analysis), a standard 'Mini Prep' protocol was used. For medium to large amounts of DNA (up to 0.5 mg) with only reasonable purity (e.g. for restriction digests to obtain segments of plasmids), a 'Maxi Prep' was used. Again, using a Maxi Prep kit could improve yield and quality, but at higher expense. Finally, for a large amount (up to 5 mg) of highly purified DNA (e.g. for transfection of eukaryotic cells or for permanent stocks of DNA) a 'caesium chloride density centrifugation' method was employed.

For all these DNA preparations, a culture of bacterial cells, carrying the plasmid of interest, was grown as described, in a medium selective for the plasmid.

Alkaline lysis forms the basis of all these plasmid preparations, commercial and non-commercial, as the means of releasing the DNA from the bacterial cells whilst precipitating and pelleting all the protein and carbohydrate as well as the genomic DNA.

When not using commercial kits, 3 solutions were used: solution 1 was a resuspension buffer, containing 50 mM glucose, 25 mM Tris/HCl pH 8.0 and 10 mM EDTA, sterilised by filtering. Solution 2 was a lysis buffer, best made freshly, which was a 0.2 M NaOH solution with 2% (w/v) SDS. Solution 3 was a precipitation buffer, made up from 60 ml of 5M K-acetate, 11.5 ml glacial acetic acid and 28.5 ml H_2O.

The ways in which the cells were incubated with each of these solutions in turn varied according to each protocol, and are described below.

2.9.2 Small preparations of plasmids- 'Mini Preps'

2.9.2.2 Preparing plasmid DNA using a 'Mini Prep' protocol

This method was adapted from Sambrook *et al.* (1989). Only 1.5 ml of an overnight cell culture were needed for this method. The required volume of culture was pipetted into an eppendorf and centrifuged at 5000 rpm for 5 min.

The bacteria were resuspended in 100 µl solution 1 and left for 5 min at room temperature. Solution 2 (200 µl) was added and mixed by pipetting the solutions up and down gently with a sterile, wide nozzle tip. It was then left on ice for 5 min, before 150 µl solution 3 was added and mixed as before. After 5 min on ice, the precipitate was pelleted by centrifugation at 12000 rpm in a bench top centrifuge, for 10 min.

Equal amounts of phenol and chloroform (230 µl) were added to the

supernate in a new eppendorf tube. The phases were mixed by inverting the tube a few times, and then separated again by centrifuging at 12000 rpm for 10 min (see phenol extraction, 2.9.6.2). The aqueous phase was pipetted into another sterile eppendorf, and 1 ml 100% ethanol was added and mixed. This was then incubated at -70°C for 15 min before pelleting the precipitated DNA by centrifugation in a benchtop centrifuge at 12000 rpm for 10 min. The supernate was removed with a pipette, and the pellet gently washed with 70% (v/v) ethanol. If the pellet was dislodged at this stage, the tube could be recentrifuged; otherwise, the 70% ethanol could be removed without re-pelleting the DNA. The pellet was dried, either by air drying or *in vacuo*, and resuspended in 40 µl TE buffer (pH 7.4- 8.0).

2.9.2.1 Preparing plasmid DNA using a 'Mini Prep' kit

The preferred kit was the Wizard™ *Plus* Mini Preps DNA purification system. A bacterial culture (1-3 ml), taken from 5 ml overnight culture, was used for each preparation. The cells were pelleted at high speed in a benchtop centrifuge for 2 min, the supernate was discarded, and the pellet resuspended into 200 µl Cell Resuspension Solution. After lysis with 200 µl Cell Lysis Solution, the proteins and genomic DNA were precipitated with 200 µl Neutralisation Solution. The lysate was centrifuged for 5 min in a benchtop centrifuge at high speed. The supernate was carefully removed, placed in a fresh eppendorf tube, and 1 ml resin was added to the cleared lysate. This solution was then pipetted into a syringe barrel, which was attached to a minicolumn. The solution was filtered through the minicolumn by either applying a vacuum to the column to draw the solution through, or by pushing the solution through with a syringe plunger. Resin was washed by pushing through 2 ml Column Wash Solution in the same manner. The resin was dried by centrifuging the column placed in a microfuge tube in a benchtop

centrifuge for 2 min at high speed. The DNA was finally eluted from the resin by applying 50 µl dH_2O or TE buffer to the column, allowing the DNA 1 min for solubilisation, then centrifuging the column in a fresh eppendorf to recover the DNA in the water or TE buffer.

2.9.3 Large preparations of plasmids- 'Maxi Preps'

2.9.3.1 Preparing plasmid DNA using a 'Maxi Prep' protocol

This method was adapted from Sambrook *et al.* (1989). Overnight bacterial culture (200 to 500 ml) was centrifuged at 6000 rpm in a JA-14 rotor in a Beckman J2-21 centrifuge at 4°C, for 15 min. The pellet was resuspended in 10 ml solution 1, lysed with a further 10 ml solution 2 and then neutralised by addition of 7.5 ml solution 3. The lysate was then centrifuged at 10000rpm in a JA-14 rotor, for 15 min at 4°C.

The supernate was carefully removed, the DNA and RNA was precipitated with 1 vol isopropanol, incubated for 15 min, and then pelleted by centrifuging for 15 min at 13000 rpm at 4°C in a JA-14 rotor. The DNA pellet was dried and redissolved in 2 ml TE buffer, and 2 ml solution 4 (5 M LiCl, 50 mM Tris-HCl, pH 8.0) added. This was incubated on ice for 10 min to precipitate the RNA only, which was pelleted at 13000 rpm in a Corex tube in a JA-17 rotor, in the Beckman J2-21 centrifuge for 15 min at 4°C. The DNA was precipitated from the supernate by addition of 2 vol absolute ethanol, incubated for 15 min on ice, then centrifuged again at 13000 rpm for 15 min as previously. The ethanol was removed, the pellet dried, and the DNA then resuspended in 500 µl TE buffer, and pipetted into an eppendorf tube. To this was added 30 µl RNase (10mg/ml filter sterilised solution of RNase), prior to incubation at 37°C for 1 h.

To remove the RNase, a phenol:chloroform extraction was carried out (see 2.9.6.2), using 2 washes of phenol:chloroform, and a final wash with

chloroform. The DNA was precipitated with 0.5 vol 3 M sodium acetate (pH 5.4) and 2 vol 100% ethanol. This was incubated at -70°C for 10 min before being centrifuged at high speed in a benchtop centrifuge at 4°C. The supernate was carefully removed, and the pellet washed with 70% (v/v) ethanol. DNA was repelleted by centrifuging for 10 min at high speed as before, the supernate was removed, and the pellet dried for 5-10 min *in vacuo* to remove all traces of ethanol. The DNA was redissolved in 100-500 µl TE buffer, depending on expected yields and preferred concentration.

2.9.3.2 Preparing plasmid DNA using a 'Maxi Prep' kit

The preferred kit was the Wizard™ *Plus* Maxi Prep DNA purification system. Cells from an overnight bacterial culture (500 ml), were pelleted at 5000 xg for 10 min in a centrifuge at 22°C; the supernate was discarded, and the pellet resuspended in 15 ml Cell Resuspension Solution and lysed with 15 ml Cell Lysis Solution. Cell lysis was monitored by ensuring that the solution became clear and viscous before proceeding. Proteins and genomic DNA were then precipitated with 15 ml Neutralization Solution. The lysate was centrifuged for 15 min at 14000 xg at 22°C. The supernate was carefully removed, and filtered through a sheet of muslin, 0.5 vol isopropanol added, and the DNA pelleted at 14000 xg at 22°C in a JA-14 rotor in a Beckman J2-21 centrifuge. The supernate was carefully discarded and the pellet resuspended in 2 ml TE buffer. Resin (10 ml) was added to the DNA, and the solution was pipetted into a maxicolumn. The solution was filtered through the maxicolumn by applying a vacuum to the base of the column. The resin was then washed with 25 ml Column Wash Solution in the same manner. The resin was dried by maintaining the vacuum for 1 min, centrifuging the column in a 50 ml microfuge tube in a benchtop centrifuge for 5 min at 1300 xg, and finally drying by replacing the vacuum for another 5 min. The DNA was

eluted from the resin and the column by applying 1.5 ml dH_2O or TE buffer (at 65°C) to the column, allowing the DNA 1 min to solubilise, and then centrifuging the column in a fresh 50 ml tube to recover the DNA in dH_2O or TE buffer. To ensure that resin fines were not present, the DNA was filtered through a sterile Nalgene filter, included with the kit.

The other kits used were the Qiagen Midi and Maxi Prep kits, expected to yield 100 µg and 500 µg DNA respectively. The kits were not used frequently since they were found to be variable and unreliable. The protocol was as follows for both kits;

The bacterial pellet from a 50 ml (500 ml for Maxi) culture was resuspended in 4 ml (10 ml for Maxi) buffer P1, then lysed with 4 ml (10 ml) buffer P2, and precipitated with 4 ml (10 ml) buffer P3, and kept on ice for 15 min (20 min). The precipitated material was pelleted at 3000 xg and the supernatant retained. The supernate was poured down a pre-equilibrated Qiagen column containing a resin which bound the DNA. The column was washed with 2x10 ml (or 2x30 ml for Maxi) of buffer QC, and the DNA then eluted with buffer QF (5 ml for Midi Prep, 15 ml for Maxi). The DNA was precipitated by adding 0.7 ml of isopropanol to the elute, and centrifuged at 15000 xg for 30 min at 4°C. The DNA pellet was washed with 70% (v/v) ethanol, dried, and then redissolved in TE buffer.

2.9.3.3 Preparing plasmid DNA using caesium chloride density centrifugation

This method was adapted from Sambrook *et al.* (1989). Cultures of cells were grown overnight and lysed by the Alkaline Lysis Method as described above in the 'Maxi Prep' protocol. After centrifugation of the lysate to pellet the precipitated carbohydrate and protein, the supernate was taken and all polynucleotides precipitated by adding 2.5 vol of isopropanol and incubating

on ice for 1 h. The various polynucleotides were then pelleted by centrifuging at 14000 xg (as described above) for 30 min at room temperature. The supernate was discarded, and the pellet air dried.

This pellet was then resuspended gently in 2 ml TE buffer at pH 7.8. Into this was dissolved 2 g caesium chloride, and then 200 µl ethidium bromide (10 mg/ml) added. This was carefully pipetted into a puncturable ultracentrifuge tube (Beckman Polyallomer Quickseal^{TM} tubes) by means of a syringe and needle, and the tube filled with mineral oil before balancing tubes to within 50 mg of each other, and sealing. Tubes were then spun at 45000 rpm for 24 h in a TLN 100 rotor in a Beckman benchtop ultracentrifuge. The tubes were carefully removed, and the tube then clamped to a retort stand. A needle punctured the top of the tube to allow air into the vessel, and then the plasmid was taken out by means of a syringe and needle puncturing the side of the tube to remove the relevant band (Fig. 2.1) .

The ethidium bromide and caesium chloride were both removed from the DNA. To extract the ethidium bromide, washes with a butanol solution were carried out. The butanol solution was made by saturating butan-1-ol with sterile distilled water first, and then saturating the water with caesium chloride. An equal volume of the top, butanol, layer was added to the DNA to wash out the ethidium bromide. The solutions were gently mixed, and the bottom, aqueous phase, was retained. This was repeated until the pink colour could no longer be seen in the butanol layer, and then a final wash was performed.

Ice cold ethanol (2.5 vol) was added, the sample was incubated on ice for a further 1h, and then centrifuged at 12500 rpm for 30 min to obtain the DNA pellet. The DNA pellet was resuspended in 20 ml of ice cold 70% (v/v) ethanol, and recentrifuged as before. This procedure was repeated until the pellet turned from a hard, salty pellet, into a fluffier, salt-free pellet (usually 2 or 3 ethanol washes). After this, the DNA was air or vacuum dried,

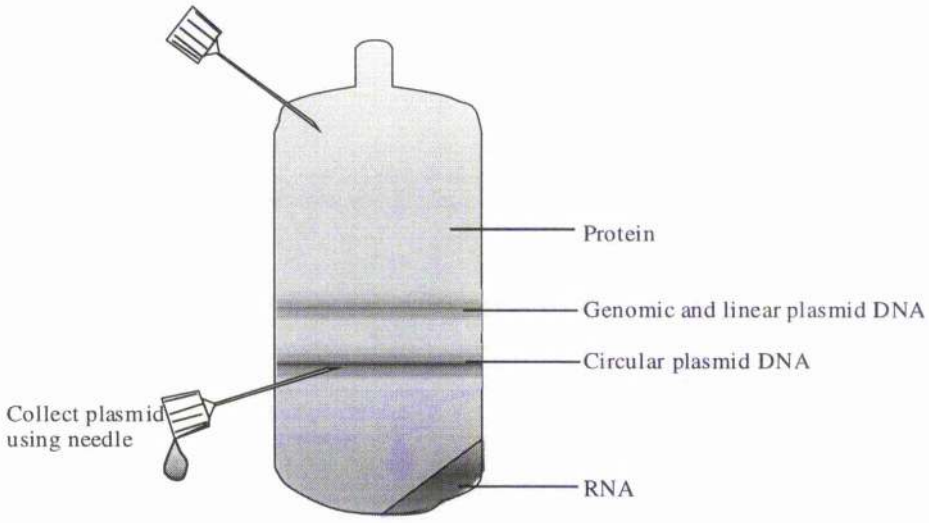


Figure 2.1 Collecting plasmid DNA from CsCl/ethidium bromide centrifugation

Needles are used to puncture the polyallomer tube containing the DNA at two points. the first needle allows air to enter from the top. The second needle allows the DNA to flow out, where it is collected for the next stage in the purification.

resuspended in 1 ml TE buffer pH 7.8, and the spectrum and absorbance readings taken for the solution for determining purity and quantity of DNA, as described below.

2.9.4 Extraction of linear DNA from agarose gels

Four different methods for the extraction of DNA from gels were used. The original method to extract DNA from gels was the Qiaex DNA Gel Extraction kit. Later, however, 3 methods were adopted for individual applications. To isolate small fragments of DNA (90 bp or less), the MERmaid® kit was used, to extract larger fragments, the Wizard™ PCR prep was used. The last method, electroelution, was used to isolate a large amount of DNA from a large piece of agarose. In all applications, the band of interest was located in the gel by UV light and working as quickly as possible. To minimise damage to the DNA, the band was excised using a scalpel.

2.9.4.1 Isolation of DNA from agarose gels with the Qiaex kit

The Qiaex kit is similar in methodology to the other two kits, utilising a bead technology to bind DNA whilst solubilising the agarose. The beads could then be pelleted with the DNA, leaving the agarose and other contaminants in solution.

After excising the DNA fragment from the rest of the gel, 300 µl buffer QX1 was added for every 100 mg of gel, followed by 10 µl Qiaex. The solution was incubated for 10 min at 50°C, then centrifuged in a benchtop centrifuge at high speed for 30 s to pellet the DNA/Qiaex. The pellet was washed twice in 500 µl buffer QX2, repelleting the DNA by centrifuging as before each time. In the same manner, the pellet was washed twice with 500 µl buffer QX3. The pellet was air dried to remove all traces of ethanol, and

finally eluted from the Qiaex in 20 µl TE by resuspending the Qiaex in the TE, pelleting the Qiaex as before and removing the supernate, which contained the DNA.

2.9.4.2 Isolation of small DNA fragments with the MERmaid® kit

The MERmaid® kit is optimised for the isolation of 10-200 bp of single or double stranded DNA fragments from gels, using a silica bead technology to bind the DNA in solution and pellet it by centrifugation.

Using the agarose and running buffer included with the kit, a 3% (w/v) gel was used to separate fragments, electrophoresing at high voltage (16 volts/cm) for 15 min to minimise diffusion of DNA. The DNA was visualised under UV, and the band of interest excised. High Salt Binding Solution (3 vol) was added to the gel slice, and 8 µl Glassfog (DNA binding solution) added for every 1µg DNA expected to be present. The solution was vortexed continuously for 10 min and DNA was pelleted at high speed in a benchtop centrifuge for 1 or 2 sec. The supernate was discarded and the pellet resuspended in 300 µl Ethanol Wash by vortexing for a few seconds. The DNA was repelleted by centrifuging as previously, and the ethanol wash repeated two more times. The last time the DNA was pelleted, care was taken to remove all the ethanol, and the pellet was dried under vacuum for 5 min.

The DNA was then eluted from the Glassfog by resuspending the pellet in 2 vol dH_2O, and heating to 45°C for 5 min. The Glassfog was pelleted by centrifuging at high speed for 1 min, and the DNA, now in solution, was removed from the Glassfog pellet with the supernate, and transferred to a new tube. The elution step was repeated, and the two elutions combined.

2.9.4.3 Isolation of large DNA fragments with the Wizard™ PCR prep kit

The Wizard™ kit is optimised for the isolation of 200-3000bp of double stranded DNA from gels, using a resin to bind the DNA in solution. The resin beads are designed to be too large to pass through a filter included with the kit, whereas the DNA being eluted is not.

Using standard protocols, either a normal or a low melting point agarose gel, in TAE buffer, was used to separate fragments. The DNA was visualised under UV, and the band of interest excised. The gel was estimated for volume, and for every 300 µl of gel, a separate preparation was carried out. The gel slice was incubated at 65°C until melted (approx. 5 min), and 1 ml resin was added. The mixture was gently pushed, using a syringe, or sucked, using a vacuum, into a Wizard™ minicolumn and 80% (v/v) isopropanol (2 ml) was pushed through the column to wash the DNA. The minicolumn was centrifuged for 20 sec at high speed in an benchtop centrifuge to remove all traces of the isopropanol. The DNA was then eluted from the column and the resin by adding 50 µl dH_2O or TE buffer to the column, and leaving for 1 min before centrifuging the column in a fresh eppendorf tube, as previously. The column was disposed of, and the DNA solution in the tube retained and checked for purity and concentration.

2.9.4.4 Isolation of DNA from gels using electroelution

This method was taken from Sambrook *et al.* (1989). DNA was electrophoresed through a TAE-agarose gel, as described in 2.11.2.1. The DNA was visualised by using 0.5µg ethidium bromide per ml of agarose in the gel. The band of interest was located and the band was excised. The gel slice was then placed in a dialysis bag filled with 1xTAE buffer (the dialysis tubing

was previously pre-treated by boiling for 1h in 2%(w/v) $NaHCO_3$, 1 mM EDTA. Dialysis tubing was stored in ethanol and rinsed in dH_2O before use). Most of the buffer was removed, leaving just enough to surround the gel slice. The bag was sealed and placed in an electrophoresis tank filled with TAE buffer. The dialysis bag was then electrophoresed for 2-3h at 4-5volts/cm. After this, the current was reversed for 1 min to force the DNA back from the wall of the dialysis bag and into the buffer. After opening the bag, the buffer was carefully removed, leaving the agarose behind. The DNA was isolated by extracting several times with butan-2-ol, another 2 times with phenol:chloroform (see 2.9.6.2), and then precipitated with 0.2 vol 10M ammonium acetate and 2 vol absolute ethanol at 4°C. The DNA was incubated 10 min at room temperature, and centrifuged at high speed for 20 min in a benchtop centrifuge. The pellet of DNA was then resuspended in 200 µl TE buffer, and reprecipitated with 25 µl 3M sodium acetate and 2 vol absolute ethanol at 4°C, and centrifuged as previously. This pellet was retained and dried under vacuum before resuspending in a small amount of TE buffer.

2.9.5 λgt10 cDNA purification

These methods are taken from Sambrook *et al.* (1989).

2.9.5.1 λgt10 cDNA library amplification

An overnight culture of *E. coli* C600 (1 ml), grown in LB broth, was added to 500 ml LB medium, pre-warmed to 37°C, and incubated, with shaking at 37°C for 3-4h. To this, 10^{10} plaque forming units of λgt10 phage were added, and incubated further for 3-5h, or until bacteria appeared to be lysed (overnight if necessary).

2.9.5.2 λgt10 phage purification

Chloroform (10 ml) was added to the culture described in the previous section and incubated a further 10 min. The lysed culture was then cooled to room temperature and left for 30 min after adding 500µg RNase and DNase. NaCl was added to give a final concentration of 1M. The solution was slowly stirred for 1h on ice, before centrifuging at 11000 xg for 10 min at 4°C, using a JA-14 rotor in a Beckman J2-21 centrifuge. The supernate was retained, and PEG 8000 was added as a solid to a final concentration of 10% (w/v). This solution was slowly stirred on ice for 1 h, and again centrifuged at 11000 xg for 10 min at 4°C. The pellet was retained, and the fluid drained from it completely. The pellet was then gently resuspended in 8 ml SM medium (5.8g NaCl, 2g $MgSO_4.7H_2O$, 50 ml 1M Tris/HCl, pH 7.5, 5 ml 2% gelatin solution, H_2O to 1l, sterilised by autoclaving), and mixed with an equal volume of chloroform. This was centrifuged in glass Corex tubes at 3000 xg for 15 min at 4°C in a JA-17 rotor using a Beckman J2-21 centrifuge. The aqueous phase was removed with a pipette and retained.

2.9.5.3 Purifying DNA from λgt10 phage

0.5M EDTA (320 µl) and 400µg of proteinase K were added to the aqueous phase obtained in section 2.9.5.2, and mixed; 400 µl 10% (w/v) SDS was then added, and mixed. After incubation at 56°C for 1h, the solution was cooled to room temperature. An equal volume of Tris-phenol was added, mixed by inverting the tube and allowed to settle. The aqueous phase was removed, and an equal volume of phenol/chloroform was added, mixed in, and allowed to separate as before. The aqueous phase was dialysed overnight in 1000 vol TE buffer at a pH between 7.4 and 8.0. The dialysed DNA was stored at -20°C until required.

2.9.6.1 Decontamination of DNA

DNA which was found, through spectrophotometry or gel electrophoresis, to be contaminated by solvents, salts or protein, could be purified by precipitating the DNA. If the DNA was in an inappropriate buffer, containing too much salt from a previous precipitation, or contaminated with ethanol from a precipitation, then the DNA could be re-precipitated, as described below. If, however, the DNA was contaminated with protein from either a plasmid preparation, a DNA modifying enzyme, or if an oil overlay from a PCR reaction was present, it was necessary to include a phenol/chloroform extraction before proceeding to the DNA precipitation.

2.9.6.2 Phenol/chloroform extraction

This protocol was adapted from Sambrook *et al.* (1989). The DNA solution was mixed with an equal volume of Tris-phenol, and then spun in eppendorf tubes at 14000 rpm for 1 min in a bench top centrifuge. The aqueous phase was kept and pipetted into a new eppendorf tube, and this phenol extraction step was repeated. A final extraction with an equal volume of chloroform:isoamyl alcohol (24:1) followed, the solutions were mixed, centrifuged for 1 min at 14000 rpm in a bench top centrifuge, and the aqueous phase retained for DNA precipitation.

An alternative method was to use one or two phenol/chloroform washes (phenol:chloroform:isoamyl alcohol, 24:24:1) as above, followed by a final chloroform:isoamyl alcohol wash, as above.

2.9.6.3 DNA precipitation

0.1 vol 3 M sodium acetate (pH 5.4) was added, and then 2.5 vol ice cold absolute ethanol prior to mixing by pipetting the solution up and down. The DNA solution was then incubated either at -70°C for 15 min or -20°C for 1 h. After the incubation, the solution was centrifuged at 14000 rpm for 30 min at 4°C in a refrigerated benchtop centrifuge. The supernate was carefully removed and replaced with half the original volume of cold 70% (v/v) ethanol. Unless the pellet was very small, and unlikely to contain much salt, the pellet was resuspended in the ethanol. In either case, this was then re-centrifuged as before. The supernate was removed and the pellet dried in air or *in vacuo*, and finally redissolved in TE buffer.

Occasionally 4 M ammonium acetate was used in place of sodium acetate, such as in the pCRScript protocol (2.10.4.2), to precipitate large DNA fragments, leaving very small DNA molecules and dNTPs in the supernate following centrifugation.

2.10 DNA manipulation

2.10.1 Polymerase chain reaction

The polymerase chain reaction is a powerful technique which allows the *in vitro* amplification of DNA, and can be used to detect the presence of specific sequences, to clone genes from a population of DNA molecules and to create mutated DNA (Arnheim and Erlich, 1992).

During the course of the research all the above uses of PCR were employed. It was found that different PCR reactions required different reaction conditions, so for each of the desired products the necessary reaction conditions had to be optimised and modified from the more general protocol below.

'Primers', oligonucleotides based on a known DNA sequence, were synthesised either by Pharmacia, by Dr. Veer Math at the Division of Biochemistry and Molecular Biology (IBLS), Glasgow University, or latterly by Gibco BRL, Inchinnan Business Park, Paisley. In order to amplify DNA, two primers were required- one which would anneal to one 3' end of the DNA sequence of interest, and another to bind to the other end of the sequence on the other strand of DNA. Primers also frequently incorporated restriction sites at their 5' ends, so that the subsequent reaction product, as well as having the DNA sequence of the original DNA, carried extra DNA at either end that allowed easy cloning of the DNA into restriction sites present in other DNA sequences, such as plasmids.

To carry out a reaction, the following components were added to a 0.5 ml eppendorf tube: 0.1-1 µg each primer, 10-500 ng template DNA, 5 µl 10x reaction buffer, 1 µl nucleotide mix ('dNTPs'- 25 mM dATP, 25 mM dCTP, 25 mM dGTP, 25 mM dTTP) and 0.5 µl *Taq* DNA polymerase or 1 µl $Vent_R$® (exo-) DNA polymerase.

If Taq polymerase was used then 3 µl of 25 mM $MgCl_2$ was added, but if $Vent_R$® (exo-) was used then no $MgCl_2$ was added. The volume was made up to 50 µl with sterile distilled water. The contents were mixed by tapping the tube and centrifuging momentarily in a bench top microfuge. If required, 40 µl mineral oil was layered over the reaction.

The tubes were then placed in a Techne PHC-3 Thermal Cycler or a Genetic Research Instruments PTC 100 Thermal Cycler, where they were heated and cooled to allow denaturation of all the double stranded DNA, annealing of primers to the DNA and then extension of the primers by the polymerase to give the required DNA sequence.

Since different methodologies were used, different cycling conditions were employed. However, cycling conditions were usually based on the following scheme:

Initial denaturation

95°C

4 min

30 cycles of the following with a 1°C per second ramp rate

1) Denaturation 95°C

1 min

2) Reannealing 40-60°C 1

min

3) Extension 72°C

1 min

This was followed by a final extension step:

1) Final extension 72°C

5 min

2) Cooling 4°C

Products were then taken out and kept either at 4°C or -20°C until further use.

2.10.2 Restriction digestion of DNA

Restriction enzymes could be used to cut DNA so that it could be ligated to a different piece of DNA. It could also be used in diagnostic procedure; detecting changes in the DNA by examining the pattern of DNA digestion before and after a manipulation.

Restriction digestion of DNA was performed by diluting or dissolving DNA in an appropriate buffer which optimised salt concentration and pH for the enzyme being used. If DNA was pure and dissolved in TE buffer, then simple addition of a 10x buffer, usually included with the enzyme by the supplier, would provide the correct conditions. A final volume of 10 µl per 1 µg DNA was usually needed, and to this, 1-10 units of enzyme per 1 µg of

DNA were added. This was incubated for 1 h at a temperature that depended on the enzyme used (usually 37°C). These conditions were occasionally varied depending on circumstances.

To digest DNA with more than one enzyme, all enzymes could be used at the same time if they required similar conditions. However, if enzymes required different conditions, the DNA would be incubated with one according to its requirements, the DNA would be precipitated and redissolved in the new buffer, as described above, and digested with the second enzyme in its required buffer.

2.10.3 Dephosphorylation of DNA

This protocol was based on Sambrook *et al.* (1989). Purified cut DNA, in either water or TE buffer, was dephosphorylated at 5' protruding termini by incubating with calf intestinal phosphatase (CIP). For every 100 pmol of DNA (approx. 1.4 µg of a 5 kb plasmid) 1 unit of enzyme was added, along with an appropriate volume of the supplied CIP 10x buffer. This was incubated at 37°C for 30 min.

CIP was inactivated by adding EDTA solution (pH 8.0) to 5 mM and heating the solution to 74°C for 10 min.

2.10.4 Ligation of DNA

Fragments of DNA which were of interest could be ligated into a plasmid; once in the plasmid the fragment could be amplified, sequenced and manipulated more easily.

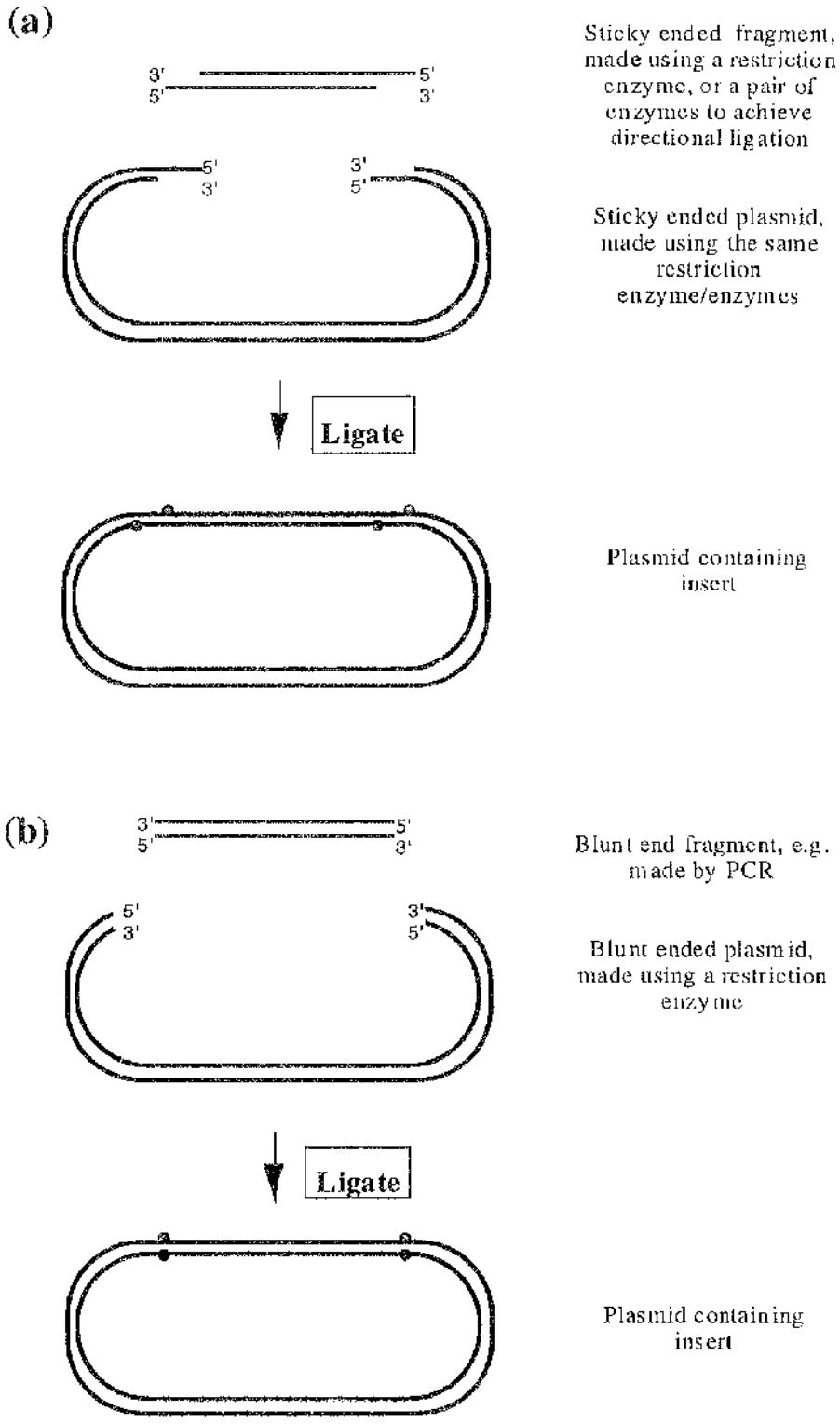
DNA could either be ligated into a plasmid 'sticky ended' or 'blunt ended'. Sticky ended ligation involved the ligation of a piece DNA, cut at both ends using restriction enzymes, to a plasmid which has also been cut with the

same enzymes. The choice of enzymes used previously dictated how this ligation was performed. Typically, the plasmids were cut so that they presented two cohesive termini cut by separate enzymes. Ligated into this would be a DNA fragment which also had cohesive termini, generated by the same enzymes, so that ligation of the DNA fragment into the plasmid could only take place in one direction, that is, when the correct cohesive termini lined up with each other (Fig. 2.2a).

Alternatively, blunt ended ligation could be used, where the fragment to be cloned and the vector into which the fragment was being ligated were both blunt ended, with no cohesive termini (Fig. 2.2b). Ligation of one to the other could still occur, but proceeds at a slower rate, and with a higher background of self ligations of the original vector. A kit used to clone PCR products utilised blunt end ligation combined with a restriction enzyme which would re-cut non-recombinants.

2.10.4.1 Sticky end ligation

This protocol was adapted from Sambrook *et al.* (1989). Vector DNA, cut with the appropriate enzymes, was mixed with similarly treated foreign DNA in equimolar quantities. Greater amounts of foreign DNA could be used if necessary. Typically 0.1 µg vector DNA was employed, and 1 µl 5 mM ATP, 1 µl T4 DNA ligase and 1 µl 10x buffer was added. The total volume was made to 10 µl with dH_2O, and left for 1-12 h, then used to transform bacteria.



Sticky ended fragment, made using a restriction enzyme, or a pair of enzymes to achieve directional ligation

Sticky ended plasmid, made using the same restriction enzyme/enzymes

Plasmid containing insert

Blunt end fragment, e.g. made by PCR

Blunt ended plasmid, made using a restriction enzyme

Plasmid containing insert

Figure 2.2 Blunt and sticky end ligation

The diagram at the top (Fig. 2.2a) represents the situation where a piece of DNA has "sticky ends" or "cohesive termini", i.e. a restriction enzyme has cut each strand at a different position, leaving a small 2-4 base overhang. This overhang can interact, through base-pairing, with another sticky end which has the complementary sequence. A plasmid cut with the same restriction enzymes, for instance, would bind in this way. The ligation proceeds quickly, since the interaction is specific.

The diagram below it (Fig. 2.2b) represents the situation where a piece of DNA with "blunt ends", i.e. 5' and 3' termini which end at the same point ligates to a plasmid with similar termini. Such is the case when a PCR product ligates into a plasmid cut with an enzyme that leaves blunt ends. This ligation proceeds slowly since it is non-specific.

2.10.4.2 The pCRScript™ cloning protocol (blunt end ligation)

The pCRScript™ cloning protocol was a kit and methodology devised and supplied by Stratagene. This procedure allows relatively easy cloning of PCR products into a pBluescript based vector, and subsequent transformation into *E. coli* XL1 Blue MRF' Kan. The protocol below was a slightly modified version of that given by Stratagene®, to give more transformants per plate, and optimised for the particular PCR reactions.

A PCR product could either be isolated and purified through agarose gel electrophoresis followed by a gel extraction, as described above, or, if the PCR product was pure enough, a selective precipitation was used to remove excess PCR primers. STE buffer (0.1 vol) was added to the PCR product, plus an equal volume of 4 M ammonium acetate. To this new solution was added 2.5 vol (approx. 5 vol original PCR product) of absolute ethanol, at room temperature, was added, mixed, and immediately centrifuged at 14000 rpm for 20 min in a bench top centrifuge. The pelleted DNA was then washed with 200 µl 70% (v/v) ethanol, and centrifuged at 14000 rpm for 10 min as before. After drying the DNA pellet *in vacuo*, it was finally resuspended in TE buffer. Though the protocol advised resuspending the pellet in the original volume (50-100 µl), it was found to be more helpful if the DNA was concentrated during this step by resuspending in 10-50 µl of TE buffer, since PCRs were often found to yield less DNA than was supposed by the manufacturers of the kit. The DNA was then stored at 4°C until further use.

If the PCR product was generated using Taq, then a 'polishing' step was required, since Taq polymerase can leave ends which are not blunt-ended, but have 3' overhangs (Hu, 1993). To a 0.5 ml eppendorf tube, 10 µl purified DNA was added, and then the following supplied reagents; 1 µl 10 mM dNTP mix (2.5 mM each), 1.3 µl 10 x polishing buffer, 1 µl cloned *Pfu* DNA

polymerase (0.5 units). This was mixed by pipetting up and down, a 20 µl overlay of mineral oil was added, and the reaction mix incubated at 72°C for 30 min. The DNA product was stored at 4°C until further use.

The amount of PCR product used in the cloning protocol was optimised to give insert end to vector end ratios of between 40:1 and 100:1. This was deduced using the calculation:

$$\text{pmol ends/mg of DNA} = \frac{(2 \times 10^6)}{(\text{number of base pairs} \times 660)}$$

Since 10 ng of vector was used in the cloning reaction, this entails using 36-90 ng of a 300 bp insert.

The following components which, with the exception of the insert DNA and distilled water, were supplied by the manufacturer, were added in order to a sterile 0.5 ml eppendorf tube; 1 µl pre-digested pCRScript vector (10 ng), 1 µl 10x reaction buffer, 0.5 µl 10 mM rATP, 2-4 µl PCR reaction/polishing reaction, 1 µl of *Srf1* restriction enzyme, 1 µl T4 DNA ligase and dH_2O to make the volume up to 10 µl. This was mixed by pipetting the solutions up and down, and then incubated for 1 h at room temperature. The reaction was stopped and enzymes inactivated by heating the reaction mixture to 65°C for 10 min; this was stored at 4°C until it was required for transforming the *E. coli* XL1 Blue MRF' Kan.

E. coli XL1 Blue MRF' Kan were supplied by the manufacturer of the kit as frozen 'supercompetent' cells, stored at -80°C. The cells were thawed on ice, and then aliquoted into 40 µl amounts in sterile 1.5 ml eppendorfs. To each aliquot was added and mixed 0.7 µl β-mercaptoethanol (a final concentration of 25 mM). Cell suspensions were incubated on ice for 10 min, with occasional swirling, until addition of 2 µl of the previous cloning reaction. This was then mixed by gently swirling the eppendorf, and left on ice for 30 min. The cells were heat pulsed for 45 sec at 42°C, put on ice for 2 min,

and 0.45 ml of SOC medium, preheated to 42°C, was pipetted into the cells. The transformation mix was incubated, with shaking, at 37°C for 1 h, before spinning down the cells at 6500 rpm for 1 min, resuspended in 100 µl SOC medium, and then plated onto LB ampicillin/X-gal/IPTG plates. These were incubated overnight at 37°C, and white colonies taken to investigate for the presence of cloned DNA inserts.

2.11 DNA detection and analysis

2.11.1 Spectrophotometry

This protocol was taken from Sambrook *et al.* (1989). DNA absorbs UV light with a maximum absorbance around 260 nm, whereas protein absorbs more at 280nm. To assess the purity of DNA, absorbance readings at 260 and 280nm could be taken using a Shimadzu UV-Visible scanning spectrophotometer. Routinely, 5 µl DNA sample was added to 995 µl TE buffer or water in a quartz cuvette, which had already been used to zero the absorbance readings on the spectrophotometer. The 280 nm and 260 nm readings were then measured. When the ratio of 260 nm / 280 nm was greater than 1.8, the DNA was assumed to be pure. The reading at 260 nm could be used to estimate the DNA concentration; an absorbance reading of 1.0 related to 50 µg/ml of double stranded DNA, 40 µg/ml of single stranded DNA and 20 µg/ml of dNTPs.

2.11.2 Gel electrophoresis

2.11.2.1 Agarose gel electrophoresis

This method was adapted from Sambrook *et al.* (1989). DNA could be separated and visualised by gel electrophoresis- a method which allowed an estimation of the size and concentration of DNA, and, if required, separation of a specific DNA fragment away from a population of DNA molecules. For this, various gel sizes and concentrations were used, depending on how many samples of DNA needed to be analysed, and the expected sizes of the DNA fragments. Volumes of agarose varied between 50 ml and 125 ml, and concentrations of agarose in buffer varied between 0.7% and 2% (w/v). The type of agarose and buffers used depended on the use of the gel; for visualisation of DNA only, the buffer used was TBE, with normal agarose (0.7-1.5% (w/v)), whereas gels which were used to isolate a DNA fragment from a population used low melting point agarose (1%-2% (w/v)) in TAE buffer.

To make a gel, a solution of melted agarose was prepared. The required volume of 1x buffer was poured into a conical flask, and to this was added powdered agarose to the desired concentration. This was boiled in a microwave oven to melt the agarose. The solution was allowed to cool slightly before being poured into a gel base. A well-forming comb was then inserted at one end, and the gel was allowed to set at room temperature. Once set, the tape which sealed either end was removed, the comb taken out, and the gel, with base, placed in an electrophoresis tank. The tank was filled with the same buffer as was used in the gel, to a level where the gel became submerged.

To each DNA sample was added 0.2 vol DNA gel loading buffer (from Promega) and mixed. The samples were then gently pipetted into the wells in the gel along with a DNA ladder of known DNA fragment sizes to compare to the unknowns of the samples. Elecrophoresis was carried out with the negative

electrode nearest the wells, at a constant voltage of 1-8 volts/cm between the two electrodes. The progress of the electrophoresis was monitored by the two dyes in the loading buffer with bromophenol blue travelling approx. as a 300 bp fragment would and xylene cyanol moving at approx. the same speed as a 4000 bp fragment.

DNA could be visualised by staining the gel with a 0.5 µg/ml solution of ethidium bromide in water for 30 min, followed by destaining for a further 30 min in water only, then viewing the gel on a UV transilluminator. The gel was photographed using either a camera or video imaging equipment. Calculation of sizes of DNA was carried out by comparing to DNA markers.

2.11.2.2 Acrylamide gel electrophoresis

This protocol was adapted from Sambrook *et al.* (1989). When an agarose gel was not precise enough to estimate sizes of DNA fragments, a non denaturing acrylamide gel could be used. The gel was of a similar type to the SDS-PAGE, substituting the buffer system and removing the SDS.

Gels, again with dimensions of 170 mm x 145 mm x 1.5 mm, were made from the following stock solutions; acrylamide (29% (w/v) acrylamide, 1% (w/v) N,N'-methlenebis acrylamide), 5xTBE (445 mM Tris-borate, 10 mM EDTA (pH 8.0), and a 10% (w/v) ammonium persulphate solution. The 5-12% gel was made with 16.7-40% acrylamide solution, 1xTBE (from a 5xTBE stock) and 0.7% (w/v) ammonium persulphate stock solution. Polymerisation was initiated with 0.035% (v/v) TEMED.

Gels were placed in vertical gel electrophoresis kits (BRL), and 1xTBE buffer poured into the electrode tanks. The samples, with an appropriate loading buffer, could then be applied to wells formed in the gel, and a voltage of 1-8V/cm applied. Electrophoresis was stopped when the DNA of interest was presumed to be in the middle of the gel, as judged by the dyes in the

loading buffer.

DNA was visualised in the same manner as agarose gels, namely by staining the gel with a 0.5 µg/ml solution of ethidium bromide in water for 30 min, followed by destaining for a further 30 min in water only. Again, the gel was viewed on a UV transilluminator, photographed using either a camera or video imaging equipment, and the DNA sizes calculated by comparison with the DNA markers, either roughly by eye or graphically.

2.11.2.3 Denaturing (sequencing) acrylamide gel electrophoresis

Two kinds of sequencing gel were used, one for automated sequencing and one for manual. The glass plates for both sequencing gel types had to be cleaned prior to use, firstly with a detergent and secondly with 70% (v/v) ethanol. The sequencing gel used for manual sequencing was a Bio-Rad Sequi-Gen sequencing cell.

For manual sequencing, the main concern in cleaning the plates was to prevent the formation of bubbles whilst the gel was poured. Soap or washing up liquid was used to clean the plates first, and an isopropanol wash to finish after an ethanol wash was found useful to clean the plates of dust, using a lint free paper towel. The spacers were placed between the plates, and the plates clamped together.

A sealing gel was made to seal the base of the assembly, prior to pouring the gel solution into the plates. The gel solution used for the main gel, Gel-Mix®6 (7 M urea, 100 mM Tris-borate (pH 8.3), 1.0 mM Na_2EDTA, 5.7% acrylamide, 0.3% (w/v) N,N'-methylene bisacrylamide), was also used to seal the base rapidly. To 50 ml gel solution was added 200 µl 25% (w/v) ammonium persulphate and 200 µl TEMED. This was poured into the sealing gel casting tray, and was absorbed into the filter paper lining the base of the

tray. The plate assembly was placed into the tray, on top of the filter paper, and the sealing gel rose 3-4 cm into the plates by capillary action, and set within 5 min. The casting tray was removed, leaving the sealing gel at the end of the plates. Into the top of the plate assembly was poured the main gel, made with 150 ml gel solution, 150 µl 25% (w/v) ammonium persulphate and 150 µl TEMED. The assembly was held at a slight angle to allow the gel solution to flow down. If any air bubbles were generated, a Promega 'Bubble getter' could pull the bubbles out. The comb or the well former for the sharktooth comb was then put in place at the top of the gel, and the top of the assembly was clamped together using bulldog clips. The gel was allowed to polymerise overnight.

A more stringent protocol was used to clean the plates of the automated sequencing gel for two reasons. The plates had to be dust free so there were no air bubbles, as with the manual sequencing. Secondly, however, the plates had to be optically clear for the laser scanning the gel for the labelled DNA (see 2.11.3); the laser was particularly sensitive to any dirt on the glass. The detergent wash was carried out with Alconox^{TM}, using plenty of lint free paper towels to eliminate dust. The 70% (v/v) ethanol wash also used many lint free towels. The spacers were placed between the plates, and the plates sealed together using Permacel^{TM} tape.

The automated sequencing required manufacturer's gel solutions to be used, and the method followed was that of the instructions on the bottle; 20 ml buffer reagent were mixed with 80 ml monomer solution for each 100 ml required, with 0.8 ml 10% (w/v) ammonium persulphate added to initiate polymerisation. The gel mix was poured into the gap between the two plates. Whilst pouring, the plates would be held at an angle to allow the gel to flow down. Air bubbles could be pulled out with the 'bubble getter'. Once the plates were filled with gel solution, the assembly was laid flat, and the well former for the sharktooth comb was slotted into the gel. The gel was allowed

to polymerise for 3-4 h.

2.11.3 Sequencing

Both methods used employed the Sanger dideoxynucleotide termination method for sequencing. This method uses *in vitro* synthesis of DNA based on the template DNA, terminated at specific points along the DNA molecule, depending on the sequence.

In manual sequencing, the DNA is radiolabelled. Four reactions are used, each one corresponding to DNA which is terminated by a different dideoxynucleotide, either ddATP, ddCTP, ddGTP or ddTTP. The reactions are electrophoresed through a gel, and separate according to size. The sequence is then read by exposing the gel to a photographic film.

In automated sequencing, there is only one reaction, and the 4 different dideoxynucleotides are in the same tube. The different bases are identified through fluorescent probes; each dideoxynucleotide species has a different fluorescent label attached. When the reaction is electrophoresed, a computer reads the sequence as it passes through the gel by exciting the fluorescent label with a laser.

2.11.3.1 Manual sequencing

The protocol followed was that of Amersham's Sequenase version 2.0. Single stranded DNA (approx. 1 µg) could be used directly in the sequencing reaction, however, double stranded DNA had to be denatured first.

To denature double stranded DNA, 3-5 µg of DNA was incubated at 37°C for 30 min with 0.1 vol 2 M NaOH, 2 mM EDTA. The DNA was then precipitated by 0.1 vol of 3 M sodium acetate (pH 5.5) and 3 vol ethanol. The DNA was pelleted by centrifuging in a benchtop centrifuge for 20 min at high

speed, the supernate discarded and the pellet washed with 70% (v/v) ethanol, prior to redissolving in 7 µl dH_2O.

The DNA, single stranded or double stranded, could then be used in a sequencing reaction. The protocol included with Sequenase Version 2.0 was adhered to strictly, using a thermal cycler to achieve the temperature variations.

Primer DNA was added to the template DNA, in a 1:1 molar stoichiometry, along with reaction buffer. The sample was heated to 65°C for 2 min, then the temperature was reduced to room temperature over a period of 30 min, and finally put on ice. Labelling mix, [^{35}S]dATP, polymerase and DTT were added to the annealed DNA, and incubated at room temperature for 5 min. This reaction mixture was added to 4 separate tubes, each containing one of the 4 dideoxynucleotides. The extending DNA was allowed to incorporate these nucleotides for 5 min before the reaction was stopped with the addition of a 'stop solution'.

2.11.3.2 Automated Sequencing

Here, the Taq DyeDeoxy™ Terminator Cycle Sequencing Kit protocol was followed. Single stranded DNA (500 ng) could be used, or 1 µg double stranded DNA.

Firstly, a reaction premix was made, which contained the buffer, the polymerase, each of the labelled ddNTPs and dNTP mix. This mixture was added to the template DNA, along with 0.8-3.2 pmol of primer. The reaction was then cycled through these temperatures:

1) 96°C for 30 sec

2) 50°C for 15 sec

3) 60°C for 4 min

This cycle was repeated 25 times, then chilled to 4°C. The reactions were purified by phenol:choroform extraction twice, followed by an ethanol/sodium acetate precipitation. The precipitated DNA was finally boiled in 4 µl formamide for 2 min, then kept on ice until loading onto the automated sequencer gel.

Chapter 3

Constructing a Novel Vector

3.1 Introduction

The E2 subunits from all the oxoacid dehydrogenase complexes possess unusually elongated mitochondrial matrix targeting presequences, varying in humans from 61 amino acids for BCDC-E2 (Lau *et al.*, 1992) to 73 amino acids for PDC-E2 (Patel *et al.*, 1994). Most human matrix-targeted polypeptides utilise a targeting sequence comprising only 19-40 amino acids, and as few as 12 amino acids have been shown experimentally to target a polypeptide (Hurt *et al.*, 1985). Other extended mitochondrial targeting presequences are known to carry additional information, such as the cytochrome b_2 polypeptide which carries an N-terminal mitochondrial targeting signal followed by a inter-membrane space targeting signal activated once inside the mitochondrion, and the ATPase subunit 6 which has an extended presequence to aid in antifolding. Does the extended presequence of PDC-E2 also carry additional information?

To uncover the properties and physiological roles of a protein presequence it is common to fuse the presequence to the N-terminus of a 'reporter protein' and note the change in the behaviour of the reporter effected by the presequence. To achieve this a chimaeric gene construct is made using DNA coding for the presequence fused to a reporter gene, held within an expression vector. This vector can be introduced into suitable cells, where the chimaeric construct is then expressed by the host cells, and the effect of the protein presequence can be evaluated.

A commonly used reporter gene is that coding for the bacterial enzyme chloramphenicol acetyltransferase (CAT). It can be detected relatively easily by enzymatic assay and immunological approaches. It is also a prokaryotic enzyme and is not easily confused with any proteins naturally present within eukaryotic cells, the common subject of such work. As such, this enzyme is most commonly used for promoter/repressor analyses (see Bryans et *al.*, 1992

or Kamrud *et al.*, 1995 for an example).

In the research detailed here, where the presequence of human PDC-E2 is under investigation, the use of CAT as a reporter gene is particularly appropriate, since the behaviour of CAT and PDC-E2 are so similar. Both are acetyltransferases, possessing similar 3-D structures, as demonstrated by Mattevi *et al.* (1992). As mentioned previously, it is also thought that the presequence of PDC-E2 may be involved in preventing premature assembly of the E2 into trimers (McCartney *et al.*, 1997; Hunter and Lindsay, 1986) in the cytosol, but once cleaved allows PDC-E2 to form trimers before assembling into a 60-mer; here again CAT is useful, since it is only enzymatically active as a trimer (Fig. 1.5). CAT has also been shown to be an effective reporter for mitochondrial targeting studies in plants (Boutry *et al.*, 1987), and therefore stands as a tried and tested method for this work.

After expressing an E2 presequence-CAT construct in host cells, the role of the presequence in targeting and assembly can easily be deduced. Targeting can be monitored by cell fractionation, followed by detection of CAT in various subcellular fractions by immunological approaches. Assembly of CAT into enzymatically active trimers can be followed by quantifying the enzymatic activity of CAT.

3.1.1 Reasons for constructing a new vector

A number of methods have previously been used to make constructs for investigating presequences. Two common approaches involve linking the DNA coding for the presequence to a reporter gene. The original method employed DNA restriction fragments isolated from plasmids containing the relevant genes. The excised presequence fragment from one plasmid was ligated to the reporter gene excised from another plasmid, and the ligation product was subsequently cloned into an expression vector (see Hurt *et al.*,

1984).

The second, newer method, follows a similar approach, linking presequence DNA to a reporter gene by overlap extension by PCR, a method for joining DNA molecules exactly with no restriction sites or intervening DNA sequence (Horton *et al.*, 1989). Again, the chimaeric product is cloned into an expression vector to express the new gene.

Both approaches require starting the procedure afresh for each presequence construct. When making many presequence-reporter gene constructs, or a range of presequence mutants, such methodologies can be labour intensive. A simpler method would be to obtain a plasmid with the reporter gene, and ligate various presequence DNA species into a multiple cloning site at the beginning of the reporter gene. For each construct, one would use the same plasmid, but exchange the presequence DNA by excising the old sequence and ligating in the new.

However, expression vectors containing the CAT gene are not usually designed for presequence work, and do not contain any restriction sites near the start of the gene. Instead, most of the unique restriction sites are present as a result of plasmid construction, or specifically designed to be in the promoter region. The nearest available site found for insertion of the presequence DNA in any of the plasmids investigated was a *CpoI* site, 60 bp upstream of the initiator codon for CAT in pBLSV CAT. Inserting a presequence at this point would have caused the translation of a portion of promoter sequence and introduced 20 extra amino acids to the presequence, which would have compromised the validity of the proposed study by substantially altering the presequence.

It was therefore decided, for this research and for general targeting studies on N-terminal presequences, to construct a plasmid which had restriction sites at the start of the CAT gene. A strategy was devised to mutate a currently existing CAT expression vector which would generate a suitable

cloning site for presequence insertion. To minimise the amount of DNA between the presequence and the CAT open reading frame, it was decided to put the cloning site after the first coding ATG of the CAT gene, eliminating the incorporation and translation of promoter sequence. To allow directional cloning of sequences, 2 restriction sites would be required. The mutations generating these 2 new sites were also designed to minimise the changes to both the CAT protein and the presequence.

In summary, this new vector would be designed specifically for presequence analyses, using a simple digestion/ligation method to obtain constructs with minimal alteration to the presequence or the reporter gene, and utilising a reporter enzyme which can relay information about the effect a presequence has on expression, intracellular location and assembly.

3.1.2 Method for mutagenesis

A plasmid, named pBLSV CAT, was kindly provided by Mr. Grant Scotland, along with a restriction map and a full DNA sequence. This plasmid is a variant of pGS7, missing the 300 bp RD1 insert; the history of the plasmid can be found in Scotland & Houslay (1995) and Luckow & Shätz (1987). In brief, the plasmid is based on pBLCAT2, an expression vector for CAT constructed in pUC18. An SV40 promoter drives the expression of CAT in mammalian cells, with an SV40 termination signal after the CAT gene. An ampicillin resistance gene provides a selectable marker for cloning the plasmid in bacterial host cells (Fig. 3.1).

Since no restriction sites lay near the start of the CAT gene, it was decided to mutate the DNA sequence to give 2 unique restriction sites for directional cloning of DNA fragments directly behind the initiator codon of CAT. The restriction sites chosen were designed to result in minimal alteration to the amino acid sequence of CAT whilst allowing directional cloning of

presequence DNA. The two restriction enzymes chosen, and their corresponding sites, were *Nhe*I and *Hin*dIII, as shown in Fig. 3.2. This causes minimal alteration to the expressed polypeptide, but does require the mutation of 8 bases in the DNA.

To change the original DNA to the new version, it was decided to excise a length of DNA containing the region to be mutated, using restriction enzymes, and replace it with a corresponding DNA molecule with the desired alterations, as shown in Fig. 3.3. A and B were the only usable sites in the region. A and B were both sites for *Bsi*MI cleavage; site A also had a sequence suitable for *Cpo*I digestion which overlapped with the site for *Bsi*MI (Fig. 3.4).

To allow directional insertion of the mutated form of this DNA section, it was proposed to digest pBLSV CAT first with *Cpo*I. This would cut the vector, and inactivate the 'A' site of *Bsi*MI, making the 'B' site unique. To show that this could be done, 1 µg plasmid was digested for 1 h at 30°C with 40 units (2 µl) *Cpo*I (to guarantee digestion) in 10 µl CpoI buffer, followed by addition of 20 units (2 µl) of *Bsi*MI and 10 µl BsiMI buffer and incubated for 1h at 55°C. *Cpo*I was first shown to linearise the plasmid, cleaving at a single site. The double digest with *Cpo*I and *Bsi*MI was then compared to a single digestion with *Bsi*MI only, on an acrylamide gel (not shown). A difference in the size of the smaller fragments (the DNA to be discarded) made by each digest was seen, indicating that *Cpo*I had inactivated the A site of *Bsi*MI as predicted.

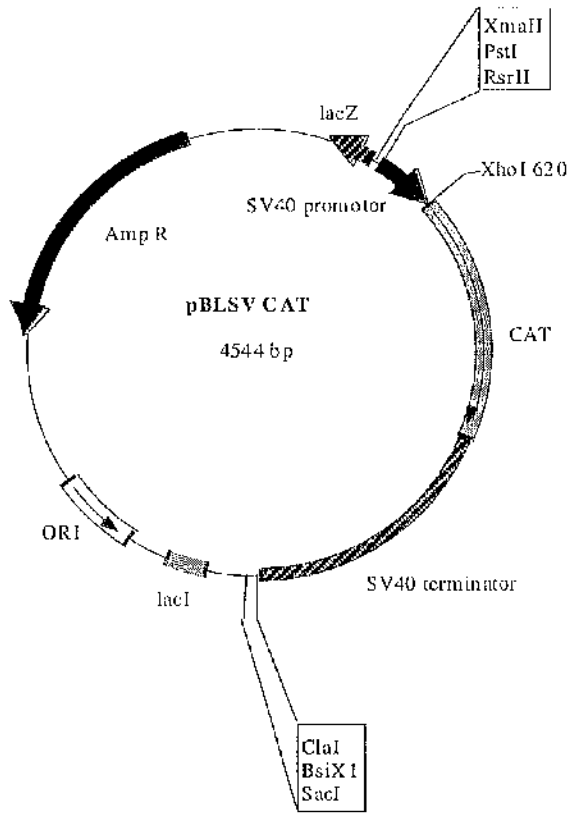


Figure 3.1 Diagram of pBLSV CAT

This plasmid diagram illustrates how the SV40/CAT construct has been ligated into the multiple cloning site of pUC18.

The origin of replication ('ORI') allows replication of the plasmid in host bacteria. The ampicillin resistance gene ('AmpR') allows selection of the plasmid in bacteria.

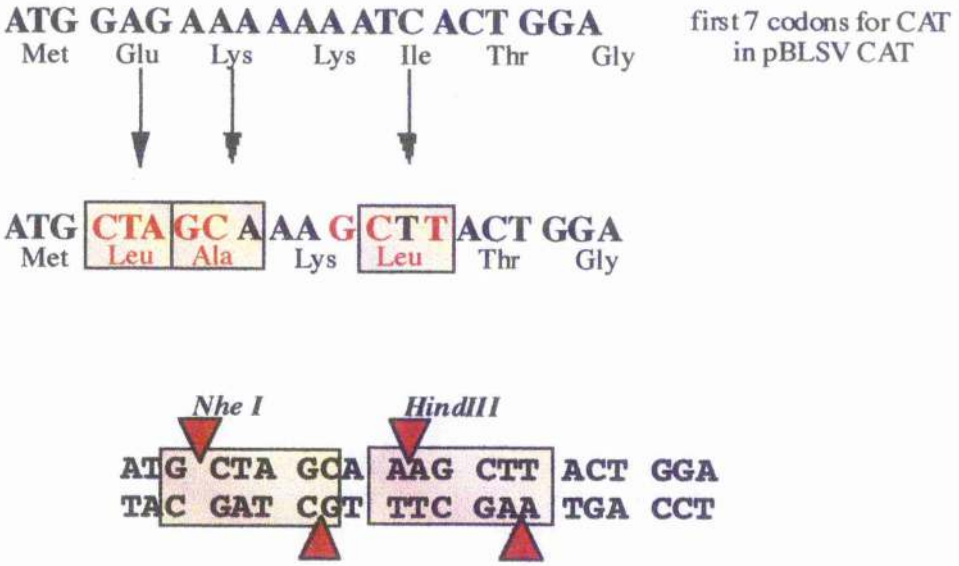


Figure 3.2 Mutations incorporated into pBLSV CAT to generate new restriction sites

The top diagram shows how many mutations must be made in order to obtain the 2 restriction sites shown in the bottom diagram. In all, 8 base changes are needed. This results in 3 amino acid substitutions, glutamine to leucine (a major change), lysine to alanine (a major change) and isoleucine to leucine, which is a conservative substitution.

Changes in the first few amino acids are expected to have little or no effect on the behaviour of the native enzyme since the N-terminal region has minimal structure, and lies outside the rest of the protein (Fig. 1.5).

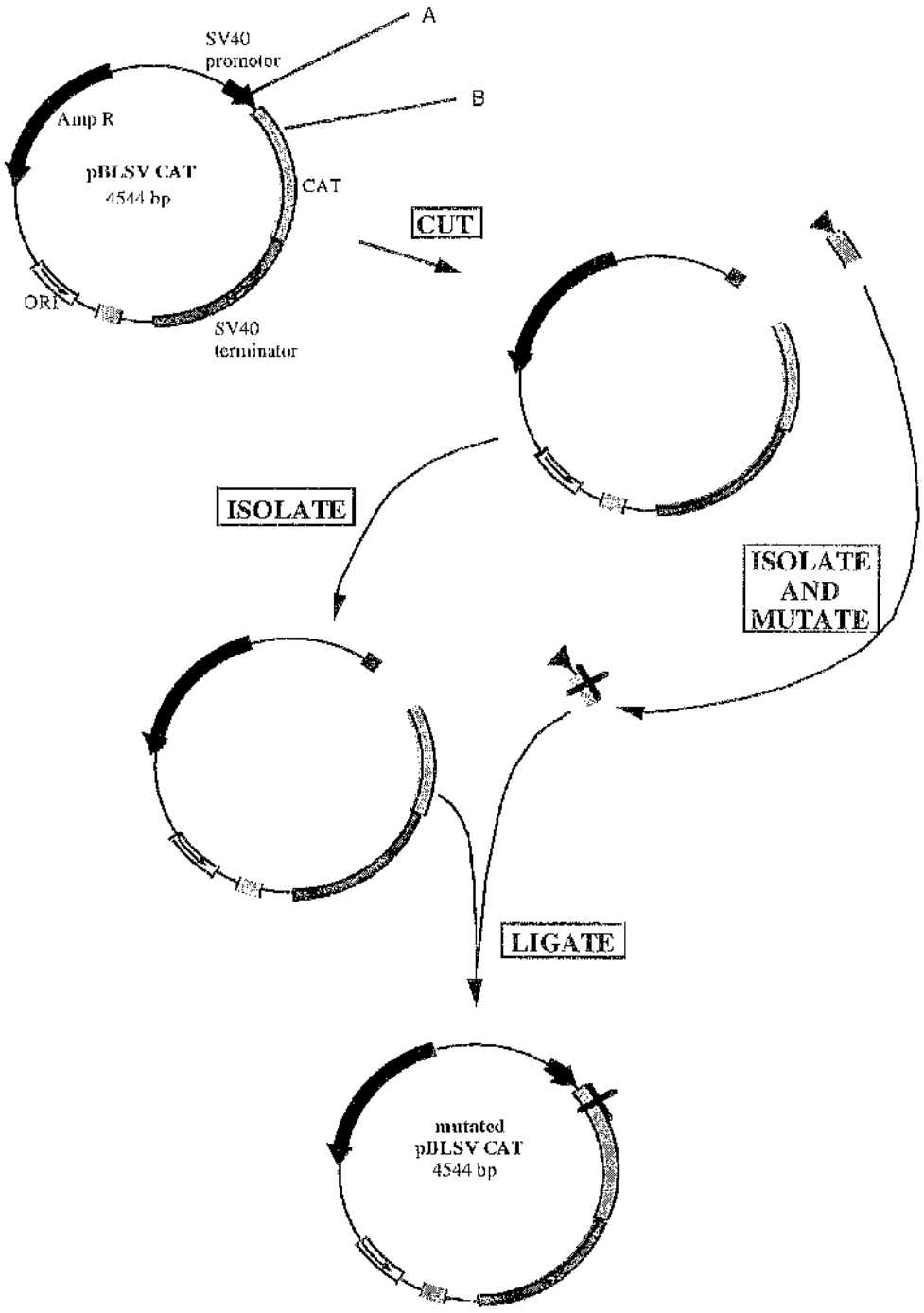

A
SV40
promotor
B
Amp R
pBLSV CAT
4544 bp
CAT
ORI
SV40
terminator
CUT
ISOLATE
ISOLATE
AND
MUTATE
LIGATE
mutated
pBLSV CAT
4544 bp

Figure 3.3 How to mutate the plasmid

Restriction sites near the point of mutation are chosen (A and B), and the plasmid digested with these enzymes. The treated plasmid is isolated from the fragment released by digestion. A mutated version of the fragment is made, and then ligated back into the digested plasmid. The plasmid is now restored to its original form, but with the mutations incorporated into the fragment.

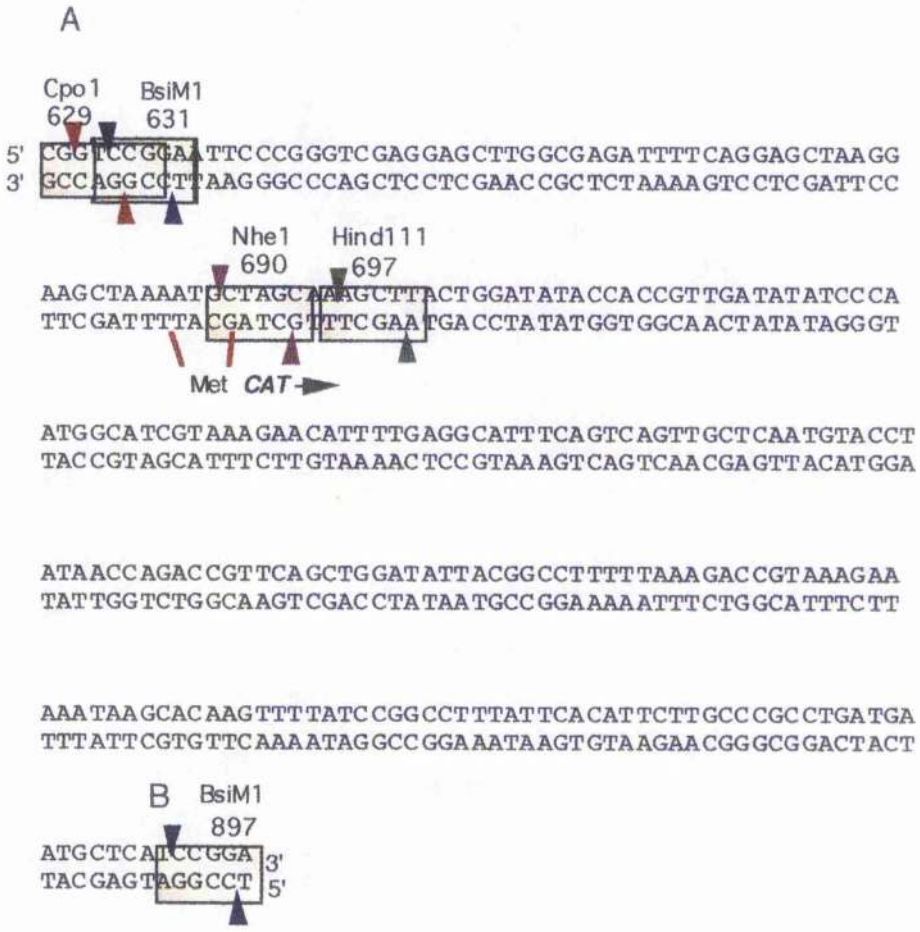


Fig 3.4 The region chosen in pBLSV CAT for mutation and replacement

The sequence above shows the positions of the restriction sites in the fragment removed from pBLSV CAT. The positions of the planned *Nhe*I and *Hin*dIII sites are shown.

It can be seen that *BsiMI* sites are at positions A and B, and a *Cpo*I site at A. Cleavage of pBLSV CAT by *Cpo*I linearises the plasmid, and inactivates the A site of *Bsi*MI cleavage by removing part of the recognition sequence. As the diagram shows, only two bases of the recognition sequence are left on the other strand after cleavage with *Cpo*I.

3.2 Site Directed Mutagenesis by Polymerase Chain Reaction

To generate the new insert, it was decided to utilise the polymerase chain reaction, or PCR, to synthesise a mutant version of the DNA fragment.

PCR is an *in vitro* enzymatic method for the amplification of specific sequences of DNA (see Bej *et al*, 1991; Arnheim & Erlich, 1992, for a review). Two oligonucleotides are designed to anneal to specific sequences of DNA flanking the region to be amplified, one anneals to the 3' end of one strand, and the other binds to the 3' end of the complementary strand. A DNA polymerase will then treat these oligonucleotides as primers and synthesise daughter strands complementary to the template DNA.

In PCR this process is repeated cyclically, denaturing the DNA at high temperature (at or above 94°C), annealing the primers at their melting temperature, then extending the primers at a temperature optimal for the polymerase. Since most polymerases are inactivated at the high temperatures required for denaturing DNA, thermostable polymerases, such as *Taq*, Vent$_R$® and *Pfu* polymerase are used. The optimal temperature for extension with these is between 72°C and 75°C. In repeated cycles, the synthesised daughter strands become the template for further synthesis, and the region of DNA between the two primers is therefore amplified exponentially.

This protocol can be used for site directed mutagenesis in two ways. The simplest way is to carry out a normal PCR with an altered primer (Fig. 3.5). Primers do not need to be exact complements to the target sequence DNA; there can be room for a small number of differences between them and the sequence they will bind to specifically. During PCR, a slightly altered primer can be incorporated into the daughter strands, giving a product mutated from the original.

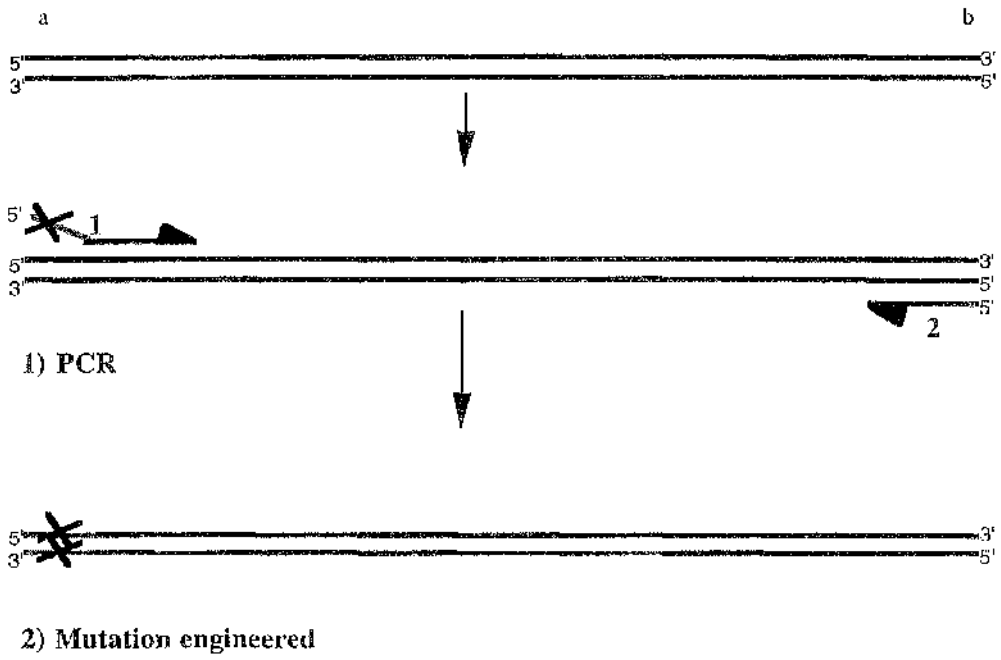


Figure 3.5 Site directed mutagenesis by PCR, using a one stage method

A PCR product is generated based on the template DNA (1). However, incorporated into one or both primers are mutations, shown in red with a cross. This leads to a product which is identical to the original template DNA, but with mutations where the primers diverged from the original sequence (2).

The second method uses a two stage PCR (Fig. 3.6), as described by Higuchi *et al.* (1988). Two PCRs are carried out on the same template DNA, one amplifies one section of the DNA, and the other amplifies the rest. The products of two PCRs overlap in the middle, in the region of the 'inner' primers. The inner primers both contain the mutations and complement each other. When both PCR products are mixed, denatured and allowed to anneal, some of the strands of one PCR product will anneal to strands of DNA generated by the other PCR, joining in the region of overlap. A third PCR to amplify the entire template sequence will use such annealed DNA as a template, and will exponentially amplify the joined product. The result is that the original DNA sequence has given rise to a product which is mutated in the middle of the sequence (Ho *et al.*, 1988).

As mentioned, this research requires a CAT expression vector with restriction sites at the start of the CAT gene. To do this, a region of DNA around the start of the gene is to be excised by *CpoI* and *BsiMI*, and replaced by a copy of the DNA possessing the two new restriction sites in the middle of the fragment. Since multiple mutations were required at these sites located in the middle of the fragment, the second, two stage mutagenic PCR was used in the course of the plasmid construction.

3.3 Design of Primers

The PCR primers were based on the sequence provided by Mr. Grant Scotland (Division of Biochemistry and Molecular Biology, IBLS, University of Glasgow), modified with mutations to give the desired restriction sites, as described above. The design is shown in Fig. 3.7. Primers 1 and 2 (22 and 36 bp respectively) generated the 'left hand' product, from a to b on Fig. 3.6. Incorporated into primer 2 was an extension based on the mutated sequence

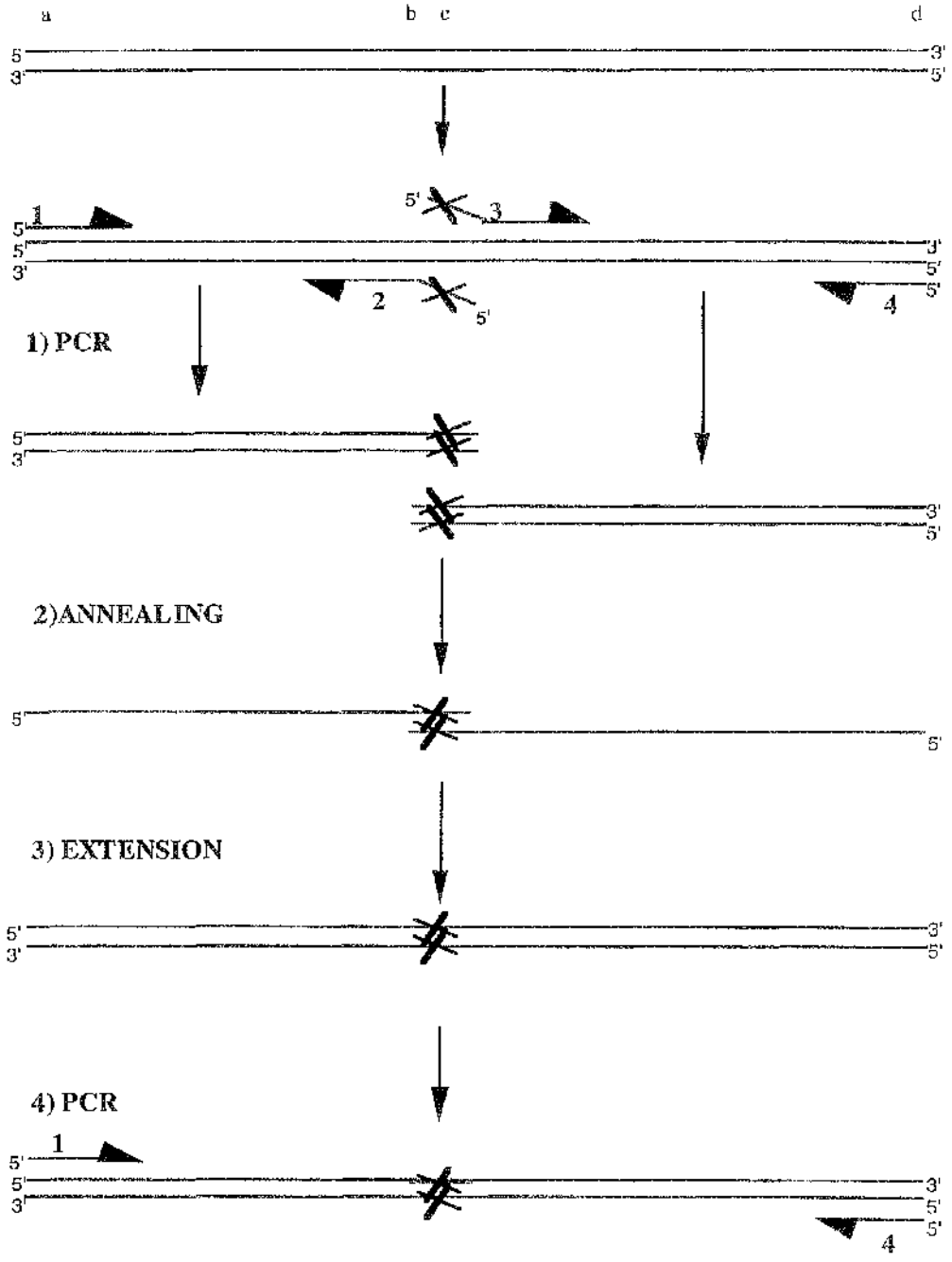

a
b c
d
5'
3'
3'
5'
1
5'
3
5'
5'
3'
3'
5'
2
5'
4
5'
1) PCR
5'
3'
3'
5'
2)ANNEALING
5'
5'
3) EXTENSION
5'
3'
3'
5'
4) PCR
1
5'
5'
3'
3'
5'
5'
4

Figure 3.6 Site directed mutagenesis by PCR, using a two stage method

In the first stage, 2 PCRs are made based on the template DNA (1) from a to b and from c to d. Incorporated into primers 2 and 3 are mutations, shown in red with a cross. This leads to two products which have identity at the region where the mutation was made (at b and c). The primers are designed so that this region of overlap will permit annealing (2). The annealed product is then extended to yield duplex DNA (3), using each template strand as a primer for the other strand. The DNA is finally re-amplified by PCR (4) using primers 1 and 4 to obtain a full length product similar to the original template, but possessing the mutations.

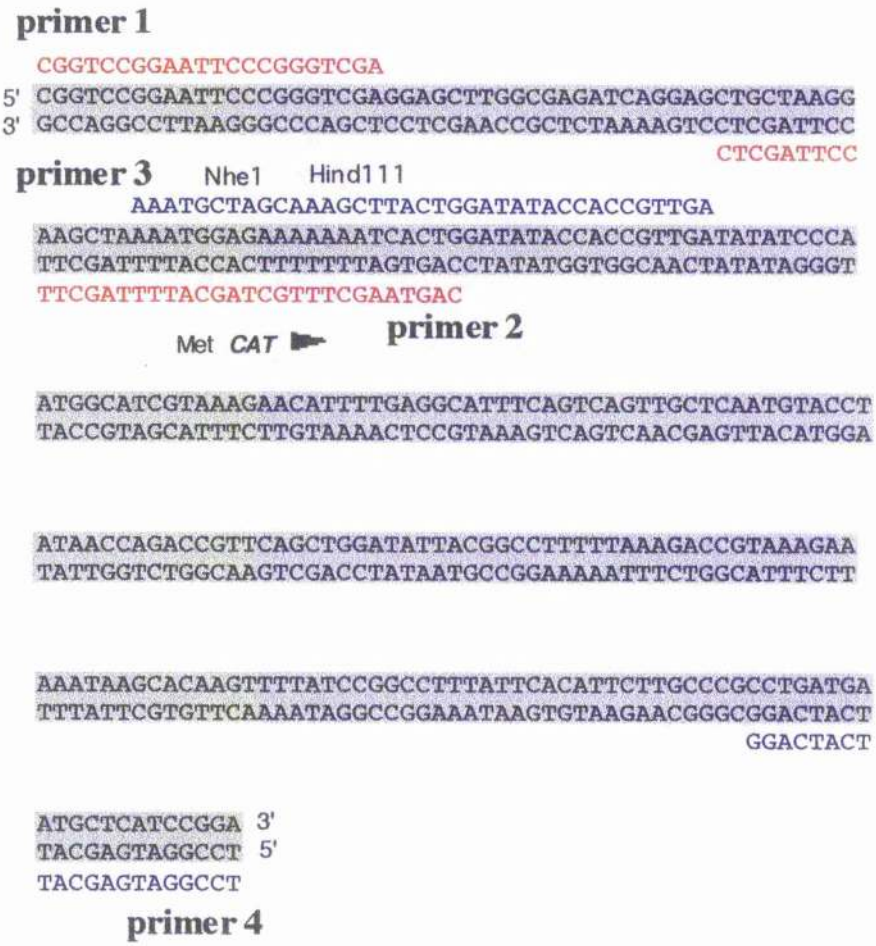


Fig 3.7 The primers chosen to mutate and replace in pBLSV CAT

The sequence above shows the sequence from Fig. 3.4b, the fragment removed from pBLSV CAT. The 4 primers chosen to generate the mutated fragments are numbered according to the scheme in Fig. 3.6. The sequence of the planned *Nhe*I and *Hin*dIII sites are shown in the primers and the original sequence is shown in the grey box to demonstrate mismatch between the template DNA and the 2 inner primers.

and not the original, incorporating the 8 base changes. Primers 3 and 4 (35 bp and 21 bp respectively) generated the 'right hand' product, from c to d. Primer 3 also contained sequence based on the mutated form of the DNA, complementary to primer 2, overlapping by 21 bp (Fig. 3.7). This would allow annealing of the left and right hand PCR products together, as well as introducing the mutations to the PCR products.

3.4 Optimising conditions for PCR

The PCR was optimised based on the performance of a positive control, the amplification of the template DNA from a to b using primer 1 and 4. Sufficient product was made when using 200 ng of each primer, 800 ng of plasmid DNA, 100 µM dNTPs, 1.5 mM Mg^{2+}, Taq polymerase and an annealing temperature of 60°C (all other parameters as described in Ch. 2.10.1).

The right hand reaction was found to proceed under the same conditions but with an annealing temperature lowered to 45°C. The left hand reaction also preferred the above conditions, annealing only at 40°C. Both these annealing temperatures are well below the theoretical values for the primers, but since the mutagenic primers 2 and 3 have very long regions which may bind loosely, or not at all, to the template, this probably lowers the melting temperature considerably by encouraging the primer to remain free in solution. The control (full length) PCR, the right hand and the left hand PCRs can be seen in Fig. 3.8.

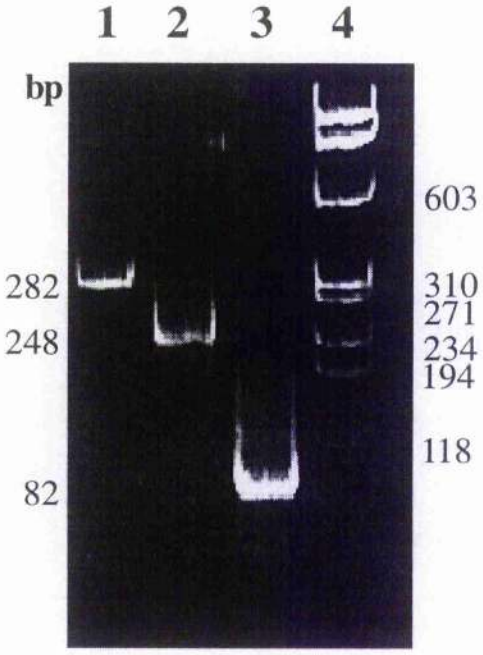


Figure 3.8 Acrylamide gel analysis of PCR products

5% (w/v) acrylamide/ TBE gel of PCR products.

1) Full length PCR, 282 bp

2) 'Right hand' PCR, 248 bp

3) 'Left hand' PCR, 82 bp

4) 'DRIgest III' markers (λ *Hin*dIII digest and ϕ *Hae*III digest)

3.5 Initial attempt to make mutated pBLSV CAT

3.5.1 First attempt to make full length mutated PCR product

PCR is expected to yield an amount of product which greatly exceeds the level of original template. It should therefore be possible to carry out final annealing and PCR of the two fragments straight from small aliquots of the previous reactions, without significant template competing in the reaction.

Each PCR reaction was used (2.5 µl, approx. 25 ng each product) to generate the final full length product in a second stage PCR, using primers 1 and 4. The final PCR utilised the same conditions as the PCRs for the left and right reactions, annealing at 40°C. This yielded a PCR product of the correct size.

3.5.2 Ligation of mutated product into pBLSV CAT

The PCR product was then isolated from a gel, and purified using Qiaex. The purified product was digested with *Cpo*I, at 30°C for 1 h, and subsequently with *Bsi*MI at 55°C for 1 h.

PBLSV CAT was similarly cut with *Cpo*I and *Bsi*MI, and the larger part of the vector purified in the same manner, using Qiaex, to isolate it from the insert. The linearised vector was then dephosphorylated using calf intestinal alkaline phosphatase (CIP) to ensure no religation of any plasmid which may have been cut only with *Bsi*MI, and had escaped digestion with *Cpo*I.

A ligation was carried out using 16 ng of the PCR generated fragment and 16ng of the linearised plasmid, for 18 h at room temperature with 3 units of T4 ligase, 0.5 mM ATP, and 1 µl ligase buffer in 10 µl total volume. The reaction mixture was then transformed into *E. coli* XL1 Blue, made competent by $CaCl_2$ treatment. The transformed cells were incubated for 1 h at 37°C before plating out onto LB agar plates with ampicillin (50 µg /ml).

Few clones were obtained, so only 2 colonies were selected to be analysed by sequencing the plasmid DNA. Overnight cultures (50 ml) were grown from the colonies, and the plasmids harvested using Qiagen Midi Preps. The DNA was then sequenced using the ABI '*Taq* Dyedeoxy Terminator Cycle Sequencing Kit', and read from an ABI automated sequencer. The sequences revealed that the clones contained wild type plasmid.

It was reasoned that the problem lay with the PCR product. It was likely that the PCR product had not been digested fully since the restriction sites lay at the ends of the primers, which can often inhibit digestion. This would therefore prevent ligation into the plasmid. As a result, the only clones obtained were the small numbers of self-ligations, plasmids which had only been cut once, and were able to religate to give the original pBLSV CAT.

To overcome these difficulties it was decided to blunt-end ligate the PCR product into another vector, allowing easy amplification of the product, and permit its full digestion. To prevent self-ligation of pBLSV CAT, more enzyme was to be used to ensure complete digestion.

3.5.3 Ligating and sequencing PCR product in pCRScript™

The PCR product was then blunt end ligated and subsequently transformed into supercompetent *E. coli* XL1 Blue using the pCRScript™ protocol.

White colonies were then selected from the agar plates, small cultures of these were grown overnight and plasmids purified from the cultures using the Mini prep method (Ch. 2.9.2).

These plasmids were then sequenced manually using Sequenase V2.0. The results revealed that the PCR product was wild-type, and did not contain the mutations.

This was a result of amplifying the wild type sequence, in preference to

the annealed PCR products, during the third and final PCR. In the final PCR very small amounts of the original template DNA would be present, carried over from the first 2 PCRs. From this we can deduce that the annealing step when the two PCRs interact with each other over a region of 21 bases is a relatively inefficient one.

It was therefore decided to isolate the left and right hand reaction products away from the template plasmid DNA first, then carry out the final PCR on these products in the absence of original template plasmid DNA, sequencing the final product before ligation into pBLSV CAT.

3.6 Purification of left and right hand PCR products to make full length PCR product

Three methods of isolating the primary PCR products from their template were tried simultaneously, to guarantee success in creating the full length PCR product. This proved fortuitous, since a reagent in the pCRScript^{TM} kit, used in all three methods, proved to be faulty. After testing of the kit it was found that the supplied T4 ligase had lost activity, and was replaced in time to complete the last of the three protocols.

3.6.1 Isolating the left and right hand PCR products by cloning into pCRScript™

The first method used to isolate the PCR products from the template involved cloning the products of each PCR into a vector and selecting for this vector over the wildtype template plasmid. The kit used to clone the PCR products was the pCRScript™ kit (Ch. 2.10.4.2), which uses blunt end ligation of the PCR products into a pre-digested cloning vector.

DNA from the left and right hand *Taq* generated PCRs was purified by chloroform/phenol extraction and precipitation, and redissolved in water. From a total of 20 µl, 4 µl each PCR was used in the pCRScript™ method, including the 'polishing' reaction to ensure all PCR products had blunt ends. A number of white colonies were obtained, and approx. 10% of these colonies were isolated for analysis.

PCR was used as a method for detecting the presence of the left, right and full length products in the plasmids. Some of the clones obtained were expected to be false positives, generated by transformation of the *E. coli* XL1 Blue with pBLSV CAT which had been carried over from the PCRs. The presence of a full length PCR product would indicate the presence of the original plasmid, and not pCRScript™.

PCR demonstrated that only the original template plasmid was present in any of the clones. Although some clones were expected to have pBLSV CAT, others were expected to have been transformed with the pCRScript™/PCR product. The protocol was repeated, using Vent$_R$® (exo-) generated fragments, on the possibility that ragged *Taq* generated ends were not being polished, and therefore not ligating. Again, left and right Vent$_R$® (exo-) generated PCR products were cleaned by phenol/chloroform extraction and precipitation, and redissolved in water. Using 4 µl (from a total of 90 µl, as described in the Stratgene protocol) in the ligation with pCRScript™. The subsequent

transformation generated a number of white colonies for each product.

Again, only template plasmid was found in any of the clones tested. This result may be indicative of a faulty kit; possibly the T4 ligase supplied was not active. Alternatively, this result could be due to the large amount of template plasmid used in the PCR. In the ligation, there could be up to 32 ng of pBLSV CAT carried over from the PCR, compared to 10 ng of pCRScript vector in the reaction. It is possible therefore that the majority of white colonies were derived from the PCR template DNA, even if the PCR product had been cloned into pCRScript. It was expected that the denaturation of pBLSV CAT during PCR would render the plasmid unable to transform cells, but this appeared to be not the case.

In view of the high level of template plasmid contamination, it was decided to purify the PCR products away from the template prior to cloning.

3.6.2 Isolating the left and right hand PCR products by electroelution from gels

3.6.2.1 Electroeluting PCR products

The second method used to isolate the left and right PCR products was the electroelution method. This method was chosen in light of the above result, demonstrating the requirement for purified PCR product before cloning.

Each PCR product (100 µl total) was electrophoresed through a TAE gel over a number of lanes, and PCR products excised, leaving the template plasmid in the gel. The DNA was then electroeluted and precipitated. Approx. 500ng of each product was recovered and redissolved in 50 µl TE buffer. Of this, 5 µl was then used in a PCR to show the fragment could be reamplified, under the conditions previously described. Both PCR products could be amplified satisfactorily (not shown). Since both products had been isolated and were capable of amplification, it was decided to proceed to make full

length product.

3.6.2.2 Generating the full length PCR product from electroeluted DNA

Electroeluted left and right products (5 µl each) were used in a PCR reaction, (conditions as previously described) to obtain the final full length product using $Vent_R^{®}$ (exo-) polymerase. The full length product was made, as can be seen in Fig. 3.9.

3.6.2.3 Cloning full length product made from electroeluted DNA into PCRScript™

The PCR product was subsequently cloned into pCRScript™ as described (Ch. 2.10.4.2). No white colonies were obtained; only a small number of blue, non-recombinant colonies were found, indicating that vector was present, but that the ligation itself had not been efficient. Control ligations also failed. The cells themselves were shown to be competent with a positive control.

It was concluded that the kit was faulty (specifically the T4 DNA ligase), and replacement reagents were obtained. It was also concluded that the kit was probably responsible for the failure of the previous attempt to ligate the 2 PCR products.

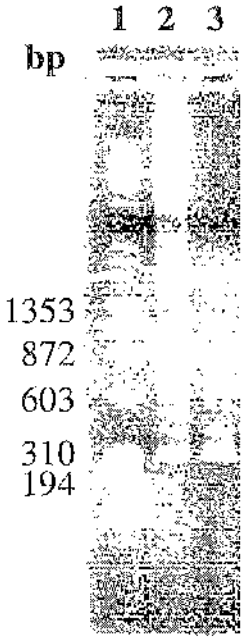


Figure 3.9 Agarose gel analysis of PCRs of electroeluted DNA fragments

1.5% (w/v) agarose gel of full length PCR product made from electroeluted PCR products. Though PCR has occurred, the products were too weak to be visualised without long exposure.

1) 'DRIgest III' markers (λ *Hin*dIII digest and ϕ *Hae*III digest)

2) Full length control PCR from original template, 282 bp

3) Full length control PCR from 10 µl electroeluted right hand PCR and 10 µl electroeluted left hand PCR, 282 bp

3.6.3 Direct PCR on gel slices of left and right PCR products to make full length product

3.6.3.1 Isolation of DNA in gel slices

The third method which was used in conjunction with the methods described in sections 3.6.1 and 3.6.2 was that of direct PCR on gel fragments containing DNA. The left and right PCR products were each electrophoresed through a 1.5%(w/v) agarose gel. PCR products were cut out of the gel, again leaving the template plasmid in the gel. Each gel slice (1-10 µl, melted at 65°C to mix the DNA throughout the agarose) was used in a final annealing/PCR.

The first time this was attempted, no final product was made. It was suspected that insufficient template DNA was present in the gel slice, since only 50-200 ng would be present in the entire slice, and of that, only a small amount (less that 25%), was used in the PCR. To use more of the agarose slice could be detrimental to the PCR, since agarose contains contaminants which are known to affect the quality of the reaction.

Therefore 4 PCRs of both left and right fragments were carried out, and electrophoresed through an agarose gel as before. On this occasion less agarose was taken with each DNA band, and considerably more DNA was recovered. After melting (65°C) the agarose to mix the DNA through the solution, 5 µl of each gel slice was used in a PCR reaction to show the presence of the PCR product and to demonstrate that direct PCR could be achieved successfully. The results of this experiment showed that the PCR products were purified, and were capable of amplification by PCR with agarose present (Fig. 3.10).

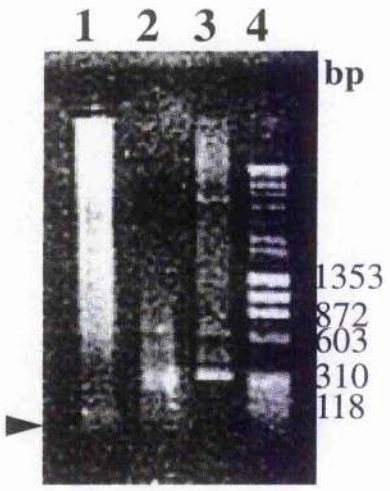


Figure 3.10 Agarose gel analysis of direct PCRs on gel isolated DNA fragments

1.5% (w/v) agarose gel of PCR products made from gel isolated PCR products to show the presence of the DNA fragments, and their ability to be reamplified by PCR.

1) Left hand PCR reamplified from 5 µl gel isolated left hand PCR, 82 bp

2) Right hand PCR reamplified from 5 µl gel isolated right hand PCR, 248 bp

3) Full length control PCR from original template, 282 bp

4) 'DRIgest III' markers (λ *Hin*dIII digest and ϕ *Hae*III digest)

Although faint and contaminated, a light band can be seen towards the bottom of the gel in lane 1, indicating that some of the 82 bp product is present (indicated by the arrow).

3.6.3.2 Generating the full length PCR product by direct PCR on gel slices of left and right PCR products

Using the previous method, with 5 µl of gel from both PCR products, PCRs were carried out to make the full length DNA using Vent$_R$® (exo-) polymerase.

3.6.3.3 Analysing the PCR product

The PCR product was checked by digesting with *Hin*dIII, one of the enzymes that should cleave the mutated form of DNA, but not original wild-type DNA. Fig. 3.12 shows that the full length PCR product had been generated and could be cleaved by *Hin*dIII into smaller, left and right PCR sized fragments, indicating that it contained a new *Hin*dIII site, and therefore behaves as expected for the desired mutated fragment.

3.6.3.4 Cloning full length product from direct PCR on gels into pCRScript™

The PCR products were then cleaned by Wizard™ PCR Preps, prior to cloning. From the 50 µl of cleaned product obtained, 4 µl was used to clone the PCR product into pCRScript™ as previously described. The PCR product was ligated and a large number of white colonies were obtained.

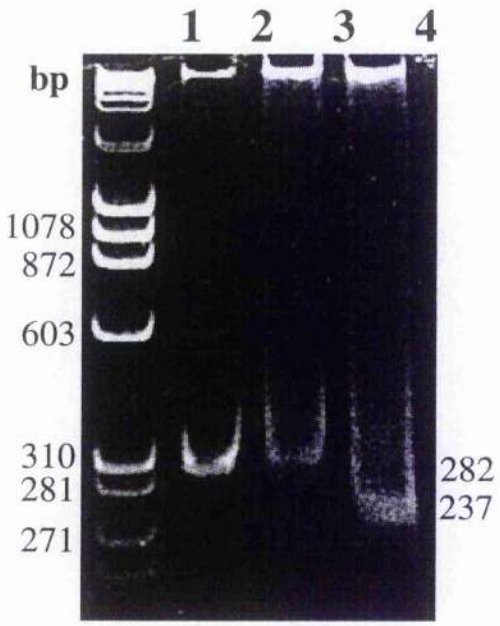


Figure 3.11 Restriction analysis of full length PCR product

The figure shows a 5% (w/v) acrylamide gel of full length PCR product which has been cut by *Hin*dIII.

1) 'DRIgest III' markers (λ *Hin*dIII digest and ϕ *Hae*III digest)

2) Control full length PCR from template plasmid

3) Full length PCR made from excised gel pieces of left and right PCR

4) PCR from lane 3 cut with *Hind*III

The shift in mobility in lane 4 is caused by a loss of approx. 45 bases, though this fragment is too small to be detected in the gel. The mobility shift is due to cleavage of the PCR product by *Hind*III into left and right hand fragments.

Again, the sizes of the PCR products and digestion fragments appear to be larger than expected. This may be due partly to the presence of PCR and restriction enzyme buffers not present within the DNA markers.

3.7 Analysing clones

3.7.1 Restriction digests of the clones

Mini plasmid preps were performed on 3 clones and on the original pBLSV CAT. These plasmids were then used for 2 restriction digests, as shown in Fig. 3.12. Lane 3 shows that the full length PCR is present, cutting at two *Bsi*MI sites that did not previously exist and giving the correct sized fragment, whilst lane 4 indicates the presence of an insert with an extra *Hin*dIII site.

3.7.2 PCR of clones

The clones were then tested for their ability to yield products from the left, right, and full length PCRs, as were initially carried out on the original plasmid. Fig. 3.13 shows a typical result, confirming the restriction digest data that the cloned product is the correct, mutated version of the desired DNA fragment.

3.7.3 Sequencing of the clones

Single stranded DNA was made from the plasmids using the M13 single stranded DNA prep. This was sequenced manually with the Sequenase V2.0 kit, using M13 -40 primer. The results showed that the plasmid contained the correct full length PCR with the new restriction sites in the appropriate place.

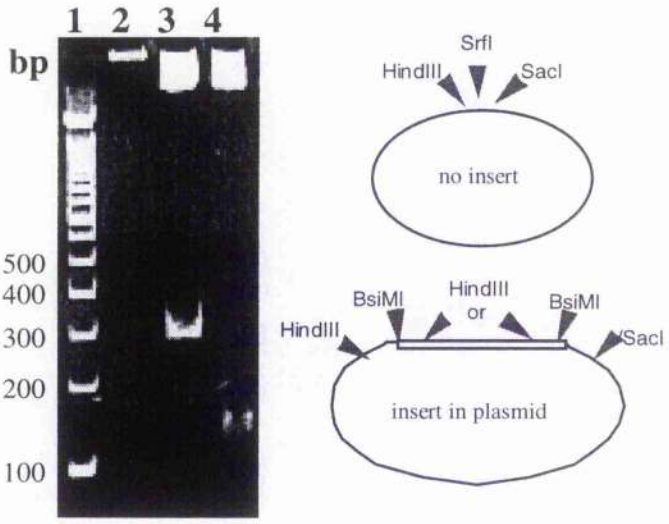


Figure 3.12 Acrylamide gel analysis of digestions of cloned full length PCR

The results of 4 digestions of the cloned mutated PCR product in pCRScript^{TM} are shown on an 8% (w/v) acrylamide/TBE gel. Expected results are calculated for either orientation of the inserted PCR product.

1) Markers; 100 bp ladder

2) Uncut plasmid

3) Plasmid cut with *Bsi*MI only. Expected 266 bp for either orientation. Fragment is approx. 300 bp

4) Plasmid cut with *Hin*dIII only. Expected 109 bp or 238 bp fragment. Fragment is approx. 150 bp

The presence of two *Hin*dIII sites alone in lane 4 would suggest the presence of the insert since the unmutated PCR product does not cut with *Hin*dIII, and pCRScript^{TM} only cuts once (not shown). The sizes of the fragments are close to the expected size, it is therefore concluded that the cloned product is the correct one.

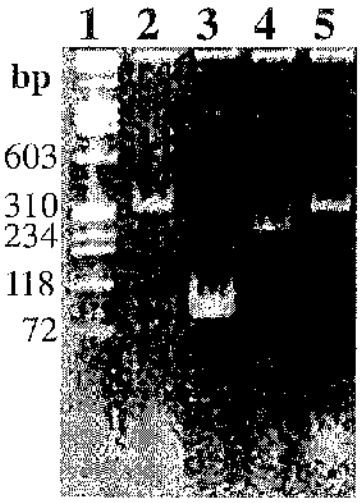


Figure 3.13 Acrylamide gel analysis of PCR products

5% (w/v) acrylamide/ TBE gel of PCR products made from pCRScript with cloned full length PCR product

1) 'DRIgest III' markers (λ *Hin*dIII digest and ϕ *Hae*III digest)

2) Full length control PCR from pBLSV CAT, 282 bp

3) 'Left hand' PCR, 82 bp

4) 'Right hand' PCR, 248 bp

5) Full length PCR, 282 bp

3.8 Excising and purifying mutated fragment

PCRScript plasmid (5 µg) containing the PCR product was made, and then digested overnight at 30°C with 20 µl (100 units) *Cpo*I in a total vol of 500 µl with appropriate buffer. After this, the buffer was changed by precipitating the DNA with 4 M ammonium acetate. The DNA was redissolved in 300 µl *Bsi*MI buffer and digested overnight at 55°C by 20 µl (200 units) *Bsi*MI. In this way, maximum digestion of the DNA was achieved, ensuring that there was enough DNA to recover from gels later.

The cut DNA was loaded onto 2 TAE agarose gels and electrophoresed to permit isolation of the 282 bp fragment. The fragments were excised from the gels and purified using Wizard™ PCR Preps. DNA (200 µl) was isolated at 0.1 mg/ml, but was impure. It was repurified using a phenol/chloroform extraction and a sodium acetate precipitation. The DNA fragment was checked for plasmid contamination by gel analysis (Fig. 3.14).

3.9 Digesting and purifying pBLSV CAT

PBLSV CAT (5 µg) was treated in the same way as above, to remove the region of DNA being replaced. The plasmid was also purified by Wizard™ PCR Preps from a TAE agarose gel, leaving the insert, and retaining the vector. A total of 75 µg vector was isolated.

To ensure that plasmid which had only been cut by one of the enzymes was not allowed to religate, the cut pBLSV CAT was then dephosphorylated by CIP.

Enzymes were removed by extraction with phenol/chloroform, and the DNA precipitated by sodium acetate and ethanol. The DNA was then taken up

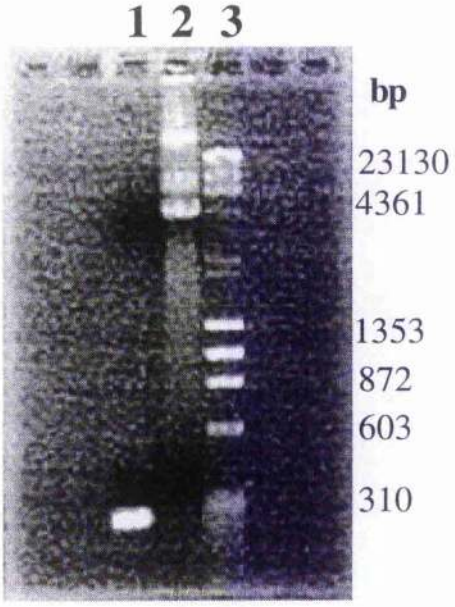


Figure 3.14 Agarose gel analysis of purified insert from pCRScript^{TM}

1.5% (w/v) agarose/TBE gel of purified insert and pBLSV CAT

1) Isolated full length fragment excised from pCRScript^{TM} vector using *Cpo*I and *Bsi*MI

2) pBLSV CAT plasmid, uncut for comparison

3) 'DRIgest III' markers (λ *Hin*dIII digest and ϕ *Hae*III digest)

The gel shows that the cloned PCR product has been excised and isolated with no contaminating template or vector plasmid

into 200 µl CIP buffer, and 20 µl CIP added. Enzyme treatment was performed for 1h at room temperature before heat inactivation. The DNA was then repurified using a extraction with phenol/chloroform and a sodium acetate precipitation.

3.10 Ligating mutated fragments into pBLSV CAT

Ligations were performed in a total vol of 10 µl, both using 250 ng cut vector and 250 ng excised PCR fragment. The ligations were carried out overnight or over the weekend, using Stratagene T4 ligase used in conjunction with 1 µl 10 mM rATP. The ligations were then transformed into DH5α cells supplied by Dr. S. Khan. The transformed cells were spread on LB agar plates containing 50 µg/ml ampicillin. It was discovered that ligating the DNA overnight gave rise to only 9 colonies in total, whereas ligation over a weekend provided a total of 90 clones.

3.11 Digests of pSoUP

This new variant of pBLSV CAT, designed to support experiments upon uncharacterised presequences was called 'pSoUP' (plasmid supporter of uncharacterised presequences).

To find those ligations which had successfully yielded the correct pSoUP, a small number of colonies were taken from ligation 3. Wizard™ Mini Preps of the clones were used in restriction digests to generate a restriction map of the plasmid, to ensure greater chance of sequencing the correct ligation.

All three selected clones behaved as predicted for pSoUP when digested with *Pst*I, *Cpo*I, *Nhe*I with *Pst*I and *Hin*dIII with *Sac*I (not shown). A further restriction digest determined that the PCR product had ligated in the correct

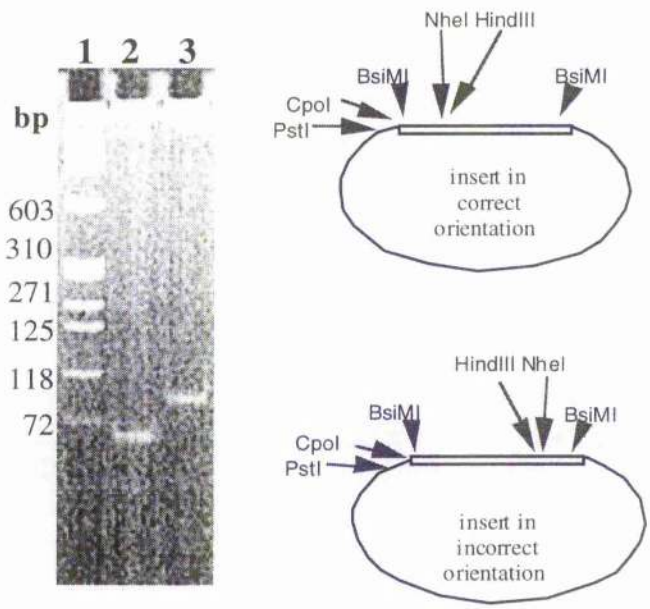


Figure 3.15 Restriction analysis of pSoUP to show orientation of mutated DNA.

The results of 2 digestions on pSoUP are shown on an 8% (w/v) acrylamide gel. Expected results are calculated for the correct orientation of the mutated DNA and the incorrect.

1) 'DRIgest III' markers (λ *Hin*dIII digest and ϕ *Hae*III digest)

2) Plasmid cut with *Cpo*I and *Nhe*I. In the correct orientation, the fragment expected would be 61 bp. Incorrectly inserted, the fragment would be 204 bp. Fragment is approximately 70 bp.

3) Plasmid cut with *Hin*dIII and *Pst*I. In the correct orientation, the fragment would be 75 bp, incorrectly oriented, it would be 204. Fragment is approximately 95 bp.

Incorrectly inserted, both digestions would produce a fragment of the same size (204 bp). Since each digestion produces a fragment of a different size, it

Figure 3.15 cont'd.

would appear that the PCR product is in the correct orientation. The reason why both fragments appear to be 10-20 bp larger than expected may be due to the presence of digestion buffers or traces of salt and protein affecting running of the gel.

orientation, as can be seen in Fig. 3.15.

3.12 Sequencing of pSoUP

One of the clones was sequenced. The plasmid was prepared by Wizard™ Mini Prep and then sequenced using the ABI automated sequencer and sequencing kit. All four of the primers which were initially used to generate the PCR product were used as sequencing primers. The sequence obtained demonstrated that the plasmid was exactly as predicted, containing both the new restriction sites without any other alterations to the CAT gene in the region. This clone was subsequently used to establish a permanent frozen stock.

3.13 Discussion

One of the difficulties faced in this series of experiments which isolated the left and right hand products was that of size. At the beginning of this research, it was not anticipated that small fragments of PCR amplified DNA would often not precipitate or purify by many of the available methods. The 82 bp fragment in particular was difficult to purify in quantity, and there was frequently less material to use than desired. This problem led to all subsequent DNA manipulations being performed on a larger scale than would otherwise be necessary.

Another problem encountered seems to have been caused by the unusual PCR conditions. After optimisation, up to 10 times more plasmid was employed in the PCR than is often required. In the case of primer 2 and 3 inefficient binding of the primers with long extensions could lead to more template being required, but this does not explain the requirement in control reactions where the primers match exactly. It is likely that the plasmid is

especially resistant to heat denaturation, and a longer period of heating was required at the start of the PCR. This initial step was never examined since the standard time used was considered adequate before the other reaction conditions were optimised. As a result of using large quantities of template plasmid to overcome this denaturation problem, wild-type pBLSV CAT has contaminated many of the efforts to use reaction products directly. This could also explain the number of false positives seen in section 3.6.1. In this experiment, pBLSV CAT was carried over from the initial PCR, through the ligation reaction, to the transformation. If the pBLSV CAT was not fully denatured during the PCR, then it would retain the potential to transform bacteria at this stage, as appears to be the case. The annealing temperatures for the PCR were also unusually low. The left and right hand reactions were over 10°C lower than expected, probably due to the extended 5' overhang. Many site directed mutagenesis PCRs do not require the lower temperature seen here. This is because, in place of the more common single restriction site, 2 sites (14 and 17 bp for left and right hand reactions respectively) were required.

Due to the large amount of plasmid required, isolation of the mutated PCR product was necessary. Three methods were used to isolate DNA species, all of which utilised a kit called pCRScript™. Two of these methods failed due to a faulty ligase in the kit, undetected due to the omission of a suitable positive control. Until the failure of the method to clone the electroeluted DNA it was thought that plasmid contamination was the sole cause of failure in previous experiments, since the pCRScript™ kit had previously given excellent results (see chapter 4). The full length clone had been generated by the electroelution method (section 3.6.2.2) from isolated left and right PCR products, in the absence of a full length template DNA. If the faulty ligase had been detected earlier, the electroelution method would probably have succeeded also.

What other methods for DNA mutation were available, and why were these not used? The first method considered was based on similar approach. The theory was to excise a fragment around the mutation site small enough to be replaced by a synthesised oligomer. This method would have been ideal if 2 unique sites had been available close together, near the first coding methionine of the CAT gene. However, the nearest sites found were 282 bp apart, and such an oligomer could not be made satisfactorily.

It was also suggested that using a long range polymerase, such as *pfu*, one could take 2 primers, with the appropriate mutations, and polymerise daughter strands completely round the plasmid, a technique now commonly referred to as 'inverse PCR'. This method would avoid a second stage PCR, and avoid any ligations. However, only a small number of mutations had ever been attempted at one time (Dr. Alan Wise, private communication). It was unlikely that the primers, with the 8 mutations required, would anneal easily enough to permit this difficult procedure.

Finally, the most common method for site directed mutagenesis, the site directed mutagenesis kit (e.g. Amersham's Sculptor^{TM} kit), was discounted for the same reason as above. It may be that such a kit would have succeeded in less time, by generating the mutations one at a time. As site directed mutagenesis becomes more common place, so newer kits make such procedures easier, and should this be done again, such a kit may be the method of choice.

Chapter 4

Preliminary Investigation into Cloning of the E2 Presequence DNA

4.1 Introduction

As discussed in chapter 3, to investigate the possible roles for the extended presequence of PDC-E2, a presequence/CAT reporter gene construct was to be made and expressed in mammalian cells. Vital to the creation of a presequence/reporter gene construct was the isolation of the DNA coding for the presequence. However, no clone which included the complete PDC-E2 presequence was available to us. There were two published sequences which existed at the time this research commenced (Coppel *et al.*, 1988; Thekkumkara *et al.*, 1988), as shown in Fig. 4.1. The presequences of the pre-E2 polypeptide are identical over their C-terminal region (amino acid 59-75 in the figure). However, from amino acid 1 to 59 the two presequences are totally divergent. It was revealed to us in communication with R. Coppel and co-workers that a recombination event had taken place, and that the rest of the sequence was derived from the cloning vector (Dr. M.E. Gershwin, private communication). The sequence published by Thekkumkara *et al.* (1988), from a cDNA clone, did not have the entire presequence since it possessed no 5' initiation codon, so could not be used in this investigation.

During the course of this research it was found that the sequence for the entire human PDC-E2 presequence had been elucidated, but left unpublished. Sequencing of a 'genomic clone' had been carried out by Dr. P. Leung (Dr. M.S. Patel, private communication) to obtain the sequence of the upstream promoter region. Subsequently, the full presequence had also been determined, and this was made available to us, although the clone, however, was not available.

We therefore had the DNA sequence but no clone of the PDC-E2 presequence; we had to obtain this ourselves from a cDNA library or genomic DNA.

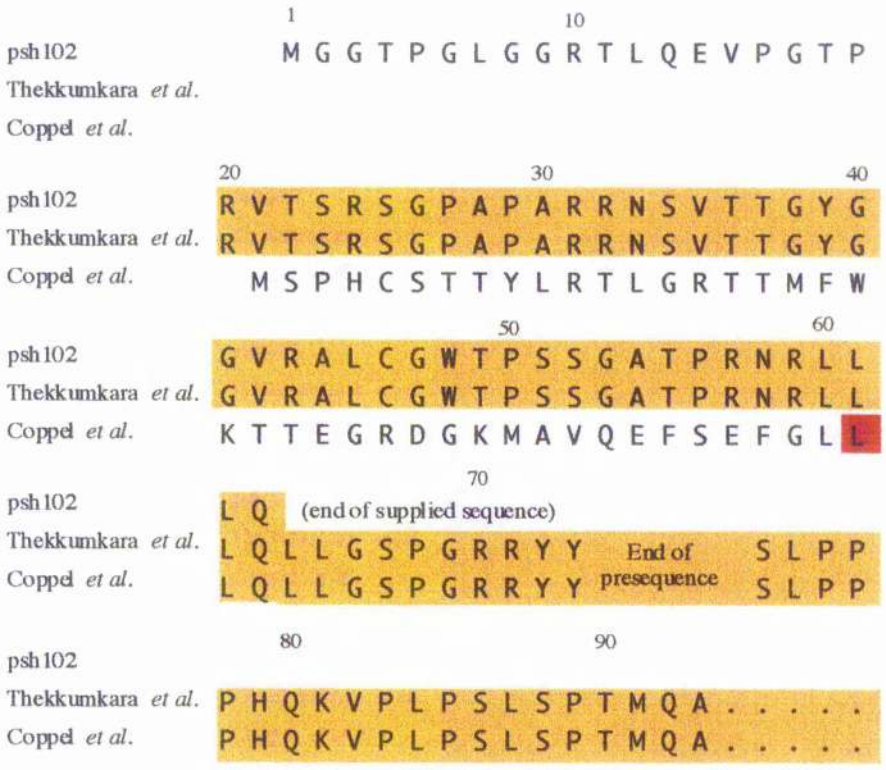


Figure 4.1 Sequence alignment of the 3 PDC-E2 presequences

The three available PDC-E2 presequences are aligned above to give the correct overall sequence. Regions of identity are shaded yellow. The two published sequences (Thekkumkara *et al.*, 1988; Coppel *et al.*, 1988) agree over the entire DNA sequence coding for mature PDC-E2, and for the final 14 amino acids of the presequence. However, the two sequences are completely divergent over amino acids 1-59 of the presequence. The point at which the sequence by Coppel *et al.* (1988) diverges from the other two is thought to be due to genetic recombination in the cDNA library, and this point is shaded in red.

Sequence from the genomic clone (a plasmid called pSh102), shown on the top line, agrees with the sequence published by Thekkumkara *et al.* (1988).

4.2 Designing primers to amplify the presequence DNA

Primers were designed to amplify the PDC-E2 presequence DNA coding for the region between the N-terminus of the presequence and the N-terminus of the mature protein (designated 'presequence only'), based on the sequence provided. Both primers incorporated restriction sites for directional cloning into pSoUP (Fig. 4.2). A third primer was also designed after the discovery that this PCR generated multiple products when amplifying from cDNA. This new primer was identical to a region within the open reading frame of mature PDC-E2. This would be employed along with the primer designed to the N-terminus of the presequence to amplify a 'partial clone', coding for the presequence and part of the mature PDC-E2 polypeptide, as shown in Fig. 4.2. Being able to amplify this product would demonstrate that the presequence was adjacent to the rest of the PDC-E2 gene, and would support the genomic sequence supplied. This amplified product could then be used as the template for the original PCR to obtain the presequence alone, with appropriate restriction sites.

4.3 Obtaining E2 presequence by PCR from a λgt10 cDNA library

4.3.1 Obtaining the E2 presequence from λgt10 by PCR

A number of attempts were made to amplify the 'presequence only' DNA from a human foetal cDNA library housed in λgt10, assuming heat denaturation of the λgt10 coat protein would release the viral DNA for amplification. This is a standard approach occasionally used to amplify from whole bacterial cells. Each primer (0.3 μg) was used to amplify the presequence from 10^9 pfu of λgt10.

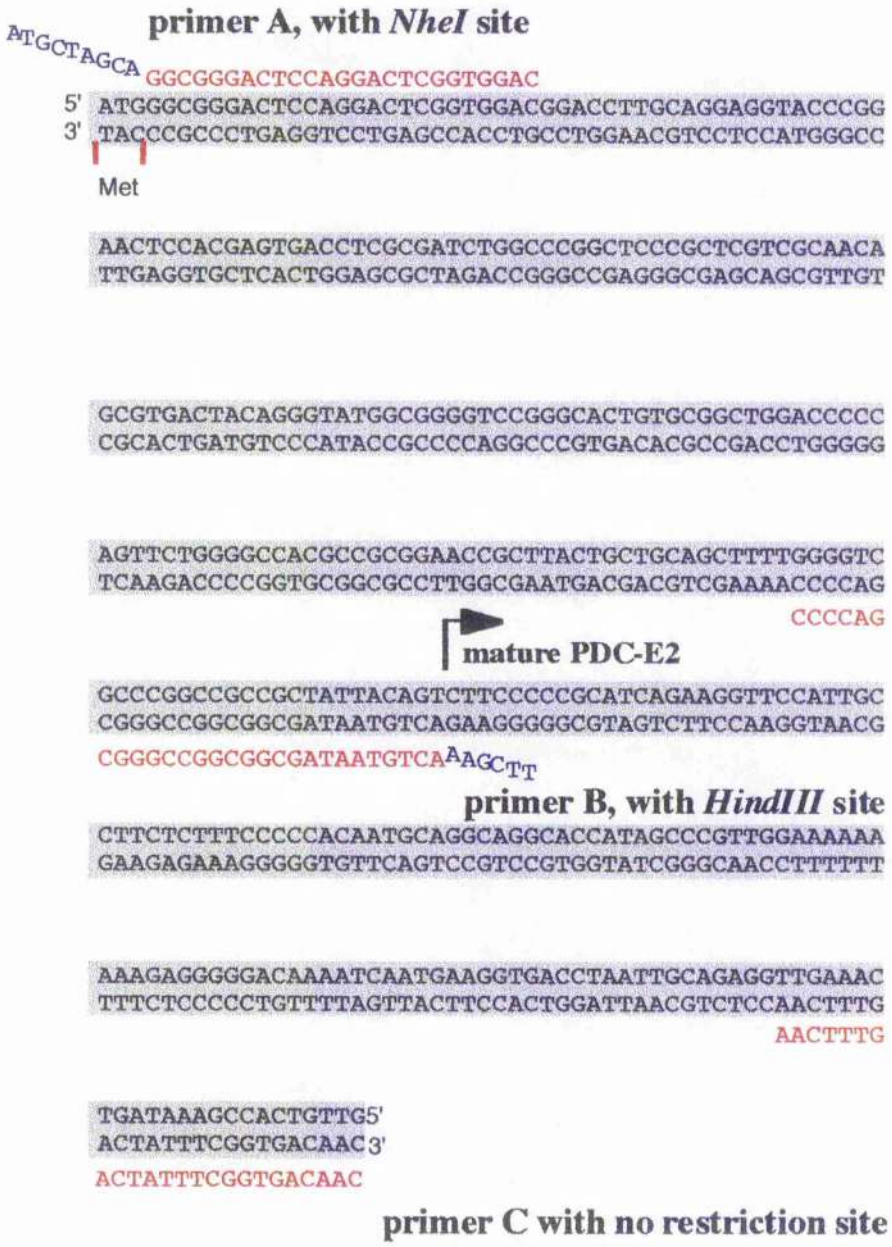


Figure 4.2 The sequence for PDC-E2 up to the internal primer site, and the 3 primers designed to amplify the DNA

The two original primers (A and B) designed to amplify the presequence are shown, both possessing one of the restriction sites which would allow directional

Figure 4.2 cont'd.

cloning of the presequence into pSoUP. Primer A was identical to the sequence shown, up to the first codon (shown in red) with an additional restriction site (shown in blue), designed for in-frame insertion into pSoUP. Primer B was designed to the reverse complement of the 3' end of the DNA coding for the presequence (in red), also including a restriction site (in blue) for directional cloning into pSoUP. The third primer (Primer C) was designed to work with primer 'A' to amplify the 'partial clone', but possessed no restriction site.

The PCR was carried out using $Vent_R^{®}$ (exo-) polymerase using the following steps:

1) 94°C for 3 min initial denaturation
2) 30 cycles of
    94°C 1 min
    45°C 1 min 30 sec
    72°C 1 min 30 sec
3) 72°C for 5 min final extension

This did not yield any products, so the method was repeated, using a lower annealing temperature (40°C instead of 45°C). This gave faint multiple PCR products (Fig. 4.3). Whilst heat denaturation followed by PCR is an accepted technique for amplifying DNA directly from bacterial cells, it appeared to be unsatisfactory for amplifying DNA held in viral capsids. It was decided to use a standard protocol to isolate the DNA from the phage particles, and PCR directly from the DNA.

4.3.2 Using purified λgt10 DNA in the E2 presequence PCR

DNA was purified from the phage for use in PCR applications, as described in the methods section (2.9.5). Using previous conditions, annealing at 45°C, the 'presequence only' PCR was attempted. This generated faint, multiple products. A third primer, primer 'C' in Fig. 4.2, was made to amplify a 'partial clone' DNA (as described in section 4.2). PCRs using primer A and C to amplify the 'partial clone' also gave rise to multiple products, but, in contrast to the reaction using primers A and B, one band predominated. The contrast between the two PCRs can be seen in Fig. 4.4, where 'presequence

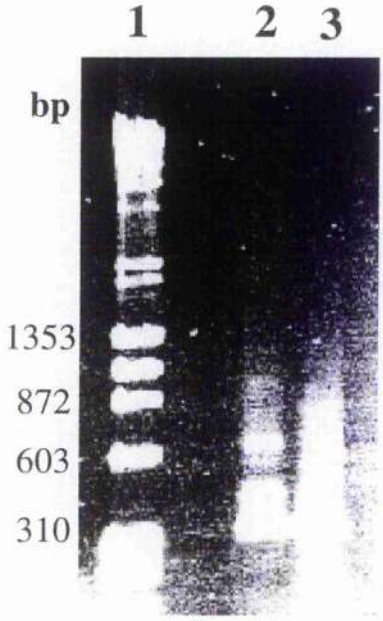


Figure 4.3 Agarose gel analysis of the first PCRs for PDC-E2 presequence

Agarose gel 1.5% (w/v) of two PCR reactions carried out on λgt10 demonstrates the variability of the PCR carried out on whole phage. The products were too weak to be visualised without long exposure. Expected size of correct PCR product is 234 bp.

1) 'DRIgest III' markers (λ *Hin*dIII digest and ϕ *Hae*III digest)

2 and 3) Independent PCRs carried out as described on heat-treated phage to amplify the 'presequence only' DNA, demonstrating variability as well as a lack of specific product of the correct size.

only' and 'partial clone' PCRs are compared over a range of template DNA concentrations. The expected size of the product from PCR with primers A and C was 343 bp, PCR with primers A and B should give a product of 228 bp. The predominant band found in the 'partial clone' PCRs was larger than expected, but it was decided to pursue investigation of this product despite the large size. Unfortunately, attempts to clone this product into pUC118 failed, due to a contaminated stock of the plasmid. All other cloning experiments were therefore carried out using pCRScript™.

4.4 Cloning and analysis of PDC-E2 PCR product

4.4.1 Cloning 'presequence only' PCR product into pCRScript™

Difficulties were encountered at this point, caused by a batch of faulty *Taq* polymerase. During this time, the products of the 'presequence only' PCR made above were ligated into pCRScript™, according to the manufacturer's protocol. A number of recombinant clones were isolated, and five of these were selected for further analysis.

4.4.2 Analysing 'presequence only' PCR product in pCRScript™

Wizard™ Mini Preps of plasmid DNA from each of the 5 selected clones were made, and 20 µl plasmid (1 µg) was digested with 10 units *Nhe*I and 20 units *Hin*dIII for 3 h, 37°C. The restriction digestions were then analysed on a 1.5% (w/v) agarose gel. Of the 5 digestions, 4 yielded products, which were all estimated to be approx. 300 bp, which corresponds approx. to the predicted size of 226 bp (Fig. 4.5).

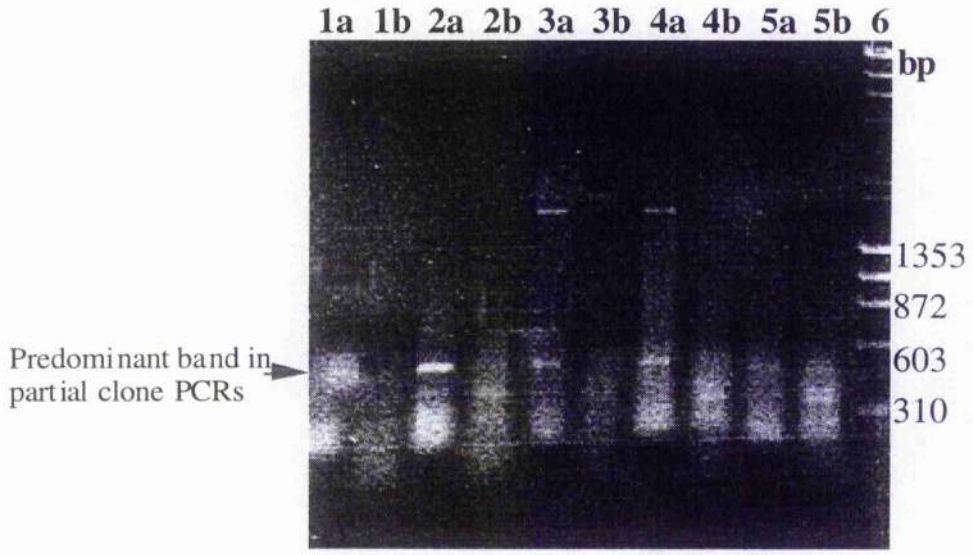


Figure 4.4 Agarose gel analysis of PDC-E2 presequence PCRs on purified phage DNA

This 1.5% (w/v) agarose gel of 'presequence only' and 'partial clone' PCRs carried out on λgt10 DNA shows the variability of the PCR with different amounts of initial template. Expected size of correct 'presequence only' PCR product is 234 bp, expected size of 'partial clone' PCR product is 343 bp.

PCRs are labelled 'a' for 'partial clone' PCR and 'b' for 'presequence only' PCR.

1a,b) PCRs using 3 ng template

2a,b) PCRs using 30 ng template

3a,b) PCRs using 150 ng template

4a,b) PCRs using 325 ng template

5a,b)PCRs using 650 ng template

6) 'DRIgest III' markers (λ *Hin*dIII digest and ϕ *Hae*III digest)

4.4.3 Sequencing 'presequence only' PCR product in pCRScript™

Since identical fragments were excised in all the restriction digests, only one of the clones was sequenced. Single stranded DNA was made from the clones, and was manually sequenced with M13 -40 primer. The insert was detected, but the sequence bore no homology to the expected sequence, except in the region of the primers.

4.5 Obtaining E2 presequence by PCR from human genomic DNA

At this point it was felt that the λgt10 cDNA library available to us may represent a poor source of the template DNA required. Many researchers have had difficulties obtaining full length clones from cDNA libraries, including the authors of the two published human PDC-E2 sequences, who were unable to find correct full length cDNA clones in their original publications of the PDC-E2 gene sequence.

It was decided to use human genomic DNA in further PCRs, which had proven to be a useful source of template DNA for related PCR work.

Blood was kindly given by Prof. Gordon Lindsay, and DNA purified from it by Susan Richards, using a proteinase K treatment followed by a phenol/chloroform purification and sodium acetate precipitation. The DNA was found to be at a concentration of approx. 0.47 mg/ml.

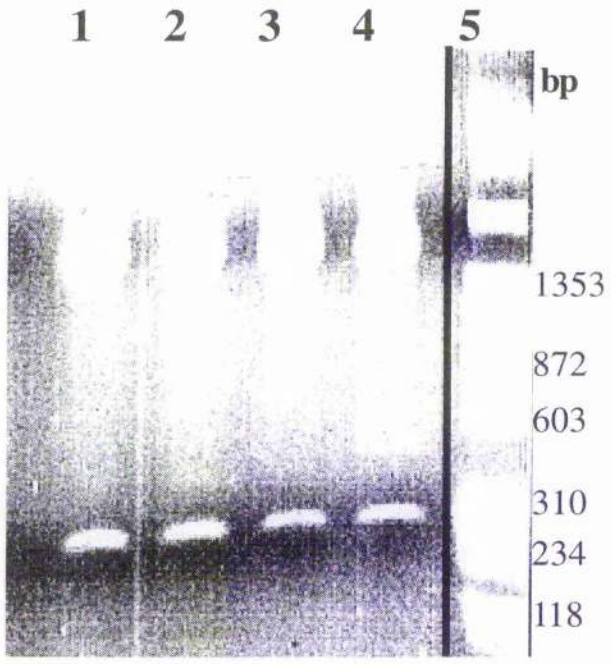


Figure 4.5 Agarose gel analysis of digestion of 4 clones

The plasmids obtained from the 'presequence only' clones were digested with *Nhe*I and *Hin*dIII, and were then examined on a 1.5% (w/v) agarose gel. The expected size of the fragment generated by digestion of the correct product in pCRScript^{TM} would be 228 bp. Variation in the running of individual lanes of the gel made accurate sizing of the fragments difficult, but the product appears to be approx. of the correct size.

1-4) Successful digestions, all yielding products of approx. 250-300 bp.

5) 'DRIgest III' markers (λ *Hin*dIII digest and ϕ *Hae*III digest)

PCRs for the presequence, using both 'presequence only' and 'partial clone' primers, were optimised to provide a reasonable yield of PCR product. Optimum conditions were similar to before, using 5 µl genomic DNA in 50 µl reactions with 0.3 µg each primer. *Taq* polymerase was preferred to $Vent_R$® (exo-), since *Taq* gave more product with this template. All PCRs gave multiple products (Fig. 4.6). This is in contrast to a control PCR, using primers based on confirmed published sequence to amplify the lipoyl domains and E3 binding site of PDC-E2. This PCR gave a much cleaner product at precisely the size expected.

4.6 Cloning and analysing all PCR reaction products in pCRScript™

4.6.1 Cloning PCR into pCRScript™

Since the PCR had a number of DNA species within the DNA mixture it was decided to clone the mixture of reaction products into pCRScript™. From this, individual clones could be selected, and the fragments cloned could be identified by restriction digestion. The correct clone should yield a fragment of 228 bp.

Only the 'presequence only' PCR products were used. The possibility was noted that the presequence may be on a different exon to the rest of the gene; the correct 'partial clone' PCR product could therefore be very much larger than predicted. A number of 'presequence only' PCRs were purified by sodium acetate and ethanol precipitation. The purified DNA mixture, at 10 ng/µl, was ligated into pCRScript™ as previously, and used to transform *E.coli* XL1 Blue with recombinant plasmid. A large number of recombinants were made, 20 of which were selected for analysis to find a clone for each PCR product that was made.

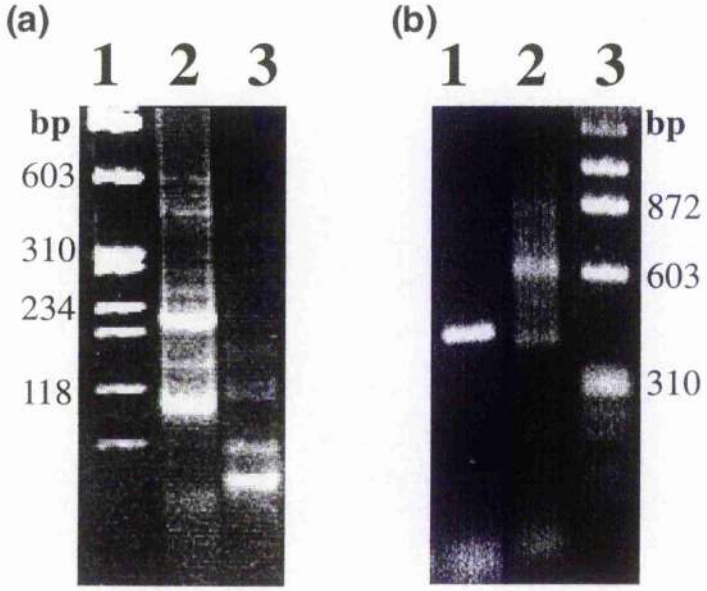


Figure 4.6 Gel analysis of PCRs from human genomic DNA

Figure 4.6(a)

The products obtained from the presequence PCRs were visualised on an 8% (w/v) acrylamide gel. Whilst there are a large number of amplified products in each reaction, there are some major bands present. In the 'presequence only' reaction, there is one candidate at approx. 220 bp, which is the expected size for the correct PCR product.

1) 'DRIgest III' markers (λ *Hin*dIII digest and ϕ *Hae*III digest)

2) PCR to amplify 'presequence only' DNA. Expected size 234 bp

3) PCR to amplify 'partial clone' DNA. Expected size 343 bp

Figure 4.6(b)

The control PCR was compared to the products of a 'presequence only' PCR on a 1.5% (w/v) agarose gel. The control PCR, amplifying a section of DNA from the PDC-E2 gene, gives a single, clean product, whilst the experimental PCR gives a range of products.

1) PCR to amplify DNA coding for E3 binding domain and lipoyl domain of PDC-E2. Expected size 543 bp

2) PCR to amplify 'presequence only' DNA. Expected size 234 bp .

3) 'DRIgest III' markers (λ *Hin*dIII digest and ϕ *Hae*III digest)

There was a degree of variability between PCRs (such as 'presequence only' PCR in panels a and b). However, the same DNA bands were usually present in each PCR, in varying concentrations.

4.6.2 Analysing PCR clones by restriction digest analysis

Wizard™ Mini Preps were made from the selected clones, and the plasmids obtained digested with *Nhe*I and *Hin*dIII. Each plasmid (20 µl, approx. 1 µg) was digested with 10 units *Nhe*I and 20 units *Hin*dIII overnight, 37°C. A sample of the 20 digestions is shown in Fig. 4.7. The result obtained was surprising since it was expected that cloning the whole PCR reaction should yield clones of the various products found in the PCR. From the 20 clones analysed, only 2 varieties of recombinant emerged; one which did not digest with the 2 enzymes used, and one which gave a fragment of approx. 400 bp, which was regarded as too large to be the desired presequence, and was therefore discarded.

4.7 Cloning and analysis of individual PCR products into pCRScript™

4.7.1 Isolation of individual PCR products

It was not known why a 400 bp product from this PCR should be selectively ligated so it was decided to isolate all candidate PCR products on a gel, and clone each species into pCRScript™ separately. Again, the 'presequence only' PCR products were used.

The 'presequence only' PCR products were precipitated using sodium acetate and ethanol, and the washed, purified DNA pellet resuspended in half the original volume for loading onto a 1.5% (w/v) agarose/TAE gel. All the major PCR products (a, b, c and d), as shown in Fig 4.8, were excised from the gel, and the DNA isolated using the Wizard™ PCR prep kit. Of the 4 bands purified, 1 did not yield sufficient DNA to continue (band 'a'), and another was approx. 150 bp too large to be the presequence (band 'd'). The other two bands (band 'b' and 'c') were isolated both individually and together in a

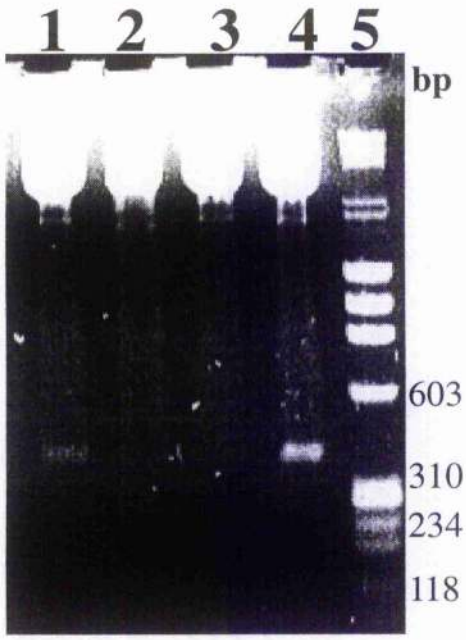


Figure 4.7 Agarose gel analysis of digested plasmid clones of human genomic PCR

The plasmids obtained from the recombinant clones were digested and analysed on an 1.5% (w/v) agarose gel. The figure shows a representative sample of the 20 digestions, which all gave patterns identical to 1 of the 4 above.

1 and 4) partial digestion of the plasmid to yield a 400 bp fragment

2 and 3) no digestion of the plasmid

5) 'DRIgest III' markers (λ *Hin*dIII digest and ϕ *Hae*III digest)

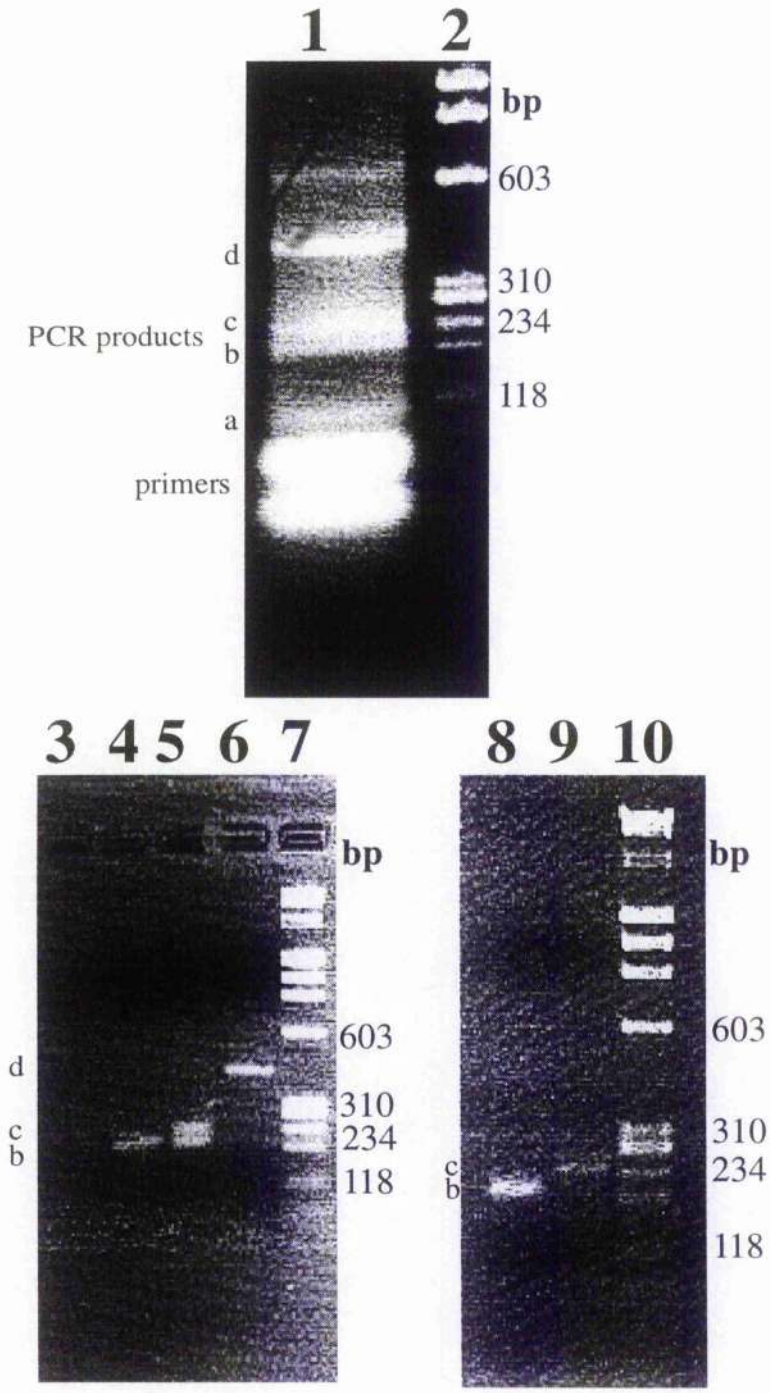

1
2
bp
603
310
234
118
d
c
b
a
PCR products
primers
3 4 5 6 7
8 9 10
bp
603
310
234
118
d
c
b
c
b

Figure 4.8 Agarose gels showing isolation and analysis of various PCR products from human genomic DNA

The precipitated DNA was electrophoresed (lane 1, top gel), and 4 major bands isolated (a, b, c, d). The isolated products are shown in the lower left gel. After discovering band 'c' was contaminated, it was re-isolated, as shown in the lower right gel. All gels were 1.5% (w/v) agarose. Band 'a' was not re-isolated, since, on reflection, it was too small to be a suitable candidate.

1) PCR DNA prepared for isolation
2) 'DRIgest III' markers (λ *Hin*dIII digest and ϕ *Hae*III digest)
3) Purified band 'a'
4) Purified band 'b'
5) Purified band 'c'
6) Purified band 'd'
7) 'DRIgest III' markers (λ *Hin*dIII digest and ϕ *Hae*III digest)
8) Purified band 'b'
9) Purified band 'c'
10) 'DRIgest III' markers (λ *Hin*dIII digest and ϕ *Hae*III digest)

mixture of the two.

These purified PCR products were then precipitated from 50 µl to 20 µl, and 10 µl of this used in the pCRScript™ protocol, at approx. 20 ng/ml. The 'polish' step was included to promote blunt-end ligation of the *Taq* generated DNA. The ligation was transformed into *E. coli* XL1 Blue, according to the protocol, and recombinants from the procedure selected for further analysis.

4.7.2 Analysing various isolated PCR products

Wizard™ Mini Preps were made from some of the clones of both products, and were tested for the presence of a PCR generated insert with *NheI* and *HindIII* sites, by digesting with the 2 enzymes as previously. It was found that both selected products had been cloned individually. A clone was also found which had been generated from a ligation of the mixture of the two products into pCRScript™ (Fig. 4.9). Single strand DNA was made from each clone, which was then manually sequenced, using the M13 -40 primer. None of the sequences were found to have homology to the presequence DNA expected.

4.8 Cloning and analysis of 'partial clone' PCR

Since none of the PCR products utilising primers A and B had yielded correct products, it was decided to check the sequence of the products made by the PCR using primers A and C. For simplicity, it was assumed that the DNA sequence for the 'partial clone' would be on the same intron, and the expected product would be 343 bp. A 'partial clone' PCR was electrophoresed on a 1.5% (w/v) agarose TAE gel, and a band seen at approx. 350 bp (data not shown) was isolated using the Wizard™ PCR prep, into a total of 50 µl. Approx. 5 ng purified DNA was cloned using the pCRScript™ method, and a

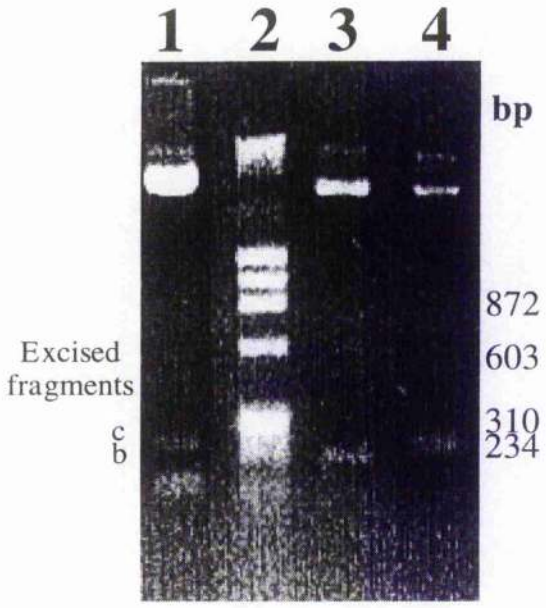


Figure 4.9 Agarose gel analysis of digested clones of individual PCR products

The isolated PCR products cloned and digested with *Nhe*I and *Hin*dIII, analysed on a 1.5% (w/v) agarose gel. The expected size of a 'presequence only' PCR product is 234 bp; digested this product would be 228 bp.

1) Clone of b and c mixture. Band c is cloned (approx. 260 bp), but a shorter fragment, unlikely to be band b, can also be seen.

2) 'DRIgest III' markers (λ *Hin*dIII digest and ϕ *Hae*III digest).

3) Clone of band b. Fragment approx. 220 bp

4) Clone of band c. Fragment approx. 260 bp

number of recombinant clones selected for analysis.

Wizard™ Mini Preps were made from the selected clones, and these were tested for the presence of a *Pst*I site by restriction digestion along with *Nhe*I (found in the primer) as previously described. A *Pst*I site would be present within the correct 'partial clone' sequence.

When the products of the digestion were analysed by electrophoresis through a 1.5% agarose gel, no fragments were obtained, indicating the lack of a *Pst*I site.

Single stranded DNA was made from the clones, which were manually sequenced using M13 -40 primer. The sequence data revealed the presence of the primers, but the sequence between the primers exhibited no homology to the PDC-E2 presequence DNA.

4.9 Using shorter primers to PCR the presequence

It was suggested that the restriction sites incorporated into the primers were causing priming at incorrect locations. Upon this advice, new primers were ordered. These corresponded to primers A and B, but possessed no restriction sites and were only 17 bases in length. It was expected that these new primers would not generate so many products, only the PDC-E2 presequence. The presequence would be cloned into pCRScript™ , and from there it would be re-amplified using the older primers.

All parameters for the PCR were varied to optimise conditions, but again only a wide range of multiple products could be made. Since this method had proven to yield no purer PCR products than previous attempts, it was decided not to pursue this line of investigation any further.

4.10 Discussion

This chapter describes a number of attempts to amplify a segment of DNA coding for the presequence of PDC-E2 from a λgt10 human foetal cDNA library and human genomic DNA by PCR.

As mentioned already, it is not surprising to find that the λgt10 human foetal cDNA library did not yield the correct product. Since cDNA is generated by reverse transcriptase, long stretches of DNA can frequently be truncated by the enzyme, and can also cause deletions where the mRNA being transcribed possesses strong secondary structure.

More difficult to understand is how genomic DNA, the source of the original, full length sequence, failed to give the correct PCR product. It is possible that the PCR had not been fully optimised. Although optimised for annealing temperature, DNA concentrations, dNTP concentrations and enzyme type, it could be that other factors, such as length of time spent in the initial denaturation, were not optimal. If the region of DNA supposed to be amplified is left in a native state with large regions of secondary structure, this could prevent the correct PCR product being synthesised whilst many incorrect ones were made. This would not be the case for the original investigator, who possessed this DNA in a plasmid which would aid such work.

It was not expected that the presequence is encoded on more than one exon, but it may be that an intron lies after the end of the presequence, separating the DNA coding for the presequence and the mature part of PDC-E2. This region is known only by cDNA sequence and not covered by the genomic sequence given to us. If this were the case the 'partial clone' PCR products would be considerably larger than expected, or could be too long to be completed by the polymerase during the extension cycle. When the research commenced it was considered unlikely that this short stretch of DNA

would overlap an intron , especially since it has been shown that the nearby E3 and lipoyl domains are encoded on one exon. However, in the light of the difficulties encountered, the possibility that an intron towards the end of the presequence exists must be considered.

Though we have been unable to amplify the presequence based on the given sequence, we have successfully carried out a control PCR reaction on a segment of the PDC-E2 gene near to the presequence, based on other sequence data. It is possible therefore that the sequence given to us was incorrect, but this does not seem likely. The sequence in question agrees with previously published data, is in frame with this sequence, and has been obtained from genomic sources. A possible explanation for the difference between the presequence PCRs and the control reaction may be a large intron in the PDC-E2 gene, separating these two regions. This would allow the control reaction to proceed, but, as mentioned above, could interfere with the experimental PCRs. It must also be noted that the primers could also be at fault. A number of minor flaws in the design (e.g. 5' GC rich regions) were found which could lead to mispriming at temperatures below the the T_m of the primer. Possibly this is the cause of the multiple bands in the presequence PCRs, in direct contrast to the single strong band generated by the control reaction.

The employment of a new PCR based application called 'rapid amplification of cDNA ends' (RACE) could aid in obtaining this presequence in future. By combining cDNA synthesis and PCR techniques, this method can be used to discover the sequences at the 3' and 5' ends of a gene. Subsequent to this, the newly sequenced terminals of the gene can be used as primers in a standard PCR to amplify the whole gene.

In conclusion, a number of different primers designed to amplify the DNA coding for the presequence of PDC-E2 have been able to amplify many products from genomic DNA, none of which appear to be the correct target

sequence. Whilst this may generate doubt over the validity of the proposed sequence, it is felt that the problem lies in the template DNA, and not in the DNA sequence or the design of the primers.

Chapter 5

Making PDC-E2 Presequence Segments

5.1 Introduction

The aim of this research was to determine possible functions of the PDC-E2 presequence in targeting and preventing premature folding/assembly of the polypeptide, and specifically to discover if a limited section of the presequence was sufficient for targeting. In other extended targeting presequences, additional information is often incorporated into the presequence. For example, the bipartite presequence for yeast cytochrome b contains firstly an N-terminal mitochondrial-matrix targeting signal. C-terminal to that is another sequence which redirects the polypeptide from the matrix to the intermembrane space. Another example is the subunit 6 of the ATPase in yeast. This 81 amino acid long polypeptide possesses a 66 amino acid presequence. The presequence is not only responsible for targeting the subunit to the mitochondrion, but it is also required for keeping this hydrophobic polypeptide in a soluble form during targeting.

It was intended to use sections of the presequence to assess the functions played by various parts of the polypeptide, using a PCR based approach to amplify sections of the cloned presequence. In the absence of such a clone, the presequence segments had to be made *de novo*, without the use of a template DNA, using synthetic oligonucleotides instead.

A long oligomer, up to 90 bp, can easily be synthesised by most custom DNA suppliers; DNA coding for presequence mutants could therefore be manufactured as a single stranded oligomer. Transforming this oligomer into double stranded DNA could be performed by annealing a short primer to the 3' end of the oligonucleotide and extending the primer using a DNA polymerase, such as *Taq* or Vent$_R$® (exo-), as shown in Fig. 5.1. This DNA could then be cloned into the *Nhe*I and *Hin*dIII sites of pSoUP, and subsequently used in transfection studies.

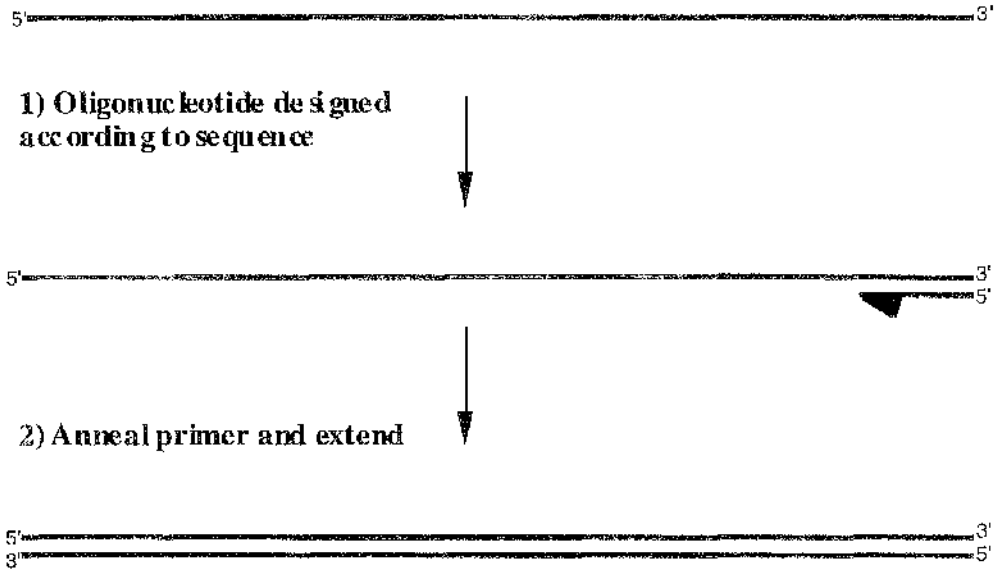


Figure 5.1 Schematic diagram showing how a single stranded oligomer could be made double stranded

An oligomer is made which is identical to the sequence being tested. A primer is also made which is the complement of the 3' end of the oligomer. The primer is then extended by a polymerase.

5.2 Analysis of the primary structure of PDC-E2 presequence

Sequence analysis of the presequence segments was also carried out prior to the transfection studies. These sequence analyses were designed to detect the presence of any mitochondrial targeting motifs suitable for further study.

A comparison of the targeting presequence of PDC-E2 with previously characterised presequences was made. Sequences targeting proteins to the mitochondrion are often 20-30 amino acids long and are frequently shown to have an amphiphilic nature when folded into an α-helix, with polar residues (especially hydroxylated and positively charged residues) on one side of the helix and non-polar or hydrophobic residues on the other. They also have a tendency to contain alanine, arginine, leucine and serine, and a strong bias against aspartate and glutamate (Rusch and Kendall, 1995). By comparison, the sequence for the targeting signal of PDC-E2 is 73 amino acids long, with no overall amphiphilic nature and containing an aspartate residue. It does however contain a large number of leucine, serine and arginine residues, especially in the C-terminal region of the sequence.

A number of different programmes were employed to find any sub-domains which had the features typical of a targeting polypeptide within the presequence. Two programmes were used: GCG on UNIX, from the Genetics Computer Group (Genetic Computer Group, Inc., Madison, Wisc., USA), is a piece of DNA/protein sequence analysis software. GCG was used to generate alpha-helical wheel plots, which can reveal amphiphilic nature over short stretches of sequence (Fig. 5.2). Sybyl, from Tripos, Inc (St. Louis, Mo., USA), on the Silicon Graphics Indy workstation, is a programme for modelling and analysing the 3-D structures of molecules. Sybyl was used to generate 3-D models of the presequence, forced into an α-helix, to give further insights into the polypeptide (Fig. 5.3), such as spiralling ridges of polar

molecules, which can be associated with the C-terminal 10-12 amino acids of presequences (Heijne *et al.*, 1989).

From the figures, it can be seen that the PDC-E2 presequence is not an 'ideal' mitochondrial signal, having no overall amphiphilic nature. However, in the N-terminal and C-terminal thirds of the polypeptide, there is some similarity to a mitochondrial signal, with these regions possessing many of the features normally associated with such sequences. The centre of the E2 presequence, between 2 proline rich regions, appears to have less amphiphilicity. This is in comparison to the presequences of cytochrome c-oxidase subunit IV and OGDC-E2. In cytochrome c-oxidase subunit IV, all the charged residues cluster together to give a defined amphiphilic helix. In the OGDC-E2 presequence, the N-terminal region of the sequence has considerable similarity to a mitochondrial targeting signal. The C-terminal region of that sequence, however, does not have a recognisable mitochondrial targeting motif.

To assess whether the N-terminal region of the PDC-E2 presequence was sufficient for targeting to the mitochondrion, it was decided to analyse the first 22 amino acids. The first few amino acids, being the first released into cytosol from the ribosome, should logically contain the mitochondrial targeting signal (Hurt *et al.*, 1985). The last 24 amino acids were also selected for testing since they too formed a convicing amphiphilic helix, and were therefore considered as interesting to study. The lengths of both sequences match the regions where the helix appears to have a dual nature.

Figure 5.2 Alpha-helical wheel plots of E2 presequence segments

Alpha-helical wheel plots were generated by the GCG programme on a UNIX computer based on the whole presequence and segments of it. The angle between each residue is taken as 100°. The programme highlights 'hydrophobic' residues in grey boxes. Charged residues have also been illustrated to aid identification of targeting motifs (circled positive and negative symbols). Glycine, interpreted by GCG as hydrophobic, has been assigned no status in these diagrams, since glycine is a flexible and amphiphilic residue.

The alpha-helical wheel plot of the entire PDC-E2 presequence (not shown) revealed no obvious overall characteristics; plots were subsequently performed on thirds of the presequence, shown in Fig. 5.2b. These plots are compared to (a) a typical mitochondrial presequence and (c) the OGDC-E2 presequence.

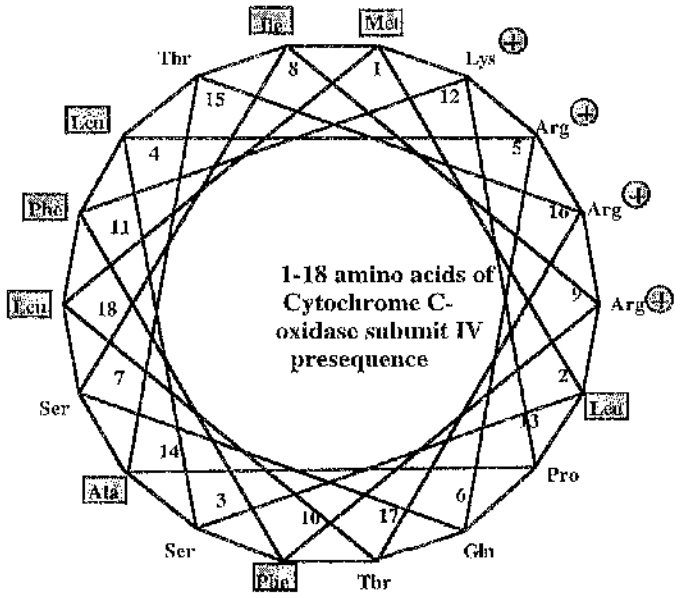
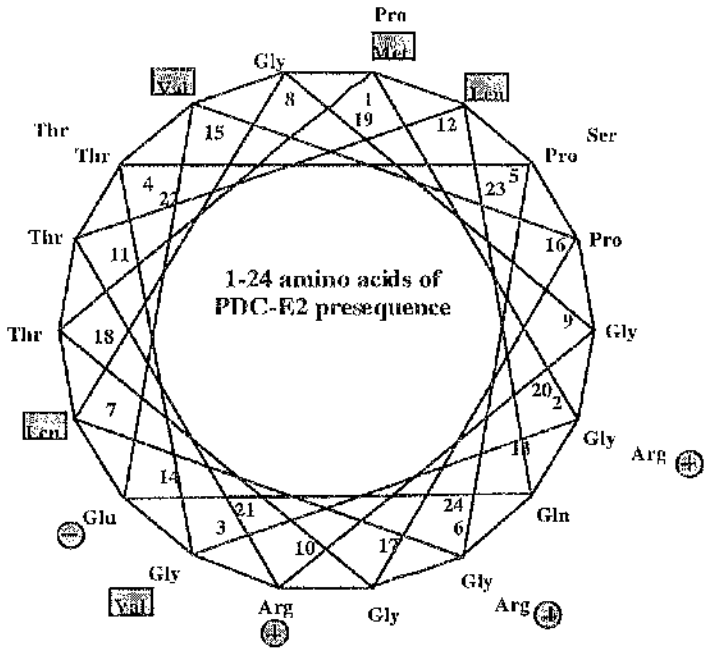


Figure 5.2a Alpha-helical wheel plot of a typical mitochondrial presequence: the cytochrome c-oxidase subunit IV presequence

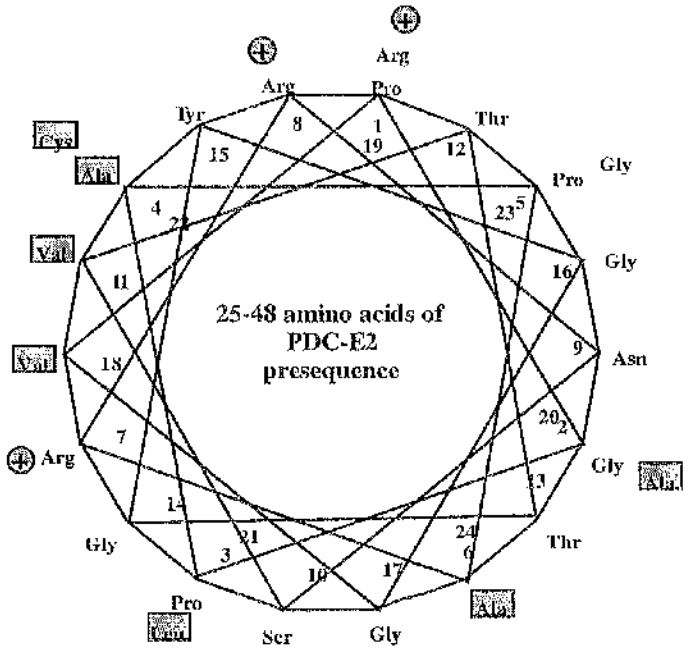
The first 18 amino acids of the presequence are shown as an example of an 'ideal' mitochondrial presequence. Although polar residues are scattered throughout the hydrophobic section of the helix, the charged residues are clustered in a single region, giving the helix a typical amphiphilic nature. It should be noted that the first methionine in the sequence will carry a positive charge, though GCG recognises all methionine residues as hydrophobic.

Figure 5.2b Alpha-helical wheel plots of PDC-E2 presequence segments



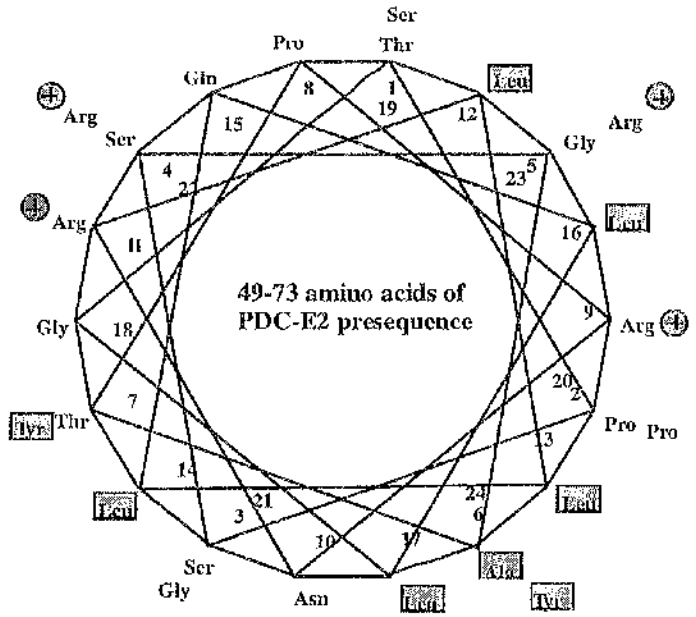
N-terminal 1-24 amino acids

The plot reveals poor amphiphilicity, with a large number of positively charged residues oposite a small number of hydrophobic residues. In the first 18 amino acids (inner circle) there is clustering of the charged residues towards the bottom of the diagram. This may indicate the presence of a targeting signal, although the first methionine, as mentioned, will carry a positive charge and disrupt this pattern. The negatively charged glutamate is atypical of mitochondrial presequences.



Amino acids 24-48

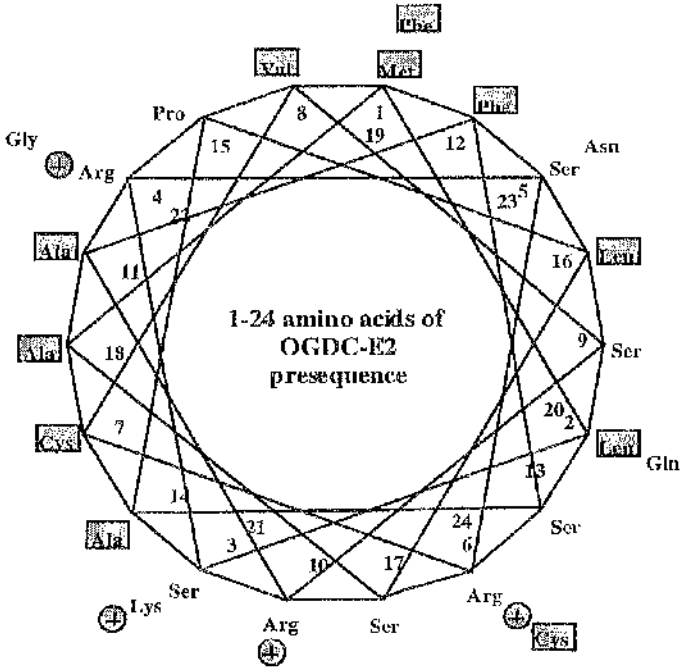
The plot shows a partly hydrophobic sequence of amino acids, with some scattered polar residues and a polar strip made from tyrosine, arginine, and threonine. There appears to be a very slight amphiphilic tendency to this section of the presequence.



Amino acids 49-73

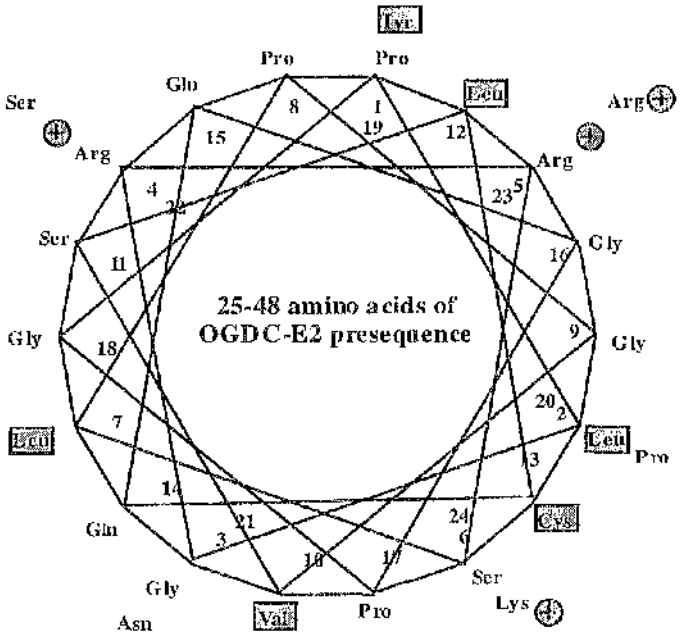
The plot reveals a number of arginine residues scattered over one half of a largely hydrophobic helix. This then has a slightly amphiphilic nature, and may be worth investigation as a putative matrix targeting signal.

Figure 5.2c Alpha-helical wheel plots of OGDC-E2 presequence segments



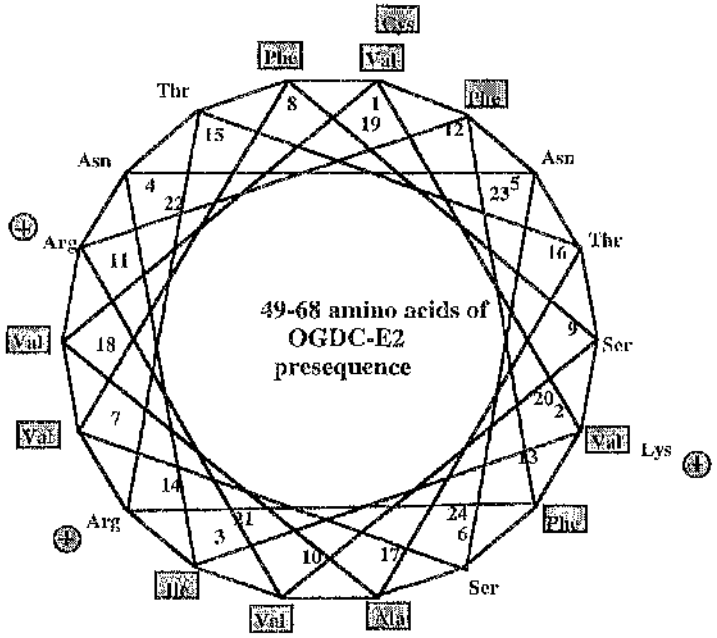
N-terminal 1-24 amino acids

The alpha-helical wheel plot for the first 24 amino acids of the OGDC-E2 presequence shows more similarity to the typical mitochondrial presequence than the corresponding amino acids in PDC-E2, with the majority of charged residues clustered in a small part of the helix. Note again however that the first methionine, which is charged, falls outside this cluster of charged residues.



Amino acids 25-48

The middle portion of this sequence has a large number of charged residues scattered among a number of hydrophobic amino acids. The amphiphilic nature noted in the first 24 amino acids is not present within this section.



Amino acids 49-68

This final part to the OGDC-E2 presequence also does not appear to be amphiphilic, since it has a number of charged arginine residues scattered throughout a largely hydrophobic helix.

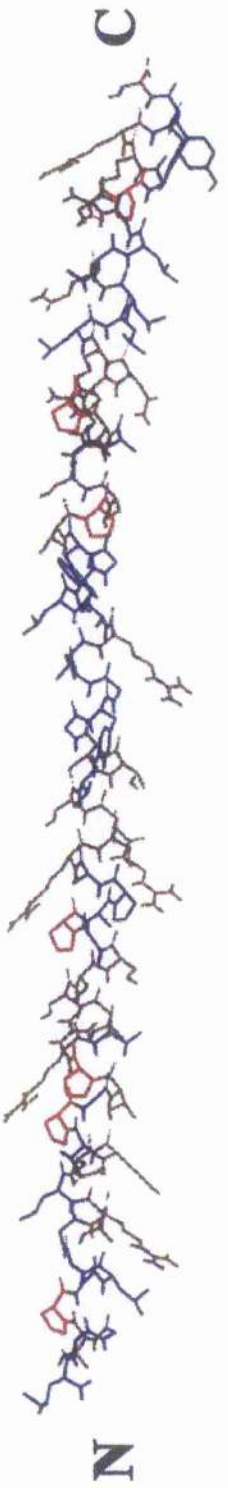

C
N

Figure 5.3 Molecular model of PDC E2 presequence

The amino acid sequence for the PDC-E2 presequence was used by the Sybyl molecular modeling software to generate this drawing. The sequence was forced into an alpha-helix, with proline residues breaking the structure, to search for targeting motifs. Residues were coloured brown for hydrophilic and blue for hydrophobic, using the same definitions for hydrophobic/hydrophilic used in the GCG programme. Proline is coloured red. Glycine has been treated here as hydrophobic. No part of the sequence exhibits strong amphiphilicity.

The illustration shows that the N-teminal region displays the properties of a potential targeting motif for the first 16 amino acids. The helix then changes from a hydrophobic helix with a hydrophilic strip to a hydrophilic helix with a hydrophobic strip for a single turn.

The middle segment does not have the amphiphilicity that the alpha-helical wheel plot suggested, here being drawn as a sequence of hydrophobic and hydrophilic segments.

The C-terminal region has 2 segments of amphiphilic helix according to the model, the first 6 amino acids are largely hydrophilic, the last segment of the presequence also forms a helix with a reasonably convincing dual nature.

5.3 Method for making presequence segments

Each section of the presequence was to be made from single stranded oligomers as described. The double stranded DNA would then be blunt end ligated into pCRScript™ for amplification of the fragment. It could then be excised by restriction digestion, purified, and finally cloned into pSoUP (Fig. 5.4).

5.4 Making the DNA for the N-terminal segment of the PDC-E2 presequence in pSoUP

The DNA corresponding to the N-terminal 22 amino acids was made first, to ensure the feasibility of the technique before continuing with the design of further constructs. The DNA fragment corresponding to the N-terminal 22 amino acids was named 'E2n'.

5.4.1 Making E2n

An oligonucleotide was made coding for the first 22 amino acids of the PDC-E2 presequence, however, an extra 2 amino acids was introduced directly after the first methionine by the addition of an *Nhe*I restriction site for cloning into pSoUP. The *Hin*dIII site included for cloning into pSoUP would not affect the presequence, and was already present within pSoUP. The oligonucleotides made to generate this DNA can be seen in Fig. 5.5.

To make the DNA double stranded, 40 µg primer was added to 11.6 µg 81 bp oligonucleotide. This gave an excess of primer, ensuring that all the 81 bp oligonucleotide was converted into double stranded DNA. The DNA was diluted to 50 µl total, in Vent$_R$® (exo-) buffer, and was heated to denature at 75°C for 1 min. The DNA was then annealed by cooling to 30°C over a 15

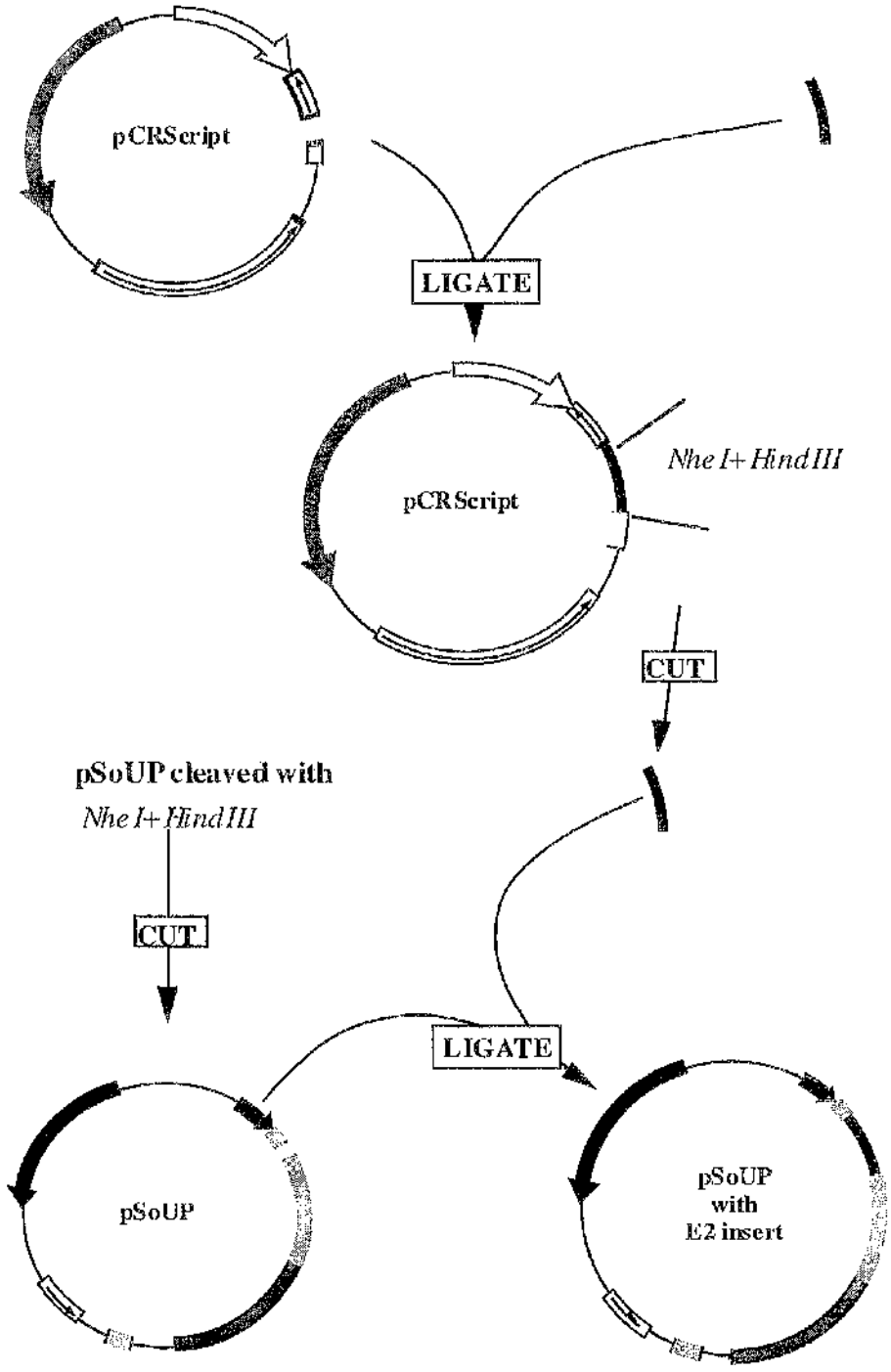

pCRScript
LIGATE
pCRScript
Nhe I+ Hind III
CUT
pSoUP cleaved with
Nhe I+ Hind III
CUT
LIGATE
pSoUP
pSoUP
with
E2 insert

Figure 5.4 Using pCRScript to amplify oligomer and allow sticky end ligation into pSoUP

After making a single stranded oligomer double stranded, it can be blunt end ligated into pCRScript. Once cloned, the oligomer can be sequenced and amplified in the plasmid. In this way, large amounts of the DNA can be made, and the oligomer can be excised from the plasmid using *Nhe*I and *Hin*dIII, ready for ligation into the expression vector, pSoUP.

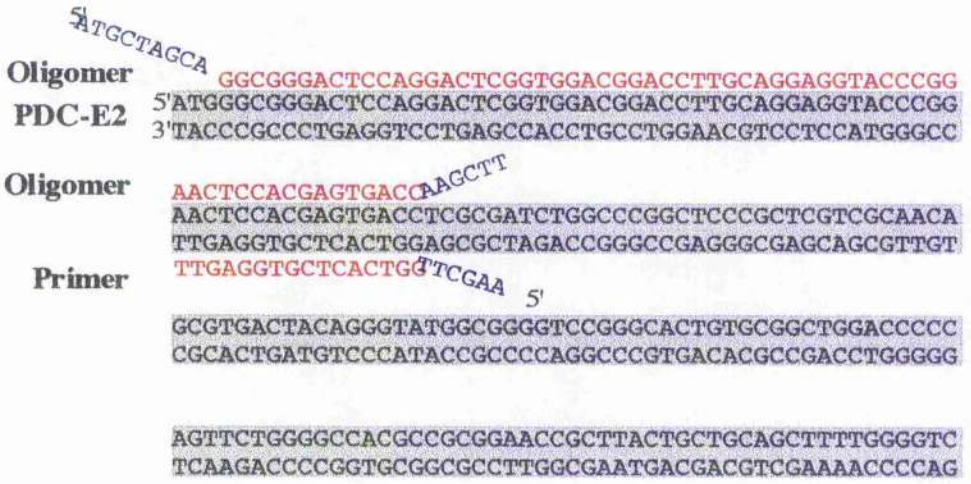


Figure 5.5 Design of oligomer DNA and primer to make first PDC E2 presequence segment

The first 200 bases of the DNA coding for the PDC E2 presequence are shown, along with the 2 oligomers designed to make the presequence segment which would code for the first 21 amino acids (after the initial methionine), as described in Section 5.1. The PDC-E2 presequence is shaded in grey. The oligomer and primer are shown in red, with the additional restriction sites coloured blue.

15 min period, after which 12.5 mM dNTPs and 2 µl $Vent_R^{®}$ (exo-) polymerase were added to start polymerisation of the second strand. To ensure the polymerase performed adequately, the reaction was slowly warmed over 10 min to 50°C, and left to incubate for 30 min.

Gel electrophoresis was employed to detect any shift in mobility as a result of changing the piece of DNA from single stranded to double stranded. The polymerase treated DNA was compared with the original oligomer on a 10% (w/v) acrylamide gel (Fig. 5.6). Such analysis is capable of demonstrating whether or not the DNA had been altered.

5.4.2 Cloning and analysis of E2n in pCRScript™

The E2n DNA fragment was cloned into pCRScript™, according to the manufacturer's protocol. A large number of recombinants were made, 4 of which were tested for the presence of an insert. Wizard™ Minipreps were made of each of the 4 selected clones. These plasmids (5 µg) were then digested for 2 h with 20 units *Nhe*I and *Hin*dIII at 37°C. The results are shown in Fig. 5.7. The 3 clones which were found to have inserts were used to make single stranded DNA, which was sequenced with the Sequenase kit using M13 -40 primer. One of the clones which was found to have the correct sequence was taken forward for cloning into pSoUP whilst the others were discarded.

5.4.3 Cloning of E2n into pSoUP

A fresh Wizard™ Miniprep was made of the E2n clone in pCRScript™. The E2n clone (5 µg) was digested in 50 µl *Nhe*I buffer with 20 units *Nhe*I overnight at 37°C, to ensure digestion. To this was added 20 µl *Hin*dIII buffer, along with 20 units *Hin*dIII. The reaction was then incubated for 3 h at 37°C. The excised fragment was isolated and purified, using the

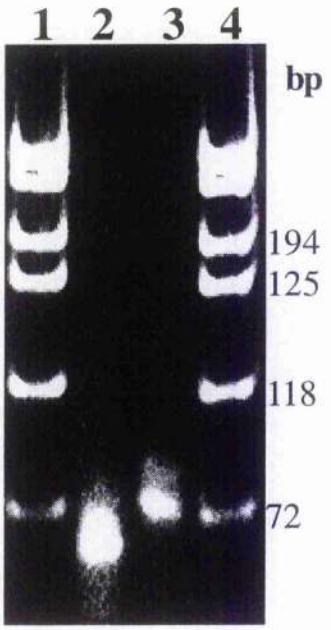


Figure 5.6 Acrylamide gel analysis of double stranded oligomer

The new double stranded oligomer was compared to the original single stranded oligomer on a 10% (w/v) acrylamide gel. The gel shows that the oligomer has lost mobility, an indication that it has become double stranded.

1) 'DRIgest III' markers (λ *Hin*dIII digest and ϕ *Hae*III digest)

2) Single stranded oligomer

3) Double stranded oligomer

4) 'DRIgest III' markers (λ *Hin*dIII digest and ϕ *Hae*III digest)

MERmaid® kit. The DNA was quantified to be at 900 ng in 30 µl.

pSoUP (12 µg) which had been purified by density centrifugation was digested with 50 units *NheI* for 6 h (to ensure digestion) at 37°C in a total of 50 µl *Nhe*I buffer. This was diluted with 50 µl *Hin*dIII buffer, and redigested with 50 units *Hin*dIII for 6 h at 37°C. Full digestion of the plasmid was checked by ligating 50 ng for 5 h with 1 unit T4 DNA ligase, and transforming into supercompetent cells (supplied by Stratagene). A low background was obtained, presumably from religation of a small amount of the excised linker.

It was decided to ligate the E2n DNA fragment into the cut vector without further purification of the vector away from its own linker DNA. Excised E2n DNA (80 ng) was ligated to pSoUP (60 ng), an 80x excess of insert, using 1 unit ligase into 10 µl total volume. The reaction was incubated at room temperature for 4 h, and transformed into supercompetent cells. Of the colonies obtained, 5 were tested for the presence of an E2n-pSoUP construct, called 'pE2nSoUP'.

5.4.4 Analysis of pE2nSoUP

Plasmids were prepared from the clones by Wizard™ Minipreps, and 2 µg of each digested with 10 units *Hin*dIII and *Nhe*I for 1 h at 37°C. The result, shown in Fig. 5.8, shows that 4 of 5 clones tested were suitable. The DNA was then manually sequenced with the Sequenase™ kit, utilising the PCR primers 1 and 4 (see Ch3) to sequence pSoUP in both directions from outside the cloning site. A clone was found with absolutely correct sequence, which was grown overnight and frozen down in glycerol to establish a permanent stock. With the procedure established, the second presequence segment could be made.

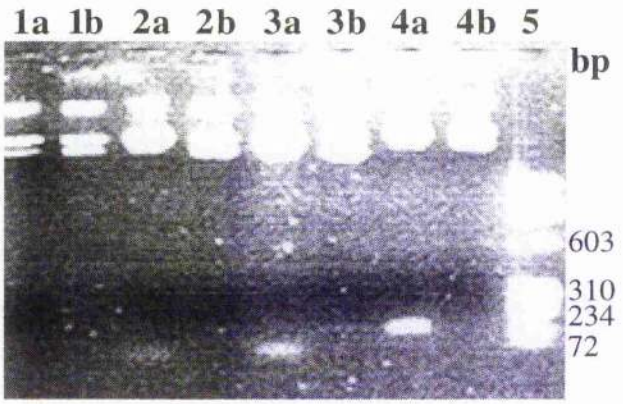


Figure 5.7 Agarose gel analysis of restriction digests of oligomer clones

Restriction digests of the 4 tested clones, analysed by electrophoresis through a 2% (w/v) agarose gel. Plasmids were digested with *NheI* and *Hin*dIII (a), or left undigested (b).

Lanes 1-4) The digested and undigested clones

5) 'DRIgest III' markers (λ *Hin*dIII digest and ϕ *Hae*III digest)

From the gel, it can be seen that there are 2 different types of ligation product. The clone in lanes 4a and b, gives a restriction fragment in lane 4a which is considered to be too large to be the correct ligation. The origins of this fragment are unknown; it is possible that the *Nhe*I digestion has not occurred, and a *Hin*dIII fragment has been excised.

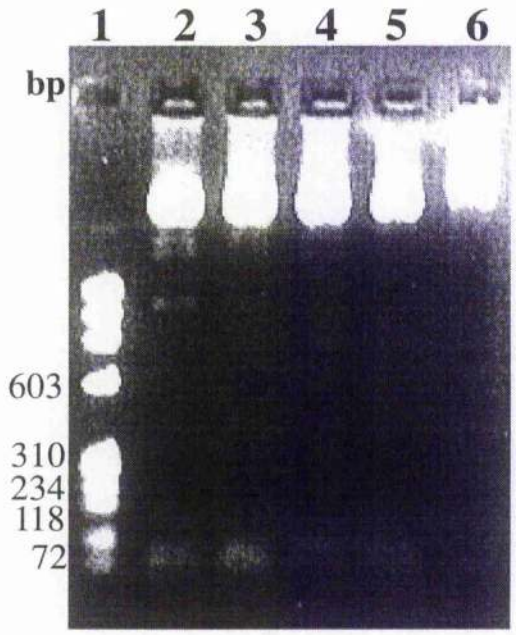


Fig 5.8 Agarose gel analysis of restriction digests of pE2n SoUP clones

Restriction digests of the 5 clones tested for the presence of an insert by digestion with *Hin*dIII and *Nhe*I were analysed by electrophoresis through a 2% (w/v) agarose gel.

1) 'DRIgest III' markers (λ *Hin*dIII digest and ϕ *Hae*III digest)

2-6) Digestions of 5 selected clones

From the gel it can be seen that 4 of the 5 clones yield the correct size of fragment upon digestion. Those clones which yielded a fragment were suitable for sequencing.

5.5 Making the second PDC-E2 presequence segment in pSoUP

5.5.1 Making the DNA coding for the C-terminal presequence segment

The last 24 amino acids in the presequence were selected for the second set of experiments. In addition to the 24 amino acids, 3 amino acids caused by the inclusion of an *Nhe*I linker sequence in the DNA would be encoded by the oligomer. In addition to this, DNA coding for a single serine at the 3' end of the presequence DNA was included. This serine, from the start of the mature PDC-E2, is the point at which the mitochondrial processing peptidase cleaves the presequence. Including it allows us to test whether this sequence is sufficient for signal cleavage. Again, a *Hin*dIII site is included for cloning into pSoUP, but does not form part of the presequence. This segment of DNA was called 'E2c'.

The oligonucleotide sequence desiged to code for this polypeptide is shown in Fig. 5.9, along with the primer used to make the single stranded DNA double stranded. The DNA was made double stranded as before, using 40 µg primer and 12 µg 87 bp oligonucleotide (an excess of primer). The DNA was diluted to 50 µl total, in Vent$_R$® (exo-) buffer, and was denatured at 75°C for 1min. The DNA was then annealed by cooling to 30°C over a 15 min period, after which 12.5 mM dNTPs and 2 µl Vent$_R$® (exo-) polymerase were added to start polymerisation of the second strand. To ensure the polymerase performed adequately, the reaction was slowly warmed over 10 min to 50°C, and left to incubate for 30 min. The DNA was electrophoresed on an acrylamide gel as before, but not all DNA had converted from the single stranded form to the double stranded. It was decided to proceed, and select for clones in pCRScript by sequencing.

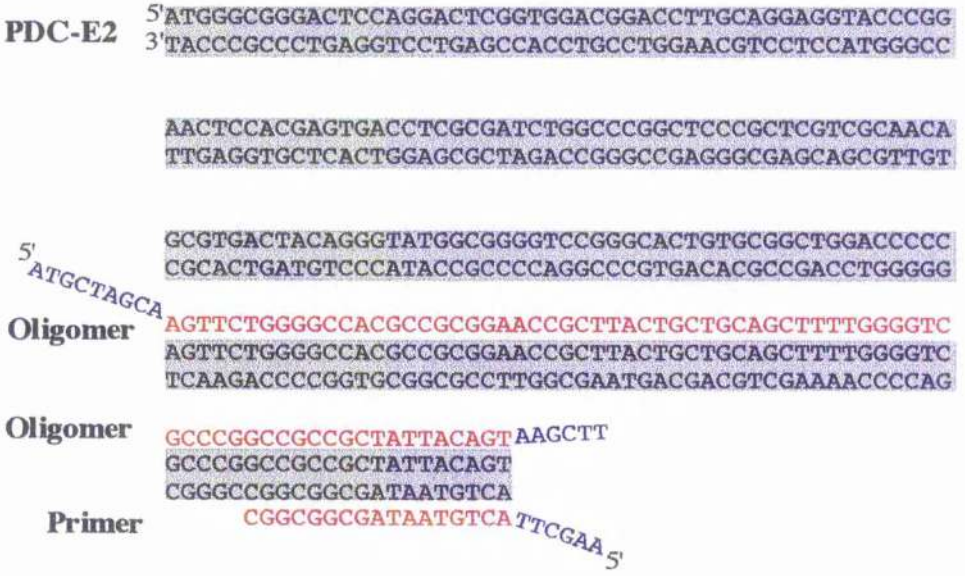


Figure 5.9 Design of oligomer DNA and primer to make second PDC E2 presequence segment

The DNA coding for the PDC-E2 presequence is shown, along with the 2 oligomers designed to make the second presequence segment. The mutant DNA codes for the last 24 amino acids of the PDC E2 presequence and the first serine of the mature PDC-E2. The PDC-E2 presequence is shaded in grey. The oligomer and primer are shown in red, with the additional restriction sites coloured blue.

5.5.2 Cloning and analysis of E2c in pCRScript™

The E2c DNA fragment was cloned into pCRScript™, as described previously. A large number of recombinants were found, 27 of which were tested for the presence of an insert. Wizard™ Minipreps were made of each of the selected clones, and digested with *Nhe*I and *Hin*dIII as before. Of the clones which were found to have inserts, 2 were used to make single stranded DNA, which was manually sequenced with the Sequenase kit using M13 -40 primer. Both clones were found to have the E2c insert, and one was taken forward for cloning into pSoUP.

5.5.3 Cloning of E2c into pSoUP

Plasmid was made from the E2c clone by Wizard™ Miniprep. The plasmid (5 µg) was then digested in 50 µl *Nhe*I buffer with 20 units *Nhe*I overnight at 37°C, to ensure complete digestion. To this was added 20 µl *Hin*dIII buffer, along with 20 units *Hin*dIII. This was incubated for 3 h at 37°C. The excised fragment was then isolated and purified, using the MERmaid® kit. pSoUP was also digested as before.

Excised E2c DNA (100 ng) was ligated to cut pSoUP (70 ng), using 1 unit ligase in 10 µl total volume. Again this ligation uses approx. 80x excess of insert to vector. The reaction was incubated at room temperature overnight, and transformed into supercompetant cells. A large number of colonies were obtained by this method, 10 of which were tested for the presence of an E2c-pSoUP construct, called 'pE2cSoUP'.

5.5.4 Analysis of pE2cSoUP

Wizard™ Minipreps were used to make plasmid DNA of the clones, which were digested with *Nhe*I and *Hin*dIII as before. Clones which posessed an insert were frozen down to form permanent stocks, and were then sequenced using the Sequenase kit, utilising the PCR primers 1 and 4 (see Ch. 3) to sequence pSoUP on either side of the cloning site. This did not give satisfactory results, so Wizard™ Maxipreps were carried out to obtain better quality DNA. A number of clones verified the presence of the desired insert, and one of these was sequenced fully, revealing absolutely correct sequence for the insert except for a single base deletion. This deletion was interpreted as a sequencing error since the missing base was in the region of DNA synthesised by the manufacturers of both the primer and the oligonucleotide, and could not have been introduced by the polymerase.

5.6 Discussion

Analysis of the presequence DNA revealed regions possessing targeting signal-like motifs and regions where the sequence was distinctly unlike a targeting signal. The N-terminus of the sequence has the capacity to form an amphiphilic helix, and bears some (though poor) similarity to typical targeting signals. Interestingly, the C-terminal region also has a similar motif, whilst the middle portion of the presequence has no such amphiphilicity nature. This is in contrast to the mitochondrial signal for the E2 subunit of OGDC, which had a more pronounced amphiphilic nature at the N-terminal region in Fig. 5.2c, and no recognised motif at the C-terminus, which is the expected nature of a targeting presequence.

The aim of this research was to determine the reason for and functions of the extended presequence on immature PDC-E2 polypeptides. This domain-

like structure to the presequence suggests that there is more than one function and that the targeting signal may be held in one part of the presequence whilst additional information is held elsewhere.

Chapter 6

Transfection of COS-7 Cells with pSoUP Constructs

6.1 Introduction

The aim of this research was to characterise the presequence of PDC E2, by elucidating whether the N-terminal region of the presequence carried all the information for targeting and antifolding for the E2 polypeptide. This would be decided based on the ability of the presequence to cause translocation of CAT to the mitochondrion. In addition, the nature of the C-terminal region, which had an amphiphilic nature similar to the N-terminal region, was to be tested for similar targeting properties. A fourth version of the plasmid was also tested in parallel with those described in Chapter 5; an oligonucleotide construct of the entire PDC-E2 presequence, made by Dr. L. Briand, was used in pSoUP to replace the cloned product which could not be obtained in Chapter 4. This plasmid would be used to test the properties of the native PDC-E2 presequence.

Transfection studies using COS-7 cells, a monkey kidney fibroblast cell line commonly used in such work, would reveal the behaviour of the CAT polypeptide constructs. The plasmids made in chapter 3 and 5 could be transfected into the cells and, once they had expressed the various chimaeric proteins, the cells would be harvested. These cells could subsequently be analysed to find the position and enzyme activity of CAT; the subcellular location of CAT would reveal the targeting properties of the presequence, and the enzyme activity would demonstrate whether or not the CAT trimer had been able to assemble.

To detect assembly of the CAT trimer, the CAT assay of Seed and Sheen (1988) was used, as detailed in the methods section (2.7.1.6). This assay can establish assembly since CAT is only active once it has folded correctly and formed a trimer. To discover the location of the CAT, cells were fractionated into various subcellular compartments using digitonin, and the expressed CAT detected immunologically. Digitonin is a useful detergent which interacts

specifically with cholesterol. Digitonin can therefore selectively disrupt membranes based on their cholesterol content. Since the plasma membrane of mammalian cells contains 17% cholesterol (liver cells) in comparison to mitochondria which have only 3% cholesterol (inner and outer membranes combined), the plasma membrane of cells such as COS-7 will be ruptured at a lower digitonin concentration than that required to disrupt mitochondria. This allows efficient fractionation of a small number of cells into cytosol and cytosol/mitochondrial matrix fractions without the need for density centrifugation, as is commonly used to isolate mitochondria.

It is worth pointing out that most mitochondrial targeting studies have been carried out on isolated mitochondria *in vitro*, or in yeast. Only recently have transfection technologies allowed the routine introduction of foreign genes into mammalian cells. The work of Boutry *et al.* (1987) is most relevant to this work, where CAT was linked to a 90 amino acid N-terminal segment from the β-subunit of the mitochondrial ATPase of *Nicotiana plumbaginifolia*. This construct was then introduced into and expressed in tobacco cells, where the chimaeric polypeptide was targeted to the mitochondrion. This demonstrates the suitability of CAT for mitochondrial targeting studies (at least in plants) and the ability of *in vivo* targeting studies to reveal the nature of presequences.

6.2 Fractionating COS-7 cells

For the transfection/fractionation experiments to proceed, it was necessary to optimise all the appropriate conditions for these procedures. A protocol for cell fractionation was initially devised, to ensure that the plasma membrane could be disrupted without rupturing the mitochondrial membranes. Following this, a suitable method for transfection of the plasmid was selected based on maximal expression. The three chosen methods were compared using

optimal conditions for each protocol.

6.2.1 Culturing and harvesting cells

COS-7 cells were cultured (in 60 mm transfection dishes) as described in the methods section, until confluence was reached. Cells were harvested as described, washed with sterile ice-cold PBS and scraped off the dish to be resuspended in 1-1.5 ml ice-cold PBS. The cells were then centrifuged for 5 min at 2500 rpm in 50 ml sterile plastic tubes, and resuspended to 10^6 cells/ml in PBS.

6.2.2 Fractionating cells

Cells (at 10^6 cells/ml) were incubated for 1 min at 4°C in the presence of a range of digitonin concentrations before centrifuging at high speed in a benchtop centrifuge to pellet the membrane, whole cells and unruptured organelles. The supernate was carefully removed, leaving the pellet undisturbed. Both were then stored at -20°C.

6.2.3 Identifying fractions by enzyme assay

By assaying for enzymes specific for the fractions of interest, it is possible to identify which subcellular compartments are present in each lysate by the presence or absence of an enzmye. The marker enzymes chosen were lactate dehydrogenase (LDH), which is found solely in the cytosol, citrate synthase, which is found only in the mitochondrion, and malate dehydrogenase (MDH), which is found in both the cytosol and mitochondrion. As the plasma membrane is solubilised, so LDH and MDH are released into solution. When more digitonin is used, the mitochondria are ruptured along

with the plasma membrane. In this case, mitochondrial MDH and citrate synthase are released along with the cytosolic MDH and LDH.

In Fig. 6.1 it can be seen that LDH and MDH are released at approx. 0.05 mg digitonin/10^6 cells. Citrate synthase and mitochondrial MDH are released at approx. 0.5 mg/10^6 cells. From this it was decided to use a range of digitonin concentrations between 0 and 1.0 mg/10^6 cells which ensured that the plasma membrane was completely solublised without any damage occurring to the mitochondrion (i.e. >0.05 mg/10^6 cells, but <0.5 mg/10^6 cells), and at least 1 fractionation carried out at a concentration of digitonin >0.5 mg/10^6 cells.

6.2.4 Identifying fractions by Western blots

Cell extracts (40 µl, 10^6 cells/ml) were subjected to SDS-PAGE, and transferred onto nitrocelluose as described. LDH and citrate synthase were then detected by ECL (Fig. 6.2), using rabbit polyclonal antisera raised 'in house'. The results confirmed the data generated by the enzyme assays, that cytosolic enzymes are released when digitonin concentrations exceed 0.05 mg/10^6 cells, and that the mitochondrial membranes are disrupted at 0.5 mg digitonin/10^6 cells.

6.2.5 Optimising transfection

Three transfection protocols were tested, following optimised protocols described for each method. Plasmid (pBLSV CAT) was purified using the Qiagen midi prep kit, and transfected into cells growing in 4x60 mm transfection dishes, using one of the three protocols. Two negative controls were included; 'plasmid only' was to monitor for background uptake of the plasmid by adding plasmid but no transfection reagent, and a negative control

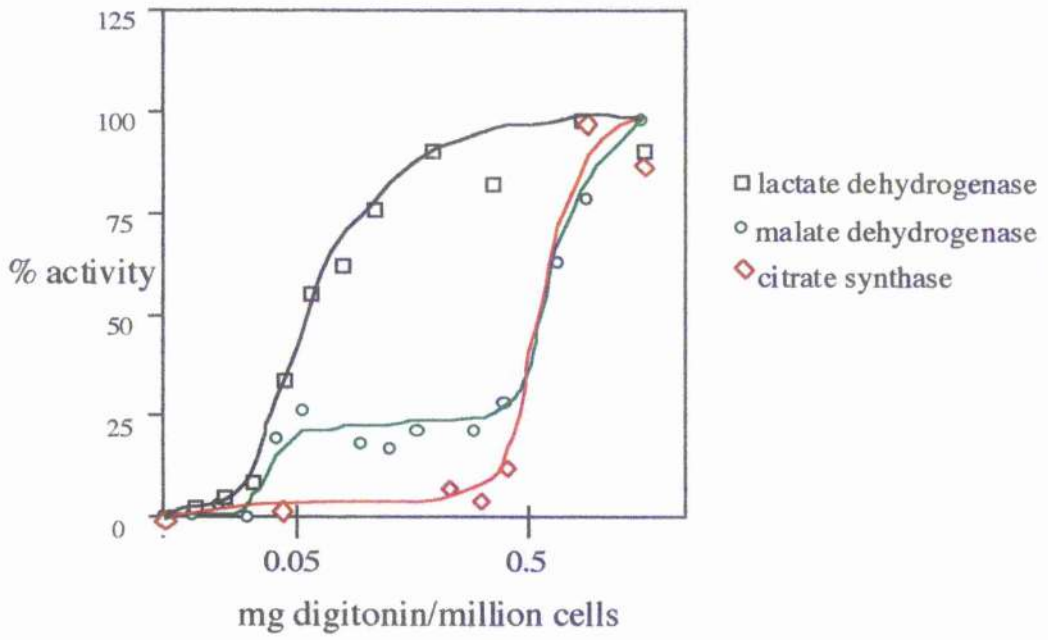


Figure 6.1 Fractionation of cells and release of marker enzymes

Graph showing release of various enzymes with increasing digitonin concentration.

It can be seen from the graph that cytosolic enzymes are released when digitonin concentrations exceed 0.05 mg/10^6 cells. Citrate synthase, and the mitochondrial malate dehydrogenase are found in the supernate once the mitochondrial membranes are disrupted at 0.5 mg digitonin/10^6 cells.

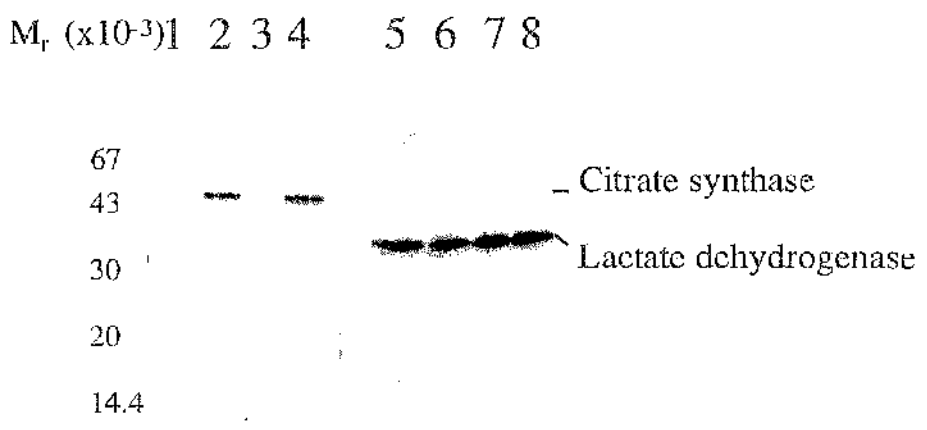


Figure 6.2. Western blot of fractionated cells to detect released enzymes

Fractionated cell extracts were electrophoresed through a 10% (w/v) SDS-polyacrylamide gel, and then transferred to nitrocellulose for ECL detection of lactic dehydrogenase and citrate synthase.

Lanes 1-4; Fractionated cells tested for the presence of citrate synthase

1 and 3) Cytosolic fraction only (0.05 mg digitonin/10^6 cells)

2 and 4) Cytosol and mitochondrial matrix (0.5 mg digitonin/10^6 cells)

Lanes 5-8; Fractionated cells tested for presence of LDH

5 and 7) Cytosolic fraction only (0.05 mg digitonin/10^6 cells)

6 and 8) Cytosol and mitochondrial matrix (0.5 mg digitonin/10^6 cells)

The results of the Western blot confirm that at 0.05 mg digitonin/10^6 cells, the plasma membrane is ruptured without mitochondria being lysed, but at 0.5 mg digitionin/10^6 cells, mitochondrial disruption has also taken place.

negative control with no plasmid and no reagent. Cells were harvested after 48 h, and diluted to 10^6 cells/ml. Cell extracts were then made by adding 1 mg/ml digitonin or 1% (w/v) Triton X-100.

The results are tabulated below (Table 6.1), and demonstrate that lipofection is the most efficient method for mammalian cell transfection.

Table 6.1

Transfection method used	**CAT activity (cpm/assay)**
Lipofection	23,074 +/- 597
DEAE-Dextran	7,936 +/- 2117
Calcium chloride	2,955 +/- 732
Plasmid only	505 +/- 125
Negative control	889 +/- 220

Both detergents were found to cause some difficulties in obtaining a clear phase separation during the CAT assay, particularly when using Triton X-100. It was therefore decided to use an additional aliquot which was given a freeze-thaw treatment in later experiments to avoid such complications.

6.2.6 Developing a standard protocol for cell transfection and fractionation

The enzyme assays and Western blotting analysis revealed that COS-7 cells could be fractionated into cytosol only and cytosol + mitochondrial matrix extracts. When digitonin was at a concentration of 0.05 mg/10^6 cells (in 1 ml PBS), only the plasma membrane was solubilised, but when the concentration of digitonin was 1 mg/10^6 cells (in 1 ml PBS), both the plasma membrane and the mitochondrial membranes were broken, releasing their

contents.

CAT assays demonstrated the efficiency of lipofection over calcium chloride and DEAE dextran treatments, using prescribed protocols for each method.

A standard protocol for transfection of cells, their harvesting and fractionation was therefore as follows: 5 µg pure DNA per dish transfected by lipofection and incubated with the lipofectin^{TM}/DNA overnight, after which time the medium is replaced and the cells left to incubate for 48 h. The cells are harvested in PBS, and fractionated with digitonin at 3 concentrations- to leave the cells intact, 0 mg/10^6 cells is used, to rupture the plasma membrane, 0.1 mg/10^6 cells guarantees full solubilisation without damage to the mitochondrion; and for full disruption, 1 mg/10^6 cells ensures complete solubilisation of both membranes. A fourth aliquot of cells was freeze-thaw treated to obtain detergent-free cell extracts for CAT assays.

6.3 Testing pSoUP against original pBLSV CAT

Both plasmids were purified by Qiagen midi preps, and used to transfect cells using the above standard procedure. CAT assays performed on freeze-thaw lysed cells demonstrated that CAT activity was unaffected by the mutation of 2 N-terminal amino acids in CAT expressed from pSoUP (Table 6.2).

Table 6.2

Plasmid tested	**CAT activity (cpm/assay)**
pBLSV CAT	166,299 +/- 3834
pSoUP	125,627 +/- 22372
negative control	572.5 +/- 28.5

Variability between pSoUP and pBLSV CAT activities is likely to be due to expression levels, caused by slight differences in quantity and quality of plasmid DNA transfected, or by random errors introduced by the experimental method.

A Western blot was then performed to discover if pSoUP also expressed a cytosolic CAT, as expected. Fig. 6.3 demonstrates that the location of CAT is also unaffected by the amino acid substitutions, with CAT released after plasma membrane lysis. Minimal CAT was detected associated with the membrane pellet after lysis of cells (not shown). The identity of the fractions was confirmed by the use of enzyme assays (section 6.2.3).

6.4 Testing the PDC E2 presequence/CAT gene constructs

The presequence-CAT gene constructs were then tested using the above protocols. The plasmids tested were pSoUP (CAT expression vector), pE2n SoUP (CAT expression vector with the DNA for the first 22 amino acids of the E2 presequence), pE2c SoUP (CAT expression vector with the DNA for the last 25 amino acids of the E2 presequence) and the full length E2 presequence construct, pE2 SoUP.

Fig. 6.4 shows the Western blot of the fractionated samples taken from the cells transfected with pBLSV CAT and pE2n SoUP. pE2c SoUP did not express any CAT antigen, and is not shown in the figure. Results for pSoUP

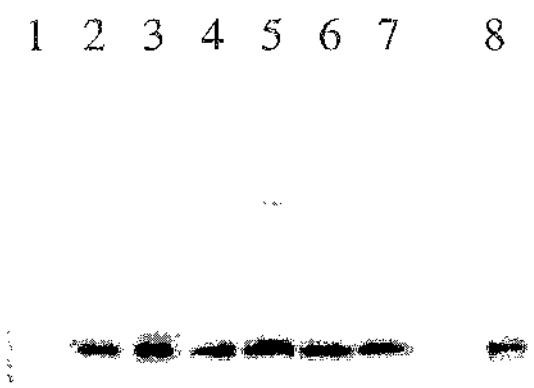


Fig 6.3 Fractionation of expressed CAT from pSoUP

Western blot of fractionated cells electrophoresed through a 10% (w/v) SDS-acrylamide gel.

1) 0 mg digitonin/10^6 cells

2) 0.02 mg digitonin/10^6 cells

3) 0.04 mg digitonin/10^6 cells

4) 0.1 mg digitonin/10^6 cells

5) 0.2 mg digitonin/10^6 cells

6) 0.4 mg digitonin/10^6 cells

7) 0.8 mg digitonin/10^6 cells

8) Purified CAT standard

The blot shows that CAT is detectable in the supernate once the digitonin concentration exceeds 0.02 mg/10^6 cells.

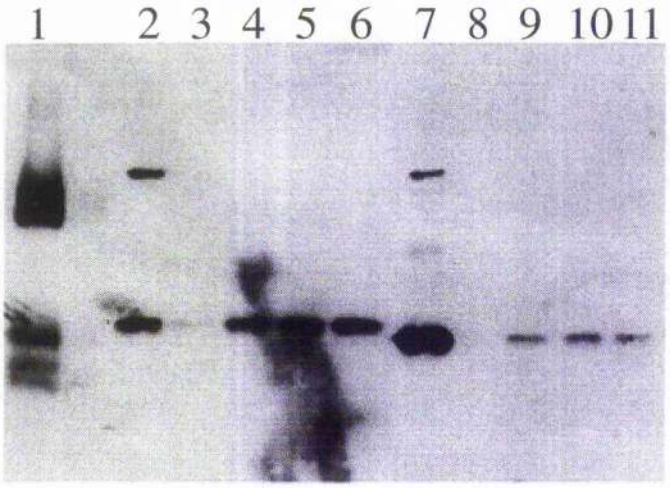


Figure 6.4 Fractionation of expressed pE2n SoUP and pBLSV CAT

Western blot of fractionated cells electrophoresed through a 10% (w/v) SDS-polyacrylamide gel.

1) Purified CAT standard

Lanes 2-6, cells expressing pE2n SoUP

2) Freeze-thaw lysed cells

3) 0 mg digitonin/10^6 cells

4) 0.1 mg digitonin/10^6 cells

5) 0.25 mg digitonin/10^6 cells

6) 1 mg digitonin/10^6 cells

Lanes 7-11, cells expressing pBLSV CAT

7) Freeze-thaw lysed cells

8) 0 mg digitonin/10^6 cells

Figure 6.4 cont'd

9) 0.1 mg digitonin/10^6 cells

10) 0.25 mg digitonin/10^6 cells

11) 1 mg digitonin/10^6 cells

This demonstrates that CAT expressed by pE2n SoUP is a cytosolic enzyme (released at 0.1 mg digitonin/10^6 cells), as is pBLSV CAT and pSoUP (Section 6.3). It can also be seen that the CAT expressed by pE2n SoUP is of a higher subunit M_r value than that expressed by pBLSV CAT, as expected with the presence of the presequence on the CAT.

were similar to pBLSV CAT (not shown). Further Western blot data on the behaviour of pE2 SoUP from Loïc Briand revealed a higher M_r CAT antigen present in the cytosolic fractions after fractionation of the cells. It appeared therefore that CAT expressed from all plasmids was cytosolic , except pE2c SoUP, which was not detected.

Shifts in mobility in CAT can be attributed to the presence of the presequence. CAT assays performed on freeze-thaw lysed cells demonstrated that CAT activity was unaffected by the presence of any of the presequences, except pE2c SoUP, which had no detectable activity. These results are tabulated below (Table 6.3).

Table 6.3

Plasmid tested	**CAT activity (counts per minute)**
pE2 SoUP	22,728 +/- 2274
pE2n SoUP	146,688 +/- 13976
pE2c SoUP	513 +/- 21
pSoUP	125,627 +/- 22372
pBLSV CAT	166,299 +/- 3834
negative control	572.5 +/- 28.5

Variability in expression levels, caused by slight differences in quantity and quality of plasmid DNA transfected, has again caused the differences seen here between those transfections which can be seen from ECL to be expressing CAT. It is worth noting that the low activity seen in the full length presequence-CAT construct (pE2 SoUP) was reflected in Western blots (not shown) where the levels of expressed CAT were low.

Since pE2c SoUP had not expressed any CAT, it was decided to repeat the transfection of COS-7 with freshly prepared pE2c SoUP, made by Wizard™ Maxipreps. This transfection also failed to express CAT, as detected by immunological and CAT activity assays. Since the plasmid was unable to express any CAT, it was considered that what had been regarded as a sequencing error in pE2c SoUP could have been a genuine base deletion in the region of the synthesised oligonucleotide. Sequence analysis was repeated on the plasmid. The results confirmed that a base was missing, and that any expressed polypeptide would terminate prematurely due to a frame shift mutation.

6.5 Discussion

In this chapter, the presequence segments integrated into the plasmid pSoUP were expressed in COS-7 cells to test the ability of the presequences to halt folding and intiate targeting of CAT to the mitochondrion.

It was found that the N-terminal section of the presequence was insufficient to prevent folding of CAT, or to trigger its translocation to the mitochondrion. Furthermore, it was found that a base deletion had occurred in the DNA for the C-terminal segment of the presequence, causing a frame shit in the rest of CAT, which invalidated that part of the experiment. This frame shift was confirmed by the absence of any activity in any cell extracts in a number of separate transfections.

The plasmid constructed by L. Briand, which contained the entire presequence in pSoUP, gave some surprising results. Expression of CAT was poor and was untargeted, a fact confirmed by the observation that the presequence was not cleaved from CAT when analysed by SDS-PAGE analysis (not shown). Table 6.3 confirms that there was low expression of active CAT from pE2 CAT. This observation also demonstrates that the

presequence did not prevent CAT from folding properly, since a significant level of CAT activity was present, despite the presequence being attached.

Of note are other experiments carried out in this laboratory which demonstrate the antifolding activity of the BCDC-E2 presequence. Expression of mature BCDC-E2 in *E. coli* produces folded, active cores. However, N-terminal segments of the BCDC-E2 presequence as short as 17 amino acids prevented assembly, and led to degredation of the polypeptide. This raises a number of issues; firstly, the difference in behaviour between the BCDC-E2 expression studies and the PDC-E2/CAT targeting studies would suggest that CAT in mammalian systems is not an accurate model for E2, since folding is not prevented and import is not initiated. Secondly, it demonstrates that E2 presequences have antifolding capabilities in heterologous systems, whether it be by directly preventing the rest of the polypeptide from reaching a native conformation, or by interacting with the chaperones of the host bacterium.

It is disappointing that the PDC-E2 presequence was unable to target CAT. A number of reasons why the sequence failed to target CAT exist. It is possible that the presequence interacts with the lipoyl domain in PDC-E2; since this is lacking from CAT, the behaviour of the presequence would not be the same. Alternatively, it may be that the assembly of CAT has prevented uptake. A polypeptide has been found which has both a peroxisomal targeting signal and a weak mitochondrial targeting signal. This polypeptide was translocated to peroxisomes when present as a dimer, but imported into the mitochondrion when a point mutation inhibited dimerisation (reviewed by Danpure, 1997). These discoveries highlight the importance of the assembly status of a polypeptide on its ablility to target to the mitochondrion.

Further work to elucidate the role of the presequence should be based around a repetition of the experiments with BCDC-E2 expression in *E. coli*, using instead PDC-E2. Investigation into the interaction of the presequence with the lipoyl domain may also aid understanding of this system.

Chapter 7

Determining the Behaviour of PDC Kinase by Electrospray Mass Spectroscopy of PDC E1α

7.1 Introduction

PDC is a key enzyme in glucose homeostasis because it catalyses the committed step in carbohydrate metabolism linking glycolysis to the citric acid cycle, thus governing the flux of 2 carbon units into the cycle. Since PDC represents such a pivotal point in metabolism, it is regulated by a number of factors, including hormonal signals and the end products of the reaction. Regulation is primarily achieved by four isoforms of a dedicated complex-bound kinase which inhibits the enzyme by phosphorylation; a specific phosphatase reactivates it. The kinase and phosphatase exert their control over the E1 component, which catalyses the first, irreversible, rate limiting step, so that by inhibiting this component the entire reaction sequence is prevented.

Kinase activity is controlled by a number of metabolites which act as positive or negative effectors; pyruvate, ADP, CoA and NAD^+ inhibit phosphorylation while ATP, acetyl CoA and NADH are positive effectors. In response to activation, the kinase abolishes almost all PDC activity by phosphorylating up to 3 different serine residues on E1α (Fig. 7.1). Recent site directed mutagenesis studies (Nemerya *et al.*, 1996) have shown that the phospho-serine residues generated by the kinase are likely to exert their effect by competing with pyruvate for the active site arginine residues. The 3 serines undergo phosphorylation at different rates, site 1 being phosphorylated more rapidly than site 2 and site 2 more rapidly than site 3. Each site is sufficient, when phosphorylated, to inactivate E1 without the other 2 sites being phosphorylated (Korotchkina *et al.*, 1996). It has also been shown, through ^{32}P incorporation (Yeaman *et al.*, 1978 and Korotchkina *et al.*, 1996), that only 1 phosphorylation event per E1 tetramer needs to occur for inactivation of the entire tetramer, suggesting an element of positive cooperativity between subunits. Cooperativity between subunits in the tetramer has already been proposed for other aspects of the enzyme's activity, such as the alternating site

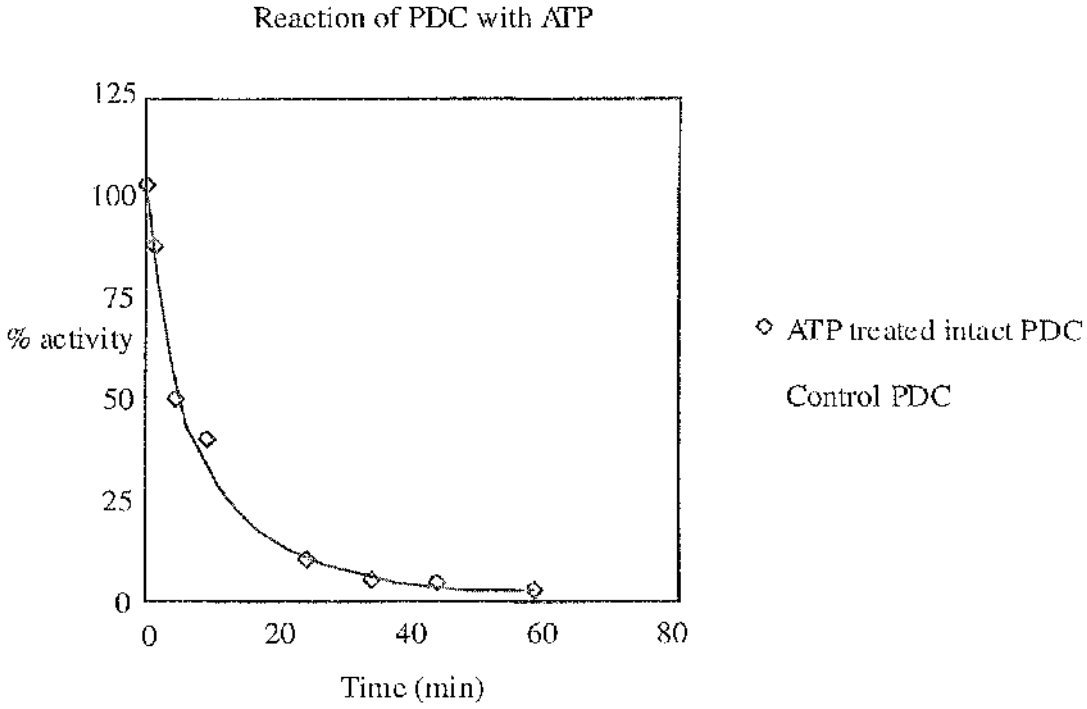


Figure 7.1 Effect of ATP on PDC activity

ATP (4 mM) was added to PDC in kinase buffer (1 mg/ml, methods section 2.7.6). PDC assays (methods section 2.7.1.4) at regular intervals show the loss of activity due to phosphorylation. After 20 min, only 10% activity is left, and minimal activity is left after 60 min.

demonstrated by Khailova *et al.* (1989).

A number of related PDC kinase genes have been found (currently 4 in the human genome numbered PDCK 1-4, Rowles *et al.*, 1996) which all have sequences similar to the prokaryotic protein histidine kinase family, and lack motifs usually associated with the eukaryotic serine/threonine protein kinases (Popov *et al.*, 1993). The kinase isoenzymes possess varying characteristics, with differing specific activities and K_i values for synthetic analogues of its substrates (Bowker-Kinley *et al.*, 1997). The isoenzymes are found in varying quantities in different tissues (e.g. PDCK3 is found exclusively in heart and skeletal muscle, Gudi *et al.*, 1995), indicating that the subtle differences in the functions of each kinase are important to the metabolism of those tissues. Work by Harris *et al.* (1997) has demonstrated that PDCK 4 levels increase in the heart of rats when starved, and is restored to normal levels when the rat is fed again, whereas PDCK 1 and 2 levels are unaffected by the treatments.

The prokaryotic protein histidine kinases are known to be dimers which correlates well with existing evidence suggesting that PDC kinase is also a dimer. This would explain how 3-5 kinase molecules can inactivate 60 E1 subunits without leaving the complex. PDC kinase subunits are able to bind the inner lipoyl domain of PDC-E2, and Liu *et al.* (1995b) have demonstrated that the kinase repeatedly releases and binds lipoyl domains. The kinase dimer is therefore thought to move over the complex by swinging "hand over hand", alternating between states where the kinase is bound to 1 and 2 lipoyl domains. An active kinase molecule could therefore deactivate the entire complex very quickly, moving over the surface of the E2 core, gaining access to all the bound E1.

The focus of this research was on the nature of the multiple phosphorylations. Whilst all 3 phosphorylations take place when both the E1 and the kinase are bound to the complex, this is not the case when both subunits are free in solution. When kinase and E1 are freed from the complex,

only the first, most rapidly phosphorylated serine, is modified by the kinase. The question to be asked, therefore, is this: is attachment of the kinase, E1 or both kinase and E1 to the E2 core necessary for the multiple phosphorylation of E1?

To help in answering this question, it is possible to utilise a discovery made in this laboratory regarding the lipoyl domains. Digesting intact PDC with the enzyme collagenase leads to the cleavage of E2 at the linker region between the inner lipoyl domain and the E1 binding domain (Fig. 7.2a,b), giving 2 lighter fragments termed $E2_L$ (the lipoyl domains) and $E2_I$ (the inner domains). Collagenase does this by cleaving a tetrapeptide (Pro Leu Gly Pro), unique in PDC-E2, a sequence which occurs frequently in the structural protein, collagen. The result of cleaving E2 at this point is a loss of PDC activity due to the loss of the E2 lipoyl groups (Fig. 7.2c). A residual activity of 10%-15% is retained by the complex since the lipoyl domains of protein X are unaffected, and are also able to participate in the reaction. This cleavage also releases the kinase molecules from the complex.

As a consequence of being in solution rather than being attached to the complex, the kinase inactivates PDC 10-20 times more slowly. In this research, it was intended to compare free with bound kinase, quantifying how many phosphorylations took place throughout the inactivation curve of PDC while E1 remained bound. In this way, we could discover whether the multiple phosphorylations were promoted by the kinase being localised to the complex, or whether this pattern was solely due to the attachment of the E1 component to core of PDC.

To discover how many phosphorylation events have taken place in each experiment, a recently developed approach was adopted, using the power of electrospray mass spectrometry (ESMS) to distinguish between the species. Electrospray mass spectrometry permits the accurate determination of the mass of large polypeptides, which could not be determined using a more

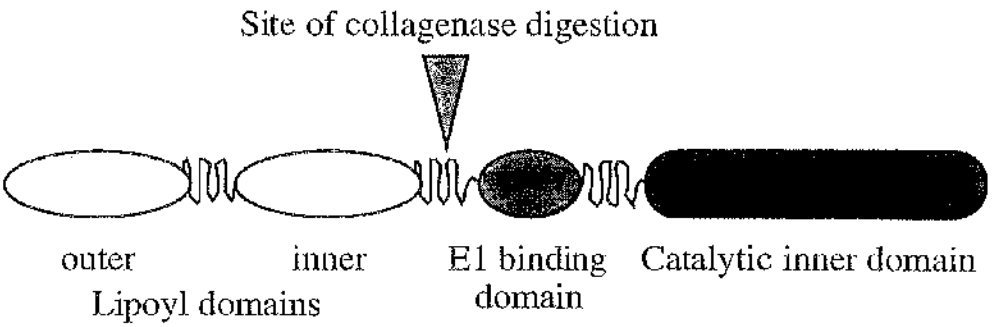


Figure 7.2a Domain diagram of E2 and site of collagenase digestion

A diagramatic representation of E2 is shown, highlighting the domain structure. The catalytic, core forming domain is C-terminal, the outer lipoyl domain is N-terminal. The site where collagenase cleaves the lipoyl domains from the E2 core is shown by the arrow head.

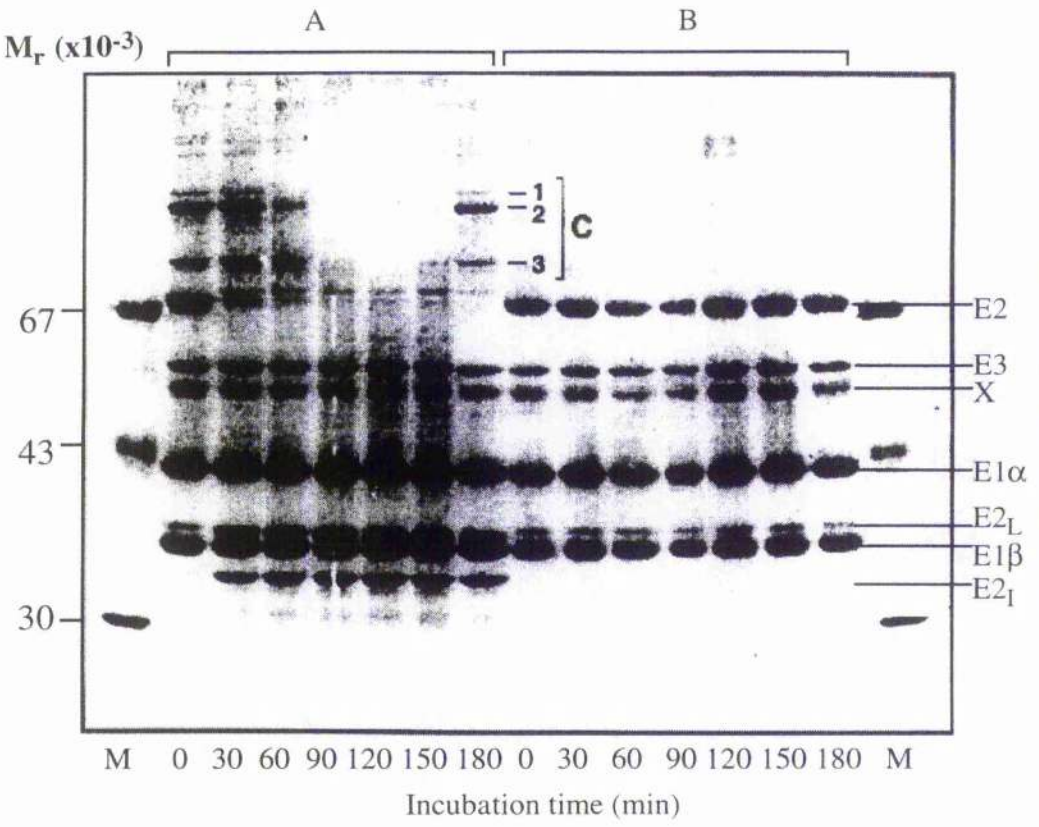


Figure 7.2b SDS-PAGE analysis of collagenase digested PDC

SDS-PAGE (10% (v/v)), shows the digestion of PDC by collagenase over time (A) compared to untreated control PDC (B). The collagenase digests E2, yielding lighter $E2_L$ (E2 lipoyl domains) and $E2_I$ (Inner core and E3 binding domains) polypeptide fragments. It can be seen that digestion is almost complete by 60 min. It is worth pointing out that the bands in the undigested 'B' lanes showing up as $E2_L$ are degredation products. Whilst $E2_L$ is cross-reactive with E2 antisera in immunoblots, these bands in the undigested material do not show up on coomassie stained gels (data not shown).

M=M_r markers C = collagenase bands.

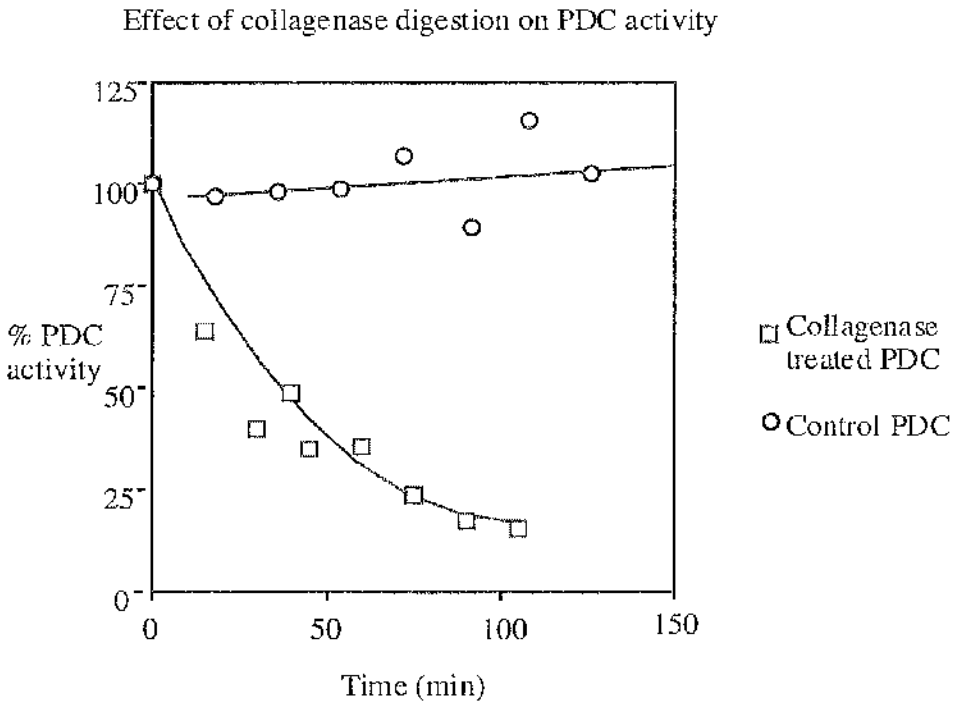


Figure 7.2c Graph showing effect of collagenase on PDC activity

PDC was incubated with collagenase as described, and PDC assays taken at intervals to reveal the loss of activity due to the loss of lipoyl domains. After full digestion, approx. 15% activity remains.

conventional mass spectrometer. In brief, the ESMS generates small, ionised droplets of solute carrying the polypeptides. The solvent surrounding each polypeptide is rapidly dried off ('desorption'), leaving the molecule in the gas phase, in a charged state. The ionised peptide then enters the mass spectrometer, where the mass to charge ratio of the molecule is measured (Fig. 7.3). The measurement data generated are finally processed by computer to determine the mass of the polypeptide (for a review, see Mann and Wilm, 1995). This technique is sensitive and precise enough to detect the small percentage change in mass that a single phosphorylation can produce in a large polypeptide and distinguish between multiple phosphorylated forms of subunits in a mixed sample (Nairn *et al.*, 1995).

In addition to this area of research, it was also hoped to confirm the results of Yeaman *et al.* (1978) and Korotchkina *et al.* (1995), demonstrating negative cooperativity between subunits using ESMS. If only 1 phosphorylation was required per tetramer, then ESMS would reveal an equal number of phosphorylated and unphosphorylated E1α subunits in PDC which was completely inactivated.

7.2 Devising a protocol for ESMS analysis of PDC-E1

Since PDC preparations were found to contain too many species on which to conduct ESMS analysis directly, a protocol had to be found to obtain purer E1 preparations which would allow such analysis.

7.2.1 ESMS directly on PDC E1/E3 fraction

It was first decided to isolate the E1/E3 fraction of PDC, using a high salt treatment followed by purification of the separated enzymes using FPLC gel filtration, as described in the methods section (2.6.3). Absorbance readings

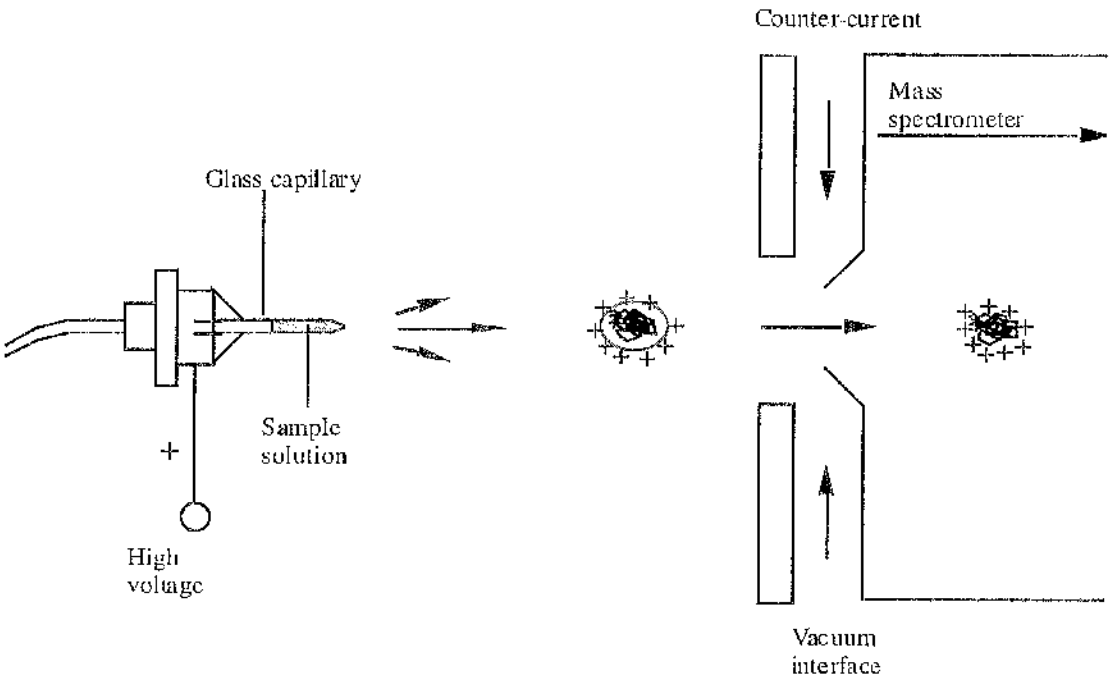


Figure 7.3 The 3 steps in electrospray mass spectrometry

Diagram taken from Mann and Wilm, (1995). Firstly, the solution is passed through a needle which is kept at a high voltage. The solution is then dispersed from the glass capillary into a mist of highly charged droplets. The droplets are then dried using a hot counter-current of hot N_2 gas, which releases the protonated protein molecules into the gas phase. The change from a liquid to gas phase leaves the protein intact, with at least one solvation shell. The droplets then enter a vacuum in the mass spectrometer, and are separated and detected according to their mass/charge ratio.

measured at 280 nm were used to identify those fractions containing E2 with protein X and E1 with E3 (Fig. 7.4). SDS-PAGE revealed which polypeptides were in each fraction (Fig. 7.5). During the salt treatment it was noted that a small amount of protein precipitated. From 7 mg PDC, approx. 2.5 mg of E1/E3 was obtained in a total volume of 13 ml (70% recovery of E1/E3 from the column).

The isolated E1/E3 fraction was injected directly into the ESMS to discover if further purification would be necessary, or whether the ESMS would be able to detect the various species in the mixture. Using recommended amounts of protein (approx. 100 μg, enough to guarantee injecting 20 μg E1α) in 50%(v/v) acetonitrile, 0.1%(v/v) TFA, the sample was injected into the ESMS at 0.02 ml min^{-1}. It was discovered that a large number of species were present in the sample, and it was felt that the data could not be trusted to give accurate mass data for any polypeptides present within the sample. It is not known if the sample was contaminated heavily or if the E1 and E3 components had been precipitated and lost, reducing their concentrations to the level of the contaminants.

From SDS-PAGE analysis, it was noted that some contaminants were present in the preparation. Based on the premise that contaminants had caused the difficulties, it was decided individual species would have to be isolated before injection into the ESMS.

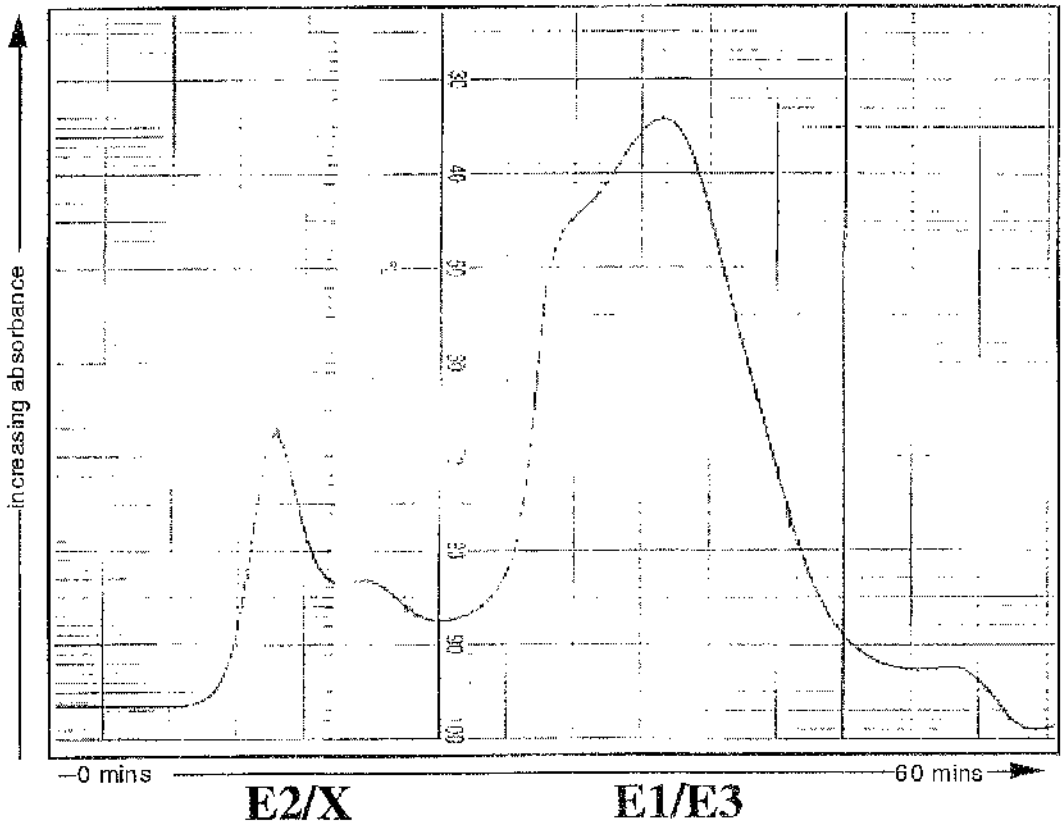


Figure 7.4 Separation of E2/X and E1/E3 by FPLC

The absorbance trace, taken at 280 nm, of the eluate from the Superose column (as described in methods chapter 2.6.3).

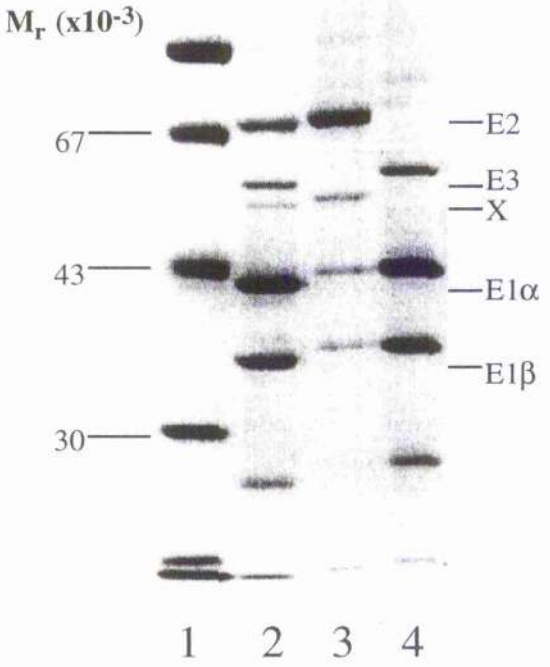


Figure 7.5 SDS-PAGE analysis of fractions obtained from FPLC

Fractions obtained from FPLC gel filtration of dissociated PDC were analysed by SDS-PAGE (10% acrylamide), and stained with Coomassie blue.

1) M_r markers

2) Native PDC

3) E2/protein X pooled fractions.

4) E1/E3 pooled fractions.

The gel shows that the E1/E3 is free of contamination with E2 or X. Only Sp-22 (an anti-oxidant protein), associated with PDC, contaminates the fraction.

7.2.2 Isolation of E1α by HPLC

Since it was apparent that E1α would have to be isolated from other components, it was decided to use ESMS analysis on polypeptide fractions as they were eluted from a reverse phase HPLC column; a technique suggested by Dr. Tino Krell. For this method approx. 200 µg of E1/E3 would be required.

HPLC was assessed for its ability to separate the various polypeptides. After dissociating PDC as described above, the E1/E3 fraction was dialysed into HPLC grade dH_2O by ultrafiltration from buffer into dH_2O. Washing the sample 3 times with HPLC grade dH_2O was considered sufficient to remove all traces of salt. During the cycles of washing and concentrating the E1/E3 fraction, precipitation of the polypeptides was noticed.

An HPLC column was pre-equilibrated with buffer A (98:2 dH_2O:acetonitrile, 0.1% (v/v) TFA), washed with 100% buffer B (10:90 dH_2O:acetonitrile, 0.1% (v/v) TFA), and re-equilibrated in buffer A. The E1/E3 (200 µg in 500µg of dH_2O) was injected onto the column, and buffer (flowing through the column at 0.4 ml min^{-1}) was changed from 100% A to 100% B over the period of 1 h. The absorbance of the eluate (at 215 nm) was measured over time to detect the elution of the various polypeptides (Fig. 7.6). The resolution between the absorbance peaks was deemed sufficient for isolation of each unknown species prior to ESMS analysis.

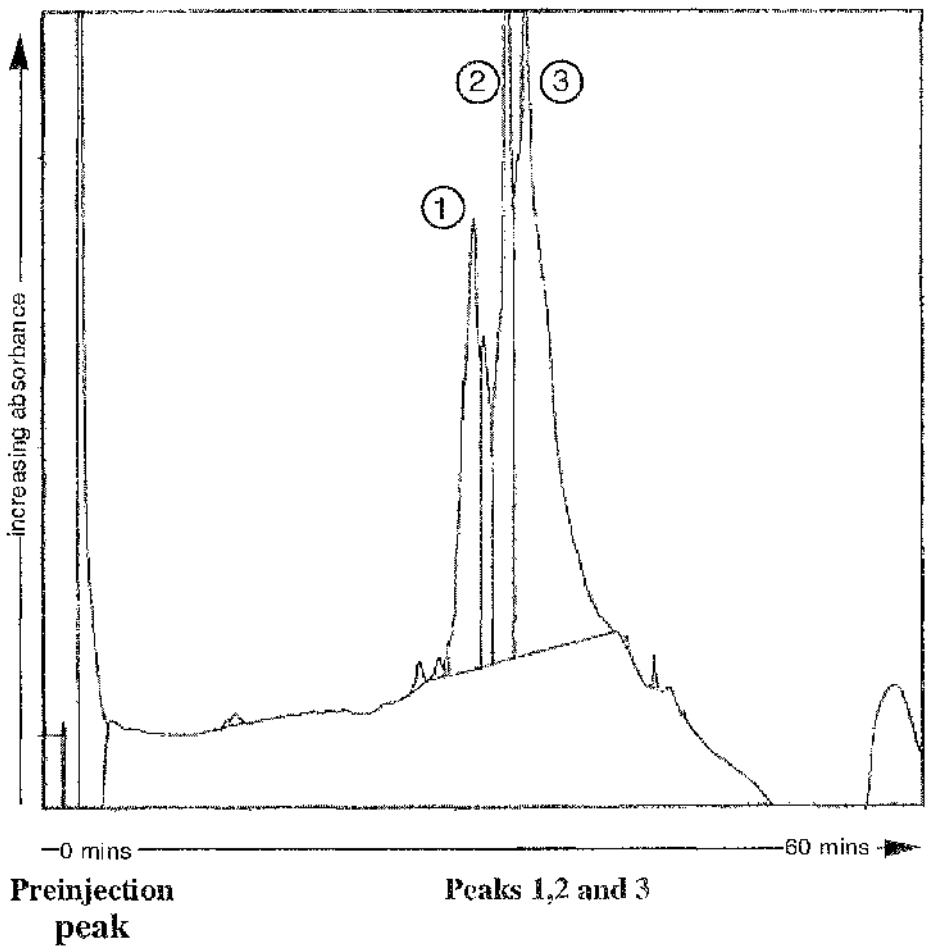


Figure 7.6 Separation of E1α, E1β and E3 by HPLC

The absorbance measurements read at 215 nm of the eluate coming from the reverse phase HPLC column. Peaks of absorbance were numbered 1-3, and later identified by ESMS, comparing the obtained data with mass data for human PDC equivalents. The identity of each peak is discussed in section 7.2.3

7.2.3 Analysis of E1α, E1β and E3 by ESMS

Since the isolation of each protein species (and therefore E1α) from purified E1/E3 was apparently possible, the ability to resolve the various phosphorylated forms of PDC by ESMS was checked before proceeding with the collagenase experiments.

This presented an early opportunity to answer the latter question, whether or not 1 phosphorylation per tetramer was sufficient for complete inactivation of PDC, since phosphorylated PDC was required for this procedure anyway.

Phosphorylation of PDC was carried out; this was performed at a higher PDC concentration than normal for purification of sufficient E1/E3, to compensate for losses expected during their dissociation/purification and concentration, and because only a small volume could be injected onto the FPLC column. PDC at 14 mg/ml was diluted 1:1 with kinase buffer. ATP (0.4 mM) was added to the PDC and the inactivation monitored over time using PDC assays, with 1 ml aliquots (7 mg) removed at various points during the inactivation (Fig. 7.7). Treatment was stopped in each aliquot by the addition of 5 mM EDTA; samples were then stored at -20°C. Each aliquot of treated PDC (including an untreated negative control) was dissociated and the E1/E3 fraction isolated by FPLC, as described. It was noted that slight precipitation of the protein occurred during salt denaturation in each of the aliquots. Once the E1/E3 fractions from each aliquot had been isolated, they were washed and concentrated into 100 µl HPLC grade dH_2O, as previously. This procedure too caused precipitation of the polypeptides. The yeild of the E1/E3 isolated from each aliquot (as determined by BCA assay) was between 78 µg (time=0 control PDC) to 190 µg (half inactivated PDC). As mentioned, 200 µg was the recommended amount for HPLC/ESMS. This meant that 94-98% (for time=0

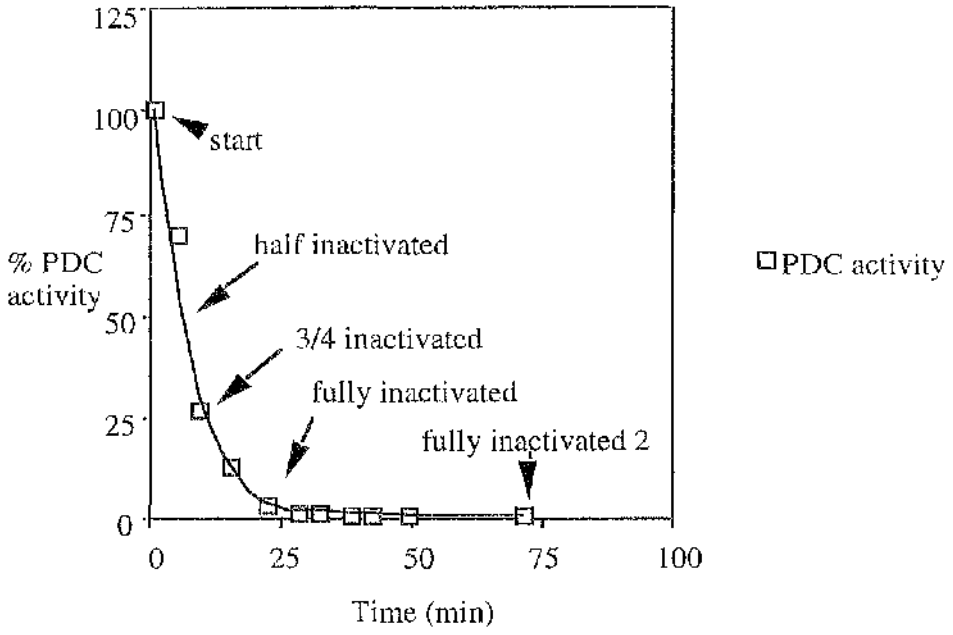


Figure 7.7 Effect of ATP on PDC activity

PDC was treated as described, and PDC assays carried out to monitor loss of activity with time. At 4 points on the graph (indicated by the arrow heads) an aliquot of PDC was removed and the treatment halted by chelating Mg^{2+} with EDTA. The aliquot was then frozen for later analysis.

and half inactivated PDC, respectively) of E1 and E3 had precipitated from the original 7 mg PDC, or was otherwise lost from the start of the PDC dissociation until the final isolation of the two components.

HPLC was performed on each of the samples as previously. The HPLC was connected to the ESMS, allowing the eluate from the HPLC to be injected directly into the ESMS for real-time analysis of the polypeptides eluted from the column. Unfortunately it was found that the flow rate used to elute the polypeptides from the HPLC was too high for the ESMS. With high flow rates, the ESMS is unable to dry off all the solvent before it enters the mass spectrometer. The result was that the mass spectrometer could not detect any polypeptides above the background noise generated by the solvent.

It was decided to separate the remaining samples on the HPLC as before, collecting peak fractions to be injected later into the ESMS at a lower flow rate. Fractions (100 µl) containing the peak concentrations of separated polypeptides were injected into the ESMS at 0.02 ml min^{-1}, in contrast to the previous rate of 0.3 ml min^{-1}.

Such were the losses at each stage of protein purification that most fractions collected produced no ESMS data. However, the ATP treated PDC which was removed at the half point on the inactivation curve (Fig. 7.7) was in sufficient quantity for data to be collected. The data were processed by subtracting background noise and smoothing the graphed data, using the MassLynx software. The software could then be used to calculate approximate mass values for the ions detected (by manually detecting components after peak centering), before using the Maximum Entropy deconvolution procedure (stopped after 8 iterations) to obtain more accurate data on the same ions. The data obtained are shown in Fig. 7.8. From this data, identification of the polypeptides could be made, based on the known masses for the human PDC components.

Figure 7.8 Graphs generated by the Maximum Entropy programme

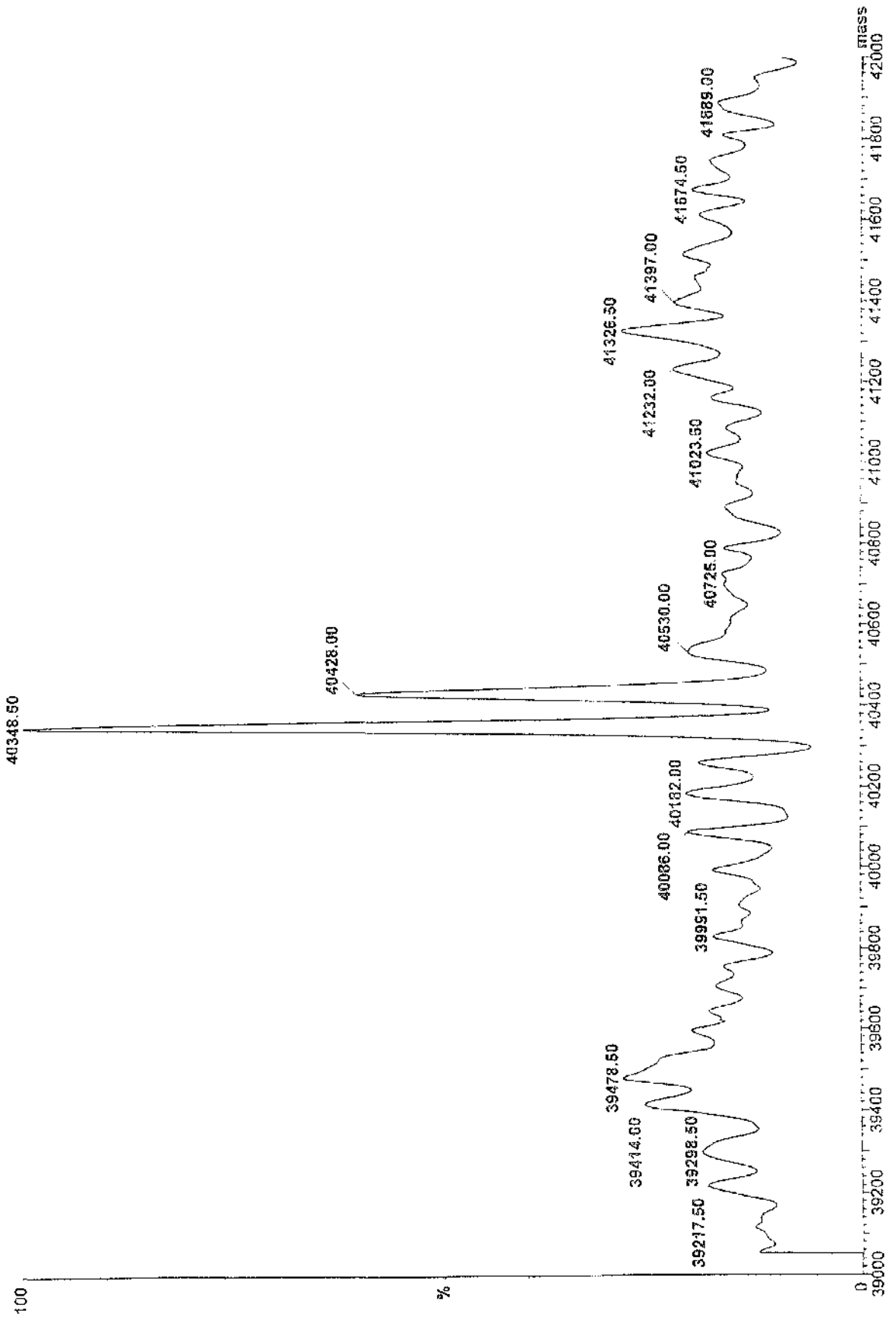
The data for the 3 fractions isolated in Fig. 7.8 were processed by the Maximum Entropy programme. The 3 graphs generated show the presence of ions with masses of:

Peak 1) 40348.5 and 40428 units, corresponding to E1α and singly phosphorylated E1α

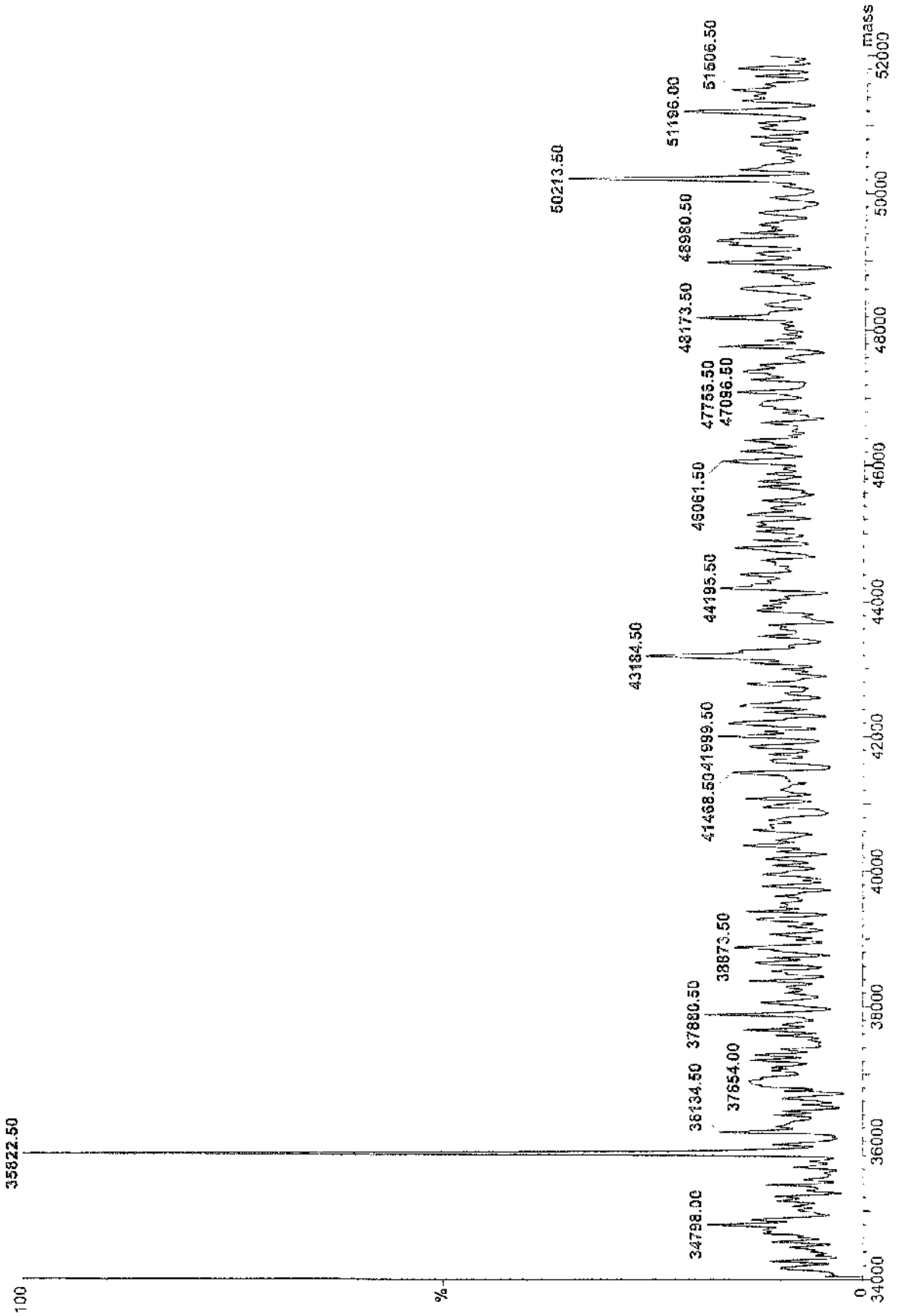
Peak 2) 35822 and 50213.5 units, corresponding to E1β and E3

peak 3) 35809 units, corresponding to E1β

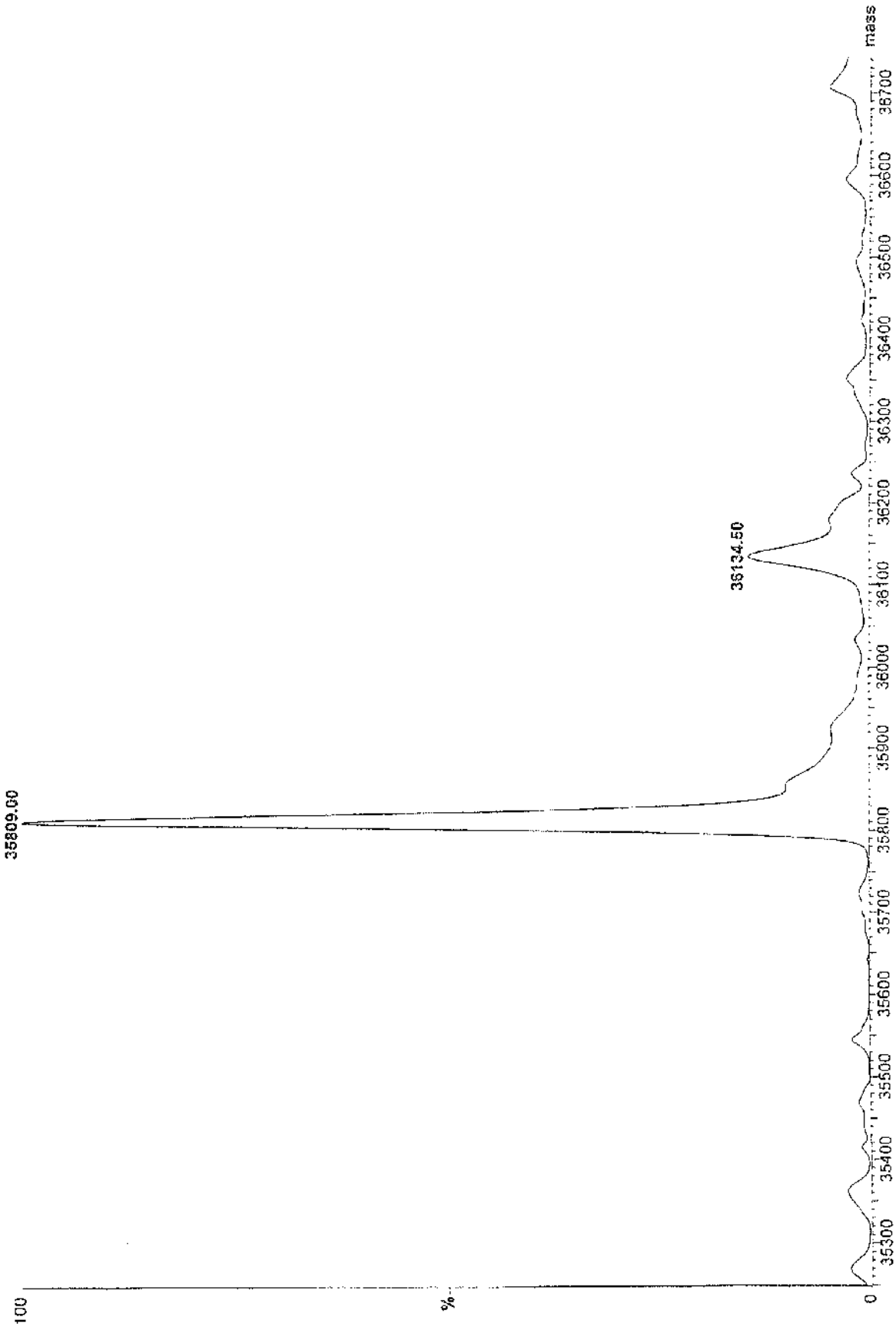
E1-E3 Peak 1


40348.50
40428.00
40530.00
40725.00
41023.50
41232.00
41326.50
41397.00
41674.50
41685.00
40182.00
40086.00
39951.50
39478.50
39414.00
39298.50
39217.50
100
%
0
39000
39200
39400
39600
39800
40000
40200
40400
40600
40800
41000
41200
41400
41600
41800
42000
Mass

E1-E3 Peak 2


35822.50
36134.50
37654.00
37880.50
38873.50
41468.50
41999.50
43184.50
44195.50
46061.50
47096.50
47755.50
48173.50
48980.50
50213.50
51196.00
51506.50
34798.00
100
%
0
34000
36000
38000
40000
42000
44000
46000
48000
50000
52000
mass

E1-E3 Peak 3


35809.00
36134.50
100
%
0
35300
35400
35500
35600
35700
35800
35900
36000
36100
36200
36300
36400
36500
36600
36700
mass

By comparison with Fig 7.6 it can be seen that E1α elutes in peak (1), E3 in peak (2), E1β in peak (3). E1β is also found in peak (2), possibly due to the presence of two allelic forms of E1β which each elute in different peaks. Alternatively, the difference between the 2 E1βs found could be caused by an event such as deamidation. The masses for bovine E1α, E1β and E3 are therefore E1α = 40348.5 and 40428.0 ; E1β = 35809 and 35822.5 ; E3 = 50213.5 mass units. These compare favourably with masses calculated from cloned human genes (E1α = 40183, E1β = 35863 and E3 = 50216). In the peak (1) fraction, a peak at 40428 represents the singly phosphorylated E1α (40348.5 + 79.5 units, the mass of phosphate being 78 units). The viability of the technique is therefore demonstrated, but no further conclusions can be drawn from the experiment since the E1α of interest, from the fully inactivated PDC, was not in sufficient quantity for mass spectrometery.

7.3 Using less PDC to avoid precipitation

It was clear that precipitation was caused by using PDC at such high concentrations. An attempt was made to carry out the purification at lower concentrations (3 mg/ml). Fractions (1.2 mg) were separated as before, using 24 ml bed volume Superose 12 column (1x30 cm) instead, to minimise losses on the column. Fractions of E1/E3 were collected, washed in dH_2O and concentrated as before. At this point a BCA assay was used to measure the amount of protein present. It was found that only 50 µg of protein was isolated, which was not enough to continue. Whilst using less protein was more efficient in terms of %yield, it did not yield enough protein on a single run; to make enough E1/E3 for HPLC using this method, 6 FPLC purifications would be performed for each time point on the inactivation curve.

7.4 Optimising purification protocol for analysis of treated PDC

It was decided to continue adapting earlier procedures (section 7.2) for the analysis of collagenase treated PDC, and its response to ATP. Further research was designed to continue optimising yields of purified E1/E3 with the buffering system required for the collagenase treatment.

7.4.1 Optimising the buffering system

The procedure was initially carried out using the collagenase Tris buffer and the kinase phosphate buffer as described in methods. PDC was split into 4 parts, to be treated with either ATP, collagenase, collagenase and then ATP, and an untreated control. All PDC was first diluted with 2 vol collagenase buffer (to 8 mg/ml), and incubated in the presence or absence of collagenase.

Following this, all PDC was diluted with 2 vol kinase buffer. However, some precipitation occurred upon addition of the buffer, and addition of ATP also caused a large amount of precipitation.

The protein remaining in solution was dissociated and purified by FPLC. After dialysing the E1/E3 fractions into HPLC grade dH_2O, only 100 µg protein, on average, was obtained. This was below the recommended amount of protein required for HPLC/ESMS. The E1/E3 samples were processed by HPLC/ESMS to ensure the recommended limits were correct. No results were obtained, confirming the need for approx. 200 µg of E1/E3 fraction for this experiment.

Since the precipitation of PDC during the kinase treatment was thought to be due to Ca^{2+} in the collagenase buffer reacting with phosphate in the kinase buffer, protocols to eliminate the phosphate were devised.

The previous experiment was repeated with an imidazole buffer (20 mM imidazole, pH 7.5, 2 mM $MgCl_2$, 2 mM DTT), in place of a kinase buffer, to avoid precipitation of the protein as before.

Upon addition of the buffer there was less precipitation of PDC. However, upon addition of ATP, the protein again precipitated. It was thought that this was due to the ATP complexing with the calcium, so a protocol to eliminate the calcium was devised.

To prevent calcium complexing with ATP when performing the kinase experiment, EGTA was added after the collagenase treatment. This would afford the collagenase the required calcium, but remove it subsequently. The previous experiment, with the modified kinase buffer, was repeated using a 5 molar excess of EGTA to chelate all the calcium after collagenase treatment. This did not prevent precipitation of the protein, and the experiment was halted.

7.5 Discussion

It is already known that E1 does not respond well to concentration when separated from the rest of the complex, probably due to its high hydrophobicity. It is expected, therefore, that after FPLC dissociation, precipitation of E1 during successive concentration steps should occur to some extent. The degree of precipitation however was not expected; at least 50% of the polypeptides appear to have been lost at this stage. Slight precipitation of PDC when treated with ATP and the incompatibility of kinase treatment of PDC with the collagenase buffer system was also unexpected.

Previous experiments in this area (Anderson, 1995) had not noted any great difficulty with precipitation of ATP. This is certainly due to the differences between previous protocols with ones used here; previous

investigators have used larger volumes of buffer to dilute the PDC, and dilution of the collagenase buffer 10 fold appears to avoid such precipitation, as does dilution of PDC with kinase buffers.

To repeat these experiments successfully, the 3 stages at which material is lost must be modified to minimise losses. Protein is lost through precipitation during the collagenase/kinase treatment, through gel filtration purification of the E1/E3 fraction, and through concentration of the E1/E3 into a small volume of dH_2O ready for HPLC purification of E1α. It is obviously the third stage which causes the greatest losses.

If the precipitation during this final stage is due to the hydrophobic nature of the polypeptides, this could be avoided by trypsin digesting the polypeptides. The smaller fragments should stay in solution during concentration, which would have to be performed by dialysis instead of ultrafiltration. However, there would be a very large number of fragments generated by this method. Whilst HPLC - ESMS has been used to show phosphorylation of polypeptides from their tryptic digest fragments (Krell, 1997), the large number of fragments generated by digesting 3 different polypeptides may be too many to resolve. If this experiment was possible it would have the advantage of relaying information about which sites were phosphorylated, and the ratio of phosphorylated to unphosphorylated serine at each site.

The data found here gives accurate mass determinations for bovine E1α E1β and E3, which have not yet been cloned and sequenced. The viability of the technique for discovering the behaviour of the PDC kinase was evaluated, and improvements discussed.

Chapter 8

General Discussion

8.1 Role of extended presequences

Long presequences are common to all the E2 components of the 2-oxoacid dehydrogenase complexes, and protein X of mammalian PDC. Where standard mitochondrial matrix targeting signals are approx. 15-30 amino acids long in human, these presequences are between 61 and 73 amino acids in length. In mitochondrial polypeptides which have such presequences, additional information is usually included, such as those polypeptides targeted for the inner membrane, or the 66 amino acid presequence of the 81 amino acid ATPase subunit 6, required to hold this hydrophobic polypeptide in solution.

An additional feature of all 2-oxoacid dehydrogenase complex E2 components is the ability to self assemble into large symmetrical multimeric structures, as demonstrated recently for the PDC-E2/X subcomplex where full activity could be regained from separated denatured components (Sanderson *et al,* 1996). It is this capacity to assemble spontaneously which may be regulated outside the mitochondrion by the extended presequences. This hypothesis is supported by research demonstrating that OGDC-E2 with a presequence cannot be recognised by antibodies raised against the mature form, and further studies carried out in this laboratory with BCDC-E2, where expressed BCDC-E2 in *E. coli* could only assemble when no presequence was present. Current research into protein folding has demonstrated this theme in other systems, where parts of a polypeptide may prevent or control the folding of other parts of the peptide. Such sequences within polypeptides are described as 'intra-molecular chaperones'.

8.2 Summary of research

In order to investigate the roles of the presequence, a plasmid was designed to aid research into this sequence, and presequences in general. The majority of presequence mutagenesis studies are performed using vectors where presequence DNA has been inserted in-frame and upstream of the reporter gene. Such methods lead to the expression of the promotor region within the construct, compromising the validity of the subsequent results. A vector was designed to eliminate this problem. The gene for CAT was utilised, based on the common usage of CAT as a reporter protein and the great similiarity of CAT with the core domain of PDC-E2. To advance targeting studies from the *in-vitro* work originally employed by Hurt *et al.* (1984) and to further improve the *in-vivo* methods used by workers such as Thomas *et al.* (1994), it was decided to alter the CAT plasmid to allow insertion of DNA fragments into the CAT gene, behind the first codon of the coding sequence. However, mutagenesis of the CAT plasmid was hampered by unusual PCR conditions and loss of low molecular weight PCR products during purification. These complications were incurred due to the choice of restriction sites at which the replacement DNA was to be ligated, and by the decision to mutate all 8 bases simultaneously. This degree of mismatch lowered the annealing temperature considerably more than estimated. Despite these temporary difficulties, the plasmid was finally altered to provide the desired restriction sites, and the encoded CAT was demonstrated in chapter 5 to behave in precisely the same manner as the CAT expressed by the original plasmid.

Following creation of the plasmid, a number of attempts to amplify the DNA coding for the presequence of PDC-E2 were carried out. PCRs were performed on both a λgt10 human foetal cDNA library and human genomic DNA. Difficulties were encountered in all cases. Reasons for this were dicussed in chapter 4, including the possiblity that the genomic DNA sequence

for PDC-E2 contained an intron near the end of the presequence. Other methods could be used in future endeavours to isolate the entire presequence, including new techniques such as 5' RACE.

In place of a full clone of the entire presequence, segments of the presequence were selected for inital research within this work, which were made synthetically. The choice of segments was based on analysis of the presequence using variety of computer methods. From this close scrutiny of the amino acid sequence, the N and C terminals of the presequence were selected for further analysis by expression in pSoUP. These sections were synthesised as single stranded precursors, then made double stranded by a methodology based on a single PCR cycle. Each DNA fragment was then cloned into the new plasmid vector, ready for transfection into COS-7 cells. It was fortunate that at this point a co-worker in the lab, Dr Loic Briand, managed to produce a full length PDC-E2 presequence. This too was incorporated into the vector pSoUP.

Finally, the presequence segments integrated into pSoUP were expressed in COS-7 cells to test the ability of the presequences to halt folding and intiate targeting of CAT to the mitochondrion.

8.3 Summary of findings

Surprisingly, the N-terminal sequence could not target CAT to the mitochondrion, but allowed CAT to fold in the cytosol uninhibited; one could compare this *in vivo* experiment with a similar *in vitro* experiment where a reduced presequence was able to target a non-mitochondrial enzyme to the mitochondrion (Hurt *et al.*, 1985).

The full presequence in pSoUP, made by L. Briand, was also unable to target CAT *in vivo*, or prevent folding in the cytosol. Although CAT activity was low, it did correspond to the level of expression of CAT as seen by gel

analysis, indicating that the low activity was not due to any antifolding. It is not surprising that import of CAT was not detected since CAT had already folded in the cytosol in both cases. These results are in contrast to other experiments carried out by L. Briand, where the antifolding activity of the BCDC-E2 presequence was revealed; expression of mature BCDC in *E. coli* produces folded, active cores, but incorporation of segments of the BCDC-E2 N-terminal presequence as short as 17 amino acids were able to prevent this folding.

These findings suggest that either the presequence is incorrect, unable to prevent folding, or, more likely, that CAT, as a prokaryotic cytosolic enzyme, does not form a good model for mammalian mitochondrial PDC-E2. It perhaps also illustrates the difference between *in vitro* and *in vivo* experiments, where inside the cell, under normal circumstances, the mitochondrion appears to be more discerning. *In vitro* targeting experiments often require the organelles to operate with a high concentration of precursors to mitochondria, favouring import. *In vivo* experiments are often carried out in yeast and there is not always direct correlation between mitochondrial behaviour in yeast and mammalian cells; hence the need to perform mitochondrial targeting studies in higher eukaryotes to complete the picture.

It is worth noting that not only was CAT not translocated into the mitochondrion by the N-terminal section of the presequence, but it also remained soluble within the cytoplasm. If the N-terminal and the full length presequences were simply incapable of keeping CAT unfolded, then it could still be targeted to the mitochondrion. In this event, CAT would be trapped, held to the mitochondrion in a translocation-incompetent state by the partially translocated N-terminal sequence, and would not appear in the cytosolic fraction. Instead, CAT activity would be present in the membrane pellet in all digitonin fractionations. Of interest, however, are the findings of Leiper and Danpure (1997, also reviewed in Danpure, 1997) where a polypeptide was

targeted to the mitochondrion when monomeric, and not targeted when the polypeptide dimerized. It was noted that this polypeptide dimerized faster than it could be commited to mitochondrial import. It is possible that a similar event takes place with CAT; trimerisation (as indicated by the presence of CAT activity) could also prevent both targeting and uptake.

In summary, the experiments show that this PDC-E2 presequence is insufficient alone for targeting CAT to the mitochondrion. It is possible some property is held by the rest of PDC-E2 which also aids mitochondrial uptake.

8.4 Further work

Further work to elucidate the role of the presequence will be based around an extension of the experiments of Dr L. Briand with BCDC-E2 expression in *E. coli*, using instead PDC-E2. An alternative method to determine the behaviour of the presequence would be the use of 'epitope tagging'. If a full length cDNA for PDC-E2 were available, the mature part of E2 could be labelled with an epitope from a non-eukaryotic source. The result should be a more accurate representation of what takes place *in vivo*, yet retaining the capacity to follow the targeting of the polypeptide by detecting the epitope tag.

To further investigate the suitability of the plasmid, a number of characterised presequences should be used to target CAT *in vivo*. One of these should be the BCDC-E2 presequence, already characterised in *E. coli* as described. It may be possible that CAT requires the lipoyl didomain to accurately simulate any E2 component. Since the lipoyl domains are the next to be synthesised from the ribosome after the presequence, it would make sense if the intra-molecular chaperone activity of the presequence acted upon this section of the polypeptide. Further work on the interaction of the presequence with the lipoyl domains of E2 is currently being undertaken

within the laboratory. A more lateral approach would be to substitute the CAT gene with the DNA sequence for mature PDC-E2. The presence of PDC-E2 in the cytosol would demonstrate the inability of this presequence to target to the mitochondrion. Proof of targeting could only be demonstrated by epitope tagging, as described above.

8.5 Experiments with PDC Kinase

The ability of PDC kinase to phosphorylate three sites on E1 was also investigated. Initially, a method for isolation and analysis of E1 samples was devised and tested prior to its application to phosphorylated E1. Once this method had been tested, it was applied to PDC treated with ATP. Although other workers in this area had not noted any problems caused by precipitation of PDC during or after phosphorylation, the differences between the new method and previous ones were sufficiently large to compromise the experiments. Additional loss of material at a number of other points usually rendered mass spectroscopy of the final sample worthless. Fortunately, however, some data was acquired which gave accurate mass determinations for bovine PDC-E1α, PDC- E1β and PDC-E3, which have not yet been cloned or sequenced.

Despite the apparent unsuitability of these procedures to analyse the behaviour of the kinase, the ability of the technique to differentiate between phosphorylated and non-phosphorylated forms of E1α was demonstrated. Alternative methods for treatment and purification of E1 could be devised; some of these possiblities were discussed in chapter 7. For example, if a dilute sample of PDC could be processed on an affinity column instead of a chromatography column, it may be possible to purify and analyse E1, as in this thesis, but without repeated concentration and washing steps, or without the high starting concentration for PDC during the kinase/collagenase

treatment.

8.6 Discussion of Research

In each results chapter the research has been described and alternative approaches discussed. We can now look at the work as a whole and other ways in which the research could have been carried out, using methods that are now available.

Initially, a plasmid (pSoUP) was designed for general presequence analysis, the construction of which took more time than was suitable for this portion of the experiment. On reflection, a more general vector could have been used, such as a β-galactosidase or luciferase vector with a cloning site near, but not necessarily at, the start of the gene. This would have increased the amount of time available to carry out the other parts of the research. A new reporter gene coding for a fluorescent polypeptide called 'Green fluorescent protien' (GFP) is becoming popular, and in addition to being a highly effective reporter, the use of GFP with immunfluorescence microscopy could replace digitonin fractionation as a means of discovering the position of the targeted polypeptide. Furthermore, the folding of GFP could be monitored more accurately than CAT, by the use of laser scanning confocal microscopy and associated software to pinpoint location and quantitate the folding of the polypeptide from a non-fluorescent species to a folded one.

Isolation of the PDC-E2 presequence proved difficult; other presequences could have been tested, such as the E2 presequences of BCDC and OGDC. PCR of the OGDC-E2 presequence based on published data was also attempted, but failed. Only recently has a previously published full length clone of the BCDC-E2 gene been made available to this laboratory, and presequence experiments upon it are currently being undertaken. A positive control presequence could also have been utilised. A plasmid containing one

such sequence was expected, but like the BCDC-E2 gene, this was available only in the later stages of this research.

As mentioned previously, it could be that regions of the polypeptide after the presequence are needed for targeting, such as the lipoyl didomain. It is known that the presequence interacts with the mature part of the polypeptide to prevent folding. It is possible therefore that without this interaction the polypeptide cannot be transported, if, like BSC1, the presequence must fold into a hairpin loop before translocation. Experiments to show this could use either epitope tagged full length PDC-E2, or pSoUP modified with the DNA coding for the presequence-lipoyl didomain in the cloning site.

It may be more interesting now to investigate the presequence of Protein X, which has been postulated to be the dominant autoantigen in PBC. The obvious experiment to conclude the various studies discussed here would be the transfection of PBC epithelial cells with GFP constructs containing the various presequences discussed. This would demonstrate which presequence, if any, mistargets polypeptides to the plasma membrane during PBC.

References

Anderson, L. (1995) Honours Project; The Interaction and Regulatory Properties of Mammalian Pyruvate Dehydrogenase Complex Kinase.

Ali, S. T. and Guest, J. R. (1990) Biochem. J. **271**, 139-145.

Arnheim, N. and Erlich, H. (1992) Annu. Rev. Biochem. **61**, 131-156.

Behal, R. H., DeBuysere, M. S., Demeler, B., Hansen, J. C. and Olson, M. S. (1994) J. Bio. Chem. **269**, (50), 31372-31377.

Bullock, W. O., Fernandez, J. M. and Short, J. M. (1987) Biotechniques **5**, 376-379.

Bolliger, L., Junne, T., Schatz, G. and Lithgow, T. (1995) EMBO J. **14**, (24), 6318-6326.

Boutry, M., Nagy, F., Poulsen, C., Aoyagi, K. And Chua, N. H. (1987) Nature **328**, (6128), 340-342.

Brookfeild, D. E., Green, J., Ali, S. T. Machado, R. S. and Guest, J. R. (1991), FEBS Lett. **295**, 13-16.

Bryans, M., Kass, S., Scivwright, C. and Adams, R. L. P (1992) FEBS Lett. **309**, (1), 97-102.

Carothers, D. J., Pons, G. and Patel, M. S. (1989) Arch. Biochem. Biophys **268**, 409-425.

Cha, S., Leung, P. S. C., Coppel, R. L., Van de Water, J., Ansari, A. A. and Gershwin, M. E. (1994) Hepatology **20**, (3), 574-582.

Chen, Q., Rowley, M. J. and MacKay, I. R. (1993) Liver **13**, 130-135.

Clarkson, G. H. D., King, T. E. and Lindsay, J. G. (1987) Biochem. J. **244**, 15-20.

Clarkson, G. H. D. and Lindsay, J. G. (1991) Eur. J. Biochem. **196**, 95-100.

Conner, M., Krell, T. and Lindsay, J. G. (1996) Planta **200**, 195-202.

Coppel, R. L. and Gershwin, M. E. (1995) Immunological Reviews **144**, 17-49.

Coppel, R. L., McNeilage, L. J., Surh, C. D., Van de Water, J., Spithill, T. W., Whittingham, S., Gershwin, M. E. (1988) Proc. Natl. Acad. Sci. **85**, 1-5.

Damuni, Z., Humphries, J. S. and Reed, L. J. (1984) Biochem. Biophys. Res. Commun. **124**, 95-99.

Danpure, C. J. (1997) Bioessays **19**, (4), 317-326.

Dardel, F., Laue, E. D. and Perham, R. N. (1991) Eur. J. Biochem. **201**, (1), 203-209.

De Marcucci, O. G. L. (1985) Ph.D. Thesis, Immunology and Biosynthesis of the Mammalian Pyruvate Dehydrogenase Complex.

De Marcucci, O. G. L., DeBuysere, M. S. and Olson, M. S. (1995) Archives of Biochemisrty and Biophysics **323**, (1), 169-176.

DeRosier, D. J. (1992) Nature **357**, 196-197.

Dey, R., Naik, S. and Patel, M. (1996) Biochim. Biophys. Acta **1305**, 189-195.

Endo, T., Mitsui, S., Nakai, M., and Roise, D. (1996) J. Bio. Chem. **271**, (8), 4161-4167.

Felgner, P. L., Gadek, T. R., Holm, M., Roman, R., Cha, H. W., Wenz, M., Northrop, J. P., Ringold, G. M. and Danielsen, T. M. (1987) Proc. Natl. Acad. Sci. USA **84**, 7413-7417.

Fölsch, H., Guiard, B., Neupert, W. and Stuart, R. A. (1996) EMBO. J., **15**, (13), 479-487.

Fuller, C. C., Reed, L. J., Oliver, R. M. and Hackert, M. L. (1979) Biochem. Biophys. Res. Commun., **90**, 431-437.

Fullerton, R. E. (1995) Ph.D. Thesis, Biochemical Characterisation of Dihydrolipoamide Dehydrogenase from Potato.

Fussey, S. P. M., Ali, S. T., Guest, J. R., James, O. F. W., Bassendine, M. F. and Yeaman, S. J. (1990) Proc. Natl. Acad. Sci., USA **87**, 3987-3991.

Gershwin, M. E. (1994) Private communication.

Gershwin, M. E., MacKay, I. R., Sturgess, A. and Coppel, R. L. (1987) J. Immunology **138**, (10), 3525-3531.

Glick, B. and Schatz, G. (1991) Annu. Rev. Genet. **25**, 21-44.

Glover, L. A. and Lindsay, J. G. (1992) Biochem. J. **284**, 609-620.

Green, D. E. (1994) Ph.D. Thesis; The Lipoate Protein Ligase of *Escherichia coli*..

Gopalkrishnan, S., Rahmatulla, M., Radke, G. A., Powers-Greenwood, S. and Roche, T. E. (1989) Biochem. Biophys. Res. Commun. **160**, 715-721.

Griffin, T. A., Lau, K. S. and Chuang, D. T. (1988) J. Bio. Chem. **263**, (28), 14008-14014.

Gruhler, A., Ono, H., Guiard, B., Neupert, W. and Stuart, R. A. (1995) EMBO. J. **14**, (7), 1349-1359.

Hajek, P., Koh, J. Y., Jones, L. and Bedwell, B. M. (1997) Molecular and Cellular Biology **17**, (12), 7169-7177.

Hammen, P. K., Gorenstein, D. G. and Weiner, H. (1994) Biochemistry **33**, 8610-8617.

Hanahan, D., (1983) J. Mol. Biol. **166**, 557-580.

Harris, R. A., Bowker-Kinley, M. M., Wu, P. F., Jeng, J. J. and Popov, K. M. (1997) J. Biol. Chem. **272**, (32), 19746-19751.

Hartl, F. U., (1996) Nature **381**, 571-580.

Hartmann, E., Sommer, T., Prehn, S., Gorlich, P., Jentisch, S., Rapoport, T. A. (1994) Nature **367**, 654-657.

Haucke, V., Lithgow, T., Rospert, S., Hahne, K. and Schatz, G. (1995) J. Bio. Chem. **270**, (10), 5565-5570.

Heijne, G. V., Steppuhn, J. and Herrmann, R.G. (1989) Eur. J. Biochem. **180**, 535-545.

Hendrick, J. P. and Hartl, F. U. (1995) FASEB. J. **9**, 1559-1556

Higuchi, R., Krummel, B. and Saiki, R. (1988) Nucleic Acids Research **16**, (15), 7351-7368.

Hirano, M. and Paklavis, S. G. (1994) J. of Child Neurol. **19**, (1), 4-13.

Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K. and Pease, L. R. (1989) Gene **77**, (1), 51-59.

Horton, R. M., Hunt, H. D., Ho, S. N., Pullen, J. K. and Pease, L. R. (1989) Gene **77**, (1), 61-68.

Hucho, F., Randall, D. D., Roche, T. E., Burgett, M. W., Pelly, J. W. and Reed, L. J. (1972) Arch. Biochem. Biophys. **151**, 328-340.

Hummel, K. B., Litwer, S., Bradford, A. D., Aitkin, A., Danner, D. J. and Yeaman, S. J. (1988) J. Biol. Chem. **263**, (13) 6165-6168.

Hunter, A. and Lindsay, J. G. (1986) Eu. J. Biochem. **155** (1), 103-109.

Hurt, E. C., Pesold-Hurt, B., and Schatz, G. (1984) FEBS Lett. **178**, (2), 306-310.

Hurt, E.C., Pesold-Hurt, B., Suda, K., Oppliger, W. and Schatz, G. (1985) EMBO J. 4, (8), 2061-2068.

Jacnicke, R. (1995) Phil. Trans. R. Soc. Lond. B. **348**, 97-105.

Jacnicke, R. (1996) Current Topics in Cellular Regulation **348**, 97-105.

Jaussi, R. (1995) Eur. J. Biochem. **228**, 551-561.

Jones, D. E. J., Palmer, J. M., James, O. F. W., Yeaman, S. J., Bassendine, M. F. and Diamond, A. G. (1995) Hepatology **121**, (4), 995-1002.

Joplin, R. E., Johnson, G. D., Matthews, J. B., Hamburger, J., Lindsay, J. G., Hubscher, S. G., Strain, A. J. and Neuberger, J. M. (1994) Hepatology **19**, (6), 1375-1380.

Joplin, R. E., Lindsay, J. G., Johnson, G. D., Strain, A. J. and Neuberger, J. M. (1992) The Lancet **339**, 93-94.

Kamrud, K. I., Powers, A. M., Higgs, S., Olson, K. E., Blair, C. D., Carlson, J. O. and Beatty, B. J. (1995) Exptl. Parasitol. **81**, 394-403.

Karplus, P. A. and Schulz, G. E. (1987) J. Mol. Biol. **195**, 701-729.

Kiebler, M., Becker, K., Pfanner, N. and Neupert, W. (1993) J. Mol. Biol. **135**, 191-207.

Khailova, L. S., Korochkina, L. G. and Severin, S. E. (1989) Annal. N.Y. Acad. Sci. **573**, 36-54.

Koike, K., Urata, Y. and Goto, S. (1992) Proc.Natl.Acad.Sci., USA **89**, 1963-1967.

Korotchkina, L. G., Khailova, L. S and Severin, S. E. (1995) FEBS Lett. **364**, 185-188.

Kubota, H., Hynes, G. and Willison, K. (1995) Eur. J. Biochem. **230**, 3-16.

Kubrich, M., Dietmeier, K. and Pfanner, H. (1995) Curr.Genet. **27**, (5), 393-403.

Kun, E., Eanes, R. Z. and Volfin, P. (1967) Nature **214**, 1328-1329.

Kuriyan, J., Krishna, T. S. R., Wong, L., Guenther, B., Pahler, A., Williams, C. H. and Model, P. (1991a) Nature **352**, 172-174.

Kuriyan, J., Kong, X.P., Krishna, T. S. R., Sweet, R. M., Murgolo, N. H., Field, H., Cerami, A. and Henderson, G. B. (1991b) Proc. Natl. Acad. Sci., USA **88**, 8764-8768.

Lan, K. S., Herring, W. J., Chuang, J. L., McKean, M., Danner, D. J., Cox, R. P. and Chuang, D. T. (1992) J. Biol. Chem. **267**, (33), 24090-24092.

Leiper, J. M. and Danpure, C. J. (1997) Clinica Chimica Acta **266**, (1), 39-50.

Leung, P. S. C., Iwayama, T., Coppel, R. L. and Gershwin, M. E. (1990) Hepatology, **12**, (6), 1321-1328.

Lessard, I. A. D. and Perham, R. N. (1995) Biochem. J. **306**, 727-733.

Lessard, I. A. D. Fuller, C. and Perham, R. N. (1996) Biochemistry, **35**, 16863-16870.

Lindsay, J. G. (1989) Annal. N.Y. Acad. Sci. **573**, 254-266.

Lindsay, J. G., Lever, T. M., Sanderson, S. J. and Taylor, A. E. (1993) Mitochondria. In: Methods of Immunological Analysis (eds. Masseyeff, R.F., Albert, W.H. and Staines, N.A.) Chapter 6.3.4, VCH, Weinheim.

Lithgow, T., Cuezva, J. M. and Silver, P. A. (1997) Trends Biochem. Sci. **22**, (4), 110-113.

Liu, S., Baker, J. C., Andrews, P. C. and Roche, T. E. (1995a) Arch. Biochem. Biophys. **316**, (2), 926-940.

Liu, S., Baker, J. C. and Roche, T. E. (1995b) J. Biol. Chem., **270**, (2), 793-800.

Luciano, P. and Geli, V. (1996) Experimentia **52**, (12), 1077-1082.

Luckow, B. and Shütz, G. (1987) Nucleic Acids Res. **15**, 5490-5493.

Ma, Y., Ooi, T. C., Liu, M-S., Zhang, H., McPherson, R., Edwards, A. L, Forsythe, I. J., Frohlich, J., Brunzell, J. D. and Hayden, M. R. (1994) J. Lipid Res. **35**, 1066-1075.

Machado, R. S., Guest, J. R. and Williamson, M. P. (1993) FEBS Lett. **323**, (3), 243-246.

Maeda, T., Onishi, S., Tanaka, H., Saibara, T., Iwamura, S., Miyazaki, M. and Yamamoto, Y. (1993) Intl. Hepatol. Comm. **1**, 104-108.

Maeng, C. Y., Yazdi, M. A., Niu, X-D., Lee, H. Y. and Reed, L. J. (1994) Biochemistry **33**, 13801-13807.

Mann, M. and Wilm, M. (1995) TIBS **20**, 219-224.

Marsac, C., Stansbie, D., Bonne, G., Cousin, J., Jehenson, P., Benelli, C., Leroux, J-P. and Lindsay, J. G. (1993) J. Pediat. **123**, (6), 916-920.

Martinus, R. D., Ryan, M. T., Naylor, D. J., Herd, S. M., Hoogenraad, N. J. and Hoj, P. B. (1995) FASEB. J. **9**, 371-378.

Mattevi, A., de Kok, A. and Perham, R. N (1992) Curr. Opin. Str. Biol. **2**, 877-887.

Mattevi, A., Obmolova, G., Schulze, E., Kalk, K. H., Westphal, A. H., de Kok, A. and Hol, W. G. J. (1992) Science **255**, 1544-1550.

M^{c}Cartney, R. G., Sanderson, S. J. and Lindsay, J. G. (1997) Biochemistry, **36**, (22), 6819-6826.

M^{c}Connell, B. B., Burkholder, B. and Danner, D. J. (1997) Biochim. Biophys. Acta. **1361**, (3), 263-271.

M^{c}Cully, V., Burns, G. and Sokatch, J. R. (1986) Biochem. J. **223**, 737-742.

M^{c}New, J. A. and Goodman, J. M. (1994) J. Cell Biol. **127**, (5), 1245-1257.

Mesnel, M., DeKroon, A. I. P. M. and Marsh, D. (1995) Biochemistry **34**, 3605-3613.

Milton, H. and Saier, J. R. (1994) Microbiol. Rev. **58**, (6), 71-91.

Murakami, K. and Mori, M. (1990) EMBO J. **9**, (10), 3201-3208.

Nairn, J., Krell,T., Coggins, J. R., Pitt, A. R., Fothergill-Gillmore, L. A., Walter, R. and Price, N. C. (1995) FEBS Lett. **359**, 192-194.

Nakano, K., Matuda, S., Sakamoto, T., Takase, C., Nakagawa, S., Ohta, S., Ariyama, T., Inazawa, J., Abe, T. and Miyata, T. (1993) Biochim. Biophys. Acta. **1216**, 360-368.

Nakano, K., Takase, C., Sakamoto, T., Nakagawa, S., Inazawa, J., Ohta, S. and Matuda, S. (1994) Eur. J. Biochem. **224**, 179-189.

Nakanuma, Y., Tsuneyama, K., Kono, N., Hoso, M., Van de Water, J., and Gershwin, M. E. (1995) Human Pathology. **26**, (1), 92-98.

Neagle, J. C. and Lindsay, J. G. (1991) Biochem. J. **278**, 423-427.

Nemerya, N. S., Zemskova, M. A., Nyukhalkina, I. A. and Khailova, L. S. (1996) FEBS Lett. **394**, 96-98.

Niu, X. D., Browning, K. ., Behal, R. H. and Reed, L. J. (1988) Proc. Natl. Acad. Sci., USA **85**, 7546-7550.

O'Donohue, J. and Williams, R. (1996) Qjm-Monthly Journal Of The Association Of Physicians **89**, (1), 5-13.

Oliver, R. M and Reed, L. J. (1982) in Electron Microscopy of Proteins (ed. R. Harris), Academic Press, London, **2**, 1-48.

Pares, S., Cohen-Addad, C., Sieker, L., Neuburger, M. and Douce, R.

(1994) Proc. Natl. Acad. Sci., USA **91**, 4850-4853.

Patel, M. S. and Harris, R. A. (1995) FASEB J. **9**, 1164-1172.

Patel, M. S., Leung, P. and Gershwin, M. E. (1994) Personal communication.

Patel, M.S. and Roche, T. E. (1990) FASEB J. **4**, 3224-3231.

Pesce, A., McKay, R. H., Stolzenbach, F., Cahn, R. D. and Kaplan, N. O. (1964) J. Biol. Chem. **239**, 1753-1761.

Perham, R. N., (1991) Biochemistry **30**, (35), 8501-8512.

Povov, K. M., Kedishvii, N. Y., Zhao, Y. U., Shimomura, Y., Crabb, D. W. and Harris, R. A. (1993) J. Biol. Chem. **268**, (35), 26602-26606.

Popov, K. M., Zhao, Y. U., Shimomura, Y., Kuntz, M. J. and Harris, R. A. (1992) J. Biol. Chem. **267**, (19), 13127-13130.

Pratt, M. L., Mahler, J. F. and Roche, T. E. (1982) Eur. J. Biochem. **125**, 349-355.

Quinn, J., Diamond, A. G., Palmer, J. M., Bassendine, M. F., James, O. F. W. and Yeaman, S. J. (1993) Hepatology **18**, (6), 1384-1391.

Radke, G. A., Ono, K., Ravindran, S. and Roche, T. E. (1993) Biochem. Biophys. Res. Commun. **190**, (3), 982-991.

Rahmatullah, M., Radke, G. A., Andrews, P. C. and Roche, T. E. (1990), J. Biol. Chem. **265**, (24), 14512-14517.

Rahmatullah, M. and Roche, T. E. (1988), J. Biol. Chem. **263**, (17), 8106-8110.

Rapaport, D., Neupert, W. and Lill, R. (1997) J. Bio. Chem. **272**, (30), 18725-18731.

Ravindran, S., Radke, G. A., Guest, J. R. and Roche, T. E. (1996) J. Biol. Chem., **217**, (2), 653-662.

Rice, J. E., Dunbar, B. and Lindsay, J. G. (1992) EMBO J., **11**, (9), 3229-3235.

Richarme, G. and Heine, H-G. (1986) Eur. J. Biochem., **156**, 399-405.

Roche, T. E., Powers-Greenwood, S. L., Shi, W. F., Zhang, W. S., Ren, S. Z., Roche, E. D., Cox, D. J. and Sorensen, C. M. (1993) Biochemistry, **32**, 5629-5637.

Reed, L. J. and Yeaman, S. J. (1987) in "The Enzymes" (eds. P. D. Boyer and E. G. Krebs) Academic Press, Orlando, F.L **18**, 77-95.

Rosch, S. L. and Kendall, D. A. (1995) Molec. Membr. Biol. **12**, 295-307.

Runswick, M. J., Powell, S. J., Nyren, P. and Walker, J. E. (1987) EMBO J. **6**, (5), 1367-1373.

Rowles, J., Xi, T., Scherer, S. and Prochazka, M. (1996) Diabetes 45, (S2), 343.

Rusch, S. L. and Kendall, D. A. (1995) Molec. Membr. Biol. **12**, 295-307.

Russell, G. C. and Guest, J. R. (1991) Biochim. Biophys. Acta **1076**, (2), 225-232.

Russell, G. C., Machado, R. S. and Guest, J. R. (1992) Biochem. J. **287**, 611-619.

Ryan, K. R. and Jensen, R. E. (1995) Cell **83**, 517-519.

Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) "Molecular Cloning- a Laboratory Manual" 2nd Edition (eds. N. Ford, C. Nolan and M. Ferguson) Cold Spring Harbour Laboratory Press.

Sanderson, S. J., Khan, S. S., McCartney, R. G., Miller, C. and Lindsay, J. G. (1996) Biochem., J. **319**, 109-116.

Schiering, N., Kabsch, W., Moore, M. J., Distefano, M. D., Walsh, C. T. and Paiu, E. F. (1991) Nature **352**, 168-172.

Schlesinger, M. J. (1994) Pediatric Research **36**, (1), 1-6.

Scotland, G. and Houslay, M. D. (1995) Biochem. J. **308**, 673-681.

Schatz, G. and Dobberstein, B. (1996) Science **271**, 1519-1526.

Seed, B. Sheen, J. Y. (1988) Gene **67**, 271-277.

Seki, N., Moczko, M., Nagase, T., Zufall, N., Ehmann, B., Dietmeier, K., Schafer, E., Nomura, N and Pfanner, N. (1995) FEBS Lett. **375**, (3),

307-310.

Shakur, Y., Pryde, J. G. and Houslay, M. D. (1993) Biochem. J. **292**, 677-686.

Sherlock, S. (1994) Amer. J. Med. **96**, 1A/27S-1A/33S.

Shore, G.C., McBride, H. M., Millar, D. G., Steenart, N. A. E. and Nguyen, M. (1995) Eur. J. Biochem. **227**, 9-18.

Sirrenberg, C., Endres, M., Folsch. H, Stuart, R.A., Neupert, W. and Brunner, M. (1998) Nature **391**, (6670), 912-915.

Smith, D. B. and Johnson, K. (1988) Gene **67**, 31-60.

Snel, M. M. E., DeKroon, A. I. P. M. and Marsh, P. (1995) Biochemistry **34**, 3605-3613.

Song, H. and Komuniecki, R. (1994) J. Biol. Chem. **269**, (50), 31573-31578.

Stehle, T., Ahmed, S. A., Claiborne, A. and Schulz, G. E. (1991) J. Mol. Biol. **221**, 1325-1344.

Stepp, L. R. and Reed, L. J. (1985) Biochemistry **24**, (25), 7187-7191.

Stuart, R. A. Cyr, D. M., Craig, E. A. and Neupert, W. (1994) TIBS **19**, (2), 87-92.

Stuart, R. A., Gruhler, A., Vanderklei, I., Guiard, B., Koll, H. and Neupert, W. (1994) Eur. J. Biochem. **20**, (1), 9-18.

Teague, W. M., Pettit, F. H., Wu, T-L, Silberman, S. R. and Reed, L. J. (1982) Biochemistry **21**, 5585-5592.

Thekkumkara, T. J., Ho, L., Wexler, I. D., Pons, G., Liu, T-C., Patel, M. S. (1988) FEBS Lett. **240**, (12), 45-48.

Thekkumkara, T. J., Pons, G., Mitroo, S., Jentoft, J. E., Patel, M. S. (1989) Annal. N. Y. Acad. Sci. **573**, 113-129.

Thomas, D., Leterme, S. and Boutry, M. (1994) Plant Mol. Biol. **24**, 631-641.

Tian, G., Vainberg, I. E., Tap, W. D., Lewis, S. A. and Cowan, N. J. (1995) Nature **375**, 250-253.

Tsuneyama, K., Van de Water, J., Leung, P. S. C., Cha, A., Nakanuma, Y., Kaplan, M., Delellis, R., Coppel, R., Ansari, A. and Gershwin, M. E. (1995) Hepatology **21**, (4), 1031-1037.

Von Heijne, G. (1983) Eur. J. Biochem. **133**, 17-21.

Von Heijne, G., Steppuhn, J. and Herrmann, R. G. (1989) Eur. J. Biochem. **180**, 535-545.

Wagenknecht, T., Grassucci, R., Radke, G. A. and Roche, T. E. (1991) J. Biol. Chem. **266**, (36), 24650-24656.

Weiland, O. (1983) Rev. Physiol. Biochem. Pharmacol. **96**, 123-170.

West, S. M., Rice, J. E., Beaumont, E. S., Kelly, S. M., Price, N. C. and Lindsay, J. G. (1995) Biochem. J. **308**, 1025-1029.

Wise, A. (1994) Private communication.

Wu, T. L. and Reed, L. J. (1984) Biochemistry **23**, (2), 221-226.

Wynn R. M., Davie, J. R., Zhi, W., Cox, R. P. and Chung, D. T. (1994) Biochemistry **33**, 8962-8968.

Yang, Y-S., Datta, A., Hainfeld, J. F., Furuya, F. R., Wall, J. S., Frey, P. A. (1994) Biochemistry **33**, 9428-9437.

Yanisch-Perron, C., Vieira, J., Messing, J. (1985) Gene **33**, 103-119.

Yeaman, S. J., Hutcheson, E. T., Roche, T. E., Pettit, F. H., Brown, J. R., Reed, L. J., Watson, D. C. and Dixon, G. H. (1978) Biochemistry **17**, 2364-2370.

Yeaman, S. J. (1989) Biochem. J. **257**, 625-632.

Zwizinski, C., Schleyer, M. and Neupert, W. (1983) J. Biol. Chem., **258**, (7), 4071-4074.