Molecular and structural studies on proteins involved in lipid biosynthesis in plants and bacteria

Simon, Josiah William

Molecular and Structural Studies on Proteins Involved in Lipid Biosynthesis in Plants and Bacteria.

A thesis submitted to the University of Durham in accordance with the regulations for admittance to the Degree of Doctor of Philosophy

by Josiah William Simon.

School of Biological and Biomedical Sciences

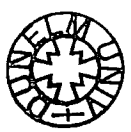
University of Durham

Science Laboratories

South Road

Durham DH1 3LE

March 2002



Molecular and Structural Studies on Proteins Involved in Lipid Biosynthesis in Plants and Bacteria.

J. W. Simon (2002)

Abstract

Fatty acid synthesis in plants and most bacteria is catalysed by a type II, dissociable fatty acid synthetase, critically involving the participation of acyl-carrier protein (ACP) during synthesis and elongation reactions. This provides the substrate for glycerolipid biosynthesis, the first reaction of which in plants, is catalysed by a soluble glycerol-3-phosphate acyltransferase (G3PAT).

A candidate gene for malonyl CoA:ACP transacylase (MCAT) from plants has been identified using homology to known MCAT sequences. A full-length cDNA was isolated from a *Brassica napus* embryo library and used to complement an *E.coli* mutant defective in MCAT activity, so providing proof of function.

E.coli ACP has been investigated in attempts to obtain crystals which can be used in structural investigations to solve its structure via X-ray crystallography. Wild type butryl-ACP can be successfully crystallised and defracts to 2.0 Å resolution, and is thus a candidate for solving the structure of the protein. The complete structure of butryl-ACP was elucidated following the introduction of new methionine residues into the protein using site directed mutagenesis, and the production of recombinant proteins containing selenomethionine as heavy metal derivatives.

Squash recombinant G3PAT was over-expressed and used to obtain crystals which defract to 1.8 Å. In order to solve the phasing problem selected cysteine residues were removed from the protein, resulting in the identification of histidine 279 as an additional heavy metal binding site. The structure of the enzyme was solved to 1.8 Å resolution and potential substrate binding sites have been modelled into it on the basis of the conserved $H(X)_4D$ acyltransferase domain. The site for G3P binding has been confirmed following site directed mutagenesis and novel substrate selectivity has been introduced into the protein by PCR mediated mutagenesis.

These studies have resulted in the first X-ray structure of two components of lipid biosynthesis, increased our understanding of the reactions they catalyse and successfully identified an authenticated cDNA for MCAT from plants.

Acknowledgements

My first acknowledgement must go to Professor Toni Slabas for allowing me the opportunity to carry out this research on a "part-time" basis whilst working in his laboratory. He has shown a great interest in the research, has enthusiastically encouraged the furthering of my academic achievements, and has given a great deal of his time and effort throughout the research period.

Next I would like to thank the School of Biological and Biomedical Sciences for allowing me the use of its facilities and the time to carry out this research whilst employed in the School.

To all the members of the Plant Lipid Molecular Biology Group I would like to say thank you for your help particularly Tony, Adrian, Johan, and Paul, thanks for the many useful conversations. I would also like especially to acknowledge John Gilroy who as a friend has helped on a daily basis with my research and has always been willing to chat and offer useful suggestions. To Paul Sidney for the photography and to other friends and colleagues in the department I would also say many thanks for the help.

Finally to my wife, Sarah who has been a constant inspiration, always-offering the love, encouragement and support necessary to keep me going. Thank you for your help towards the completion of this thesis.

Declaration

The work reported in this thesis has not been submitted previously for any degree in this or any other University. It is entirely the result of my own research except where otherwise acknowledged.

J.W. Simon (2002)

Abbreviations

ACP	Acyl-Carrier Protein
Amp	Ampicillin
BSA	Bovine serum albumin
cDNA	Complimentary DNA
CoA	Coenzyme A
DNA	Deoxyribose nucleic acid
DAG	Diacylglycerol
dNTPs	Dioxynucleotide triphosphates
DTT	Dithiothreitol
EDTA	Ethylenediaminetetracetic acid
ESIMS	Electrospray ionisation mass spectrometry
FAS	Fatty acid synthase
G3PAT	Glycerol-3-phosphate acyltransferase
HPLC	High performance liquid chromatography
IPTG	Isopropyl -β-D- thiogalactopyranoside
Kan	Kanomycin
kb	Kilobases
kDa	Kilodaltons
LB	Luria-Bertani medium
MALDItof-MS	Matrix assisted laser desorption time of flight mass spectrometry
MCAT	Malonyl-CoA:ACP transacylase
MWCO	Molecular weight cut off
MQ	Milli-Q water
mRNA	Messenger RNA
OD	Optical density

PA	Phosphatidic acid
PC	Phosphatidylcholine
PCR	Polymerase chain reaction
PE	Phosphatidylethanolamine
pfu	Plaque forming units
PG	Phosphatidylglycerol
PI	Phosphatidylinositol
RNA	Ribose nucleic acid
rpm	Revolutions per minute
SDM	Site directed mutagenesis
SDS	Sodium dodecoyl sulphate
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
SSC	Salt sodium citrate buffer
TAE	Tris-acetate buffer
TAG	Triacyglycerol
TE	Tris-EDTA buffer
TEMED	N'N'N'N'-tetramethylethylenediamine
TCA	Trichloroacetic acid
Tm	Annealing (melting) temperature
Tris	tris-(hydroxymethyl)-aminomethane
U	Units of enzyme activity
UV	Ultraviolet
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

Table of Contents

List of Figures

List of Tables

Page Number

Chapter 1

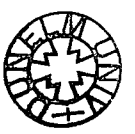
General Introduction

1.1 The Importance of Plant Fatty Acids and Lipids

Fatty acids and lipids are important components of all living organisms. They serve an essential role in all cells as integral component of the bilayer of biological membranes (phospholipids, galactolipids and sterol esters), as signalling molecules and as triacylglycerols (TAGS) which are one of the major storage reserves, deposited as oil bodies in seeds and other lipid rich tissues.

Fatty acids are organic molecules consisting of a hydrocarbon chain, ranging from one to thirty carbon atoms in length attached to a terminal carboxyl group (**Figure 1.1**) (Gunstone *et al.*, 1994).

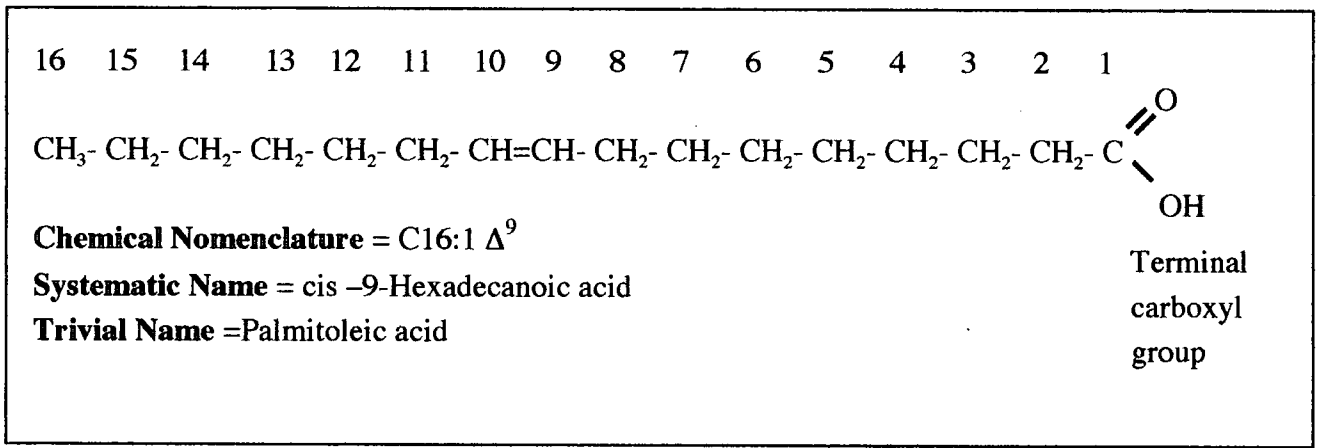
There are a variety of "common names" for fatty acids, the derivation of which is largely historic. For example hexadecanoic acid is commonly known as palmitic acid due to the fact that its origin was from oil palm. A short hand nomenclature has been developed for fatty acids. The methyl carbon is termed ω and the carboxyl carbon Δ and they are known by the chemical abbreviation C followed by the number of carbon atoms in the chain beginning with the C of the carboxyl group. Fatty acids can contain a variable number of double bonds which can considerably influence their physical properties, such double



bonds occur in specific positions and are predominantly *cis* in conformation. When written the double bond is indicated by a colon and represented by a $\Delta^{(n)}$ showing the position of the double bond when counted from the carboxyl end of the fatty acid **(Figure 1.1)**.

In plants the most common fatty acids found are between 12 and 22 carbon atoms in length, mainly C16 and C18 containing up to three double bonds.

Figure 1.1 Structure and nomenclature of fatty acids.



Most fatty acids are found in living cells as constituents of lipid molecules, glycerolipids, sterols and waxes.

Lipids are made up of a diverse array of different chemical classes, some of which serve household functions, and some, which serve luxury functions within an organism.

There is little variation in the fatty acid composition of the structural lipids within plant membrane systems, (mainly C16, C18 acyl chain length, with 1,2, or 3 degrees of unsaturation) (for review see Millar *et al.*, 2000). However there is wide variation in the

composition of the fatty acids found in the triacylglycerols of seed oils. It is this wide variety, with more than 300 different fatty acids occurring in seed oils (Harwood 1980,van de Loo 1993), many including unusual fatty acids, such as hydroxy, epoxy, and cyclic (Broun *et al.*, 1998), with chain lengths varying from C8 to C24, that make these seed oils of commercial value. The typical fatty acid composition of some of the common seed oils is shown in **Table 1.1**. Many important commercial crops such as rape, flax, sunflower etc. can contain as much as 45% of the seed weight as oil.

Estimates suggest that over 9000 kilo-tonnes of oils and fats a year are utilised in the worlds chemical industries (Somerville, 1993). Plant fatty acids and lipids may provide a renewable alternative to the use of some fossil fuels in the future. Vegetable oils are the major source of all edible lipids, they account for more than 75% of all the fats consumed in the world (Salas *et al.*, 2000) with the total expected global production in the year 2000 reported to be around 82 million metric tons (Kaufman and Ruebusch 1990).

Seed oils which predominantly contain medium chain (C10 to C14) fatty acids such as palm (mesocarp) oil and coconut are used in the manufacture of soaps and detergents. Those containing longer chain (C16, C18) saturated and unsaturated fatty acids such as soybean, sunflower, olive and rapeseed are generally edible and are used in food production. The longer chain fatty acids i.e. (C22:1) from high erucic acid rapeseed are used as lubricants and plasticisers.

Table 1.1 Typical fatty acid composition of common plant seed oils (% by weight).

Fatty Acid	Castor	Coconut	Linseed	Palm		Rape seed
				Oil	kernel	
Caproic (C6)		0.5			0.3	
Caprylic (C8)		7.5			3.9	
Capric (C10)		7.0			4.0	
Lauric (C12)		48.0			49.6	
Myristic (C14)		16.5		1.0	16.0	
Palmitic (C16)	1.0	8.0	5.5	47.0	8.0	4.9
Palmitoleic (C16:1)	1.0					0.5
Stearic (C18)	1.0	4.0	3.5	4.0	2.4	2.0
Dihydroxystearic (C18)	1.5					
Oleic (C18:1)	3.0	5.0	19.0	37.5	13.7	47.9
Ricinoleic (C18:1) (hydroxy oleic acid)	89.0					
Linoleic (C18:2)	4.0	2.5	15.5	10.0	2.0	25.2
Linolenic (C18:3)			57.0			15.2
Arachidic (C20)	0.5				9.1	1.2
Arachidonic (C20:1)				0.5		1.9

There is an increasing awareness that many of these lipid molecules have wider roles within biological systems. They are involved in signal transduction (plant – bacterial communication) between *Rhizobium meliloti* and it's legume partner during root nodule development (Lerouge *et al.*, 1990, Truchet *et al.*, 1991) and phospholipid signalling, (Lerouge *et al.*, 1990). They also play a major role in the chilling sensitivity of plants, where the degree of un-saturation of the acyl group of glycerolipids present in the lipids of the chloroplast membrane effects the fluidity of the membrane and hence a plants

ability to resist cold temperatures (Murata *et al.*, 1992 and Murata & Wada 1995). Plants such as *Arabidopsis* and spinach with a high proportion of *cis* unsaturated fatty acids in their membranes are chilling resistant, whereas those with a low proportion of unsaturated fatty acids, such as squash and bean are chilling sensitive (Murata *et al.*, 1992 and Murata & Wada 1995).

Because of these diverse roles for fatty acids and lipids there is extensive interest in the enzymology of this biosynthetic pathway. Research in this area has become much more intense since it has been recognised that several of these enzymes, (particularly the reductases) could be targets for pharmaceuticals (Banerjee *et al.*, 1994) and agrochemicals (Ashton *et al.*, 1994).

The central core of the reactions leading to both membrane and storage lipids in all organisms is *de novo* fatty acid biosynthesis (FAS). A pathway which in plants is plastid located and which results in the synthesis of C16 and C18:1 fatty acids linked to acyl carrier protein (acyl-ACPs). Once synthesised the acyl-ACPs are then either hydrolysed to free fatty acids by the action of acyl-ACP thioesterase and exported from the plastid to the cytoplasm or directly incorporated into plastid membrane lipids.

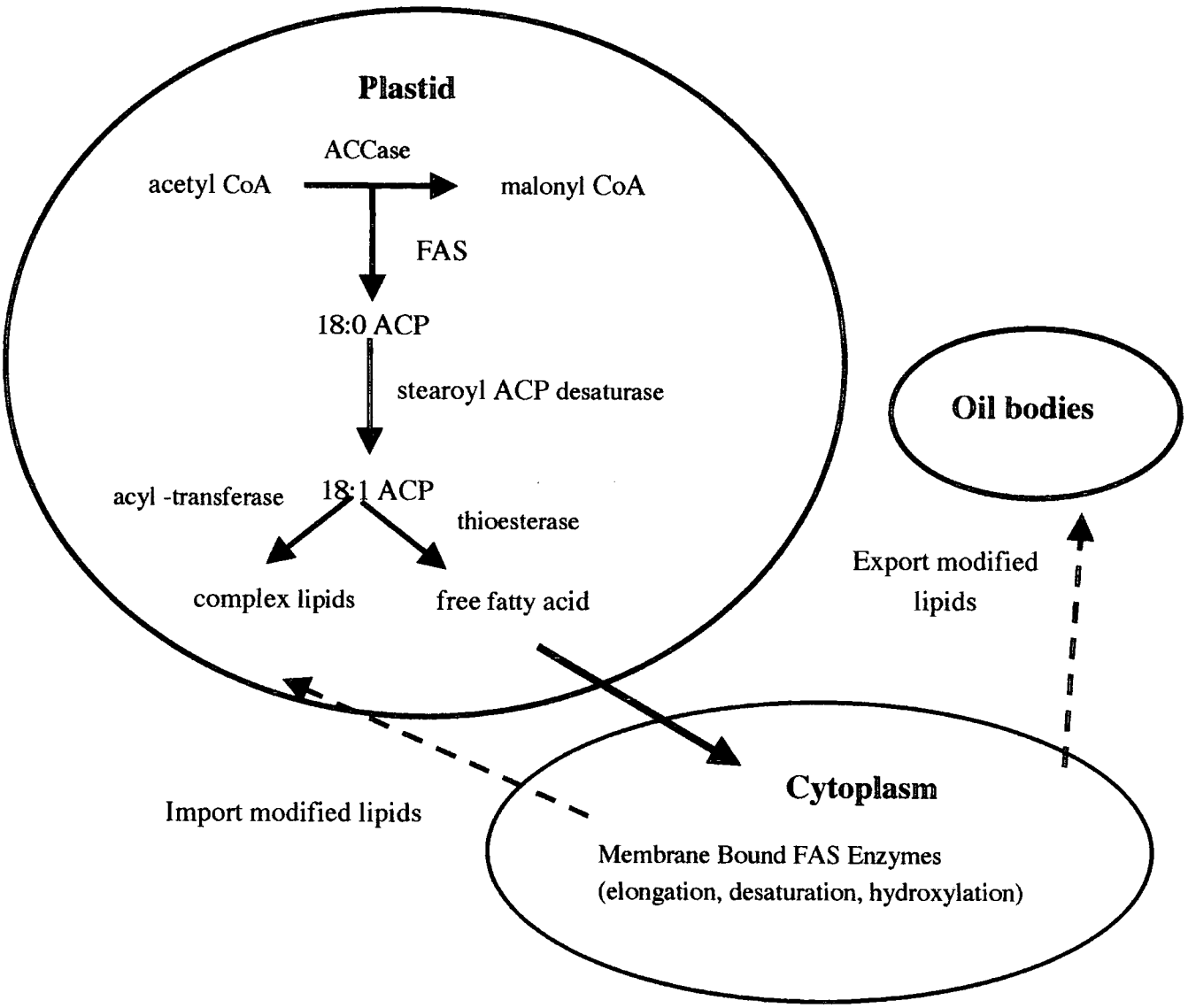
The main biosynthetic pathway for the synthesis of complex lipids in the cytoplasm involves an acyl-CoA route and it is not entirely clear how the fatty acids are exported from the plastid or where they are converted to acyl-CoAs. Both the desaturation

reactions, converting fatty acids from C18:1 to C18:3 and elongation reactions, converting C18 to C22 fatty acids occur in the cytoplasm. These modified fatty acids are incorporated into membrane lipids or TAGS by the appropriate acyltransferase. These cytoplasmic acyltransferases are all membrane bound. The initial step of the incorporation of fatty acids into plastid lipids is catalysed by the action of a soluble plastid located glycerol–3-phosphate acyltransferase (G3PAT).

This two-compartment system of plant lipid synthesis is outlined in **Figure 1.2**

The principal difference in the lipid produced by these two compartments is the nature of the fatty acid esterified to the *sn 2* position of the glycerol backbone (Roughan and Slack 1982, Browse & Somerville 1991). The plastid lipids contain 16:0 and 18:1 acyl groups at the *sn1* position and almost exclusively 16:0 acyl groups at the *sn 2* position. This is typical of the lipids found in prokaryotic photosynthetic organisms such as cyanobacteria, the presumed forerunner of the chloroplast, and hence is known as a prokaryotic lipid. The lipids synthesised at the endoplasmic reticulum however contain both 16 and 18 acyl groups at the *sn1* position and predominantly C18 acyl groups at the *sn2* position these are termed eukaryotic lipids.

Figure 1.2 The two compartment model of fatty acid synthesising and modifying enzymes in plants.



Within the plastid fatty acids up to C18 in length are synthesised as their ACP derivatives by the action of the fatty acid synthetase (FAS) enzymes. Desaturation can then occur by the action of soluble stearoyl-ACP desaturase before they are either incorporated directly into complex plastid lipids by acyltransferases or alternatively hydrolysed to free fatty acids by acyl-ACP thioesterase. These free fatty acids are then exported from the plastid where further elongation, desaturation and hydroxylation can occur (predominantly on microsomes) before the lipid is either returned to the plastid or incorporated into triacylglycerols.

1.2 Plant Fatty Acid Synthetase (FAS).

Plant FAS is made up of at least eight catalytic components and a central acyl carrier protein (ACP) (**Figure 1.3**). These components are soluble and dissociable from each other in the same way as the FAS enzymes of the bacterial systems, typified by *E.coli*. Such a dissociable system, which is stimulated by the addition of ACP, is termed type II FAS. This is in contrast to the arrangement found in the yeast and animal systems, where all of the components exist on one or two large polypeptide chains (Lynen 1969, Bloch and Vance 1977) and known as type I FAS. This organisational difference makes the separate catalytic components of the plant system much more amenable as targets for classical protein, molecular, and biochemical studies.

There is currently considerable interest in how these individual components may interact with each other in *vivo*, and how the individual substrates are channelled through the pathway. Recent work with isolated chloroplasts from spinach and pea (Roughan and Ohlrogge 1996) suggest that these individual enzyme components, together with their co-factors, exist *in vivo* as a complex or metabolon. The metabolon is present in the chloroplast stroma and acetate is directly channelled through this complex into the lipid biosynthetic pathway. This implies that acetyl-CoA synthetase and acetyl-CoA carboxylase must also be included in this complex (Roughan and Ohlrogge 1996, Roughan 1997). A model has been proposed whereby the multi-enzyme complex spans

the stroma linking both the thylakoid and chloroplast envelope, with the thylakoid providing the energy for converting acetate into long chain fatty acids and the inner envelope accepting the products for processing into glycerolipids (Roughan and Ohlrogge 1996). However there has been no structural evidence to support this.

Several of the enzymes of the fatty acid biosynthesis (FAS) pathway in plants and bacteria have been the subject of crystallisation and structural studies using X-ray crystallography over the last few years. These include enoyl-ACP reductase (ENR) (Rafferty *et.al.*,1994 and 1995, Baldock *et al.*, 1998),and βketo-ACP reductase (βKR), (Fisher *et al.*, 1999 and 2000) from *Brassica napus* and *E.coli* (Rafferty *et al.*, 1998), where the complete atomic structure for both enzymes has been resolved. They are now the targets, based on the inhibition of the metabolic pathway of lipid synthesis, of important novel drug discovery. In particular ENR has been shown to be one of the target sites for the antibacterial agents ionazid and diazaborine in *Mycobacterium tuberculosis* and *E.coli* (Banerjee *et.al.*,1994, Baldock *et al.*,1996 and de Boer *et al.*,1999). It has also been shown to be the target site in several micro-organisms for the biocide triclosan (Roujeinikova *et al.*, 1999, Heath *et al.*, 1999 and 2000), a ubiquitously used additive in household cleaning agents. These studies now confirm that triclosan has a specific intracellular target, ENR, and it is a pico-molar inhibitor of this enzyme. Previously its anti-microbial activity was believed to be via the non-specific disruption of cellular

membranes (Regos *et al.*, 1979). The ability of a single amino acid mutation (G93V) in ENR (*fab* I) of *E.coli* to give the organism an acquired resistance to triclosan activity (Heath *et al.*, 1999) suggest that the wide spread use of this bacteriocide will eventually lead to the occurrence of resistant organisms and the need for new anti-microbial agents. More recently the crystal structure of acyl carrier protein synthase (AcpS) from *Streptococcus pneumoniae* (Chirgadze *et al.*, 2000), β-ketoacyl-ACP synthase II from *Synechocystis* (Moche *et al.*, 2001), and a crystal complex of holo-acyl carrier protein synthase together with ACP from *Bacillus subtilis* have been reported (Parris *et al.*, 2000).

1.3 Comparison of the FAS Pathway with the Polyketide Synthase Biosynthetic Pathway.

A similar, and equally important metabolic system to that of FAS, is the polyketide synthase (PKS), responsible for the biosynthesis of a large family of secondary metabolites, known as polyketides (Donadio *et al.*, 1991). These are produced by bacteria (mostly actinomycetes), fungi and plants and are extremely important as antibiotics and chemotheraputic agents (Sherman *et. al.*,1989, Donadio *et al.*, 1993). Although the two pathways contain many of the same enzymatic functions they differ in their substrate requirements, mode of action and the nature and complexities of the products they

produce. FAS begins with a 2 carbon (acetate) starter unit, is extended by condensation with 2 carbon units (from malonate) and processed through β-keto reduction, dehydration and enoyl reduction, for repeated cycles, to yield a symmetrical fatty acid usually C16 or C18 carbon units in length. PKS however can use a wide variety of starter and extender molecules (acetate, propionate, butyrate and in some cases more complex molecules). It can omit or curtail the cycle at any of the reduction or dehydration steps, and it yields a wide range of molecules containing methyl or ethyl side groups on an acyl chain of similar length to common fatty acids. Depending on the nature of the polyketide being produced the catalytic components of PKS consist of up to six covalently linked modules, each containing a number of putative FAS like domains. These can be organised in either a type I (Cortes *et al.*, 1990, Donadio *et al.*, 1991) or type II (Sherman *et al.*, 1989) structural arrangement. Putative domains for acyl carrier protein (ACP), acyltransferase (AT), dehydratase (DH), enoyl-reductase (ENR), β-ketoreductase (βKR), ketoacyl-ACP synthase (KAS), and thioesterase (TE) have all been identified within PKS systems and for a number of these components amino acid motifs conserved throughout both PKS and FAS systems have been found. Because of the use of polyketides in the pharmaceutical industry in a wide variety of roles, including antibiotics, immuno-supressants, and tumour suppressers there is a great interest in elucidating the structure of the PKS. However due to its size crystallisation efforts have not been successful. It may be that insight into its

structure may be gained from structural information obtained from studies on type II FAS of plants and bacteria.

1.4 The Biosynthetic Reactions of Plant FAS.

1.4.1 The Key Biochemical Steps.

Early studies on the biosynthesis of fatty acids in plants involved the use of [^{14}C] labelled acetate in metabolic labelling experiments. These experiments established that acetate can be incorporated into fatty acids and that these fatty acid are synthesised from acetyl-CoA and malonyl-CoA (for reviews see Stumpf, *et al.*, 1980 and 1987).

The key biochemical steps in this pathway can be considered to have three major components.

1) Supply of the primer acetyl-CoA.

The exact source of this could differ depending on the organelle and tissue, but it is believed to be via one of two possible routes. (1)The pyruvate dehydrogenase reaction inside the plastid (Rawsthorne 2002), or (2) from acetate production outside the plastid followed by its activation inside the plastid by acetyl-CoA synthetase (Harwood, 1979; Roughan and Ohlrogge 1996). In certain tissues clearly both pathways exist.

A possible third route in plastids via acetyl carnitine has been proposed by Masterson *et al* (1990) although this observation has not been confirmed by any other researcher worker and in fact Roughan *et al.*,1993 showed that 0.4mM L-acetyl carnitine could not compete with 0.2mM acetate as a substrate for FAS. Neither could they detect carnitine acetyl transferase activity in preparations of chloroplasts.

2) Supply of the chain extender malonyl CoA.

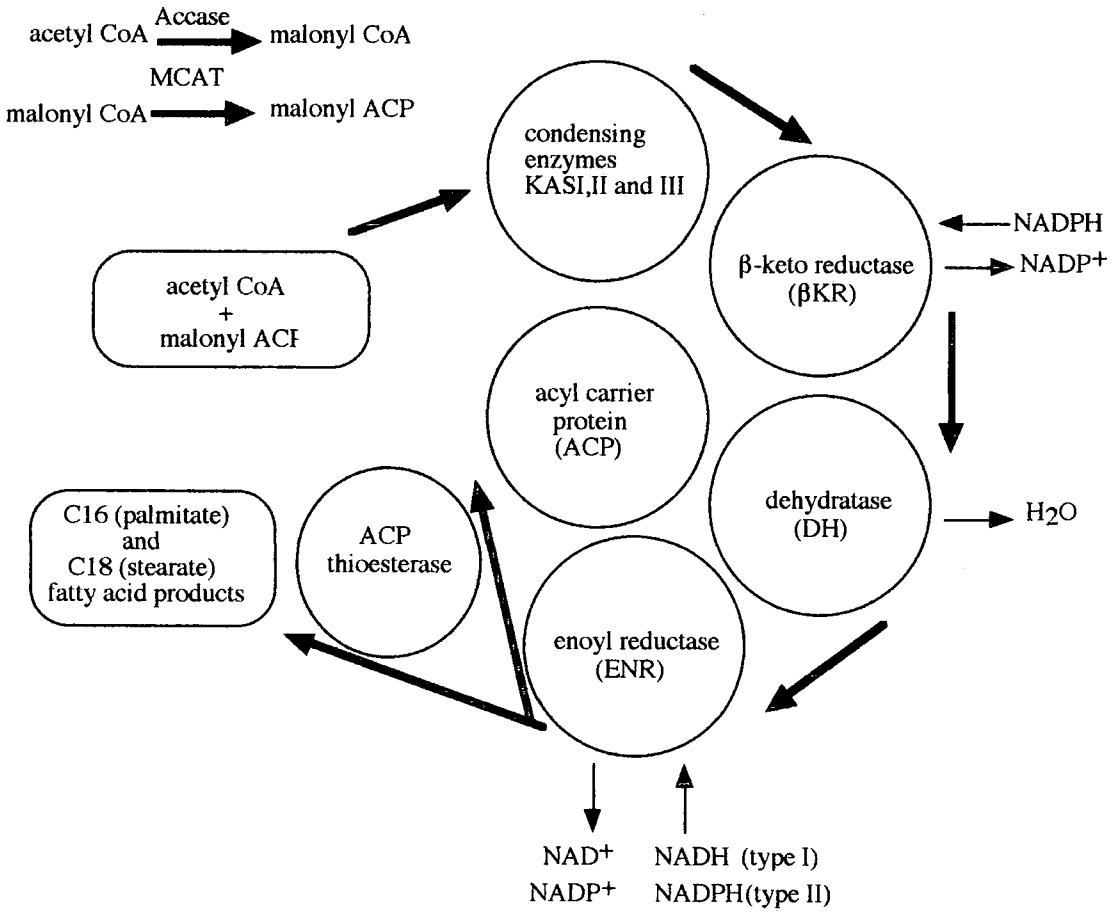
The conversion of acetyl-CoA to malonyl-CoA catalysed by acetyl-CoA-carboxylase (ACCase).

3) Growth of the acyl chain.

The synthesis of long chain fatty acids from acetyl-CoA and malonyl-CoA catalysed by the plastid located multi-enzyme system, fatty acid synthetase (FAS).

The current model describing the enzymatic reactions of plant FAS is shown in **Figure 1.3** and **Table 1.2** identifies the gene encoding the individual proteins, using the *E.coli* nomenclature. A comprehensive analysis of plant fatty acid biosynthesis and lipid synthesising genes from a diverse range of species can be found at WWW.canr.msu.edu (A catalogue of genes for plant glycerolipid biosynthesis at Michigan State University (Mekhedov *et al.*, 2000).

Figure 1.3 Diagrammatic representation of the sequential enzymatic reactions of the plant fatty acid synthetase (FAS).



Two carbon atoms are added to the growing fatty acid chain through repeated cycles of condensation (KAS), β-keto-reduction (βKR), dehydration (DH), and enoyl-reduction (ENR) until a C16 or C18 acyl-ACP is produced. The final C16 or C18 products can then be released from the ACP as free fatty acids via the action of ACP thioesterase. Details of the individual reactions are described in the text.

Table 1.2 Nomenclature of *E.coli* fatty acid and lipid biosynthesis genes.

Gene	Biosynthetic Component
aas	Acyl-ACP synthetase
accA	ACCase - carboxyltransferase α subunit.
accB	ACCase -BBCP subunit.
accC	ACCase – Biotin carboxylase subunit
accD	ACCase – carboxyltransferase β subunit
*acp*P	Acyl carrier protein
*acp*s	Holo-ACP synthetase
*fab*A	3-Hydroxydecanoyl-ACP dehydrase
*fab*B	3-Ketoacyl-ACP synthase I
*fab*D	Malonyl-CoA:ACP transacylase
*fab*E	Acetyl-CoA carboxylase
*fab*F	3-Ketoacetyl-ACP synthase II
*fab*G	3-Ketoacyl-ACP reductase
*fab*H	3- Ketoacetyl-ACP synthase III
fab I	NADH specific enoyl-ACP reductase
*fab*Z	3-Hydroxy myristol-ACP dehydratase
*pls*B	*sn*-Glycerol 3-phosphate acyltransferase
*pls*C	1-acylglycerol 3-phosphate acyltransferase
*fab*K	Flavo protein dependant enoyl ACP reductase

A brief summary of plant FAS and each of the enzyme components, β-ketoacyl-ACP synthases I,II and III (KAS I,II and III), β-ketoacyl-ACP reductase (βKR), β-ketoacyl-

ACP dehydrase (DH), enoyl-ACP reductase (ENR) and acyl-ACP thioesterase (TE) is described below.

This is followed by a more detailed review of malonyl-CoA:ACP transacylase (MCAT), acyl carrier protein (ACP) and glycerol- 3-phosphate acyltransferase (G3PAT) the three components central to this thesis.

Understanding the reactions involved in plant FAS improved rapidly in the early 1980's with several independent laboratories reporting the isolation of many of the soluble proteins from a variety of sources. In particular Shimakata and Stumpf (1982), Hoj and Mikkelsen (1983), Caughey and Kekwick (1982), and Shultz *et al.*, (1982), all reported the isolation and separation of individual enzyme components (β-ketoacyl-ACP-reductase, β- ketoacyl-ACP synthetase, acetyl-CoA:ACP transacylase, malonyl-CoA:ACP transacylase and acyl carrier protein) from different plant sources. These investigators used spinach, barley, avocado material and parsley suspension culture respectively.

In the years since this early work many laboratories have been involved in extensive research on plant FAS and lipid biosynthesis proteins and genes. Using molecular biology techniques many of the genes and cDNA's have been cloned and many of the protein gene products have been over-produced and biochemically characterised.

1.4.2 The Source of the Essential Precursor – Acetyl-CoA.

The synthesis of long chain fatty acids in the plastids of plants begins with acetyl-CoA, the exact source of which is still under debate, although ultimately it comes from carbon fixation during photosynthesis. As acetyl-CoA cannot cross membranes, it must be synthesised within the plastid and there are two main biosynthetic models proposed for this.

The first is via the pyruvate dehydrogenase (PDH) complex within the plastid (Kang and Rawsthorne 1984 and1996), which plays a central role in the conversion of imported substrates that are first converted to pyruvate and then via PDH to acetyl-CoA. Such imported substrates are glucose, phosphoenol pyruvate and pyruvate and recent papers have increased our knowledge of the relative importance of these substrates (Eastmond *et al.*, 1997, Ke *et al.*, 2000 and Rawsthorne 2002). It is unclear whether there is sufficient activity of PDH enzyme to account for *in vivo* FAS rates within the chloroplast. The second is that it is the product of the mitochondrial pyruvate dehydrogenase complex coupled to an acety-CoA hydrolase, generating free acetate (Murphy and Stumpf 1981, Liedvogel and Stumpf 1982) which is freely able to enter the chloroplast.

Once generated acetate is able to cross the chloroplast membrane where it is rapidly re-converted to acetyl-CoA by the action of plastidic acetyl-CoA synthetase. Several workers have analysed acetate concentrations in plant chloroplast, and Roughan (1995)

showed light generated levels sufficiently high to account for maximum rates of *de novo* fatty acid synthesis.

It is likely that the acetyl-CoA required for plastid localised fatty acid biosynthesis can be supplied by a combination of either / or both pathways, and that the preferred route may depend on whether the plastid is a photosynthesising chloroplast or a non green plastid (Kang and Rawsthorne 1996, Roughan and Ohlrogge1996). Plants have a high demand for acetyl-CoA for a wide variety of metabolic processes, other than the synthesis of fatty acids, these includes the synthesis of flavanoids, D-amino acids, polyhydroxybutyrates, xenobiotics and isoprenoids. This has led Wurtele and co-workers (1998) to hypothesise that plants have several mechanisms for its generation, and that its supply is differentially regulated across the various biosynthetic pathways.

Within the plastid acetyl-CoA is carboxylated to malonyl-CoA catalysed by acetyl-CoA carboxylase, before being used as substrate for the sequential addition of two carbon units during successive cycles of fatty acid synthesis yielding a final C18 acyl-ACP. (**Figure 1.3**).

The cycle begins with the condensation of malonyl-ACP with acetyl-CoA by the action of the condensing enzyme, 3- Ketoacetyl-ACP synthase III (KASIII) to produce aceto – acetyl ACP. This is subsequently reduced in a NADPH dependant reaction catalysed by β–ketoacyl-ACP reductase (βKR) to yield β–hydroxyoacyl-ACP, which is then

dehydrated by the action of β–hydroxyacyl-ACP dehydrase (DH) and finally reduced by enoyl-ACP reductase (ENR), to yield a four carbon (butyryl) acyl-ACP. In the further cycles of elongation, from C4 through to C16 acyl-ACP, the condensation reaction is catalysed by a different condensing enzyme KAS I. This condensing enzyme unlike KASIII utilises acyl-ACP as primer instead of the acyl-CoA. All of the remaining reactions of FAS elongation use the same ENR, βKR and DH components.

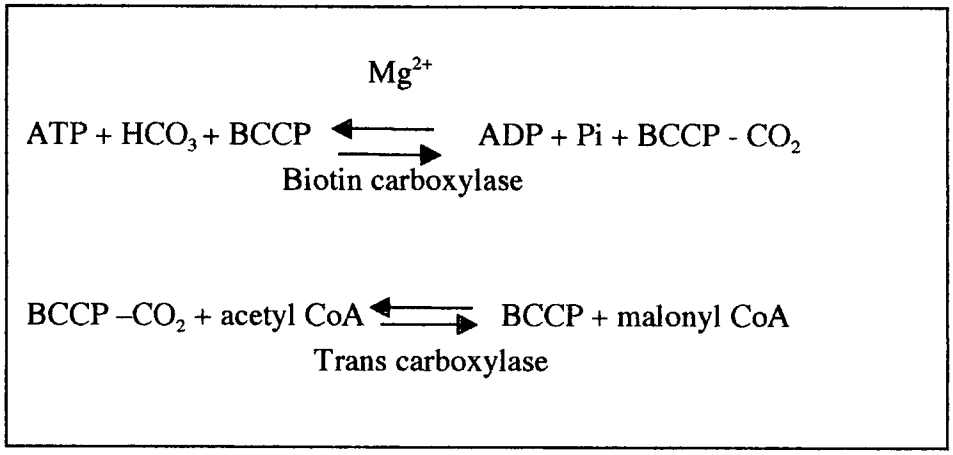
The palmitoyl ACP product of these successive cycles of elongation can be further elongated in a final cycle, yielding stearoyl–ACP, this time using the third condensing enzyme KASII, for the condensation reaction.

The final acyl-ACP products synthesised in the plastid can be utilised in one of two ways. They can either be used as a substrate for acyltransferase enzymes in the production of complex lipids within the plastid, or alternatively they can be hydrolysed to a free fatty acid and ACP by a plastid located thioesterase. These free fatty acids are then exported into the cytosol where they are converted into an acyl-CoA by a chloroplast envelope located acyl-CoA synthetase (Joyard and Stumpf 1982) and then used in the eukaryotic pathway for further lipid metabolism.

1.4.3 Acetyl CoA Carboxylase (ACCase)

ACCase (EC.6.4.1.2) is the first committed step in fatty acid biosynthesis (Turnham and Northcote 1983, Ohlrogge 1997), and catalyses the conversion of acetyl-CoA to malonyl-CoA. It is a biotin containing enzyme containing three functional domains. The domains are the biotin containing biotin carboxyl carrier protein (BCCP), the biotin carboxylase domain (BC) and the transcarboxylase domain (TC). The carboxylation of acetyl-CoA to malonyl-CoA is a two step reaction described in **Figure 1.4.**

Figure 1.4. The two step reaction of acetyl CoA carboxylase (ACCase).



In step 1 CO_2 is activated and attached to the biotin portion of BCCP in an ATP dependant reaction catalysed by biotin carboxylase .In step 2 the CO_2 is transferred from the carboxyl biotin to acetyl CoA, yielding malonyl CoA, a reaction catalysed by the transcarboxylase.

Two types of ACCase are found in plants (Sasaki *et al.*, 1995, Elborough *et al.*, 1996 and Schulte *et al.*, 1997). The first is a eukaryotic type I enzyme, similar to that found in mammals. It is composed of a single 220kDa polypetide containing all of the functional domains. The second is a multi-subunit (heteromeric) complex composed of four different protein subunits, the biotin carboxylase (BC), the biotin carboxyl carrier protein (BCCP) and the α and β subunits of the carboxyltransferase domain. This prokaryotic type II complex is plastid located and its function is the provision of malonyl CoA for *de novo* fatty acid biosynthesis.

In the cytoplasm of plants only type I ACCase has been found and its function is believed to be in providing malonyl CoA for fatty acid elongation, cuticular lipid biosynthesis (Kolattukudy, 1987) and the production of secondary metabolites, such as phytoalexins and flavanoids (Ebel *et al.*, 1984).The situation is more complex in the plastid where *de novo* fatty acid biosynthesis occurs. In the graminaceae the plastidic ACCase is type I whilst in non-graminaceae species type II is the dominant form. The basis of this difference in the organisation of ACCase between graminaceae and non-graminaceae species has been used to develop selective herbicides. Two grass specific herbicides, aryloxyphenoxypropionates and cyclohexanediones, both acting on the eukaryotic type I form of ACCase have been developed to control grass weeds in broad leaf crops (for review Harwood 1996).

1.4.4 Condensing Enzymes (3-Ketoacetyl-ACP Synthases).

Three independent condensing reactions are involved in the biosynthesis of C18 fatty acids from a C2 precursor in plants.

In the early studies on the condensing reactions of plant fatty acid biosynthesis using spinach leaf (Shimikata and Stumpf 1982and 1983), it was shown that KASI could utilise short chain fatty acids up to C16 (palmitic acid), and that the final condensation from C16 to the C18 (stearate) product was catalysed by KASII. The sensitivity of KASI to the antibiotic cerulenin was used to establish the chain length selectivity of the two enzymes (Shimakata and Stumpf 1983).

Up until 1987 the first condensation reaction of both plant and bacterial fatty acid biosynthesis was thought to between acetyl-ACP and malonyl-ACP. The acetyl-ACP being generated by acetyl-CoA-ACP transacylase (ACAT). The established model was that ACAT supplied acetyl-ACP which was condensed with malonyl-ACP by KAS I to yield acetoacetyl-ACP and that this was repeated through successive rounds of FAS until a C16 product was achieved. A final round of condensation to C18 was catalysed by KAS II.

A third, short chain utilising condensing enzyme (KAS III) was discovered in *E.coli* (Jackworski and Rock 1987) and subsequently in plants (Jaworski *et al.*, 1989, Clough *et*

al., 1992 and Gulliver and Slabas, 1994) which was cerulenin insensitive and uses exclusively acetyl-CoA and not acetyl-ACP. It is now believed that this enzyme catalyses the first condensation reaction, of plant fatty acid biosynthesis, between acetyl-CoA and malonyl-ACP **Figure 1.3**. During the course of the reaction the acetyl group is transferred from CoA to a cysteine thiol group on the enzyme, so making an acetyl enzyme intermediate.

1.4.5.Reductases.

The first reductive step in FAS is catalysed by β-ketoacyl-ACP reductase (βKR) where the β-ketoacyl-ACP product of each condensation reaction is reduced to β-hydroxyacyl-ACP. It has been purified from a number of plant sources including spinach, avocado and rape (Shimakata and Stumpf 1982 and Sheldon *et al.*, 1988 and 1992) and shows a preference for NADPH as its pyridinucleotide cofactor. cDNA clones have been isolated from *Brassica napus*, *Arabidopsis* (Slabas *et al.*, 1992) and *Cuphea* (Klein *et al.*, 1992) and the deduced amino acid sequence for all of the plant genes show strong (41-55%) identity to the *Nod* G gene product from *Rhizobium meliloti* (Slabas *et al.*, 1992). Enoyl-ACP reductase (ENR) catalyses the second and final reductive step in each cycle of FAS when the trans-double bond of enoyl-ACP is reduced to produce a saturated ACP product. Two forms of ENR (type I and II) have been identified in *E.coli* (Weeks and

Wakil 1968), safflower (Shimakata and Stumpf 1982) and rape seed (Slabas *et al.*, 1986). The two forms differ in their specificity for their hydrogen donor, type I requires NADH whereas type II requires NADPH for activity. ENR was first purified from plants when the type I form was isolated from spinach leaves (Shimakata and Stumpf 1982) and it has been subsequently isolated from avocado mesocarp (Caughey and Kekwick 1982) and *Brassica napus* seed where both type I and II were found (Slabas *et al.*, 1986). The cDNA for the *Brassica napus* enzyme was cloned and found to have a leader sequence of 73 amino acids which targeted the nuclear encoded protein to the plastid (Kater *et al.*, 1991). Four genes for ENR have been identified in *Brassica napus*, two of which were inherited from each of the ancestors *Brassica oleracae* and *Brassica campestris* (Kater *et al.*,1991). Four isoforms were identified by 2-dimensional PAGE and western blotting of *Brassica napus* seed and leaf extracts (Fawcett *et al.*, 1994). All four were present in both tissue but showed differential expression. The two most abundant seed isoforms were also found to be the most abundant in leaf. mRNA abundance studies showed that levels increased just prior to the onset of lipid deposition and that the level in seeds was 15 times the level in leaf tissue (Fawcett *et.al.*, 1994). Both the *E.coli* (Baldock *et al.*, 1998) and the *Brassica napus* (Rafferty *et al.*, 1995) enzyme have been successfully crystallised, and their structure and catalytic mechanism resolved. ENR has been shown to be one of the target sites for the antibacterial agents ionazid and diazaborine in

Mycobacterium tuberculosis and *E.coli* (Banerjee et.al.,1994, Baldock *et al.*,1996 and de Boer *et al.*,1999). It has also been shown to be the target site in several micro-organisms for the biocide triclosan (Roujeinikova *et al.*, 1999, Heath *et al.*,1999 and 2000), and in higher plants the herbicide diflufenican (Ashton *et al.*,1994).

1.4.6. β-Hydroxyacyl-ACP Dehydrase

Two forms of β-hydroxyacyl-ACP dehydrase (DH) the enzyme which catalyses the removal of water from β-hydroxyacyl-ACP to yield 2,3trans enoyl–ACP have been isolated from *E.coli*. The genes for them are known as *fab*A (Birge and Vagelos 1972) and *fab*Z (Heath and Rock 1996) and their main difference is that *fab*A shows both dehydratase and isomerase activity whilst *fab*Z only has dehydratase activity. Although their substrate specificities overlap *fab*Z prefers short chain β-hydroxyacyl-ACPs and *fab*A medium chain.

DH was purified to homogeneity from spinach and partially purified from developing safflower seeds (Shimakata and Stumpf 1982) but no plant gene has been cloned to date. Work on the cDNA cloning of DH from *Brassica napus* using a sequence derived from a *Ricinus* EST database with 75% identity to *E.coli fabZ* has been reported from this laboratory (Doig *et al.*, 1998 and Doig PhD thesis 2001).

1.4.7 Thioesterase

The termination of the acyl chain elongation process of plant fatty acid synthesis is brought about by the action of an acyl-ACP thioesterase (TE). The enzyme cleaves the fatty acid from the sulphydryl group of ACP and releases it as a free fatty acid, which is exported from the plastid to the cytosol. Here it is converted to acyl-CoA by the action of a chloroplast envelope located acyl-CoA synthetase (Joyard and Stumpf 1982) and utilised in complex lipid biosynthesis in the eukaryotic lipid synthesis pathway.

A chloroplast located TE that shows greatest activity for 18:1, less for 16:0 and little activity for 18:0 ACP has been known about for several years (Harwood 1996). It has been purified to homogeneity from soybean seeds (Kinney *et al.*, 1990), oil seed rape (Hellyer *et al.*, 1992) and squash cotyledons (Imai *et al.*, 1992) and studied in many other plant species including leek (Liu and Post-Beittenmiller 1995) and *Arabidopsis* (Dorman *et al.*, 1995).

Thioesterases with short and medium chain specificity have been identified in *Cuphea* (Dormann *et al.*, 1993) and California bay laurel *(Umbellularia californica)* (Pollard *et al.*, 1991) two plants species which have medium chain (C8-14 and C12 respectively) fatty acids as the major components in their triacylglycerols.

When the California bay medium chain length TE was transformed into *Arabidopsis* plants, which normally make long chain fatty acids the major fatty acid produced was a

C12 fatty acid (Voelker *et al.*, 1992). This has great future potential for the use of standard crops such as oil seed rape in the production of medium chain fatty acids, which have many commercial uses, including detergents, lubricants and food (Harwood 1996).

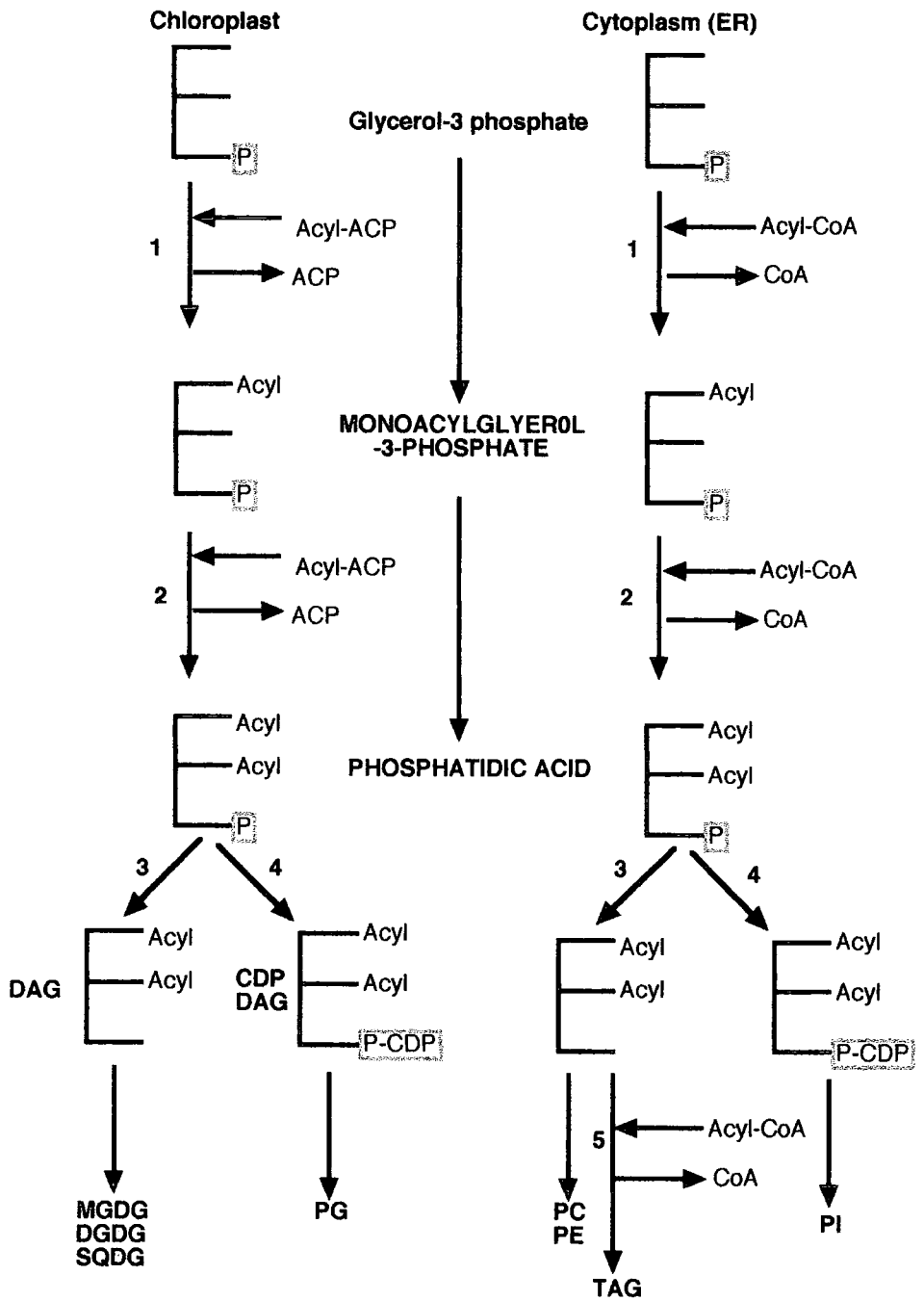
1.5 Complex Lipid Biosynthesis

Complex lipid biosynthesis in plants occurs in two separate compartments, the plastid and the cytoplasm (Heinz and Roughan 1983, Browse and Somerville 1991). In the plastid it occurs via a pathway known as the prokaryotic pathway, where the acyl-ACP products of *de novo* fatty acid biosynthesis are utilised directly. In the cytoplasm at the endoplasmic reticulum (ER), via a pathway using acyl-CoA derivatives known as the eukaryotic pathway (Kennedy 1961). Fatty acids exported from the plastid are converted to CoA derivatives at the chloroplast envelope (Sommerville and Browse 1991).

The two pathways follow essentially the same steps (**Figure 1.5**), although the enzymes involved have different substrate specificities and different sub-cellular locations. The acyl group distribution of the complex lipids synthesised by the two pathways are also different due to the acyl group specificity of the acyltransferase enzymes found in the different locations and the availability of the substrates at each location (**Figure 1.5**).

In the plastid (prokaryotic pathway) the final products, predominantly C18 (*sn1*) and C16 (*sn2*) diacylglycerols are used to synthesise phosphotidylglycerol (PG),

Figure 1.5 The two compartment (chloroplast and cytoplasm) model of complex lipid biosynthesis in plants.



1= glycerol - 3- phosphate acyltransferase (G3PAT), 2= 2lysophosphatidic acid acyltransferase (LPAAT) 3= phosphatidic acid phosphatase (PAP), 4= Phosphatidate cytidyltransferase,5= Diacylglycerol acyltransferase (DAGAT).

Abbreviations: (MGDG) monogalactosyl diacylglycerol, (DGDG) digalactosyl diacylglycerol, (SQDG) sulfoquinovosyl diacylglycerol, (PG) phosphatidylglycerol, (PC) phosphatidylcholine, (PE)phosphatidyl ethanolamine, (PI) phosphatidylinositol, (TAG) triacylglycerol.

monogalactosyldiacylglycerol (MGDG), digalactosyldiacylglycerol (DGDG) and sulphoquinovosyldiacylglycerol (SQDG), which are all utilised in plastid membrane synthesis.

In the endoplasmic reticulum (eukaryotic pathway) the products are C16 or C18 (*sn1*) and C18 (*sn2*) lipids which are used for the synthesis of phospholipids, phosphotidylcholine (PC), phosphotidylethanolamine (PE) and phosophotidylinositol (PI). These are all used in the synthesis of extra-plastid membranes.

It is also at the ER, that the triacylglycerol (TAG) storage lipids are synthesised. Here a final acyl group is added to the diacylglycerol (DAG) molecule, by a diacylglycerol acyltransferase (DAGAT), the only acyltransferase unique to the acyl-CoA dependant TAG pathway. The flux between the prokaryotic and eukaryotic pathways varies between plant species, in some higher plants PG is the only prokaryotic lipid produced and all others have to be imported from the eukaryotic pathway. In others such as *Arabidopsis* both pathways contribute equally to the synthesis of plastid membrane lipids (Browse *et al.*, 1986).

The precursor of all plant glycerolipid biosynthesis is glycerol-3-phosphate (G3P). It is to this backbone that acyl groups are attached by the action of specific acyltransferase enzymes. In plants there are a limited number of acyl groups incorporated into glycerolipids, they are mainly 18:1, 18:2, 18:3, 16:0 and 16:3 fatty acids, which together

make up almost 80% of the acyl chains found in membrane lipids (Ohlrogge and Browse 1995).

There are three separate acyltransferase enzymes which catalyse the acylation at positions 1, 2, and 3 of the backbone respectively (**Figure 1.5**). These are (I) glycerol–3 phosphate acyltransferase (G3PAT) [EC 2.3.1.15], (II)1-acyl-glycerol-3-phosphate-2-acyltransferase (LPAAT) [EC2.3.1.51] and (III) diacylglycerol acyltransferase (DAGAT) [EC 2.3.1.20] (for reviews see Frentzen 1990, and Slabas and Brough 1997).

All acyltransferase enzymes are membrane associated, except for a soluble G3PAT enzyme located in the chloroplast of plants and responsible for the acylation with acyl-ACP at the *sn1* position of glycerol-3-phosphate. It has been suggested that the competition between the activity of this G3PAT and that of the long chain acyl-ACP thioesterase, (responsible for the supply of free fatty acids for release from the chloroplast), determines whether an acyl group enters the prokaryotic or eukaryotic pathway (Heinz and Roughan 1983). Studies on this enzyme will be reported in this thesis therefore a more detailed coverage of this enzyme follows in section **1.8**

1.6.Malonyl CoA:ACP Transacylase (MCAT).

The two-carbon starter unit for successive rounds of *de novo* fatty acid biosynthesis in plants and bacteria is malonyl-ACP. This is derived from malonyl-CoA during a transfer

reaction catalysed by malonyl coenzyme A:ACP transacylase (MCAT), which converts the malonyl group from the coenzyme A to the ACP in an exchange reaction to yield malonyl-ACP.

MCAT was first purified (Ruch and Vagelos 1973a) and characterised (Ruch and Vagelos 1973b) from *E.coli*, where it was found to be a monomeric protein with a Mr of 36.6 kDa on SDS-PAGE and 37kDa by gel filtration chromatography. A reaction mechanism was proposed whereby a malonyl-enzyme intermediate is formed when the carboxyl group of the malonate is esterified to the hydroxyl group of an active site serine residue in the protein (Ruch and Vagelos 1973b). This serine residue lies within the conserved active site pentapeptide (Gly-His-Ser-X-Gly), a sequence motif reported to be a consensus sequence for transacylases, also found in type I FAS acetyl, malony, and palmityl transferase domains, (Verwoert *et al.*, 1992).

MCAT has subsequently been partially purified and characterised from a number of plant sources summarised in **Table 1.3**.

The gene for MCAT (*fab*D) was first cloned from *E.coli* by complementation cloning of the temperature sensitive mutant known as *fab*D (L2-89) (Verwoert *et al.*, 1992 and Mangnuson *et al.*, 1992). The gene encodes a protein of 358 amino acids, with a molecular mass 32 kDa, containing the conserved motif described above. Sequence

analysis revealed that the MCAT gene was part of an operon (*fab*) containing at least three fatty acid biosynthesis genes, with 3-ketoacyl-ACP synthase upstream and β-keto-ACP reductase downstream (Verwoert *et al.*, 1992).

Table 1.3 Summary of plant MCAT purification's reported in the literature.

Plant source	Molecular mass (kDa)	Reference
Avocado fruit	40.5	Caughey and Kegwick (1982)
Barley chloroplasts	41	Hoj and Mikkelsen (1982)
Safflower seeds	22	Shimakata and Stumpf (1982)
Spinach leaves	31	Stapleton and Jaworski (1984)
Soybean	43 (two isoforms both present in leaf one predominant in seeds)	Guerra and Ohlrogge (1986)
Leek leaves	38 +45 (two isoforms)	Lessire and Stumpf (1983)
Anabaena	36	Stapleton and Jaworski (1984)
Cuphea	27.5	Bruck *et al.*, (1994)

The crystal structure at 1.5 Å resolution has been reported for *E.coli* MCAT following the over-production of the *E.coli fab*D gene and the production of three (mercury, platinum and gold) heavy metal derivatives (Serre *et al.*, 1994 and Serre *et al.*, 1995).

The *E.coli* MCAT gene has been expressed in transgenic rapeseed and tobacco under the control of the rapeseed napin promoter for developmental seed expression, and linked to the leader sequence of enoyl-reductase for chloroplast targeting (Verwoert *et al.*, 1994).

Successful chloroplast targeting was confirmed by immunogold labelling studies, and up to 55 times normal MCAT activity was achieved at the end of seed development. This increased activity did not however effect the overall lipid composition of the transgenic plants indicating that MCAT does not catalyse a rate limiting step in plant fatty acid biosynthesis when grown under these conditions (Verwoert *et al.*, 1994).

Attempts to clone MCAT from a *Zea mays* cDNA expression library using complementation of the *fab*D (LA2-89) mutant failed to isolate a gene for MCAT.

A cDNA encoding a GTP binding protein of the ARF family was however isolated (Verwoert *et al.*, 1995), it remains unclear how this ARF protein can compensate for the MCAT mutation and allow the organism to grow at restrictive temperatures.

In this thesis the first cDNA cloning of a plant MCAT is reported (Simon and Slabas 1998).

1.7 Acyl Carrier Protein (ACP).

Acyl Carrier Protein (ACP) as its name suggests, is a protein that is the central carrier of the acyl chain between each of the consecutive reactions of the fatty acid synthetase. It is also a component of acyl- ACP desaturation (McKeon and Stumpf 1982) and plastid localised acyl-transferase reactions (Frentzen *et al.*, 1983). It was the first of the proteins involved in the fatty acid biosynthesis pathway to be purified from both *E.coli* (Majerus

et al., 1964) and plants (Simoni *et al.*, 1967). The *E.coli* protein was also chemically synthesised in a biologically active form (74 out of 77 residues) in an effort to understand something of its structure (Prescott and Vagelos 1972).

In the years since its first purification from a plant source ACP has subsequently been isolated from many plant species, these include, barley (Hoj and Svedson, 1983), spinach (Kuo and Ohlrogge, 1984) and oil seed rape (Slabas *et al.*, 1987). The primary amino acid sequence of all of these plant ACPs is highly conserved, and NMR structural studies have shown a high degree of structural homology between the spinach and the *E.coli* protein (Oswood *et al.*, 1997). *In vitro* experiments have demonstrated that *E.coli* ACP can substitute for plant ACP in a plant synthetic system. Curiously *E.coli* ACP stimulates the reaction rate to higher levels than the native plant ACP (Simoni *et al.*, 1967). The reason for this is still unknown, but may be resolved as crystal structural studies progress.

Multiple isoforms of ACP have been identified in a number of plant species, these include *Arabidopsis* (Hlousek-Radojcic *et al.*, 1992), rape (Safford *et al.*,1988 and De Silva *et al.*, 1990), barley (Hoj and Svendsen 1984), spinach (Ohlrogge and Kuo 1985), *Cuphea* (Schutt *et al.*, 1998) and coriander (Chung Suh *et al.*, 1999). Several of these isoforms are constitutively expressed, whereas others have been shown to be tissue specifically expressed (Battey and Ohlrogge 1990, de Silva *et al.*, 1990, Hlousek-Radojcic *et al.*,1992) and involved in the biosynthesis of unusual fatty acids, for example

the synthesis of petroselinic acid in coriander (Chung Suh *et al.*, 1999). The role of these individual isoforms is still not fully understood although it has been shown that specific isoforms are required for optimal activity of Δ4 and Δ6 desaturase enzymes present in the seeds of coriander (Chung Suh *et al.*, 1999).

Analysis of cDNA clones for ACP isolated from *Brassica napus* seed showed that they comprise a large nuclear encoded multi-gene family with six unique genes coding five different mature proteins (Safford *et al.*, 1988, de Silva *et al.*, 1990). It is suggested that these multiple genes could regulate the levels of ACP in response to the differing demands between normal constitutive levels and increased levels during triglyceride biosynthesis throughout seed development (de Silva *et al.*, 1990).

A correlation between ACP levels and the increase in lipid synthesis during oil accumulation in developing rapeseed (Slabas *et al.*, 1987) and soybean (Ohlrogge and Kuo 1984) was made and this showed that ACP levels increased dramatically just prior to lipid accumulation, suggesting that ACP genes would be targets for promoters for both tissue and temporal regulation.

Tissue specificity of isoforms of ACP have been characterised in spinach (Schmid and Ohlrogge 1990), barley (Hansen 1987) , *Ricinus* (Ohlrogge and Kuo 1985) oilseed rape and *Arabidopsis* (Hlousek-Radojcic *et al.*,1992). In spinach there are two major isoforms type I and type II. Both are found in leaf tissue whereas only type II is expressed in seed

and root. Seed specific isoforms have been identified in *Brassica napus* and *Arabidopsis* that are believed to be important in the synthesis of storage triglycerides (de Silva *et al.*, 1990, Hlousek-Radojcic *et al.*,1992). In barley three isoforoms have been identified, ACP I and II were chloroplast located but with different N-terminal sequence and type III a minor leaf isoform with no assigned function (Hoj and Svendsen 1984).

In bacteria and plants ACP is a small acidic protein with a molecular mass, depending upon the organism of between 8 and 10 kDa, (Vanaman *et al.*, 1986) whereas in yeast and animals ACP is found as a functional domain on the large multifunctional fatty acid synthetase polypeptide (Wakil *et al.*, 1983).

ACP is synthesised as an apo-protein, which in order for it to be functionally active undergoes a post-translational modification to an holo-protein. During this process a 5'phosphopantetheine prosthetic group derived from coenzyme A is attached to an essential serine residue (serine 36 in *E.coli*) (Flugel *et al.*, 2000). This pantethenylation is catalysed in a Mg^{+} dependant reaction by the action of a holo ACP synthetase enzyme (HAS) (EC 2.7.8.7) (Elovson and Vagelos 1968). Acyl groups are then attached as thioesters to the sulphydryl group at the terminus of this phosphopantetheine arm (Majerus *et al.*, 1965). Holo ACP synthetases have been identified in both the cytosol and the plastids of plants, where they play a role both in synthesising holo-ACP and also in regenerating holo-ACP from ACP which has lost its prosthetic group (Elhussein *et al.*,

1988, Fernandez and Lamppa 1990). The presence of this chloroplast located HAS provides evidence that ACP does not need to be in the holo form in order for it to be imported into the chloroplast (Fernandez and Lamppa 1990)

A mitochondrial form of ACP from plants (pea and potato) and eukaryotic micro -organsims has been reported (Chuman and Brody 1989). This was located using an antibody raised against mitochondrial ACP from *Neurospora crassa*, and although mitochondrial *de novo* fatty acid biosynthesis was not demonstrated the authors suggest that the presence of acylated forms of ACP show it may be present.

To date there is no crystal structure of holo-ACP from any source, although reports of the progression from one and two dimensional nuclear magnetic resonance (NMR) structures from *Escherichia coli* acyl-ACP have been published over the years (Mayo *et al.*, 1983, Holak and Prestergard 1986). This has lead to a complete low-resolution three-dimensional NMR structure being published in 1988 (Holak *et al.*, 1988 a and Holak *et al.*, 1988b). Also a NMR structure from spinach ACP has been produced (Kim and Prestergard 1989 and Kim *et al.*, 1990) and its structural characteristics compared to those of *E.coli* ACP.

Despite many efforts the inability to obtain robust crystals of ACP which diffract well in a X-ray beam has meant that there has been no further progress on the structural studies, using X-ray crystallography of ACP since these early reports. Except for the report

recently of a crystal structure of *Bacillus subtilis* holo-ACP in complex with holo-ACP synthase (Parris *et al.*, 2000).

ACP from a *Streptomyces glaucescens* type II polyketide (PKS) has been shown (Zhou *et al.*, 1999) to have the ability to catalyse both its own self malonylation, and also the malonylation of the separate fatty acid synthetase ACP. It therefore appears that ACP from this system has a MCAT type activity. The MCAT activity is at a level two orders of magnitude smaller than that catalysed by malonyl CoA:ACP transacylase (*fab*D) on a specific activity basis and only proceeds in the presence of high concentrations of PKS ACP. This observation is consistent with the inability to locate *fab*D homologs within most type II polyketide synthase gene clusters.

1.8 Glycerol-3-Phosphate-1-Acyltransferase (G3PAT).

Glycerol-3-phosphate-1-acyltransferase (G3PAT) [EC 2.3.1.15] exists in plants in three locations, in the cytoplasm, the mitochondria, and the plastid. The plastid form is soluble and is stromal located (Joyard and Douce 1977, M^{c}keon and Stumpf 1982). It is responsible for the acylation of the *sn1* position of glycerol-3-phosphate with an acyl group from the acyl-ACP product of *de novo* fatty acid biosynthesis, to yield

1-acylglycerol-3-phosphate (lysophosphatidic acid) (Roughan and Slack 1982, Frentzen *et al.*, 1983).

The cytoplasmic form is bound to the ER (Frentzen 1990) and the mitochondrial form is bound to the outer membrane (Frentzen *et al.*,1990) these both use acyl-CoA as their acyl donors.

The soluble G3PAT has been purified and characterised from a number of plant sources, including pea, spinach (Betrams and Heinz1981), and squash (Nishida *et al.*, 1987). In squash three isoforms were identified AT1, AT2 and AT3 with *pI*'s of 6.6, 5.6 and 5.5 and molecular masses of 30, 40 and 40 k Da respectively, with AT1 present as the most abundant isoform (Nishida *et al.*,1987). To date the precise N-terminus of the plant G3PAT has not been determined, although it has been predicted for *Arabidopsis* and squash using chloroplast target sequence predictions (Murata and Tasaka 1997). N-terminally truncated forms are however catalytically active.

A membrane bound cytoplasmic G3PAT from avocado mesocarp has also been partially purified, revealing three isoforms of this enzyme, with molecular masses of 70, 60 and 54 kDa (Eccleston and Harwood 1995).

Plant chloroplast (prokaryotic) lipids are found with either C16 or C18 fatty acids at the *sn1* position of their diacylglycerol molecules. Murata (1982) discovered a close correlation between the chilling tolerance of plants and the nature of the fatty acid at the

sn1 position of phosphatidylglycerol, an observation later confirmed by Roughan in 1985 and Kenrick and Bishop in 1986. Chilling tolerant plants were shown to have low amounts (0-20%) of 16:0 and high amounts of unsaturated C18:1 fatty acids at the *sn1* position whereas chilling sensitive plants had much higher levels (25-65%) of 16:0 and lower amounts of C18:1. This increased level of *trans* monounsaturated PG present in chilling tolerant plants allows their membranes to remain more fluid at low temperatures preventing phase partitioning of the lipids. This hypothesis was tested when the cDNAs for G3PAT from both *Arabidopsis* (chilling tolerant) and squash (chilling sensitive) where transformed into tobacco plants.

The plants containing the *Arabidopsis* construct became more chilling resistant whereas those containing the squash construct became much more sensitive (Murata *et al.*, 1992). A cDNA clone encoding the soluble G3PAT was first isolated from squash chloroplasts (Ishizaki *et al.*, 1988), and then subsequently from many other plant species, including, pea, (Weber *et al.*, 1991), cucumber (Johnson *et al.*, 1992), *Arabidopsis* (Nishida *et al.*, 1993), spinach (Toguri 1994) and french bean (Fritz *et al.*, 1995). These all show a high degree of sequence conservation at the deduced amino acid level (Slabas and Brough 1997). A chloroplast G3PAT gene (ATS1) which contains 11 introns in the protein coding region and one in the 3' un-translated region was also isolated from *Arabidopsis* (Nishida *et al.*, 1993).

Because of the importance of G3PAT substrate specificity to chilling resistance in plants, the substrate selectivity of the protein has been extensively studied from both chilling tolerant (pea, *Arabidopsis*, spinach) and chilling sensitive (squash) plants.

In competitive assays where the enzyme from pea and spinach, were presented with acyl-ACPs and acyl-CoAs simultaneously, a preference for the ACPs was observed (Frentzen *et al.*, 1983) and oleoyl-ACP was preferred from a mixture of palmitoyl, stearoyl and oleoyl-ACP.

This distinct preference for acyl-ACPs over acyl-CoAs distinguishes the soluble chloroplast located G3PAT from the membrane bound enzymes that can only utilise acyl-CoAs.

Substrate specificity and kinetic assays have been performed on the enzyme from a number of plant sources, including pea, spinach (Betrams and Heinz, 1981, Frentzen 1993) and squash (Frentzen *et al.*, 1987, Nishida *et al.*, 1987). Many of these have been performed using the commercially available acyl-CoAs rather than the natural substrate acyl-ACP, which have to be synthesised in the laboratory.

However competitive assays using both acyl-CoAs and acyl-ACPs have clearly demonstrated that acyl-ACPs are the preferred substrate (Frentzen *et al.*, 1983) consistent with the fact that both the enzyme and the acyl-ACP substrates (C16:0 and C18:1) are located in the chloroplast.

X-ray crystallography and structural studies on G3PAT from squash have lead to the resolution of the complete atomic structure of this enzyme and an understanding of the catalytic mechanism (Turnbull *et al.*, 2001), and site directed mutagenesis studies reported here (**Chapter 5**) have contributed to this work.

1.9 The Aims of this Thesis.

This thesis will concentrate mainly on trying to obtain missing information on the components of FAS, both plant and bacterial, particularly with a view to future structural studies. It will take particular advantage of the ability to produce recombinant proteins for use in site directed mutagenesis and molecular interaction experiments.

Some powerful examples of how these types of study have been used to advantage in structural studies in the past, would be, (I) the introduction of seleno-methionine into proteins to help overcome heavy metal incorporation problems (Hendrickson *et al.*,1990, Bottomley *et al.*, 1994) during phase resolution in crystallography studies, (II) the introduction of cysteine residues via mutagenesis as sites for heavy metal binding as an alternative approach for the incorporation of a specific heavy metal site into the crystal structure of *Brassica napus* β-ketoacyl-ACP reductase (Fisher *et al.*, 2000).

There are also many further examples cited in the literature where site directed mutagenesis studies have been used to alter the substrate specificity of enzymes and

determine important residues in their active site. Three representative examples relevant to this work are cited here. (I) A conserved histidine residue is essential for glycerolipid acyltransferase catalysis (Heath and Rock, 1998), (II) Mutational analysis of plant enoyl-ACP reductase in *Escherichia coli.* (Stuitje *et al.*, 1998), (III) Reaction mechanism of recombinant 3-oxoacyl-(acyl carrier protein) synthase III from *Cuphea wrightii* embryo, a fatty acid synthase type II condensing enzyme (Abbadi *et al.*, 2000).

The complimentary approaches of protein crystallography and molecular structure, site directed mutagenesis, and protein interaction studies should all combine to enhance our understanding of the nature of the interaction between the components of this pathway. Central to all these issues is a good understanding of the plant lipid synthesising enzymes particularly plant FAS and it's core protein ACP and the soluble plastid located acyltransferase glycerol–3-phosphate acyltransferase (G3PAT) which is central in the partitioning of lipids between the prokaryotic and eukaryotic pathways.

The specific aims of this thesis are as follows.

(I) Attempt to clone the first cDNA for MCAT (*fab*D) from a plant source and to prove that it is functionally active. This is because it is one of the two missing components of plant FAS that has not been successfully cloned. The other component is dehydrase (fabZ) the subject of other research in this laboratory.

(II) To find a form of ACP from either a plant or bacterial source that will crystallise and be stable in a X-ray beam and provide the first three-dimensional crystal structure of this protein. This will involve obtaining wild type and recombinant protein from a number of sources, investigating and characterising chemical and enzymatic acylation reactions, production of mutant forms via site directed mutation reactions to aid in the crystal structure solution.

(III) Use these over-expressed forms of ACP together with site directed mutagenesis of G3PAT to investigate the substrate binding sites and specificity and selectivity of G3PAT, particularly as most studies to date have used acyl-CoA's and not the natural acyl-ACP substrates. This may give clues to the nature of the interactions of the G3PAT / ACP complex and how it works.

Chapter 2

Materials and Methods.

2.1 General Biochemical and Molecular Biology Reagents and Methods.

All chemicals and biochemicals used in this work were obtained from either Sigma Cemical Co. (Fancy Road, Poole, Dorset, UK), MERCK Ltd (BDH Merck House Poole, Dorset, UK) or Fisher Scientific Ltd (Bishop Meadow Road, Loughborough, UK) unless otherwise specified. All reagents were of Analar, Molecular Biology, Ultrapure or equivalent quality.

Electrophoresis reagents including acrylamide, sodium dodecyl sulphate (SDS), ammonium persulphate, N,N,N'N'-tetramethylethylene diamine (TEMED) and coomassie blue R-250 were all obtained from Bio-Rad laboratories Ltd.,(Hemel Hempstead, Hertfordshire, UK).

Dithiothreitol (DTT), isopropyl thio-β-D-galactosidase (IPTG), 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal), ampicillin and kanamycin were all purchased from Melford laboratories Ltd., (Chelsworth, Ipswich, Suffolk, UK).

Restriction enzymes and molecular biology reagents were obtained from either Boeheringer Mannhiem UK (Diagnostics and Biochemicals Ltd. Lewes, Sussex UK), Stratagene (La Jolla, CA. USA) or Promega (Maddison, WI, USA).

BioTaq® polymerase was obtained from Bioline (Humber Road, London UK) and Ultrapure solution dNTPs were from Amersham Pharmacia Biotech (St Albans, UK).

Oligonucleotide primers used for DNA sequencing, PCR reactions and site directed mutagenesis reactions were synthesised from either MWG-Biotech GmbH, (Germany), PE-Applied Biosystems (Warrington UK) or Cruachem Ltd (Glasgow UK).

The QuickChange® Site Directed Mutagenesis Kit used in all mutagenesis experiments was from Stratagene (La Jolla, CA. USA).

MonoQ®, Hi-LoadQ®, and PD10® chromatography columns were from Amersham Pharmacia Biotech. (St Albans, Hertfordhire,UK). Perfusion® chromatography media and chromatography columns were obtained from PerSeptive Biosystems (Framingham, MA.USA).

Matrix assisted laser desorption time of flight mass spectrometry (MALDI-tof) reagents were obtained from either Applied (PE) Biosystems Ltd (Warrington, UK), Rathburn Chemicals Ltd (Walkerburn, Scotland, UK) or from Sigma / Aldrich (Poole, Dorset, UK) and kept exclusively for work with MALDI-tof.

All biochemical procedures were carried out at 4°C unless otherwise stated and Milli-Q (MQ) grade water (Millipore, UK) was used throughout all procedures.
Whenever possible biological material was snap frozen in liquid nitrogen and stored frozen at –80°C to avoid degradation.

All general molecular biology procedures including cloning, bacterial transformation and DNA electrophoresis were carried out using standard practices described in Sambrook *et al.*, 1989, or in Protocols and Applications Guide, (third Edition) 1996, Promega Corporation USA).

Aseptic techniques were used for all microbiological procedures.
Risk assessed protocols were used for all work involving hazardous material and Departmental radioactive procedures were followed for all work involving radioisotopes.

2.2 Microbiological Strains.

Glycerol stocks of DH5α, XL1blue, BL21 (DE3) and BL21 (DE3) *pLysS E.coli* host strains were obtained from Novagen Biosciences Inc. (Madison, WI.USA). The TOP10F' *E.coli* strain was supplied by Invitrogen BV. (Groningen, Netherlands).

Master stocks of all bacterial strains (host and recombinant) were stored as 15% glycerol stocks at –80°C. Short-term (1 month) working stock cultures were maintained on agar plates containing appropriate antibiotic at 4°C.

After inoculation from a stock culture liquid cultures (5ml, 50ml or 800ml) were grown with the appropriate antibiotic selection incubated at 37°C on an orbital shaker at 150 rpm.

2.3 Cloning Vectors.

Vector	Supplier	Use in this work
pBluescript^{+}	Stratagene	Library screening
pGem-T easy	Promega Ltd	Subcloning of Taq PCR products.
pET (3a, 11d, 17b, and 24a)	Novagen Inc.	Over-production of recombinant proteins.
TOPO TA	Invitrogen BV.	Subcloning of Taq PCR products.

2.4 Microbiological Growth Media

All microbiological growth media were prepared as described in Sambrook *et al.* (1989) and outlined below (**Table 2.1**). All solutions and glassware were sterilised by autoclaving at 121°C for 20 minutes. When heat liable constituents were used then solutions were filter sterilised through Minisart® 0.2 μm sterile membrane filters (Sartorius GmbH. Gottingen, Germany) into sterile containers.

Agar, bacto-tryptone and bacto-yeast extract, were all obtained from DIFCO Laboratories Ltd (West Molesey, Surrey).

2.5 Molecular Biology Stock Solutions and Buffer Solutions.

All molecular biology stock solutions outlined below (**Table 2.2**) were prepared as described in either Sambrook *et al.* (1989) or in Protocols and Applications Guide Third Edition (1996) (Promega Corporation USA).

Antibiotic solutions were added to bacterial growth media after the media had been autoclaved and immediately before it was used.

Table 2.1 Microbiological growth medium and stock solutions.

Media / Stock	**Recipe**
LB (Luria- Berani) Medium	10g bacto-tryptone, 5g bacto-yeast extract, 10g NaCl. Adjust to pH 7.0 with1N NaOH and made up to 1 litre with MQ water, Agar added to 1.5% if solid media was required.
M9 minimal medium	6g Na_2HPO_4, 3g NH_4Cl, 0.5g NaCl, made up to 1litre with MQ water. 1mM$MgSO_4$, 0.4% glucose and 100ml of 0.5% thiamine all added after autoclaving.
NZY Broth	5g bacto-yeast extract, 5g NaCl, 2g $MgSO_4 7H_2O$, 10g NZ amine (casein hydrolysate). Adjust to pH 7.0 with1N NaOH and made to 1 litre with MQ water.
SOB Medium	2% bacto-tryptone, 0.5% bacto-yeast extract, 10mM NaCl, 2.5mM KCl, 10mM $MgCl_2$, 10mM$MgSO_4$.
SOC Medium	SOB medium +20mM glucose.
Top agarose	NZY broth or LB medium with 0.7% agarose.
2 x YT Medium	16g bacto-tryptone, 10g bacto-yeast extract, 5g NaCl. Adjust to pH 7.0 with1N NaOH and made up to 1 litre with MQ water, Agar added to 1.5% if solid media was required.

Table 2.2 Molecular biology stock solutions.

Reagent	**Recipe**
Ampicillin	Stock solution at 50mg/ml in sterile MQ water stored as aliquots at –20°C. Used at a final concentration of 100µg /ml for bacterial selection.
Biotaq® PCR reaction buffer (x 10)	160mM $(NH_4)_2SO_4$, 670mM Tris:HCl pH8.8, 0.1% Tween 20.
Denaturing solution	400mM NaOH, 1M NaCl
50 x Denhardts reagent	1% ficoll, 1% polyvinylpyrrolidine (PVP), 1% bovine serum albumin (BSA).
6 x DNA gel loading buffer	0.25% bromophenol blue, 0.25% xylenecyanol FF, 15% ficoll
Kanamycin	Stock solution at 25mg/ml in sterile MQ water stored as aliquots at –20°C. Used at a final concentration of 25µg /ml for bacterial selection.
Hybridization solution	5 x SSC, 5 x Denhardts solution 0.5%SDS.
IPTG	100mM in MQ water stored as aliquots at –20°C
Neutralising solution	1.5M NaCl, 500mM Tris:HCl pH 7.4
Phosphate buffered saline (PBS)	150mMNaCl, 10mM Na_2HPO_4, 25mM KCl, 2mMKH_2PO_4
20 x SSC	3M NaCl, 300mM sodium citrate, final pH7.0
TbfI solution	30mM KOAc, 100mM RbCl, 10mM $CaCl_2$, 50mM $MnCl_2$ 15% glycerol (v/v), pH 5.8 adjusted with 0.2M acetic acid. Filter sterilised and stored dark at 4°C.
TbfII solution	10mM MOPS, 75mM $CaCl_2$, 10mM RbCl, 15% glycerol (v/v), pH 6.5 adjusted with 1.0M KOH. Filter sterilised and stored dark at 4°C.
TBS	130mMNaCl, 2mMKCl 25mM Tris, final pH7.4
X-gal	20mg/ml in dimethyl formamide.
50 x TAE (DNA electrophoresis buffer)	2M Tris base, 6% acetic acid 50mM EDTA ph8.0
TE	pH 7.4 -10mM Tris :HCl pH7.4, 1mM EDTA pH 7.6 -10mM Tris :HCl pH7.6, 1mM EDTA pH 8.0 -10mM Tris :HCl pH8.0, 1mM EDTA

2.6 Protein Electrophoresis Stock Solutions and Buffer Solutions.

SDS-PAGE and native gel protein electrophoresis was carried out using the mini ProteanII® electrophoresis apparatus from Bio-Rad using standard Laemelli methods (Laemelli 1970). The buffers and staining solutions used are described in **Table 2.3**.

Table 2.3 Protein electrophoresis stock solutions and buffer solutions.

Coomassie blue solution I	25% propan-2-ol, 10% acetic acid, 1% Coomassie blue R-250.
Coomassie blue solution II	10% propan-2-ol, 10% acetic acid, 0.125% Coomassie blue R-250.
Coomassie blue solution III	10% acetic acid, 0.125% Coomassie blue R-250.
Coomassie blue destain	10% acetic acid, 1% glycerol.
Gel running buffer	25 mM Tris:HCl pH 8.3, 250mM glycine, 0.1% SDS.
Laemelli sample loading buffer	50mM Tris:HCl pH6.8, 100mM DTT, 2% SDS, 0.1% bromophenol blue, 10%glycerol.
SDS-PAGE resolving gel (10%)	3.33 ml Acrylamide/Bis solution (37.5:1) 2.5 ml 1.5M Tris:HCl pH8.8, 100 µl of 10% SDS, 100 µl of 10% ammonium persulphate, 10µl of TEMED in a final volume of 10 ml with MQ water.
SDS-PAGE stacking gel (4%)	1.3ml Acrylamide/Bis solution (37.5:1) 2.5 ml 0.5M Tris:HCl pH6.8, 100 µl of 10% SDS, 100 µl of 10% ammonium persulphate, 10µl of TEMED in a final volume of 10 ml with MQ water.
Native gel running buffer	25 mM Tris:HCl pH 8.3, 250mM glycine,
Native gel sampling loading buffer + urea (0.25M)	50mM Tris:HCl pH6.8, 0.1% bromophenol blue, 10%glycerol, 0.312 ml 8M urea.
Native PAGE + urea (0.5M) resolving gel (18%)	6 ml Acrylamide/Bis solution (37.5:1) 2.5 ml 1.5M Tris:HCl pH9.0, 0.625 ml 8M urea, 100 µl of 10% ammonium persulphate, 10µl of TEMED in a final volume of 10 ml with MQ water.
Native PAGE +urea (0.5M) stacking gel (5%)	1.5ml Acrylamide/Bis solution (37.5:1) 2.5 ml 0.5M Tris:HCl pH6.8, 0.625 ml 8M urea, 100 µl of 10% ammonium persulphate, 10µl of TEMED in a final volume of 10 ml with MQ water.

2.7 Molecular Biology Techniques

2.7.1 Plasmid DNA Isolation.

Plasmid DNA was isolated and purified from transformed *E.coli* cells using either the Wizard® *Plus* SV mini-prep DNA purification system (Promega) or the Hybaid Recovery® plasmid min-prep isolation kit (Hybaid Ltd, Teddington, Middlesex, UK). DNA was isolated following the manufacturers published protocols.

Both of these protocols involved the alkaline lysis of bacterial cells followed by the removal of cell debris, chromosomal DNA and proteins, and the purification of plasmid DNA by binding and washing on resin beads.

Pure high quality super-coiled plasmid DNA was obtained from these procedures, which was used directly as required for DNA sequencing, bacterial transformation and site directed mutagenesis experiments.

2.7.2 DNA Agarose Electrophoresis.

DNA fragments and plasmid DNA were separated on mini (6 x 11cm) 0.7% agarose electrophoresis gels using the RunOne™ electrophoresis kit (EmbiTec USA). The gels were prepared by mixing the required amount of agarose and 1 x TAE buffer together and heating the solution gently in a microwave to fully dissolve the agarose. The gel solution was allowed to cool to approximately 60°C and ethidium bromide was added to a

final concentration of 0.5µg/ml. The gel solution was then poured into a gel mould and allowed to set with a well comb in place.

DNA samples were mixed with 6 x DNA gel loading buffer to give a final concentration of 1x buffer before loading into the sample wells. Electrophoresis was carried out in 1 x TAE running buffer at 100v until the bromophenol blue dye front had migrated to approximately two thirds of the gel length.

Following electrophoresis the DNA was visualised and a printed record of the gel image made using an ultra-violet transilluminator and gel documentation system (UV Products, Cambridge, UK).

Lambda (λ) *Hind* III and ϕ 174 (Hae III) DNA size markers were run in adjacent lanes alongside samples during electrophoresis and used to estimate the size and concentration of sample DNA.

2.7.3 Polymerase Chain Reaction (PCR) Methods.

DNA amplification was carried out using standard polymerase chain reaction (PCR) methods (Saiki *et al.*, 1988) on a Robocycler® (Stratagene) thermal cycling PCR machine in 200µl GeneAmp® (Perkin Elmer, Beaconsfield, Bucks. UK) thin walled reaction tubes.

Typically a PCR reaction contained 1x Biotaq®PCR reaction buffer, 1.5 mM $MgCl_2$,

200 μM dNTPs, 400 nM-1000 nM of each DNA primer, 1ng to 200ng template DNA and 1 to 5 units of BioTaq® DNA polymerase in a total reaction volume of 50 μl. Once prepared the reaction mixture was carefully overlaid with 50μl of mineral oil to prevent evaporation during the thermal cycling reactions.

The exact thermal cycling condition used were specific for each PCR experiment performed and are described in the relevant chapter, but they all contained the typical temperature cycling parameters described in **Table 2.4**.

Table 2.4 Typical PCR temperature cycling parameters.

1 cycle at 94°C for 5 minutes (denaturation).
25 - 30 cycles of denaturation (94°C for 45 seconds), annealing (50 – 68°C depending on the Tm of the primers for 60 seconds) and extension (74°C for 1.0 min per kilobase of expected PCR product).
1 cycle of extension at 74°C for 5.0 minutes.

PCR primers were designed to be 20-30 nucleotides long, to contain equal numbers of G+C and A+T combinations where possible and to end with a G or C anchor. Details of the specific primers used in each experiment are described in the relevant chapter.

At the end of a PCR reaction an aliquot of the reaction mixture was mixed with DNA sample loading buffer and analysed on an agarose electrophoresis gel.

When necessary PCR products and DNA fragments were purified from agarose gel slices using the GFX® PCR DNA and Gel Band Purification Kit (Amersham Pharmacia Biotech) following the manufacturers protocol.

2.7.4 Restriction Enzyme Digestion of DNA.

Restriction enzyme digestion of DNA was typically carried out by incubating 0.5 –5 μg of plasmid DNA with 1 – 5 units (U) of restriction enzyme at 37°C in the appropriate reaction buffer in a final volume of 10-40μl. When double digestion was required the reaction buffer which gave the maximum enzymatic activity for both enzymes was used. Where necessary reaction temperatures, reaction times and amounts of enzyme (U) were adjusted to give efficient DNA digestion.

2.7.5 Synthesis of Radioactive DNA Probes for cDNA Library Screening.

Radioactive DNA probes for cDNA library screening were synthesised using the *Redi*Prime® random primer DNA labelling system from Amersham Life Science (Little Chalfont UK). The labelling system uses klenow DNA polymerase I, random primers and

the incorporation of [α $-^{32}$P] dCTP to prime DNA synthesis at random sites on denatured template DNA.

In a typical labelling experiment 2.5 - 25 ng of template DNA in 45 µl of sterile MQ water was denatured by heating in a boiling water bath at 95°C for 5 minutes. Denatured DNA was then added to an eppendorf tube of *redi*prime® reaction mixture that contained klenow buffer, dATP, dGTP, dTTP nucleotides, random nonamer primers and klenow DNA polymerase I. The mixture was vortexed briefly and centrifuged to bring the contents to the bottom of the tube before 5 µl (50µCi) of [α $-^{32}$P] dCTP was added and the reaction incubated at 37°C for 10-30 minutes. Following the incubation radio-labelled DNA was separated from unincorporated nucleotides by passage through a Bio-Spin® P-6 (Bio-Rad) desalting column. Purified ^{32}P radio-labelled DNA probes were used immediately in cDNA library screening experiments and then stored for re-use at –20°C for 1-2 weeks.

2.7.6 cDNA Library Screening.

cDNA library screening by cross-hybridisation of library phage plaques (pfus)to radioactive DNA probes was carried out according to standard molecular biology protocols Sambrook *et al.*, (1989) and described in detail in **Chapter 3 section 3.3**.

Cross-reacting pfus were purified through three rounds of library screening before cDNA clones were plasmid rescued from the library and DNA sequenced.

2.7.7 Preparation of Competent *E.coli* Cells.

Transformation competent *E.coli* cells of XL1 blue, BL21 (DE3) and BL21 (DE3) *pLysS* were all prepared as laboratory stocks using a modified rubidium chloride method (Stratagene). A culture of the required bacterial strain was grown overnight in LB + antibiotic at 37°C on an orbital shaker (150rpm). A 1.0 ml inoculum from this overnight culture was transferred to 100ml of fresh LB medium + antibiotic and grown at 37°C on an orbital shaker (150 rpm) until an OD550 of 0.5 was reached. The culture was chilled on ice for 5.0 minutes, transferred to 50 ml Falcon® tubes and the cells gently pelleted from the medium by centrifugation at 4000g at 4°C for 5 minutes in an Avanti® bench top centrifuge (Beckman). The supernatant was removed from the cells and the cells were re-suspended in 40ml ice-cold TfbI solution and incubated on ice for 5.0 minutes. The cells were re-pelleted as above, the supernatant removed, and the cells re-suspended in 4.0 ml of ice-cold TbfII solution and incubated on ice for 15 minutes. The cell suspension was then dispensed into 100 µl aliquots, snap frozen in liquid nitrogen and stored at –80°C until required.

2.7.8 Transformation of Competent *E.coli* Cells with Plasmid DNA.

An aliquot of competent *E.coli* cells was removed from the –80°C freezer and thawed on ice for 5.0 minutes. For each transformation a 50 µl aliquot of thawed cells was transferred to a chilled Falcon® 2059 polypropylene tube and 1-2 µl (approximately 10 ng) of DNA was added and gently mixed. The cells were then incubated on ice for 30 minutes. The reaction was removed from ice and given a 45 seconds heat shock at 42°C by placing the tubes in a pre-heated water bath. The tubes were then transferred back to ice for 2 minutes before 0.9 ml of pre-heated (37°C) LB medium was added and the cultures incubated at 37°C on an orbital shaker (150 rpm) for one hour. Following this incubation a 100µl aliquot of cells and the remaining cells pelleted following centrifugation in a microfuge, were plated on separate LB agar plates containing the appropriate antibiotic. The plates were incubated overnight at 37°C and the transformation efficiency assessed by colony counting following this incubation. Single transformants were selected and grown as liquid cultures for plasmid isolation, DNA sequencing and protein over-expression studies.

2.7.9 DNA sequencing.

Automated DNA sequencing was carried out on an Applied Biosystems 377 sequencer using dye-dideoxy terminator chemistries. The DNA Sequencing Service of the Department of Biological Sciences, University of Durham performed all DNA sequencing.

2.7.10 Site Directed Mutagenesis.

Site directed mutagenesis (SDM) of plasmid DNA was carried out using the QuickChange® Site Directed Mutagenesis kit from Stratagene.

Detailed reaction conditions, and a description of the mutagenesis primers used in individual experiments are described in the relevant sections of **Chapters 4 + 5.**

Reactions were carried out in a total volume of 50 µl in 200µl GeneAmp® thin walled reaction tubes (Perkin Elmer, Beaconsfield, Buckinghamshire, UK) using a Robocycler® thermal cycling PCR machine (Stratagene).

In the reaction *Pfu*Turbo® DNA polymerase II, a high fidelity proof reading polymerase capable of synthesising large DNA products was used, together with complimentary mutant oligonucleotide primers and temperature cycling (**Table2.5**) to synthesise mutant DNA from double stranded pET plasmid DNA template.

Following temperature cycling a 5 µl aliquot of the mutagenesis reaction mixture was mixed with DNA sample loading buffer and analysed on an agarose electrophoresis gel.

Table 2.5 Typical mutagenesis reaction temperature cycling parameters.

1 cycle at 95°C for 30 seconds.
12-18 cycles of 95°C for 30 seconds 55°C for 60 seconds 68°C for 60 seconds / kb of total plasmid length.

Once the mutant DNA was synthesised the methylated parental DNA template was removed from the non-methylated mutant DNA reaction product by incubating the reaction mixture at 37°C with 1µl (10U) of the restriction enzyme D*pn*I, an enzyme specific for methylated DNA.

A 1-4 µl aliquot of the resulting D*pn*I treated mutant DNA was then used to transform super-competent XL1-blue® *E.coli* cells (Stratagene). Transformed cells were plated on LB agar plates containing the appropriate antibiotic and grown overnight at 37°C. Typically between 50 and 200 colonies were obtained following the overnight incubation.

Two single transformants were picked from the plate and re-grown in 5.0 ml LB containing antibiotic overnight at 37°C. Plasmid DNA was prepared from these overnight cultures and DNA sequencing of the purified plasmid DNA was used to confirm the presence of the correct amino acid mutation.

During all mutagenesis experiments a control reaction was included to monitor the efficiency of mutagenesis and to ensure that the procedure was working. In this control reaction *p*Whitescript® a 4.5 kb mutant plasmid containing a stop codon in the β-galactosidase gene of pBluescript®IISK was used together with mutagenesis primers to revert this stop codon back to its correct glutamine residue. Following mutagenesis and transformation of super-competent *E.coli* XL1-blue® cells the bacteria were plated on LB agar indicator plates containing ampicillin (100μg /ml), IPTG (0.2mM) and x-gal (40μg/ml) and incubated overnight at 37°C. Transformants containing the *p*Whitescript® plasmid mutated back to contain the correct glutamate residue and therefore containing an active β-galactosidase protein appeared blue on these indicator plates. Routinely greater than 80% of these control transformants were blue indicating a mutagenesis efficiency of greater than 80%.

2.7.11 Over-expression of Recombinant Proteins in *E.coli.*

Recombinant ACP, HAS and G3PAT proteins were over-expressed in *E.coli* using the pET Protein Expression System (Novagen) and a full description of the individual vectors used and the methods of cloning and over-expression are described in the relevant sections of **chapter 4 and chapter 5**. The system utilises derivatives of the original pET vectors constructed by Studier and colleagues (Studier and Moffat 1986, and Studier *et al.*, 1990). Target genes are cloned into the pET vectors under the control of the bacteriophage T7 transcription signals so that when the recombinant vector is transformed into a bacterial host strain containing T7 RNA polymerase under *lac* UV5 control (BL21 DE3 derivatives) and IPTG is added to the growth media, target gene expression is induced.

To enable the sub-cloning of recombinant DNA into a pET expression vector, compatible cloning sites were synthesised via PCR at the 5' and 3' ends of insert DNA. This allowed the DNA to be cloned in frame into a restriction site within the multiple cloning site (MCS) of the appropriate pET vector. This recombinant vector was then transformed into super-competent XL1-blue® cells, plasmid DNA was purified from the transformants and the DNA sequenced. Once authenticated by DNA sequencing this pET plasmid DNA was then re-transformed into competent BL21(DE3) cells or into competent BL21 (DE3) *pLysS* cells in preparation for protein over-expression.

For a typical induction experiment a single colony was transferred from an agar plate of transformants into 5.0 ml of LB medium + antibiotic and grown overnight at 37°C on an orbital shaker (150 rpm). From this overnight starter culture 2.0 ml was inoculated into a 2 litre conical flask containing 800 ml of LB medium + antibiotic and grown at 37°C on an orbital shaker (150 rpm) for 3-4 hours until an OD@ 650 nM of 0.6 was reached. At this point a 1.0 ml sample was removed pelleted in a microfuge and snap frozen for later SDS-PAGE analyses as a pre-induced control. Once a culture OD of 0.6 was achieved 4.0 ml of 100mM IPTG was added, to give a final concentration of 0.5mM and the culture was further grown for 3 hours to allow the production of recombinant protein. A post induction 1.0 ml sample was removed at this stage, pelleted and snap frozen for later SDS-PAGE analyses.

The cells were harvested by centrifugation at 4000g in a J-Lite® rotor at 4°C in a J2HS centrifuge (Beckman). The supernatant was removed and the cell pellet was snap frozen in liquid nitrogen and stored at –80°C.

2.8 Protein Methods.

2.8.1 Freeze Thaw Extraction of Recombinant Proteins.

Over-expressed ACP, HAS, and G3PAT proteins were selectively extracted from *E.coli* cells using repeated cycles of freezing and thawing (Johnson and Hecht 1994).

Following IPTG induction of recombinant mRNA and protein overexpression, *E.coli* cells were pelleted from the growth medium by centrifugation at 5000g for 10 minutes in 500ml bottles using a J-lite rotor in a Beckman J2-HS centrifuge. The medium was decanted and the pellets were transferred to 50 ml tubes, snap frozen in liquid nitrogen and stored at –80°C until required.

Pellets were removed from –80°C storage and allowed to thaw slowly on ice before being incubated in a dry ice / ethanol bath for 8 minutes. Following this incubation, the pellets were placed in an ice water bath for 8 minutes. This freeze thaw cycle was repeated twice more and then the pellets were placed on ice. Approximately 5 ml of ice cold extraction buffer (10mM potassium phosphate pH6.2 for ACP, or 25 mM Tris:HCl pH 7.5 for G3PAT) was added to each pellet and the pellets were gently re-suspended. The extracts were pooled together in a 50 ml centrifuge tube and incubated on ice for 60 minutes with periodic gently agitation. Following centrifugation at 40,000g in a Beckman Avanti® centrifuge the supernatant containing the extracted recombinant protein was carefully decanted away from the cell pellet, aliquoted and snap frozen in liquid nitrogen and stored at –80°C.

A 5 to 10 µl aliquot of the supernatant containing the recombinant protein was mixed together with Laemelli sample loading buffer and analysed by SDS-PAGE.

2.8.2 SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE).

SDS-PAGE analysis of protein extracts was carried out using mini Protean II® electrophoresis apparatus (BioRad) according to the methods described by Laemelli (Laemelli, 1970). Gels were polymerised between glass plates separated with 0.75mm spacers using a 37.5:1 acrylamide / bis acrylamide solution (BioRad) to give a 10% resolving gel overlaid with a 3% stacking gel. Samples were boiled in Laemelli sample loading buffer for 3 minutes and microfuged prior to loading into sample wells. Electrophoresis was carried out in gel running buffer at 100 volts through the stacking gel and 200 volts through the resolving gel until the bromophenol blue dye front just reached the end of the gel. In order to estimate protein size and concentration a lane containing 5µl of SDS VII (1µg of protein / band) molecular weight markers (Sigma) was included in all electrophoresis gels.

Following electrophoresis the gels were removed from the gel plates and immediately stained with coomassie blue to visualise the protein bands.

2.8.3 Conformational Gel Analysis of ACP and Acyl-ACP Derivatives (Native - PAGE).

Purified ACP and acyl ACP derivatives, which were synthesised either enzymatically using holo-ACP synthetase (HAS) or chemically, using n-acylimidazoles, were analysed

on native PAGE gels containing urea, using the methods described by Post-Beittenmiler *et al.*, (1991). Apo, holo and butyryl ACP samples were separated on 18% PAGE + 0.5M urea resolving gels, 5% PAGE + 0.5M urea stacking gels using the mini Protean II® kit (BioRad). Gels were run using native gel running buffer. The gels, sampling loading buffer, and the gel running buffer were all prepared without SDS and DTT. The ACP samples were mixed with a modified loading buffer containing 0.25M urea and were loaded directly into the sample wells without prior boiling. Electrophoresis was carried out at 150 volts until the bromophenol dye front had reached the end of the gel. Following electrophoresis gels were stained with coomassie blue to enable the visualisation of the protein bands.

2.8.4 Coomassie Blue Protein Staining Procedures.

On completion of protein electrophoresis a modified coomassie blue staining method was used for the visualisation of protein bands.

Gels were removed from the electrophoresis apparatus and immediately transferred to a container of coomassie I staining solution that had been previously heated to 60°C in a microwave oven, and incubated with gentle agitation for 60 minutes. The coomassie I solution was poured off the gels and replaced with hot (60°C) coomassie II solution and again incubated with agitation for 60 minutes. The coomassie II solution was then

replaced with hot coomassie III solution for a final incubation of 60 minutes. Once staining was complete the gels were incubated in coomassie blue de-stain solution until the gel background had cleared leaving blue stained protein bands visible in the gel.

2.8.5 Purification of Wild Type *E.coli* ACP.

Wild type ACP was purified from cell paste of fermenter grown (UCL London) *E.coli* K12 strain using a modification of the acid precipitation method originally described by Majerus *et al.*, (Majerus *et al.*, 1964). Approximately 1.0 kg of cell paste was removed from –80°C storage and slowly defrosted at 4°C overnight. Once defrosted 4 x volumes of ice cold MQ water was added and the mixture stirred gently to give an even suspension. The cell suspension was passed once through a cell disrupter (Constant Systems Ltd. UK) at 25 psi to disrupt the cells, and the broken cell suspension was collected in 2 litre conical flasks on ice. The required volume of 1M Tris:HCl (pH 8.0) was added to the suspension to give a final concentration of 20mM. βmercaptoethanol (0.1%), MnCl (3mM) and Dnase I (50mg / litre) were also added and the suspension gently stirred on ice for 30 minutes. This removed long strands of DNA and resulted in a non-viscous suspension.

The broken cell suspension was transferred to 250 ml centrifuge bottles and centrifuged (28000g) for 30 minutes at 4°C in a JA14 rotor using a J2HS centrifuge (Beckman). The supernatant was carefully removed from the loosely pelleted cell debris, the volume accurately measured and transferred to a glass beaker on ice. A stirrer bar was added and the solution stirred continuously. A 60% ammonium sulphate cut (Dawson *et al* 1994) was performed on the supernatant by adding 361g/litre of (60% saturation) finely ground ammonium sulphate, and stirring on ice for 60 minutes. The solution was transferred to 250 ml centrifuge bottles and centrifuged (28000g) for 30 minutes at 4°C in a JA14 rotor using a J2HS centrifuge. The supernatant containing the ACP was carefully decanted from the pellets into a cold (4°C) 2 litre conical flask. The pH was adjusted to pH 1.0 using concentrated HCl and the acidified solution was left at 4°C for four hours to allow the ACP to precipitate. The bulk of the supernatant was siphoned off the pellets and the remainder was removed following centrifugation at (28000g) for 30 minutes at 4°C in a JA14 rotor using a J2HS centrifuge. The pellets, containing the ACP, were carefully re-suspended in a small volume of 1.0M Tris:HCl pH 8.0 and dialysed overnight at 4°C in Spectrapore® (6000-8000 kDa cut off) dialyses membrane against 3 x 2litre changes of 10mM potassium phosphate buffer pH 6.2, 0.1% β-mercaptoethanol. After dialyses the supernatant was centrifuged at 20000g using an Avanti® centrifuge, poured off the resulting pellets, and snap frozen in liquid nitrogen and stored at –80°C.

ACP was purified on a Highload MonoQ® anion exchange column (Amersham Pharmacia Biotech.) using a Highload® chromatography system. The supernatant was defrosted from –80°C, membrane filtered through a 0.2 μm filter and loaded onto a pre-equilibrated (pH 6.2 in 10mM potassium phosphate buffer, 0.1% β–mercaptoethanol) column. Once loaded the column was washed with 10 column volumes of buffer and the bound ACP was eluted with a 500ml linear gradient of 0 – 500mM lithium chloride (LiCl) in equilibration buffer. Fractions were collected during the elution and analysed by SDS-PAGE. The fractions containing ACP were assessed for purity, pooled together and fully reduced by incubation with DTT (5mM final concentration) on ice for 30 minutes. The fully reduced ACP was dialysed using Spectrapore® (6000-8000 kDa cut off) membrane against MQ water adjusted to pH7.0 by the addition of few grains of ammonium bicarbonate. Following dialyses the ACP was aliquoted into 5.0 mg lots, lyophilised, snap frozen in liquid nitrogen and stored at –80°C. A typical purification would yield approximately 150 mgs of > 95% pure ACP from 1.0 kg of cell paste. This was shown to be <90% holo ACP by electrospray mass spectrometry analysis (ESIMS).

2.8.6 Purification of Recombinant ACPs.

Following over-expression and freeze thaw extraction recombinant ACPs were purified using Porous HQ® anion exchange chromatography on a Biologic Chromatography System® (BioRad UK) running in a cold cabinet at 4°C.

Porous HQ® media is a fast flow medium allowing flow rates of up to 10 ml / minute and very rapid (< 5.0 minute) chromatography runs. This was purchased in bulk from PE Biosystems (Framingham USA) and packed into 1.0 ml and 10 ml stainless steel columns by Jones Chromatography (UK).

An aliquot of freeze thaw extract containing the recombinant ACP was removed from the –80°C and defrosted on ice. Once defrosted it was transferred to 50ml centrifuge tubes and centrifuged at 40000g for 30 minutes at 4°C in an Avanti® centrifuge. The clarified supernatant was then passed through a 0.2 μm membrane filter before being loaded directly onto a pre-equilibrated (pH 6.2 in 10mM potassium phosphate buffer, 0.1% β–mercaptoethanol) 1.0 ml Porous HQ® column. The column was washed with ten column volumes of buffer and the bound ACP was eluted using a 20ml linear gradient of 0-500 mM LiCl in equilibration buffer. 1.0ml fractions were collected throughout the gradient, several runs were carried out in a single day and sequential runs were collected in the same fraction collector tubes.

Fractions were analysed by SDS-PAGE and those containing ACP were pooled together dialysed using Spectrapore®(6000-8000 kDa cut off) membrane against MQ water adjusted to pH7.0 by the addition of few grains of ammonium bicarbonate. Following dialyses the ACP was aliquoted into 5.0mg lots, lyophilised, snap frozen in liquid nitrogen and stored at –80°C. A typical purification would yield between 10-20mg of >95% pure ACP from a bacterial culture (800ml) of *E.coli* containing recombinant plasmid. Purified recombinant ACP was checked by native-PAGE conformational gels and MALDItof mass spectrometry and found to be greater than 90% apo protein, indicating that post-translational modification with 4'phosphopantatheine does not occur during protein over-expression under the conditions used. This turned out to be advantageous as it allowed the enzymatic synthesis of acyl-ACPs using holo-ACP synthetase (HAS), a more specific acylation reaction than that obtained with chemical acylation methods.

2.8.7 Purification of Over-expressed Holo -ACP Synthetase.

Over-expressed holo – ACP synthetase (HAS) was purified from freeze thaw extracts of *E.coli* using a PorousHS® cation exchange column on a Biologic liquid chromatography system running in a cold cabinet at 4°C. The column was a self-packed 1.0ml PEEK column packed under pressure at 8.0 ml/min on a FPLC® system (Amersham Pharmacia

Biotech.) A freeze thaw extract containing recombinant HAS was defrosted from –80°C on ice, and then transferred to 50 ml centrifuge tubes and centrifuged at 40000g for 30 minutes at 4°C in an Avanti® centrifuge. The supernatant was filtered through a 0.2μm membrane filter before it was loaded directly onto a pre-equilibrated (50mM Tris:HCl pH8.0, 10mMMgCl_2, 1mMDTT) column. The column was washed in 10 column volumes of buffer and the bound HAS was eluted in a 15 ml linear gradient of 0 – 1.0 M NaCl in equilibration buffer. Fractions (1.0 ml) were collected throughout the gradient and analysed on a 15% SDS-PAGE gel. The fractions containing HAS were assessed for purity, pooled together, aliquoted into 200μl lots, snap frozen in liquid nitrogen and stored at –80°C.

2.8.8 Chemical Synthesis of Acyl-ACPs.

Purified wild type *E.coli* ACP was used for the chemical synthesis of acyl-ACP in acylation reactions with C4:0 and C8:0 n-acylimidazoles (Cronan and Klages 1981). For each reaction a freeze-dried aliquot containing 5.0 mgs of purified and fully reduced ACP was removed from the –80°C freezer, dissolved in 2.0 ml of cold MQ water and transferred to a 5.0 ml reacti-vial® (Pierce, Rockford USA) containing a magnetic stirrer. The reacti-vials were placed in an ice bath and the solution stirred vigorously while 20μl of 500mM imidazole buffer pH6.5 and 10μl of 50mM EDTA was added. The vigorous

stirring was continued and 200µl of a 500mM C4:0 or C8:0 acyl-imidizole (made up in HPLC grade isopropanol) was added drop-wise to the buffered ACP solution. Following this addition the reactions were incubated at room temperature for 60 minutes with continuous stirring. At the end of the incubation period the reaction mixture was diluted to a final volume of 2.5 ml with MQ water and desalted through a PD10 desalting column into 3.5 ml of MQ water. The desalted acyl-ACP product of each reaction was confirmed and quantified by native-PAGE and electrospray mass spectrometry (ESMS).

2.8.9 Enzymatic Synthesis of Acyl ACPs.

Purified recombinant *E.coli* and *Brassica napus* apo ACPs were used for the enzymatic synthesis of holo and acyl-ACPs in reactions catalysed by recombinant holo-ACP synthetase (HAS). This reaction is a modification of the reactions described by Carreras *et al.*, 1997.

The details and results of specific experiments are presented in **Chapter 4.**

A typical synthesis was carried out in a 2.0 ml reaction containing 350µmole apo ACP, 700µmole butryl CoA, and 3.5µmole HAS at pH7.3 in 50mM potassium phosphate buffer. The reaction was stirred in a reacti-vial® at room temperature for 60 minutes. Following synthesis the reaction mixture was diluted to 3.5 ml with MQ water, desalted

through a PD10 desalting column into MQ water, freeze-dried, snap frozen in liquid nitrogen and stored at –80°C.

A small aliquot of the desalted reaction products were removed and confirmed by confirmational gel analyses and by accurate mass measurement on a Matrix Assisted Laser Desorption time of flight mass spectrometer (MALDItof-MS).

2.8.10 Isolation of Recombinant Squash Glycerol-3-Phosphate-1- Acyltransferase (G3PAT) from *E.coli.*

Squash G3PAT was overproduced in *E.coli* in several different forms, pNA4, Q24a, Q17b and numerous site directed mutations. These are all described in detail in **Chapter 5**. Following their over-production in *E.coli* they were isolated using one of two methods depending on their intended subsequent use.

For activity and substrate selectivity assays they were prepared as crude cell free extracts using freeze thaw extraction and quantified against SDS VII standards on coomassie stained SDS-PAGE gels.

For crystallography trials and structural studies a rigorous purification protocol involving, ion exchange, hydrophobic interaction and gel filtration chromatography was followed which yielded 8-13 mg of 90% pure G3PAT from 1500ml of bacterial culture. This

purification was carried out by Sveta Sedelnikova, one of our collaborators in the crystallography group of Professor Rice at the University of Sheffield (Turnbull *et al.*, 2001a).

2.8.11 Preparation of Radio-labelled 16:0 and 18:1 Acyl ACP Substrates.

Radio-labelled C16:0 (^{14}C) and C18:1 (^{3}H) acyl-ACP substrates for use in G3PAT activity assays were prepared using purified, recombinant and enzymatically (HAS) pantethenylated *E.coli* holo ACP, radio-labelled free fatty acids and recombinant *E.coli* acyl-ACP synthetase.

The acyl-ACP synthetase was over-produced and partially purified using a construct kindly provided by Dr J.Shanklin at the Brookhaven National Laboratory USA.

Following synthesis the radio-labelled acyl-ACPs were purified from the excess fatty acids and ACP present in the reaction mixture using anion exchange and octyl-sepharose chromatography (Rock and Garwin 1979), snap frozen in liquid nitrogen and stored at –80°C.

Dr Ted Scheirer and Matthew Hayman (PhD student) carried out the synthesis of these substrates whilst working in this laboratory.

2.8.12 G3PAT Activity Assay.

The biological activity and the substrate selectivity of recombinant and site directed mutations of squash G3PAT were measured using modifications to the standard methods previously reported in the literature (Betram and Heinz 1981, Frentzen *et al.*, 1983). Because of their availability and the difficulty of synthesis of the physiological substrates acyl-ACP, many of these previous studies have used acyl-CoAs as substrates instead of acyl-ACPs. Various concentrations of BSA and various pH conditions have also been used in earlier assays.

The conditions used for assay in this work have been optimised to reflect the physiological conditions, (glycerol 3-phosphate, acyl-ACP concentrations and pH) expected in the plant chloroplast. This modified assay is the result of extensive studies by Drs T Scheirer, Dr Johan Kroon and Matthew Hayman in this laboratory (Hayman *et al.*, 2000 and Slabas *et al.*, in press).

A standard assay involves the G3PAT catalysed incorporation of radioactive acyl-ACP into the *sn-1* position of glycerol-3-phosphate to yield radioactive lysophosphatidic acid (LPA) which is then extracted into organic solvent and counted in a liquid scintillation counter.

The standard reaction mixture contained 250 mM HEPES pH8.0, 5 mg/ml BSA, 300μM glycerol-3-phosphate and 1 μM each of ^{14}C 16:0 and ^{3}H 18:1 acyl ACP in a final volume of 180μl. An aliquot of diluted cell free extract of recombinant G3PAT was added to initiate the reaction and it was incubated at 25°C for 2 minutes. Following incubation the reaction was stopped by the addition of 710μl chlorofom : methanol (1:1) and 280μl acid salt solution (1M KCl in 200mM H_3PO_4). The mixture was then vortexed and microfuged at 13000 rpm and the radioactive LPA reaction product extracted from the lower organic phase and counted in a Packard TR liquid scintillation counter. These assays were carried out together with Matthew Hayman a fellow PhD student in our laboratory.

2.9 Crystallisation Techniques.

2.9.1 Crystallisation Trials.

Crystallography studies on squash G3PAT and *E.coli* and *Brassica napus* ACP were carried out in collaboration with Professor David Rice, Dr John Rafferty and Dr Andy Turnbull at the Krebs Institute, University of Sheffield.

For squash G3PAT crystals were grown at 17°C using the hanging drop vapour diffusion method with 14-25% PEG 4000 as the precipitant in 100mM citrate buffer pH5.6 containing 100mM ammonium acetate and 10% isopropanol. For crystal structure

determination several crystal forms were grown including native protein, cysteine to serine mutations (four single mutants), grown in the presence of 1mM ethyl mercury phosphate (EMP) and a form where seleno-methionine had been incorporated in place of methionine. A detailed description of this experimentation is described in **Chapter 5.**

For ACP the highly acidic nature and the small size of the protein resulted in great difficulty in obtaining a crystal form from any source which was stable in the X-ray beam, and which would diffract and give interpretable structural data. Many forms were tried, including wild type and recombinant *E.coli* ACP, and recombinant ACPs from *Brassica napus*, strawberry and *Mycobacterium*. Several chain length acyl-ACPs were also tried including C4 and C8 *E.coli* and *Brassica* wild type and recombinant proteins. These acyl-ACPs were synthesised both chemically and enzymatically.

A stable crystal form was obtained from both a wild type and a recombinant C4 acyl-ACP from *E.coli* using the hanging drop vapour diffusion method with 8-12% PEG20000 as the precipitant in 40mM sodium cocadylate buffer pH 6.0 containing 30mM $ZnCl_2$. Site directed cysteine and methionine mutations of the recombinant form were produced and used in trials to attempt to produce heavy metal crystal derivatives with mercury or seleno-methionine incorporated respectively. The details and results of this experimentation are reported in **Chapter 4.**

2.9.2 Growth and Selenomethionine Incorporation Into Recombinant ACP and G3PAT to Aid in Crystal Structure Determination.

The incorporation of selenomethionine into recombinant proteins is a valuable tool for protein crystallographers as a method to combine a heavy metal into a protein crystal, which is essential for the isomorphous replacement methods used to determine crystal structure. Other methods involve soaking crystals in heavy metal solutions such as mercury which bind to the side chains of cysteine residues. The intensity differences between the native X-ray data set and the heavy metal X-ray data set allows the location of the heavy metal atoms to be identified and this provides a starting point for the protein phase angles to be determined and the crystal structure to be resolved.

Selenomethionine incorporation into recombinant ACP and G3PAT was achieved using two different methods and the detailed description of these is described in **Chapter 4 and Chapter 5.**

For squash G3PAT and *E.coli* and rape ACP recombinant plasmid was transformed into a BL21 (DE3) metC$^-$ *Ecoli* strain developed by Dr A.R.Stuitje (Free University of Amsterdam) and grown in a minimal medium (M9 + amino acids except methionine) supplemented with seleno-methionine instead of methionine. Growth of the bacteria under these conditions and induction with IPTG ensured that selenomethionine was incorporated into the recombinant protein.

This method did not work well with site directed methionine mutations of *E.coli* ACP (**Chapter 4**) where the cultures did not grow well under these conditions and no over-expression of ACP was visible. An alternative protocol for seleno-methionine incorporation by metabolic inhibition of the methionine pathway in a normal BL21 (DE3) strain was used (Van Duyne *et al.*, 1993). In this method methionine was replaced in the growth medium by seleno-methionine which was added together with excess amounts of other amino acids known to inhibit methionine biosynthesis at the point of IPTG induction of the recombinant ACP.

Recombinant plasmids containing *E.coli* ACP with introduced methionine mutations were transformed as normal into BL21 (DE3) cells, and 5.0 ml LB + antibiotic cultures were grown overnight at 37°C from single transformants. The cells from 4.0 ml of these overnight cultures were gently pelleted by centrifugation at 1300 g in a bench-top microfuge, resuspended in 1.0 ml of M9 medium and used to inoculate a 2 litre conical flask containing 800 ml of M9 + 0.4% glucose medium without amino acid supplements. These 800 ml cultures were grown at 37°C on an orbital shaker (150 rpm) until an OD of 0.6 @ 600nm was achieved. Typically this took 9 hours. At this point the following final concentrations of amino acids and seleno-methionine were added to the cultures as solid powders: lysine (100mg/l), phenylalanine (100mg/l), threonine (100mg/l), isoleucine (50mg/l), leucine (50mg/l), valine (50mg/l) and seleno-methionine (50mg/l). The cultures

were re-incubated at 37°C for 15 minutes to allow the inhibition of methionine biosynthesis to start. At this point 0.4 mM IPTG was added to induce the production of recombinant ACP and the cultures were re-incubated at 37°C on an orbital shaker (150 rpm) for 6 hours. A post induction 1.0 ml sample was removed at this stage, pelleted and snap frozen for later SDS-PAGE analyses.

The cells were harvested by centrifugation at 4000g in a J-Lite® rotor at 4°C in a J2HS centrifuge (Beckman). The supernatant was removed and the cell pellet was snap frozen in liquid nitrogen and stored at –80°C.

When required the seleno-methionine modified ACP was purified and enzymatically derivatised using the same methods described for recombinant un-modified ACP.

2.10 Mass Spectrometry Methods.

2.10.1 Electrospray Mass Spectrometry Analysis of C4 and C8 ACP.

Prior to mass spectrometry facilities becoming available at Durham (MALDItof-MS), ACP and chemically synthesised acyl-ACP derivatives where characterised by their mass measurement using electrospray mass spectrometry (ESMS) facilities at the University of Cambridge Department of Chemistry (Phillip Lowden) or at Zeneca Biotechnology Section (Steve Rayner).

Following purification and derivatisation samples were desalted into MQ water using a PD10 desalting column (Amersham Pharmacia Biotech) to give a final concentration of 1.0 mg/ml (100pmole/μl). This was diluted with 90% acetonitrile /1.0% formic acid and 5μl (25 pmole) of this solution was injected onto the mass spectrometer. A 5μl aliquot (50pmole) of horse myoglobin (Mr 16952) was used to calibrate the instrument.

2.10.2 Matrix Assisted Laser Desorption Time of Flight Mass Spectrometry (MALDItof-MS) Analysis of Acyl-ACP Derivatives and Site Directed Mutants of ACP.

Following purification and derivatisation, apo, holo, acyl-ACP and site directed mutants of ACP were all confirmed by accurate mass measurement using Matrix Assisted Laser Desorption Time of Flight Mass Spectrometry (MALDItof MS) on a PE Biosystems Voyager DE-STR® Biospectrometry Workstation.

For analysis the ACP was diluted in MQ water to give a final concentration of 1.0 μM and 0.5 μl of this solution was spotted onto a sample well of a MALDI target plate. This was overlaid with 0.5 μl of 3,5-dimethoxy-4-hydroxycinnamic acid (sinapinic acid) (10mg/ml in 50% acetonitrile + 0.3% TFA) matrix solution and allowed to air dry. The target plate was placed into the mass spec and ionisation of the ACP was achieved using a laser intensity of 3000 volts. Mass measurements were made with the mass

spectrometer in linear positive ion mode, with delayed extraction switched on and with an accelerating voltage of 25 kvolts. Data were acquired between 4000 and 20000 Da and data from 150 laser shots were accumulated for each spectrum acquired.

For each ACP sample analysed the mass spectrometer was calibrated for mass accuracy using a 0.5µl aliquot of a sequazyme® 3 calibration mixture (**Table 2.6**) (ABI Framingham) spotted together with 0.5µl of matrix solution on the MALDI target plate adjacent to the individual ACP sample.

Table 2.6 Constituents of Sequazyme 3™ MALDItof calibration mixture

Protein	Charge (m)	(M+mH)n+ average	Concentration when mixed with matrix and spotted on MALDI target
Insulin (bovine)	+1 +2	5734.59 2867.48	50 fmole/µl
Thioredoxin *(E.coli)*	+1 +2	11674.48 5837.74	275 fmole/µl
Apomyoglobin	+1 +2	16952.56 8476.78	400 fmole/µl

2.11 Computer Analysis.

2.11.1 DNA Sequence Analysis.

Sequence data files from DNA sequencing reactions were checked for accuracy, assembled and analysed using the computer software package DNA Strider (Marck 1988)

on an iMAC personal computer. This program was also used to check the suitability of oligo-nucleotide primers for Quick Change® mutagenesis reactions.

2.11.2 Protein Multiple Sequence Alignments.

The sequence alignment of MCAT (plant and bacterial), ACP (plant and bacterial) and G3PAT (plant) amino acid sequences was carried out using either the Clustal W program (Higgins and Sharp 1989) on an iMAC personal computer or the Multialign (Corpet 1988) program on the World Wide Web (WWW). This program is available as part of the Network Protein Sequence Analysis Server (NPS@ http://npsa-pbil.ibcp.fr/.) (Combet *et al.*, 2000). In all cases the protein amino acid sequences were gathered together in a FASTA (Pearson and Lipman 1988) formatted text file before being submitted to the program for alignment.

Chapter 3.

Malonyl-CoA ACP:Transacylase (MCAT).

Identification of a Putative EST, cDNA Cloning, and Demonstration of Function by Complementation of an *E.coli* MCAT Mutant.

3.1 Introduction

MalonylCoA:ACP Transacylase [MCAT] catalyses one of the first steps in fatty acid biosynthesis in a reaction which transfers the malonyl group of malonyl CoA to ACP producing malonyl ACP (**Figure 3.1**). Malonyl ACP is then used as the two-carbon extender unit for successive rounds of elongation during fatty acid biosynthesis.

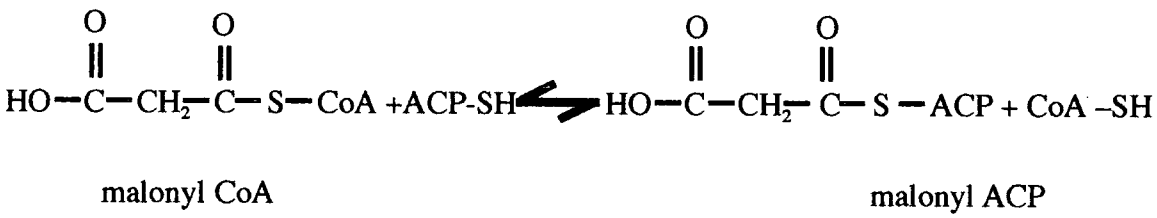


Figure 3.1. Reversible MCAT reaction.

MCAT was first purified from *E.coli* (Ruch and Vagelos 1973a) using ion exchange chromatography, gel filtration chromatography and preparative SDS-PAGE. It is a monomeric enzyme with a subunit Mr of 36kDa when analysed on SDS-PAGE. The enzyme is fully soluble, acidic, with an isoelectric point of pH4.65 and is sensitive to

phenylmethanesulfonyl fluoride (PMSF) inhibition indicating the presence of an essential serine residue in the catalytic domain of the protein (Ruch and Vagelos 1973a). It is also sensitive to both n-ethylmaleimide and iodacetamide treatment, and is inactivated, indicating that a thiol group on the protein is required for maximum catalytic activity (Ruch and Vagelos 1973a).

The malonyl-binding site of the enzyme in *E.coli* was identified and characterised by the isolation of a malonyl-enzyme intermediate (Ruch and Vagelos 1973b). The intermediate was found to be capable of transferring a covalently bound malonyl group to either CoA or ACP, and the formation of the malonyl-enzyme appeared to be readily reversible. [^{14}C] malonyl CoA was used to form a malonyl-enzyme intermediate and subsequent peptide mapping, using paper and Dowex 50 chromatography, following thermolysin digestion was used to locate two major [^{14}C] malonyl labelled peptides. Further proteolytic digestion and mapping of these two peptides resulted in the identification of a radioactive product with identical chromatographic properties to synthetic malonyl-*O*-serine, indicating that a serine residue in *E.coli* MCAT is the covalent binding site for the malonyl group.

The serine identified was located within the tetrapeptide sequence Ala-Gly-His-Ser a sequence identical to that found in the pentapeptide, Leu-Ala-Gly-His-Ser reported by Schweizer *et al.*,(1970) in the MCAT domain of yeast FAS. The Gly-His-Ser sequence has subsequently been shown to be conserved within the substrate binding regions of

known acetyl, malonyl and palmityl transferase domains of type I and II FAS and the predicted active site regions of related polyketide enzymes (Verwoert *et al.*, 1992).

The gene was cloned from *E.coli* by complementation of the *fab*D LA2-89 *E.coli* strain (Magnuson *et.al.*, 1992, Verwoert *et al.*, 1992), which is a temperature sensitive mutant in MCAT activity (Harder *et al.*, 1974). The gene was then over-expressed in *E.coli* using the pET expression system and the resultant MCAT protein assayed for activity. A 1000 fold increase in MCAT activity was found in extracts from the recombinant strain (Verwoert *et al.*, 1992).

The recombinant protein was also used in an X-ray crystallography study where X-ray quality crystals were produced and preliminary data at 2 Å resolution were collected (Sere *et al.*, 1994). The preparation of heavy metal derivatives allowed the complete structure to be resolved at 1.5 Å resolution (Sere *et al.*, 1995).

The *E.coli* MCAT gene has been used in studies in transgenic rape and tobacco. These were performed under the control of the rapeseed napin promoter, which gave high expression during seed development. The structural gene was linked to the leader sequence of enoyl reductase to obtain chloroplast targeting (Verwoert *et al.*, 1994). The gene was targeted successfully to the chloroplast, which was confirmed by immunogold labelling studies, and up to 55 times normal MCAT activity was achieved at the end of seed development. This increased activity did not effect the overall lipid composition of

the transgenic plants indicating that MCAT does not catalyse a rate limiting step in plant fatty acid biosynthesis when grown under these conditions (Verwoert *et al.*, 1994).

There have been few detailed studies on the plant enzyme, although it has been purified from a number of sources. These include avocado, barley, leek, spinach soybean, *Cuphea* and the filamentous cyanobacteria, *Anabaena* (summarised in **Table 3.1**).

Table 3.1 Characteristics of MCAT purified from plant sources.

Plant source	Molecular mass (kDa)	km ACP	km malonyl CoA	Reference
Avocado fruit	40.5	42μM	3.26μM	Caughey and Kekwick (1982)
Barley chloroplasts	41			Hoj and Mikkelsen (1982)
Safflower seeds	22			Shimakata and Stumpf (1982)
Spinach leaves	31	0.4mM	0.5mM	Stapleton and Jaworski (1984)
Soybean	43 (two isoforms both present in leaf one predominant in seeds)			Guerra and Ohlrogge (1986)
Leek leaves	38 + 45 (two isoforms)			Lessire and Stumpf (1983)
Anabaena	36	0.4mM	0.3mM	Stapleton and Jaworski (1984)
Cuphea	27.5			Bruck *et al.*, (1994)

No cDNA clone has been isolated for MCAT from any plant source and no amino acid sequence data is available for the purified protein apart from a literature report of an N-terminal sequence (VAVAELQVE-FI) from *Cuphea lanceolata* (Bruck *et al.*, 1994).

With the intensification of research on whole genomes and the establishment of expressed sequence tag (EST) data-bases a number of putative MCAT genes and cDNA's have been identified (**Table 3.2**), however no definitive proof of function has been shown for any of these plant genes. It is important to furnish such evidence when making gene assignments. Proof of function could be obtained by complementation of function, as shown for the lysophosphatidate acyltransferase (LPAT gene) from maize isolated by complementation cloning of a temperature sensitive *E.coli* LPAT mutant JC201 (Brown *et al.*, 1994). However complementation carried out in this way can give rise to false positives for example in an early attempt to clone plant MCAT by complementation of *fab*D-89 an *E.coli* temperature sensitive MCAT mutant, failed to isolate a gene for MCAT. Instead a cDNA encoding a putative GTP binding protein of the ARF family was isolated (Verwoert *et al.*, 1995).

In order to confirm the function of genes isolated by complementation detailed sequence analyses of the predicted translational product of the complementing cDNA at the amino acid level needs to initially be performed.

Direct proof of function could also be shown by the replacement of a gene in the test organism, following deletion, and the restoration of the lost activity. Kater *et al.*, (1994) demonstrated this for the *Brassica napus* enoyl-ACP reductase (ENR) using an *E.coli* ENR deletion mutant. Such replacement studies can however identify genes which have minimal activity in a biological function providing the selection pressure is high enough.

For example it would not be possible to differentiate between a general amino acid, and a specific amino acid transporter using yeast mutants, as both cDNAs would complement growth of the organism. Perhaps the most stringent method to identify gene function is by over-expression of the gene in a suitable system, purification and direct assay of the purified enzyme.

Table 3.2 Putative MCAT genes and cDNAs within public databases.

Plant Source	Number and type of sequence
Glycine max (soybean)	2 ESTs
Lycopersicon.esculentum (potato)	2 ESTs
Orysa sativa (rice)	1 EST
Ricinus communis (casor bean)	1EST
Zea mays (maize)	4 ESTs
Arabidopsis thaliana	2 BAC 5 GSS

Data from the NPLC catalogue of lipid genes at www.msu.edu/lgc
(Mekhedov *et al.*, 2000)
EST = expresed sequence tag, BAC = genomic clones in bacterial artificial chromosome vectors, GSS = genomic survey sequences.

In plants MCAT is localised to the plastid and there is every indication that it is nuclear encoded and that like ENR, ACP and other FAS components the translational product would have a leader sequence on it which is required for its targeting and import into the plastid. Such chloroplast targeting sequences vary in length depending on the protein but they are typically rich in hydroxylated amino acid residues, low in acidic residues (von

Heinje *et al.*, 1989) and often contain the semi-conserved motif (I/V)-X-(A/C) close to the cleavage site (Gavel and von Heinje 1990). Import into the chloroplast is via an ATP dependant route and involves translocon protein complexes at the outer and inner chloroplast envelope (Sol and Tein 1998). During or immediately after entry into the chloroplast the targeting sequence is cleaved off by a stromal processing peptidase releasing the mature protein (Robinson and Ellis 1984).

It is the intention of this part of the investigation to attempt to clone a plant MCAT cDNA and prove its function using a complementation strategy. The use of database searches to identify plant EST homologs of the *E.coli* MCAT sequence should allow the assignment of a potential putative cDNA / gene or EST. PCR, cDNA library screening and complementation of the *fab*D-89 *E.coli* temperature sensitive MCAT mutant could then be used to isolate the first cDNA clone for MCAT from a plant source and to confirm its identity and that it is functionally active.

3.2 RESULTS.

3.2.1 Alignment of DNA Sequences for MCAT and Database Searching.

Using the *E.coli* amino acid sequence for MCAT (SwissProt accession number P25715) as a starting point, all of the putative bacterial MCAT sequences were extracted from the SwissProt and trEMBL databases and a sequence alignment was made using the multiple alignment tool CLUSTAL W (Thompson *et al.*, 1994) (**Figure 3.2**). This was to look at sequence similarity, and identify conserved regions which could be used for the interrogation of plant databases. The sequences have a high degree of similarity with 40% of their amino acids identical and approximately 67% of the amino acids homologous. There are several blocks of highly conserved regions including a region of 12 amino acids which contains the tetrapeptide Ala-Gly-His-Ser previously identified as the malonyl binding domain in *E.coli*.

The *E.coli* amino acid sequence was used in a Basic Local Alignment (tBLASTn) (Altschul *et al.*, 1990 and 1997) sequence similarity search of the plant Expressed Sequence Tag (EST) and GenBank databases (http://www.ncbi.nlm.nih.gov/) to search for a putative plant MCAT sequence. An EST from Maize [GenBank accession number AA030706] was identified which had strong homology (59%) to the *E.coli* sequence and therefore looked a possible candidate for a plant MCAT (**Figure 3.3a**). This maize EST is 272 base pairs (bp) long (**Figure 3.3b**), although obviously not full length it appeared to code for a sequence homologous to the region between amino acids 102 and 171 of the

Figure 3.2 CLUSTAL W (1.8) multiple sequence alignment of bacterial MCAT amino acid sequences.

```
FABD_ECOLI            -TQFAFVFPGQGSQTVGMLADMAASYPIVEETFAEASAALGYDLWALTQQGPAEELNKTW
FABD_SALTY            -TQFAFVFPGQGSQSVGMLAEMAANYPIVEETFAEASAALGYDLWALTQQGPAEELNKTW
FABD_HAEIN            MKKFAMVFPGQGSQTVGMLADLATEYPIVIETFKQASDALGYDLWYLVQQGPAEELNKTW
FabD_Pasteurella      MKKFAMVFPGQGSQAVGMLAELATEYPVVEETFKQASDVLGYDLWQLVQQGPAEELNKTW
FABD_Vibrio           MSKYAVVFPGQGSQTIGMLADLAAEHSVVEQTFAQASEMLGYDLWDLVQHGTVEELSQTH
FABD_BACSU            MSKIAFLFPGQGSQFIGMGKELYEQVPAAKRLFDEADETLETKLSSLIFEGDAEELTLTY
                       .: *.:******* :**  ::  . . . . * :*.  *  .*  *  .* .***. *
FABD_ECOLI            QTQPALLTASVALYRVWQQQ-GGKAPAMMAGHSLGEYSALVCAGVIDFADAVRLVEMRGK
FABD_SALTY            QTQPALLTASVALWRVWQQQ-GGKMPALMAGHSLGEYSALVCAGVINFADAVRLVEMRGK
FABD_HAEIN            QTQPALLAASVAIYRVWKEKFPQLKPEVMAGHSLGEYSALVCAGVLDFQDAIKLVELRGK
FabD_Pasteurella      QTQPALLAASVAIYRVWQEKYPHLKPDVMAGHSLGEYSALVCAGALDFQDAVKLVELRGK
FABD_Vibrio           ITQPALLATSVALWRIAAAK-EDFKPALVAGHSLGEYSALVCAGVIKFEDAIALVELRGQ
FABD_BACSU            NAQPALLTTSIAVLEKFKES--GITPDFTAGHSLGEYSALVAAGALSFKDAVYTVRKRGE
                       :*****:;*:*: .    .     * . ************.**.:.* **:  *. **:
FABD_ECOLI            FMQEAVPEGTGAMAAIIGLDDASIAKACEEAAE-GQVVSPVNFNSPGQVVIAGHKEAVER
FABD_SALTY            FMQEAVPEGTGGMSAIIGLDDASIAKACEESAE-GQVVSPVNFNSPGQVVIAGHKEAVER
FABD_HAEIN            LMQQAVPEGTGAMYAIIGLDNEAIINACKQAEE-GEVVSAVNFNSPGQVVIAGAKAAVER
FabD_Pasteurella      LMQQAVPEGTGAMYAIIGLDNEAIISACADAAQ-GEVVSAVNFNSPGQVVIAGAKAAVER
FABD_Vibrio           LMQQAVPQGIGAMAAVIGLDNDAIIAACATAAE-DEVVSAVNFNSPGQVVIAGNKAAVVR
FABD_BACSU            FMNEAVPAGEGAMAAILGMDAEALKQVTDKVTEEGNLVQLANLNCPGQIVISGTAKGVEL
                      :*:;*** * *.* *:;*:*  ::  .     : .:;*. .*:*.***:**:*   .*
FABD_ECOLI            AGAACKAAGAKRALPLPVSVPSHCALMKPAADKLAVELAKITFNAPTVPVVNNVDVKCET
FABD_SALTY            AGAACKAAGAKRALPLPVSVPSHCALMKPAADKLAVELAKITFSAPTVPVVNNVDVKCET
FABD_HAEIN            AAALCKEAGAKRALPLAVSVPSHCALMKPAAEQLAVTLENIQINTPTISVLNNVDVKAET
FabD_Pasteurella      AAAACKDAGAKRALPLAVSVPSHCALMKPAADQLAVSLDNIAIRTPTTAVINNVDVACET
FABD_Vibrio           ASELCVEAGARRVMPLPVSVPSHCSLMKPAADELKLALEKVTFNTPVIKLINNVDVAAPV
FABD_BACSU            ASELAKENGAKRAIPLEVSGPFHSELMKPAAEKLKEVLDACDIKDADVPVISNVSADVMT
                      *.  .   **:*.:** ** * *. ******;:*   *    :  .   ::.**..   .
FABD_ECOLI            NGDAIRDALVRQLYNPVQWTKSVEYMAAQGVEHLYEVGPGKVLTGLTKRIVDTLTASALN
FABD_SALTY            DAAAIRDALVRQLYNPVQWTKSVEFIAAQGVEHLYEVGPGKVLTGLTKRIVDTLTASALN
FABD_HAEIN            EGTEIRTALVRQLYSPVRWTETVEKMAQDGVLVLAEVGPGKVLNGLTKRIVGDLQAISVN
FabD_Pasteurella      ENSEIRHALVRQLYSPVRWTETVERMAKDGVQVLVEVGPNKVLTGLTKRIVADLQATAVN
FABD_Vibrio           DAESIKDALVRQLYMPV----------------------------------------
FABD_BACSU            EKADIKEKLIEQLYSPVRFEESINKLIAEGVTTFIEIGPGKVLSGLVKKVNRRLKTIAVS
                      :   *:  *:.*** **
FABD_ECOLI            EPSAMAAALEL--------
FABD_SALTY            EPAALSAALTQ--------
FABD_HAEIN            DVASFNAVEEFLV------
FabD_Pasteurella      DLTSLNAVDELFA------
FABD_Vibrio           -------------------
FABD_BACSU            DPETIELAIQTLKEENDNA
```

(*) conserved residues in all sequences, (:) conserved amino acid substitutions (.) amino acids differ, (-) absence of amino acids. The conserved tetrapeptide AGHS containing the malonyl binding serine residue is highlighted in bold red text.
ECOLI: *Escherichia coli* (P25715), SALTY:*Salmonella typhimirium* (O85140), HAEIN: *Haemophilus influenzae* (P43712), Pasteurella: *Pasteurella multocida* (Q9CJS7), Vibrio: *Vibrio harveyi* (Q9RA34) and BACSU: *Bacillus subtillus* (P71019).
Numbers in parentheses are SwissProt / trEMBL database accession numbers.

Figure 3.3 BLAST sequence search result obtained using *Ecoli* MCAT amino acid sequence to search against the GenBank database.

```
E.coli    102 AGVIDFADAVRLVEMRGKFMQEAVPEGTGAMAAIIGLDDASIAKACEEA 150
              AG   F D ++LV++RG+ MQ+       AM ++IGLD   + + C+ A
EST         5 AGAFXFEDGLKLVKLRGEAMQDXSDAANSAMVSVIGLDSEKVQELCDAA 151

E.coli    146 ACEEAAEGQVVSPVNFNSPGQVVIAG 171
              A +E  E   V   NF  PG   ++G
EST       149 ANDEXDENDRVQIANFLCPGNYAVSG
```

The search resulted in a candidate EST for Maize MCAT (Accession number AA030706).

Figure 3.3b The full nucleotide sequence (272bp) of the maize EST identified as a putative MCAT.

atttgccggt gcttttannt **ttgaggatgg actgaagc**tt gtcaagctaa gaggagaagc tatgcaggat nnttcagatg ctgccaatag tgcgatggtt agtgtgattg gtctggattc agaaaaggtg caagaactat gtgatgctgc taatgacgaa nnngatgaaa acgatagagt tcaaatagca **aactttctgt gccctgggaa tta**tgcagtt tctggtggtg taaaaggtat tgaagtagtc gaagccaaag caaagtcctt ca

The nucleotides highlighted in bold show the positions of the internal oligonucleotide primers designed and used for the PCR amplification of the EST from a maize pUC13 cDNA library.

E.coli protein. It was however a useful starting point to design primers and attempt to obtain via PCR, a potential cDNA probe for cloning. In order to clone a full length cDNA heterologous screening of an existing *Brassica napus* cDNA library was going to be used.

3.2.2 PCR Amplification of the EST (AA030706) From a Maize Plasmid Library and Its Use to Obtain a Longer Clone.

Internal oligonucleotide primers JWSMCAT1 (5'-GAGGATGGACTGAAGCT-3') and JWSMCAT2 (ATAATTCCCAGGGCACAG-3') were designed from the EST sequence (**Figure 3.3b**) and used in a PCR reaction with DNA from a pUC13 maize cDNA library to confirm the presence of the EST sequence within the cDNA library. Reactions were set up in a total volume of 50µl containing 1ng of DNA,10 pmole of each primer, 200 µM dNTPs (50µM of each) Bioline *Taq*® reaction buffer and 1 unit of Bioline *Taq*® DNA polymerase. PCR was carried out in 200µl reaction tubes in a Stratagene Robocycler® for 25 cycles of denaturation at 94°C for 45 seconds, annealing at 50°C for 30 seconds and extension at 72°C for 2.0 minutes. A final extension cycle of 72°C for 5 minutes was used to complete the PCR reaction. 5µl of the reaction product was mixed with DNA gel loading buffer and analysed on a 0.7% agarose DNA electrophoresis gel containing 0.5mg/ml ethidium bromide

A 160bp PCR reaction product was obtained (**Figure 3.4a**). The DNA was sequenced and confirmed to be the expected product for the maize EST sequence (**Figure 3.4b**). Having confirmed the presence of the EST sequence within the cDNA pool of the pUC13 library a further round of PCR was carried out, using the internal primers JWSMCAT1 and JWSMCAT 2 in conjunction with forward and reverse primers JWSMCAT3 (5'-ATGCTTCCGGCTCGTATCTTGTGT-3') and JWSMCAT4 (5'-GCGCTTAAGTTGGGTAACGCCAGG-3') specific to the pUC13 vector sequence flanking the muli-cloning site of the vector, in order to obtain a larger MCAT cDNA fragment suitable for use as a probe in heterologous screening. Primers were used in the combinations JWSMCAT1 + 3, JWS2 + 3, JWS1 + 4 and JWS 2+ 4 to allow for either orientation of cDNA inserts within the vector. The PCR was carried out under the same conditions described above except that four reactions were set up for each primer combination and annealing was carried out on one of the 4 tubes at 58°C, 60°C, 62°C and 64°C. Reaction products were obtained for three of the four primer combinations (**Figure 3.5a**) and from these a 350 bp fragment generated with the primer combination JWSMCAT2 +3 was sequenced and found to be identical to the 3' sequence of the maize EST (**Figure 3.5b**).

Figure 3.4 PCR amplification of putative MCAT EST from Maize.

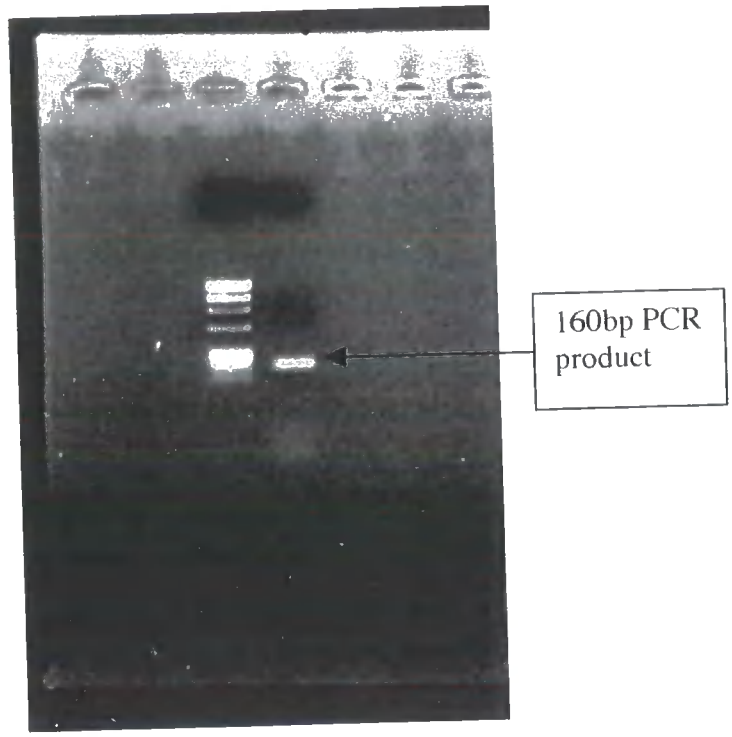


Figure 3.4a A 0.7% agarose gel containing 0.5µg/ml ethidium bromide showing the 160bp product from the PCR reaction using 1.0ng of DNA from a pUC13 maize cDNA library and maize MCAT EST specific primers (JWSMCAT1 and JWSMCAT2). PCR primers (JWSMCAT1 forward primer) 5' GAGGATGGACTGAAGCT 3" and (JWSMCAT2 reverse primer) 5' ATAATTCCCAGGGCACAG 3' correspond to the 5' end and 3'end respectively of the data base sequence.

Figure 3.4b Nucleotide sequence alignment between the 160bp PCR product shown in **Figure 3.4a** and the nucleotide sequence of the maize EST.

```
PCR product    1 GGAGAAGCTATGCAGGATGCTTCAGATGCTGCCAATAGTGCGATGGTTAGTGTGATTGGT 60
                 |||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Zea mays      53 GGAGAAGCTATGCAGGATNNTTCAGATGCTGCCAATAGTGCGATGGTTAGTGTGATTGGT 112

PCR product   61 CTGGATTCAGAAAAGGTGCAAGAACTATGTGATGCTGCTAATGACGAAGTAAATGAAAAC 120
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Zea mays     113 CTGGATTCAGAAAAGGTGCAAGAACTATGTGATGCTGCTAATGACGAANNNGATGAAAAC 172

PCR product  121 GATAGAGTTCAAATAGCAAACTTTCTGTGCCCTGGGAATTAT 162
                 |||||||||||||||||||||||||||||||||||||||||
Zea mays     173 GATAGAGTTCAAATAGCAAACTTTCTGTGCCCTGGGAATTAT 214
```

Figure 3.5 PCR amplification of Maize MCAT DNA for use as a cDNA library probe.

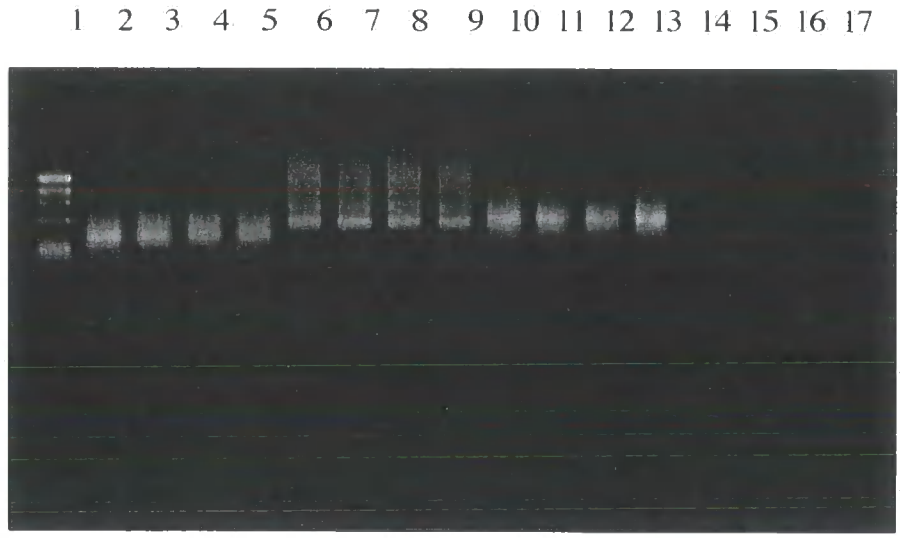


Figure 3.5a A 7% agarose gel of the PCR reaction products obtained using pUC13 maize library DNA with vector and MCAT EST specific primers.

PCR products from reactions carried out at annealing temperatures of 58°C, 60°C, 62°C and 64°C. Lane 1:θ174 DNA size markers, lanes 2-5:products obtained with primers JWSMCAT1+3 lanes 6-9: products obtained with primers JWSMCAT2+3, lanes 10-13:products obtained with primers JWSMCAT1+4 and lanes 14 –17: products obtained with primers JWSMCAT2 +4.
The products from lanes 6-9 were pooled and sequenced

Figure 3.5 b A 350 bp sequence was obtained for the pooled reaction product which had a 60 bp overlap to the maize EST 3' sequence (highlighted in bold).

GCAGTTTCTGGTGGTGTAAAAGGTATTGAAGTAGTCCAAGCCAAAGCAA
GTCCTTCGAGGGCCAAATGACGGTTCGCCTAGCTGTTGCTGGCGCTTTCC
ATACTAGCTTCATGCAACCAGCTGTTTCAAGATTGGAATTTGCGTTGGCT
GCCACTGAGATTAGAACACCTAGAATTTCCGGTCATTTTCCCAATGTTGA
TGCGCAGCCCCACTCAGATCCTAACACCATCAAGCAGATTTTAGCTCAGC
AGGTAACCTTTCCGGTGCAATGGGAAACCACTGTTAAGAATCTTATGGGC
AAGGGGCTTGAGATAAGTTATGAACTCGGA

3.2.3 cDNA Library Screening and Isolation of a Full Length Plant MCAT cDNA.

A *Brassica napus* (λZAP II Jet neuf developing embryo) and a *Zea mays* (λZAP young maize seedling) cDNA library were screened for the presence of a full length MCAT sequence using standard molecular biology protocols (Sambrook *et al.*, 1989). The *Brassica* library was one available within our laboratory and the maize library was kindly supplied by Dr Sean Coughlan, Pioneer Hi-bred, USA.

To provide a radioactive probe for the library screening the 350 bp PCR product obtained in **3.2** was gel purified from the PCR reaction mixture using the GFX® PCR DNA and Gel Band Purification Kit (Amersham Pharmacia Biotech) following the manufacturers protocol. Following purification, approximately 10 ng of the purified product was then [^{32}P] labelled using the *Redi*Prime® random primer DNA labelling system from Amersham Life Science. *E.coli* XL1 blue cells were inoculated with an aliquot of the library and mixed with sterile LB top agarose kept molten at 37^{0}C. This was carefully poured over the surface of a LB agar plate and allowed to set before incubating overnight at 37°C until phage plaques had formed.

In all 28000 and 150000 plaque forming units (pfu) were screened for the *Brassica napus* and maize libraries respectively in the first round of the library screen.

Duplicate plaque filter lifts were taken for each plate and following denaturation, neutralisation and washing these were hybridised against the radioactive probe overnight

at 60°C. The probe was removed and the filters washed to a final stringency of 1 x SSC at 60°C before being exposed to autoradiography film.

A single positive cross-reacting plaque was isolated from the first round screen of the maize library and two positive plaques were isolated from the *Brassica* library.

Phage from these positive plaques was taken through secondary and tertiary rounds of library screening before being excised from the lambda vector and DNA sequenced.

Following DNA sequencing and analyses of the sequence data obtained the maize clone (MCATmaize1) isolated (**Figure 3.6**) was found to be a 800bp clone with a 669 bp ORF which coded for a protein of 223 amino acids. The encoded protein showed strong sequence homology (48%) between amino acid residues 82 and 305 of the *E.coli* protein (**Figure 3.8**). The sequence also contained the conserved serine residue of the Gly,His,Ser,Leu, motive although in this sequence the histidine is a leucine residue (**Figure 3.8**) Thus a putative partial cDNA for maize MCAT had been isolated.

The two sequences obtained for the *Brassica napus* clones isolated were identical. They contained an insert of 1277bp (MCATrap1) containing an open reading frame coding for a protein of 367 amino acids (**Figure 3.7**) and like the maize sequence contained the malonyl binding serine residue within a Gly,Leu,Ser,Leu conserved motif. The translated *B.napus* sequence showed 47% homology to the sequence of the *E.coli* protein and 90% homology to the translated sequence of the partial cDNA clone isolated from maize (**Figure 3.8**).

Figure 3.6 Sequence data obtained for the putative maize MCAT clone (MCAT maize 1).

```
1/1                                   31/11                                 61/21
ccg cgg tgg cgg ccg ctc ta* aat agt gga tcc ccc ggg ctg cag gaa ttc ggc acg agt aac tct gta gat gtt aca tgt ggt ctc agt
                                                                            S   N   S   V   D   V   T   C   G   L   S
91/31                                 121/41                                151/51
ttg gga gag tat acc gct ctt gca ttt gcc ggt gct ttt agc ttt gag gat gga ctg aag ctt gtc aag cta aga gga gaa gct atg cag
L   G   E   Y   T   A   L   A   F   A   G   A   F   S   F   E   D   G   L   K   L   V   K   L   R   G   E   A   M   Q
181/61                                211/71                                241/81
gat gct tca gat gct gcc aat agt gcg atg gtt agt gtg att ggt ctg gat tca gaa aag gtg caa gaa cta tgt gat gct gct aat gac
D   A   S   D   A   A   N   S   A   M   V   S   V   I   G   L   D   S   E   K   V   Q   E   L   C   D   A   A   N   D
271/91                                301/101                               331/111
gaa gta gat gaa aac gat gga gtt caa ata gca aac ttt ctg tgc cct ggg aat tat gca gtt tct ggt ggt gta aaa ggt att gaa gta
E   V   D   E   N   D   G   V   Q   I   A   N   F   L   C   P   G   N   Y   A   V   S   G   G   V   K   G   I   E   V
361/121                               391/131                               421/141
gtc caa gcc aaa gca aag tcc ttc gag ggc caa atg acg gtt cgc cta gct gtt gct ggc gct ttc cat act agc ttc atg caa cca gct
V   Q   A   K   A   K   S   F   E   G   Q   M   T   V   R   L   A   V   A   G   A   F   H   T   S   F   M   Q   P   A
451/151                               481/161                               511/171
gtt tca aga ttg gaa ttt gcg ttg gct gcc act gag att aga aca cct aga att tcc ggt cat ttt ccc aat gtt gat gcg cag ccc cac
V   S   R   L   E   F   A   L   A   A   T   E   I   R   T   P   R   I   S   G   H   F   P   N   V   D   A   Q   P   H
541/181                               571/191                               601/201
tca gat cct aac acc atc aag cag att tta gct cag cag gta acc ttt ccg gtg caa tgg gaa acc act gtt aag aat ctt atg ggc aag
S   D   P   N   T   I   K   Q   I   L   A   Q   Q   V   T   F   P   V   Q   W   E   T   T   V   K   N   L   M   G   K
631/211                               661/221                               691/231
ggg ctt gag ata agt tat gaa ctc gga cct gga aag gtt ata gca ggt att ttc aag agg atc aac aaa ggc act agc att gag aac atc
G   L   E   I   S   Y   E   L   G   P   G   K   V   I   A   G   I   F   K   R   I   N   K   G   T   S   I   E   N   I
721/241                               751/251                               781/261
ggg gct taa acg tct ggg atc tga aga act ctt gat gga acc aag tta cac agc aga gga ttg aaa cat gcc att att gtg gaa cga
G   A
```

The clone is 800 bp long and contains an open reading frame coding for a protein of 223 amino acids. The sequence has 48% homology to amino acid residues between residues 82 and 305 of the *E.coli* protein. The putative malonyl binding serine residue is highlighted in bold red text.

Figure 3.7 The full cDNA sequence and amino acid translation of the *Brassica napus* MCAT clone (MCATrap1).

```
1/1                                   31/11                                  61/21
atc atc cgc acg ctt tcc agc gcc acc acc ATG GCC ACC TCC TCC TCC TCT TCC TTG CTC CTC CCT TCC GTA TCT CTC AAC AAC  CTC TCC
                                       M  A  T  S  S  S  S  S  L  L  L  P  S  V  S  L  N  N  L  S
91/31                                 121/41                                 151/51
TCC TAT AGA AAT GCC TCC TCC CTC GGA TTC TCC GTC AAG AAT CTC ACC CGA TCC AGG GTC TCC ATA AGC GTC TCA GCT GCA TCT CAC ACT
S  Y  R  N  A  S  S  L  G  F  S  V  K  N  L  T  R  S  R  V  S  I  S  V  S  A  A  S  H  T
181/61                                211/71                                 241/81
GCT GTC AAC GAC TCT TTG TTC GCC GAC TAC AAA CCC ACC TCT GCT TTT CTA TTT CCC GGT CAG GGA GCT CAa gca GTA GGG ATG GGC AAA
A  V  N  D  S  L  F  A  D  Y  K  P  T  S  A  F  L  F  P  G  Q  G  A  Q  A  V  G  M  G  K
271/91                                301/101                                331/111
GAA GCT CTC AGT GTT GCA GCA GCT GGA GAG TTG TAT AAC AAA GCT AAC CAT ATC TTG GGG TAT GAT CTT GTT GAC GTA TGT GTT AAT GGA
E  A  L  S  V  A  A  A  G  E  L  Y  N  K  A  N  H  I  L  G  Y  D  L  V  D  V  C  V  N  G
361/121                               391/131                                421/141
CCT AAA GAG AAG CTT GAT TCC ACT GTC ATA AGC CAG CCT GCT ATT TAC GTC ACA AGT TTA GCA GCT GTT GAA TTG CTC CGT GTT CGT GAA
P  K  E  K  L  D  S  T  V  I  S  Q  P  A  I  Y  V  T  S  L  A  A  V  E  L  L  R  V  R  E
451/151                               481/161                                511/171
GGC GGC GAA AAA ATC ATT AAC TCG GTT GAT GTT ACT TGT GGT CTC AGT TTG GGA GAG TAC ACT GCT CTG GCC TTT GCT GGA GCC TTC AGC
G  G  E  K  I  I  N  S  V  D  V  T  C  G  L  S  L  G  E  Y  T  A  L  A  F  A  G  A  F  S
541/181                               571/191                                601/201
TTT GAG GAT GGG CTA AAG CTT GTA AAA CTT AGA GGA GAA GCC ATG CAG GCG GCT GCG GAT GCT GCT AAG AGT GCC ATG GTT AGT ATC ATA
F  E  D  G  L  K  L  V  K  L  R  G  E  A  M  Q  A  A  A  D  A  A  K  S  A  M  V  S  I  I
631/211                               661/221                                691/231
GGG TTG GAC TCA GAA AAA GTT CAG CAG CTA TGT GAT GCA GCA AAT CAG GAA GTA GAT GAA GCT GAC AAA GTT CAG ATT GCA AAT TAC TTA
G  L  D  S  E  K  V  Q  Q  L  C  D  A  A  N  Q  E  V  D  E  A  D  K  V  Q  I  A  N  Y  L
721/241                               751/251                                781/261
TGT CCG GGT AAC TAC GCA GTA TCT GGA GGT CTT AAG GGA ATC GAA GTT GTT GAA GCC AAA GCT AAG TCA TTC AAG GCT CGA ATG ACG GTT
C  P  G  N  Y  A  V  S  G  G  L  K  G  I  E  V  V  E  A  K  A  K  S  F  K  A  R  M  T  V
811/271                               841/281                                871/291
CGG CTA GCT GTT GCG GGT GCT TTC CAC ACA AGT TTC ATG GAA CCA GCA GTG TCT AGA TTA GAA GCT GCA TTG GCG GAA ACG GAG ATC AGA
R  L  A  V  A  G  A  F  H  T  S  F  M  E  P  A  V  S  R  L  E  A  A  L  A  E  T  E  I  R
901/301                               931/311                                961/321
AGT CCA AGG ATC CCA GTG ATC TCC AAT GTG GAT GCA CAG CCT CAT GCA GAT CCA GAC ACA ATC AAG AAG ATA CTT GCA CGC CAG GTG ACA
S  P  R  I  P  V  I  S  N  V  D  A  Q  P  H  A  D  P  D  T  I  K  K  I  L  A  R  Q  V  T
991/331                               1021/341                               1051/351
TCT CCA GTC CAA TGG GAG ACA ACA GTA AAG ACA CTC TTA TCC AAA GGA CTC AAG AGC AGC TAT GAA TTG GGA CCT GGA AAG GTG ATT GCA
S  P  V  Q  W  E  T  T  V  K  T  L  L  S  K  G  L  K  S  S  Y  E  L  G  P  G  K  V  I  A
1081/361                              1111/371                               1141/381
GGG ATA TTC AAG AGA GTA GAT AAA AGT GCT AGC GTC GAA AAC ATC AGT GCT TGA gac tct *tt tgc tgc ttt aga tgt ttg gaa ttt gaa
G  I  F  K  R  V  D  K  S  A  S  V  E  N  I  S  A  *
1171/391                              1201/401                               1231/411
agt ttt gtt atc gaa tat tta tgt aac act tag cac gac gat ata tgt tgt atc ttt gtt taa aat ttc agt tat tga aag tg* aaa aaa
1261/421
aaa aaa aaa aaa aaa aa
```

The cDNA clone is 1277bp long and contains an ORF encoding a protein of 367 amino acids beginning at the start methionine and ending at the first stop codon. The translated amino acid sequence has 47% homology to the MCAT protein sequence from *E.coli* (**Figure 3.8**) and contains the putative malonyl binding serine residue (S166) highlighted in bold red text (S166).

Two possible processing sites for plastid targeting were identified within the translated MCATrap1 sequence. These are (counting from the start methionine) serine 45 using the SigPep program (von Heijne 1989) and leucine 56 determined manually, using the features ((V/I)-X-(A/C):A with an R at between –6 an –10) identified by von Heijne as a cleavage motif (**Figure 3.8**).

The N-terminal sequence reported in the literature for MCAT from *Cuphea* VAVAELQVE-FI (Bruck *et al.*, 1994) was compared against the translated MCATrap1 sequence and no homology was found. This was surprising as there is a relatively high conservation of the amino acid sequence at the N-terminus of the mature protein of MCATrap1 and the *E.coli* protein (**Figure 3.8**). It is possible the amino acid sequence from *Cuphea* represents another protein.

3.2.4 Complementation of the *E.coli* MCAT Mutant *fabD-89* As Proof of Function For the *B.napus* MCAT Clone MCATrap1.

Having isolated a putative MCAT cDNA from *B.napus* and carried out detailed sequence analysis and comparisons to show that it is a full length clone with a high degree of homology to bacterial MCAT proteins and a putative *Arabidopsis* homologue direct proof of function was still needed. For this a complementation strategy using *fab* D-89, an *E.coli* MCAT temperature sensitive mutant, was employed.

Figure 3.8 Amino acid sequence homology of the translated maize cDNA clone (MCATmaize1), *Brassica napus* cDNA clone (MCATrap1) and *E.coli* MCAT protein sequence.

```
MCATmaize1      ------------------------------------------------------------
MCATrap1        MATSSSSSLLLPSVSLNNLSSYRNASSLGFSVKNLTRSRVSISVSAASHTAVNDSLFADY
Ecoli MCAT      ------------------------------------------------------------


MCATmaize1      ------------------------------------------------------------
MCATrap1        KPTSAFLFPGQGAQAVGMGKEALS-VAAAGELYNKANHILGYDLVDVCVNGPKEKLDSTV
Ecoli MCAT      MTQFAFVFPGQGSQTVGMLADMAASYPIVEETFAEASAALGYDLWALTQQGPAEELNKTW
                    ++ +++++ + +++             +    +   +++++      ++ + +  +

MCATmaize1      -------------------------SNSVDVTCGLSLGEYTALAFAGAFSFEDGLKLVK
MCATrap1        ISQPAIYVTSLAAVELLRVREGGEKIINSVDVTCGLSLGEYTALAFAGAFSFEDGLKLVK
Ecoli MCAT      QTQPALLTASVALYRVWQQQGG-----KAPAMMAGHSLGEYSALVCAGVIDFADAVRLVE
                  +++    + +               ::  : .* *****:**. **.:.* *.::**:

MCATmaize1      LRGEAMQDASDAANSAMVSVIGLDSEKVQELCDAANDEVDENDGVQIANFLCPGNYAVSG
MCATrap1        LRGEAMQAAADAAKSAMVSIIGLDSEKVQQLCDAANQEVDEADKVQIANYLCPGNYAVSG
Ecoli MCAT      MRGKFMQEAVPEGTGAMAAIIGLDDASIAKACEEA----AEGQVVSPVNFNSPGQVVIAG
                :**: ** *   ...**.::****. .: : *: *     * : *. .*: .**: .::*

MCATmaize1      GVKGIEVVQAKAKSFEGQMTVRLAVAGAFHTSFMQPAVSRLEFALAATEIRTPRISGHFP
MCATrap1        GLKGIEVVEAKAKSFKARMTVRLAVAGAFHTSFMEPAVSRLEAALAETEIRSPRIP-VIS
Ecoli MCAT      HKEAVERAGAACKAAGAKRALPLPVSVPSHCALMKPAADKLAVELAKITFNAPTVP-VVN
                  :.:* . * .*:  .: :: *.*: . * ::*:**..:*   **   :.:* :.  .

MCATmaize1      NVDAQPHSDPNTIKQILAQQVTFPVQWETTVKNLMGKGLEISYELGPGKVIAGIFKRINK
MCATrap1        NVDAQPHADPDTIKKILARQVTSPVQWETTVKTLLSKGLKSSYELGPGKVIAGIFKRVDK
Ecoli MCAT      NVDVKCETNGDAIRDALVRQLYNPVQWTKSVEYMAAQGVEHLYEVGPGKVLTGLTKRIVD
                ***.: .:: ::*:. *.:*:  **** .:*: : .:*::  **:*****::*: **: .

MCATmaize1      GTSIENIGA----------
MCATrap1        SASVENISA----------
Ecoli MCAT      TLTASALNEPSAMAAALEL
                  : . :.
```

- indicates the absence of an amino acid, * indicates where all three sequences are identical ,: show conserved substitutions across all three sequences and + indicates areas of the N-terminus where *B.napus* and *E.coli* are identical.

The translated sequence for maize shows 48% amino acid sequence homology to the *E.coli* sequence and 90% homology to the *B. napus* sequence. The translated *B.napus* sequence shows 47% homology to the sequence of the *E.coli* protein. The residues in red are the conserved residues around the active site serine and the two residues in blue indicate putative chloroplast target sequence cleavage sites.

The *E.coli* strain *fab*D-89 is unable to grow at growth temperatures above 35°C due to a single G to A mutation at bp 1301 in the sequence. This mutation alters the tryptophan (257) codon (TGG) into an amber codon (TAG) which results in premature termination and a truncated (27 kDa) inactive protein (Verwoert *et al.*, 1994). However when complemented with an *E.coli* gene for MCAT the strain was able to grow at the non-permissive temperature of 39°C (Verwoert *et al.*, 1992). Complementation experiments can be quite difficult when using a plant gene to complement a bacterial strain, as the plant gene may not function correctly, due to lack of post–translational modifications or the gene product may be lethal to the host strain.

Approximately 15 ng of MCATrap1, pMIC6, a positive control containing an *E.coli* chromosomal MCAT insertion (Dr A.R.Stuitje) or pBSKS$^+$control plasmid without insertion were transformed into rubidium chloride competent fabD-89 *E.coli* cells using the standard transformation protocol. Two transformations of each plasmid were prepared. The transformed cells were incubated for 60 minutes at 30°C on an orbital shaker (150 rpm). Following this incubation both a 100µl aliquot of cells and the remaining cells pelleted following centrifugation in a microfuge, were plated on separate LB ampicillin (100µg/ml) agar plates containing 0.2% NaCl. One plate of each plasmid transformation was incubated at 30°C (permissive temperature) and one was incubated at 39°C (non-permissive temperature) and allowed to grow overnight. The plates were

removed from the incubators and the resulting colonies counted (**Table 3.3**). A photographic comparison of the colony growth on the plates is shown in **Figure 3.9.**

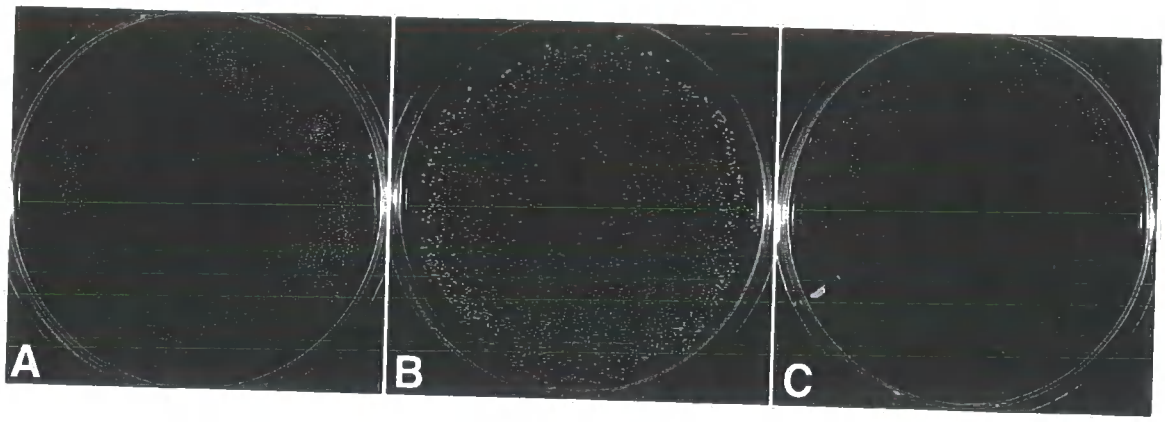
Table 3.3 Growth and colony number following complementation of *fab*D-89 at permissive and non-permissive temperatures.

Transformant	Colony number at 30°C		Colony number at 39°C	
	100µl	pellet	100µl	pellet
pBSK negative control	1000	complete plate$^+$	0	<100
pMIC6 positive control	250	complete plate$^+$	250	complete plate$^+$
pMCATrap1	500	complete plate$^+$	100	complete plate$^+$

Complete plate$^+$ indicates an overgrown plate where the colony numbers were too numerous to count.

Although a similar number of colonies were present on the plates for pMCATrap1 and pMIC6 grown at 39°C a comparison of the colony size showed the colonies formed when *fab*D-89 is complemented with the *B.napus* MCAT cDNA are much smaller than those complemented with the *E.coli* gene (**Figure 3.9**). This is probably the result of a heterologous complementation of an *E.coli* strain with a plant protein containing a full length cDNA complete with a plastid target sequence. Removal of this leader sequence may make the plant gene more compatible and result in better growth. Confirmation of

Figure 3.9 Complementation of the *E.coli* temperature sensitive mutant *fab-D89* with *Brassica napus* MCAT (MCATrap1).



Competent cells of *E.coli fabD-89* were transformed with [A] *Brassica napus* MCAT (MCATrap1), [B] *E.coli* MCAT (pMIC6) positive control and [C] pBSKS plasmid without insert, negative control. Transformed cells were plated onto duplicate plates of LB + ampicillin (100 µg / ml) and grown at the non-permissive temperature of 39°C for 24 hours. A similar number of colonies were present on both the MCATrap1 plate (A) and the pMIC6 positive control plate (B), although the bacteria complemented with the plant gene produced smaller colonies. This complementation was confirmed by the further growth of the complemented cells following re-plating and re-growth at the non-permissive temperature. None of the small number of colonies present on the negative control plate was able to grow on replating.

this complementation was achieved by further streaking eight colonies from each of the plates onto fresh LB ampicillin plates and re-incubating at the non-permissive temperature. Five out of the eight streaks grew well for the MCATrap1 and pMIC6 whereas none of the pBSK showed any further growth.

The successful complementation of the bacterial MCAT mutant provides direct proof that the cDNA isolated in this study is a functional MCAT. It also shows that a plant MCAT can complement a bacterial MCAT mutant. This has previously been demonstrated for another FAS component where plant enoyl reductase complemented bacterial enoyl reductase (Kater *et al.*, 1994).

3.2.5 Identification of a Corresponding *Arabidopsis* MCAT Genomic Sequence in the GenBank Database.

The full length cDNA sequence of *B.napus* MCATrap1 was used to search the *Arabidopsis* genome sequence database and identified a genomic sequence for a putative *Arabidopsis* MCAT (GenBank accession number AC002338).

The sequence was identified within the BAC clone T27E13 and located on chromosome II near the molecular marker m283. The sequence assignment within the database made by alignment with the *E.coli* MCAT sequence, was found to contain ten putative intron regions.

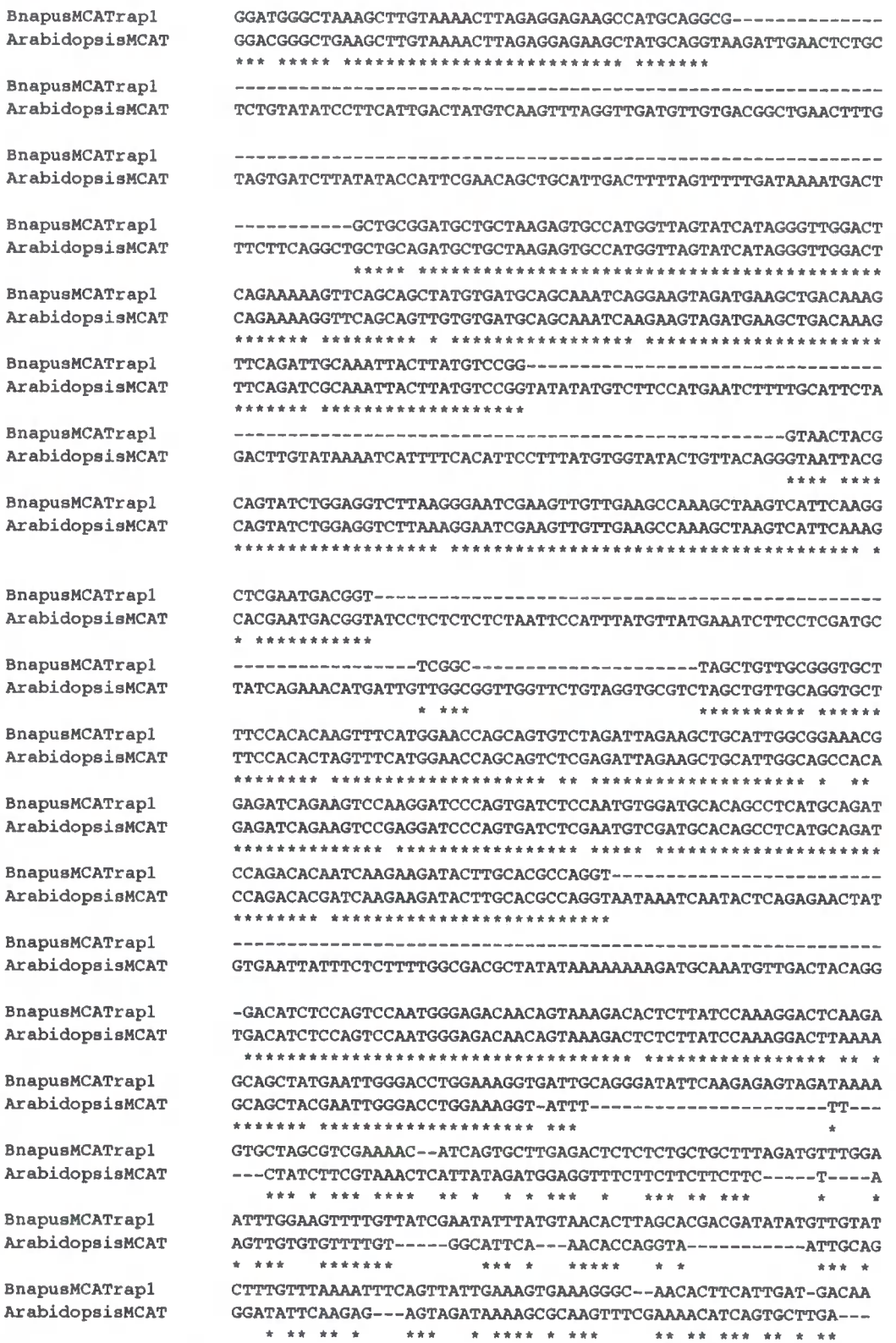
The MCATrap1 sequence was aligned against the full genomic sequence and confirmed that the intron assignments originally made were correct and the *Arabidopsis* gene contains ten introns (**Figure 3.10**).

The above sequence identification was made while the *Arabidopsis* genome sequencing program was still in progress (May 1998). Recently the *Arabidopsis* genome-sequencing project has been completed and the translation product for putative MCAT has been identified (Accesssion number AAC16926). The sequence is 1339 bp long with an open reading frame of 1060 bp coding for a protein of 367 amino acids with a molecular weight of 38.6kDa.

The *Arabidopsis* amino acid sequence was aligned against the translated MCATrap1 sequence (**Figure 3.11**) and shows 92% identity. It also contains the conserved malonyl binding serine residue within the motif Leu, Ser, Leu, Gly.

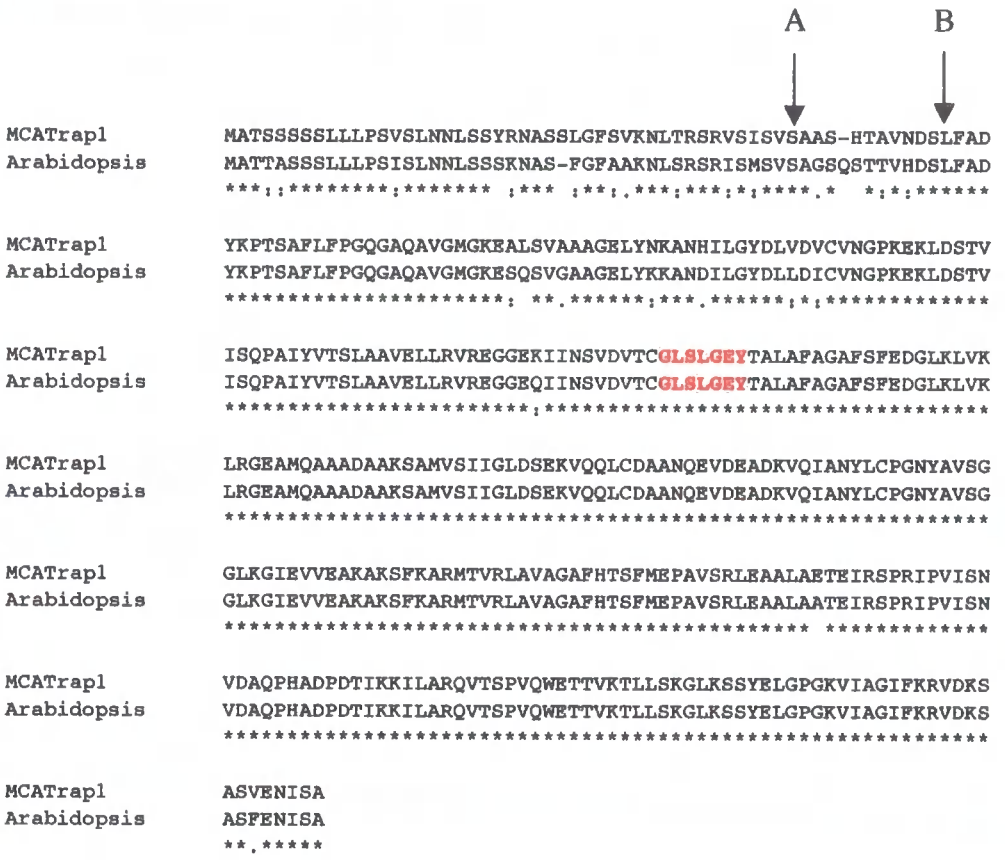
Figure 3.10 CLUSTAL W (1.8) DNA sequence alignment of the *Brassica napus* cDNA clone MCATrap1 and the genomic clone for *Arabidopsis* MCAT.

```
BnapusMCATrap1      ATCATCCGCACGCTTTCCAGCGCCACCACCATGGCCACCTCCTCCTCCTCTTCCTTGCTC
ArabidopsisMCAT     ------------------------------ATGGCCACCACCGCCTCCTCTTCCTTGCTC
                                                   ********* ** *****************

BnapusMCATrap1      CTCCCTTCCGTATCTCTCAACAACCTCTCCTCCTATAGAAATGCCTCCTCCCTCGGATTC
ArabidopsisMCAT     CTCCCTTCCATATCTCTCAACAATCTCTCCTCCTCTAAAAATGCCTCCT---TTGGCTTC
                    ********* ************* ********** ** ***********   * ** ***

BnapusMCATrap1      TCCGTCAAGAATCTCACCCGATCCAGGGTCTCCATAAGCGTCTCAGCTGCATCTCA---C
ArabidopsisMCAT     GCCGCCAAGAATCTCAGCCGATCTAGGATTTCTATGAGCGTCTCTGCTGGATCTCAGAGT
                     *** ********** ****** *** * ** ** ******** **** ******

BnapusMCATrap1      ACTGCTGTCAACGACTCTTTGTTCGCCGACTACAAACCCACCTCTGCTTTTCTATTTCCC
ArabidopsisMCAT     ACTACTGTTCACGATTCTCTGTTCGCTGATTACAAACCCACCTCTGCTTTTCTCTTTCCC
                    *** ****  **** *** ******** ** ************************ ******

BnapusMCATrap1      GGTCAGG-------------------------------------------------------
ArabidopsisMCAT     GGTCAGGTAACTTCCGATTACATGTTATCAATCTGAACAATGAAAATGTGGACTCTGATT
                    *******

BnapusMCATrap1      ------------------------------------------------------------
ArabidopsisMCAT     GATTTCGTATCAATTTAGGAAATTTGGAATTGCCCAGATTGAAAAGTTTGCTAATTTTGA

BnapusMCATrap1      ------------------------------------------------------------
ArabidopsisMCAT     TCTGATTGATGTGATTTTAATTACAATGTAGGGTCTTCATCATAACGTCATATTCTTTGC

BnapusMCATrap1      ------------------------------------------------------------
ArabidopsisMCAT     TTGTAATCTCTCAGAACTCGTTTCAATTTTTAGATCAAATGTGACAAAGTTTAGATGAAT

BnapusMCATrap1      ---------------GAGCTCAAGCAGTAGGGATGGGCAAAGAAGCTCTCAGTGTTGCAG
ArabidopsisMCAT     TTGCTATGTCTCAGGGAGCTCAAGCAGTAGGAATGGGAAAAGAGTCTCAGAGTGTTGGAG
                              ***************** ***** *****  ***  ******* **

BnapusMCATrap1      CAGCTGGAGAGTTGTATAACAAAGCTAACCATATCTTGGGGTATG---------------
ArabidopsisMCAT     CAGCTGGAGAGTTGTATAAGAAAGCTAATGATATCTTAGGGTATGTTCCTTTGTTTAAAG
                    ******************** ********  ******* *******

BnapusMCATrap1      ------------------------------------------------------------
ArabidopsisMCAT     TTGAAACCTTTTTGTTTGTTTTTGTGAATGAATAAACTATATGGTGGCTTTCTGCAGGTA

BnapusMCATrap1      --ATCTTGTTGACGTATGTGTTAATGGACCTAAAGAGAAGCTTGATTCCACTGTCATAAG
ArabidopsisMCAT     TGATCTTTTGGATATTTGTGTTAATGGACCAAAAGAGAAGCTTGATTCTACGGTCATAAG
                      ***** * **  * *************** ****************** ** ********

BnapusMCATrap1      CCAG----------------------------------------------------------
ArabidopsisMCAT     CCAGGTATCTTTGCAACTTTTGGTTAGGGACAACTTCAGATTTATGATCTTCTTTCTTGT
                    ****

BnapusMCATrap1      ------------------------------------------------------------
ArabidopsisMCAT     CTGTTGTCACCTAACTCTAGCAAAACTTAATTTTTATGTGATGTTTCCTTTTAGTATGTA

BnapusMCATrap1      ---------------------------------------------CCTGCTATTTACGTCA
ArabidopsisMCAT     CATTTTGGTATGATGTCTAAGTTATTGGTCTATCATTTTCTCAGCCTGCTATTTATGTCA
                                                               *********** ****

BnapusMCATrap1      CAAGTTTAGCAGCTGTTGAATTGCTCCGTGTTCGTGAAGGCGGCGAAAAAATCATTAACT
ArabidopsisMCAT     CAAGTTTAGCAGCAGTTGAATTGCTCCGTGTTCGTGAAGGCGGAGAACAAATAATTAACT
                    ************* ****************************** *** **** *******

BnapusMCATrap1      CGGTTGATGTTACTTGTGGTCTCAGTTTGGGAGAGTACACTGCTCTGGCCTTTGCTGGAG
ArabidopsisMCAT     CGGTTGATGTGACTTGCGGTCTCAGTTTGGGAGAGTATACTGCTCTGGCTTTTGCTGGAG
                    ********** ***** ******************** *********** **********

BnapusMCATrap1      CCTTCAG-------------------------------------------------------
ArabidopsisMCAT     CCTTCAGGTTAGTTAACATCTGAATCAATATCCTTAACTGAATGATAGAGGTATCAAATC
                    *******

BnapusMCATrap1      ------------------------------------------------------------
ArabidopsisMCAT     TCGGATTTTCAAATATCATTGTAAGAAAGTATGTCACTTTGTTCCCTTTTACCTATGTGC

BnapusMCATrap1      ------------------------------------------------------CTTTGA
ArabidopsisMCAT     ATCATCACCTAACTTACTTTTGTTCTTTTCTCTGTTACTGTTTTTATTCTCAAGCTTCGA
                                                                     *** **
```



* indicates blocks of sequence identity where the MCATrap1 clone aligns with the 10 putative exon assignments made for the *Arabidopsis* genomic clone.

Figure 3.11 CLUSTAL W sequence alignment of the amino acid sequences of *Brassica napus* MCATrap1 and the putative *Arabidopsis* MCAT (Accession number AAC16926).



The sequences show 92% identity. (*) conserved residue, (:) conserved substitution (strong homology), (.) different amino acid (weak homology) and (–) absence of an amino acid.
The *Arabidopsis* sequence like both the MCAT maize1 and the MCATrap1 sequence contains the conserved malonyl binding serine residue within the motif LSLGEY highlighted in bold red text. (A) indicates the serine residue predicted by SigPep and (B) indicates the leucine residue predicted manually as the potential signal peptide cleavage site.

3.2.6 Overexpression of *Brassica napus* MCATrap1

An attempt was made at over-expression of MCATrap1 protein using the pET Protein Expression System from Novagen. Oligonucleotide primers MCATox1 (5'GGAATTCCATATGGCATCTCACACTGCTGTCAACG3") and MCATox2 (5" GCGCAGATCTGCAGCAGAGAGAGTCTCAAGC3") were designed to introduce a 5' NdeI site and a 3" BglII to the MCATrap1 sequence. The primers were designed to produce a mature protein of 321 amino acids beginning at the predicted target peptide cleavage site Ser 45 and ending at the stop codon of the MCATrap1 sequence. The primers were used together in a standard PCR reaction with MCATrap1 template DNA and Taq®DNA polymerase for 30 cycles annealing at 60°C. The use of Taq® polymerase in this reaction resulted in a PCR product containing an A overhang at the 3'end of the DNA strand suitable for ligation into a T vector for sub-cloning.

The PCR product obtained was gel purified and a1µl aliquot ligated into the TOPO TA vector (Invitrogen). The ligation mixture was transformed into TOPO one shot *E.coli* cells following the manufacturers protocol and the cells were plated on LB ampicillin plates containing X-gal and IPTG for blue / white colony selection. Transformants were checked for the presence of inserts of the correct size in a PCR reaction using M13 forward and reverse primers. These priming sites flank the T/A sites on the TOPO TA vector. Positive transformants were re-grown for plasmid isolation and inserts were cut from the purified plasmid in a double restriction digest with NdeI and BglII. The released

insert was gel purified and ligated into pET 24a vector which had been previously digested with NdeI and BamHI. The ligated vector was transformed into *E.coli* X11 blue cells plated onto LB kanamycin plates and grown overnight at 37°C. Plasmid DNA was isolated from single transformants DNA sequenced and re-transformed into *E.coli* BL21cells for over-expression.

Several attempts where made at over-expression using standard induction protocols) but no over-expressed protein band of the correct molecular weight were visualised on SDS-PAGE either in freeze thaw cell free extracts or when whole cells were extracted with SDS gel loading buffer.

There are several possible reasons why this over-expression did not work, these include problems with the sub-cloning between the T/A and pET24a vectors, problems with the codon usage between plant and bacterial systems or the over-expressed product is toxic in the host strain. More detailed studies are required to resolve this expressional problem but are outside the present study.

3.3 Discussion

During this study a 272 bp putative EST cDNA for maize MCAT was identified by homology to the amino acid sequence of the *E.coli* protein. Using this information as a starting point PCR was used to generate a cDNA probe, which was used to isolate a partial cDNA clone from a λZAP young maize seedling library. The isolated cDNA clone was 800bp long with a 669 bp ORF encoding a protein of 223 amino acids with 48% homology to the amino acid sequence of the *E.coli* protein. This clone was not full length, but represents a partial cDNA for plant MCAT.

The maize probe was also used in the heterologous screening of a *Brassica napus* λZAPII Jet Neuf developing embryo library and a 1200 bp cDNA clone was isolated. This contained a 1101 bp ORF encoding for a protein of 367 amino acids with 90% homology to the maize translated amino acid sequence and 47% homology to the *E.coli* protein. Both the maize and the *Brassica napus* cDNA clones contained the conserved malonyl binding serine residue within the conserved sequence motif GHSXGEY, found in known acetyl, malonyl and palmityl transferase domains of type I and II FAS and the predicted active site regions of related polyketide enzymes (Verwoert *et al.*, 1992). Sequence analyses allowed two (serine 45 and leucine 56) possible processing sites for a plastid targeting sequence within the protein to be identified using SigPep computer prediction and manual identification.

A genomic sequence for a putative *Arabidopsis* MCAT was identified in the GenBank database, which contained ten presumptive introns assigned by computer prediction. *Arabidopsis* is a close relative of *Brassica napus*, both being members of the *Cruciferae* (Slabas *et al.*, 1992) so an alignment was made between the *Brassica napus* cDNA sequence and the *Arabidopsis* genomic sequence which confirmed that the intron assignment made by computer prediction was correct.

Finally, transformation with plasmid DNA containing the isolated *Brassica napus* cDNA for MCAT allowed successful growth of the MCAT *fab*D-89 *E.coli* mutant at the non-permissive temperature of 39°C. This demonstrated for the first time proof of function for a plant cDNA for MCAT and that plant MCAT can complement for bacterial MCAT.

Future work would involve a more detailed attempt to over-express the encoded protein in a bacterial expression system and the purification of the recombinant protein for use in interaction studies and co-crystallisation studies with other plant FAS components, particularly ACP. It is clear that these will possibly be aided by an understanding of the order of binding and kinetic mechanism of the enzyme.

Chapter 4

Structural Studies on Acyl Carrier Protein (ACP) from Bacteria and Plants.

4.1 Introduction.

Acyl carrier protein (ACP) is the essential co-factor for all of the reactions of the fatty acid synthetase system in bacteria and plants (Prescott and Vagelos 1972, Slabas and Fawcett 1992). It is also a component of acyl-ACP desaturation (McKeon and Stumpf 1982), plastid localised acyl-transferase reactions (Frentzen *et al.*, 1983) and in many other reactions involving acyl transfer steps, including the synthesis of polyketide antibiotics (Shen *et al.*, 1992). It is a small acidic protein with a molecular mass of <10kDa and is one of the most abundant proteins (~6 x 10^4 molecules per cell) found in *E.coli* (Magnuson *et al.*, 1993). ACP is synthesised as an apo protein, which undergoes a post-translational modification to an holo form when a 4'posphopantetheine group is transferred from coenzyme A to an essential serine residue within the ACP (Flugel *et al.*, 2000). Pantothenylation of apo-ACP is catalysed by holo-ACP synthetase in a Mg^+ dependant reaction (Elovson and Vagelos 1968). Acyl groups are then attached as thioesters to the terminal group of this phosphopantotheine arm of the holo protein in a reaction catalysed by acyl-ACP synthetase (Acps) (Majerus *et al.*, 1965).

ACP was the first of the proteins involved in the fatty acid biosynthesis pathway to be purified from both *E.coli* (Majerus *et al.*, 1964) and plants (Simoni *et al.*, 1967). The *E.coli* protein was also chemically synthesised in a biologically active form (74 of the 77 residues) in an effort to gain an understanding of its structure and the mechanism of its acylation (Prescott and Vagelos 1972). It has subsequently been isolated from many plant species, including barley (Hoj and Svedson, 1983), spinach (Kuo and Ohlrogge, 1984) and oil seed rape (Slabas *et al.*, 1987). The primary amino acid sequence of the mature protein of all of the plant ACPs is highly conserved, particularly in the mid-region surrounding the phosphopantethylated serine (**Figure 4.1**). NMR structural studies have also shown a high degree of structural homology between the spinach and the *E.coli* proteins (Oswood *et al.*, 1997). Tissue specific isoforms of ACP have been identified in a number of plant species. These include *Arabidopsis* (Hlousek-Radojcic *et al.*, 1992), rape (Safford *et al.*, 1988 and De Silva *et al.*, 1990), barley (Hoj and Svendsen 1984), spinach (Ohlrogge and Kuo 1985), *Cuphea* (Schutt *et al.*, 1998) and coriander (Chung Suh *et al.*, 1999). Although the role of these different isoforms is not fully understood at present it has been shown that specific isoforms are involved in the activity of the Δ4 and Δ6 desaturases present in coriander seeds (Chung Suh *et al.*, 1999).

In recent years there have been major advances in our understanding of the structure of the components of type II FAS from both bacteria and plants. The crystal structure of

Figure 4.1 Clustal W amino acid sequence alignment of plant ACP isoforms.

```
ACP1_BRANA      -MSTTFCSSVSMQAT-----SLAAT---------------TRISFQKPALVS---RTNLS
ACP2_BRANA      -MSTTFCSSVSMQAT-----SLAAT---------------TRISFQKPGLVS---RTNLS
ACP3_BRANA      -MATTFSASVSTLAT-----SLATP---------------TRISFQKPALVS---RTNLS
ACP1_ARATH      -MATQFSASVSLQTS-----CLAT----------------TRISFQKPALISNHGKTNLS
ACP2_ARATH      MASIAASASISLQARPRQLVSLF---PRISHLVLLVLL-WQAIAASQVKSFSNGRRSSLS
ACP3_ARATH      MASIATSASTSLQARPRQLVSLFSSSPRISHLFLTEFLSWQVIGAKQVKSFSYGSRSNLS
ACP1_CUPLA      MASAAAGAS-ICIKS-----ASFS------PLA------PGRISSLRSVSLPVSRKSFPS
ACP4_CUPLA      MASAAAGAS-ICIKS-----ASCS------PLA------PGRISSLRSVSLPVSRKSFPS
ACP2_CUPLA      MASAAPASS-ICIKS-----ASCS------ALA------PGRISSLRSVSLPVSRKSFFS
ACP3_CUPLA      MATAAAGSSLICIKS-----ASCSL---NRAQV------PSGLSSLRSVSLPISGKIFPS
ACP1_HORVU      MAHCLAAVS-SFSPS-----AVRRR---LSSQV------ANVVSSRSSVSFHSRQMSFVS
ACP2_HORVU      MASAAASAV-----------SFAR----------------PVKAICVNSVSFSALRKDNVS
ACP3_HORVU      MASIAGSAV-----------SFAK----------------PVKAINTNSLSFSGARRGNAF
ACP2_SPIOL      MASITGSSV-----------SFKC----------------APLQSSFNSKNYAL--KSSVT
ACP1_SPIOL      MASLSATTTVRVQPSSS---SLHK----------------LSQGNGRCSSIVCLDWGKSS
                                    .

ACP1_BRANA      FNLSRSIPTR-LSVSC------AAKPETVEKVSKIVKKQLSLKDDQNVVAETKFADLGAD
ACP2_BRANA      FNLRRSIPTR-LSVSC------AAKPETVEKVSKIVKKQLSLKDDQKVVAETKFADLGAD
ACP3_BRANA      FNLRRSIPTR-LSVSC------AAKPETIEKVSKIVKKQLSLKDDQKVVAETKFADLGAD
ACP1_ARATH      FNLRRSIPSRRLSVSC------AAKQETIEKVSAIVKKQLSLTPDKKVVAETKFADLGAD
ACP2_ARATH      FNL-RQLPTR-LTVSC------AAKPETVDKVCAVVRKQLSLKEADEITAATKFAALGAD
ACP3_ARATH      FNR-RQLPTR-LTVYC------AAKPETVDKVCAVVRKQLSLKEADEITAATKFAALGAD
ACP1_CUPLA      LKS-SKSSFA-LRVSC------QAKPETVAKVCGIVKKQLALPDDSEVNGLSKFSALGAD
ACP4_CUPLA      LRS-SKGSFA--RVSC------QAKPETVAKVCRIVKKQLALPDDSEVNGLSKFSALGAD
ACP2_CUPLA      R----RGSFS-LRVNC------QAKPETVTKVCNIVKKQLALPDDSDVSGVSKFSALGAD
ACP3_CUPLA      LRS-SRGPLS-FRVCC------QAKQETVTRVCEIVKKQLALPEDSEVNGLSKFSALGAD
ACP1_HORVU      ISS-RPSSLR-FKICCAAMGEAQAKKETVDKVCMIVKKQLAVPDGTPVTAESKFSELGAD
ACP2_HORVU      FRL-QPVPQR-FSVCC------AAKKETVEKVCDIVKSQLALSDDTEVSGSSTFADLGAD
ACP3_HORVU      LRL-QPVPMR-FAVCC------SAKQDTVEKVCEIVKKQLAVPEGTEVCGTTKFSDLGAD
ACP2_SPIOL      FWRRTPVMPRGLSVSC------AAKPEMVTKVSDIVKSQLALAEDAKVTGETKFSEIGAD
ACP1_SPIOL      FPTLRTSRRR--SFIS------AAKKETIDKVCDIVKEKLALGADVVVTADSEFSKLGAD
                                .  .       ** : : :*. :*:.:*::     : . : *: :***

ACP1_BRANA      SLDTVEIVMGLEEEFHIEMAEEKAQKITTVEEAAELIDELVQAKK-
ACP2_BRANA      SLDTVEIVMGLEEEFDIEMAEEKAQKIATVEEAAELIEELVQAKK-
ACP3_BRANA      SLDTVEIVMGLEEEFDIEMAEEKAQKIATVEEAAELIEELVLLKK-
ACP1_ARATH      SLDTVEIVMGLEEEFNIQMAEEKAQKIATVEQAAELIEELINEKK-
ACP2_ARATH      SLDTVEIVMGLEEEFGIEMAEEKAQSIATVEQAAALIEELLFEKAK
ACP3_ARATH      SLDTVEIVMGLEEEFGIEMAEEKAQSIATVEQAAALIEELLLEKAK
ACP1_CUPLA      SLDTVEIVMGLEEEFGISVEEESAQSIQTVQDAADLIEKLMEKKGH
ACP4_CUPLA      SLDTVEIVMGLEEEFGISVEEESAQSIQTVQDAADLIEKLMEKKGH
ACP2_CUPLA      SLDTVEIVMGLEEEFGISVEEESAQSIQTVQDAADLIEKLMEKKGH
ACP3_CUPLA      SLDTVEIVMGLEEEFGISVEEESAQSIQTVQDAADLIEKLVGNKK-
ACP1_HORVU      SLDTVEIVMGLEEEFNITVDETSAQDIATVQDAANLIEKLVTEKTA
ACP2_HORVU      SLDTVEIVMGLEEAFGISVEESSAQTIATVEDAANLIDSLVGK---
ACP3_HORVU      SLDTVEIVMGLEEEFQISVEETSAQAIATVEDAATLIDKLVSAKSS
ACP2_SPIOL      SLDTVEIVMKLEEEFGVTVEEENAQTITTIQEAADMIEALQQNK--
ACP1_SPIOL      SLDTVEIVMNLEEEFGINVDEDKAQDISTIQQAADVIESLLEKKA-
                ********* *** * : : * .** * *:::** :*: *
```

(*) conserved amino acids, (:)conserved substitutions, (.)where amino acids differ, (-) amino acids missing.
ARATH: *Arabidopsis thaliana* (ACP1=P11829, ACP2= P25701, ACP3= P25702), BRANA: *Brassica napus* (ACP1=P10352, ACP2= P17650, ACP3= P32887), CUPLA: *Cuphea lanceolata* (ACP1=P52411, ACP2=P52412, ACP3=P52413), HORVU:*Hordeum vulgare* (ACP1= P02902, ACP2= P08817, ACP3= P15543), SPIOL: *Spinacia oleracea* (ACP1= P07854, ACP2= P23235).
Numbers in parentheses are SwissProt database accession numbers.

The primary amino acid sequence is highly conserved across plant species and different isoforms of ACP, particularly in the middle and C-terminal regions surrounding the phosphopantethenylated serine residue (highlighted in bold red type).

several of the enzymes has been solved. These include enoyl-ACP reductase (ENR) (Rafferty *et al.*, 1994 and 1995, Baldock *et al.*, 1998), βketo-ACP reductase (βKR), (Fisher *et al.*, 1999 and 2000) acyl carrier protein synthase (AcpS) (Chirgadze *et al.*, 2000), and β-ketoacyl-ACP synthase II (Moche *et al.*, 2001).

Nuclear magnetic resonance (NMR) solution structures of *E.coli* holo-ACP (Holak *et al.*, 1988 and 1989 and Kim and Prestegard 1990), *Streptomyces* polyketide synthase actinorhodin apo-ACP (Crump *et al.*, 1997) and *Bacillus subtilis* apo and holo-ACP (Xu *et al.*, 2001) have all been determined. A crystal complex of holo-acyl carrier protein synthase together with ACP from *Bacillus subtilis* has also been reported (Parris *et al.*, 2000). Detailed crystallographic studies are still required however in order to gain a better understanding of how the various forms of ACP, (apo holo and acyl) interact with their modifying enzymes such as HAS and Acps. X-ray crystallographic studies with apo, holo and acyl-ACPs will also provide information on crystallisation conditions which will aid in co-crystallography and protein interaction studies with other components of the bacterial and plant FAS pathway. It is the intention of this part of the investigation to obtain ACP from *E.coli* and a number of plant sources and embark on a detailed study aimed at the resolution of the full atomic structure of ACP from either a plant or bacterial source. This will be a prelude to future co-crystallography studies and protein interaction studies with other FAS components aimed at the re-constitution of a plant FAS metabolon.

For this detailed study 'wild type' *E.coli* protein which can be relatively easily purified (Majerus *et al.*, 1964) in mg quantities will be used as a starting material to determine reaction conditions for chemical acylation and initial crystallography trials. Recombinant bacterial and plant ACPs will be over-expressed and purified and used for extensive crystallographic studies. These will include enzymatic acylation reactions with HAS to produce holo and various acyl chain length ACPs and the generation of site directed ACP mutants in order to be able to produce sites for heavy metal derivatives to aid in the resolution of the crystal structure.

Conformational gel analyses (Post-Beittenmiler *et al.*, 1991) together with electrospray and MALDItof mass spectrometer analysis will be used to characterise the various derivatives of ACP produced prior to crystallography trials being carried out. It is hoped that a form of ACP can be produced which can be successfully crystallised, is stable in the X-ray beam and produces diffraction data that will enable the resolution of the crystal structure of an acyl-ACP to be resolved for the first time.

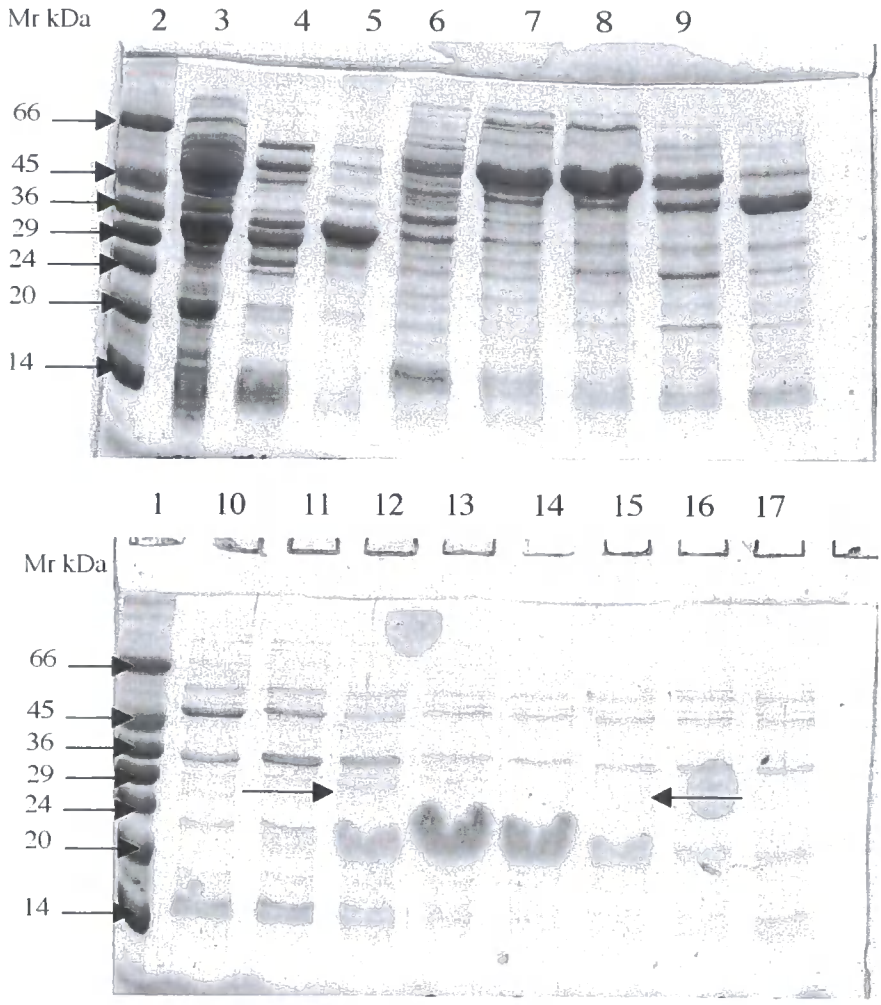
4.2 Results.

Purification of Wild-Type ACP from *E.coli*, its Acylation Using N-acylimidazoles and Preliminary Crystallography Trials with acyl-ACPs.

4.2.1 Purification of Wild-Type ACP from *E.coli*

E.coli ACP was purified from fermenter grown (UCL) *E.coli* K12 cells using a modification of the acid precipitation method originally described by Majerus *et al.*, (Majerus *et al.*, 1964). The steps in the purification are outlined in detail in section **2.8.5** of materials and methods in **chapter 2**, they included cell lysis, ammonium sulphate precipitation, acid precipitation and final purification on a Highload MonoQ® anion exchange column. The purified ACP was eluted from the column using a linear gradient of 0 – 500mM lithium chloride and fractions were collected during the elution and analysed for purity using SDS-PAGE. A coomassie stained SDS-PAGE gel of fractions collected from a typical purification run is shown in **Figure 4.2**. The fractions containing ACP were pooled together and the ACP was fully reduced by incubation with DTT (5mM final concentration) on ice for 30 minutes. The fully reduced ACP was dialysed using Spectrapore® (6000-8000 kDa cut off) membrane against MQ water and following dialyses aliquoted into 5.0 mg lots, lyophilised, snap frozen in liquid nitrogen and stored at –80°C until required. A typical purification would yield approximately 150 mgs of

Figure 4.2 SDS-PAGE analyses of the PorousQ ® anion exchange column fractions collected during the purification of wild type *E.coli* ACP.



Coomassie blue stained 18% SDSPAGE gels of the sequential fractions eluted during the 0-500mM LiCl gradient. Lane 1 on both gels: SDSVII molecular weight markers (Sigma) lane 2 on the top gel: the supernatant from the acid precipitation step of the purification. Lanes 3-17: sequential fractions collected during the gradient. ACP eluted from the column over four fractions (lanes 12-15) at approximately 300mM LiCl. These fractions containing ACP (arrowed) were pooled dialysed and used in acylation and crystallisation experiments.

> 95% pure ACP from 1.0 kg of cell paste. An aliquot of this purified material was checked by native–PAGE (**Figure 4.3a**) and by electrospray mass spectrometery (ESMS) (**Figure 4.3b**) and found to be ~90% holo -ACP (Mr 8850).

4.2.2 Synthesis of C4 and C8 acyl ACP Using Wild-Type *E.coli* holo ACP.

C4 and C8 acyl-ACPs were synthesised using the purified and reduced wild type *E.coli* ACP in chemical acylation reactions with C4 and C8 n-acylimidazoles. Reaction conditions are described in section **2.8.8** of materials and methods in **chapter 2.**

In each reaction a 5.0mg (500 nM) aliquot of ACP was reacted with 50mM of n-acyl-imidazole at room temperature for 60 minutes with continuous stirring. Following desalting of the reaction products through a PD10 desalting column an aliquot was removed and the extent of acylation confirmed by native-PAGE (**Figure 4.4a**) and ESMS (**Figure 4.4b**).

On native-PAGE gels containing urea, acylated-ACP migrates to a greater distance than apo and holo forms. The migration distance is determined by the chain length of the acyl ACP and the concentration of urea in the gel and the sample loading buffer (Post-Beittenmiler *et al.*, 1991). The purified ACP and the desalted reaction products from both the C4 and C8 reactions were separated on 18% PAGE + 0.5M urea resolving gels, 5% PAGE + 0.5M urea stacking gels. The purified ACP showed a major band (90%) at the expected position of holo-ACP and a minor band (10%) at the expected position of

Figure 4.3 Native-PAGE and electrospray mass spectrometric (ESMS) analyses of purified wild type *E.coli* ACP.

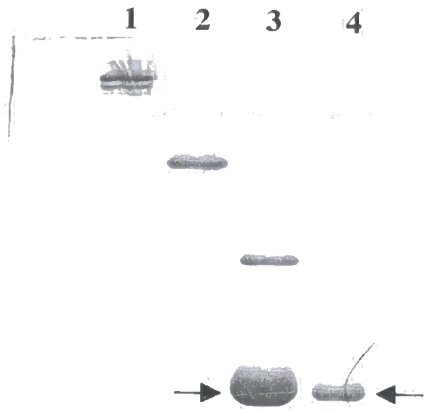


Figure 4.3a Coomassie blue stained 20% native –PAGE + 0.5M urea gel showing purified wild type *E.coli* ACP. Lane 1: carbonic anhydride standard (29 kDa), lane 2: α lactoglobulin standard (14 kDa), lane 3: 20μl of purified *E.coli* ACP and lane 4: 5μl of purified *E.coli* ACP. The major band (arrowed) is holo-ACP, the faint band immediately above it is apo-ACP, and the higher Mr band is dimmerised ACP.

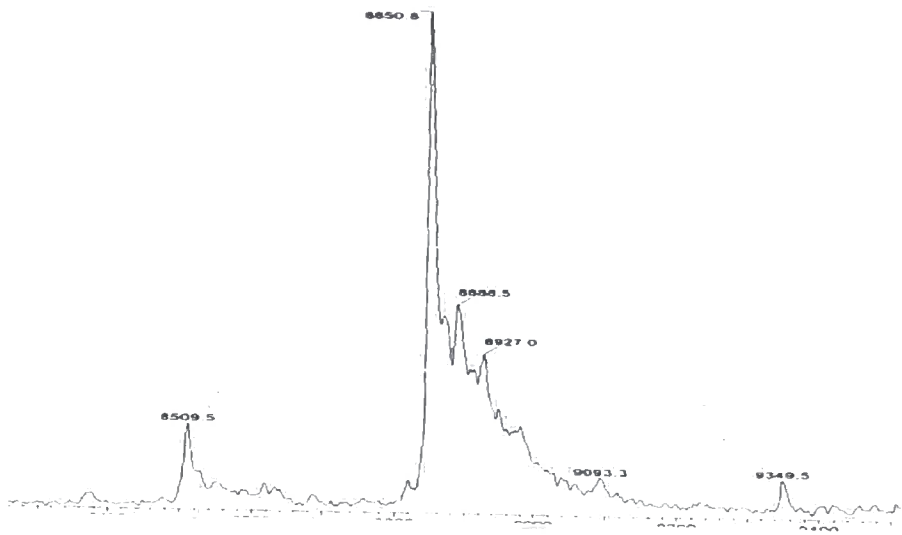
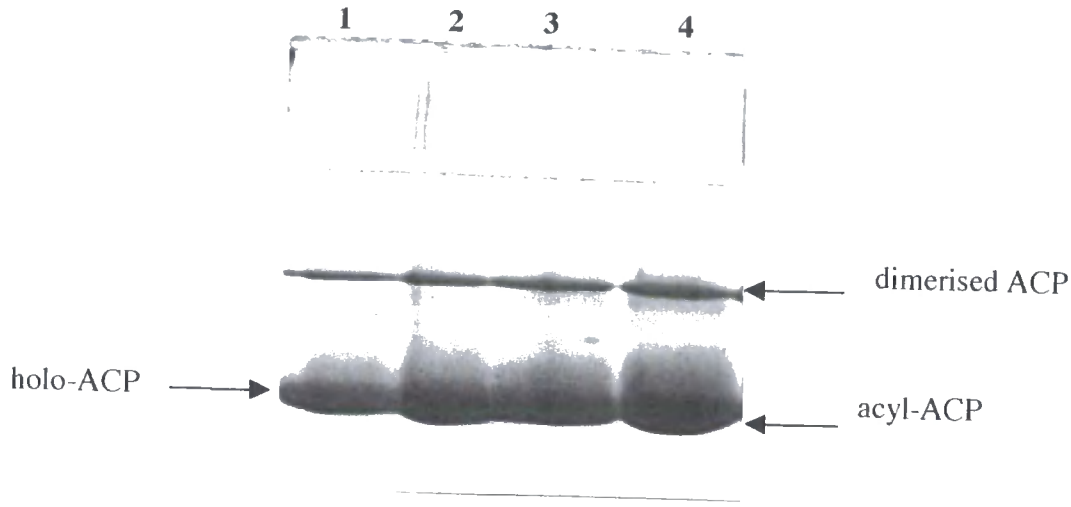


Figure 4.3b ESMS spectra of purified wild type *E.coli* ACP. The major peak at Mr 8850.8 represents holo-ACP and the minor peak at Mr 8509 represents apo-ACP. It was estimated from this data that the purified material was at least 90% holo-ACP.

Figure 4.4a. Native-PAGE analyses of acylated *E.coli* ACP reaction products following chemical synthesis with n-acylimidazoles.



Coomassie blue stained 18% native –PAGE + 0.5M urea gel showing *Ecoli* holo-ACP and acyl ACP reaction products. Lane 1 + 3: holo-ACP starting material, lane 2:C4-acyl-ACP reaction product and lane 4: C8-acyl-ACP reaction product. The acyl-ACP migrates further than the holo-ACP starting material in each case.

apo -ACP (**Figure 4.4a**). The results for the acyl-ACP reaction products showed that >90% of the ACP migrated away from the position of holo-ACP starting material to the position of the acyl-ACP product (**Figure 4.4a**). This indicated an almost quantitative derivatisation of *E.coli* wild type ACP in both reactions.Accurate mass measurements of purified *E.coli* holo-ACP and C4 and C8 acyl-ACPs were made by ESMS (**Figure 4.3b and 4.4b**). The ACP was diluted with 90% acetonitrile in 1.0% formic acid to give a final concentration of 5pmole /µl, 5µl (25 pmole) of this solution was injected into the mass spectrometer. A 5µl aliquot (50pmole) of horse myoglobin (Mr 16952) was used to calibrate the instrument. Mass spectrometry was carried out at the University of Cambridge Department of Chemistry (Phillip Lowden) and at Zeneca Biotechnology Section (Steve Rayner). The data for the purified *E.coli* ACP showed a single major peak at 8853 mass units (holo-ACP) and a smaller peak at 8512 mass units (apo-ACP). Although in mass spectrometry peak height for different molecular species is not a quantitative measurement, as apo and holo-ACP are chemically very similar, in a first approximation, signal intensity of these closely related species could be used to estimate the ratio of each. From the spectral data it is concluded that the purified wild type *E.coli* ACP was ~90% holo -ACP. This correlated well with the data obtained from the native-PAGE analyses. The spectra for the C4 and C8 reaction products showed a quantitative shift of the major holo peak of the starting material to 8921 mass units for the C4 product (**Figure 4.4b I**) and to 8976 mass units for the C8 product (**Figure 4.4b II**).

Figure 4.4b Electrospray mass spectrometry (ESMS) spectra of acylated *E.coli* ACP reaction products following chemical synthesis with n-acylimidazoles.

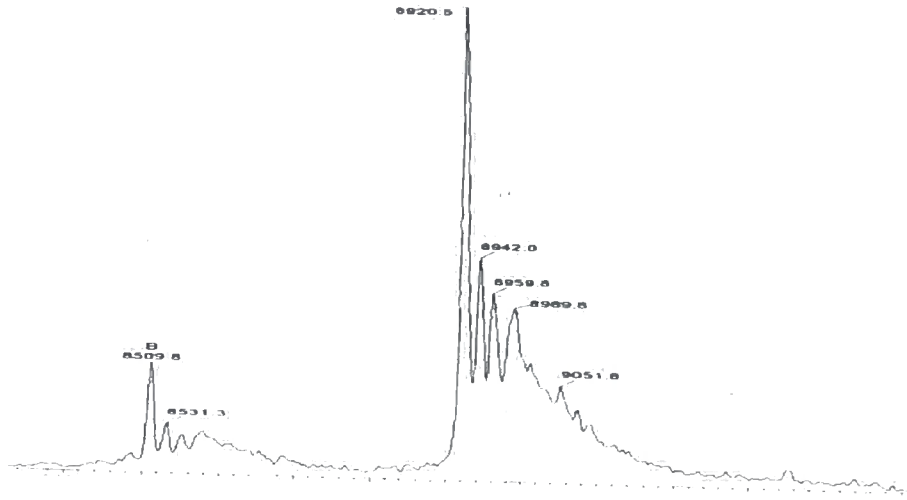


Figure 4.4b (I) ESMS spectra of *E.coli* C4 acyl-ACP. The major peak at Mr 8920.5 represents the C4-acyl-ACP reaction product. The minor peak at Mr 8509 represents the apo-ACP present in the starting material.

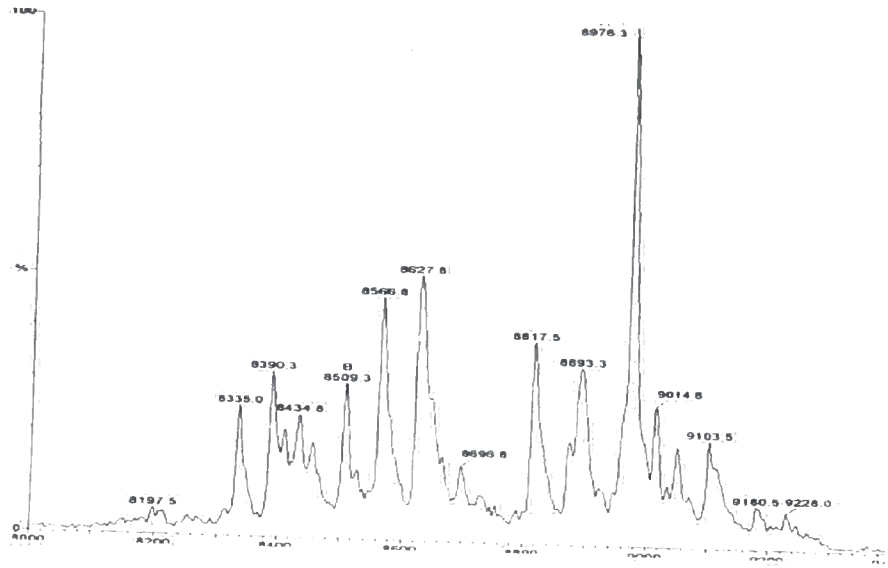


Figure 4.4b (II) ESMS spectra of *E.coli* C8 acyl-ACP. The major peak at Mr 8976.3 represents the C8 acyl-ACP reaction product. The other peaks in the spectra represent non-specific side reactions.

The C8 spectra showed a number of smaller contaminating peaks that made the spectra a little more difficult to interpret.

Together the native-PAGE and the ESMS analyses confirmed that the purified *E.coli* ACP was 90% holo protein and that under these reaction conditions with C4 and C8 n-acylimidazoles, a quantitative derivatisation of the starting material to the acyl-ACP product could be achieved.

4.2.3 Crystallisation Trials with Wild-Type *E.coli* acyl-ACPs.

Aliquots of *E.coli* holo-ACP and the C4 and C8 acyl-ACPs synthesised in the chemical derivatisation procedures were used in a preliminary crystallisation screen to attempt to establish conditions under which X-ray stable crystals of ACP could be produced. This crystallisation trial was set up by Dr Anna Roujeinikova at the Krebs Institute for Biomolecular Research at the University of Sheffield.

Only the acylated forms of the protein crystallised, the holo-ACP did not. The C4-acyl-ACP produced crystals which were stable in the X-ray beam and which gave diffraction data at 2Å resolution. No X-ray stable crystals were obtained for the C8 derivatised protein. The crystals were obtained by the hanging drop method in a solution of 8-12% PEG 20000, 30mM $ZnCl_2$ and 40mM sodium cocadylate pH6.0 at 17°C.

It was not possible using this preliminary diffraction data and molecular replacement techniques using the NMR structural data available for *E.coli* ACP (Holak *et al.*, 1988

and 1989 and Kim and Prestegard 1990), to determine the atomic structure of the protein. We therefore set out to obtain suitable heavy metal derivatives of the crystal in order that isomorphous replacement methods between the native crystal and the metal derivatised protein could be used to determine the complete crystal structure. For these studies either a selenomethionyl protein or a protein with modified amino acid residues suitable for heavy metal binding would be needed. A first requirement for this was to be able to use recombinant proteins. For this ACP from several plant and bacterial sources was cloned and the protein over-expressed and purified. Using the recombinant proteins three approaches were followed in an attempt to produce a heavy metal derivative of ACP. The first was the introduction of cysteine residues into the protein by site directed mutagenesis so that mercury derivatives could be made by soaking the native crystals with ethyl-mercury phosphate (EMP). The second was to substitute the natural methionines within the protein with seleno-methionine by over-expressing the recombinant ACP in the presence of seleno-methionine. The third was to introduce additional methionine residues into the protein using site directed mutagenesis and substitute these introduced residues with seleno methionine.

4.2.4 Overexpression and Purification of Recombinant ACP from Plants and Bacteria.

cDNA encoding for ACP from a number of plant and bacterial sources was obtained and over-expressed in pET bacterial protein expression vectors.

4.2.4 (a) *E.coli* ACP.

pET11d Plasmid DNA containing a 600 bp genomic DNA insert which included a 234 bp ORF coding for the mature amino acid sequence of *E.coli* ACP was obtained from Dr A.R.Stuitje (free University of Amsterdam).

4.2.4 (b) *Brassica napus* ACP.

pET 11d plasmid DNA containing a 200 bp ORF encoding the *Brassica napus* seed specific ACP amino acid sequence was obtained from Dr Tony Fawcett.

4.2.4 (c) Strawberry ACP.

pET 24a plasmid DNA containing a 355 bp ORF encoding Strawberry ACP amino acid sequence was obtained from Matthew Themis at the Horticultural Research Institute (HRI).

4.2.4 (d) *Mycobacterium* ACP.

The whole genome of the gram positive bacteria *Mycobacterium tuberculosis* has been sequenced and putative functional assignments have been made for its genes based on their homology to known sequences from other organisms (Cole *et al.*, 1998). Analyses of the genome showed that five of the FAS II genes (*fab*D, *acp*M, *kas*A, *kas*B and *acc*D)

involved in the biosynthesis of the very long chain mycolic acids found in this organism cluster together in the genome. The *acp*M gene was isolated in a PCR reaction using mycobacterial cosmid DNA MTCY427 (obtained from The Pasteur Institute, Paris) and oligonucleotide primers specific to the ACP sequence (MycACPox1(forward primer) 5'cgaattccatatggtgcctgtcactcagggaa 3' and MycACPox2 (reverse primer) 5'gcgcagatcttcacttggactcggcctcaag3') . The primers were designed to introduce a 5' NdeI site and a 3' BglII restriction site onto the ACP sequence to allow the PCR product to be ligated into the multiple cloning site of the pET 24a over-expression plasmid.

A 370bp product was obtained from the PCR reaction which was sub-cloned via the TOPO TA® (Invitrogen BV) cloning vector into the pET24a over-expression plasmid. Following transformation into XLlblue *E.coli* cells plasmid DNA was isolated and DNA sequenced. The DNA sequence obtained confirmed the pET clone contained an in frame sequence for over-expression of mycobacterium ACP.

4.2.5 Over-expression and Purification of Recombinant ACP's.

The DNA from all these over-expression vectors was transformed into transformation competent *E.coli* BL21 (DE3) *plys*S cells using the standard protocol and over-expression of recombinant ACP was induced with IPTG. The extent of over-expression of recombinant ACP's was analysed by SDS-PAGE (**Figure 4.5**) and in all cases a clear induced protein band was visible. Following induction the *E.coli* cells were harvested and

the recombinant ACP was selectively extracted from the cells following repeated cycles of freezing and thawing.

The recombinant ACP was purified from the freeze thaw extract on a 1.0 ml Porous Q® ion exchange column. The column was equilibrated in 5 ml of 10mM potassium phosphate buffer, pH 6.2 + 0.1% β mercaptoethanol and the ACP was eluted from the column using a 15 ml linear gradient, 0 - 500 mM lithium chloride in 10mM potassium phosphate buffer, pH 6.2 + 0.1% β mercaptoethanol

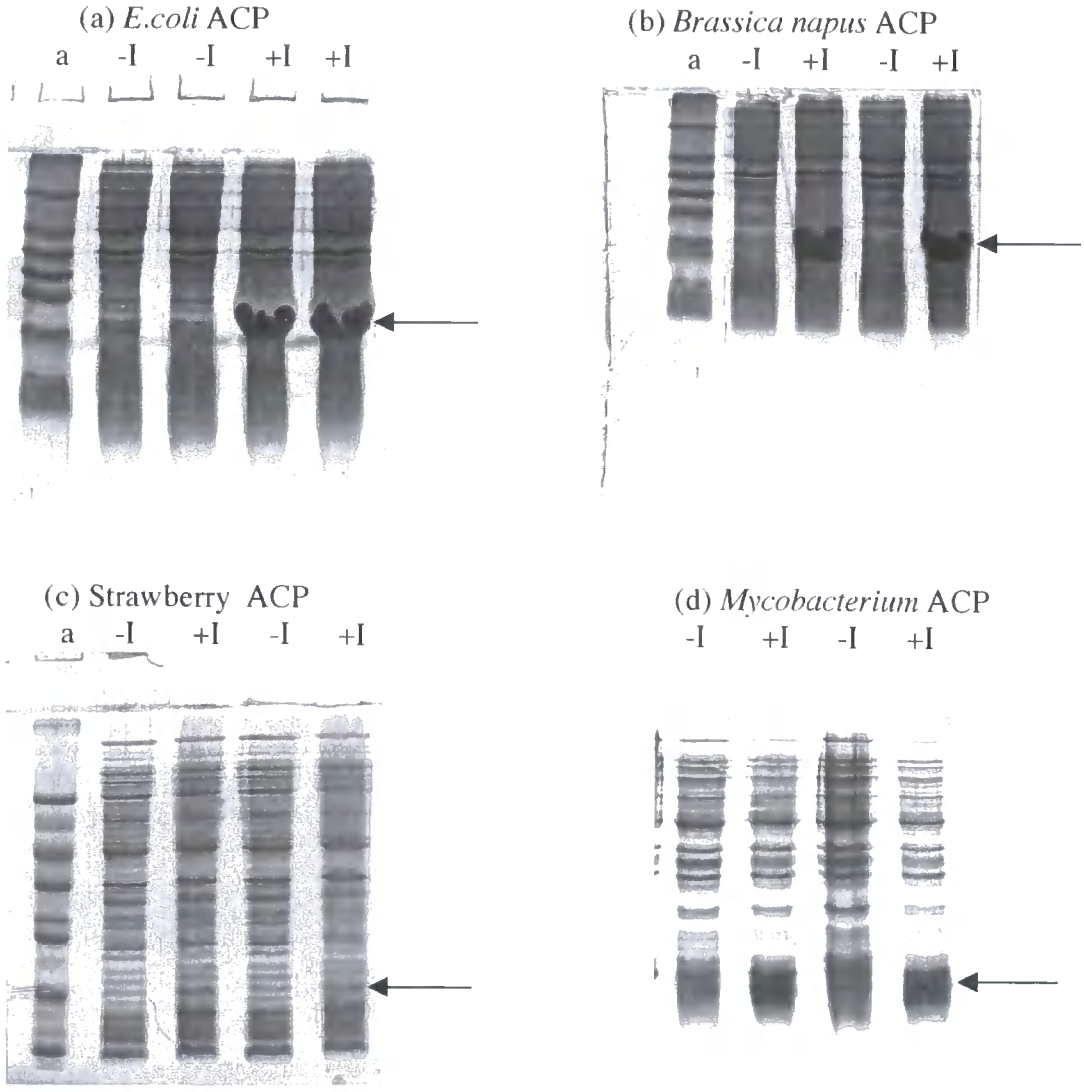
Fractions collected during the elution were analysed by SDS-PAGE (**Figure 4.6**) and fractions containing ACP were pooled together dialysed using Spectrapore® (6000-8000 kDa cut off) membrane against MQ water and following dialyses aliquoted into 5.0 mg aliquots, lyophilised, snap frozen in liquid nitrogen and stored at –80°C until required.

E.coli ACP eluted from the column with approximately 300mM LiCl, *Brassica napus* ACP with 250 mM LiCl , strawberry ACP with 300 mM LiCl and mycobacterium ACP 500 mM LiCl.

In the case of *E.coli, Brassica napus* and strawberry this freeze thaw extraction and single chromatographic purification step yielded > 95% pure ACP (**Figure 4.6**). The protein from *mycobacterium* however did not purify well using this procedure and was therefore not used in further acylation and crystallography trials.

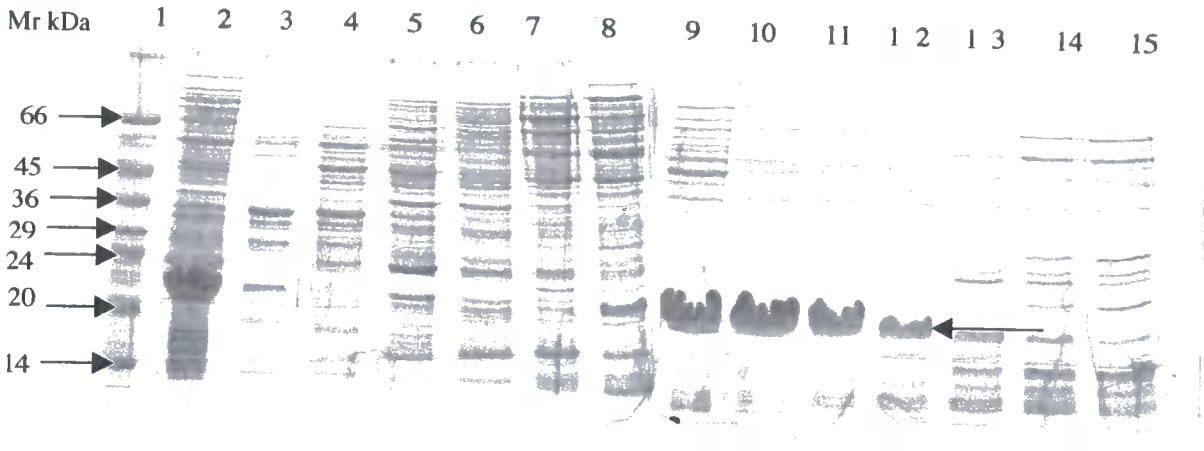
Native gel (**Figure 4.7**) and MALDItof mass spectrometry analyses (**Figure 4.8**) showed that recombinant ACP from *E.coli* and *Brassica napus* expressed in this bacterial system

Figure 4.5 SDS-PAGE analyses showing the over-expression of recombinant ACP from (a)*E.coli*, (b)*Brassica napus*, (c)*Fragaria* (strawberry) and (d)*Mycobacterium tuberculosis*.

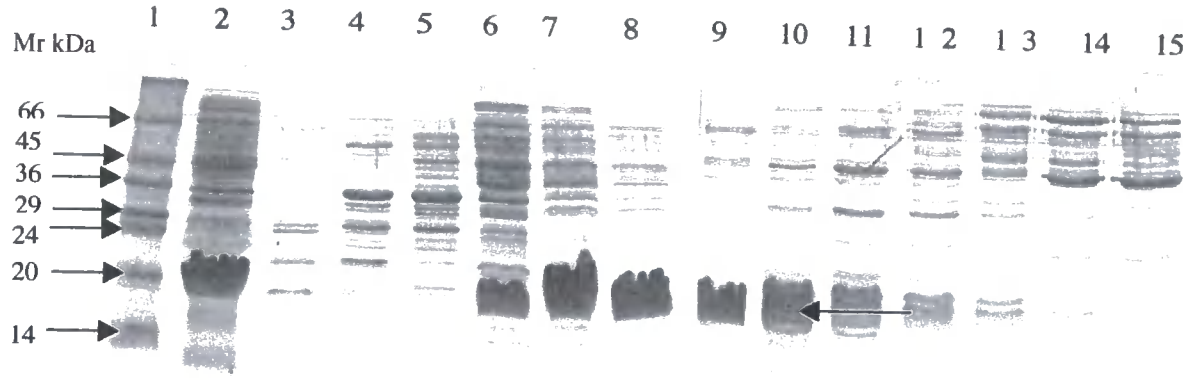


The position of the over-expressed ACP band is arrowed. (-I) is the un-induced protein extract, (+ I) is the protein extract following IPTG induction. Lane a:SDS-VII molecular weight markers (66,45,36,29,24,20+14 kDa). For both strawberry and *Mycobacterium* the ACP band became much more obvious following freeze-thaw extraction.

Figure 4.6 a + b SDS-PAGE analyses of the PorousQ ® anion exchange column fractions collected during the purification of recombinant ACPs from (a) *E.coli*, (b) *Brassica napus*.

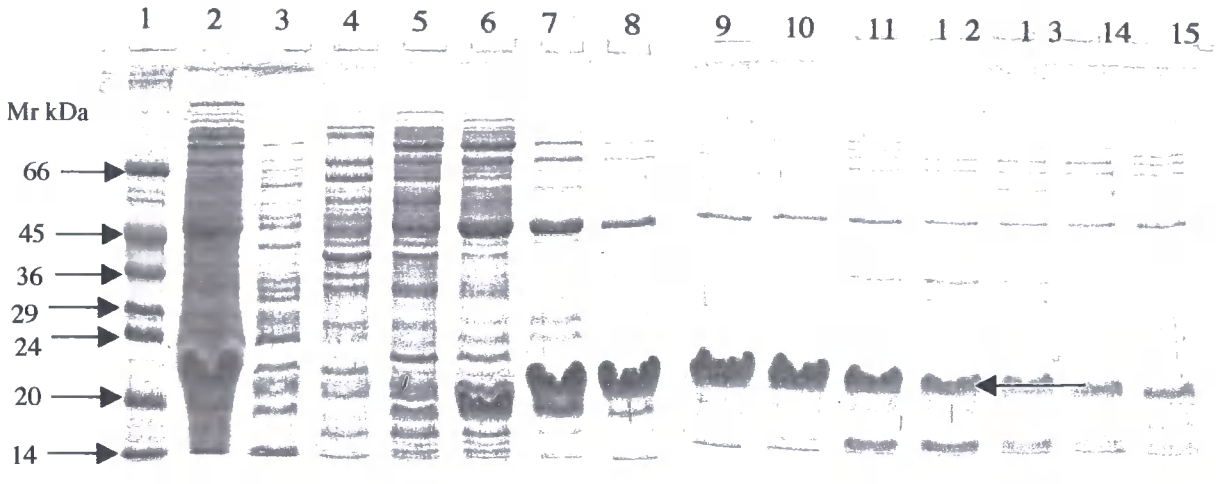


(a) Sequential fractions collected during the purification of *E.coli* ACP. Lane 1: SDSVII Mr markers, lane 2:freeze thaw cell free extract (column load), Lanes 3-15 sequential fractions eluted from the column during the gradient. Fractions (9-12) containing ACP (arrowed) eluted from the column with approximately 300mM LiCl.

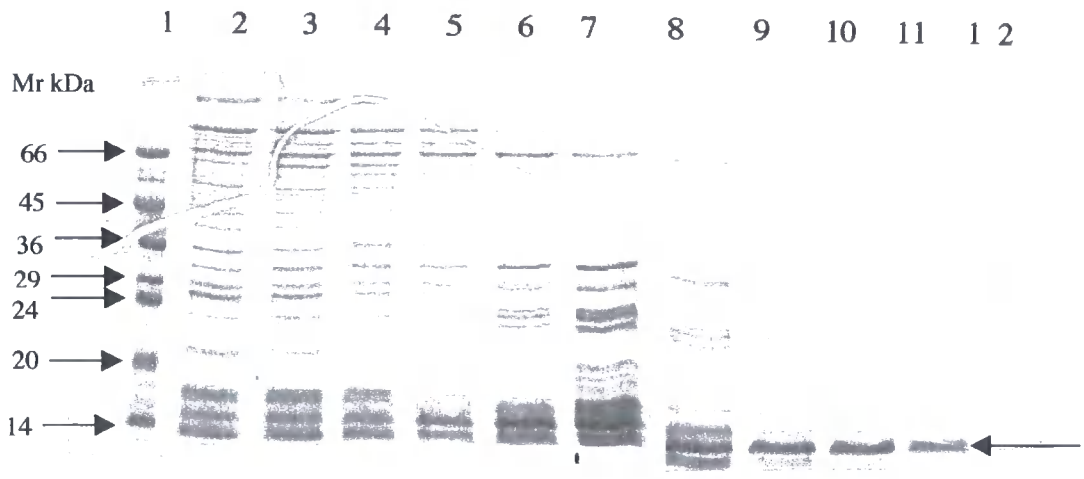


(b) Sequential fractions collected during the purification of *Brassica napus* ACP. Lane 1: SDSVII Mr markers, lane 2: freeze thaw cell free extract (column load), Lanes: 3-15 sequential fractions eluted from the column during the gradient. Fractions (6-11) containing ACP (arrowed) eluted from the column with approximately 250mM LiCl.

Figure 4.6 c + d SDS-PAGE analyses of the PorousQ ® anion exchange column fractions collected during the purification of recombinant ACPs from (c)*Fragaria* (strawberry) and (d)*Mycobacterium tuberculosis*.



(c) Sequential fractions collected during the purification of strawberry ACP. Lane 1: SDSVII Mr markers, lane 2: freeze thaw cell free extract (column load), Lanes: 3-15 sequential fractions eluted from the column during the gradient. Fractions (6-12) containing ACP (arrowed) eluted from the column with approximately 300mM LiCl.



(d) Sequential fractions collected during the purification of *Mycobacterium* ACP. Lane 1: SDSVII Mr markers, lane 2: freeze thaw cell free extract (column load), Lanes: 3-15 sequential fractions eluted from the column during the gradient. Fractions (7-11) containing ACP (arrowed) eluted from the column with approximately 500mM LiCl.

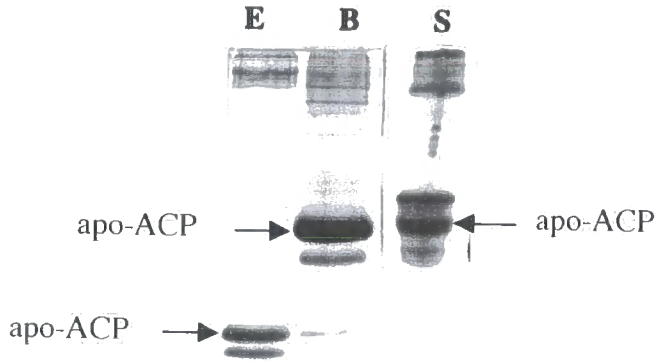
was >90% in the apo form and did not contain the 5'phosphopantetheine post-translational modification. The protein from strawberry showed two bands on native gel analyses, the lower most intense band corresponded to apo-ACP and the higher weaker band dimmerised ACP (**Figure 4.7**).

Enzymatic Synthesis of Recombinant Acyl-ACPs.

4.2.6 Over-expression and Purification of Holo-ACP Synthetase (HAS) from *E.coli.*

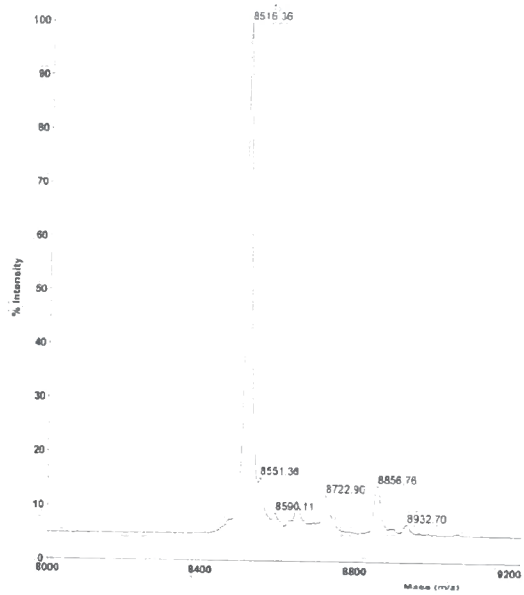
Holo-ACP synthetase (HAS) catalyses the post-translational modification of apo-ACP to holo-ACP in a Mg^{+} dependant reaction where the 4'phosphopantetheine group from coenzyme A (CoA) is transferred to an essential serine residue within the ACP. In studies on the actinorhodin polyketide synthase enzymes Carreras *et al.*, (1997) demonstrated that *E.coli* HAS could not only carry out this phosphopantethenylation reaction from apo to holo-ACP, but could also catalyse the synthesis of acyl-ACPs directly, utilising acyl-CoAs (acetyl, butyryl, benzoyl, phenyl-acetyl and malonyl). This enzymatic acylation reaction was more specific than chemical reactions with no side reactions occurring, could be carried out directly on apo-ACP and produced fully functional acyl-ACPs. We therefore decided to obtain recombinant *E.coli* HAS and attempt this enzymatic synthesis of acyl-ACP's for all of our further studies with recombinant ACPs.

Figure 4.7 Native-PAGE analyses of purified recombinant ACPs from *E.coli* (**E**) *Brassica napus* (**B**) and strawberry (**S**).

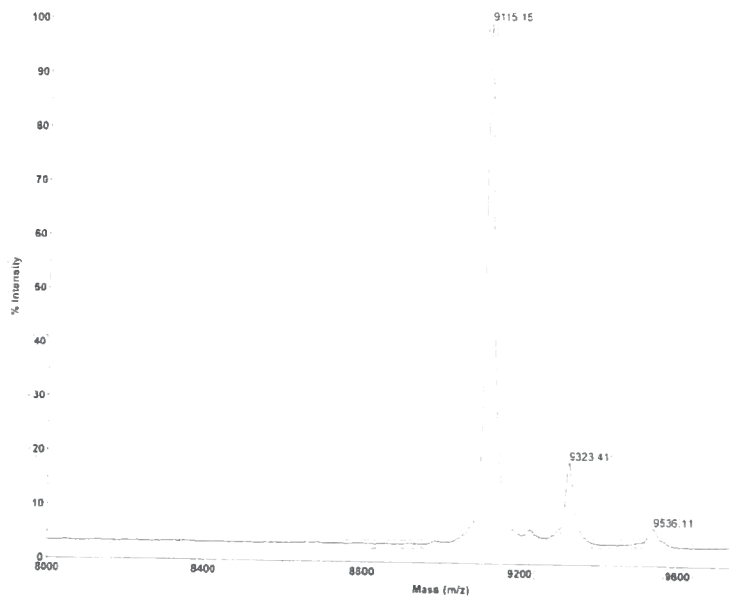


Coomassie blue stained 18% native –PAGE + 0.5M urea gel showing purified recombinant *E.coli* (**E**), *Brassica napus* (**B**) and strawberry (**S**) ACP. The major band visible in the *E.coli* and *Brassica napus* lanes (arrowed) corresponds to the position of apo-ACP, the faint band visible below it is holo-ACP. In the strawberry lane the lower most intense band (arrowed) is apo-ACP and the band above is believed to be either a dimer or an aduct of ACP.

Figure 4.8 MALDItof mass spectrometry analyses of purified recombinant ACP from *E.coli* and *Brassica napus*.



MALDItof-MS spectra of purified *E.coli* apo-ACP. The major peak at 8516 m/z represents *E.coli* apo-ACP. The measured mass is within 0.07% (6Da) of the expected mass of 8510. A minor holo-ACP peak can be seen at 8856 m/z. The spectra shows that >90% of the purified ACP is in the apo form and is not post-translationally phospho-pantethenylated.



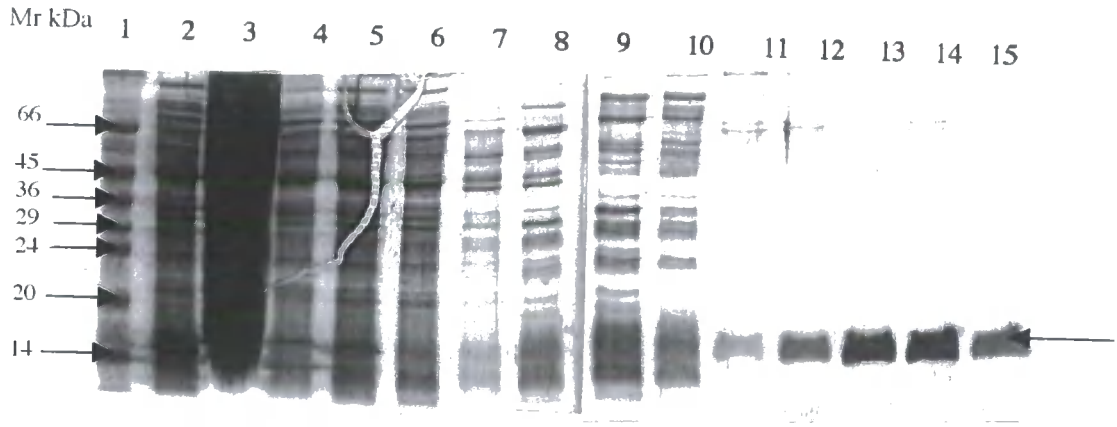
MALDItof-MS spectra of purified *Brassica napus* apo-ACP. The major peak at 9115 m/z represents *Brassica napus* apo-ACP. The measured mass is within 0.07% (7 Da) of the expected mass of 9108. No holo-ACP was detected at 9464m/z.

pET24a plasmid DNA containing the 420 bp ORF for *E.coli* holo-ACP synthetase was obtained from Simon Doig (this laboratory). Following transformation and induction in *E.coli* BL21cells the induced protein was selectively released from the harvested cells by freeze thaw extraction. The freeze thaw supernatant was applied to a pre-equilibrated (50mM Tris:HCl pH 8.0 buffer+10mM Mg Cl_2 , 1mM DTT) Porous HS® cation exchange column and purified HAS was eluted from the column with a 15 ml 0 – 1.0 M NaCl linear gradient in equilibration buffer. Fractions were collected during the gradient and analysed for purity on SDS-PAGE (**Figure 4.9**). HAS eluted from the column with approximately 700 mM NaCl. A typical purification yielded approximately 20 mgs of > 95% pure HAS which was snap frozen in liquid nitrogen and stored in aliquots at –80°C until required.

4.2.7 Enzymatic Synthesis of Recombinant *E.coli*, *Brassica napus*, and Strawberry Holo and Butyryl ACP's.

The purified recombinant apo-ACPs from *E.coli*, *Brassica napus*, and Strawberry were used in enzymatic pantethenylation and acylation reactions catalysed by purified *E.coli* HAS. A typical synthesis was carried out in a 2.0 ml reaction containing 350µmole apo-ACP, 700µmole CoA or butyryl CoA, and 3.5µmole HAS at pH7.3 in 50mM potassium phosphate buffer in a stirred reacti-vial at room temperature for 60 minutes. The reaction

Figure 4.9 SDS-PAGE analyses of the PorousHS™ cation exchange column fractions collected during the purification of recombinant *E.coli* holo-ACP synthetase (HAS).

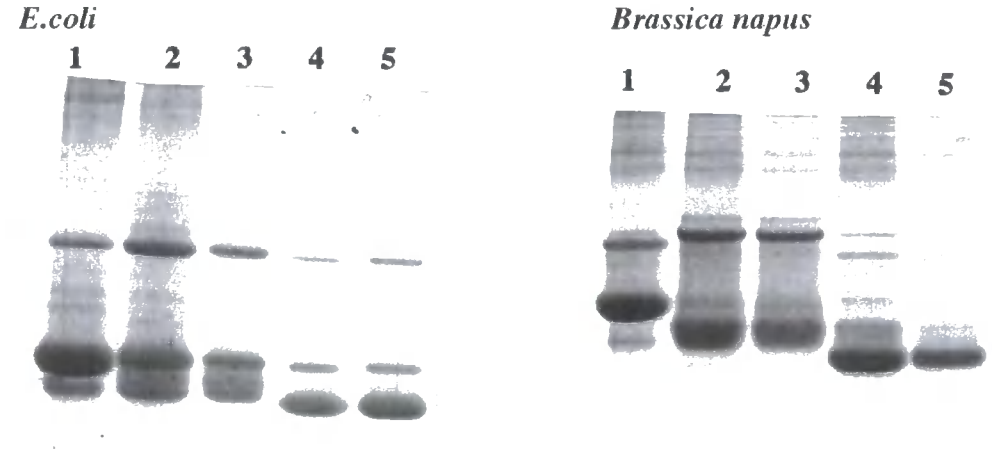


Coomassie blue stained 12% SDSPAGE gel of the sequential fractions eluted during the 0-1.0M NaCl gradient. Lane 1:SDSVII molecular weight markers (Sigma), lane 2:freeze thaw supernatant following centrifugation at 40000g (column load), lane 3: pellet from the 40000g spin, lanes 4 –15: sequential fractions eluted from the column during the gradient. HAS eluted from the column over five fractions (lanes 11-15) at approximately 700 mM NaCl. These fractions containing HAS (arrowed) were pooled aliquoted, snap frozen in liquid nitrogen and stored at –80°C until required for acylation reactions.

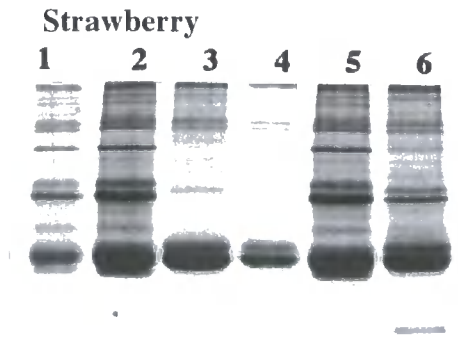
products were desalted through a PD10 desalting column before analysis by native-PAGE and MALDItof mass spectrometry.

The *E.coli*, and *Brassica napus* reactions showed quantitative derivatisation to both holo and butyryl-ACP as measured by native-PAGE (**Figure 4.10**) and by MALDI-tof mass spectrometry analyses (**Figure 4.11 + 4.12**). The strawberry protein did not pantethenylate or acylate successfully. Analyses by native-PAGE (**Figure 4.10**) showed extensive dimerisation of the strawberry ACP under these reaction conditions and it was assumed that this prevented the reaction from proceeding.

Figure 4.10 Native-PAGE analyses of recombinant ACPs from *E.coli*, *Brassica napus* and strawberry pantethenylated and acylated in enzymatic reactions with HAS.

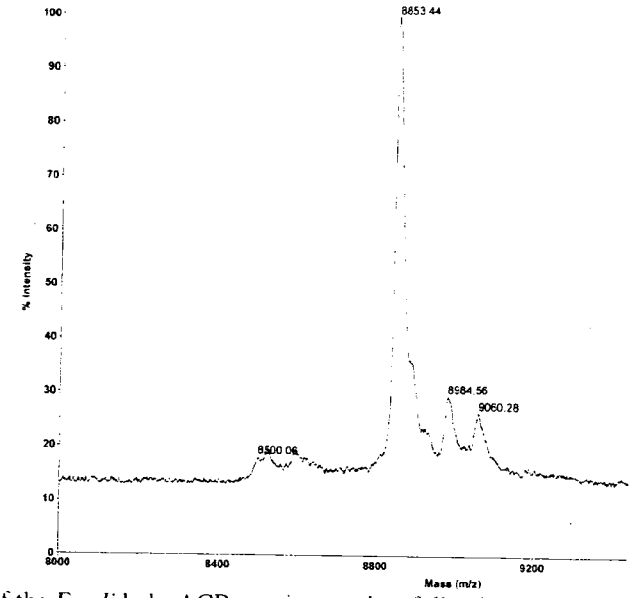


Lane 1:apo-ACP starting material, lane 2 + 3: holo-ACP product from a reaction containing Coenzyme A, lane 4 + 5: butyryl-ACP product from a reaction containing butyryl-CoA. For both *E.coli* and *B.napus* the holo and acyl-ACP products migrate further in the gel than the apo starting material. Based on the amount of starting material remaining in each case the reactions yielded approximately 90% product.

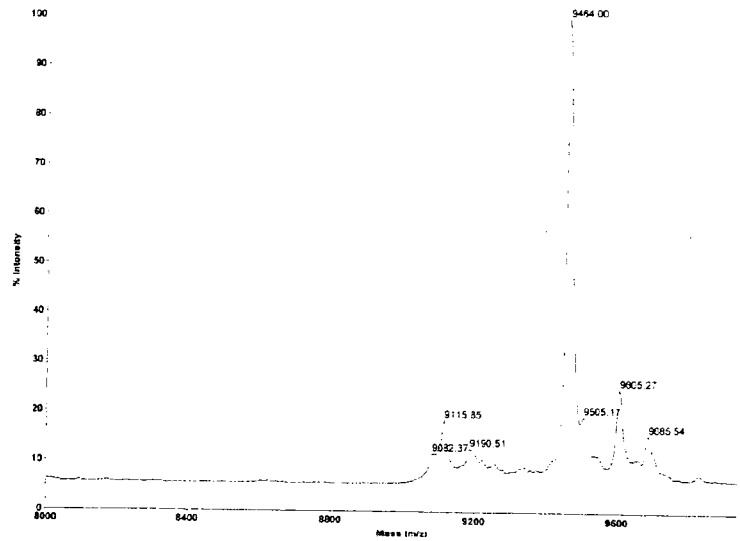


Lane 1:apo-ACP starting material, lane 2 + 3:holo-ACP product from reactions containing Coenzyme A, lane 4: apo-ACP starting material, lane 5 + 6:butyryl-ACP product from a reaction containing butyryl-CoA. Due to extensive dimmerisation the results of the reactions could not be determined.

Figure 4.11 MALDItof mass spectrometry analyses of recombinant *E.coli* and *Brassica napus* ACP reaction products following enzymatic phosphopantethenylation with holo-ACP synthetase (HAS) and Coenzyme A.

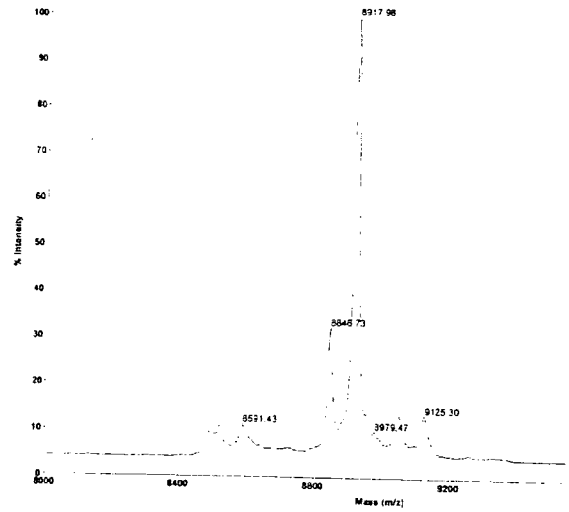


MALDItof-MS spectra of the *E.coli* holo-ACP reaction product following enzymatic pantethenylation with HAS. The major peak in the spectra at 8853 m/z represents holo-ACP and the spectra shows a quantitative conversion from the apo form at 8516 m/z shown in figure 4.8. The measured mass is within 0.03% (3Da) of the expected mass of 8850.

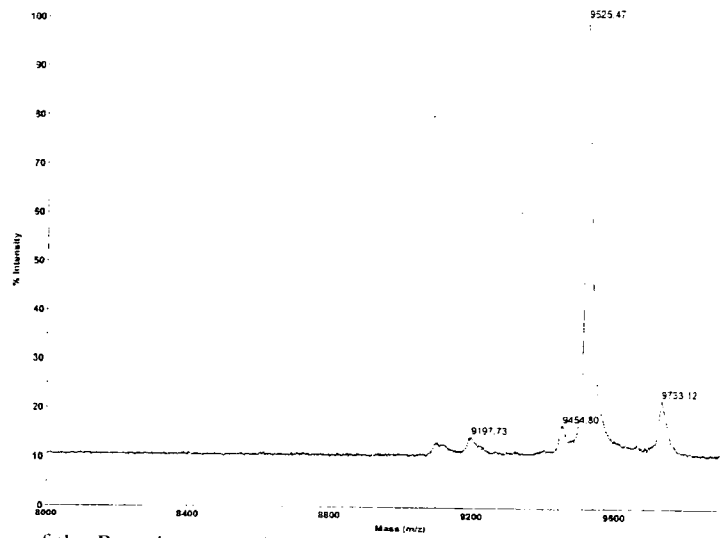


MALDItof-MS spectra of the *Brassica napus* holo-ACP reaction product following enzymatic pantethenylation with HAS. The major peak in the spectra at 9464 m/z represents holo-ACP and the spectra shows a >90% conversion from the apo form at 9115 m/z shown in figure 4.8.

Figure 4.12 MALDItof mass spectrometry analyses of recombinant *E.coli* and *Brassica napus* ACP reaction products following enzymatic acylation with holo-ACP synthetase (HAS) and butyryl-CoA.



MALDItof-MS spectra of the *E.coli* butyryl-ACP reaction product following enzymatic acylation catalysed by HAS. The major peak in the spectra at 8918 m/z represents *E.coli* butyryl-ACP and the spectra shows a >90% conversion from the apo form at 8516 m/z shown in figure 4.8 to the butyryl form. The measured mass for butyryl-ACP is within 0.03% (3Da) of the expected mass of 8921. The peak at 8548 m/z represents *E.coli* holo-ACP estimated to be <10% of the reaction product.



MALDItof-MS spectra of the *Brassica napus* butyryl-ACP reaction product following enzymatic acylation catalysed by HAS. The major peak in the spectra at 9525 m/z represents B*rassica napus* butyryl-ACP and the spectra shows quantitative conversion from the apo form at 9115 m/z shown in figure 4.8 to the butyryl form. The measured mass for butyryl-ACP is within 0.07% (7 Da) of the expected mass of 9532.

4.2.8 Crystallisation Trails with Recombinant ACP's.

Aliquots of all of the recombinant *E.coli* and *Brassica napus* holo and butyryl-ACP's synthesised in the enzymatic derivatisation procedures described above were put down to crystallography trials using the same conditions as were successfully used for the wild type ACP. Crystals were obtained for only the acylated forms of both the *E.coli* and *Brassica napus* ACP and only those for the *E.coli* butyryl-ACP were stable in the X-ray beam and gave diffraction data.

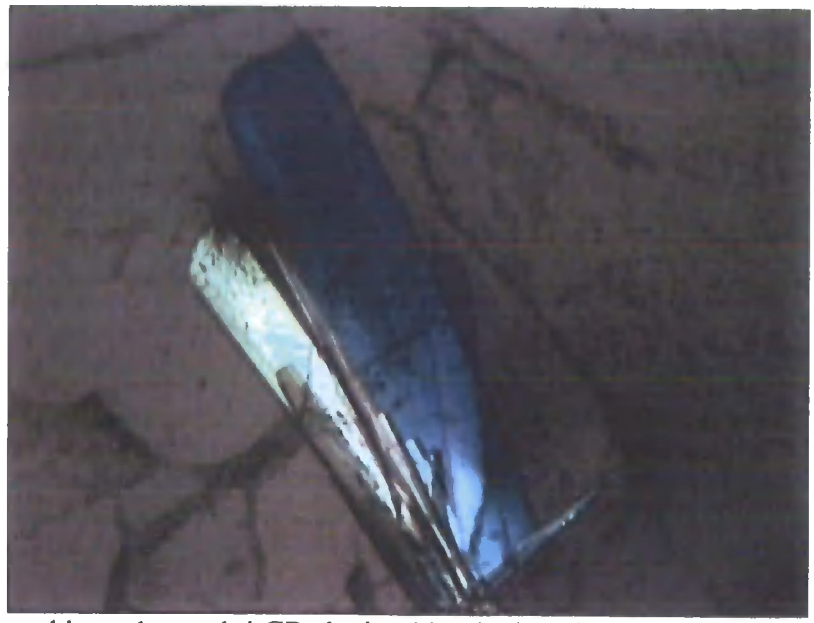
These crystals were smaller in size than those obtained for the wild type *E.coli* butyryl-ACP but were of the same form.

This result showed that the recombinant *E.coli* ACP would crystallise under the same conditions as the wild type protein, and resulted in similar crystals and yielded comparable diffraction data (**Figure 4.13**). This provided a mechanism to generate heavy metal derivatives of *E.coli* ACP to aid in the resolution of the structure.

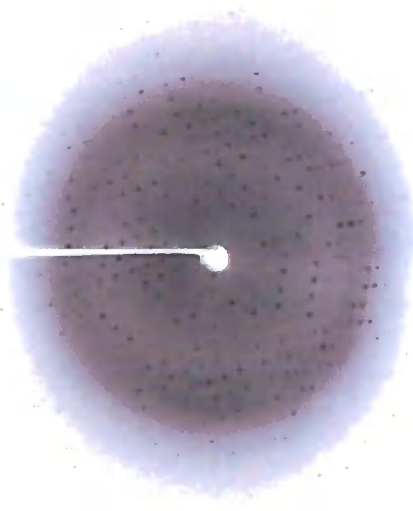
4.2.9 Site Directed Mutagenesis of Recombinant *E.coli* ACP.

Site directed cysteine and methionine mutations were introduced into recombinant *E.coli* ACP as sites suitable for the introduction of mercury and selenium derivatives to aid in the phase resolution of the diffraction data and the resolution of the crystal structure. Structural information obtained from the diffraction data from the recombinant crystal was used together NMR structural information and primary sequence alignment data to

Figure 4.13 Crystals and diffraction data of *E.coli* recombinant butyryl-ACP



Crystal of recombinant butyryl-ACP obtained by the hanging drop vapour diffusion method in a solution of 8-12% PEG20000, 30mM $ZnCl_2$ and 40mM sodium cocadylate pH6.0 at 17°C. The crystals belong to space group $P_2 1_2 1_2 1$ with cell dimensions a=27, b=42, c=65 Å.



Diffraction data collected from a cryo-cooled hexagonal crystal of *E.coli* butyryl-ACP collected using a CCD detector on the ESRF beamline BM30. Resolution at the edge of the detector is 2.0 Å.

determine suitable positions within the amino acid sequence of the protein for the introduction of mutations.

4.2.10 Introduction of Cysteine Residues to Generate Mercury Derivatives of *E.coli* ACP Crystals.

Approximately 20 ng of pET11d plasmid DNA containing a DNA insert coding for *E.coli* ACP was used together with complimentary mutagenesis primers (**Table 4.1**) in QuickChange®(Stratagene) mutagenesis reactions to generate recombinant DNA containing the cysteine mutations, T23C, A26C, S27C, and D51C.

Following removal of the parental plasmid DNA by digestion with the restriction enzyme *Dpn*I the mutated DNA produced in the reactions was transformed into super-competent Xl1blue *E.coli* cells and plated onto LB amp plates and grown overnight at 37°C. Two individual colonies were taken from each of the plates and grown overnight in 5 ml liquid cultures for plasmid production. Plasmid DNA was isolated and DNA sequenced using the T7 promoter sequencing primer to confirm the presence of the mutation. Sequence data was obtained confirming the presence of the cysteine mutation for all four of the residues attempted (T23C, A26C, S27C, and D51C).

Table 4.1 Nucleotide sequence of the oligonucleotide primers used for the cysteine mutagenesis of *E.coli* ACP. Nucleotides highlighted indicate the position of the mutation.

Primer name	Primer function	Sequence
ACPT23CF	T23C forward primer	5'gttaagcaggaagaagtt**tgc**aacaatgcttcctttcg 3'
ACPT23CR	T23C reverse primer	5'cgaaagaagcattgtt**gca**aacttcttcctgcttaac 3"
ACPA26CF	A26C forward primer	5'gaagttaccaacaat**tgt**tctttcgttgaag3'
ACPA26CR	A26C reverse primer	5'cttcaacgaaaga**aca**attgttggtaacttc3'
ACPS27CF	S27C forward primer	5'gaagttaccaacaat**gc**ttgcttcgttgaag3'
ACPS27CR	S27C reverse primer	5'cttcaacgaa**gca**agcattgttggtaacttc3'
ACPD51CF	D51C forward primer	5'gctctggaagaagagttt**tgt**actgagattccggacg3'
ACPD51CR	D51C reverse primer	5'cgtccggaatctcagt**aca**aaactcttcttccagagc3'

Following sequence confirmation an aliquot of the DNA containing the mutation was transformed into BL21 (DE3) *plys*S *E.coli* cells and induction freeze thaw extraction and purification was carried out using the same conditions as were used for recombinant *E.coli* ACP without the mutations. A similar level of induction was achieved for all four mutations as that typically achieved for the recombinant protein which did not contain mutations.

The mutated ACPs eluted from the Porous Q™ ion exchange column in the same position (300mM LiCl) and with a comparable purity (>95%) as the recombinant protein without the mutations. Approximately 15 mgs of each mutant ACP was purified from a 800ml culture. A sample of each mutant was analysed by MALDItof-MS and the mass

measured for the ACP peak in each spectra confirmed the presence of mutation in each case. (**Table 4.2**).

Table 4.2 MALDItof analyses of cysteine mutants of *E.coli* ACP.

***E.coli* ACP**	**Expected mass (m/z)**	**Measured mass (m/z)**
Apo ACP	8510	8510
T23C	8512 (+2)	8513
A26C	8542 (+32)	8540
S27C	8526 (+16)	8525
D51C	8498 (-12)	8498

0.5µl of purified ACP was loaded together with 0.5µl sinapinic acid matirx solution directly onto a MALDI target plate. Spectra were calibrated using a calibration mixture containing bovine insulin (m/z 5734), E.coli thioredoxin (m/z 11674) and horse myoglobin (16952).

4.2.11 Enzymatic Acylation of Cysteine Mutants of *E.coli* ACP and Crystallisation Trials.

Purified aliquots of the four *E.coli* ACP cysteine mutants were used in enzymatic acylation reactions with purified *E.coli* HAS and butyryl-CoA using the identical reactions conditions outlined for recombinant ACPs. Following synthesis the acyl-ACP reaction products were desalted through a PD10 desalting column and analysed by native-PAGE.

The gel profiles showed a more complex banding pattern than was normally seen for recombinant *E.coli* ACP. This was likely because of dimmerisation of the ACPs at the introduced cysteine residues.

However for all of the mutations a band shift was observed between the position of the apo-ACP starting material and the butyryl-ACP product indicating that the acylation of the apo-ACP had taken place.

A range of crystallisation trials were set up using the exact conditions which had previously produced diffraction stable crystals of both the wild type and the recombinant *E.coli* ACP, but despite several attempts we were unable to obtain crystals with any of the butyryl-ACP cysteine mutants. This resulted in the need to try to introduce a heavy metal site into the protein by incorporation of seleno-methionine into the position of the naturally occurring methionines or alternatively by first introducing extra methionines by mutagenesis and then incorporating seleno-methionine into these positions.

4.2.12 Seleno-methionine Incorporation Into Recombinant *E.coli* ACP.

E.coli ACP contains two methionine residues within its amino acid sequence, one at the N-terminus of the protein and one at amino acid position 45. The N-terminal methionine of the recombinant form is almost totally removed (80-90% as estimated by sequencing of the native protein) during synthesis, leaving a single methionine residue (met 45) as a site for heavy metal incorporation. The incorporation of selenium into this methionine

site was achieved by transformation of plasmid DNA containing *E.coli* ACP into a BL21 (DE3) metC$^-$ *E.coli* strain and growth in minimal media supplemented with seleno-methionine in place of methionine.

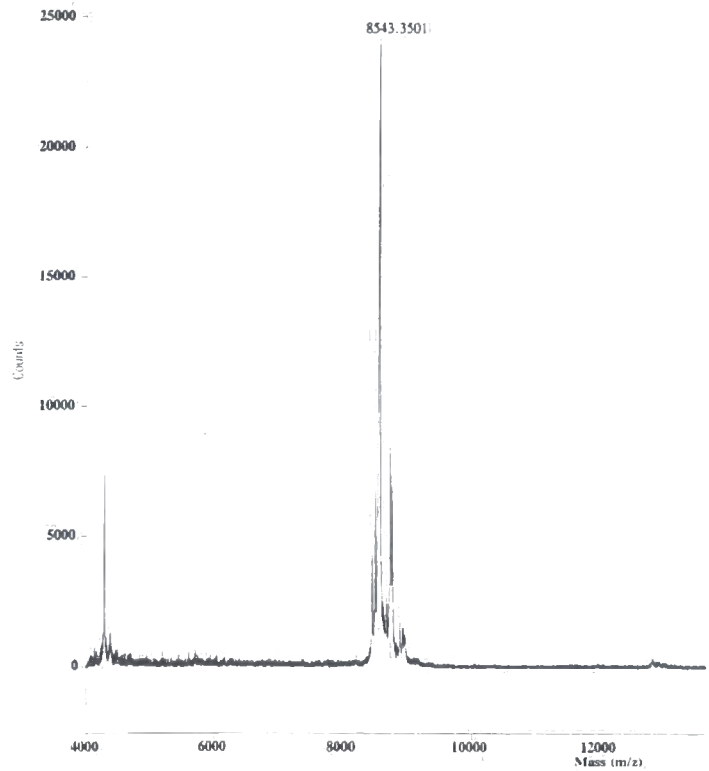
pET11d plasmid DNA containing a DNA insert coding for *E.coli* ACP was transformed into a BL21 (DE3) metC$^-$ *Ecoli* strain developed by Dr A.R.Stuitje (Free University of Amsterdam) using the standard transformation protocol. The transformants were grown in 800 ml liquid cultures of M9 minimal medium containing all amino acids except methionine, which was replaced in the medium with seleno-methionine. Growth of the transformed bacteria in this minimal medium was much slower than in LB medium, it took 9 hours growth at 37°C on an orbital shaker (150 rpm) to reach an OD 550 nM of 0.6. At this point recombinant protein synthesis was induced with IPTG and growth was continued for a further 6 hours. The cells were harvested from the growth medium and freeze thaw extraction and purification of the seleno-methionine incorporated ACP was carried out using the same conditions as were used for recombinant *E.coli* ACP. Following purification the column fractions containing ACP were pooled together, dialysed using Spectrapore® (6000-8000 kDa cut off) membrane against MQ water and following dialyses aliquoted into 5.0 mg lots, lyophilised, snap frozen in liquid nitrogen and stored at –80°C until required. A small aliquot of the dialysed material was analysed by MALDItof mass spectrometry and confirmed to be the correct mass (8557 Da) for *E.coli* apo ACP containing a single seleno-methionine incorporation (**Figure 4.14**).

Approximately 12 mgs of > 95% pure seleno-methioine incorporated ACP was purified from a 800ml culture.

4.2.13 Enzymatic Acylation of Seleno-methionine *E.coli* ACP and Crystallisation Trials.

Purified aliquots of the Seleno-methionine *E.coli* ACP were used in enzymatic acylation reactions with purified *E.coli* HAS and butyryl-CoA using the identical reaction conditions outlined for recombinant ACPs. Following synthesis the acyl- ACP reaction products were desalted through a PD10 desalting column and analysed by native-PAGE and MALDItof mass spectrometry. The gel profile (**Figure 4.15a**) showed an almost quantitative conversion from apo-ACP to butyryl-ACP and the MALDItof spectra (**Figure 4.15b**) showed a single peak at mass (8972 Da) the correct mass unit for *E.coli* butyryl-ACP containing single selenium incorporation. This butyryl-selenomethionyl *E.coli* ACP was put down to crystallography trials using exactly the same conditions as were used successfully for recombinant *E.coli* ACP. Small reproducible crystals were obtained after one month, which gave diffraction data at 2 Å resolution but this data from a single heavy metal site was still insufficient to resolve the phasing problems and to interpret coordinates and successfully determine the structure. A requirement for additional heavy metal sites was identified and a strategy for the introduction of methionine residues by site directed mutagenesis were planned.

Figure 4.14 MALDItof mass spectrometry analyses of purified recombinant *E.coli* selenomethionyl apo-ACP.



MALDItof-MS spectra of purified recombinant *E.coli* selenomethionyl apo-ACP. The major peak at 8543 m/z represents *E.coli* apo-ACP containing a single seleno-methionine residue. The measured mass is within 0.16% of the expected mass of 8557 Da.

Figure 4.15 Native-PAGE and MALDItof mass spectrometry analyses of purified seleno-methionyl *E.coli* holo and butyryl-ACP.

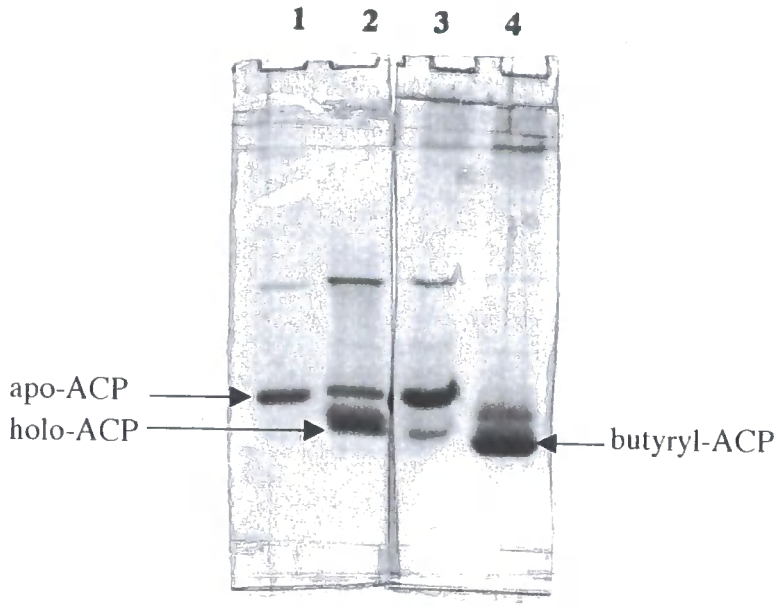
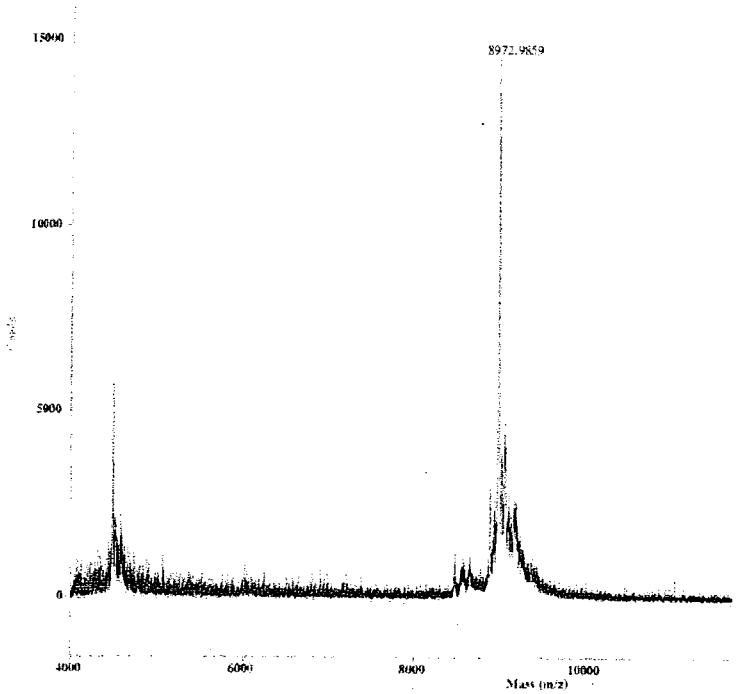


Figure 4.15a Coomassie blue stained 18% native –PAGE + 0.5M urea gel showing purified recombinant seleno-methionyl *E.coli* ACP and acylation reaction products. Lane 1:*E.coli* apo-ACP, lane 2: *E.coli* holo-ACP, lane 3: *E.coli* seleno-methionyl apo-ACP and lane 4: enzymatically synthesised butyryl seleno-methionyl *E.coli* ACP.

Figure 4.15b MALDItof mass spectrometry analyses of the recombinant *E.coli* selenomethionyl butyryl-ACP reaction product following enzymatic acylation with holo-ACP synthetase (HAS) and butyryl-CoA.



MALDItof-MS of the *E.coli* selenomethionyl butyryl-ACP reaction product following enzymatic acylation with holo-ACP synthetase (HAS). The major peak in the spectra at 8972 m/z represents *E.coli* selenomethionyl butyryl-ACP. The measured mass is within 0.04% (4 Da) of the expected mass of 8968.

4.2.14 Introduction of Methionine Residues into *E.coli* ACP by Site Directed Mutagenesis.

Using the diffraction data obtained from both the recombinant *E.coli* ACP, the selenomethionyl *E.coli* ACP crystals and the NMR structural data available for *E.coli* ACP suitable isoleucine or similar residues were identified at appropriate positions within the protein for conversion to methionine residues. These positions were identified as I12, T40, I55, I63 and I70. The mutations were introduced into recombinant *E.coli* ACP pET11d plasmid DNA using the QuickChange®(Stratagene) mutagenesis kit and complimentary mutagenesis primers (**Table 4.3**) using the same reaction conditions as outlined for the cysteine mutagenesis.

Following the mutagenesis reactions the plasmid DNA was isolated and sequenced using the same procedures outlined for the cysteine mutagenesis. The sequence data obtained confirmed that all of the mutations had been correctly introduced and that five new methionine mutants of *E.coli* ACP proteins had been obtained.

Table 4.3 Nucleotide sequence of the oligonucleotide primers used for the methionine mutagenesis of *E.coli* ACP. Nucleotides highlighted (bold) indicate the position of the mutation.

Primer name	Primer function	Sequence
ACPI12MF	I12M forward primer	5'gttaagaaaatt**atg**ggcgaacagctgggc3'
ACPI12MR	I12M reverse primer	5'gcccagctgttcgcc**cat**aattttcttaac3"
ACPT40MF	T40M forward primer	5'gattctcttgac**atg**gttgagctggtaatg3'
ACPT40MR	T40M reverse primer	5'cattaccagctcaac**cat**gtcaagagaatc3'
ACPI55MF	I55M forward primer	5'gagtttgatactgag**atg**ccggacgaagaag3'
ACPI55MR	I55M reverse primer	5'cttcttcgtccgg**cat**ctcagtatcaaactc3'
ACPI63MF	I63M forward primer	5'gaagctgagaaa**atg**accaccgttcag3'
ACPI63MR	I63M reverse primer	5'ctgaacggtggt**cat**tttctcagcttc3'
ACPI70MF	I70M forward primer	5'gttcaggctgcc**atg**gattacatcaacggc3'
ACPI70MR	I70M reverse primer	5'gccgttgatgtaatc**cat**ggcagcctgaac3'

4.2.15 Growth and Incorporation of Selenomethionine into the *E.coli* ACP Methionine Mutants.

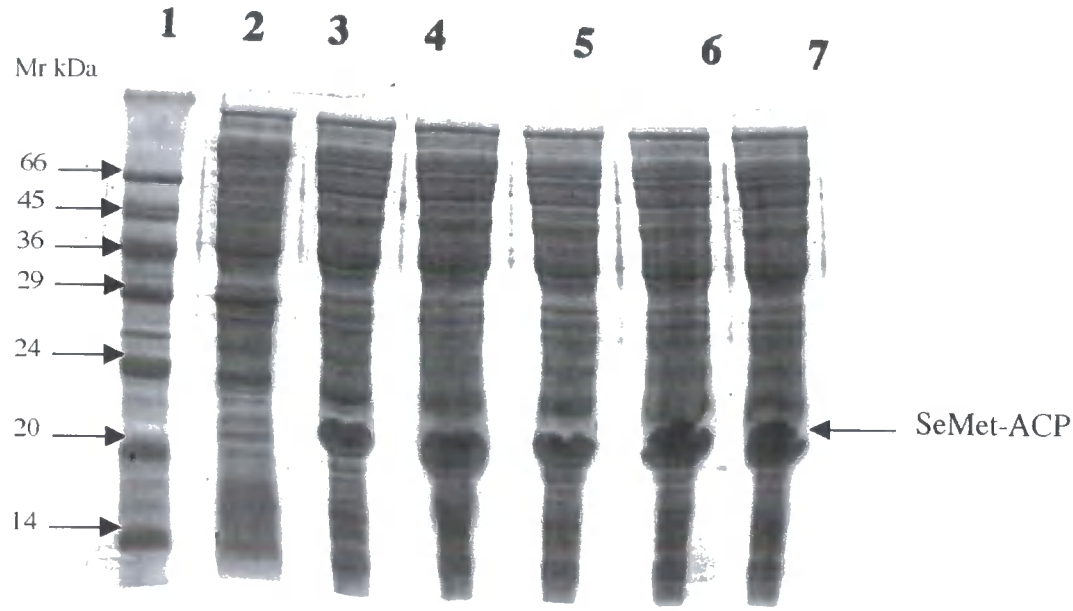
Plasmid DNA for the five methionine mutants of *E.coli* ACP was transformed into the BL21(DE3) metC$^-$ *E.coli* strain and the cultures were grown in M9 minimal medium and induced using the same conditions used for recombinant *E.coli* ACP without methionine mutations. The growth rate and induction of protein for the mutants was similar to that of the non-mutated protein grown under these conditions i.e. approximately 9 hours growth at 37°C to reach an OD 550 nM of 0.6 and induction for further 6 hours following the

addition of IPTG. However there was no obvious sign of an induced protein band corresponding to the position of ACP on SDS-PAGE gel analyses following freeze thaw extraction of the bacterial cells. To attempt to overcome this a alternative protocol (for seleno-methionine incorporation by metabolic inhibition of the methionine pathway in a normal BL21 (DE3) *E.coli* strain was used (Van Duyne *et al.*, 1993). In this method methionine was replaced in the growth medium by seleno-methionine which was added together with excess amounts of other amino acids (phenylalanine, threonine, isoleucine, leucine and valine) known to inhibit methionine biosynthesis at the point of IPTG induction of the recombinant ACP.

DNA for the five mutants was transformed into *E.coli* BL21(DE3) plysS cells using the standard transformation protocol. The transformants were grown in 800 ml liquid cultures of M9 minimal medium with no amino acid enrichment on an orbital shaker (150 rpm) at 37°C for 9 hours until an OD550 nM of 0.6 was reached. At this point the following final concentrations of amino acids and seleno-methionine were added to the cultures as solid powders: lysine (100mg/l), phenylalanine (100mg/l), threonine (100mg/l), isoleucine (50mg/l), leucine (50mg/l), valine (50mg/l) and seleno-methionine (50mg/l). The cultures were re-incubated at 37°C for 15 minutes to allow the inhibition of methionine biosynthesis to start. 0.4mMIPTG was added to induce the production of recombinant ACP and the cultures re-incubated at 37°C on an orbital shaker (150 rpm) for a further 6 hours. The cells were harvested from the growth medium and following freeze thaw

extraction the protein profile was analysed by SDS-PAGE. An induced band corresponding to the profile of *E.coli* ACP could be seen for all five mutations (**Figure 4.16**). The selenomethionyl methionine mutants were all purified acylated (butyryl) and checked by MALDItof analyses using exactly the same conditions as described for the cysteine mutants before being set down for crystallography trials. The MALDItof mass data for the purified mutants showed that they all contained the expected methionine mutation, that selenium was incorporated at both the naturally occurring (met 45) and at the introduced methionine residue and that quantitative derivatisation to butyryl-ACP occurred in each case (**Table 4.4**).

Figure 4.16 SDS-PAGE analyses showing the over-expression of recombinant *E.coli* ACP methionine mutants grown on minimal medium in the presence of selenomethionine



Coomassie stained 15% SDS-PAGE gel of freeze thaw protein extracts of methionine mutants of *E.coli* ACP grown and induced in the presence of seleno-methionine. Lane 1: SDS VII Mr markers, lane 2: un-induced protein extract, lane 3: I12M, lane 4: T40M, lane 5: I 55M, lane 6: I63M and lane 7: I70M. 2µl of freeze thaw cell free extract was loaded in each sample lane. The position of the over-expressed ACP band is arrowed.

Table 4.4 MALDItof analyses of methionine mutants of *E.coli* ACP grown in the presence of seleno-methionine and acylated in enzymatic synthesis using butyryl-CoA.

E.coli ACP	apo ACP (m/z)	butyryl ACP (m/z)
I12M	8627	9034
T40M	8632	9045
I55M	8621	9038
I63M	8626	9038
I70M	8625	9032

0.5µl of purified ACP or acyl-ACP reaction product was loaded together with 0.5µl sinapinic acid matirx solution directly onto a MALDI target plate. Spectra were calibrated using a calibration mixture containing bovine insulin (m/z 5734), E.coli thioredoxin (m/z 11674) and horse myoglobin (16952).

Expected mass (m/z) for selenomethyl-apo-ACP with I-M mutation = 8622, and T-M mutation = 8634.

Expected mass (m/z) for selenomethyl-butyryl-ACP with I-M mutation = 9033 and T-M mutation = 9049.

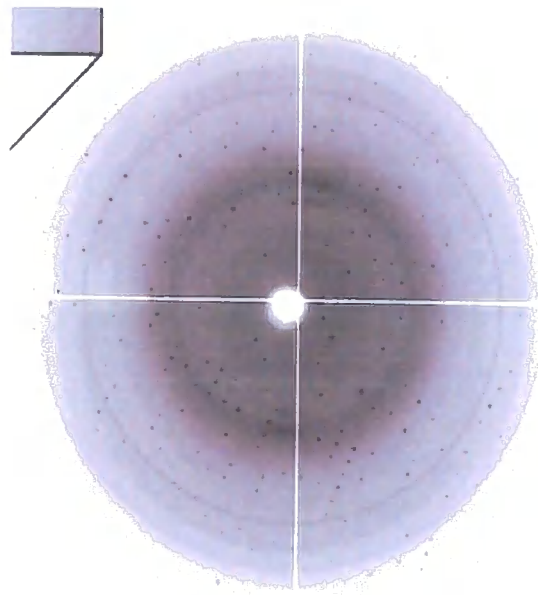
4.2.16 Resolution of the X-ray Crystal Structure of *E.coli* Butyryl-ACP.

Crystals were obtained for the I63M mutant of *E.coli* ACP (**Figure 4.17a**) which were stable in the X-ray beam and which gave diffraction data at 1.2 Å resolution (**Figure 4.17b**). This data together with the data collected for the recombinant butyryl protein and the selenomethionyl recombinant butyryl protein were used in molecular replacement experiments to resolve for the first time the complete crystal structure of *E.coli* butyryl ACP (**Figure 4.18a**). The protein structure is composed of four α-helices, and is broadly similar to the NMR model available for *E.coli* ACP. It contains 77 of the expected 78 amino acid residues with the acylated 4'phosphopantetheine group attached to serine 36. A hydrophobic cavity within the structure close to the phosphopantethenylated serine (36) (**Figure 4.18b**) has been identified as a putative acyl-chain-binding site large enough to accommodate up to eight carbon units. It is proposed that this may function to stabilise the acyl-chain, by protecting the most reactive part of the growing chain from side reactions during the fatty acid synthesis cycle (Roujeinikova *et al.*, in press).

The electron density in the structure around this hydrophobic cleft was weak and therefore the position of the butyryl prosthetic group within the cleft could not unambiguously be assigned. Further investigations with a range of acyl-ACPs of different chain lengths will continue on from this work.

Figure 4.17 Crystals and X-ray diffraction data for *E.coli* I63M seleno-methionine-butyryl-ACP.

A group of SeMet butyryl-ACP I63M crystals. These crystals have the same morphology as recombinant *E.coli* ACP and belong to space group $P_2 1_2 1_2 1$ with cell dimensions a=27, b=42, c=65 Å.



A 1° rotation diffraction pattern collected at 1.2 Å resolution for a cryo-cooled orthorhombic crystal of *E. coli* SeMet butyryl-ACP I63M.

Figure 4.18 The crystal structure of *E.coli* butyryl-ACP.

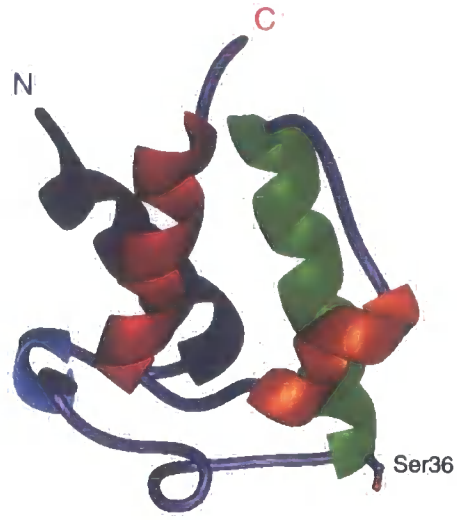


Figure 4.18a Ribbon diagram representing the crystal structure of *E.coli* ACP. The structure is composed of four α helices that form a right twisted bundle. α_1 is formed by residues 3-15, α_2 residues 36-50, α_3 residues 56-61 and α_4 residues 65-75. A long loop region (residues 16-35) links helices α_1 and α_2. The position of the phosphopantethenylated serine residue (S36) is shown.

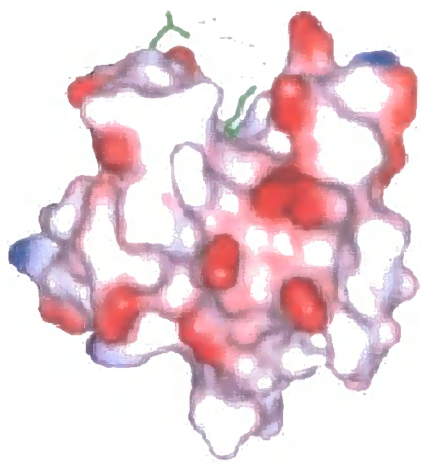


Figure 4.18b Solvent-accessible surface representation of *E.coli* ACP showing a deep hydrophobic cleft close to the phosphopantethenylated serine (S36) residue, which reaches deep into the four helix bundle of the structure. This cleft is believed to be involved in the binding of the acyl-chain to protect it during sequential cycles of FAS. It is shown with a butyryl-prosthetic group modelled as a stick representation.

4.3 Discussion

During this study wild type *E.coli* ACP was purified to homogeneity using a modified method of Majerus *et al.*, 1964. The purified protein was acylated with C4 and C8 acyl groups in chemical derivitisation procedures with N-acylimidazoles. These acyl-ACPs were authenticated by electro-spray mass spectrometry (ESMS) and conformational gel analyses (Post-Beittenmiler *et al.*, 1991) and used to obtain the first crystals of the acylated protein to give X-ray diffraction data.

This preliminary crystallography work resulted in conditions under which *E.coli* ACP would crystallise and gave diffraction patterns to 1.9 Å resolution. It also highlighted the need for a source of recombinant protein that could be used to introduce heavy metal sites into the structure to aid in solving the phasing problem.

Recombinant protein was obtained from *E.coli, Brassica napus*, strawberry and *Mycobacterium tuberculosis* and all were over-expressed in *E.coli*, purified and authenticated by native-PAGE and matrix assisted laser desorbtion time of flight (MALDItof) mass spectrometric analyses. Typically approximately 15 mgs of 95% pure apo-ACP was purified from 800ml batch cultures of each of these recombinant ACPs. Enzymatic reaction conditions were established with purified recombinant holo-ACP synthetase from *E.coli* that allowed these recombinant ACPs to be either pantethenylated or acylated directly with coenzyme A (CoASH) or acyl-CoA in a single step. This

resulted in a single acyl-ACP product rather than the side reaction products observed with chemical acylation reactions of ACP.

Once acylated these recombinant ACPs were put down to crystallography trails and conditions were found under which butyryl-ACP from *E.coli* crystallised and diffraction data was collected at 1.9 Å resolution.

Site directed mutagenesis studies were performed to introduce cysteine and methionine residues into the recombinant *E.coli* ACP in order to make heavy metal (mercury and selenium respectively) derivatives to aid in the phase resolution of the X-ray diffraction data. In all four cysteine (T23C, A26C, S27C, and D51C) and five methionine (I12M, T40M, I55M, I63M and I70M) mutants of *E.coli* ACP were produced.

Growth conditions for the over-expression of recombinant *E.coli* ACP in the presence of seleno-methionine were optimised, and selenomethionyl recombinant protein of all of the methionine mutants was produced, purified, acylated and authenticated by MALDtof-MS. X-ray diffraction data sets at high resolution were collected for recombinant butyryl-selenomethionyl-ACP and I63M butyryl-selenomethionyl-ACP. This data together with the data for the butyryl wild type protein was combined to solve for the first time the crystal structure of *E.coli* butyryl–ACP and to identify putative acyl-chain binding sites within the protein. The ability to crystallise *E.coli* butyryl-ACP and resolve its structure has important implications for future work on co-crystallography studies between ACP and other fatty acid synthetase (FAS) components. With the availability of

the data presented here it is likely that conditions may now be found under which ACP can be crystallised together with one of its partner enzymes. Crystal structures for a number of these partner enzymes exist but to date obtaining co-crystals has presented a major challenge. This is now a step closer having successfully obtained crystallisation conditions for *E.coli* butyryl-ACP and resolved its complete crystallographic structure. Structural details of how both the acyl chain and ACP itself interact with these enzymes will also help in interaction studies aimed at understanding the nature of the proposed metabolon complex of plant and bacterial type II FAS.

Chapter 5

Site Directed Mutagenesis Studies on the Soluble 1-Acyltransferase (G3PAT) From Squash (*Cucurbita moschata*) to Enable the Solution Of Its Atomic Structure and the Alteration Of Its Substrate Specificity.

5.1 Introduction.

The biosynthesis of glycerol lipids involves three successive acylation reactions and a de-phosphorylation step outlined in **Figure 1.5 (Chapter 1).** This biosynthetic pathway is known as the Kennedy pathway (Kennedy 1961). The first acyltransferase (Glycerol-3-phosphate -1– acyltransferase (G3PAT)(EC 2.3.1.15) in the pathway is present in two different forms. A soluble form located in the chloroplast in plants, which has also been reported in the cytosol of an oleagenous yeast (Gangar *et al.*, 2001), and a membrane associated form that has been identified from both bacteria and a number of plant sources. G3PAT catalyses the acylation of the *sn1* position of glycerol-3-phosphate incorporating either saturated or unsaturated fatty acids of C16 or C18 chain length to yield 1-acylglycerol-3-phosphate (lysophosphotidic acid) (Roughan and Slack 1982, Frentzen *et al.*, 1983). In bacteria the membrane bound enzyme has been isolated and the gene *pls*B has been cloned and the protein over-expressed in *E.coli* giving rise to microtubules in an expanded endomembrane system of the bacteria (Lightner *et al.*, 1980).

The plant soluble chloroplast located enzyme has been purified from a number of sources including pea, (Weber *et al.*, 1991), spinach, (Frentzen *et al.*, 1983) and squash (Nishida *et al.*, 1987). In squash three isoforms were identified AT1, AT2, and AT3 all with different p*I*s and molecular masses, AT1 is the most abundant form (Nishida *et al.*, 1987). G3PAT is nuclear encoded and its import into the chloroplast requires a signal peptide which following import is removed to yield a mature protein. The exact location of the processing site for the signal peptide within the pre-protein has not been determined experimentally although it has been predicted for *Arabidopsis* and squash using chloroplast target sequence prediction programmes and N-terminal sequence data obtained from the purified protein from squash (Nishida *et al.*, 1987). The original prediction (Ishizaki *et al.*, 1988) gave a mature squash enzyme 27 amino acids shorter than the new proposed mature protein derived from alignments of full length cDNA sequences from plants (Murata and Tasaka 1997). Transformation experiments using the truncated form of the squash enzyme in tobacco (Murata *et al.*, 1992) and enzyme assays using over-expressed protein (Slabas *et al.*, in press) have shown that this form of the enzyme is biologically active both *in vivo* and *in vitro*.

Soluble G3PAT can utilise both acyl-ACPs and acyl-CoAs as acyl donors, however when presented with both in competitive assays they show a strong preference for acyl-ACPs (Frentzen *et al.*, 1983). This is consistent with the location of both G3PAT and C16 (palmitate) and C18:1 (oleate) acyl-ACP within the chloroplast (Frentzen *et al.*, 1983).

This distinct preference for acyl-ACPs over acyl CoAs distinguishes the soluble chloroplast located enzyme from the membrane bound form, which can only utilise acyl-CoA. The ready commercial availability of acyl-CoAs and the relative difficulty in synthesising acyl-ACPs has however led to the use of acyl-CoAs rather than acyl-ACPs in early selectivity and kinetic assays on soluble G3PAT. The use of this substrate analogue, and the use of a single substrate instead of competitive assays with dual substrates, may have led to erroneous conclusions on the selectivity of the enzyme between saturated and unsaturated acyl substrates. Most of the early substrate selectivity studies have been performed with little consideration of the known metabolite concentrations present in the chloroplast. They have largely failed to consider changes in pH (from pH8.0 to pH 7.4) associated with light to dark transitions. They have also not taken into account estimates for the glycerol–3-phosphate concentrations in chloroplast. These vary between 65-90μM in spinach (Frentzen 1983) and 450 –620μM in *Amaranthus* (Cronan and Roughan 1987). Acyl-ACP concentrations in the plastid vary between 0.6-1.9μM for C16:0 and 0.4-2.0μM for 18:1 dependant on assumptions made of chloroplast volume (Soll and Roughan 1982, Roughan and Nishida 1990). Not withstanding the complications on the physiological relevance of assays for substrate selectivity performed by many investigators the purified G3PAT enzyme from a number of plant sources has been assayed for its substrate selectivity between 18:1 and C16:0 acyl-ACP *in vitro*. These assays have demonstrated that in broad terms this enzyme can

be classified into types which are *oleagenous*, preferring 18:1, or *non-selective*, incorporating 16:0 and18:1 at about the same rate. Selective enzymes have been isolated from pea, spinach (Betrams and Heinz1981) and *Arabidopsis* (Nishida *et al.*, 1993) and non-selective from squash (*Cucurbita moschata*) (Frentzen *et al.*, 1987). A close correlation between the chilling tolerance of plants and the nature of the fatty acid at the *sn-1* position of their phosphatidylglycerols (PGs) was discovered by Murata (1983) and later confirmed by Roughan (1985) and Kendrick and Bishop (1986). Plants which are chilling sensitive have a non-selective G3PAT and incorporate large quantities of saturated (16:0) fatty acids into the *sn1* position of their PGs, whilst those which are chilling resistant preferentially incorporate unsaturated fatty acids (18:1). This has led to the hypotheses that PG with increased level of unsaturated fatty acids allows the membranes of chilling tolerant plants to remain more fluid at low temperatures and hence they are less sensitive to chilling damage (Murata *et al.*,1992). This was supported by tansgenic studies in tobacco when constructs for *Arabidopsis* and squash G3PAT were targeted to the chloroplast (Murata 1992). The transgenes containing *Arabidopsis* G3PAT became more chilling resistant whereas those containing squash became much more chilling sensitive. A similar transgenic experiment was performed when the *plsB* gene from *E.coli* was transformed into *Arabidopsis*, which again resulted in increased levels of 16:0 fatty acid in the glycerolipids of the plant and a consequent increase in chilling sensitivity (Wolter *et al.*, 1992).

The primary amino acid sequence of all plant G3PATs is highly conserved. A comparison of the mature protein sequence from, *Arabidopsis*, bean, cucumber, pea, spinach and squash (**Figure 5.1**) showed 39% identity, 61% homology and only identified 16 residues amongst the conserved residues in the sequence from the chilling resistant plants which were different in the squash sequence. Therefore it is highly unlikely that primary sequence alignments alone would indicate individual residues which may be important in substrate selectivity.

Amino acid sequence alignments of all glycerolipid acyltransferases have however identified a conserved **H(X)$_4$D** motif believed to be involved in binding of the glycerol-3-phosphate. Site directed mutations of the histidine residue within this motif of the *E.coli plsB* sequence resulted in total loss of biological activity (Heath and Rock 1998). Mutation of the D to an E however results in only a partial loss of activity. This same series of conserved residues are found in the soluble G3PAT from plants, in the chilling resistant plant sequences (*Arabidopsis*, spinach and pea) it is **HQSEAD** whilst in the chilling sensitive plants (squash, kidney bean and cucumber) sequences the serine is a threonine (**Figure 5.1**).

Figure 5.1 CLUSTAL W (1.8) amino acid sequence alignment of G3PAT sequences from chilling tolerant and chilling resistant plants.

.

(*): conserved residue, (:) conserved substitution (strong homology), (.) different residue (weak homology) and (-) missing residues.

Garden pea:*Pisum sativum* (P30706), Kidney bean:*Phaseolus vulgaris* (Q43822), Squash:*Cucurbita moschata* (P10349), Cucumber:*Cumis sativus* (Q39639), Spinach:*Spinacea oleracea* (Q43869) and Arabidopsis:*Arabidopsis thaliana* (Q43307). SwissProt database accession numbers are in parentheses.

The sequences contain 166 conserved amino acids. The conserved $H(X)_4D$ acyltransferase domain is highlighted in bold red text. In the chilling resistant plant sequences (pea, *Arabidopsis* and spinach) it is HQSEAD whilst in the chilling sensitive plant sequences (squash, bean and cucumber) the serine is a threonine. The four cysteine residues present in the amino acid sequence from squash are highlighted in bold blue text.

Figure 5.1 CLUSTAL W (1.8) amino acid sequence alignment of G3PAT sequence from chilling tolerant and chilling resistant plants.

```
Gardenpea       -----MTDSFAHCASHIN------YRHKMKTMFIFSTP---CCSPSTAFFSPFRASNSKP
Kidneybean      ---MSMTGSSAYYVAHAIPPF---LRLSNKTMLLLSTPPTTFFPTSTTPRVTLLSSTSSS
Squash          ------------------------------------------------------------
Cucumber        MFILSAVSSSSSSSSSVPSSLPPFSLSPSISLSFSRVSLPPSSSSSSSSLKLFLPLSLHF
Arabidopsis     ---MTLTFSSSAATVAVAAAT---VTSS------ARVPVYPLASSTLRGLVSFRLTAKKL
Spinach         ---MLVLSSSAPPVLEVCKDR---VSSSFSTSSSSSSSAFSAVVFRRSFFTRFNSSLICC

Gardenpea       LRSTLSLRSS----ISSS-SITSTSHCSLAFNIVKHKEKN-VVSANMTSSVSSRTFLNAQ
Kidneybean      SSSSISLRSS----TAPSPSCSSVTPKDNCLASAKHS------PPNMSASVSSRTFLNAQ
Squash          ------------------MAELIQDKESAQSAATAAAAS-SGYERRNEPAHSRKFLDVR
Cucumber        TPPKLSSPHSFLRFSASRAMAELIQDKESAHTPSTTDVT------R-NDPPHSRAFLDLR
Arabidopsis     FLPPLRSRGG----VSVRAMSELVQDKESSVAASIAFNEA-AG-ETPSELNHSRTFLDAR
Spinach         CSSKLKLMADTALPSSSSSTSASASYSAAAKSVEEENHEIPVKKEDDNQLLRSRTYRNVR
                                                .                 .    ** : : :

Gardenpea       NEQDVLSGIKKEVEAGTLPASIAAGMEEVYLNYKSAVIKSGDPKANEIVLSNMTALLDRI
Kidneybean      SEQDVFAGIKKEVEAGSLPANVAAGMEEVYNNYKKAVIQSGDPKANEIVLSNMIALLDRV
Squash          SEEELLSCIKKETEAGKLPPNVAAGMEELYQNYRNAVIESGNPKADEIVLSNMTVALDRI
Cucumber        SEEELLSCIRRETEAGKLPSNVAAGMEELYQNYKNAVFESGNPKADEIVLSNMTVALDRI
Arabidopsis     SEQDLLSGIKKEAEAGRLPANVAAGMEELYWNYKNAVLSSGASRADETVVSNMSVAFDRM
Spinach         SAEELISEIKRESEIGRLPKSVAYAMEGLFHYYRNAVLSSGISHADEIVLSNMSVMLDFV
                . ::::: *:;* * * ** .:* .** ::  *:.**:.** .:*:* *;*** . :* :

Gardenpea       FLDVKEPFVFEAHHKAKREPFDYYMFGQNYIRPLVDFETSYVGNMPLFIQMEEQLKQGHN
Kidneybean      FLDVTDPFVFQPHHKAKREPFDYYVFGQNYIRPLVDFKNAYVGNMPLFIEMEEKLKQGHN
Squash          LLDVEDPFVFSSHHKAIREPFDYYIFGQNYIRPLIDFGNSFVGNLSLFKDIEEKLQQGHN
Cucumber        LLDVEDPFMFSPHHKAIREPFDYYTFGQNYVRPLIDFENSFVGNLSLFKDIEEKLHQGHN
Arabidopsis     LLGVEDPYTFNPYHKAVREPFDYYMFVHTYIRPLIDFKNSYVGNASIFSELEDKIRQGHN
Spinach         LLDIEDPFVFPPFHKAIREPADYYSFGQDYIRPLVDFGNSYVGNIAIFQEMEEKLKQGDN
                :*.: :*; * ..*** *** *** * ; *;***:** .;;*** .:* ;;*;;;;**.*

Gardenpea       IILMSNHQSEADPAIIALLLEMRLPHIAENLIYVAGDRVITVPLCKPFSIGRNLICVYSK
Kidneybean      IILMSNHQTEADPAIISLLLETRLPYIAENLTYVAGDRVITDPLSKPFSIGRNLICVYSK
Squash          VVLISNHQTEADPAIISLLLEKTNPYIAENTIFVAGDRVLADPLCKPFSIGRNLICVYSK
Cucumber        VVLISNHQTEADPAIISLLLEKTNPYIAENMIYVAGDRVIADPLCKPFSIGRNLICVYSK
Arabidopsis     IVLISNHQSEADPAVISLLLEAQSPFIGENIKCVAGDRVITDPLCKPFSMGRNLICVYSK
Spinach         IILMSNHQSEADPAVIALLLEKTNSLIAENLIYIAGDRVITDPLCKPFSMGRNLLCVYSK
                ::*:****:*****:*:****   . *.**   :*****;; **.****;****:*****

Gardenpea       KHMLDNPELVDMKRKANTRSRKEMAMLLRSGSQIIWIAPSGGRDRPVANSGEWAPAPFDS
Kidneybean      KHMLDDPALVEMKRTANIRALKEMAMLLRNGSQLVWIAPSGGRDRPDAQTREWVPAPFDI
Squash          KHMFDIPELTETKRKANTRSLKEMALLLRGGSQLIWIAPSGGRDRPDPSTGEWYPAPFDA
Cucumber        KHMLDIPELAETKRKANTRSLKEMALLLRGGSQLIWIAPSGGRDRPDPSTGEWYPAPFDA
Arabidopsis     KHMNVDPELVDMKRKANTRSLKEMATMLRSGGQLIWIAPSGGRDRPNPSTGEWFPAPFDA
Spinach         KHMYDDPELVDVKKRANTRSLKELVLLLRGGSKIIWIAPSGGRDRPDAVTGEWYPGTFDF
                ***   * *.; *; ** *; **:. :**.*.:::********** . ; ** *..**

Gardenpea       SSVDNMRRLVDHSGPPGHIYPLAILCHDIMPPPLKVEKEIGEKRIISYHGTGISTAPEIS
Kidneybean      SSVDNMRRLVEHSGPPGHVYPLAILCHDIMPPPLKVEKEIGEKRIICFHGAGISVAPAIS
Squash          SSVDNMRRLIQHSDVPGHLFPLALLCHDIMPPPSQVEIEIGEKRVIAFNGAGLSVAPEIS
Cucumber        SSVDNMRRLLQHSGAPGHLYPLALLCYDIMPPPSQVEIEIGEKRVISFNGTGLSVGPEIS
Arabidopsis     SSVDNMRRLVEHSGAPGHIYPMSLLCYDIMPPPPQVEKEIGEKRLVGFHGTGLSIAPEIN
Spinach         AALDNMRRLVEHAGRPGHIYPLALLCYDIMPPPAQVEKEIGEKRVMSFHGVGVSVEPEIN
                :::******;;*;. ***;;*;;;**;****** ;** ******;; ::*.*;*  * *.

Gardenpea       FSNTTAACENPEKAKDAYTKALYDSVTEQYDVLKSAIHGKKGLQASTPVVSLSQPWK-
Kidneybean      FSETTATCENPEKAKEVFSKALYNSVTEQYNVLKSAIQGKKGFEASTPVVTLSQPWK-
Squash          FEEIAATHKNPEEVREAYSKALFDSVAMQYNVLKTAISGKQGLGASTADVSLSQPW--
Cucumber        FDEIAASRDNPDEVREAYSKALYDSVAKQYNVLKAAIDGKQELEASVADVSLSQPWI-
Arabidopsis     FSDVTADCESPNEAKEAYSQALYKSVNEQYEILNSAIKHRRGVEASTSRVSLSQPWN-
Spinach         YNDVSLGCKNDEEAKSVYGQALYNSVNEQYNVLKAAIHGKQGSGASTPTTSLSQPWAS
                :.: :   .. ::.:..: :**:.**  **::*::**  ::   **.. .:*****
```

Further sequence alignments, together with chemical modification and site directed mutagenesis studies on the G3PAT and LPAAT (lysophosphatidic acid acyltransferase) enzymes from bacteria, yeast and mammals have identified four blocks of sequence homology (I, II, III and IV) believed to be important in catalysis (Lewin *et al.*, 1999). The importance of specific amino acid residues in the *E.coli* enzyme were investigated by site directed mutagenesis (Lewin *et al.*, 1999). Conserved residues in blocks I, III, and IV were found to be important for acyltransferase activity and conserved residues in blocks II and III have a role in G3P binding. The soluble plant G3PAT enzymes were not included in the alignment, as blocks III and IV could not be found in the plant sequences. Chimeric proteins constructed from three portions (N-terminus, central region and C-terminus) of G3PAT from spinach and squash (Ferri and Toguri 1997) and *Arabidopsis* and squash (Kroon PhD thesis 2000) have been tested for altered substrate specificity. The results were difficult to interpret as the individual blocks of amino acids were quite large and in Ferri and Toguris work acyl-CoAs at high concentrations were used as substrates. Also essential amino acids involved in catalysis are likely to be conserved and would therefore not be altered in the construction of the chimeras, and chimeric proteins would not be expected to fold in their correct conformation. The highly conserved central domain was however identified as containing the most important structural elements for activity. All of this work was carried out without any structural insight, as there was no crystal structure available for a G3PAT enzyme from any source.

At the commencement of the research reported in this section of this thesis our collaborators at the University of Sheffield had grown crystals of a truncated construct of the squash protein, and diffraction data had been collected at 1.9Å resolution. The crystal structure could not be resolved however as data for heavy metal derivatives, generated by soaking ethyl mercury phosphate (EMP) into the four cysteine residues within the protein indicated that the heavy metal derivative crystals were non-isomorphous. This was interpreted as being due to the binding of metal to one of the cysteines causing a major conformational shift. We anticipated that by changing individual cysteine residues by mutagenesis that the structure would be conserved and isomorphous heavy metal derivatives could be obtained which would allow the determination of the X-ray structure of this enzyme.

This section of the thesis will describe experimentation carried out to attempt to overcome this problem, and enable the crystal structure of G3PAT from a plant source to be resolved for the first time. It involves the over-expression of the soluble protein from a number of plant sources and the use of site directed mutagenesis studies to generate mutants where individual cysteine residues have been eliminated to attempt to solve the non-isomorphism problem.

It is anticipated that following the elucidation of the structure of the enzyme that the structural data obtained would be used, together with knowledge available on conserved domains and residues within the protein to make targeted decisions for site directed

mutagenesis studies. These will be directed at identifying catalytically important amino acid residues and potentially altering the substrate specificity of the protein from squash. Changes in the catalytic activity of the mutants will be measured by *in vitro* competitive assays using radioactive acyl-ACPs as substrates. These will be synthesised from recombinant ACP prepared and pantethenylated enzymatically as described in **chapter 4**. Eventually it is hoped to use these studies to further our understanding of the interaction of G3PAT with its substrate and progress research in this area towards co-crystals of G3PAT with bound substrate. Clearly the availability of structural data for both the substrate acyl-ACP and the enzyme G3PAT will make this possible.

5.2 Results.

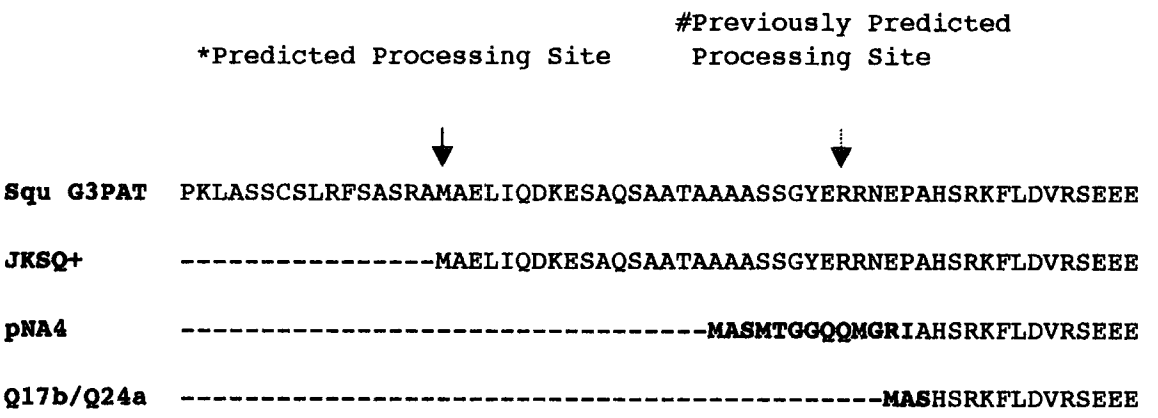
5.2.1 Over-expression of Squash 1-AT's

Four pET over-expression cDNA constructs encoding squash G3PAT were used for crystallography and mutagenesis studies aimed at solving the structure and locating the binding sites of G3P and acyl-ACP substrates within the protein. These are described in **Table 5.1.** and shown schematically in **Figure 5.2.**

Table 5.1 Squash G3PAT pET over-expression constructs used for crystallography and mutagenesis studies.

Construct	**Vector**	**Details**	**Antibiotic selection**
pNA4	pET3a	1.22 kb *NaeI – EcoR1* fragment encoding for a fusion protein of 14 amino acids originating from the vector and 366 amino acids encoding squash G3PAT beginning at the previously predicted processing site (Ishizaki 1988).	Ampicillin
Q17b	pET17b	1.107 kb *NheI – EcoR1* fragment encoding squash G3PAT beginning at the same processing site as pNA4 but with a 3 amino acid vector extension (MAS) instead of the 14 amino acids in pNA4.	Ampicillin
Q24a	pET24a	Same as Q17b but cloned into the pET24a vector to give more stable kanamycin selection.	Kanamycin
JKSQ+	pET17b	1.191kb *NheI – EcoR1* fragment encoding squash G3PAT (397 amino acids) beginning at the new predicted processing site (Murata and Tasaka 1997)	Ampicillin

Figure 5.2 N-terminal region of the pET over-expression cDNA constructs encoding squash G3PAT used for crystallography and mutagenesis studies.



Amino acids highlighted in bold were introduced from the vector during cloning.

Squ G3PAT represents the N-terminal region of the original cDNA clone showing the predicted processing sites. # original prediction (Ishizaki et al 1988), * new processing site predicted by Murata and Tasaka 1997.

pNA4 was the first over-expression construct made, and cloned into the pET3a vector.

Q17b was an over-expression construct made to represent what was believed to be the mature protein, and cloned into the pET17b vector.

Q24a was the same construct as Q17b but cloned into pET24a a kanamycin resistant vector.

JKSQ$^+$ is a construct made to represent the mature protein beginning at the new predicted processing site.

At the commencement of this work the constructs pNA4, Q17b and Q24a were over-expressed in *E.coli* BL21 (DE3) *pLysS* cells, and the protein purified using anion exchange chromatography in our laboratory. The purified protein from the truncated constructs Q17b and Q24a had successfully crystallised (**Figure 5.3a**) in trials set up at the University of Sheffield by Dr Andy Turnbull and Dr J.Rafferty and diffraction data (**Figure 5.3b**) at 1.9Åresolution had been collected. A heavy metal derivative of both the Q17b and Q24a constructs was produced by transformation of the pET plasmid DNA into the *E.coli* BL21 met C^- strain and growth and induction in the presence of seleno-methionine. Following purification the selenomethionyl proteins crystallised and diffraction data at 2.3Å resolution was collected.

A mercury derivative generated by soaking ethyl mercury phosphate into the four cysteine residues present in the protein was also generated and a diffraction data set at 2.05 Å resolution was collected for this derivative.

Using multiple isomorphous replacement techniques the data sets from both the native and the heavy metal derivatives were analysed to locate the position of the heavy metal sites, determine the protein phase angles and to begin to determine the structure. There was extensive non-isomorphism between the three data sets collected, caused by at least one of the cysteine heavy metal sites or by a fifth metal binding site within the structure. This resulted in the Patterson map being un-interpretable and therefore the crystal structure could not be resolved. To overcome this problem a site directed mutagenesis

Figure 5.3 Crystals and X-ray diffraction data pattern obtained with recombinant squash Q24a G3PAT.

Figure 5.3a . Crystals of recombinant squash G3PAT grown by the hanging drop vapour diffusion method using PEG 4000 as the precipitant. The crystals have a needle like formation with approximate dimensions 0.15 x 0.1 x 0.7 mm^3 and belong to the space group $P2_12_12_1$.

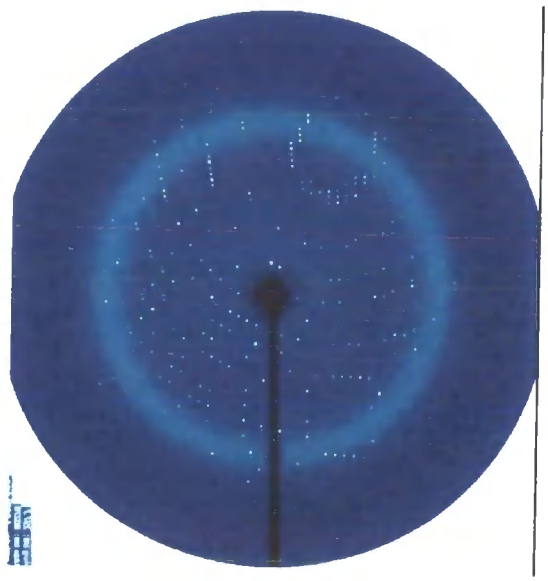


Figure 5.3b. Diffraction data pattern collected from crystals of squash Q24a G3PAT. A complete data set was collected at 1.9Åresolution from a single flash frozen crystal.

strategy was planned to replace each cysteine residue in the protein individually, to try to locate the rogue residue and to attempt to produce a heavy metal derivative which would give an isomorphous data set and allow the structure to be determined.

5.2.2 Site Directed Mutagenesis of Cysteine Residues In Squash G3PAT to Aid in the Resolution of the Crystal Structure.

The amino acid sequence of squash G3PAT contains four cysteine residues, two of which are absolutely conserved in all of the available plant G3PAT sequences. The other two are not conserved, one is a glycine in *Arabidopsis*, pea and bean and a glutamic acid in spinach and the other is a serine in the kidney bean sequence (**Figure 5.1**). Defining the methionine residue at the N-terminus of the Q17b and Q24a constructs as amino acid one the two conserved cysteines are C188 and C278 and the two which are not conserved are C20 and C177.

Each of these cysteine residues was mutated to a serine in the squash construct Q24a to yield four mutant proteins each with a different cysteine residue replaced.

Approximately 20 ng of pET24a plasmid DNA containing the Q24a DNA insert coding for squash G3PAT was used together with a pair of complimentary mutagenesis primers (**table 5.2**) in standard Quick Change ™ reactions to generate recombinant DNA.

Table 5.2. Nucleotide sequence of the oligonucleotide primers used for the cysteine mutagenesis of Squash G3PAT (Q24a). Nucleotides highlighted in bold text indicate the position of the mutation.

Primer name	Primer function	Sequence
1ATC20SF	C20S forward primer	5'gttgctctc**cag**catcaagaagg3'
1ATC20SR	C20S reverse primer	5'ccttcttgat**gct**ggagagca3"
1ATC177SF	C177S forward primer	5'gcagaccctctt**agc**aagcccttc3'
1ATC177SR	C177S reverse primer	5'gaagggctt**gct**aagagggtctgc3'
1ATC188SF	C188S forward primer	5'ggaatcttatt**agt**gtttattc 3'
1ATC188SR	C188S reverse primer	5'gaataaac**act**aataagattcc 3'
1ATC278SF	C278S forward primer	5'cttgctttatt**aagt**catgacatc3'
1ATC278Sr	C278S reverse primer	5'gatgtcatg**act**taataaagcaag3'

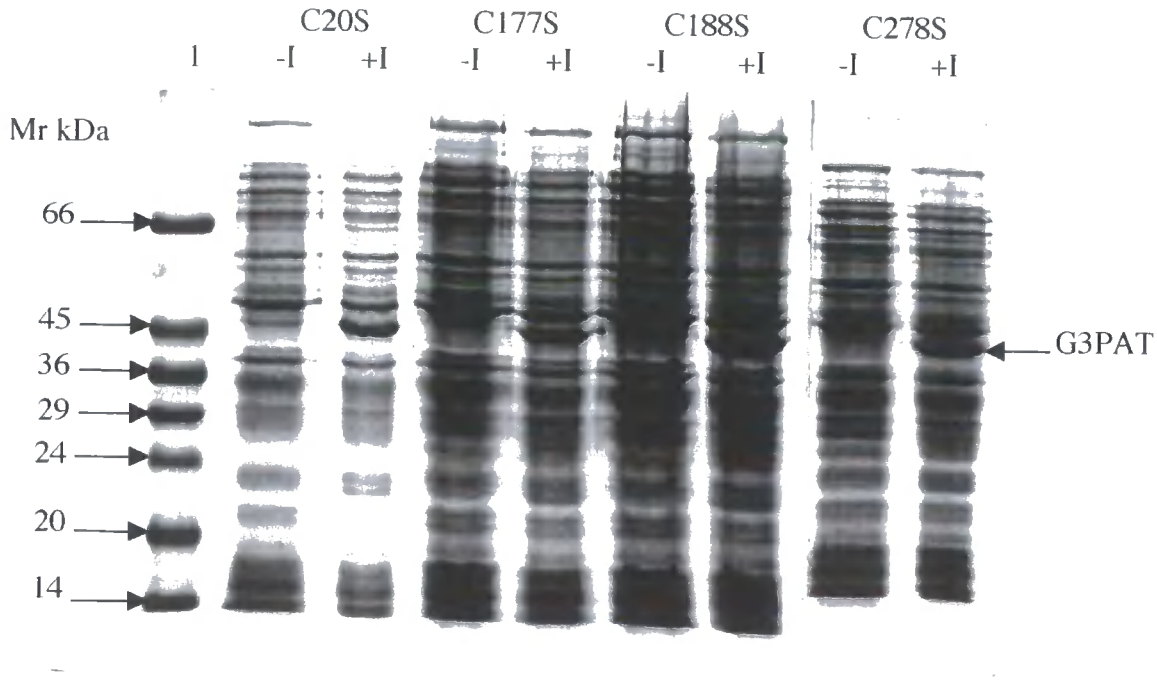
Parental plasmid DNA was removed from the mutant DNA reaction product by digestion with the restriction enzyme *Dpn*I and the mutated DNA was then transformed into super-competent Xl1blue *E.coli* cells and plated onto LB kanamycin plates and grown overnight at 37°C. Two individual colonies were selected from each of the plates and grown overnight in 5 ml liquid cultures for plasmid production. Plasmid DNA was isolated and DNA sequenced against the T7 promoter forward primer for C20, the internal G3PAT primer PQseq1 (5'gcagaaaacacgatc3' (begins at A159 in the sequence)) for C177 and C188 and the T7 terminator reverse sequencing primer for C278. These sequencing primers were chosen to ensure that good quality sequence data was obtained through the site of each mutation. Sequence data was obtained which confirmed that each of the cysteine residues had been mutated to serine.

Following this confirmation plasmid DNA for each of the mutations was transformed into BL21 (DE3) *plys*S *E.coli* cells. The transformants were grown, induced with IPTG and the over-expressed mutant G3PAT selectively released from the harvested cells by freeze thaw extraction.

The extent of over-expression was confirmed by analyses of the freeze thaw extracted protein on SDS-PAGE (**Figure 5.4**). All four mutants showed a high level of over-expression, with the stained band for G3PAT representing between 15 and 20% of the coomasssie stained protein visible on the gel (**Figure 5.4**).

The extracts were purified using ammonium sulphate fractionation, and anion exchange and gel filtration chromatography and the purified protein was crystallised using the hanging drop vapour diffusion method using 14-25%PEG 4000 as the precipitant. These conditions were the same as were used for squash recombinant G3PAT without the mutations. Heavy metal derivatives of each of the mutants were produced by soaking the crystals with 1mM ethyl-mercury phosphate (EMP) and diffraction data sets collected. The data collected for the four cysteine mutants together with that collected for the protein without the mutations and that for selenomethionyl squash G3PAT were all used in multiple isomorphous replacement analysis to solve the structure. The quality of the electron density map was good for 359 of the expected 368 residues. Each of the four heavy metal cysteine sites were located and a fifth heavy metal site, a rogue site in the original native heavy metal diffraction data, was identified as Histidine 279. The solution

Figure 5.4 SDS-PAGE analyses of the freeze thaw protein extracts of squash Q24a G3PAT cysteine mutants.



Lane 1 SDS VII Mr markers, (-I): protein extract from cells without IPTG induction, (+I): protein extract from cells following IPTG induction. The position of the induced G3PAT protein band is arrowed. Induced G3PAT represents between 15 and 20% of the coomassie stained protein visible in the lanes loaded with protein extracts from induced cells.

of the structure was previously impaired by not knowing that H279 bound heavy metal, this was clearly identified once the data for the four cysteine mutants was analysed. Cell free extracts of all four cysteine mutants were assayed for G3PAT activity and all were found to be fully active indicating that they are not required for substrate binding and catalysis.

5.2.3 Crystal Structure of G3PAT from Squash.

The resolution of the crystal structure of squash G3PAT represents the first high - resolution structural determination of any G3PAT. It was facilitated by the ability to produce over-expressed recombinant protein and the ability to make modifications to the protein by mutagenesis so overcoming problems with interpretation of diffraction data. This work was collaboration between workers in our laboratory and those at the Krebs Institute for Biomolecular Research at the University of Sheffield (Drs Andy Turnbull, and John Rafferty and Professor David Rice).

The data shows the monomeric G3PAT protein has overall dimensions of 55 x 65 x 75 Å (**Figure 5.5a**). It is constructed from 9 β strands and 14 α helices and is organised into two compact domains (**Figure 5.5b**). The first domain (I) is a four helix bundle made up of the first 77 N-terminal residues, this is linked by a loop region (residues 78-84) to the larger second domain (II) (residues 85-368), made up of mixed parallel and anti-parallel β sheets, and the remaining α helices. A noticeable feature of domain II is the presence

Figure 5.5 The crystal structure of squash G3PAT showing the structural domains and the position of modelled G3P and acyl-ACP substrates.

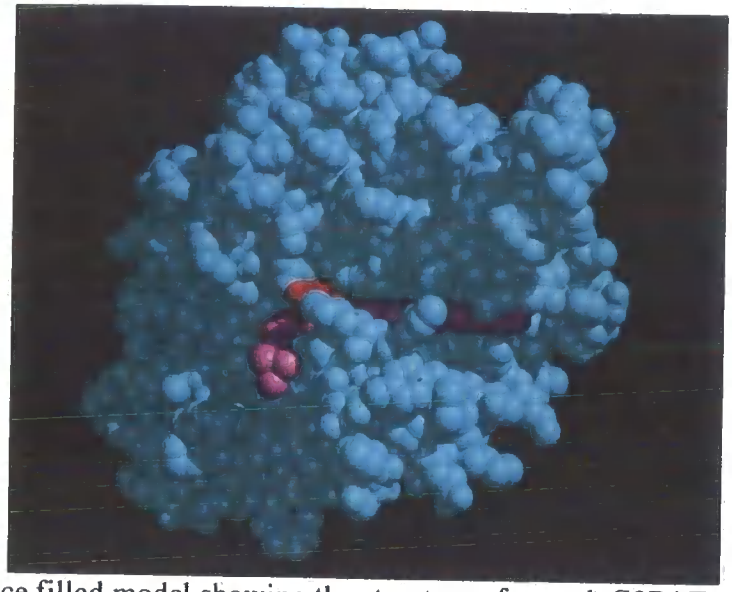


Figure 5.5a Space filled model showing the structure of squash G3PAT (light blue) with modelled C16 acyl-ACP (purple) and G3P substrate (red). A single monomeric unit has overall dimensions of 55 x 65 x 75 Å.

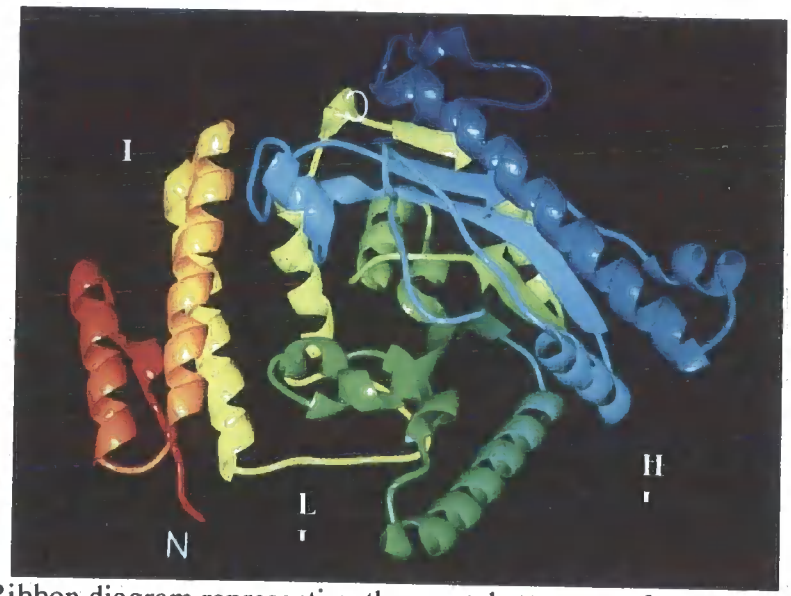


Figure 5.5b Ribbon diagram representing the crystal structure of squash G3PAT. The protein is made up of 9 β strands and 14 α helices and is organised into two compact domains. N:shows the position of the N-terminus of the protein. The first domain (I) is a four helix bundle made up of the first 77 N-terminal residues, this is linked by a loop region (L) (residues 78-84) to the larger second domain (II) (residues 85-368), made up of mixed parallel and anti-parallel β sheets, and the remaining α helices.

of a cleft ~ 20 Å in length, formed between one face of the β sheet and residues from two of the α helices. This cleft is lined with hydrophobic residues and has the overall appearance of a tunnel.

It has not been possible to date, to obtain crystals grown in the presence of either the G3P or an acyl-ACP substrate. However the use of the structural data with modelled substrates, together with sequence conservation data and the position of the residues of the $H(X)_4D$ (HQTEAD) motif within the crystal structure, has allowed putative G3P and acyl-ACP binding domains to be assigned (**Figure 5.5c and 5.5d**). The predicted G3P binding site lies within domain II of the protein with the phosphate group positioned at one end of the hydrophobic cleft in a positively charged pocket formed by the side chains of two arginine residues (R235 and R237), a lysine (K193) and a histidine residue (H194). Modelling of an acyl-ACP substrate (palmitoyl) into the crystal structure (**Figure 5.5d**) with the reactive group of the fatty acid close to the *sn-1* position of the modelled G3P, shows it to lie within the hydrophobic cleft, with 14 amino acid residues making Van der Waal contact (**Figure 5.5d**). These residues are H139, E142, V166, A167, G168, R170, V189, H194, L213, I229, R235, N257, M258 and L251 all of which are conserved throughout plant G3PAT sequences except V166 which is an isoleucine in spinach and L213 which is an arginine in the pea sequence (**Figure 5.1**). In order to examine the importance of these residues in substrate binding a series of site directed mutagenesis experiments were designed.

Figure 5.5c Close up of the position of the putative glycerol–3-phosphate binding site with modelled G3P. Amino acid residues R235, R237 and K193 form a positively charged pocket which binds the phosphate, H139 and D144 both conserved residues within the G3PAT $H(X)_4D$ domain, are arranged as a catalytic triad with the substrates C1-hydroxyl group.

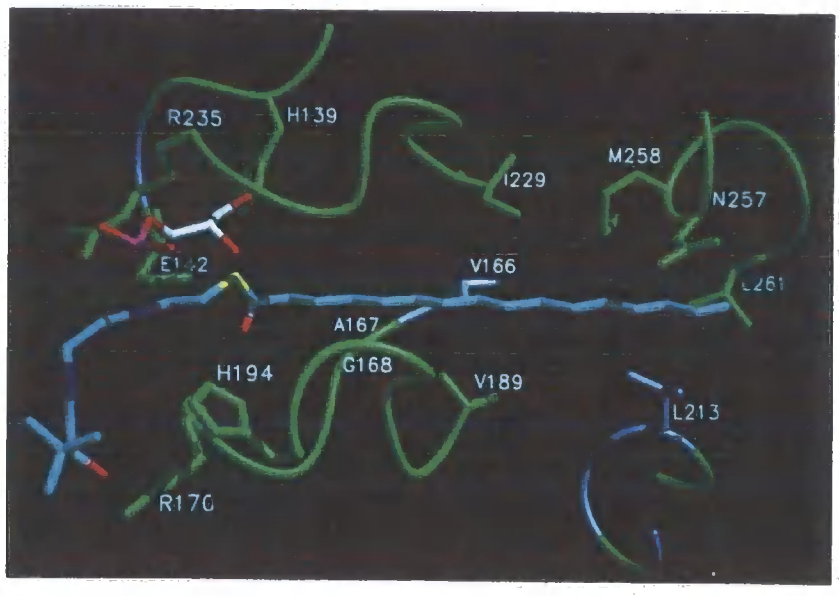


Figure 5.5d Close up of the position of putative acyl-ACP binding domain with modelled palmitoyl substrate. There are 14 residues (H139, E142, V166, A167, G168, R170, V189, H194, L213, I229, R235, N257, M258 and L251) making Van der Wall contacts with the modelled acyl chain.

5.2.4 Mutagenesis of Amino Acid Residues Predicted to be Involved in G3P Binding.

All four of the amino acid residues which were predicted to be involved in the binding of G3P (K193, H194, R235 and R237) were mutated in QuickChange ™ site directed mutagenesis reactions. Q24a plasmid DNA was used together with complimentary mutagenesis primers (**Table 5.3**) in reactions to replace the naturally occurring amino acid with a serine. Following transformation of the reaction product into *E.coli* XL1 blue cells, plasmid DNA was isolated and sequencing. DNA sequencing was carried out using the PQseq1 sequencing primer to ensure accurate sequence coverage through each mutation site. In all cases sequence data was obtained which confirmed the presence of the mutation.

Following sequence confirmation, plasmid DNA was transformed into competent *E.coli* BL21 (DE3) *PlysS* cells and mRNA for the mutant proteins were over-expressed following IPTG induction. Freeze thaw cell free extracts for each of the mutant proteins were prepared and these were analysed and the G3PAT quantified by coomassie staining on SDS-PAGE gels.

Table 5.3. Nucleotide sequence of the oligonucleotide primers used to generate mutations of the amino acid residues predicted to be involved in G3P binding.

Nucleotides highlighted in bold indicate the position of the mutation.

Primer name	Primer function	Sequence
1ATK193SF	K193S forward primer	5' gtttattcaaaa**agt**cacatgttcgatattcctg3'
1ATK193SR	K193S reverse primer	5' caggaatatcgaacatgt**act**tttttgaataaac3"
1ATH194SF	H194S forward primer	5' gtttattcaaaaaag**tcc**atgttcgatattcctg3'
1ATH194SR	H194S reverse primer	5' caggaatatcgaagat**gga**cttttttgaataaac3'
1ATR235SF	R235S forward primer	5' gcacccagtggtggt**tcg**gaccggccg3'
1ATR235SR	R235S reverse primer	5' cggccggtc**cga**accaccactgggtgc3'
1ATR237SF	R237S forward primer	5' ggtggtagggac**tcg**ccggatccctcg3'
1ATR237SR	R237S reverse primer	5' cgagggatccgg**cga**gtccctaccacc3'

An aliquot of each mutant protein, equivalent to 200 ng of recombinant G3PAT was assayed for activity in a standard dual substrate G3PAT assay (**2.8.12**) using ^{14}C 16:0 and ^{3}H 18:1 acyl-ACP at 1.1μM each. Activity measured for over-expressed Q24a protein without a mutation was used to give a control rate (**Table 5.4**). These assays were carried out in conjunction with Matthew Hayman a fellow PhD student in our laboratory. Mutations K193S, R235S and R237S showed no G3PAT activity even in assays performed for over an hour instead of the normal 2-3 minutes, thus confirming the importance of these residues in G3P binding. In separate binding studies carried out in our laboratory these mutants were still shown to be capable of binding acyl-ACP. The H194S mutant protein had approximately 80% of the control Q24a activity indicating that this residue is not critical for substrate binding and catalysis.

Table 5.4. G3PAT activity of squash G3PAT proteins containing amino acid mutations of residues predicted to be involved in G3P binding.

Mutation	G3PAT activity expressed as a percent of Q24a control activity. (standardised protein amounts were used in each assay)
Q24a	100% (arbitrary)
K193S	0%
H194S	79%
R325S	0%
R237S	0%

Assays were carried out for the standard 2 minute incubations and also repeated for 60 minute incubations to confirm the absence of activity in mutants K193S, R235S and R237S.

This mutagenesis data confirmed the structural predictions that all of these residues are involved in G3P binding and that at least three of the four are critical for biological activity.

5.2.5 Mutagenesis of Amino Acid Residues Predicted to be Involved in Acyl-ACP Binding.

Five of the amino acid residues (E142, V166, R170, L213 and N257) which were predicted to be within 5 Å of the modelled acyl chain, and lie within the hydrophobic cleft of domain II of the crystal structure were also mutated to either serine or alanine residues. As with the G3P binding mutants, these mutations were made using Q24a as the template DNA together with complimentary primers (**Table 5.5**) in QuickChange™ mutagenesis reactions.

Table 5.5. Nucleotide sequence of the oligonucleotide primers used to generate mutations of the amino acid residues predicted to be involved in acyl-ACP binding. Nucleotides highlighted indicate the position of the mutation.

Primer name	Primer function	Sequence
1ATE142AF	E142A forward primer	5' gatatcaaatcatcagactgca**gca**gatccagc3'
1ATE142AR	E142A reverse primer	5' gctggatc**tgc**tgcagtctgatgaattgatatc3"
1ATV166AF	V166A forward primer	5' gaaaacacgatctt**tgcg**gcaggggatag 3'
1ATV166AR	V166A reverse primer	5' ctatcccctg**ccgc**aaagatcgtgttttc3'
1ATR170SF	R170S forward primer	5' gtggcaggggat**agt**gttcttgcagaccctctttg 3'
1ATR170SR	R170S reverse primer	5' caaagagggtctgcaagaac**act**atcccctgccac 3'
1ATL213SF	L213S forward primer	5' gcaaacacacgaagt**tct**aaggagatg 3'
1ATL213SR	L213S reverse primer	5' catctcctt**aga**acttcgtgtgtttgc 3'
1ATN257SF	N257S forward primer	5' gcttcttcagtggac**gcc**atgagaagg 3'
1ATN257SR	N257S reverse primer	5' ccttctcat**ggc**gtccactgaagaagc 3'
1ATT141SF	T141S forward primer	5' gtcttgatatcaaatcatcag**agt**gaagcagatccag3'
1ATT141SR	T141S reverse primer	5' ctggatctgcttc**act**ctgatgatttgatatcaagac3'

Following DNA sequence confirmation using the T7 forward sequencing primer for E142A, V166A, R170S, and the PQSeq1 primer for L213S and N257S, plasmid DNA was transformed into *E.coli* BL21(DE3) *plys S* cells and the mRNA for the mutant proteins over-expressed.

Freeze thaw cell free extracts for each of the mutant proteins were prepared, quantified by coomassie staining on SDS-PAGE gels and an aliquot of each was used in a dual acyl-ACP substrate assay for G3P activity.

Mutations V166A, L213S and N257S did not show altered biological activity compared to that of the Q24a control without mutations (**Table 5.6**). The mutation E142A resulted in an inactive protein, indicating that this residue plays a critical role in catalysis. This residue lies in the HQTEAD motif of the squash sequence and is a conserved residue in all plant G3PAT enzymes. Although R170S still retained a small amount of biological activity it was drastically reduced to 5% of that of the Q24a control (**Table 5.6**) indicating that this is also a critical residue for catalysis or that its mutation prohibits the correct folding of the enzyme.

Table 5.6. G3PAT activity of squash G3PAT proteins containing amino acid mutations of residues predicted to be involved in acyl-ACP binding.

Mutation	G3PAT activity expressed as a percent of Q24a control activity. (standardised protein amounts were used in each assay)
Q24a	100% (arbitrary)
E142A	0%
V166A	97%
R170A	5.6%
L213A	104%
N257S	106%

Assays were carried out for the standard 2 minute incubations and also repeated for 60 minute incubations to confirm the absence of or greatly reduced activity in mutants E142A and R170A.

Amino acid sequence alignments of G3PAT sequences from both chilling sensitive and chilling resistant plants (**Figure 5.1**) identified a significant amino acid difference within the H(X)4D sequence of the chilling sensitive plants compared to that of the chilling resistance plants. In chilling resistant plants (*Arabidopsis*, spinach and pea) the sequence is **HQSEAD** whereas in chilling sensitive plants (bean, cucumber and squash) the serine is replaced by a threonine. This single amino acid difference within a conserved domain of the two plant types suggested that this residue might be important in substrate selectivity. To investigate this the threonine was mutated to a serine in a QuickChange ™

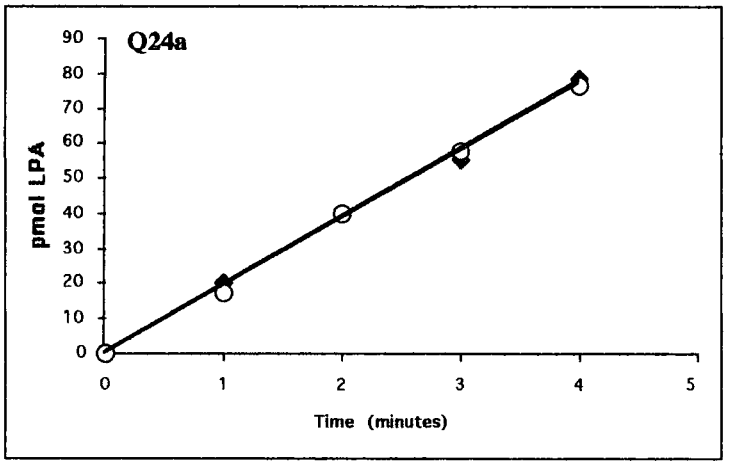
reaction using Q24a template DNA and oligonucleotide primers T141SF and T141SR (**Table 5.5**). Following DNA sequence confirmation the mRNA for the mutant protein was over-expressed, freeze-thaw extracted and assayed for any change in its substrate preference in a dual acyl-ACP substrate assay. The T141S mutant protein showed almost identical levels of activity to the wild type Q24a protein control and showed no preference for 18:1 acyl-ACP over 16:0 acyl-ACP. This was a surprising result to discover that mutation of a residue which appeared important from primary sequence alignment, actually had no effect on biological activity.

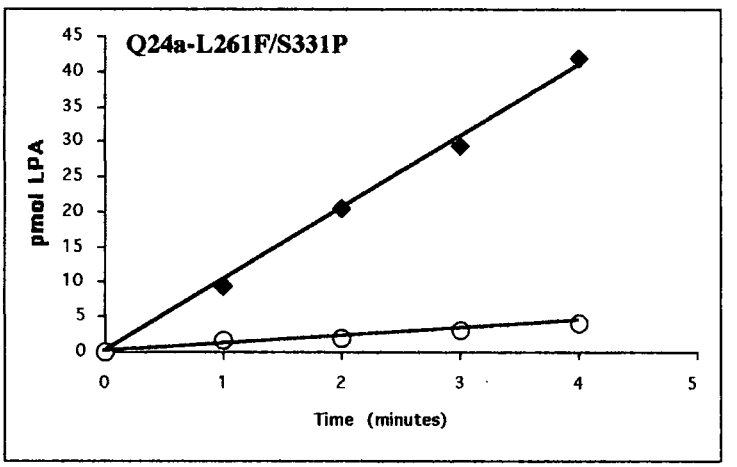
5.2.6 Altered Substrate Selectivity of G3PAT from Squash and Oil Palm.

The Q17b and Q24a pET over-expression constructs of squash G3PAT were generated using a PCR based protocol with a proof reading DNA polymerase (Vent™ DNA polymerase). Whilst assaying resultant transformants from this procedure using the standard dual acyl-ACP substrate assay it was noticed that one of them showed a preference for the 18:1 substrate. This was in marked contrast to the "wild type" plasmid which has no acyl group preference (**Figure 5.6**).

DNA sequencing of the entire coding region revealed two point mutations (L261F and S331P) that had occurred during the PCR reaction and which altered the substrate selectivity of the new protein. In order to identify which of these two residues had caused the alteration, site directed mutagenesis were carried out to mutate each of the residues

Figure 5.6 Substrate selectivity of squash G3PAT Q24a and the mutant G3PAT protein containing L261F and S331P double mutation.



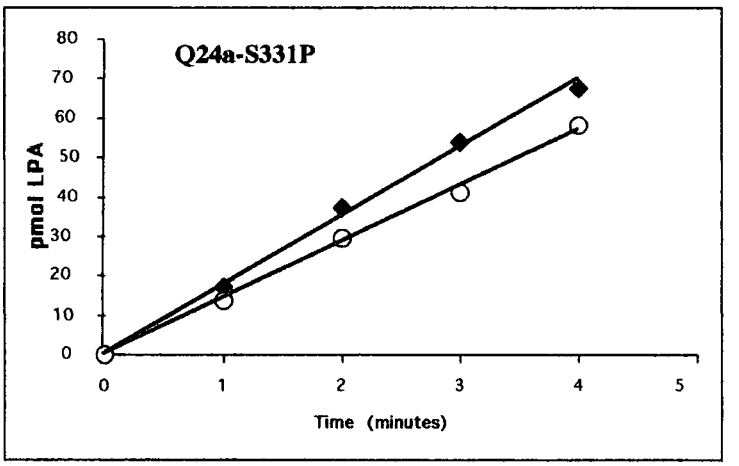


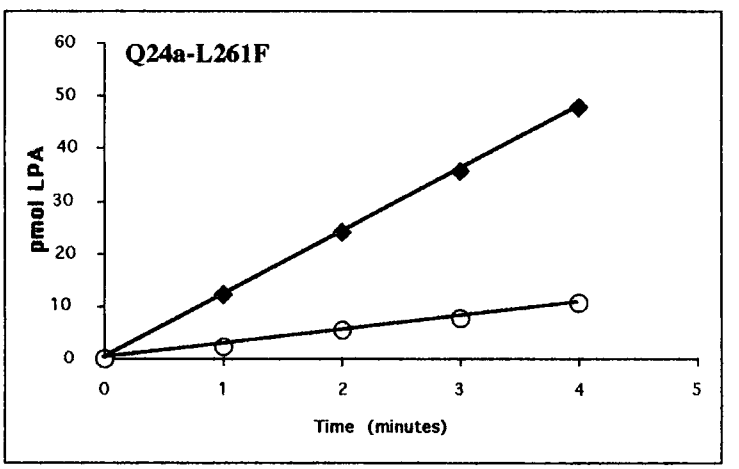
Substrate selectivity was measured as incorporation of differently labelled acyl-ACP thioesters into LPA under the standard dual substrate competitive assay conditions. ◆ = 18:1 LPA formation and O =16:0 LPA formation. The squash G3PAT protein containing the double mutation showed almost 4:1 preference for 18:1 acyl-ACP over 16:0 acyl-ACP whereas squash Q24a showed no acyl-ACP preference.

back to their correct wild type amino acid, and generate proteins containing only a single amino acid mutation.

QuickChange ™ mutagenesis reactions were carried out with template DNA and primer combinations 1ATF261LF (5'gtggacaacatgagaaggtttattcaacattcgg3') + 1ATF261LR (5'ccgaatgttgaataaaccttctcatgttgtccac3') and 1ATP331SF (5'gaggaggttagggaggcatacccaaaggcactgtttg3') + 1ATP331SR (5'caaacagtgcctttgggtatgcctccctaacctcctc3'). An aliquot of reaction product was transformed into *E.coli* Xl1blue cells and plasmid DNA was isolated and DNA sequenced completely in both directions to confirm that each point mutation had been corrected. Following DNA sequence confirmation plasmid DNA was transformed into *E.coli* BL21 (DE3) *plysS* cells and the mRNA for the mutant proteins were over-expressed and assayed for selectivity in a standard dual acyl-ACP substrate assay. The protein containing the single point mutation L261F showed marked substrate specificity for 18:1 ACP (**Figure 5.7**) whilst the protein with the S331P mutation had reverted back to the typical "wild type" activity, using both 18:1 and 16:0 acyl-ACPs at the same rate (**Figure 5.7**). Kinetic studies carried out in our laboratory (Matthew Hayman-PhD studies) have demonstrated that the preference of the L261F mutant for 18:1 acyl-ACP is a consequence of a much-increased km for the 16:0 acyl-ACP substrate. The crystal structure of squash G3PAT with a modelled acyl-ACP substrate showed this residue to

Figure 5.7 Substrate selectivity of squash G3PAT containing single point mutations S331P and L261F.





Substrate selectivity was measured as incorporation of differently labelled acyl-ACP thioesters into LPA under the standard dual substrate competitive assay conditions. ◆ = 18:1 LPA formation and O = 16:0 LPA formation. The mutant G3PAT protein containing the single mutation S331P showed no substrate selectivity i.e. showed wild type Q24a activity, whereas the protein containing the single mutation L261F showed almost 4:1 preference for 18:1 acyl-ACP over 16:0 acyl-ACP.

This data shows that the single point mutation L261F within squash G3PAT altered it from a non-selective enzyme, typical of that found in chilling sensitive plants, to one with a marked preference for 18:1 acyl-ACP, typical of that found in chilling tolerant plants.

lie within 10 Å of the terminal carbon atom of the fatty acyl chain and that it could therefore play a role in substrate selectivity. In order to test whether this single point mutation could cause a similar alteration in another non-selective plant G3PAT, a sequence alignment was made between the squash amino acid sequence and a recently cloned oil palm G3PAT cDNA (Dr J Kroon PhD thesis 2000). The corresponding residue (L352) was identified in the oil palm sequence. Complimentary primers OPL352FF (5'gacaacatgagaaggtttgtggagcattctagtg3') and OPL352FR (5'cactagaatgctccacaaaccttctcatgttgtc3') were designed and used in a QuickChange ™ reaction with pET24a plasmid DNA containing a cDNA insert for the mature oil palm G3PAT. Following transformation and plasmid isolation the presence of the mutation was confirmed by DNA sequencing and the mRNA for the mutated protein was over-expressed and assayed. As is the case with the squash protein, making this single point mutation in oil palm G3PAT dramatically altered the substrate selectivity of the enzyme changing it from a non-selective one into one with a marked substrate preference for 18:1 ACP.

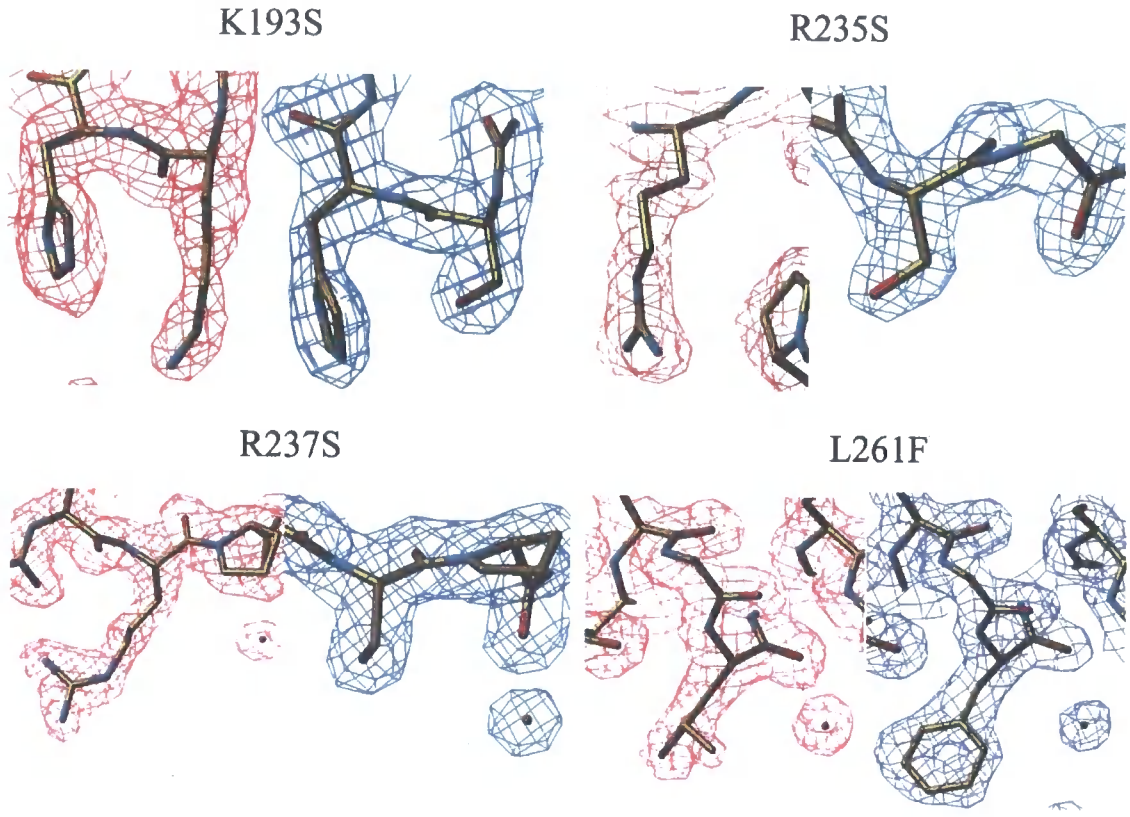
5.2.7 Structural Analysis of Mutant G3PAT Proteins, K193S, R235S, R237S and L261F.

Crystallisation of mutant G3PAT proteins K193S, R235S, R237S and L261F was, carried out in order to evaluate whether the loss of biological activity or altered substrate

specificity (L261F) was due to the modification of a critical residue or to a perturbation of the overall crystal structure of the protein. X-ray diffraction data at 1.9 Å resolution was collected for a crystal from each, and the three-dimensional structures compared to the wild-type G3PAT structure. The three-dimensional structure of all four mutant proteins was fully isomorphous and topographically equivalent to the wild-type G3PAT structure, with only minor movements observed in the peripheral loops of the R235S structure (**Figure 5.8**). Therefore the replacement of R235, R237 or K193 with a serine residue inactivates the enzyme without perturbing the three-dimensional structure (**Figure 5.8**). This structural data confirms the enzymatic and substrate modelling data, that these residues must be important in G3P binding.

The data collected for L261F mutant also confirms that the altered substrate specificity shown by the mutant protein was not caused by a conformational change of the overall structure. It could however result from the increased hydrophobic nature of the bulky aromatic phenylalinine residue enhancing the binding affinity for longer chain (18:1) acyl-ACP substrates compared to that for 16:0 substrates.

Figure 5.8 Structural analysis of squash G3PAT mutants K193S, R235S, R237S and L261F.



Comparison of the electron density maps in the region of the amino acid mutations K193S, R235S, R237S and L261F of squash G3PAT. The map highlighted in red is the wild type Q24a structure and the map highlighted in blue is the structure for the mutant enzyme. In all cases the three-dimensional structures of the mutant proteins are isomorphous with the wild type G3PAT structure. This data confirms that inactivation of mutant proteins K193S, R235S and R237S and the altered substrate specificity observed for the mutant protein L261F is not caused by significant perturbations of the crystal structure.

5.3Discussion

During this study recombinant squash G3PAT protein was over-expressed and used in crystallography studies to obtain a biologically active form of the protein which crystallised and gave X-ray diffraction data (Turnbull *et al.*, 2001a). Four mutant proteins were produced via site directed mutagenesis, each with an individual cysteine residue removed, which allowed heavy metal derivatives of the protein to be generated. These derivatives were used to resolve phasing problems in the diffraction data and to identify a histidine residue (His 279) which bound metal and was responsible for non-isomorphism between the data sets. The data collected for wild type protein and the heavy metal derivatives were combined and used in multiple isomorphous replacement analysis to solve the complete crystal structure of a soluble plant G3PAT for the first time (Turnbull *et al.*, 2001b).

G3P and acyl-ACP (palmitoyl) substrates were modelled into the structure and a number of residues were identified as important in the binding of the substrates. To confirm the importance of these residues site directed mutagenesis was carried out and mRNA for the mutant proteins were over-expressed and the new proteins assayed for biological activity and substrate selectivity. Acyl-ACP substrates generated using purified recombinant *E.coli* ACP were used for this analysis. In all four putative G3P (K193, H194, R235 and R237) and six putative acyl-ACP binding residues (E142, V166, T141S, R170, L213 and N257) were mutated and the mutant proteins over-expressed and assayed for alterations

in biological activity. Mutant proteins for three of the residues (K193S, R235S and R237S) predicted to be involved in G3P binding were all found to have completely lost biological activity even in assays carried out for extended time periods. Thus confirming their importance in G3PAT catalysis. Further evidence for the importance of these individual residues was obtained following crystallisation of the mutant proteins and comparison of the crystal structure with the wild-type Q24a enzyme. In all cases the mutant structures were fully isomorphous with the wild type structure confirming that the inactivation was a result of replacement of the residue and not by a perturbation of the structure. Mutant protein E142A was also completely inactive and R170S retained less than 5% of the level of wild type activity. Both of these residues lie in the hydrophobic cleft within domain II of the protein structure and are predicted to be within 5 Å of the modelled acyl chain. This lack of or greatly reduced biological activity again confirms the importance of these residues in G3PAT catalysis.

The point mutations L261F and P331S introduced during PCR cloning into over-expression vectors resulted in squash G3PAT with altered substrate specificity. Site directed mutagenesis was used to revert each of these residues individually and to show that the single mutation L261F, in the squash protein, resulted in a mutant protein with a marked preference for 18:1 acyl-ACP. Effectively turning a protein with a chilling sensitive profile into one typical of a protein from a chilling tolerant plant. Introduction of this mutation into a recombinant protein from oil palm gave the same result enforcing this

observation. Structural analysis of the mutant squash protein showed that the mutant structure was fully isomorphous with the wild type structure and therefore suggests that the altered specificity was a result of increased preference for longer chain acyl-ACP caused by the enhanced hydrophobicity of the introduced phenylalinine residue.

The putative assignments for substrate binding sites for mutagenesis studies were all based on the position of modelled substrates in relation to the position of the conserved $H(X)_4D$ domain within the structure. This study has confirmed residues involved with G3P binding but further experimentation is required to fully define the acyl-thiol binding site. Future work being undertaken in our laboratory will involve protein interaction studies between plant G3PATs and acyl-ACPs to aid further attempts to obtain co-crystals of G3PAT with a bound acyl-ACP substrate. The use of acyl-CoA and acyl-ACP affinity probes, which can be cross-linked into the active site of G3PAT will also be investigated to try to identify other important amino acid residues involved in acyl-ACP binding.

Chapter 6

General Discussion

At the commencement of this research many of the soluble enzyme components of the type II fatty acid synthetase (FAS) from plants had been cloned and biochemically characterised. In several cases, including enoyl-reductase and β-keto-reductase this characterisation included the determination of their crystal structure (Rafferty *et.al.*,1994 and 1995, Fisher *et al.*, 1999 and 2000) and the elucidation of their catalytic mechanism (Fawcett *et al.*,2000). The notable exceptions being deydratase (*fab* Z) and malonyl Co-A:ACP transacylase (*fab* D) two enzymes for which no cDNA from plants had ever been cloned. A putative partial cDNA clone for MCAT from maize and a full-length cDNA from *Brassica napus* has now been isolated as part of the research reported here. The *Brassica napus* cDNA clone was used to complement the temperature sensitive *E.coli* MCAT mutant (fabD-89) and restored its ability to grow at the non-permissive temperature of 42°C. This complementation of growth of the bacterial MCAT mutant with a plant cDNA offers proof of MCAT function and confirms that a plant MCAT cDNA has been cloned. This is the first cloning of a plant MCAT from any source. This together with the isolation of a cDNA clone for dehydratase (Doig PhD thesis 2001) means that cDNA clones for all of the enzymes in plant FAS are now available.

With the potential ability to produce large quantities of type II FAS proteins using bacterial and yeast expression systems a fully functional reconstitution of plant FAS now becomes a reality. This will allow the screening and targeting of potentially new agro-chemicals using the entire protein complement of plant FAS as a screen. Interaction studies using affinity techniques such as Biocore, SELDI and affinity chromatography can also be used to begin to look at the nature and strength of the interactions, the stochiometry and the interaction sites of the protein machinery involved in lipid biosynthesis in plants.

ACP the central component of FAS is crucial in all of these processes and the ability to over-express and purify it in its apo form from a number of different sources has enabled derivatisation procedures to be developed for the synthesis of acyl-ACPs. This synthesis is based on the conversion of apo-ACP to acyl-ACP using holo-ACP synthetase (HAS) and the appropriate acyl-CoA.

These procedures have eliminated t he need for the lengthy reductive procedures involving pH 9.0 hydroxylamine and DTT treatment, required for wild type ACP and the cumbersome acid precipitation steps involved in recovering the ACP following this treatment (Cronan and Klages 1981). Both chemical and enzymatic acylation reactions have been utilised and the acyl-ACP products have been compared and authenticated using conformational gel electrophoresis, electrospray mass spectrometry (ESMS) and

MALDItof mass spectrometry (MALDItof-MS). Enzymatic acylation reactions catalysed by holo-ACP synthetase (HAS), using saturated acyl-CoAs and purified recombinant apo-ACP resulted in quantitative specific acylation of ACP from *E.coli*, and *Brassica napus*. C4 and C8 acyl-ACP derivatives were produced and authenticated by MALDItof-MS and native-PAGE. Crystallisation trials on these derivatives have resulted in X-ray stable crystals from *E.coli* butyryl-ACP. Recombinant ACP cDNA clones where used as targets for site directed mutagenesis to produce *E.coli* ACP with new cysteine and methionine residues. These mutations were made to enable the generation of heavy metal crystal derivatives of *E.coli* ACP to enable the phase resolution of the diffraction data to be solved. The crystal structure of *E.coli* butyryl-ACP was determined for the first time during this research. A putative acyl-chain binding pocket large enough to accommodate up to eight carbon units has been identified within the structure (Roujeinikova *et al.*, in press). It is proposed that this pocket may function to stabilise the acyl-chain, by protecting the most reactive part of the growing chain from side reactions during the fatty acid synthesis cycle.

The enzymatic synthesis and crystallography of longer chain length acyl-ACPs and site directed mutagenesis of amino acids located around this pocket will be used in future work to investigate its putative role in acyl-binding.

Purified recombinant *E.coli* apo ACP and various chain length acyl-CoAs and fatty acids have been used in reactions catalysed by holo-ACP synthetase (HAS) and acyl-ACP synthetase (AcpS) to generate acyl-ACP (C16:0 and C18:1)substrates. Radioactive (^{14}C 16:0 and ^{3}H 18:1) acyl-ACP substrates generated in this way have enabled the development of dual substrate competitive assays for G3PAT activity to be developed. These competitive assays using the natural ACP substrate instead of the acyl-CoA substrate analogues previously used in G3PAT assays have been crucial in examining the substrate specificity (Slabas *et al.*, in press), and kinetic properties of G3PAT from both chilling sensitive (squash) and chilling tolerant (*Arabidopsis*) plants (Hayman *et al.*, in press).

Site directed mutagenesis of the four individual cysteine residues found in squash G3PAT enabled the non-isomorphism between the Patterson maps of the native crystal and that of the heavy metal crystal to be resolved. This allowed the crystal structure of the enzyme to be solved for the first time. Modelling of both the G3P and acyl-ACP (palmitoyl) substrates into the structure allowed putative substrate binding sites to be located and amino acid residues important in interacting with these substrates to be identified. Site directed mutagenesis of these residues was carried out to determine their importance in substrate binding. Four residues K193, H194, R235 and R237 predicted to be involved in G3P binding were mutated to serine and the mutant proteins assayed for G3PAT activity.

Proteins containing mutations, K193, R235 and R237 were all completely inactive confirming their importance in G3P binding. The H194 mutant retained 80% of wild type activity suggesting that this residue is not critical for substrate binding.

A putative acyl-ACP binding site was identified as a hydrophobic cleft within domain II of the structure and five of the amino acid residues (E142, V166, R170, L213 N257) which line this cleft and lie within 5 Å of the modelled ACP substrate were mutated. Of these mutant protein E142S was completely inactive and R170S retained <5% of wild type activity, confirming both of these residues as important in G3PAT catalysis. Proteins containing mutations of the remaining residues showed no altered biological activity suggesting these may not be critical residues. Other residues (H139, A167, G168, V189, H194, I229, M258) which also line this hydrophobic cleft and make van der Wall contact with the acyl-chain have been identified and future work using site directed mutagenesis will investigate their importance in acyl-ACP binding.

The use of dual substrate G3PAT assays with C16:0 and C18:1 acyl-ACP allowed the identification of a single amino acid point mutation (L261F) within squash G3PAT that altered it from a non selective enzyme, (typical of G3PAT from chilling sensitive plants), to one with a marked preference for 16:1 acyl-ACP (typical of that found in chilling tolerant plants).

A close correlation between the chilling tolerance of plants and the nature of the fatty acid at the *sn-1* position of their phosphatidylglycerols (PGs) was discovered by Murata *et al.*, (1983). Plants which are chilling sensitive have a non-selective G3PAT and incorporate large quantities of saturated (16:0) fatty acids into the *sn1* position of their PGs, whilst those which are chilling resistant preferentially incorporate unsaturated fatty acids (18:1). This increased level of unsaturated PG allows the membranes of chilling tolerant plants to remain fluid at low temperatures and hence less sensitive to chilling damage (Murata *et al.*,1992). Transgenic studies with tobacco using constructs containing selective [*Arabidopsis*] and non-selective (squash) G3PATs targeted to the chloroplast have supported this view (Murata *et al.*, 1992).

Understanding the nature of the acyl-ACP and glycerol-3-phosphate (G3P) binding sites and catalytic mechanism of plant G3PAT and the identification of critically important amino acid residues will aid in elucidating the mechanism of the substrate selectivity of this enzyme and hence its ability to effect the chilling tolerance of plants.

The advances made during this work have allowed the crystal structure of two of the components of lipid biosynthesis (*E.coli* acyl-ACP and squash G3PAT) to be resolved for the first time. The structural information obtained, together with the identification via substrate modelling and site directed mutagenesis of amino acid residues important in substrate binding and catalysis, have provided the basis for future investigations on how

these and other FAS components interact. These future studies will include, the co-crystallography of acyl-ACP with other soluble plant FAS components, now a step closer having determined the crystallisation conditions required for acyl-ACP.

To date most of the biochemical studies on the proteins involved in fatty acid biosynthesis, like those on components of other biosynthetic pathways, have tended to concentrate on the individual components in isolation. Whereas clearly in nature proteins rarely act alone, rather they interact as constituents of a pathway or protein complex. With the rapid advances made in recent years in proteomic (Aebersold and Goodlett 2001) and protein interaction studies (Grant and Husi 2001) using mass spectrometry, bioinformatics and affinity techniques, such as protein tagging (Rappsilber *et al.*, 2000) and protein chip techniques (Catimel *et al.*, 2001), the stage is now set for a detailed investigation of these interactions.

Building on the work reported in this thesis, future work will include elements of all of these approaches to enhance our knowledge of fatty acid and lipid biosynthesis in bacteria and plants and perhaps move us closer to understanding how the biochemistry of the cell really works.

References

Abbadi, A., Brummel, M., Schutt,B.S., Slabaugh, M.B., Schuch, R., and Spener, F., (2000). Reaction mechanism of recombinant 3-oxoacyl-(acyl carrier protein) synthase III from *Cuphea wrightii* embryo, a fatty acid synthase type II condensing enzyme. *Biochem. J.* **345**:153-160.

Aebersold, R., and Goodlett, D.R., (2001). Mass spectrometry in proteomics. *Chem. Rev.* **101**:269-295.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. *J.Mol.Biol.* **215**:403-410.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**: (17) 3389-3402.

Ashton, I.A., Abulnaja, K.O., Pallet, K.E., Cole, D.J and Harwood, J.L (1994). The mechanism of inhibition of fatty acid synthase by the herbicide diflufenican. *Phytochemistry.* **35**:(3)587-590.

Baldock, C., Rafferty, J.B., Sedelnikova, S.E., Baker, P.J., Stuitje, A.R., Slabas, A.R., Hawkes, T.R., and Rice, D.W. (1996). A mechanism of drug action revealed by structural studies of enoyl reductase. *Science* **274**:2107-2110.

Baldock, C., Rafferty, J.B., Stuitje, A.R., Slabas, A.R., and Rice, D.W. (1998). The X-ray structure of *Escherichia coli* enoyl reductase with bound NAD^+ at 2.1Angstrom resolution. *J, Mol. Biol.* **284**:1529-1546.

Banerjee, A., Dubnau,E., Quemard, A., Balasubramanian, V., Sun Um, K., Wilson, T., Collins, D., de Lisle, G., Jacobs, W.R., (1994). *inh*A a gene encoding a target for isoniazid and ethionamide in *Mycobacterium tuberculosis*. *Science.* **263**:227-230.

Battey, J.F., and Ohlrogge, J.B., (1990). Evolutionary and tissue-specific control of expression of multiple acyl-carrier protein isoforms in plants and bacteria. *Planta.* **180**:352-360.

Betrams, M., and Heinz, E., (1981). Positional specificity and fatty acid selectivity of purified sn-glycerol 3-phosphate acyltransferase from chloroplasts. *Plant Physiol.* **68**:653-657.

Bottomley, M.J., Robinson, R.C., Driscoll, P.C., Harlos, K., Stuart, D.I., Aplin, R.T., Clement, J.M., Jones, E.Y., and Dudgeon, T.J., (1994) Crystallisation and preliminary X-ray diffraction characterisation of both a native and selenomethionyl VLA-4 binding fragment of VCAM-1. *J. Mol. Biol.* **224**:464-468.

Birge, C.H., and Vagelos, P.R., (1972). Acyl Carrier Protein XVII Purification and properties of β-hydroxyacyl-ACP dehydrase. *J. Biol. Chem.* **247**:4930-4938.

Broun P., Shanklin, J., Whittle, E., and Somerville, C., (1998). Catalytic plasticity of fatty acid modification enzymes underlying chemical diversity of plant lipids. *Science* **282**:1315-1317.

Brown, A.P., Coleman, J., Tommey, A, Watson,M., Slabas, A.R., (1994). Isolation and characterisation of a maize cDNA that complements a 1-acyl sn-glycerol-3-phosphate acyltransferase mutant of *E.coli* and encodes a protein which has similarities to other acyltransferases. *Plant Mol. Biol.* **26**:211-223.

Browse, J. and Somerville, C. (1991). Glycerolipid synthesis: biochemistry and regulation.*Annu.Rev. Plant Physiol.Plant.Mol.Biol.* **42**:467-506.

Browse, J., Warwick, N., Somerville, C.R., and Slack, C.R. (1986). Fluxes through the prokaryotic and eukaryotic pathway of lipid synthesis in the 16-3 plant *Arabidopsis* thaliana. *Biochem. J.* **235**:25-31.

Bruck, F.M., Scuch, R., and Spener, F. (1994). Malonyl-CoA-acyl carrier protein transacylase from developing seeds of *Cuphea-lanceolata. J. Plant. Physiol.* **143**: 550-555.

Carreras, C.W., Gehring, A.M., Walsh, C.T., and Khosla, C., (1997). Utilization of enzymatically phosphopantetheinylated acyl carrier proteins and acetyl-acyl carrier proteins by the actinorhodin polyketide synthase. *Biochem.* **36**:11757-11761.

Catimel, B., Rothacker, J., and Nice, E., (2001). The use of biosensors for microaffinity purification: an integrated approach to proteomics. *J. Biochem. Biophys. Methods* **49**:289-312.

Caughey, I.,and Kegwick, R.G.O. (1982). The characteristics of some components of the fatty acid synthetase from the mesocarp of avocado (*Persea americana*). *Eur. J. Biochem.* **123**:553-561.

Chirgadze, N.Y., Briggs, S.L., McAllister, K.A., Fischl, A.S., and Zhao, G. (2000). Crystal structure of *Streptococcus pneumoniae* acyl carrier protein synthase: an essential enzyme in bacterial fatty acid biosynthesis. *EMBO Journal.* **19**:5281-587.

Chuman, L., and Brody, S., (1989). Acyl carrier protein is present in the mitochondria of plants and eucaryotic micro-organisms. *Eur. J. Biochem.* **184**:643-649.

Chung Suh, M., Schultz, D.J., and Ohlrogge, J.B., (1999). Isoforms of acyl carrier protein involved in seed specific fatty acid synthesis. *Plant Journal.* **17**:(6)679-688.

Clough, R.C., Matthis, A.L., Barnum, S.R., and Jackowski, J.G. (1992). Purification and characterisation of 3-ketoacyl carrier protein synthase III from Spinach - a condensing enzyme utilising acetyl coenzyme A to initiate fatty acid synthesis. *J. Biol. Chem.***267**:20992-20998.

Combet C., Blanchet C., Geourjon C. and Deléage G. (2000). NPS@: Network Protein Sequence Analysis. *TIBS.* **291**:147-150.

Cortes, J., Haydock, S.H., Roberts, G.A., Bevitt, D.J. and Leadlay, P.F. (1990). An unusually large multifunctional polypeptide in the erythromycin-producing polyketide synthase of *Saccharopolyspora - erythraea. Nature* **348**: (6297) 176-178.

Corpet, F., (1988). Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res.* **16:**10881-10890.

Cronan, J,E,Jr, and Klages,A,L., (1981). Chemical synthesis of acyl thioesters of acyl carrier protein with native structure. *PNAS*(USA). **78**:5440-5444.

Cronan, J,E,Jr and Roughan, P.G., (1987). Fatty acid specificity and selectivity of the chloroplast sm-glycerol-3-phosphate acyltransferase of the chilling sensitive plant Amaranthus lividus. *Plant Physiol.* **83**:676-679.

Crump, M.P., Crosby, J., Dempsey, C.E., Parkinson, J.A., Murray, M., Hopwood,D.A., and Simpson, T.J., (1997). Solution structure of the actinorhodin polyketide synthase acyl carrier protein from *Streptomyces coelicolor. Biochemistry* **36**:6000-6008.

Dawson, R.M.C., Elliott, D.C., Elliott, H.E., and Jomes, K.M., (1994). Data for Biochemical Research 3rd Ed. Oxford Scientific Publications.

de Boer, G.J., Pielage, G.J.A., Nijkamp, H.J.J., Slabas, A.R., Rafferty, J.B., Baldock, C., Rice, D.W., and Stuitje, A.R. (1999). Molecular genetic analysis of enoyl-acyl carrier protein reductase inhibition by diazaborine. *Molecular Microbiology* **31**:443-450.

de Silva, J., Loader, N.M., Jarman, C., Windust, J.H.C., Hughes, S.G., and Safford, R. (1990). The isolation and sequence analysis of two seed-expressed acyl carrier protein genes from *Brassica napus*. *Plant Mol. Biol.* **14**:537-548.

Donadio, S., Staver,M.J., McAlpine, J.B., Swanson,S.J., and Katz,L. (1991). Modular organisation of genes required for complex polyketide biosynthesis. *Science:* **252**:675-679.

Donadio, S., McAlpine, J.B., Sheldon,P.L., Jackson, M, and Katz,L (1993). An erythromycin analog produced by reprogramming of polyketide synthesis. *PNAS* (USA). **90**:7119-7123.

Dormann, P., Spener, F., and Ohlrogge, J.B., (1993). Characterisation of two acyl-acyl carrier protein thioesterases from developing *Cuphea* seeds for medium chain and oleoyl-acyl carrier protein. *Planta* **189**:425-432.

Dormann, P., Voelker, T.A., and Ohlrogge, J.B., (1995). Cloning and expression in *Escherichia coli* of a novel thioesterase from *Arabidopsis thaliana* for long chain acyl-acyl carrier proteins. *Arch. Biochim. Biophys.* **316**:612-618.

Eastmond, P.J., Dennis, D.T., and Rawsthorne, S., (1997). Evidence that malate / inorganic phosphate exchange translocator imports carbon across the leucoplast envelope for fatty acid synthesis in castor seed endosperm. *Plant Physiol.***114**:851-856

Ebel, J., Schmidt, W.E., and Loyal, R., (1984). Phytoalexin synthesis in soybean cells: elicitor induction of phenylalanine ammonia-lyase and chalcone synthase mRNAs and correlation with phytoalexin accumulation. *Arch. Biochem. Biophys.* **232**:240-248.

Eccleston, V.S., and Harwood, J.L., (1995). Solubilisation, partial purification and properties of acyl CoA:glycerol-3-phosphate acyltransferase from avocado (*Persea americana*) fruit mesocarp. *Biochim. Biophys Acta.* **1257:**1-10.

Elborough, K.M., Winz, R., Deka R.K., Markham, J.E., White, A.J., Rawsthorne, S., and Slabas, A.R., (1996). Biotin carboxyl carrier protein and carboxyltransferase subunits of the multi-subunit form of acetyl CoA carboxylase from *Brassica napus* – cloning analysis and expression during oil seed rape embryogenesis. *Biochem. J.* **315:** (1)103-112.

Elhussein, S.A., Miernyh, J.A., and Ohlrogge, J.B., (1988). Plant holo-(acyl carrier protein)-synthase. *Biochem. J.* **252:** 39-45.

Elovson, J., and Vagelos, P.R., (1968). Acyl carrier protein. X. Acyl carrier protein synthase. *J. Biol. Chem.* **243**:3603-3611.

Fawcett, T., Copse, C.L., Simon, J.W., and Slabas, A.R., (2000). Kinetic mechanism of the NADH-enoyl reductase from *Brassica napus*. *Febs Letts.* **484:**65-68.

Fernandez, M.D., and Lamppa, G.K., (1990). Acyl carrier protein (ACP) import into chloroplasts does not require phosphopantethein – evidence for a chloroplast holo ACP synthase. *Plant Cell.* **2**:195-206.

Ferri, S.R., and Toguri, T., (1997). Substrate specificity modification of the stromal glycerol-3-phosphate acyltransferase. *Arch. Biochem.Biophys.***337**:202-208.

Fisher, M., Kroon, J.T.M., Martindale, W., Stuitje, A.R., Slabas, A.R., and Rafferty, J.B. (2000). The x-ray structure of *Brassica napus* β-keto acyl carrier protein reductase and its implications for substrate binding and catalysis. *Structure* **8:**339-347.

Flugel, R.S., Hwangbo, Y., Lambalot, R.H., Cronan Jr, J.E., and Walsh, C.T., (2000). Holo-(acyl carrier protein) synthase and phosphopantetheinyl transfer in *Escherichia coli. J.Biol.Chem.* **275**:(2)959-968.

Frentzen, M., (1990). Comparison of certain properties of membrane bound and solubilized acyltransferase activities of plant microsomes. *Plant Sci.* **69**:39-48.

Frentzen, M., (1993). Acyltransferases and triacylglycerols. *In Lipid Metobolism in Plants. Ed T.S.Moore, Jr. CRC Press.* 195-231.

Frentzen, M., Heinz, E., Mckeon, T.A., and Stumpf, P.A., (1983). Specificities and selectivities of glycerol-3-phosphate acyltransferase and monoacylglycerol-3-phosphate acyltransferase frompea and spinach chloroplasts. *Eur.J.Biochem.* **129:**629-636.

Frentzen, M., Nishida, I., and Murata, N., (1987). Properties of the plastidial acyl-(acyl-carrier-protein):glycerol-3-phosphate acyltransferase from the chilling-sensitive plant squash (*Cucurbita moschata*).*Plant Cell Physiol.* **28**:1195-1201.

Frentzen, M., Neuburger, J., Joyard, J., and Douce, R., (1990). Intra-organelle localization and substrate specificities of the mitochondrial acyl-CoA:*sn*-glycerol-3-phosphate *O*-acyl-transferase and acyl-CoA:1-acyl-*sn*-glycerol-3-phosphate *O*-acyl-transferase from potato tubers and pea leaves. *Eur.J.Biochem.* **187**:395-402.

Fritz, M., Heinz, E., and Wolter, F.P., (1995). Cloning and sequencing of full length cDNA coding for a sn-glycerol-3-phosphate acyltransferase from *Phaseolus vulgaris*. *Plant Physiol.***107**:1039-1040.

Gangar. A., Karande, A.A., and Rajasekharan, R., (2001). Isolation and localization of a cytosolic 10 S triacylglycerol biosynthetic multienzyme complex from oleaginous yeast. *J.Biol.Chem.* **276** (13): 10290-10298.

Gavel, Y., and von Heinje, G., (1990). A conserved cleavage site motif in chloroplast transit peptides. *FEBS Lett.* **261**:455-458.

Grant, S.G.N., and Husi, H., (2001). Proteomics of multiprotein complexes: answering fundamental questions in neuroscience. *Trends Biotech.***19**: (supp) 49-54.

Gulliver, B.S., and Slabas, A.R. (1994). Acetoacetyl-acyl carrier protein synthase from avocado – its purification, characterisation and clear resolution from acetyl CoA:ACP transacylase. *Plant Mol. Biol.* **25**:179-191.

Gunstone, F.D., Harwood, J.L., and Padley, F.B., (1994). The lipid handbook (second edition) *Pub. Chapman and Hall,* ISBN 0412433206

Hansen, L., (1987). Three cDNA clones for barley leaf acyl carrier proteins I and III. *Carlsberg Res.Commun.* **52**:381-92.

Harder, M.E., Ladenson, R.C., Schimmel, S.D., and Silbert, D.F., (1974). Mutants of *Escherichia coli* with temperature sensitive malonyl CoA- acyl carrier protein transacyase. *J.Biol.Chem.* **249**:7468-7475.

Harwood, J.L., (1979). The synthesis of acyl lipids in plant tissues. *Prog.Lipid Res.***18**: 55-86.

Harwood, J.L. (1980) Plant acyl lipids: Structure, distribution and analysis. *In The Biochemistry of Plants, Vol 4, ed Stumpf, P.K., Academic press NY.* 2-56.

Harwood, J.L. (1996). Recent advances in the biosynthesis of plant fatty acids. *Biochim. Biophys. Acta.* **1301**:7-56.

Hayman, M.W., Fawcett, T., Schierer, T.F., Simon, J.W., Kroon, J.T.M., Gilroy, J.S., Rice, D.W., Rafferty, J., Turnbull, A.P., Sedelnikova, S.E., and Slabas, A.R., (2000). Mutagenesis of squash *(Cucurbita moschata)* glycerol-3-phosphate acyltransferase to produce an enzyme with altered substrate selectivity. *Biochem. Soc.Trans.* **28**:(6) 680-681.

Heath, R.J., and Rock, C.O., (1998). A conserved histidine residue is essential for glycerolipid acyltransferase catalysis. *J.Bacteriol.* **180**:(6)1425-1430.

Heath, R.J., Rubin, J.R., Holland, D.R., Zhang, E, Snow, M.E., and Rock, C.O. (1999). Mechanism of triclosan inhibition of bacterial fatty acid synthesis. *J.Biol.Chem.* **274**:(16)11110-11114.

Heath, R.J., Li, J., Roland, G.E., and Rock, C.O., (2000). Inhibition of the *Staphylococcus aureus* NADPH - dependant enoyl-acyl carrier protein reductase by triclosan and hexachlorophene. *J. Biol. Chem.* **275**:(7)4654-4659.

Heinz, E., and Roughan, P.G., (1983). Similarities and differences in lipid metabolism of chloroplasts isolated from 18:3 and 16:3 plants. *Plant Physiol.* **72**:273-279.

Hellyer, A., Leadlay, P.F., and Slabas, A.R. (1992). Induction, purification and characterisation of acyl-ACP thioesterase from developing seeds of oil seed rape (*Brassica napus*). *Plant Mol. Biol.***20**:763-780.

Hendrickson, W.A., Horton, J.R., and LeMaster, D.M., (1990). Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure.
EMBO Journal: **9**:(5)1665-1672.

Higgins, D.G., and Sharp, P.M., (1989). Fast and sensitive multiple sequence alignments on a microcomputer. *Comput. Appl. Biosci.* **5**:151-154.

Hlousek-Radojcic, A., Post-Beittenmiller, D., and Ohlrogge, J.B., (1992). Expression of constitutive and tissue specific acyl carrier protein isoforms in *Arabidopsis*. *Plant Physiol.* **98**:206-214.

Hoj, P.B., and Mikkelsen, J.D., (1982). Partial separation of individual enzyme activities of an ACP dependant fatty acid synthetase from barley chloroplasts. *Carlsberg Res. Commun.* **47**:119-141.

Hoj, P.B., and Svedson,I., (1983). Barley acyl carrier protein. Its amino acid sequence and assay using malonyl CoA:ACP transacylase. *Carlsberg Res. Commun.* **48**: 285-303.

Hoj, P.B., and Svedson,I., (1984). Barley chloroplasts contain 2 acyl carrier proteins coded for by different genes. *Carlsberg Res. Commun.* **49**:483-492.

Holak, T.A., Kearsley, S.K., Kim,Y., and Prestegard, J.H., (1988a). Three dimensional structure of acyl carrier protein determined by NMR pseudoenergy and distance geometry calculations. *Biochemistry* **27**:6235-6142.

Holak, T.A., Nilges, M., Prestergard, J.H., Gronenborn, A.M., and Glore, G.M., (1988b). Three-dimensional structure of acyl carrier protein in solution determined by nuclear magnetic resonance and the combined use of dynamical simulated annealing and distance geometry. *Eur. J. Biochem.* **175**:9-15.

Holak, T.A., and Prestergard, J.H., (1986). Secondary structure of acyl carrier protein as derived from two-dimensional H-1-NMR spectroscopy. *Biochemistry* **26**:5766-5774.

Ishizaki, O., Nishida, I., Agata, K., Eguchi, G., and Murata, N., (1988). Cloning and nucleotide sequence of cDNA for the plastid glycerol-3-phosphate acyltransferase from squash. *FEBS Lett.* **238:**424-430.

Jackowski, S., and Rock, C.O., (1987). Acetoacetyl-acyl carrier protein synthetase: a potential regulator of fatty acid biosynthesis in bacteria. *J.Biol.Chem.***262**:7927-7931.

Jaworski, J.G., Clough, R.C., and Barnum, S.R., (1989). A crulenin insensitive short chain 3-ketoacyl carrier protein in *Spinach oleracea* leaves. *Plant Physiol.* **90**:41-44.

Johnson, B.H., and Hecht, M.H., (1994). Recombinant proteins can be isolated from *E.coli* cells by repeated cycles of freezing and thawing. *Biotechnology* **12**:1357-1360.

Johnson, T.C., Schneider, J.C., and Somerville, C., (1992). Nucleotide sequence of acyl–acyl carrier protein:glycerol-3-phosphate acyltransferase from cucumber. *Plant physiol.* **99:**771-772.

Joyard, J., and Douce, R., (1977). Site of synthesis of phosphatidic acid and diacylgleycerol in spinach chloroplasts. *Biochim. Biophys. Acta.* **486**:273-85.

Joyard, J., and Stumpf, P.K., (1982). Synthesis of long chain acyl-~CoA in chloroplast envelope membranes. *Plant Physiol.* **67**:250-256.

Kang, F., and Rawsthorne, S., (1996). Metabolism of glucose–6-phosphate and utilization of multiple metabolites for fatty acid synthesis by plastids from developing oilseed rape embryos. *Planta* **199**:321-327.

Kater, M.M., Koningstein, G.M., Nijkamp. H.J.J., and Stuitje, A.R., (1994). cDNA cloning and expression of *Brassica napus* enoyl-acyl carrier protein reductase in *Escerichia coli. Plant. Mol. Biol.* **17**:895-909.

Kaufman, A.J., and Ruebusch, R.J., (1990). Oleochemicals a look at world trends. *Inform* **1:**(12)1034-1048.

Ke, J., Behal, R.H., Back, S.L., Nikolau, B.J., Wurtele, E.S., and Oliver, D.J., (2000). The role of pyruvate dehydrogenase and acetyl-coenzyme A synthetase in fatty acid synthesis in developing *Arabidopsis* seeds. *Plant Physiol.* **123**:497-508.

Kennedy,E.P., (1961). Biosynthesis of complex lipids. *Fed. Proc. Fed. Am. Soc. Exp. Biol.* **20**:934-940.

Kenrick, J.R., and Bishop, D.G., (1986). The fatty acid composition of phosphotidylglycerol and sulfoquinovosyldiacylglycerol of higher plants in relation to chilling sensitivity. *Plant Physiol.* **81**:946-949.

Kim,Y., and Prestergard, J.H., (1989). A dynamic model for the structure of acyl carrier protein in solution. *Biochemistry* **28:** 8792-8797.

Kim,Y., Ohlrogge, J.B., and Prestergard, J.H., (1990). Motional effects on NMR structural data – comparison of spinach and *Escherichia coli* acyl carrier proteins. *Biochem.Pharmacol.* **40** :7-13

Kolattukudy, P.E., (1987). Lipid-derived defensive polymers and waxes and their role in plant–microbe interaction. *In The Biochemistry of Plants a Comprehensive Treatise. Volume 9: Lipids Structure and Function. eds.* P.K.Stumpf and E.E.Conn. *Academic Press.*

Kuo, T.S., and Ohlrogge, J.B., (1984). The primary structure of spinach acyl carrier protein. *Arch. Biochem. Biophys.* **234**:290-296.

Laemmeli, U.K., (1970). Cleavage of structural proteins during the assembly of the bacteriophage T4. *Nature* **227**:680-685.

Lerouge. P., Roche, P., Faucher, C., Maillet, F., Truchet, G., Prome, J.C., and Denarie. J., (1990). Symbiotic host specificity of *Rhizobium meliloti* is determined by a sulphated and acylated oligosaccharide signal. *Nature* **344**:781-784.

Lewin,T.M., Wang,P., and Coleman, R.A., (1999). Analysis of amino acid motifs diagnostic for the sn-glycerol-3-phosphate reaction. *Biocemistry* **38**:5764-5771.

Lightner, V.A., Larson, T.J., Tailleur, P., Kantor, G.D., Raetz, C.R.H., Bell, R.M., and Modrich, P., (1980). Membrane phospholipid synthesis in *Escherichia coli*: cloning of a structural gene (*plsB*) of the *sn*-glycerol-3-phosphate acyl-transferase. *J. Biol.Chem.* **255**:9413-9420.

Liu, D., and Post-Beittenmiller, D., (1995). Discovery of an epidermal stearoyl-acyl carrier protein thioesterase. *J. Biol. Cem.* **270**:16962-16969.

Magnuson, K., Oh, W., Larson, T.L., and Cronan Jr, J.E., (1992). Cloning and nucleotide sequence of the *fab*D gene encoding malonyl coenzyme A-acyl carrier protein transacylase from *Escherichia coli*. *FEBS Letts.* **299**:262-266.

Magnuson, K., Jackowski, S., Rock, C.O., and Cronan, J.E., Jr. (1993). Regulation of fatty acid biosynthesis in *Escerichia coli*. *Microbiol.Rev.* **57**:522-542.

Majerus P.W., Alberts A.W., and Vagelos P.R. (1964). The acyl carrier protein of fatty acid synthesis: purification, physical properties and substrate binding site. *PNAS.* (USA) **51**:1231-1238.

Majerus P.W., Alberts A.W., and Vagelos P.R. (1965). Acyl carrier protein IV. The identification of the 4" phosphopantetheine as the prosthetic group of the acyl carrier protein. *PNAS.* (USA) **53**:410-417.

Marck, C., (1988). A C-program for the fast analysis of DNA and protein sequences on the Apple Macintosh family of computers. *Nucleic Acids Res.* **16**:10829-10832.

Masterson C., Wood, C., and Thomas, D.R., (1990). L-acetylcarnitine, a substrate for chloroplast fatty acid biosynthesis. *Plant Cell Environment* **13**:755-765.

Mayo, K.H., Tyrell, P.M., and Prestergard, J.H., (1983). Acyl carrier protein from *Escerichia coli* I. Aspects of the solution structure as evidenced by proton nuclear overhauser experiments at 500 MHz. *Biochemistry* **22**:4485-4493.

McKeon , R.T.A., and Stumpf, P.K., (1982). Purification and characterisation of the stearoyl-acyl carrier protein desaturase and the acyl-acyl carrier protein thioesterase from maturing seed of safflower . *J.Biol.Chem.* **136**:501-508.

Mekhedov, S., de Ilárduya, O., and Ohlrogge., J.(2000). Towards a functional catalogue of the plant genome: a survey of genes for lipid biosynthesis, *Plant Physiol.***122**:389-401.

Millar, A.A., Smith, M.A., and Kunst,L., (2000). All fatty acids are not equal: discrimination in plant membrane lipids. *Trends in Plant Science* **5**(3) 95-101.

Moche, M., Dehesh, K., Edwards, P., and Lindqvist, L. (2001). The crystal structure of β-ketoacyl-acyl carrier protein synthase II from *Synechocystis sp.* At 1.54 Å resolution and its relationship to other condensing enzymes. *J. Mol. Biol.* **305:**491-503.

Murata, N., Sato, N., Takahashi, Y., and Hamazaki, Y., (1982). Compositions and positional distributions of fatty acids in phospholipdids from leaves of chilling sensitive and chilling resistant plants. *Plant Cell Physiol.* **23**:1071-1079.

Murata, N, Ishizaki-Nishizawa, Q., Higashi, S., Hayashi, H., Tasaka, Y., and Nishida, I., (1992). Genetically engineered alteration in the chilling sensitivity of plants. *Nature* **356**:710-713.

Murata, N., and Wada, H., (1995). Acyl lipid desaturases and their importance in the tolerance and acclimatization to cold of cyanobaceria. *Biochem.J.* **308:**1-8.

Nishida, I., Frentzen, M., Ishizaki, O., and Murata, N., (1987). Purification of isomeric forms of acyl-(acyl carrier protein):glycerol-3-phosphate acyltransferase from greening squash cotyledons. *Plant Cell Physiol.* **28**:1071-1079.

Nishida, I., Tasaka,Y., Shiraishi, H., and Murata, N., (1993). The gene and RNA for the precursor to the plastid located glycerol-3-phosphate acyltransferase of *Arabidopsis thaliana. Plant.Mol.Biol.* **21**:267-277.

Ohlrogge, J.B., (1997). Regulation of fatty acid synthesis. *Annu Rev. Plant Physiol. Plant Mol. Biol.* **48:**109-136.

Ohlrogge, J.B., and Browse (1995). Lipid biosynthesis. *Plant Cell* 7:957-970.

Ohlrogge, J.B., and Kuo, T.M., (1984). Control of lipid-synthesis during soybean seed development-enzymic and immunochemical assay of acyl carrier protein. *Plant Physiol.* **74**:622-625.

Ohlrogge, J.B., and Kuo, T.S., (1985). Plants have isoforms for acyl carrier protein that are expressed differently in different tissues. *J.Biol.Chem.* **260:**8032-8037.

Oswood, M.C., Kim, Y., Ohlrogge, J.B., and Prestegard, J.H., (1997). Strucutural homology of spinach acyl carrier protein and *Esherichia coli* acyl carier protein based on NMR data. *Proteins* **27**:131-143.

Pearson, W.R., and Lipman, D.J., (1988). Improved tools for biological sequence analysis. *PNAS* (USA) **85**:2444-2448.

Pollard, M.R., Anderson, L., Fan, C., Hawkins, D.J., and Maelor-Davies, H., (1991). A specific acyl-ACP thioesterase implicated in medium chain fatty acid production in immature cotyledons of *Umbellularia californica. Arch. Biochem. Biophys.* **284**:306-312.

Post-Beittenmiller, D., Jaworski, J.G., and Ohlrogge, J.B., (1991). *In vivo* pools of free and acylated acyl carrier proteins in spinach evidence for sites of regulation of fatty acid biosynthesis. *J. Biol. Chem.* **266**:1858-1865.

Prescott, D.J., and Vagelos, P.R., (1972). Acyl carrier protein. *Adv. Enzymol.* **36**: 269-311.

Promega (1996). Protocols and Applications Guide (third edition) *Eds.* Doyle,K and Miles,J. *Promega Corporation USA. ISBN 1-882274-57-1.*

Rafferty, J.B., Simon, J.W., Stuitje, A.R., Slabas, A.R., Fawcett, T., and Rice, D.W., (1994) Crystallization of the NADH- specific enoyl acyl carrier protein reductase from *Brassica napus. J. Mol. Biol.* **237**:240-242.

Rafferty, J.B., Simon, J.W., Baldock. C., Artymuick, P.J., Baker, P.J., Stuitje, A.R., Slabas, A.R., and Rice, D.W., (1995) Common themes in redox chemistry emerge from the X-ray structure of oilseed rape (*Brassica napus*) enoyl acyl carrier protein reductase. *Structure*: **3**(9) 927-938.

Rafferty, J.B., Fisher, M., Langridge, S.J, Martindale, W., Thomas, N.C., Simon, J.W., Bithell, S., Slabas, A.R., and Rice, D.W., (1998). Crystallization of the NADP-dependant β-keto acyl carrier protein reductase from *E.coli. acta Crystallographica* **D54**: 427-429.

Rawsthorne, S., (2002). Carbon flux and fatty acid synthesis in plants. *Prog. Lipid Res.* **41**:182-196.

Rappsilber, J., Siniossoglou, S., Hurt, E., and Mann, M., (2000). A generic strategy to analyse the spatial organisation of multi-protein complexes by cross-linking and mass spectrometry. *Anal. Chem.* **72**:267-275.

Regos, J., Zak, O., Solf, R., Vischer, W.A., and Weirich, E.G., (1979). Triclosan action. *Dermatologica* (Basel) **158:** 72-79.

Roughan, P.G., (1985). Phosphatidylglycerol and chilling sensitivity in plants. *Plant Physiol.* **77**:740-746.

Roughan, P.G., (1995). Acetate concentrations in leaves are sufficient to drive in *vivo* fatty acid synthesis at maximum rates. *Plant Science* **107**:49-55.

Roughan, P.G., (1997). Stromal concentrations of coenzyme A and its esters are insufficient to account for rates of chloroplast fatty acid synthesis; evidence for substrate channelling within the chloroplast fatty acid synthase. *Biochem.J.* **327**:267-273.

Roughan, P.G., and Nishida, I., (1990). Concentrations of long-chain acyl-acyl carrier proteins during fatty acid syntheis by chloroplasts isolated from pea (*Pisum sativum*), safflower (*Carthamus tinctoris*) and amaranthus (*Amaranthus lividus*) leaves. *Arch. Biochem. Biophys.* **276**:38-46.

Roughan, P.G. and Ohlrogge, J.B., (1996) Evidence that isolated chloroplasts contain an integrated lipid-synthesising assembly that channels acetate into long chain fatty acids. *Plant Physiol.* **110**:1239-1247.

Roughan, P.G. and Slack, C.R. (1982) Cellular organisation of glycerolipid metabolism. *Ann. Rev. Plant Physiol.* **33**:97-132.

Ruch, F.E., and Vagelos, P.E., (1973a). The isolation and general properties of *Escherichia coli* malonyl coenzyme A-acyl carrier protein transacylase. *J. Biol. Chem.* **248**:8086-8094.

Ruch, F.E., and Vagelos, P.E., (1973b). Characterization of a malonyl-enzyme intermediate and identification of the malonyl binding site in malonyl coenzyme A-acyl carrier protein transacylase of *Escherichia coli. J. Biol. Chem.* **248:**8095-8106.

Saiki,R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B., and Erlich, H.A., (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239**:487-491.

Sambrook, J., Fritsch, E.F., and Maniatis, T., (1989). *Molecular Cloning a Laboratory Manual. Second Ed.* Cold Spring Harbor Press.

Safford, R., Windust, J.H.C., Lucas, C., De Silva, J., James, C.M., Hellyer, A., Smith, C.G., Slabas, A.R., and Hughes, S.G., (1988). Plastid localised acyl-carrier protein of *Brassica napus* is encoded by a distinct multigene family. *Eur. J. Biochem.* **174**:287-295.

Salas, J.J., Sanchez, J., Ramli, U.S., Manaf, A.M., Williams, M., and Harwood, J.L. (2000) Biochemistry of lipid metabolism in olive and other oil fruit. *Progress in Lipid Research.* **39:**151-180.

Sasaki,Y., Konishi,T., and Nagano,Y., (1995). The compartmentation of acetyl – coenzyme A carboxylase in plants. *Plant Physiol.* **108**:445-449.

Schmid, K.M., and Ohlrogge, J.B., (1990). A root acyl carrier protein from spinach is also expressed in leaves and seeds. *Plant Mol.Biol.* **15**:765-778.

Schult, W., Topfer, R., Stracke, R., Schell, J., Martini, N., (1997) Multi-functional acetyl CoA carboxylase from *Brassica napus* is encoded by a multi-gene family: indication for plastidic localisation of at least one isoform. *PNAS* (USA) **94**:(7) 3465-3470.

Schultz.R., Ebel,J., and Hahlbrock. K., (1982). Partial purification of β-keto acyl carrier protein synthetase from a higher plant. *FEBS Letts.* **140**:207-209.

Schutt, B.S., Brummel, M., Schuch, R., and Spener, F., (1998). The role of acyl carrier protein isoforma from *Cuphea lanceolata* seeds in the *de-novo* biosynthesis of medium chain fatty acids. *Planta* **205**:263-268.

Serre, L., Swenso, L., Green, R., Wei,Y., Verwoert, I.I.G.S., Verbee, E.C., Stuitje, A.R., and Derewenda, Z.S., (1994). Crystallization of the malonyl Coenzyme A-acyl carrier protein transacylase from *Escherichia coli*. *J. Mol. Biol.* **242**:99-102.

Serre, L., Verbree, E.C., Dauter, Z., Stuitje, A.R., and Derewenda, Z.S., (1995). The *Escherichia coli* Malonyl-CoA:Acyl carrier protein transacylase at 1.5 Å resolution. Crystal structure of a fatty acid synthase component. *J.Biol.Chem.* **270**:12961-12964.

Sheldon, P.S., Kegwick, R.G.O., Smith, C.G., Sidebottom,C., and Slabas, A.R. (1992). 3-oxoacyl-[ACP] reductase from oilseed rape (*Brassica napus*). *Biochim. Biophys. acta.* **1120**:151-159.

Sherman, D.H., Malpartida,F., Bibb, M.J., Kieser, H.M., Bibb, M.J., and Hopwood, D.A. (1989) Structure and deduced function of the granaticin producing polyketide synthase gene cluster of *Streptomyces violaceoruber* Tu22. *EMBO Journal* **8**:(9) 2717-2725.

Shimakata,T. and Stumpf P.K. (1982) The prokaryotic nature of the fatty acid synthetase of developing *Carthamus tinctorius* L. (safflower) seeds. *Arch. Biochem. Biophys.* **217**: 144-154.

Shimakata,T. and Stumpf P.K. (1982) Purification and characteristics of β-keto (acyl carrier protein) reductase, β-hydroxyacyl (acyl carrier protein) dehydrase and enoly (acyl carrier protein reductase from *Spinach oleracea* leaves. *Arch. Biochem. Biophys.* **218**:77-99.

Simon, J.W., and Slabas, A.R. (1998) cDNA cloning of *Brassica napus* malonyl-CoA:ACP transacylase (MCAT) (*fab*D) and complementation of an *E.coli* MCAT mutant. *FEBS Letts.* **435**:204-206.

Simoni, R.D., Criddle, R.S., and Stumpf, P.K., (1967). Fat metabolism in higher plants XXXI. Purification and partial properties of plant and bacterial acyl carrier protein. *Journal of Biol. Chem.* **242**:573-581.

Slabas, A.R., Sidebottom, C.M., Hellyer, A., Kessell, R.M.J., and Tombs, M.P., (1986). Induction, purification and characterisation of NADH-specific enoyl-acyl carrier protein reductase from developing seed of oil seed rape (*Brassica napus*). *Biochim. Biophys. acta* **877**:271-280.

Slabas, A.R., Harding, J., Hellyer, A., Roberts, P., and Bambridge, H.E., (1987). Induction, purification and characterisation of acyl carrier protein from developing seeds of oil seed rape (*Brassica napus*) *Biochem. Biophys. Acta.* **921**:50-59.

Slabas, A.R., and Fawcett, T., (1992). The biochemistry and molecular biology of plant lipid biosynthesis. *Plant Mol. Biol.* **19**: 169-191.

Slabas, A.R., and Brough, C.L., (1997). Specificity of acyltransferases and their genetic manipulation for environmental adaptation and development of new oils. *The Society for Experimental Biology (SEB1032)* 149-156.

Sol,J., and Roughan, G., (1982). Acyl-acyl carrier protein pool sizes during steady-state fatty acid synthesis by isolated spinach chloroplasts. *FEBS Letts.* **146**:189-192.

Sol,J., and Tien, R., (1998). Protein translocation into and across the chloroplastic envelope membranes. *Plant Mol.Biol.* **38**:191-207.

Sommerville, C.R., and Browse, J., (1991) Plant lipids: metabolism, mutants and membranes. *Science* **252**:80-87.

Sommerville, C.R., (1993) Production of industrial materials in trangenic plants. *Philosophical Transactions of the Royal Society* **342**: 251-257.

Stapleton, S.R., and Jaworski, J.I., (1984). Characterisation and purification of malonyl Coenzyme A:[acyl-carrier-protein] transacylase from spinach and *Anabaena variabilis*. *Biochim.Biophys. Acta* **794**:240-248.

Studier, F.W., and Moffat, B.A., (1986). Use of bacteriophage – T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* **189**:113-130.

Studier, F.W., Rosenberg, A.H., Dunn, J.J., and Dubendorff, J.W., (1990). Use of T7 RNA polymerase to direct expression of cloned genes. *Methods. Enzymol.* **185**:60-89.

Stuitje, A.R., Verbree, E,C., van der Linden, K.H., Rafferty, J.B., Rice, D.R., and Slabas, A.R., (1998). Mutational analysis of plant enoyl-ACP reductase in *Escherichia coli. In Advances in Plant Lipid Research.* Eds. Sanchez *et al.*, University of Seville Press. 77-80.

Stumpf, P.K., (1980). Biosynthesis of saturated and unsaturated fatty acids. *In The Biochemsitry of Plants* ed. P.K.Stumpf and E.E.Conn. **4**:177-203.

Thompson, J.D., Hihhins, D.G., and Gibson, T.J., (1994). CLUSTAL W improving the sensitivity of progressive multiple sequence alignment through sequence weighting. *Nucleic Acids Res.* **11**:(22) 4673-4680.

Truchet, G., Roche, P., Lerouge. P., Vasse, J., Camut, S., deBilly,F., Prome, J.C., and Denarie. J., (1991) Sulphated lipo-oligosaccharide signals of *Rhizobium meliloti* elicit root nodule organogenesis in *alfalfa. Nature* **351**:670-673.

Turnbull, A.P., Rafferty, J.B., Sedelnikova, S.E., Slabas, A.R., Schierer, T.P., Kroon, J.T.M., Nishida, J., Murata, N., Simon, J.W., and Rice, D.W., (2001a). Crystallization and preliminary X-ray analysis of the glycerol-3-phosphate 1-acyltransferase from squash (*Cucurbita moschata*). *Acta. Crystallog. section D-Biological Crystallography* **57**: 451-453.

Turnbull, A.P., Rafferty, J.B., Sedelnikova, S.E., Slabas, A.R., Schierer, T.P., Kroon, J.T.M., Simon, J.W., Fawcett, T., Nishida, I., Murata, N., and Rice, D., (2001b). Analysis of the structure, substrate specificity, and mechanism of squash glycerol-3-phosphate (1)-acyltransferase. *Structure* **9**:347-353.

Turnham, E., and Northcote, D.H., (1983). Changes in the activity of acetyl-CoA carboxylase during rapeseed formation. *Biochem. J.* **212**: 223

Vanaman,T.C., Wakil,S.J., and Hill, R.L., (1986). The complete amino acid sequence of the acyl carrier protein of *Escherichia coli*. *J. Biol. Chem.* **243**:6420-6431.

VanDune,G.D., Standaert,R.F., Karplus,P.A., Schreiber,S.L and Clardy,J., (1993). Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. *J. Mol. Biol.* **229**:105-124.

Verwoert,I.I.G.S., Verbree, E.C., van der Linden, K.H., Nijkamp, H.J.J., and Stuitje, A.R., (1992b). Cloning, nucleotide sequence, and expression of the *Esherichia coli fab*D gene encoding malonyl Coenzyme A – ACP transacylase. *J. Bacteriol.* **174:**2851-2857.

Verwoert,I.I.G.S., Vanderlinden, K.H., Nijkamp, H.J.J and Stuitje, A.R., (1994). Developmental specific expression and organelle targeting of the *Escherichia coli fab*D gene, encoding malonyl coenzyme A-acyl carrier protein in transgenic rape and tobacco seeds. *Plant Mol. Biol.* **26**:189-202.

Verwoert, I.I.G.S., Brown, A., Slabas, A.R., and Stuitje, A.R., (1995). A *Zea mays* GTP-binding protein of the ARF family complements an *Escherichia coli* mutant with a temperaturesensitive malonyl-coenzyme A-acyl carrier protein transacylase. *Plant Mol. Biol.* **27**:629-633.

Voelker, T.A., Worrell, A.C., Anderson, L., Bleibaum, J., Fran, C., Hawkins, D.J., Radke, S.E., and Maelor-Davis, H., (1992). Fatty acid biosynthesis redirected to medium chains in transgenic oil seed plants. *Science* **257**:72-74.

Von Heijne, G., Steppuhn, J., and Herman, S.G., (1989). Domain structure of mitochondrial and chloroplast targeting peptides. *Eur. J. Biochem.* **180**:535-545.

Wakil, S.J., Stoops,J.K., and Joshi, V.C., (1983). Fatty acid synthesis and its regulation. *Annu. Rev. Biochem.* **52**:537-539.

Weber, S., Wolter, F., Buck, F., Frentzen, M., and Heinz, E., (1991). Purification and cDNA sequencing of an oleate-selective acyl ACP *sn*-glycerol-3-phosphate acyltransferase from pea chloroplasts. *Plant Mol. Biol.* **17**:1067-1076.

Weeks, G., and Wakil, S.J., (1968). Studies on the mechanism of fatty acid synthesis. Preparation and general properties of the enoyl acyl carrier reductase from *Escherichia coli*. *J.Biol.Chem.* **243**:1180-1189.

Wolter, F.P., Schmidt,R., and Heinz, E., (1992). Chilling sensitivity of *Arabidopsis thaliana* with genetically engineered membrane lipids. *EMBO J.* **11**:4685-4692.

Wurtele, E.S., Behal, R.H., Cui, X., Ke, J., Johnson, J.L., Nikolau, B.J., Oliver, D.J., and Schnable, P.S., (1998). Molecular biology of acetyl CoA generation. *In Advances in Plant Lipid Research.* Eds. Sanchez *et al.*, University of Seville Press. 54-56.

Xu, G.Y., Tam, A., Lin,L., Hixon,J., Fritz, C.C., and Powers, R., (2001). Solution structure of *B.subtilis* acyl carrier protein. *Structure* **9**: 277-287.

Zhou, P., Florova, G., and Reynolds, K.A., (1999). Polyketide synthase acyl carrier protein (ACP) as a substrate and catalyst for malonyl ACP biosynthesis. *Chemistry and Biology* **6**:577-584.

Publications Submitted During this Research

Rafferty, J.B., Simon, J.W., Baldock. C., Artymuick, P.J., Baker, P.J., Stuitje, A.R., Slabas, A.R., and Rice, D.W., (1995) Common themes in redox chemistry emerge from the X-ray structure of oilseed rape (*Brassica napus*) enoyl acyl carrier protein reductase. *Structure* **3**: 927-938.

Slabas, A.R., Brown, A.P., Rafferty, J.B., Rice, D.W., Baldock, C., Kroon, J.T.M., Simon, J.W., Stuitje, A.R., and Brough, C.L., (1996). Soluble and membrane bound components of plant lipid synthesis. *C.R. Acad.Sci.Paris, Sciences de la vie.* **319**:1043-1047.

Rafferty, J.B., Fisher, M., Langridge, S.J, Martindale, W., Thomas, N.C., Simon, J.W., Bithell, S., Slabas, A.R., and Rice, D.W., (1998). Crystallization of the NADP-dependant β-keto acyl carrier protein reductase from *E.coli*. *Acta Crystallographica* **D54**: 427-429.

Simon, J.W., and Slabas, A.R. (1998) cDNA cloning of *Brassica napus* malonyl-CoA:ACP transacylase (MCAT) (*fab*D) and complementation of an *E.coli* MCAT mutant. *FEBS Letts.* **435**:204-206.

Fawcett, T., Copse,C.L., Simon, J.W., and Slabas, A.R., (2000). Kinetic mechanism of the NADH-enoyl reductase from *Brassica napus*. *Febs Letts.* **484**:65-68.

Hayman, M.W., Fawcett, T., Schierer, T.F., Simon, J.W., Kroon, J.T.M., Gilroy, J.S., Rice, D.W., Rafferty, J.B., Turnbull, A.P., Sedelnikova, S.E., and Slabas, A.R., (2000). Mutagenesis of squash (*Cucurbita moschata*) glycerol-3-phosphate acyltransferase (GPAT) to produce an enzyme with altered substrate selectivity. *Biochem. Soc. Trans.* **28:**680-681.

Slabas, A.R., Simon, J.W., Schierer, T., Kroon, J.T.M., Fawcett, T., Hayman, M., Gilroy, J.S., Nishida, I., Murata, N., Rafferty, J.B., Turnbull, A.P., and Rice. D.W., (2000). Plant glycerol-3-phosphate-1-acyltransferase (GPAT): structure selectivity studies. *Biochem. Soc. Trans.* **28**: 677-679.

Slabas, A.R., Hanley, Z., Scherier, T.P., Rice, D.W., Turnbull, A.P., Rafferty, J.B., Simon, J.W., and Brown, A.P., (2001). Acyltransferases and their role in the biosynthesis of lipids – opportunities for new oils. *J.Plant Physiol.***158**: 505-513.

Turnbull, A.P., Rafferty, J.B., Sedelnikova, S.E., Slabas, A.R., Schierer, T.P., Kroon, J.T.M., Nishida, J., Murata, N., Simon, J.W., and Rice, D.W., (2001). Crystallization and preliminary X-ray analysis of the glycerol-3-phosphate 1-acyltransferase from squash (*Cucurbita moschata*). *Acta. Crystallog. section D-Biological Crystallography* **57**:451-453.

Turnbull, A.P., Rafferty, J.B., Sedelnikova, S.E., Slabas, A.R., Schierer, T.P., Kroon, J.T.M., Simon, J.W., Fawcett, T., Nishida, I., Murata, N., and Rice, D., (2001). Analysis of the structure, substrate specificity, and mechanism of squash glycerol-3-phosphate (1)-acyltransferase. *Structure* **9**:347-353.

Slabas, A.R., Simon J.W., and Brown A.P., (2001). Biosynthesis and regulation of fatty acids and triglycerides in oil seed rape. Current status and future trends. *Eur. J. Lipid Science and Technology* **103**:455-466.

Slabas, A.R., Kroon, J.T.M., Scheirer, T.P., Gilroy, J.S., Hayman, M., Rice, D.W., Turnbull, A.P., Rafferty, J.B., Fawcett, T., and Simon, J.W., (2002). Squash Glycerol-3-phosphate (1)-acyltransferase [G3PAT] - alteration of substrate selectivity by site directed mutagenesis and identification of arginine and lysine residues involved in glycerol-3-phosphate binding. *Biochemistry (in Press).*

Roujeinikova, A., Baldock, C., Simon, J.W., Gilroy,J., Baker, P., Stuitje, A.R., Rice, D.W., Rafferty, J.B., and Slabas, A.R., (2002). Crystallization and preliminary X-ray crystallographic studies on acyl-(acyl carrier protein) from *Escherichia coli.* *Acta. Crystallog. section D-Biological Crystallography* 58:330-332.

Roujeinikova, A., Baldock, C., Simon, J.W., Gilroy,J., Baker, P., Slabas, A.R., Stuitje, A.R., Rice, D.W., and Rafferty, J.B., and (2002). X-ray crystallographic studies on butyryl-ACP reveal flexibility of the structure around a putative acyl chain binding site. (*Structure in Press*).