DOCTOR OF PHILOSOPHY

Characterisation of AAE7/ACN1 and aconitase isoforms from Arabidopsis thaliana

Nicholl, Sarah

Award date:
2011

Awarding institution:
Bangor University

Characterisation of AAE7/ACN1 and aconitase isoforms from *Arabidopsis thaliana*

Sarah Nicholl

A thesis submitted to Bangor University in candidature for the degree of Doctor of Philosophy

School of Biological Sciences

Bangor University

July 2011

Abstract

During germination plants convert stored fatty acids into sucrose via β-oxidation, the glyoxylate cycle and gluconeogenesis, in order to supply energy prior to the commencement of photosynthesis. One such fatty acid, acetyl-CoA, is produced both from β-oxidation and externally supplied acetate. Within the glyoxysome, externally supplied acetate is converted to acetyl-CoA by AAE7/ACN1, prior to entering the glyoxylate cycle. Structural analysis of AAE7/ACN1 has been conducted within this thesis to determine if AAE7/ACN1 is able to undergo a conformational change. Experimental and theoretical structural analysis has shown that AAE7/ACN1 is a dimeric protein in solution, although experimental analysis using solution scattering was unable to detect a conformational change in the presence and absence of a substrate analogue.

Previous research shows that *Arabidopsis* seedlings are able to take up and metabolise both acetate and butyrate from the growth medium and shows that *aae7/acn1* mutant growth is inhibited as a result of cytosolic acidification. The dimeric structure of AAE7/ACN1, identified within this thesis, results in a larger concentration of active sites within a localised area. This may result in the speedy detoxification of acetate and butyrate.

Other work in this thesis focuses on the role of aconitase within the glyoxylate and TCA cycles. Three isoforms of aconitase have previously been identified in the dicotyledonous plant *Arabidopsis thaliana*, namely ACO1, ACO2 and ACO3. A cytosolic location has been suggested for ACO1, whilst ACO2 and ACO3 have been identified within the mitochondria. The cytosolic isoform of aconitase is believed to be involved in the glyoxylate cycle, where it converts citrate to isocitrate, though the specific role each isoform plays in metabolism during germination is unclear. A metabolic approach has been undertaken within this thesis to determine the role each isoform plays in seedling establishment. A large reduction in aconitase activity, an increase in citrate levels and a delay in germination were observed in *aco3* mutants. These results suggest that ACO3 makes the largest contribution to citrate metabolism in germinating seedlings. This may mean that during seedling establishment the glyoxylate cycle is not as important as previously thought.

Active aconitase contains a [4Fe-4S] cluster. Under conditions of iron deficiency, the [4Fe-4S] cluster of aconitase is converted to a [3Fe-4S] cluster, which causes mammalian aconitase to lose its enzymatic function and become an RNA binding protein. A structural approach has been used within this thesis to attempt to characterise the *Arabidopsis* aconitase isoforms. ACO1 and ACO2 have been over-expressed in the eukaryotic organism *Pichia pastoris*. A novel method of over-expressing active aconitase without the need for reactivation has also been identified and implemented to show that both ACO1 and ACO2 are able to function as aconitase, confirming previously published work that used mutagenic analysis.

Contents

List of Figures

List of Tables

Abbreviations used in the text

aco (ACO)	aconitase
ACN	Acetate Non Utilising
AAE	Acyl Activating Enzyme
BSA	bovine serum albumin
cDNA	complementary deoxyribonucleic acid
Col	columbia
DEPC	diethyl pyrocarbonate
DNA	deoxyribonucleic acid
DNase	dexyribonuclease
DTT	dithiothreitol
dNTP	deoxynucleotide triphosphate
FAc	fluoroacetate
g	gravitational force (times gravity)
IAA	isoamyl alcohol
IMS	industrial methylated spirits
IRE-BP	Iron Response Element Binding Protein
IRE	Iron Response Element
LB	Luria-Bertani medium
M	molar
mRNA	messenger RNA
MS	mass spectrometry
NAD	nicotinamide adenine dinucleotide
PCR	polymerase chain reaction
RNA	ribonucleic acid
RNase	ribonuclease
RT	reverse transcription
SAXS	small angle x-ray scattering
SDS	sodium dodecyl sulphate
TCA	tricarboxylic acid
Tris	tris (hydroxymethyl) amino-methane
v/v	volume to volume ratio
w/v	weight (mass) to volume ratio

Acknowledgments

Many people deserve thanks for their guidance, assistance and support throughout my studies. Firstly, I would like to thank my Supervisors Dr M.A. Hooks and Dr L.M. Murphy for their guidance over the course of my studies. I would also like to thank Dr J. Grossmann for his assistance with small angle x-ray scattering and Dr B. McCullough and Dr H. Fuller for mass spectroscopy analysis. A special thank you is also required for Mr G. Williams without whose help radioactivity work would not have been possible.

I am extremely grateful to the University of Wales, Bangor for the Sir William Robert PhD studentship that has enabled me to pursue this PhD.

Many thanks are also due to previous lab members, Dr J. Turner, Dr B. Grail, Mrs L. Thurlow and Mrs W. Grail for bestowing upon me their many years of technical expertise.

My thanks go to the 'Women in Science' members Dr E. Allen, Dr R. Bennett, Dr S. Brittain, Dr G. Wei, Dr A. Jaendling and Dr C. Kay for their unfailing support and friendship in times of need.

I would also like to thank Mr R. Stanford, Dr C. Gertler and Dr C. Hayden for their moral and technical support during the production of this thesis.

A massive thank you is required for my mother, father and sister for all of their emotional support and understanding during the course of my PhD.

Special thanks however are reserved for Dr M.C. Malpass, without whose love and support this thesis would not have materialised. Thank you.

Introduction

1.1 Introduction

Plants are unique organisms that must adapt rapidly to changing environmental conditions in order to survive. This is important as they are organisms that provide a large proportion of the world's food, either directly or indirectly (e.g. animal products). It is therefore important to understand how these organisms develop, grow and function. However there are over 200,000 species of seed plants (Scotland and Wortley, 2003), making it impossible to conduct meaningful research on each one. The scientific community therefore conduct the majority of research using model organisms. A model organism must be similar to as many other related organisms as possible so that information learned, and conclusions drawn, can be easily transferred to a wide number of species. The most widely used model organism for laboratory plant research is *Arabidopsis thaliana*.

Arabidopsis is widely used as a model organism because of its small genome size, short life cycle, small physical size and large seed production. In addition, being a brassica, it is a relation of a number of important commercial crops. The complete sequence of this small 125 Mb genome has been known for approximately ten years, allowing every gene to be identified. An additional advantage of using *Arabidopsis* as a model organism is the ease at which the genome can be manipulated with either chemicals or radiation. This has lead to the production of a large number of mutant lines, which are readily available from stock centres. As a member of the mustard (Brassicaceae) family, *Arabidopsis* itself may not be of major agronomic significance. However, other members of this family include cabbage, cauliflower, rapeseed, radish and mustard, all of which are clearly of greater significance. Thus *Arabidopsis* is an important tool for both genetic and molecular biological research.

In order to survive plants constantly adapt to suit their changing environment. Essential nutrients are taken from the surrounding soil by a network of roots and transported through the xylem and phloem vascular tissues to the required locations. Plants are also able to influence the rhizosphere by root secretions, altering the microbial community and nutrient availability. Plants also utilise energy from sunlight and convert carbon dioxide into organic compounds in a process known as photosynthesis. During this autotrophic phase plants are known as producers; they

are able to make their own food. However, during germination seedlings must use stored energy resources from the seed for growth and development.

During the production of seeds, large quantities of carbon, nitrogen and other micro-nutrients are stored by the mother plant. These stored reserves are then utilised during seedling establishment until photosynthesis can begin. The majority of carbon stored within seeds is in the form of triacylglycerol (TAG) in oil bodies. The first step of lipid reserve mobilisation during germination is the hydrolysis of TAG into free fatty acids. These free fatty acids are subsequently transported into the glyoxysome by an ABC transporter, comatose (Baker *et al.,* 2006). Once inside the glyoxysome, fatty acids are esterified to acyl-CoA by long chain acyl-CoA synthetases (LACS) (Shockey *et al.,* 2002) in a two step reaction (Figure 1.1).

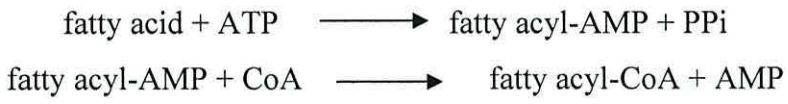


Figure 1.1: The reaction of acyl-CoA synthetase (Hisanaga *et al.,* 2004).

Acyl esters produced by LACS are then metabolised to acetyl-CoA by β-oxidation, which in turn is converted to four carbon carbohydrates by the glyoxylate cycle (Figure 1.2). Acetyl-CoA is also produced for entry into the glyoxylate cycle by AAE7/ACN1 (Turner *et al.,* 2005). This glyoxysomal enzyme converts externally supplied acetate into acetyl-CoA, which is then metabolised by the glyoxylate cycle (Turner *et al.,* 2005).

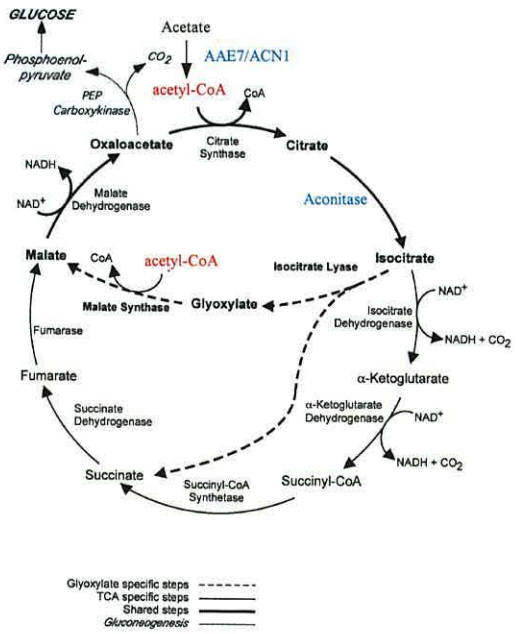


Figure 1.2: The glyoxylate cycle, TCA cycle and gluconeogenesis (Modified from Lorenz and Fink, 2002).

Acetyl-CoA enters the glyoxylate cycle at two positions where it is metabolised to citrate and malate. The produced citrate is then transported out of the glyoxysome for metabolism to isocitrate by aconitase. Malate, produced from acetyl-CoA by malate synthase, is transported out of the glyoxysome (Metter and Beevers, 1980). Malate is believed to have a number of possible uses; one of these is transport to the mitochondria where it enters the tricarboxylic acid (TCA) cycle (Figure 1.2). Malate is also believed to be transported to the cytosol where it is converted to sucrose through gluconeogenesis (Cornah *et al.,* 2004). Sucrose is an important sugar that is then transported throughout the plant to the required locations.

Succinate, produced as part of the glyoxylate cycle, is also believed to be transported to the mitochondria where it enters the TCA cycle. Once within the TCA cycle, succinate is metabolised to oxaloacetate which is converted to sucrose by gluconeogenesis (Penfield *et al.,* 2005).

The work conducted within this PhD thesis aims to investigate two steps of the TCA and glyoxylate cycles, namely AAE7/ACN1 and aconitase, which are responsible for the production of acetyl-CoA and citrate respectively.

1.2 Thesis aims

1.2.1 Experimental structural analysis of AAE7/ACN1

The initial aim of work conducted in chapter three was to determine the previously unknown three dimensional structure of AAE7/ACN1. This was firstly attempted by protein crystallisation with the view of structural determination by x-ray crystallography. Secondly, solution scattering analysis was attempted to determine the protein's overall conformation; substrate inhibitors were used to determine if AAE7/ACN1 undergoes a conformational change during catalysis.

1.2.2 Theoretical structural analysis of AAE7/ACN1

In addition to the experimental analysis, computer models were produced with the aim of determining the three dimensional structure of AAE7/ACN1 (chapter four). Homology models of AAE7/ACN1 were created using two different modelling programs, namely I-TASSER and MODELLER. The latter of the two programs utilised the crystallographic structures of a long chain fatty acyl-CoA synthetase in both an open and closed conformation. This was in order to attempt to theoretically predict the extent that AAE7/ACN1 might undergo a conformational change.

1.2.3 Structural analysis of the *Arabidopsis thaliana* aconitase isoforms

The primary objective of work conducted in chapter five was to assess the activity of the three isoforms of aconitase. An additional aim was to structuralise the aconitase isoforms by x-ray crystallography and x-ray absorption spectroscopy to determine the three dimensional shape and characterise the iron sulphur cluster of the aconitase isoforms. In order to achieve these aims purified protein is required. Attempts were made to over-express large quantities of aconitase in *Pichia pastoris* for subsequent structural analysis. Additionally, in order to compare the three aconitase isoforms, homology models of the aconitase isoforms were made based on the crystallographic structure of the *Homo sapiens* cytosolic aconitase.

1.2.4 Investigate the metabolic roles of aconitase in *Arabidopsis thaliana*

The aim of work conducted in chapter six was to investigate the metabolic roles of the three aconitase isoforms during seedling establishment. Aconitase isoforms are believed to function as part of both the TCA and glyoxylate cycle. However the location and cyclic participation of each isoform is currently unknown. In order to

attempt to determine the roles that aconitases play in metabolism, homozygous knockout mutants needed to be identified for subsequent metabolic analysis.

1.3 Cellular metabolism

The basic functional unit of living organisms are known as cells. There are two types of cells, prokaryotic and eukaryotic. Most bacteria and archaea are unicellular organisms that contain a single prokaryotic cell. In contrast plants and animals are multicellular organisms comprised of eukaryotic cells. The main differences between prokaryotic and eukaryotic cells are the absence of a membrane bound nucleus and internal organelles in prokaryotes. The absence of these compartments in prokaryotic cells results in the presence of DNA and metabolic processes within the cytoplasm.

Metabolism in eukaryotic cells is therefore much more complicated than prokaryotic cells, as transport mechanisms are required to move metabolites to and from different organelles. The presence of organelles, however, offers an advantage by allowing different metabolic processes to occur within different locations for a particular purpose. The compartmentalisation of metabolic pathways is important to separate and control metabolism. Cellular metabolism is complex and includes a degree of overlap with a single metabolite being involved in multiple cycles. Examples of this include citrate and malate which are involved in both the TCA and glyoxylate cycles.

1.4 Tricarboxylic acid cycle

This Tricarboxylic acid (TCA) cycle is an important cycle within all living organisms, converting acetyl-CoA to carbon dioxide, water and energy. Enzymes of the TCA cycle include citrate synthase, aconitase, isocitrate dehydrogenase, α-ketoglutarate dehydrogenase, succinyl-CoA synthetase, succinate dehydrogenase, fumarase and malate dehydrogenase.

1.4.1 Citrate synthase

The first step of the TCA cycle is the condensation of acetyl-CoA with oxaloacetate to form citrate and CoA. This condensation reaction is catalysed by the enzyme citrate synthase. This step in the TCA cycle is the only one that results in the formation of a carbon-carbon bond (Wiegand and Remington, 1986).

1.4.2 Aconitase

The second step of the TCA cycle is the stereo specific isomerisation of citrate to isocitrate via *cis*-aconitate. This dehydration followed by rehydration reaction is catalysed by aconitase. Active aconitase is believed to contain a [4Fe-4S] cluster that interacts directly with the substrate during catalysis.

1.4.3 Isocitrate dehydrogenase

The conversion of isocitrate with NAD^+ to α-ketoglutarate, NADH and CO_2 is the next step in the TCA cycle. The conversion of isocitrate to α-ketoglutarate is catalysed by isocitrate dehydrogenase. This reaction occurs in a two step process by the oxidative decarboxylation of isocitrate.

1.4.4 α-ketoglutarate dehydrogenase

α-ketoglutarate dehydrogenase catalyses the conversion of α-ketoglutarate, NAD^+ and CoA to succinyl-CoA, NADH and CO_2. This enzyme represents a control point in the TCA cycle, as a build up of either succinyl-CoA or NADH inhibits the enzymes activity (Smith *et al.,* 1973).

1.4.5 Succinyl-CoA synthetase

Succinyl-CoA synthetase catalyses a reversible step in the TCA cycle. This enzyme is responsible for the conversion of succinyl-CoA to succinate and CoA and vice versa under certain conditions.

1.4.6 Succinate dehydrogenase

Succinate dehydrogenase is an enzyme complex (complex II) that is bound to the inner mitochondrial membrane. This is a unique enzyme that participates in both the TCA cycle and the electron transport chain. Within the TCA cycle succinate dehydrogenase catalyses the oxidation of succinate to fumarate.

1.4.7 Fumarase

As part of the TCA cycle, fumarase catalyses a reversible hydration/dehydration reaction. During this reaction fumarate is converted to malate.

1.4.8 Malate dehydrogenase

The final step of the TCA cycle is the conversion of malate and NAD^+ to oxaloacetate and NADH. This reaction is catalysed by malate dehydrogenase, completing the TCA cycle.

1.5 The glyoxylate cycle

The glyoxylate cycle converts fats into carbohydrates in both plants and microorganisms (Kornberg and Krebs, 1957; Eastmond and Graham, 2001). Within plants, this cycle has been shown to function both during germination and senescence (De Bellis and Nishimura, 1991; Eastmond and Graham, 2001). Acetyl-CoA enters the glyoxylate cycle at two positions to produce citrate and malate. Within *Arabidopsis* acetyl-CoA can be produced from the breakdown of fatty acids by β-oxidation or from externally supplied acetate by AAE7/ACN1 (Eastmond and Graham, 2001; Turner *et al.,* 2005).

1.5.1 AAE7/ACN1

AAE7/ACN1 is an acyl activating enzyme (AAE) identified during fluoroacetate screening experiments as acetate non-utilising (ACN) (Shockey *et al.,* 2003; Turner *et al.,* 2005). This enzyme is responsible for the conversion of acetate to acetyl-CoA in the glyoxysome. Acetyl-CoA produced from this two step reaction then feeds into the glyoxylate cycle for further metabolism.

1.5.2 Citrate synthase

Oxaloacetate and acetyl-CoA are converted to citrate and CoA within the glyoxysome by citrate synthase (Nishimura *et al.,* 1986). Knockout mutants of citrate synthase have shown that citrate synthase is essential for germination and development in *Arabidopsis* (Pracharoenwattana *et al.,* 2005). The reaction of citrate synthase allows carbon from β-oxidation to be exported from the glyoxysome as citrate.

Once exported citrate may re-enter the glyoxylate cycle as isocitrate (Canvin and Beevers, 1961), be transported to the mitochondria (Salon *et al.,* 1988; Pracharoenwattana *et al.,* 2005) or be converted to acetyl-CoA within the cytosol (Fatland *et al.,* 2005).

1.5.3 Aconitase

In plants aconitase activity has not been identified within the glyoxysome. Therefore cytosolic aconitase is believed to function as part of the glyoxylate cycle (Brouquisse *et al.,* 1987; Verniquet *et al.,* 1991). Active aconitase contains a [4Fe-4S] cluster that is essential for the conversion of citrate to isocitrate via *cis*-aconitate. The produced isocitrate is then believed to re-enter the glyoxysome to continue the glyoxylate cycle.

1.5.4 Isocitrate lyase

Isocitrate lyase is responsible for the cleavage of isocitrate to succinate and glyoxylate. The produced succinate is transported from the glyoxysome to the mitochondria where it enters the TCA cycle, though glyoxylate remains in the glyoxysome where it is metabolised by malate synthase.

1.5.5 Malate synthase

A second molecule of acetyl-CoA enters the glyoxylate cycle at this stage. The acetyl-CoA and glyoxylate are converted into malate and CoA by malate synthase. In *Arabidopsis* malate synthase is a single copy gene whose product functions solely in the glyoxysome. Knockout mutants of malate synthase block the glyoxylate cycle, preventing the conversion of acetyl-CoA to sugar (Eastmond *et al.,* 2000; Smith, 2002).

1.5.6 Malate dehydrogenase

The final step of the glyoxylate cycle is the conversion of malate and NAD^+ to oxaloacetate and NADH. This reaction is catalysed by malate dehydrogenase, completing the glyoxylate cycle.

1.6 The multiple functions of metabolites

Metabolites are important in cell processes, regulating enzyme activity and gene expression in response to changes in temperature, pH, oxidative stress, pathogen attack and nutrient deficiency. In order to survive plants must closely regulate their metabolic pathways to deal with a variety of stresses, created as the environment around them changes. Metabolites can be characterised as either primary or secondary metabolites. Primary metabolites are usually essential, involved directly in

normal development, growth and reproduction, whereas secondary metabolites allow the plant to adapt to their environment (Bourgaud *et al.,* 2001).

A well studied example of metabolite regulation by gene expression is the *E. coli lac* operon (Jacob and Monod, 1961; Beckwith, 1967; Lodish *et al.,* 1995). The *lac* operon is a sequence of DNA that encodes three genes involved in lactose metabolism. These three genes are transcribed and translated to produce the three proteins β-galactosidase, β-galactoside permease and β-galactoside transacetylase. When lactose is abundant it binds to the *lac* repressor. This prevents binding to the *lac* operator allowing transcription of the *lac* operon to occur. In the absence of lactose the *lac* repressor binds to the *lac* operator preventing transcription of the *lac* operon. Binding of the *lac* repressor prevents transcription by blocking the transcription start site and preventing the binding of RNA polymerase. Repression of the *lac* operon in the absence of lactose is favourable as it results in the conservation of energy by preventing the production of unnecessary proteins.

Metabolic regulation of gene expression has also been identified in plants. The earliest example of metabolite regulation showed that sugars down regulate photosynthetic gene expression in maize (Sheen, 1990). Sugars also interact with hormones such as abscisic acid, auxin, cytokinin and ethylene to control seed and seedling development (Zhou *et al.,* 1998; Yuan and Wysocka-Diller, 2006; Dekkers *et al.,* 2008; Mishra *et al.,* 2009). In addition to their signalling role, sugars are also important metabolites in plants. Glucose is practically important with a variety of roles. This simple sugar is involved in the production of proteins, the biosynthesis of starch and cellulose and is a precursor for vitamin C.

In plants, sugars are produced from both the process of photosynthesis and the breakdown of fatty acids during germination. These sugars play an essential role in metabolism as substrates for carbon and energy metabolism. Aconitase plays a central role in metabolism converting citrate to isocitrate, as both a TCA and glyoxylate cycle enzyme. The involvement of aconitase in these cycles makes aconitase an excellent candidate for investigating metabolic signalling. In addition to its pivotal role in metabolism, aconitase contains an iron sulphur cluster within its

active site. This makes the regulation of iron important for the enzymatic functionality of aconitase.

1.7 The importance of iron regulation in organisms

The twenty sixth element in the periodic table, iron (Fe), is an abundant element that is vital to both plant and animal life. However, within organisms, iron can be both an essential nutrient and a potential toxin. This therefore requires cellular iron levels to be tightly regulated. Free iron in a cell can be potentially toxic by reacting with oxygen. This may lead to the production of reactive oxygen species (ROS) that can cause tissue damage. Harmful ROS include hydrogen peroxide (H_2O_2), superoxide and hydroxyl radicals (Cadenas, 1989; Schopfer *et al.,* 2001). In the presence of oxygen iron is in the oxidised ferric form of Fe^{3+}. However, organisms prefer iron in the reduced ferrous form of Fe^{2+}. In addition to the presence or absence of oxygen, the state of iron can also be influenced by pH levels.

The oxidation of ferrous iron (Fe^{2+}) to ferric iron (Fe^{3+}) with H_2O_2 is known as the fenton reaction (Halliwell and Gutteridge, 1984; Prousek, 2007). During this process a hydroxyl radical ($.OH^{\cdot}$) and a hydroxyl ion (OH^-) are also produced (Figure 1.3). In biological systems the superoxide ion ($O_2^{\cdot -}$) is the reducing agent that converts Fe^{3+} to Fe^{2+} (Guerinot and Yi, 1994). The hydroxyl radical can cause damage to organic molecules, DNA, protein and lipids (Shires, 1982; Halliwell and Chirico, 1993; Valko *et al.,* 2004; Apel and Hirt, 2004). For these reasons it is important for organisms to tightly regulate their cellular iron levels in order to limit damage.

$$Fe^{2+} + H_2O_2 \longrightarrow Fe^{3+} + .OH^{\cdot} + OH^-$$

$$Fe^{3+} + O_2^{\cdot -} \longrightarrow Fe^{2+} + O_2$$

Figure 1.3: Fenton reaction.

1.7.1 Iron regulation in animals

Depending on cellular levels in animals, iron can be both a potential toxin and an essential nutrient. Some negative effects of a build up of iron in tissue include skin pigmentation, diabetes mellitus and liver and heart failure. Despite these toxic effects iron is an essential component of haemoglobin, required for the transport of oxygen

and carbon dioxide throughout the body. The main problem caused by low levels of iron in humans is iron deficiency anaemia. This condition can cause extreme fatigue and weakness in adults and a heart murmur and delays in growth and development in young children. Iron deficiency anaemia can be particularly dangerous in pregnant women, increasing the risk of a premature delivery and low birth weights. The range of problems caused by either too little or too much iron requires levels to be tightly controlled.

In animals iron levels are controlled by altering the amount of iron both transported and stored within the cell. The transport protein transferrin receptor and the storage protein ferritin are both translationally regulated by an iron regulatory protein (IRP). *H. sapiens* contain two IRPs (IRP1 and IRP2) that share approximately 61 % amino acid sequence identity (Guo *et al.,* 1995a). Depending on the cytosolic iron concentration, IRP1 has the ability to function as either aconitase or an RNA binding protein (Kaptain *et al.,* 1991; Klausner and Rouault; 1993; Pantopoulos *et al.,* 1995). In contrast, IRP2 only functions as an RNA binding protein (Guo *et al.,* 1995b).

During iron deplete conditions IRP contains a [3Fe-4S] cluster with RNA binding properties (Haile *et al.,* 1992). In this state the IRP adopts an open conformation (Walden *et al.,* 2006; Figure 1.4). The movement of domains 3 and 4 expose an iron response element (IRE) binding site that allows the RNA to bind (Dupuy *et al.,* 2006; Walden *et al.,* 2006).

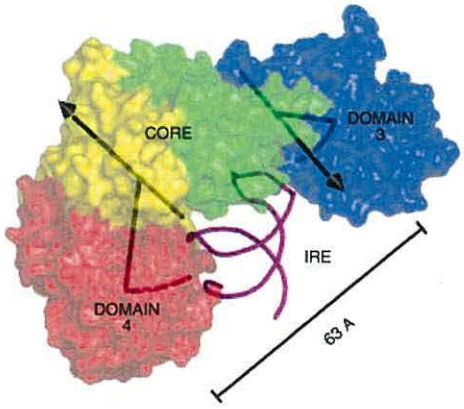


Figure 1.4: The open structure of IRP1. This image was taken directly from Walden *et al.* (2006).

IRP contains an IRE binding site that binds to IREs in the 5' or 3' UTR of mRNA. IREs are 28 nucleotides in length with a conserved central region, 5'-CAGUGH-3' (where H is any nucleotide except G) (Henderson *et al.,* 1994). IRP regulates cellular iron levels by binding to IREs in the 3' UTR of transferrin receptor mRNA and the 5' UTR of ferritin mRNA. Transferrin receptors are involved in the transport of iron across cell membranes, whereas ferritin is an iron storage protein. The binding of IRP to ferritin mRNA prevents translation, which increases the cellular concentration of iron (Harrison and Arosio, 1996; Curie and Briat, 2003).

During iron deprived conditions, IRP also binds to transferrin receptor mRNA, stabilising the mRNA. This increases the number of transferrin receptors in the plasma membrane and therefore the uptake of iron into the cell (Hirling *et al.,* 1994). Transferrin receptor mRNA is stabilised by preventing its cleavage by an endoribonuclease (Kaldy *et al.,* 1999).

IRP has been shown to adopt a closed conformation during conditions of excess iron (Figure 1.5). Movement of domains 3 and 4 shields the IRE binding site, preventing RNA binding (Hirling *et al.,* 1994). The conversion of IRP from an open to a closed conformation is believed to be dependent on cellular iron concentrations (Alen and Sonenshein, 1999; Dupuy *et al.,* 2006). Higher concentrations of iron convert the [3Fe-4S] cluster of IRP to a [4Fe-4S] cluster (Kaptain *et al.,* 1991; Basilion *et al.,* 1994b). This conversion results in the loss of RNA binding function and the gain of aconitase activity (Haile *et al.,* 1992; Klauser and Rouault, 1993; Basilion *et al.,* 1994a).

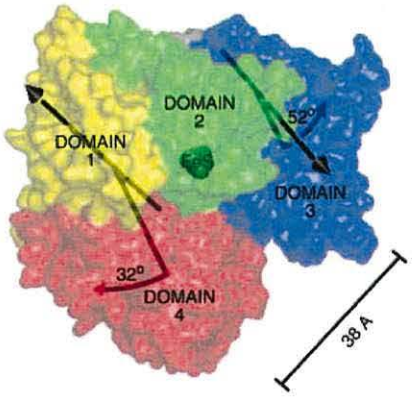


Figure 1.5: The closed structure of IRP / aconitase. This image was taken directly from Walden *et al.* (2006).

1.7.2 Iron regulation in plants

Iron is an essential element in plants as it is involved in the electron transport chains of photosynthesis and respiration (Connolly and Guerinot, 2002). This important micronutrient is also necessary for chlorophyll synthesis and as a component of cytochromes and ferredoxins. Iron deficiency can lead to a reduction in crop yields and, if severe, even complete crop failure (Guerinot and Yi, 1994). One notable symptom of iron deficiency in plants is the yellowing of new leaves (Forbes and Watson, 1994; Hell and Stephen, 2002). The yellowing of leaves, known as chlorosis, is caused by the insufficient production of chlorophyll (Brown *et al.,* 1955; Brown, 1956). Excess iron can also be problematic in plants, resulting in brown spots covering the leave surface (Kampfenkel *et al.,* 1995). Other problems caused by excess iron include a reduction in the photosynthetic rate and the production of ROS (Kampfenkel *et al.,* 1995; Caro and Puntarulo, 1996). ROS, such as hydroxyl ions, can subsequently lead to protein and DNA damage (Apel and Hirt, 2004). The generation of ROS by excess iron and the cellular damage they cause may explain the need for tight control of cellular levels by uptake, transport and storage mechanisms.

Research also shows that nitric oxide (NO) is a key signalling molecule involved in the regulation of iron in plants (Murgia *et al.,* 2002). NO has been shown to be involved in the accumulation of ferritin in *Arabidopsis* (Murgia *et al.,* 2002). Ferritin is an important iron storage protein with the ability to store up to 4500 iron atoms in

both plants and animals (Harrison and Arosio, 1996; Theil, 1987). Within animals the stability of ferritin mRNA is regulated by cytosolic aconitase. In contrast, evidence suggests that this is unlikely to occur in *Arabidopsis* (Arnaud *et al.,* 2007). However, NO has been shown to inhibit aconitase activity in *Arabidopsis* (Navarre *et al.,* 2000). This could suggest a role for plant aconitase in iron homeostasis.

The majority of iron within the soil is present in the ferric form (Fe^{3+}), making it inaccessible to plants. In order to acquire iron from the soil plants therefore need to reduce iron to its ferrous form (Fe^{2+}). The amount of iron available to plants depends largely on the pH of the soil solution (Forbes and Watson, 1994). Iron is more available at pH 6 than it is at pH 8, as Fe^{3+} is converted to Fe^{2+} as the pH decreases (Forbes and Watson, 1994). Plants manipulate this by releasing protons to acidify the soil solution surrounding the roots. This strategy (strategy I) is implemented by non-graminaceous plants using H^+ ATPases of the root plasma membrane (Fox and Guerinot, 1998; Santi *et al.,* 2005). Following acidification Fe^{3+} is reduced to Fe^{2+} by a plasma membrane bound ferric reductase oxidase (FRO), FRO2 (Robinson *et al.,* 1999). In *Arabidopsis* there are eight members of the FRO family, which are expressed in a variety of tissues (Wu *et al.,* 2005). Research shows that FRO2 transcript accumulates in roots in response to iron deficiency (Robinson *et al.,* 1999).

After reduction, iron is transported from the soil by a member of the ZIP family, iron regulated transporter 1 (IRT1) (Eide *et al.,* 1996; Guerinot, 2000). Research shows that IRT1 is located in the roots and induced by iron deficiency (Vert *et al.,* 2002). Once taken up by the roots, iron is believed to be transported throughout the plant via the xylem (Guerinot and Yi, 1994; Curie and Briat, 2003). Other methods of iron transport include the natural resistance associated macrophage protein (NRAMP) family and the yellow stripe like (YSL) family (Curie *et al.,* 2000, 2009; Thomine *et al.,* 2000).

In *Arabidopsis* there are several members of the NRAMP family involved in the transport of metal ions (Williams *et al.,* 2000). Three of these proteins have been implicated in the transport of iron (Curie *et al.,* 2000; Thomine *et al.,* 2000). Transcripts of NRAMP 1, 3 and 4 have all been shown to accumulate in response to iron deficiency (Curie *et al.,* 2000; Thomine *et al.,* 2000). NRAMP 1 contains a

plastid targeting sequence and is therefore believed to be involved in the transport of plastid iron (Curie *et al.,* 2000). NRAMP 3 and NRAMP 4 are also believed to play a role in the intracellular transport of iron (Thomine *et al.,* 2003; Lanquar *et al.,* 2005). The roles of NRAMP 3 and NRAMP 4 in mobilising iron from the vacuole have been shown to be very important during development of *Arabidopsis* seedlings (Lanquar *et al.,* 2005). Excess iron is believed to be transported into the vacuole for storage by vacuolar iron transport 1 (VIT1) (Kim *et al.,* 2006; Briat *et al.,* 2007).

A second strategy (strategy II) has been identified for the uptake of iron in grasses. These plants release low molecular weight ligands called phytosiderophores when iron levels are low (Fox and Guerinot, 1998; Jeong and Guerinot, 2009). These phytosiderophores (PS) are chelators that solubilise inorganic Fe(III), increasing the rate of iron uptake (Römheld and Marschner, 1986). Phytosiderophores, such as mugineic acid and deoxy-mugineic acid, are synthesised in the roots from L-methionine and secreted to solubilise iron (Shojima *et al.,* 1990; Ma and Nomoto, 1993; 1994). Fe(III)-PS complexes are then taken up and redistributed via yellow stripe (YS) family transporters (Curie *et al.,* 2001; Schaaf *et al.,* 2004; Murata *et al.,* 2006; Jeong and Guerinot, 2009).

A YS family has been identified in *Arabidopsis* with eight members (Curie *et al.,* 2001). YS proteins are believed to be involved in the uptake and transport of iron throughout the plant (Curie *et al.,* 2001; DiDonato *et al.,* 2004; Le Jean *et al.,* 2005). In graminaceous plants, YS proteins are believed to be responsible for the uptake of Fe(III)-PS complexes (Curie *et al.,* 2001; Hell and Stephan, 2003). However, non-graminaceous plants do not produce phytosiderophores under conditions of iron deficiency (DiDonato *et al.,* 2004; Schaaf *et al.,* 2005). Thus, YS proteins are believed to transport iron to and from vascular tissue of non-graminaceous plants bound to chelates such as nicotianamine (NA) and citrate (DiDonato *et al.,* 2004; Grotz and Guerinot, 2006; Waters *et al.,* 2006; Haydon and Cobbett, 2007; Curie *et al.,* 2009). Citrate complexed with iron is believed to be the major method of long distance transport from roots to shoots within the xylem (von Wiren *et al.,* 1999) of both strategy I and strategy II plants (Tiffin, 1966a; 1966b). Increased levels of organic acids, particularly citrate, have been observed under conditions of iron

deficiency in leaves (Landsberg, 1981) and roots (Ric De Vos *et al.*, 1986; Alhendawi *et al.*, 1997; Lopez-Millan *et al.*, 2000).

1.8 Aconitase

Aconitase is responsible for the conversion of citrate to isocitrate via the intermediate *cis*-aconitate (Figure 1.6). This dehydration followed by rehydration reaction is conducted within both the cytosol and mitochondria. Mitochondrial aconitase functions as part of the TCA cycle in all living organisms. The TCA cycle is located in the matrix of the mitochondria where it provides high energy electrons in the form of NADH and $FADH_2$. These then enter the oxidative phosphorylation system where energy in the form of ATP is produced. Cytosolic aconitase is believed to function as part of the glyoxylate cycle in both microorganisms and plants (Kornberg and Krebs, 1957; Eastmond and Graham, 2001; Regev-Rudzki *et al.*, 2005).

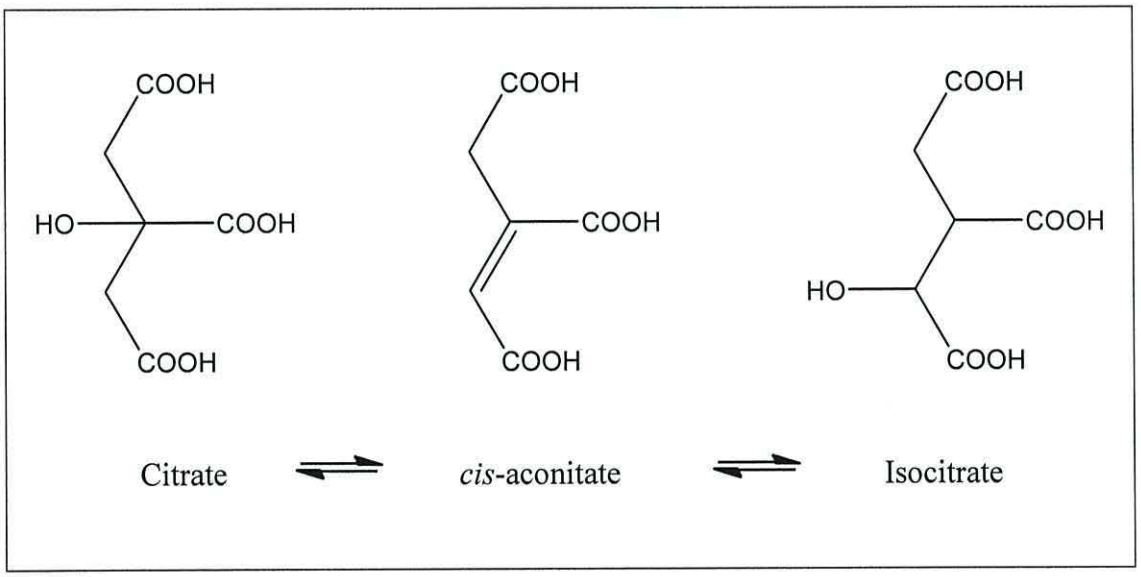


Figure 1.6: The metabolic reaction of aconitase.

1.8.1 Bacterial aconitase

Within *E. coli* there are two main isoforms of aconitase, aconitase A (acoA) and aconitase B (acoB) (Gruer and Guest, 1994). The acoA isoform is induced by iron and oxidative stress and acoB is induced during exponential growth (Jordan, 1999; Tang *et al.*, 2002). The *E. coli* aconitase isoforms have both been shown to be bifunctional proteins (Tang and Guest, 1999). Research shows that acoA and acoB

both contain enzymatic and mRNA binding functions (Tang and Guest, 1999). Bacterial aconitase isoforms from *Bacillus subtilis* and *Mycobacterium tuberculosis* have also been shown to be dual functioning proteins (Allen and Sonenshein, 1999; Kiley and Beinert, 2003; Banerjee *et al.,* 2007).

Aconitase isoforms from *B. subtilis* and *E. coli* are able to bind to an iron response element in the mRNA sequence of rabbit ferritin (Alen and Sonenshein, 1999). The binding of aconitase to the iron storage protein ferritin may increase the cellular iron content by blocking translation, reducing the amount of ferritin. Under conditions of iron depletion in *E. coli* acoB has been shown to lose enzymatic function and gain RNA binding functions (Varghese *et al.,* 2003). Increased citrate concentrations have been observed in acoB knockout mutants consistent with a reduction in aconitase activity.

E. coli acoA and acoB have also been shown to affect the stability of superoxide dismutase mRNA (Tang *et al.,* 2002). *E. coli* acoA has a positive effect on superoxide dismutase, increasing transcript stability, whereas *E. coli* acoB has a negative effect on superoxide dismutase, lowering transcript stability (Tang *et al.,* 2002; 2005).

Crystallographic analysis of *E. coli* acoB shows that this isoform of aconitase contains an additional domain (Williams *et al.,* 2002). The fifth domain of acoB (Figure 1.7) contains a HEAT like structure that is believed to be involved in protein-protein interactions (Tang *et al.,* 2005). However, both acoA and acoB form dimers in solution (Tang *et al.,* 2005; Tsuchiya *et al.,* 2008).

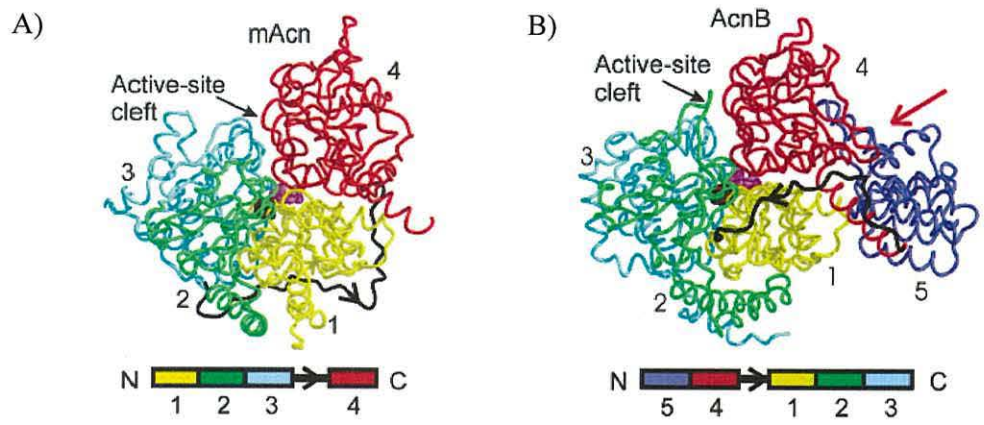


Figure 1.7: The crystallographic structure of A) mitochondrial aconitase (PDB 1ACO) and B) *E. coli* acoB (PDB 1L5J). Domains 1, 2, 3 and 4 are yellow, green, cyan and red respectively. The novel N-terminal domain of acoB is shown in purple. This image is taken directly from Williams *et al.*, 2002.

Metabolic analysis of aconitase in *Pseudomonus fluorescens* cultures has shown that aconitase is affected by elevated levels of aluminium (Middaugh *et al.*, 2005). The addition of excess aluminium to *Pseudomonus fluorescens* cultures resulted in a reduction in aconitase activity (Middaugh *et al.*, 2005). However, protein levels remained unaltered. This suggests that the addition of aluminium affects the composition of the protein's iron sulphur cluster, rendering the enzyme inactive. The oxidative environment created by aluminium generated ROS is believed to convert the [4Fe-4S] cluster into a [3Fe-4S] cluster, with the loss of aconitase activity (Boscolo *et al.*, 2003; Middaugh *et al.*, 2005).

1.8.2 Yeast aconitase

S. cerevisiae has only one isoform of aconitase, scACO1 (Gangloff *et al.*, 1990). This has been identified as a target for oxidative damage during conditions of oxidative stress (Cabiscol *et al.*, 2000). Under these conditions the presence of hydrogen peroxide resulted in a 75 % reduction in aconitase activity (Cabiscol *et al.*, 2000). This may suggest inactivation of aconitase by disruption of the iron sulphur cluster.

Other work has identified scACO1 as being within both the cytosol and mitochondria, where it participates as part of the glyoxylate shunt and TCA cycle

respectively (Regev-Rudzki *et al.,* 2005). The participation of scACO1 in metabolism is in agreement with previous research where it was identified by fluoroacetate screening (McCammon, 1996). The resistance of a scACO1 mutant to fluoroacetate suggests an involvement in acetate metabolism (McCammon, 1996).

Mitochondria contain circular strands of DNA that code for ribosomes, transfer RNA and polypeptides. Mitochondrial DNA (mtDNA) is packaged into protein-DNA complexes called nucleoids (Lodish *et al.,* 1995; Brewer *et al.,* 2000; Shadel, 2005). Studies conducted by Chen *et al.* (2005; 2007) have shown that scACO1 is involved in the maintenance of mtDNA, as a component of nucleoids. Unlike nuclear DNA, mtDNA lacks histones. The absence of histones is believed to make mtDNA more susceptible to the accumulation of point mutations and single stranded DNA breaks (O' Rourke *et al.,* 2002). Other work has shown that scACO1 is essential for mtDNA maintenance with the ability to replace Abf2p (Chen *et al.,* 2005; 2007). Abf2p has also been shown to play a key role in the structure, maintenance and expression of the yeast mitochondrial genome (Diffley *et al.,* 1991; Contamine and Picard, 2000; Brewer *et al.,* 2003).

1.8.3 Mammalian aconitase

H. sapiens contains two isoforms of aconitase, one cytosolic and one mitochondrial. These two isoforms are both located on different chromosomes. The cytosolic isoform is located on chromosome 9 and the mitochondrial isoform is located on chromosome 22 (Povey *et al.,* 1976; Slaughter *et al.,* 1978). The mammalian cytosolic aconitase isoform is also known as iron regulatory protein 1 (IRP1) (Kaptain *et al.,* 1991; Kennedy *et al.,* 1992; Klausner and Rouault, 1993). Depending on the cellular iron concentrations this isoform functions as either aconitase or IRP. In contrast, the mitochondrial isoform only functions as aconitase (Kaptain *et al.,* 1991).

The mitochondrial and cytosolic isoforms of aconitase are both involved in citrate metabolism. The mitochondrial isoform functions as part of the TCA cycle metabolising acetyl-CoA to carbohydrate. Cytosolic aconitase is involved in the production of glutamate and the regulation of glycolysis and fatty acid biosynthesis (Tong and Rouault, 2007).

Research shows that aconitase is purified in an inactive form that requires reactivation. Early work showed that inactive aconitase can be reactivated by restoring the [4Fe-4S] cluster with ferrous iron and cysteine (Dickman and Cloutier, 1950; Morrison, 1954). Subsequent research showed that aconitase can also be reactivated by incubation with ferrous iron, sodium sulphate and 2-mercaptoethanol (Suzuki *et al.,* 1976).

The crystallographic structure of mitochondrial aconitase has been determined from porcine heart (Robbins and Stout, 1989a; Robbins and Stout, 1989b; Lauble *et al.,* 1992; Lloyd *et al.,* 1999) and bovine heart (Lauble *et al.,* 1992; 1994; 1996). Structuralisation of aconitase showed that the enzymatic isoform contains four domains with a [4Fe-4S] cluster. The [4Fe-4S] cluster is located in a deep cleft between domains 1-3 and 4 (Lauble *et al.,* 1992; Beinert and Kennedy, 1993). During catalysis the fourth domain moves closer to domains 1-3, creating a closer, tighter conformation (Volz, 2008).

1.8.4 Plant aconitase

Aconitase has been identified within both cytosolic and mitochondrial fractions of plants (Brouquisse *et al.,* 1986; De Bellis *et al.,* 1993; Hayashi *et al.,* 1995). The mitochondrial isoform is believed to function as part of the TCA cycle and the cytosolic isoform as part of the glyoxylate cycle (Brouquisse *et al.,* 1986; 1987; Cornah and Smith, 2002). The glyoxylate cycle functions primarily in the glyoxysome, breaking down fatty acids (Figure 1.2). In order to metabolise fatty acids to sucrose, the glyoxylate cycle is believed to take a detour via the cytosol, where aconitase activity has been detected (Brouquisse *et al.,* 1987; Courtois-Verniquet and Douce, 1993). The absence of aconitase activity within the glyoxysome is believed to be due to the presence of ROS, as the [4Fe-4S] cluster of aconitase is susceptible to damage by ROS, rendering the enzyme inactive (Verniquet *et al.,* 1991).

Isoforms of aconitase are believed to function as part of both the TCA and glyoxylate cycle. Acetyl-CoA enters these cycles for metabolism to citrate. As part of these cycles aconitase converts citrate to isocitrate. The site of fluoroacetate toxicity in both plants and animals is believed to be aconitase. Exogenously supplied

fluoroacetate is metabolised to fluoroacetyl-CoA and subsequently fluorocitrate, which binds and inhibits aconitase activity (Figure 1.8). Fluoroacetate was used as a systemic insecticide prior to the identification of its toxicity to humans (Treble *et al.*, 1962). Plants species such as *Dichapetalum cymosum* (Marais, 1944) and *Acacia georgina* (Oelrichs and McEwanm, 1961) are both able to produce fluoroacetate. However, the mechanisms that enable these species to survive are unclear.

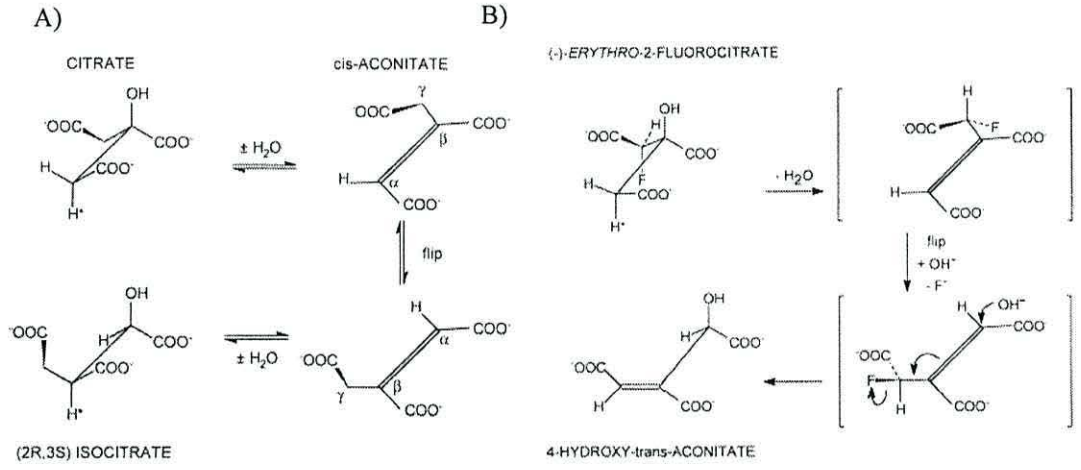


Figure 1.8: The metabolic reaction of aconitase from A) citrate and B) fluorocitrate (Images taken from Lauble *et al.*, 1996).

Lycopersicon pennellii aconitase (lpaco1) mutants showed a reduction in aconitase activity and an increase in both photosynthetic activity and fruit yield (Carrari *et al.*, 2003). The lpaco1 mutants also showed reduced levels of aspartate, glutamine, phenylalanine and leucine (Carrari *et al.*, 2003). The reduction in amino acid levels is consistent with a role in the TCA cycle, providing carbon skeletons for amino acid biosynthesis (Carrari *et al.*, 2003).

Nitric oxide (NO) is known to act as a signalling molecule in both plants and animals. Within plants NO signalling is believed to be involved in leaf expansion, root growth and activating defence responses to pathogens (Leshem, 1996; Gouvea *et al.*, 1997; Delledonne *et al.*, 1998; Durner *et al.*, 1998). Navarre *et al.*, (2000) examined the effect of NO in tobacco and showed that NO inhibits aconitase activity. The reduction in aconitase activity is believed to result in an increase in citrate (Navarre *et al.*, 2000). Other work has previously implicated citrate as a

signalling molecule, as raised levels were found to increase the level of alternative oxidase protein in tobacco cells exposed to potassium cyanide (Vanlerberghe and McIntosh, 1996; Navarre *et al.,* 2000). Aconitase within plant cell cultures has also been shown to be inhibited by salicylic acid (SA) (Ruffer *et al.,* 1995). The signalling molecule SA activates plant defence genes through elevated levels of hydrogen peroxide (Ruffer *et al.,* 1995). The binding of SA to aconitase could result in the release of iron and the production of hydrogen peroxide through the fenton reaction, reducing aconitase activity further.

Arabidopsis contains three isoforms of aconitase, ACO1, ACO2 and ACO3 (Arnaud *et al.,* 2007; Moeder *et al.,* 2007). ACO1 is believed to be located within the cytosol where it is constitutively expressed at low levels during most developmental stages (Peyret *et al.,* 1995; Arnaud *et al.,* 2007; Bernard *et al.,* 2009). However, during germination, pollen maturation and seed production the expression of ACO1 was shown to increase (Peyret *et al.,* 1995), suggesting a role within the glyoxylate cycle. Additional research has also suggested that ACO1 is located within the mitochondria (Kruft *et al.,* 2001; Arnaud *et al.,* 2007). The identification of an aconitase isoform in two locations has also been identified in *Saccharomyces cerevisiae* (Regev-Rudzki *et al.,* 2005). The additional aconitase isoforms, ACO2 and ACO3, have both been identified within the mitochondrial proteome (Millar *et al.,* 2001; Sweetlove *et al.,* 2002; Lee *et al.,* 2008). Mitochondria were isolated from *Arabidopsis* cells by percoll gradient density centrifugation. Isolated mitochondria were lysed and the resulting mitochondrial proteins separated by two dimensional gel electrophoresis. The protein spots were subsequently identified from trypsinated peptides by mass spectrometry (Millar *et al.,* 2001; Lee *et al.,* 2008).

Multiple aconitase isoforms have also been identified in maize (Wendel *et al.,* 1988), soybean (Cots and Widmer, 1999) and pumpkin (De Bellis *et al.,* 1993). Three aconitase isoforms (ACO I, ACO II and ACO III) were identified by native PAGE separation in pumpkins (De Bellis *et al.,* 1993). The ACO I and ACO II isoforms were subsequently purified by anion exchange chromatography. Native PAGE and isoelectric focusing analysis showed that ACO I and ACO II are both very similar proteins. ACO I and ACO II are both monomeric proteins with a molecular mass of approximately 98 kDa and pI values of 5.0 and 4.8, respectively (De Bellis *et al.,*

1993). The molecular mass of ACO III was also shown to be approximately 98 kDa (De Bellis *et al.,* 1995). Levels of ACO I were shown to be high in cotyledons but gradually fell as the cotyledons developed, suggesting a role in the glyoxylate cycle (De Bellis *et al.,* 1994; 1995). The ACO II and ACO III isoforms were found in cotyledons, leaves and hypocotyls (De Belli *et al.,* 1995). Subcellular localisation suggested that ACO I and ACO III are located within the cytosol and ACO II within the mitochondria (De Bellis *et al.,* 1995). This is in contrast to *Arabidopsis* as ACO1 is believed to be cytosolic and ACO2 and ACO3 are believed to be mitochondrial proteins (Arnaud *et al.,* 2007; Bernard *et al.,* 2009). However, a reduction in cytosolic aconitase activity was observed in both *aco1* and *aco3* (Arnaud *et al.,* 2007).

Within *H. sapiens* aconitase is a dual functioning protein with both aconitase and RNA binding functions. Moeder *et al.* (2007) addressed the question of a dual functioning aconitase in *Arabidopsis*. Evidence suggested that whilst aconitase was unable to bind to human ferritin IRE, it was able to bind the 5' UTR of the *Arabidopsis* chloroplastic copper / zinc superoxide dismutase 2 mRNA (Moeder *et al.,* 2007). This could suggest an involvement in oxidative stress as superoxide dismutase catalyses the conversion of superoxide to oxygen and hydrogen peroxide, and analysis of *Arabidopsis* cell cultures has shown that aconitase is inhibited by hydrogen peroxide (Sweetlove *et al.,* 2002).

Aluminium toxicity and phosphorous deficiency have also been shown to result in a reduction in aconitase activity and an increase in citric acid levels (Hoffland *et al.,* 1992; Neumann and Romheld, 1998; Neumann *et al.,* 1999; Yang *et al.,* 2004). The increase in citric acid during aluminium toxicity is believed to result from aconitase inactivation (Yang *et al.,* 2004). This could be caused by the direct interaction of aluminium with aconitase or indirectly through the production of superoxide or hydrogen peroxide (Yang *et al.,* 2004). During conditions of phosphorus deficiency citric acid levels were elevated within roots and root excludes of white lupin (Neumann and Romheld, 1999). Citric acid and protons are believed to be released from root clusters to acidify the rhizosphere and mobilise any phosphorus present, in a similar manor to iron deficiency (Neumann *et al.,* 1999). Microarray analysis of the *Arabidopsis* aconitase isoforms conducted with seedlings showed that the mRNA

levels of ACO1, ACO2 and ACO3 were all reduced in roots under phosphorous deficient conditions (Hruz *et al.,* 2008). The reduction in mRNA levels of all three aconitase isoforms could therefore result in an increase in citrate, which in turn could suggest an involvement in plant responses to phosphorus and/or iron deficiency.

Materials and Methods

2.1 General solutions and equipment

All general laboratory chemicals and biological solutions were obtained from Sigma Aldrich Ltd. (Dorset, U.K.), Fisher Scientific U.K (Leicestershire, U.K.) or VWR International Ltd. (Leicestershire, U.K.) and were of Molecular biology grade or equivalent unless otherwise stated. All centrifugation steps were either conducted with bench-top Boeco centrifuge M-25 or a 12 ml Hettich EBA12 bench-top mircocentrifuge unless otherwise stated. Optical density measurements were conducted on a UniCam UV 500 spectrophotometer with Vision 32 Software (Thermo Spectronic, New York, U.S.A). Seedling photography was conducted with a Nikon D50 with an 18-55 mm lens and selective magnifying filters when needed. All primers were synthesised by MWG biotech and all PCR reactions were carried out in 200 µl thin-walled PCR tubes (VWR, Leicestershire, U.K.) in a PTC-150 MiniCycler (MJ Research Inc., Massachusetts, U.S.A.) unless otherwise stated. The pQE31-AAE7/ACN1 *E. coli* expression vector in M15[REP4] cells has been previously described (Turner *et al.*, 2005). Aconitase full-length cDNA clones were obtained from RIKEN plant sciences (http://www.brc.riken.go.jp/lab/epd/catalog/cdnaclone.html). ACO1 (MIPS No. At4g35830), ACO2 (MIPS No. At4g26970) and ACO3 (MIPS No. At2g05710) correspond to clones pda 03083, pda 04401 and pda 09788 respectively.

2.2 Structural and enzymatic characterisation of AAE7/ACN1

2.2.1 Induction and purification of AAE7/ACN1

Selective Luria Bertani (LB) broth (1 % (w/v) tryptone, 0.5 % (w/v) yeast extract, 0.02 M NaCl) with a final concentration of 100 μg ml^{-1} ampicillin and 50 μg ml^{-1} kanamycin was prepared and inoculated with *E. coli* expressing AAE7/ACN1. A 5 ml starting culture was incubated overnight at 37°C with gentle agitation. 250 ml LB was inoculated with the starting culture creating a starting OD_{600} of 0.075 and agitated at 250 rpm at 37°C until an OD_{600} of 0.6 was reached. With an OD_{600} of 0.6 the culture was induced by addition of filter sterilized isopropyl-beta-D-thiogalactopyranoside (IPTG) to a final concentration of 1 mM and incubated for 24 hrs at 20°C with gentle agitation.

Cells were harvested at 4°C in a Beckman centrifuge using a JA-14 rotor at 4000x g for 20 min. Pelleted cells were re-suspended in binding buffer (10 mM imidazole, 0.5 M NaCl, 20 mM phosphate buffer, pH 7.4), broken by sonication with a Sanyo Soniprep 150 MSE sonicator. Cells were exposed to 2 pulses for 30 seconds at an amplitude of 5 with approximately 30 seconds cooling in between episodes. Cell debris was removed by centrifugation in a Beckman centrifuge at 15000 x g for 10 min using a JA-20 rotor pre-cooled to 4°C. The cleared lysate was loaded onto a HisTrap Ni-chelating column (GE Healthcare Life Sciences, Buckinghamshire, U.K.). The column was washed sequentially with 10 ml of 10 mM imidazole buffer (0.5 M NaCl, 20 mM phosphate buffer, pH 7.4), 5 ml of 50 mM imidazole buffer, 5 ml of 100 mM imidazole buffer, and lastly eluted with 5 ml of 250 mM imidazole buffer and collected in 1 ml fractions. Aliquots from each fraction were run on a 10 % SDS-PAGE gel (Laemmli, 1970) to check the quality of purification. Suitable fractions were combined, dialysed against a solution of 0.02 M KH_2PO_4, pH 7.4, 0.05 M KCl and 10 % glycerol and concentrated by centrifugation in a Viva Spin 2 ml concentrator, 10,000 MWCO PES (Sartorius, Surrey, U.K.). Protein concentration was determined by the Bradford method (BioRad) using BSA as a quantification standard (Bradford, 1976).

2.2.2 Visualisation of proteins by gel electrophoresis

All protein samples were mixed with loading dye (250 mM Tris / HCl pH 6.8, 2 % SDS, 10 % glycerol, 20 mM DTT and 0.01 % bromophenol blue) and visualised on a

10 % SDS-PAGE gel to check for either protein induction or purity. SDS-PAGE gels consisted of a stacking gel (0.1 M Tris / HCl pH 6.8, 3 % Acrylamide / Bis acrylamide, 0.1 % SDS, 0.05 % ammonium persulphate, 0.01 % TEMED) cast with a comb for sample application over a set resolving gel (0.4 M Tris / HCl pH 8.8, 10 % Acrylamide / Bis acrylamide, 0.1 % SDS, 0.05 % ammonium persulphate, 0.04 % TEMED). Gels were run with 300 ml of bottom tank buffer (0.05 M Tris / HCl pH 8.5) and 200 ml of top tank buffer (0.03 M Tris base, 0.2 M glycine, 0.1 % SDS). SDS-PAGE gels were stained with Coomassie stain (50 % IMS, 0.2 % coomassie brilliant blue and 10 % acetic acid) for 30 min and de-stained (20 % IMS and 10 % acetic acid) overnight with gentle agitation. SDS-PAGE gels at 8 % acrylamide concentration were constructed in a similar way.

2.2.3 Assay of AAE7/ACN1 activity

Acetyl-CoA synthetase activity was determined in a coupled assay using citrate synthase and malate dehydrogenase (Figure 2.1). The rate of reaction was followed by monitoring NADH formed by the conversion of malate to oxaloacetate, which was withdrawn into the citrate synthase reaction. All assays were conducted at room temperature with a total assay volume of 1 ml which comprised of 0.1 M Tris, pH 8.0, 8 mM $MgCl_2$, 10 mM glutathione, 0.9 mM NAD, 0.04 mM CoA, 23 units malate dehydrogenase, 21 units citrate synthase, 10 mM L-malate, 10 mM NaAc, 5 mM ATP and AAE7/ACN1. The reaction was initiated by addition of ATP at a final concentration of 5 mM. In the AMP analogue inhibitor studies, the assay mix also included the appropriate concentration of AMP analogue (see chapter 3). K_m, $K_{m\ apparent}$, V_{max} and $V_{max\ apparent}$ were determined from Lineweaver burk plots. Inhibition constants for competitive and uncompetitive inhibition, Ki_c and Ki_u respectively, were determined using $Ki_c = [I] / ((K_{m\ apparent} / K_m) - 1)$ and $Ki_u = [I] / ((V_{max} / V_{max\ apparent}) - 1)$.

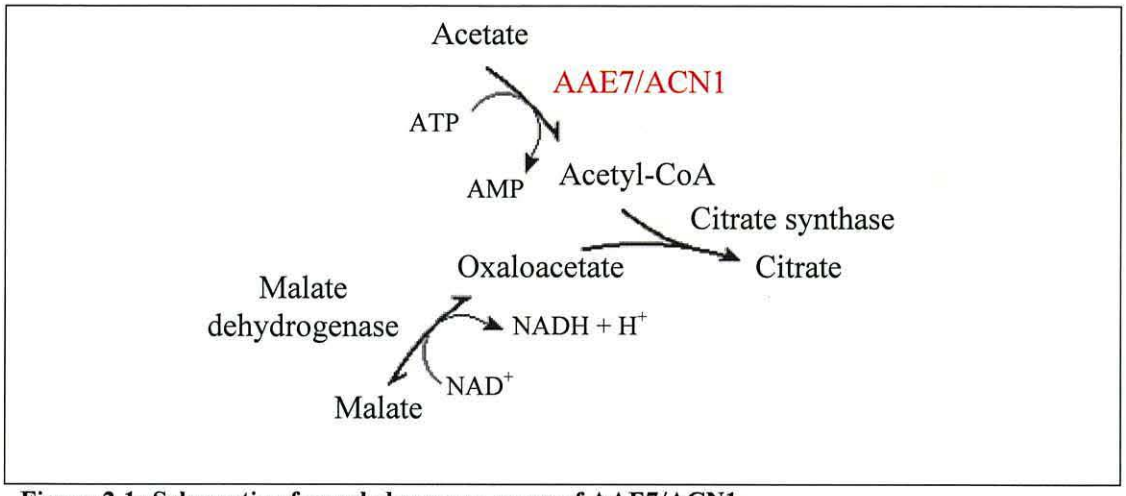


Figure 2.1: Schematic of coupled enzyme assay of AAE7/ACN1.

2.2.4 Calculating of the MgATP concentration within the enzyme assay of AAE7/ACN1

The concentration of ATP was varied in kinetic inhibition studies of AAE7/ACN1 to determine their mechanism of action. ATP forms a metal-ligand complex in solution with Mg^{2+} with a ratio of 1:1. Therefore altering the concentration of ATP affects the concentration of the MgATP complex. The stability constant of the metal-ligand complex is 20000 M^{-1} in 0.1 M Tris, pH 8.0 at 30°C (Purich and Fromm, 1972; Dawson *et al.,* 1987). The enzyme assay conditions set up for kinetic analysis of AAE7/ACN1 were also conducted with 0.1 M Tris, pH 8.0. However the lower temperature of 25°C was used. The dissociation constant of 20000 M^{-1} was corrected for the 5°C temperature difference to 19997 M^{-1} (equation 1) using the method derived by O'Sullivan and Smithers (1979).

Equation 1: Dissociation constant corrected to 25°C

The temperature at which the dissociation constant was measured at (Tm) = 303 k
The experimental temperature (Tc) = 298 k

$\Delta \log k = 750 (\Delta T)$
Therefore $\Delta T = {}^{1}/_{303} - {}^{1}/_{298}$
$\Delta T = -5.537 \times 10^{-5}$
$\Delta \log k = -5.537 \times 10^{-5}$
$= 0.9998$
$= 0.9998 \times 20000$
k at 25°C = 19997 M^{-1}

Due to the small difference observed between the dissociation constant calculated at 25°C and that obtained at 30°C the dissociation constant used to calculate the concentration of MgATP was 20 mM^{-1}. The concentration of free ATP was estimated (equation 2) and subtracted from the concentration of ATP (Dawson *et al.*, 1987).

Equation 2:
free ATP = [ATP] / (20 ([Mg^{2+}] - [ATP])

Magnesium ions bind to the nucleosides tri-, di- and mono-phosphate with decreasing degrees of tightness (Igamberdiev and Kleczkowski, 2001). This difference in affinity for tri- di- and mono-phosphates is reflected in the dissociation constants 20000, 2000 and 100 M^{-1} (Purich and Fromm, 1972). The weak binding of magnesium to the mono-phosphate results in the metabolic participation of AMP in a magnesium free form (Igamberdiev and Kleczkowski, 2001).

2.2.5 Crystallisation trials of AAE7/ACN1

Purified AAE7/ACN1 was prepared as described in section 2.2 and concentrated as detailed in section 2.2.1. AAE7/ACN1 crystallisation trials were set up using the Structure screen 1 kit (Molecular Dimensions Limited, Cambridgeshire, U.K.) at 4°C

and 18°C. All trials were set up using the hanging drop method with 2 µl protein mixed with an equal volume of reservoir solution. The protein and reservoir solution were mixed on a cover slip that was inverted and sealed over 0.5 ml of reagent. Initial crystallisation trials were conducted with 15 mg ml^{-1} purified AAE7/ACN1.

A second crystallisation trial was conducted with the Structure screen 1 kit with 16 mg ml^{-1} purified AAE7/ACN1. Crystallisation solutions were incubated at 4°C and 18°C both in the presence and absence of 1 mM AMPP and 1 mM CoA. Due to the observation of putative crystals with the reservoir solution 2.0 M ammonium sulphate and 0.1 M Tris, pH 8.5 further crystallisation trials were conducted. Trials were set up with 0.5 – 3 M ammonium sulphate (at 0.5 M increments) at 4°C and 18°C with 18 mg ml^{-1} AAE7/ACN1. Protein solutions were incubated with a) 1 mM AMPP and 1 mM CoA, b) 1 mM AMPP and c) 1 mM CoA. Crystallisation trials were repeated at 4°C with 10 mg ml^{-1} and 15 mg ml^{-1} AAE7/ACN1 incubated with 1 mM AMPP and 1 mM CoA. The concentration range of ammonium sulphate was narrowed to 2 – 2.5 M at 0.1 M increments.

The hanging drop method with 8 mg ml^{-1} AAE7/ACN1was modified from Gulick *et al.* (2003) with a reservoir solution of 10 – 14 % PEG 8000, 10 % ethylene glycol and 50 mM MES, pH 6.5. The protein solution was incubated with a) 1 mM AMPP and b) 1 mM AMPP plus 1 mM CoA. After 5 days the droplets were microseeded with a cat whisker.

2.2.6 Determining the composition of the putative crystals

Putative crystals were tested to determine if they were comprised of purified protein or salt. The crystal's strength and ability to take up the dye bromophenol blue was tested. The strength of the crystals was tested with a needle in an attempt to break them. One µl of 2.5 % bromophenol blue was added to the crystallisation solution to distinguish protein from salt. Crystals that took up the blue dye were classified as protein crystals.

2.2.7 Crystallisation of Lysozyme

Lysozyme was crystallised using the hanging drop method. Two μl of protein (10 - 40 mg ml^{-1}) was mixed with an equal volume of reservoir solution (5 - 8 % NaCl, 0.1 M sodium acetate, pH 4.2) and incubated at 18°C.

2.2.8 Small angle x-ray scattering

Small angle x-ray scattering (SAXS) is a technique used in the study of biological macromolecules to determine their overall size and shape in solution. SAXS measurements were collected at angles close the primary beam by measuring the scattered intensity I(s) as a function of the momentum transfer q ($q = 4\pi\sin\theta/\lambda$, where 2θ is the angle between incident and scattered radiation). SAXS allows the structure of a native protein under near physiological conditions to be analysed with changing external conditions at low (1-2 nm) resolution. Dilute protein solutions are random orientations of identical particles that are averaged to produce the scattering profile of a single particle. The resulting scattering profile can be used to predict a protein's overall shape. This is possible as each profile has a unique signature dependent on the size and shape of the macromolecule. Theoretical x-ray scattering curves can be predicted from crystallographic data. Predicted curves are distinguishable in the small angle region but not the high angle region (Figure 2.2). This demonstrates that SAXS data contains information about the protein's shape and fold but not its secondary structure (Svergun *et al.*, 2001).

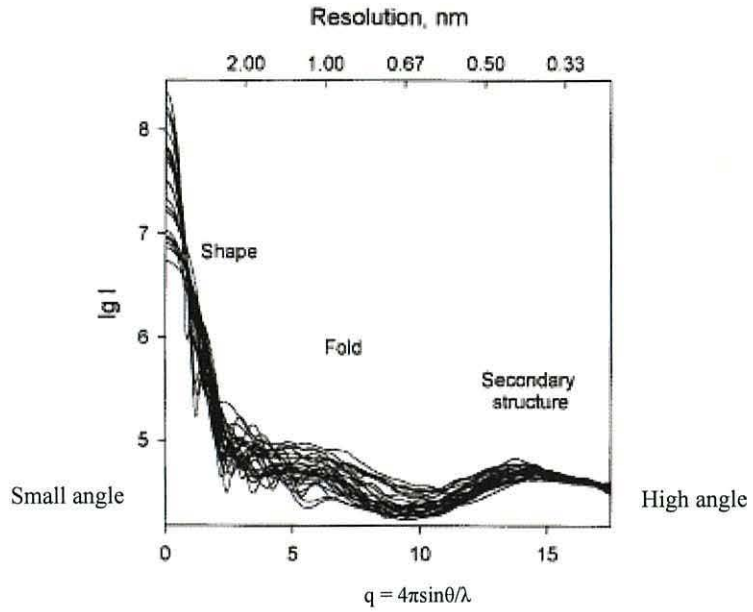


Figure 2.2: Theoretical x-ray scattering curves predicted from 25 different proteins (Svergun *et al.*, 2001).

2.2.9 Small angle x-ray scattering sample preparation and data collection of AAE7/ACN1 in the absence and presence of adenosine monophosphate propyl ester

A purified sample of AAE7/ACN1 was prepared as described in section 2.2.1. Purified, buffer exchanged protein was concentrated to 10 mg ml^{-1} and specific activity determined as described in section 2.2.3. Small angle scattering data was collected at the synchrotron radiation source Daresbury, Beamline 2.1 (Towns-Andrew *et al.*, 1989). Data was collected at a fixed wavelength of 1.54 Å with beam currents between 124 and 184 mA. Diffraction data was collected with 1 mg ml^{-1}, 2 mg ml^{-1} and 5 mg ml^{-1} purified AAE7/ACN1 at 5.5 m and with 10 mg ml^{-1} AAE7/ACN1 at 1 m camera length. Data was also collected at 5.5 m with 1 mg ml^{-1} and 2 mg ml^{-1} AAE7/ACN1 and at 1 m camera length with 10 mg ml^{-1} AAE7/ACN1 in the presence of 0.45 mM AMPP, a propyl ester of adenosine monophosphate. Low and high angle data was collected at two camera lengths (5 m and 1 m) covering a scattering vector range of 0.01 $Å^{-1} < q > 0.15$ $Å^{-1}$ ($q = 4\pi\sin\theta/\lambda$, where 2θ is the scattering angle). Frames of 60 second bursts were collected and averaged creating a total exposure time of 20 – 45 min per sample. The first and last frames of data collected were compared to ensure that no aggregation or radiation damage occurred during data collection. Samples were prepared and 100 µl transferred to a 1

mm mica cell and data collected at 4°C. The data was calibrated with either silver behenate, for high angle data (1 m camera length) or wet rat tail collagen, for low angle data (5.5 m camera length) and reduced from a 2-dimensional image to a 1-dimensional scattering profile using Beamline software (OTOKO, SRS Daresbury, U.K.). During data reduction the profile was normalised to the intensity of the incident beam, corrected for detector artefacts and the scattering buffer subtracted (OTOKO, SRS Daresbury, U.K.). Resulting data sets, collected at different camera lengths, were merged with Microsoft Excel.

2.2.10 Determining the Rg and D*max* of AAE7/ACN1

Guinier analysis gives a linear relationship between the natural log of the scattering intensity and the square of the scattering vector, at low angles (data collected at 5.5 m camera length). This region is also known as the Guinier region and is used to calculate the radius of gyration (R_g). The R_g is an estimate of the mass distribution of a molecule and is determined from the gradient of the Guinier region, $R_g^2 / 3$. The R_g was also estimated from the merged scattering profile with the indirect transform program GNOM (Svergun, 1992).

GNOM was also used to determine the maximum particle dimension (D*max*). The D*max* is the distance (r) at which the distance distribution function p(r) reaches zero (Figure 2.3). The p(r) function is very useful, unlocking valuable information about particle shape. Molecules with the same maximum size can have very different geometrical shapes. Differences in the shape of the curve and the maximum position are important structural markers. A globular protein has a bell shaped distribution with a maximum at D*max* / 2. An elongated structure has a maximum at a small distance corresponding to the particle's radius. Flat particles have a wide maximum that is smaller than D*max*. Figure 2.3 demonstrates the differences in distance distribution function of a variety of geometrical shapes.

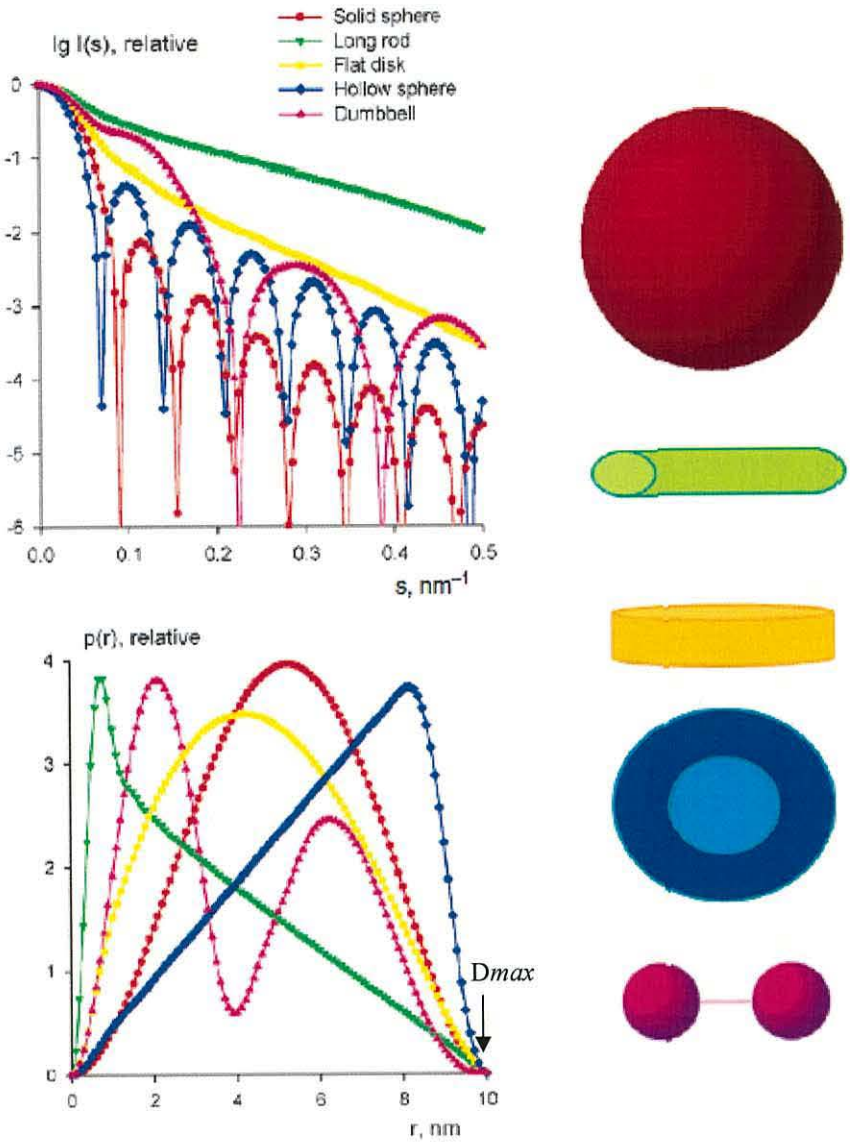


Figure 2.3: Scattering intensities and distribution functions of geometric bodies (Svergun and Koch, 2003).

2.2.11 Shape reconstruction of AAE7/ACN1

Low resolution shape reconstruction was performed using the *ab initio* program GASBOR (Petoukhov and Svergun, 2003). The program GASBOR reads the solution scattering profile and p(r) output files from GNOM (Svergun *et al.*, 2001). GASBOR utilises dummy residues to construct a chain compatible model with a specified number of amino acid residues. The dummy residues are separated by a distance of 0.38 nm at a resolution of 0.5 Å inside a sphere of diameter D*max*. The program calculates and compares scattering profiles from thousands of

configurations of dummy residues to the measured scattering pattern by randomly moving dummy residues.

Sixteen iterations of GASBOR were implemented, providing sixteen models for analysis by the DAMAVER suite (Kozin and Svergun, 2000; Volkov and Svergun, 2003). The resulting models were superimposed and averaged to obtain a representative model. DAMAVER aligns two models, represented by ensembles of points, and provides a dissimilarity measure called normalized spatial discrepancy (NSD) between them. The NSD is compared between models and a reference model, with the smallest average dissimilarity measure being selected for subsequent averaging.

2.2.12 Techniques implemented to determine the multimeric state of AAE7/ACN1

2.2.12.1 CRYSOL prediction and comparison of solution scattering profiles

CRYSOL (Svergun *et al.,* 1995) was implemented to predict the solution scattering profile from known crystal structures and compare them to AAE7/ACN1 SAXS data. Namely, these were monomeric acetyl-CoA synthetase (1PG4.pdb) from *Salmonella enterica* (seACS) (Gulick *et al.,* 2003), dimeric long chain fatty acyl-CoA synthetase from *Thermus thermophilus* (ttLC-FACS) (1ULT.pdb) (Hisanaga *et al.,* 2004) and trimeric *Saccharomyces cerevisiae* acetyl-CoA synthetase (scACS) (1RY2.pdb) (Jogl and Tong, 2004).

CRYSOL (Svergun *et al.,* 1995) calculates the theoretical x-ray scattering pattern from the atomic coordinates of a crystal structure and then alters the volume and hydration contact parameters to obtain the best fit to the experimental SAXS data.

The scattering intensity from a particle in solution is:

$$I(s) = ((A_a(s) - p_s A_s(s) + \delta p_b A_b(s))2)\Omega$$

Where $A_a(s)$ is the atomic scattering in vacuum, $A_s(s)$ is the scattering from the excluded volume, $A_b(s)$ is the scattering from the hydration shell, p_s is the electron density of the solvent and δp_b is the contrast in the salvation shell.

The level of confidence in the data fit is determined by chi-square analysis (χ). χ is a measure of the 'goodness of fit' of the theoretical data created from a crystal structure and experimental SAXS data. The lower the χ value returned by CRYSOL the better the theoretical data fits the experimental data.

χ is calculated by:

$$\chi^2 (r_0, \delta\rho) = 1/N\rho \, \Sigma \, [I_e(S_i) - cI(S_i,r_0,\delta\rho) / \sigma(S_i)]^2$$

Where $cI(S_i,r_0,\delta\rho)$ is the SAXS intensity, $I_e(S_i)$ is the experimental curve, $N\rho$ is the number of experimental points and $\sigma(S_i)$ are the experimental errors.

2.2.12.3 Native polyacrylamide gel electrophoresis to determine the molecular weight of AAE7/ACN1

Native PAGE was conducted in a similar way to SDS-PAGE (see section 2.2.2). The only modifications were the removal of SDS from the stacking and resolving gel, the loading dye and the running buffer. Native PAGE gels were run with either the running buffers stated in section 2.2.2, minus the SDS or with 0.05 M Tris / HCl at pH 7.3, 8.8 or 10.0.

2.2.12.4 Isoelectric focusing to determine the isoelectric point of AAE7/ACN1

Purified AAE7/ACN1 was prepared as described in section 2.2.1 and buffer exchanged on a PD10 column with 25 mM Tris / HCl, pH 8.0. Isoelectric focusing (IEF) was conducted with a PhastGel IEF pH 3-9 (GE Healthcare Life Sciences, Buckinghamshire, U.K.). The gel was pre-focused for 75 Vh (2000 V, 2.5 mA, 3.5 W) prior to sample application. The sample applicator, containing 5 ng AAE7/ACN1, was lowered onto the centre of the gel and run for 15 Vh (200 V, 2.5 mA, 3.5 W). The applicator was raised and the gel focused for 410 Vh (2000 V, 2.5 mA, 3.5 W).

IEF gels were fixed (20 % trichloroacetic acid) for 5 minutes, washed for 2 minutes in solution A (50 % ethanol, 10 % acetic acid) and 6 minutes in solution B (10 % ethanol, 5 % acetic acid). Gels were submerged in 8.3 % glutaraldehyde for 6 minutes, washed with solution B for 8 minutes and washed in water for 4 minutes. Gels were stained for 10 minutes with 0.5 % (w/v) silver nitrate, developed with

0.015 % formaldehyde in 2.5 % sodium carbonate for 4 minutes and stopped with 5 % acetic acid.

2.2.12.5 Determining the molecular weight of AAE7/ACN1 by gel filtration

AAE7/ACN1 was purified as described in section 2.2 and buffer exchanged as described in section 2.2.1. A 100 µl sample, containing approximately 0.26 mg AAE7/ACN1, was loaded onto a Superose 12 HR 12/30 column and run at 0.2 ml min^{-1} on a FPLC. All buffers were filtered and de-gassed prior to use. β-galactosidase, BSA and cytochrome C were run as standards.

2.2.12.6 Electro spray ionisation mass spectrometry to determine the molecular weight of AAE7/ACN1

AAE7/ACN1 samples prepared as described in section 2.2.1 were dialysed into 10 mM ammonium formate and analysed at 10 µM under non-native (50 % methanol, 0.1 % formic acid) and native (10 mM ammonium formate plus 10 % methanol) solvent conditions. Mass spectra were obtained on a Q-ToF Ultima Global and the deconvoluted mass spectrum (mass scale) from the native ESI data were generated by the transform algorithm in Masslynx (Waters Micromass Technologies, Manchester, U.K.) by Dr M. McCullough.

2.2.13 Construction of a homology model

Homology modelling is a method implemented to construct a 3D model from a protein's amino acid sequence based on similar crystallographic structures. The quality of a predicted model will depend on the quality of the template and the alignment of the primary sequences. Comparative models are useful as small changes in a protein's sequence usually results in small changes in structure (Sali and Overington, 1994). Homology models are valuable tools used in the identification of conserved residues and the prediction of important active site residues. The creation of a comparative model involves the identification of a suitable template, target - template sequence alignment followed by model creation and evaluation (Marti-Renom *et al.,* 2000).

The online server Phyre (Kelley and Sternberg, 2009) was used to identify a suitable template structure for homology modelling. This was achieved by comparing the

amino acid sequence of AAE7/ACN1 to known crystallographic structures. The amino acid sequence of the chosen template structure (ttLC-FACS) was aligned to AAE7/ACN1 with the default settings of Multalin (Corpet, 1988) and a dimeric homology model was constructed using the default settings of Modeller 9v1.

Modeller 9v1 constructs a comparative model by satisfaction of spatial restraints (Sali and Blundell, 1993; Sali *et al.,* 1995; Marti-Renom *et al.,* 2000). The restraints are derived from the alignment of the template and target sequences. Restraints include C-C distances, main chain N-O distances and main chain and side chain dihedral angles. In the core region there may be a high degree of sequence identity between the unknown protein and the template. This could result in the direct transfer of side chain conformations from the template structure. Within the core region amino acid changes are often very conservative, for example phenylalanine for tyrosine. High sequence identity between the template and target sequence also allows side chains to be easily modelled. Where there is less conservation between the amino acid sequences then the side chains must be added without reference to the template. Once the backbone conformations of the protein have been defined it is then possible to assign conformations to the side chains (Sali and Blundell, 1993).

2.3 Cloning and subsequent expression of *Arabidopsis thaliana* aconitase genes, ACO1, ACO2 and ACO3 in *Pichia pastoris*

The three full length aconitase cDNAs ACO1 (At4g35830), ACO2 (At4g26970) and ACO3 (At2g05710) were sub-cloned from the pda vector into *E. coli* prior to cloning into pPiczC and transformation into *Pichia pastoris* X-33 cells for protein expression. Glycerol stocks were made at every stage of the sub-cloning process by adding 15 % glycerol to 300 µl of an overnight liquid culture to a final volume of 1 ml. All glycerol stocks were stored at -80°C. Primer sequences used throughout the cloning process are listed in Table 2.1.

Table 2.1: Primers required for amplification and cloning of aconitase isoforms in *Pichia pastoris* (F = forward and R = reverse).

Primer name	Primer sequence 5' – 3'
AOX1-F	GACTGGTTCCAATTGACAAGC
AOX1-R	GCAAATGGCATTCTGACATCC
M13-F	GTAAAACGACGGCCAG
M13-R	CAGGAAACAGCTATGAC
ppAco26970-1F	ATATATATATGGTACCCAAAATGTCTCGACGCGCCACT
ppAco26970-1R	ATATATATATCCGCGGCTTGGCGCTCAAACTCC
ppAco35830-1F	ATATATATATGGTACCCAAAATGTCTTCCGAGAATCCTTT
ppAco35830-1R	ATATATATATCCGCGGTTGTTTGATCAAGTTCCTGATAACG
ppAco05710-1F	ATATATATATCTCGAGCAAAATGTCTTTAACCGCTTCATCTTCC
ppAco05710-1R	ATATATATATCCGCGGTTGCTTGCTCAAGTTTCTGATAAC

2.3.1 Sub-cloning and transformation into the *E. coli* vector Topo2.1

Prior to transformation into *P. pastoris* the aconitase isoforms were first sub-cloned into *E. coli*.

2.3.2 Preparation of electrocompetent *E. coli* DH5α

A volume of 200 ml of LB was inoculated with an overnight culture of DH5α to a starting density ($O.D_{600}$) of 0.05 and agitated at 250 rpm at 37°C until an $O.D_{600}$ of 0.6 was reached. Cells were harvested by centrifugation at 4000 x g for 15 min at 4°C (Beckman coulter, JA-14) and re-suspended in 50 ml 1 mM Hepes pH 7.4. Cells were centrifuged at 4000 x g for 15 min, re-suspended in 20 ml 1 mM Hepes pH 7.4 and centrifuged at 4000 x g for 15 min. Cells were re-suspended in 10 ml 10 %

glycerol and centrifuged at 4000 x g for 15 min (Beckman coulter, JA-20). The supernatant was removed, cells were re-suspended in 400 µl 10 % glycerol, dispensed in 50 µl aliquots, frozen on dry ice and stored at -80°C.

2.3.3 Amplification of aconitase genes ACO1, ACO2 and ACO3

The cDNAs of ACO1 and ACO2 were amplified with 1 unit Phusion DNA polymerase and 1 x HF PCR buffer (New England Biolabs, Hertfordshire, U.K.) and ACO3 with 2 units Accuzyme DNA polymerase and 1 x PCR buffer (Bioline, London, U.K.). The PCR mix also contained 0.5 µM of the appropriate forward and reverse aconitase gene primer pair (Table 2.1), 0.2 mM dNTP and 10 ng of the corresponding pda clone in a final volume of 25 µl. PCR cycling conditions were 1 cycle of 98°C for 30 seconds; 35 cycles of 98°C for 15 seconds, 59°C for 30 seconds and 72°C for 45 seconds. A final elongation step of 72°C for 10 min and a holding temperature of 4°C were included. PCR products were checked by agarose gel electrophoresis and the bands extracted and cleaned prior to ligation and transformation.

2.3.4 Phenol / chloroform / IAA extraction and ethanol precipitation for cleaning PCR products prior to ligation

An equal volume of phenol / chloroform / IAA (25:24:1) was added to the isolated DNA sample and mixed well by vortexing. The sample was centrifuged at 13000 x g for 4 min and the supernatant transferred to a fresh tube. An equal volume of chloroform / IAA was added, mixed well, centrifuged at 13000 x g for 4 min and the top aqueous layer transferred to a fresh tube. The previous step was repeated to remove traces of phenol. A 1/10 volume of sodium acetate, pH 5.0 and two volumes of ice-cold 100 % ethanol were added and incubated at -20°C for 2 hours to precipitate the DNA. The mixture was centrifuged for 10 min at 13000 x g, the supernatant discarded and the DNA pellet washed twice with 70 % ethanol followed by centrifugation for 3 min at 13000 x g between each wash. The pellet was air dried, re-suspended in 10 µl distilled H_2O and stored at -20°C.

2.3.5 Cloning into the TOPO2.1 vector

The protocol for T/A cloning of PCR amplified fragments was derived from the TOPO2.1 PCR cloning manual. Adenine overhangs were added to blunt ended PCR fragments by incubating in 1 x PCR buffer, 3 mM $MgCl_2$, 0.2 mM dATP and 1 unit Taq DNA polymerase (Promega Corporation Southampton, U.K.) for 20 min at 72°C. The resulting A-tailed PCR product was incubated for 5 min at room temperature with 1 µl salt solution (kit reagent) and 1 µl TOPO2.1 vector (Invitrogen, Paisley, U.K.) in a total volume of 6 µl. Two µl of the TOPO2.1 cloning reaction was added to one vial of Top10 chemically competent cells (Invitrogen, Paisley, U.K.) and incubated on ice for 30 min. Cells were heat shocked for 30 seconds at 42°C and incubated on ice for 1 min. A volume of 250 µl of S.O.C medium (20 g l^{-1} tryptone, 5 g l^{-1} yeast extract, 0.5 g l^{-1} NaCl, 250 mM KCl, 20 mM glucose, 10 mM $MgCl_2$) was added and the tubes gently agitated at 37°C for 1 hour. Two volumes of 50 µl and 100 µl were spread onto pre-warmed LB plates containing 100 µg ml^{-1} ampicillin and incubated at 37°C overnight. Resulting colonies were re-streaked onto fresh LB plates containing 100 µg ml^{-1} ampicillin and analysed for positive transformants containing the appropriate insert by colony PCR with M13-F and ppAco-1R primers (Table 2.1).

2.3.6 Colony PCR to identify positive transformants

E. coli and *P. pastoris* colonies were tested by adding a small amount of a colony to a 10 µl PCR reaction containing 1x PCR buffer, 0.8 units Taq DNA polymerase, 3 mM $MgCl_2$, 0.2 mM dNTP's, 0.5 µM forward primer and 0.5 µM reverse primer. Colony PCR conditions consisted of a denaturing step of 95°C for 2 min; 35 cycles of a denaturing step of 94°C for 30 seconds, an annealing step of 58°C for 30 seconds and an elongation step of 72°C for 2.5 min; a final elongation step of 72°C for 5 min and a holding temperature of 4°C. Agarose gel electrophoresis (1 %) was used to visualise DNA. Plasmids were purified from positive colonies with either the PureLink™ Mini Plasmid purification kit (Invitrogen, Paisley, U.K.), as manufacturer's instructions, or by alkaline lysis (see section 2.3.7).

2.3.7 Isolation of plasmids from *E. coli* using alkaline lysis

Cells were harvested by centrifugation at 13000 x g for 1 min in a 1.5 ml microfuge tube from an overnight culture incubated at 37°C. The supernatant was discarded and the pellet re-suspended in 100 µl of an ice-cold solution of Buffer 1 (50 mM glucose, 10 mM EDTA and 25 mM Tris, pH 8.0). After incubation at room temperature for 5 min, 200 µl of a freshly prepared solution (Buffer 2) containing 0.2 M NaOH and 1 % SDS was added. The contents were mixed by inverting the tube three times and the tube placed on ice for 5 min. A volume of 150 µl of an ice-cold solution of Buffer 3 (3 M potassium, 5 M acetate) was added and the tubes gently vortexed in an inverted position. The tubes were stored on ice for 5 min then centrifuged for 5 min at 13000 x g. The supernatant was transferred to a fresh microfuge tube containing 450 µl of phenol/chloroform/IAA (25:24:1), vortexed and centrifuged for 2 min. The supernatant was transferred to a fresh 1.5 ml microfuge tube containing 900 µl of 100 % ethanol, incubated on ice for 2 min, and followed by centrifugation at 13000 x g for 10 min. The supernatant was discarded and the pellet washed with 500 µl of 70 % ethanol by vortexing. The DNA was re-pelleted by centrifugation for 5 min at 13000 x g. The supernatant was discarded and the pellet air-dried by inverting the open tube for 20 min. DNA was re-suspended in 20 µl of TE, pH 8.0, and incubated at 37°C for 20 min after addition of 20 µg ml^{-1} RNase.

2.3.8 Restriction digest to create compatible ends for ligation into pPiczC

Double digests of both the purified TOPO2.1 plasmids containing an aconitase gene and the pPiczC vectors were conducted with the appropriate restriction enzymes (Table 2.2) by incubation at 37°C for 1 hr with 1 x buffer B (10 mM Tris / HCl pH 7.5, 10 mM $MgCl_2$, 0.1 mg ml^{-1} BSA), 20 µl of plasmid and 5 units of each restriction enzyme in a total volume of 50 µl. The resulting digests were visualised on agarose gels, excised, and purified by centrifugation at 13000 x g for 15 min in a Whatman microfuge tube with a 10 µm polyprolene mesh (VWR International Ltd., Leicestershire, U.K.). Several digests were combined to provide enough DNA to proceed with ligation reactions. The combined, isolated DNA samples were purified with either a phenol / chloroform / IAA extraction followed by ethanol precipitation (see section 2.3.4) or with a QIAquick PCR purification kit according to manufacturer's instructions with the following modifications (QIAGEN Ltd., West

Sussex, U.K.). DNA was eluted with two applications of 15 µl elution buffer by centrifugation for 1 min. The eluted DNA was freeze-dried (Edwards Modulyo, Pirani 10) and re-suspended in 10 µl autoclaved distilled water and an aliquot visualised on a 1 % agarose gel.

Table 2.2: Double digest combinations for cloning into the *P. pastoris* vector pPiczC and linearization of pPiczC for transformation into *P. pastoris*.

	Restriction enzyme			
	Xhol I	Sac II	Kpn I	PmeI
ACO1		X	X	
ACO2		X	X	
ACO3	X	X		
pPiczC vector	X	X		
pPiczC vector		X	X	
pPiczC plus insert				X

2.3.9 Ligation and transformation of the pPiczC vector into *E. coli*

Ligation of the purified, restricted genes and pPiczC vector was conducted overnight in a 10 µl reaction volume, incubated at room temperature. The reaction contained 4.5 µl of pPiczC, 4 µl of restricted DNA, 3 units of T4 DNA ligase (Promega Corporation Southampton, U.K.) and 1 µl of 1 x ligation buffer (30 mM Tris, pH 7.8, 10 mM $MgCl_2$, 10 mM DTT and 10 mM ATP). A 3 µl aliquot was incubated with 50 µl DH5α electrocompetent cells on ice for 30 min, and then pulse shocked using a BIO RAD Micropulser at 2.5 kv immediately followed by the addition of 1 ml of S.O.C medium. The transformed cells were incubated at 37°C with gentle agitation for 1 hr. Cells were harvested by centrifugation at 10000 x g for 1 min and the supernatant discarded. Cells were re-suspended in 250 µl of S.O.C medium and 150 µl and 100 µl were spread onto pre-warmed LB plates containing 25 µg ml^{-1} zeocin, before overnight incubation at 37°C. Resulting colonies were tested by colony PCR (see section 2.3.6). Positive plasmids were isolated and transformed into *Pichia pastoris* X-33 electrocompetent cells.

2.3.10 Preparation and transformation of electrocompetent *Pichia pastoris* X-33 cells with the pPiczC vector

A volume of 500 ml of YTP (1 % yeast extract, 2 % tryptone, 2 % glucose) was inoculated with 5 ml of an overnight culture and agitated at 200 rpm at 28°C until

and $O.D_{600}$ of 1.5 was reached. The cells were harvested by centrifugation at 1500 x g for 5 min at 4°C (Beckman coulter, JA-14) and re-suspended with 250 ml ice-cold sterile water. Cells were harvested by centrifugation at 1500 x g for 5 min, re-suspended with 20 ml ice-cold 1 M sorbitol and centrifuged at 1500 x g for 5 min in a Beckman coulter centrifuge using a pre-cooled JA-20 rotor. Cells were re-suspended in 1 ml of ice-cold sorbitol and divided into 50 µl aliquots. Freshly prepared electrocompetent cells were kept on ice and used the same day.

Prior to transformation into *P. pastoris*, the pPiczC positive constructs were linearised with PmeI (Fermetas U.K., York, U.K.) at 37°C for 1 hr by restriction digest as previously described in section 2.3.8. Two µg of linearised plasmid DNA were added to 50 µl of electrocompetent X-33 *P. pastoris* cells, incubated on ice for 20 min and pulsed on a Bio-Rad MicroPulser at 2.0 kv. One ml of ice-cold 1 M sorbitol was immediately added to the cuvette, the contents transferred to a sterile 1.5 ml microfuge tube and incubated at 28°C for 2 hours. The resulting cells were harvested by centrifugation at 10000 x g for 1 min. A volume of 800 µl of the supernatant was removed and the pellet re-suspended in the remaining 250 µl, which was subsequently spread onto two pre-heated YTPS plates (1 % yeast extract, 2 % tryptone, 1 M sorbitol and 2 % agar) containing 100 µg ml^{-1} zeocin and incubated at 28°C for two to three days. Resulting colonies were streaked onto a fresh YTPS plate and analysed by colony PCR, see section 2.3.6.

2.3.11 Small scale induction analysis for the identification of colonies expressing aconitase

P. pastoris cells were inoculated from an overnight culture of YTD (1 % yeast extract, 2 % peptone, 2 % glucose) and grown in 50 ml minimal media (1.34 % yeast nitrogen base with ammonium sulphate, 4 x 10^{-4} % biotin, 0.5 % methanol) at 28°C for 24 hrs with shaking at 200 rpm. Cells were harvested by centrifugation at 13000 x g for 5 min and the cells re-suspended in 5 ml of binding buffer (10 mM imidazole, pH 7.4, 500 mM NaCl and 20 mM phosphate). Re-suspended cells were divided into 1.5 ml microfuge tubes with an equal volume of acid washed glass beads. The mixture was vortexed four times for forty-five seconds with a thirty second pause on

ice to prevent over-heating the samples. The vortexed solution was removed from the glass beads and centrifuged for 5 min at 13000 x g at room temperature.

The resulting pellet was re-suspended in 10 µl of loading buffer and analysed by SDS-PAGE. A volume of 200 µl of supernatant was mixed with 800 µl of 100 % acetone and left on ice for 1 hr to precipitate the proteins. Precipitated proteins were pelleted by centrifugation at 13000 x g for 15 min, the supernatant removed and the protein pellet re-suspended in 20 µl of loading buffer. Samples were boiled for 5 min prior to loading onto an 8 % SDS-PAGE gel.

2.3.12 Large scale induction analysis for purification of ACO1 and ACO2

Cell cultures were agitated in minimal media (1.34 % yeast nitrogen base with ammonium sulphate, 4 x 10^{-4} % biotin, 0.5 % methanol) at 100 rpm at 28°C for 7 days. Ferrous ammonium sulphate was added to a final concentration of 0.06 mM at 48, 96 and 144 hr, and induction of the protein expression instigated with 0.5 % methanol every 24 hrs. Cells from 6 x 0.5 L of minimal media were harvested by centrifugation at 1500 x g for 5 min at 4°C (Beckman coulter, JA-14). Pelleted cells were re-suspended in binding buffer (10 mM imidazole, 0.5 M NaCl, 20 mM phosphate buffer, pH 7.4) and ruptured by sonication on a Sanyo Soniprep 150 MSE sonicator. Rupturing cells consisted of 3 pulses for 5 min at an amplitude of 10. Lysed cells were centrifuged at 15000 x g for 10 min in a Beckman Coulter centrifuge, equipped with a JA-20 rotor, and the cleared lysate was purified on HisTrap Ni-chelating columns (GE Healthcare Life sciences, Buckinghamshire, U.K.). Protein was eluted with 5 ml of 100 mM imidazole buffer (20 mM phosphate, pH 7.4, 0.5 M NaCl, 100mM imidazole) and the eluted fractions were desalted on a pd10 desalting column (GE Healthcare Life Sciences, Buckinghamshire, U.K.) equilibrated with 1 M Tris pH 7.4. The elutes were concentrated with a Viva Spin 2 ml concentrator 10,000 MWCO PES (Sartorius, Surrey, U.K.). Protein concentration was determined by the Bradford method (BioRad, U.K.) using BSA as quantification standard (Bradford, 1976).

2.3.13 Enzymatic assay for measuring aconitase activity

Aconitase activity was determined in a coupled assay modified from Rose and O'Connell (1967) using isocitrate dehydrogenase. The rate of reaction was followed by monitoring NADH formed by the conversion of isocitrate to α-ketoglutarate. All assays were conducted at room temperature with a total assay volume of 1 ml that comprised of 100 mM triethanolamine, pH 8.0, 1 mM magnesium chloride, 1 mM NADP, 0.7 units isocitrate dehydrogenase, 1 mM *cis*-aconitic acid and aconitase. The reaction was initiated by adding *cis*-aconitic acid to a final concentration of 1 mM. The assay procedure was tested using commercial porcine heart aconitase.

A crude extract of porcine heart aconitase was purchased from Sigma Aldrich Ltd (Dorset, U.K.) as a control to identify possible techniques for activation of At4g26970 and At4g35830. Half a ml of a 1.4 mg ml^{-1} solution was prepared in distilled water, activated for 1 hr with 0.02 mM ferrous ammonium sulphate and 2.3 mM cysteine at 4°C (Morrison, 1954) and assayed.

2.4 Preparation of *Arabidopsis thaliana* plant material for identification of *aco1*, *aco2* and *aco3* homozygous mutants and metabolic analysis

2.4.1 Plant material

T-DNA insertion lines were obtained through the Nottingham *Arabidopsis* Stock Centre (NASC) from the SIGnAL (Alonso *et al.*, 2003) and GABI-KAT (Rosso *et al.*, 2003) projects. Mutants in the three aconitase genes ACO1, ACO2 and ACO3 were identified from GABI-KAT for ACO1 (138A08) and ACO3 (298E10) and from SALK for ACO2 (090200). The SALK and GABI-KAT mutants were generated in a Columbia (Col-0) background using the vector pROK2 and pAC161 respectively. The pROK2 vector contains a kanamycin resistance gene and the pAC161 vector contains a sulfonamide resistance gene. Along with PCR analysis, these were used to confirm that the seedlings were mutants. Only one copy of the T-DNA was confirmed by obtaining mendelian inheritance ratios for homozygous and heterozygous lines. The SALK line 090200 (ACO2) was backcrossed to the parent line, Col-0.

2.4.2 Preparation of plant material and growth conditions

Seeds were surface sterilised with Covclor 1000 chlorine in a 1.5 ml microfuge tube. Sterilisation solution was produced by dissolving 0.125 g ml^{-1} Covclor in distilled water with 1 drop of 1 % tween20. This stock solution was diluted 1:10 with 100 % industrialised methylated sprits (IMS), allowed to stand for 5 min, and then centrifuged for 5 min at 3000 x g to remove precipitated perchlorate. Seeds were incubated in the solution at room temperature for 30 min in 1 ml with periodic agitation. The solution was decanted and seeds were washed twice in 100 % IMS in a laminar flow hood. The remainder of the IMS was removed using a pipette and the seeds allowed to dry before sowing onto agar plates.

Surface sterilised seeds were germinated onto 0.8 % agar plates containing, ½-strength Murashige and Skoog (MS) salts (Murashige and Skoog, 1962) with and without 20 mM sucrose. The media was adjusted to pH 5.7 with 0.01 M KOH prior to the addition of agar and subsequently sterilised by autoclaving at 121°C for 20 min. Plates were poured in a laminar flow cabinet once the media had cooled to approximately 55°C. Selection plates were created by the addition of 50 µg ml^{-1}

kanamycin, 50 μg ml^{-1} sulfadiazine or 500 μM fluoroacetate prior to pouring for analysis of T-DNA insertion lines.

Surface sterilised seeds were scattered on ½-strength MS plates and sealed with Micropore™ surgical tape to prevent contamination. Seeds were imbibed in the dark at 4°C for 4 days prior to germination at 20°C under constant illumination of 70 μmol of photons m^{2-1} s^{-1}. After 10-14 days, seedlings were transferred to a soil mixture compromised of 4 parts of multipurpose compost (John Innes 1, Bord Na Mona Horticulture Ltd, Kildare, Ireland, U.K.) and 1 part J. Arthur Bower's Vermiculite (William Sinclair Horticulture Ltd, Lincolnshire, U.K.). The soil was pre-sterilised by autoclaving at 121°C for at least 30 min. Prior to planting seedlings, the soil was soaked overnight in distilled water and subsequently treated with 0.2 g l^{-1} of the insecticide Intercept (Ropax, Lancaster, U.K.).

For metabolic analysis seeds were imbibed for four days in the dark prior to transfer to constant illumination of 70 μmol of photons m^{2-1} s^{-1}. Due to the varied time period required for all three mutants to reach the same growth stage the length of time under constant illumination was varied. Triplicate measurements of fifty seeds were weighted every twenty four hours for 4 days to obtain an average weight and estimate the time required for all three mutants to reach the same growth stage. Seedlings were illuminated for 56, 52, 52 and 59 hrs before harvesting for metabolic analysis for Col-0, *aco1*, *aco2* and *aco3* respectively.

2.4.3 Genomic DNA isolation

DNA was isolated from leaf or seeding material with the Gentra puregene tissue kit (QIAGEN Ltd., West Sussex, U.K.) according to the manufacturer's instructions. Approximately 100 mg fresh weight of tissue was ground in 300 μl lysis solution in a 1.5 ml microfuge tube. The ground tissue was incubated at 65°C in a water bath for 60 min. After 60 min the sample was cooled to room temperature followed by the addition of 1.5 μl RNase A solution. The solution was mixed well by inverting the tube 25 times and incubated at 37°C for 30 min. Once the solution had cooled to room temperature, 100 μl protein precipitation solution was added, the tube vortexed for 20 seconds, and incubated on ice for 1 hr. The sample was centrifuged at 13000 x

g for 9 min and the supernatant transferred to a clean microfuge tube containing 400 µl 25:24:1 phenol : chloroform : IAA. The sample was vortexed thoroughly and centrifuged for 3 min at 13000 x g. The supernatant was transferred to a clean microfuge tube containing 300 µl of 100 % isopropanol, mixed gently and kept at -20°C overnight. Precipitated DNA was centrifuged at 13000 x g for 15 min. The supernatant was discarded and the pellet washed twice with 300 µl of 70 % ethanol. The ethanol washes were discarded and the pellet was dried in air. DNA was re-suspended in 20 µl of TE (10 mM Tris / HCl pH 8, 1 mM EDTA) and incubated overnight at 4°C to rehydrate the DNA.

The quality and quantity of the isolated DNA was determined by UV spectrometry. The absorbance at 260 nm and 280 nm was read with 1 µl of sample diluted with 69 µl TE (10 mM Tris / HCl pH 8.0, 1 mM EDTA).

2.4.4 PCR identification of *aco1*, *aco2* and *aco3* homozygous mutants

Homozygous T-DNA insertion lines were created by self pollination and confirmed by PCR following the suggestions of SALK Institute (http://signal.salk.edu/tdnaprimers.html/) or GABI-KAT (http://www.gabi-kat.de/) using the primers listed in Table 2.3.

Table 2.3: Primer sets implemented to determine the presence of a T-DNA insert within the genome and determine if the mutant was heterozygous or homozygous.

	Combination	Primer sequence (5'-3')
SALK		
Lba1	-	TGGTTCACGTAGTGGGCCATCG
090200-F	LBa1	TCCCTCGTCCATTCCACGATCT
090200-R	090200-F	TGCTCAGAAGCTGCGTTTTTCACA
090200-R2	090200-F	CCCATGGTTGCGTATTTCCGTTC
037286-F	LBa1	TCGGTTTCCAGACAGCACAGCAG
037286-R	037286-F	TGCTCCAGGGTCTAGAGTTGTCGAG
030717-F	LBa1	CACCTGTGCACGTCCCCACAA
030717-R	030717-F	GGCCCCAACACTGTTCACATTCC
122247-F	LBa1	GCGTCCTCCCCAGCCTTGAA
122247-R	122247-F	GGACCGTGGTGTGATTTCGGAAG
090112-F	LBa1	TTGATTTGCCGCCATGTATCGAC
090112-R	090112-F	CCCATGGTTGCGTATTTCCGTTC
090112-R2	090112-F	TGCTCAGAAGCTGCGTTTTTCACA
079922-F	LBa1	TGCTCAGAAGCTGCGTTTTTCACA
079922-R	079922-F	CTGCGTCGTCGCTTCGCAAT
079922-R2	079922-F	GCCATGTATCGACGCGCCACT
054196-F	LBa1	ACACCCCACCCAGCGACTCCTA
054196-R	054196-F	CACTGAAAACCCCTCCAGGCCTCT
054196-R	054196-F	CAAGGTCAGAAGATGCCGCACAG
122245-F	LBa1	AACGCAGGGTGCAGACAAAGGA
122245-R	122245-F	CGATCATCCTGGCTGGTGCTG
122245-R	122245-F	AGTCGAAGGGGAAACGACGAGGT
090196-F	LBa1	CCCTCGTCCATTCCACGATCTTC
090196-R	090196-F	CCCATGGTTGCGTATTTCCGTTC
090196-R	090196-F	TGCTCAGAAGCTGCGTTTTTCACA
031352-F	LBa1	GAGGCCTGGAGGGGTTTTCAGTG
031352-R	031352-F	CGCAACCATGGGTATGTCTGTTTG
009537-F	LBa1	TTGGATCGAAGGTCTGTTGGAGTCA
009537-R	009537-F	CGGAGGAAGCACAATGGACCAA
008620-F	LBa1	TCCTGGGATCGTTCAAAGCAGGT
008620-R	008620-F	TGCCATGTGAAATTGAGAAACCTGA
008620-R2	008620-F	TCCTGGGATCGTTCAAAGCAGGT
014661-F	LBa1	CCAGAGCCAGGAGGCACAACAA
014661-R	014661-F	GGACTTTACGGGAGTCCCAGCTGTT
013368-F	LBa1	GCGTCATGCTTGGTGCTGGTCT
013368-R	013368-F	GAACCCATACCGCTTTGGTCCAT
013368-R2	013368-F	CACGATCATGTGGCCTGCAAAG
GABI-KAT		
Lb	-	ATATTGACCATCATACTCATTGC
110E12-F	Lb	ACGTATCTTTGACCGTAGCA
110E12-R	F	AGCCACGAATTGATAAGGAT
138A08-F	Lb	CCAAATGTCCCTGAAAAATA
138A08-R	F	CCAGTTGCACAAATACTTCA
298E10-F	Lb	GAGTATGGTATTGCCTGGTG
298E10-R	F	GTCAGCTTTCATTTCCTTCA
423C03-F	Lb	GATTGATCACTCGGTTCAAG
423C03-R	F	ACCAGGCAATACCATACTCA

2.4.5 RNA extraction and expression in *aco1*, *aco2* and *aco3*

Tissue was grown under constant light conditions as previously stated, harvested and frozen in 4 mg aliquots for no longer then 1 week before RNA extraction.

RNA was extracted from 4 mg of frozen tissue using the Dyna bead Kit (Invitrogen, Paisley, U.K.) according to the manufacturer's instructions. Tissue was kept frozen by dipping in liquid nitrogen and ground in a 1.5 ml microfuge tube with a pestle in 100 µl lysis / binding buffer (100 mM Tris / HCl, pH 7.5, 500 mM LiCl, 10 mM EDTA, pH 8.0, 1 % LiDS, 5 mM dithiothreitol). The lysate was cleared by centrifugation for 5 min at 14000 x g to remove cell debris. The supernatant was added to 20 µl of pre-washed Dynabeads and incubated for 5 min at room temperature with gentle agitation. The beads were harvested using the DYNAL magnetic rack and the supernatant discarded. The beads were washed twice with 100 µl buffer A (10 mM Tris / HCl, pH7.5, 0.15 M LiCl, 1 mM EDTA, 0.1 % LiDS), re-suspend in 100 µl buffer B (10 mM Tris / HCl, pH 7.5, 0.15 M LiCl, 1 mM EDTA) and transferred to a fresh microfuge tube. The beads were washed twice with buffer B and re-suspended in 50 µl 10 mM Tris / HCl (pH 7.5) at 4°C. The supernatant was removed immediately before DNase treatment.

RNA potentially contaminated with DNA was treated with DNase (Promega Corporation, Southampton, U.K.) by re-suspending the beads in a total volume of 10 µl consisting of 1 x reaction buffer plus 1 µl of DNase I solution. The mixture was incubated at room temperature for 15 min followed by addition of 1 µl stop solution and 10 min incubation at 70°C.

DNase treated RNA was added to a final volume of 20 µl cDNA synthesis reaction mixture containing 1 µl Omniscript^{TM} reverse transcriptase stock (QIAGEN Ltd., West Sussex, U.K.), 1 x of the provided RT buffer, 0.5 mM each dNTP, 1 µM Oligo dT_{15} primer and 10 units rRNAsin (Promega Corporation, Southampton, U.K.). The mixture was incubated at 37°C for 1 hr and an aliquot added to the PCR mix.

One µl cDNA was added to 10 µl PCR reactions containing, 1x buffer, 0.8 units Taq DNA polymerase (Promega Corporation, Southampton, U.K.), 3 mM $MgCl_2$, 0.2

mM dNTPs, 0.5 µM forward primer and 0.5 µM reverse primer (Table 2.4). PCR cycling conditions consisted of a denaturing step of 94°C for 60 seconds; 35 cycles of a denaturing step of 94°C for 60 seconds; an annealing temperature step of 58°C for 45 seconds and an elongation step of 72°C for 120 seconds. A final elongation step of 72°C for 5 min and a holding temperature of 4°C were added at the end. PCR products were visualised on 2 % agarose gels.

Table 2.4: Primers used in RT-PCR reaction.

	Forward primer sequence (5'-3')	Reverse primer sequence (5'-3')
ACO1	CATGTCTGGAATTGAATCTCG	CAGAGCCTGGAGCAAGACTA
ACO2	ATCTTGCGTCACCACCATTA	TTTCCTGCTGGAGAGATGTG
ACO3	GCTGACTGGCATTCATGTCT	AATTCCCGATGCATGTAGTG
Actin2	CTAAGCTCTCAAGATCAAAGGCTTA	ACTAAAACGCAAAACGAAAGCGGTT
RBCS	CTATGGTCGCTCCTTTCAACGG	TGCAACCGAACAAGGGAAGC

2.4.6 Specific activity of aconitase in wild type, *aco1*, *aco2* and *aco3*

Seedlings were weighted at 24 hr time periods after imbibition for 8 days and ground in 200 µl extraction buffer (100 mM Tris pH 7.0, 1 mM EDTA, 10 mM KCl and 10 % glycerol). The extract was centrifuged for 15 min at 4°C to remove cell debris and 170 µl was assayed immediately for aconitase activity as described in section 2.3.13.

2.4.7 Metabolic analysis of wild type, *aco1*, *aco2* and *aco3*

Radiolabel feeding studies were conducted as described in Turner *et al.* (2005). One hundred seedlings were harvested from ½-MS plates (Murashige and Skoog, 1962) and placed in 0.2 ml of 1 mM sodium [2-^{14}C] acetate (American Radiolabeled Chemicals, Missouri, U.S.A.) and 50 mM MES, pH 5.2. One µl of the incubation solution (1 mM sodium [2 - ^{14}C] acetate, 50 mM Mes, pH 5.2) was removed and added to 10 ml of Perkin Elmer^{TM} Optiphase HiSafe3 liquid scintillation cocktail (Fisher Scientific U.K., Leicestershire, U.K.) to determine the starting specific radioactivity. Seedlings were bubbled in the dark for 4 hours and evolved CO_2 was collected with two tandem traps of 200 µl of 5 M KOH. After four hours the bubbling solution was removed and the seedlings washed twice with 0.3 ml of distilled H_2O. The seedlings were re-bubbled in 0.2 ml of 1 mM unlabelled sodium

acetate and 50 mM MES, pH 5.2. After 10 min the bubbling solution was removed and the reaction stopped by addition of 1 ml boiling 80 % ethanol. The CO_2 traps were combined and added to 10 ml HiSafe3.

The ethanol was removed from the seedlings, 200 µl 80 % ethanol was added and incubated at 100°C. After 10 min the seedlings were ground using a micropestle and a further 1 ml 80 % ethanol was added. After incubation for 10 min at 100°C, the ethanol was removed and 1 ml boiling water was added. After 10 min at 100°C, the water was removed and 1 ml of 50 % ethanol was added. The sample was incubated at 100°C for another 10 min. The resulting extracts were combined and vacuum filtered through a 25 mm GF/A glass microfiber filter (Whatman®, Brentford, Middlesex, U.K.). The filtrate was dried with a Büchi Rotavapor-R to remove the ethanol and the residue re-suspended in 0.6 ml of sterile H_2O. The re-suspended extract was washed three times with 1 volume of water-saturated ethyl ether, allowed to evaporate in air, re-suspended in 150 µl 80 % ethanol, and added to 10 ml HiSafe3 liquid scintillation cocktail for determining the amount of radioactivity present. The ethanol insoluble residue retained by the filter was burnt in a Biological Material Oxidiser – OX400 (R.J. Harvey Instrument Corporation, New Jersey, U.S.A.) and $^{14}CO_2$ trapped in 15 ml of Oxosol^{TM} (National Diagnostics, Hessle, East Riding of Yorkshire, U.K.).

The remaining aqueous fraction was passed through two successive columns, firstly a cation exchange resin (AG50 W-X8, mesh 100-200, H+ form, Bio-Rad, Hertfordshire, U.K.) equilibrated with water followed by an anion exchange resin (AG1-X8, mesh 100-200, H+ form, Bio-Rad, Hertfordshire, U.K.) also equilibrated with water. Disposable columns were created for each resin by placing a small amount of deactivated borosilicate glass wool (Restek, Pennsylvania, U.S.A.) at the bottom of a 1 ml pipette tip and applying 250 µl of resin. The columns were washed well with distilled water prior to applying the sample. The neutral fraction containing the carbohydrates was eluted with 2 ml of distilled H_2O. The two columns were separated, the cation exchange column was washed with 2 ml of 1.2 M ammonium hydroxide to elute the amino acids (basic fraction), and the anion exchange column was washed with 2 ml of 4 M formic acid to elute the organic acids (acidic fraction). One ml of each eluted fraction was combined with 10 ml of HiSafe3. All samples

were dark adapted, the outside of the bottles cleaned with 100 % Methanol and then analysed on a Wallac Winspectral 1414 liquid Scintillation counter. Two background samples of both HiSafe3 and Oxosol were counted and a mean background count was determined for the calculation of net DPM values.

2.4.8 Quantification of citrate levels in wild type, *aco1*, *aco2* and *aco3*

Seedlings were surface sterilised, cold treated and germinated under constant illumination at 20°C for varying amount of time as described in section 2.4.2. Fresh tissue was harvested, frozen immediately in liquid nitrogen, freeze-dried and stored at -80°C prior to metabolite extraction.

Metabolites were extracted based on the integrated RNA, metabolite and protein extraction protocol (Weckworth *et al.,* 2004). Only the steps relevant to the extraction of metabolites are detailed here. The sample tissue was homogenised under liquid nitrogen with a pre-cooled pestle in an eppendorf. To the sample, 1 ml of a single phase solvent mixture of methanol / chloroform / H_2O (2.5:1:1 v/v/v), stored at -20°C, was added and the sample was mixed at 4°C for 30 min. The sample was centrifuged at 13000 x g at 4°C for 10 min. The supernatant was removed to a new eppendorf and chilled on ice. To the remaining pellet, 1 ml of a single phase solvent mixture of methanol / chloroform (1:1 v/v), stored at -20°C, was added and the sample vortexed. The sample was centrifuged at 13000 x g at 4°C for 10 min and the supernatant removed to a new eppendorf and chilled on ice. The two supernatants were combined and 500 µl H_2O was added to separate the hydrophobic phase from the hydrophilic phase. The sample was centrifuged at 13000 x g at 4°C for 5 min and the two phases separated using a pipette. The separated phases were dehydrated and stored under desiccation at -20°C prior to analysis.

Samples were re-suspended in water and added directly to a 1 ml assay containing citrate lyase, malate dehydrogenase and lactate dehydrogenase.

$$\text{citrate} \xrightarrow{\text{citrate lyase}} \text{acetate} + \text{oxaloacetate}$$

$$\text{oxaloacetate} \xrightarrow{\text{malate dehydrogenase}} CO_2 + \text{pyruvate}$$

$$\text{pyruvate} + \text{NADH} \xrightarrow{\text{lactate dehydrogenase}} \text{lactate} + \text{NAD}$$

The decrease in absorbance at 340 nm of NADH was monitored in a 1 ml assay containing 100 mM triethanolamine pH 7.6, 0.5 mM zinc chloride, 0.23 mM NADH, 100 units lactate dehydrogenase, 50 units malate dehydrogenase, 15 mM ammonium sulphate and 0.02 units citrate lyase (Bergmeyer *et al.,* 1974). The reaction was initiated by addition of extracted metabolites and the change in absorbance measured to determine the concentration of citrate, oxaloacetate and pyruvate present in the extract.

2.4.9 GC-MS analysis of Col-0, *aco1*, *aco2* and *aco3*

Seedlings were surface sterilised, cold treated and germinated under constant illumination at 20°C for varying periods of time as described in section 2.4.2. Fresh tissue was harvested, frozen immediately in liquid nitrogen, freeze dried and stored at -80°C prior to metabolite extraction.

Metabolites were extracted based on the extraction protocol of O'Fiehn *et al.* (2008). Prior to extraction samples were transferred to a 2 ml microcentrifuge tube, weighted and finely ground on a Retsch Mixer Mill MM 200 with a 5 mm ball bearing for 90 seconds. To the sample, 1000 µl of a single phase mixture of chloroform / methanol / water (1:2.5:1 v/v/v), stored at -20°C, and 100 µl of the internal standard ribitol (0.17 mg ml^{-1}) was added. The solution was vortexed for 10 sec and mixed at 4°C for 15 min. The sample was centrifuged at 13000 x g at 4°C for 5 min, the supernatant removed to a sterile 15 ml tube and chilled on ice. To the remaining pellet, 1000 µl of a single phase solvent mixture of chloroform / methanol / water (1:2.5:1 v/v/v), stored at -20°C, and 100 µl of the internal standard ribitol (0.17 mg ml^{-1}) was added. The solution was vortexed for 10 sec and mixed at 4°C for 15 min. The sample was centrifuged at 13000 x g at 4°C for 5 min and the supernatants combined. One ml of water was added, the sample was vortexed and centrifuged at 13000 x g at 4°C for 5 min. A 200 µl aliquot of the polar phase was transferred to a sterile tube, cooled in liquid nitrogen and stored at -80°C. The polar phase was divided into three 1000 µl aliquots, combined with 50 µl Myristic acid (0.25 mg ml^{-1} in methanol), cooled in liquid nitrogen and stored at -80°C.

Prior to data collection samples were dried by vacuum centrifugation and derivatised with methoxyamine hydrochloride and N-methyl-N-trimethylsilytrifluoroacetamide.

The resulting samples were analysed by gas chromatography mass spectroscopy at Manchester Interdisciplinary Centre by Dr J.W. Allwood. The obtained cdf files were visually compared using the freely available software, AMDIS (http://chemdata.nist.gov/mass-spc/amdis/). The raw data was processed by Dr J.W. Allwood and the abundance of the resulting compounds compared using Microsoft Office Excel.

GC-MS data was analysed using SPSS 17. The equality of variances in the samples was confirmed using a Levene's test. Data sets with equal variances were analysed with a one way analysis of variance (ANOVA) with post-hoc analysis conducted using a multiple comparison Tukey test. Data sets with unequal variances were analysed with a Kruskal-Wallis non-parametric test with a Dunnett C test used for post-hoc analysis. Statistical analysis was conducted between the entire seedling groups (Col-0, *aco1*, *aco2* and *aco3*) for each metabolite (citric acid, fumaric acid, isocitric acid, malic acid, myristic acid, pyruvic acid, succinic acid and sucrose).

Structural properties of a glyoxysomal acetyl-CoA synthetase

3.1 Introduction

The acyl-CoA synthetase, AAE7/ACN1, is a member of the AMP forming family responsible for a range of catalytic activities. Other proteins that belong to this family, recently re-named the ANL superfamily (Gulick, 2009), include firefly luciferase and non-ribosomal peptide synthetases. AAE7/ACN1 is also a member of the adenylate activating enzyme (AAE) superfamily of *Arabidopsis thaliana*. This large superfamily was identified during an amino acid based search for a second family of long-chain acyl-CoA synthetase (LACS) enzymes (Shockey *et al.*, 2002). The search identified sixty three family members, grouped phylogenetically into seven distinct categories (Shockey *et al.*, 2003). AAE7/ACN1 belongs to clade VI along with fourteen other plant specific genes.

Functional analysis conducted by Shockey *et al.* (2003) showed that AAE7/ACN1 is active against short chain substrates. When membrane fractions expressing AAE7/ACN1 were isolated, greater activity was observed against butyrate compared to acetate. This is in contrast to latter studies that show greater activity against acetate compared to butyrate in the absence of an artificial membrane (Turner *et al.*, 2005). Shockey *et al.* (2003) identified AAE7/ACN1 during the search for a second family of membrane proteins, LACS, and therefore believed that this newly identified family were also bound to the plasma membrane. AAE7/ACN1 was identified by Turner *et al.* (2005) during screening experiments with fluoroacetate, the toxic analogue of acetate, suggesting an involvement in acetate metabolism. Seedling screening experiments also identified an ABC transporter, COMATOSE, which is believed to transport acetate into the glyoxysome for catalysis (Hooks *et al.*, 2007).

Within *Arabidopsis thaliana* the monomeric plastid acetyl-CoA synthetase (atACS), characterised within clade II (Shockey *et al.*, 2003), is also responsible for the two step activation (Figure 3.1) of acetate to acetyl-CoA (Ke *et al.*, 2000). Although both enzymes are responsible for the same reaction, the respective processes are spatially separate. AAE7/ACN1 is located within the glyoxysome providing acetyl-CoA for the glyoxylate cycle (Turner *et al.*, 2005) and atACS is responsible for the biosynthesis of fatty acids within the plastid (Ke *et al.*, 2000).

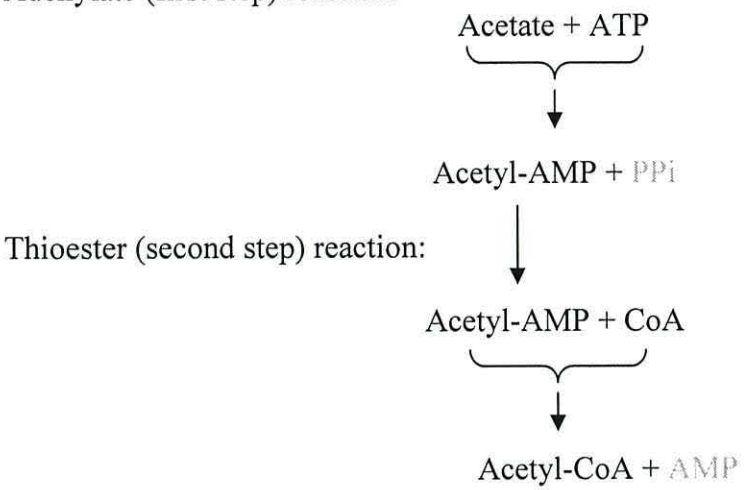


Figure 3.1: Two step reaction catalysed by acetyl-CoA synthetase. Acetate and ATP bind to the enzyme with the production of acetyl-AMP and PPi in the first reaction step with the release of PPi. CoA enters during the second reaction step with the production of acetyl-CoA and AMP.

This reaction, conducted by S*almonella enterica's* acetyl-CoA synthetase (seACS), is responsible for acetyl-CoA production at low acetate concentrations (Reger *et al.,* 2007); although at higher concentrations acetyl-CoA synthesis is carried out by acetate kinase and phosphotransacetylase (Reger *et al.,* 2007). The crystallographic structure of monomeric seACS shows a two domain structure with the active site at the domain interface (Gulick *et al.,* 2003). Other structuralised members of the AMP binding superfamily, 4-coumarate-CoA ligase, acyl-CoA ligase and firefly luciferase, also show a two domain structure attached by a flexible hinge (Conti *et al.,* 1996; Stuible *et al.,* 2000; Reger *et al.,* 2007).

The linker sequence, (also known as the A8 motif) that joins the large N-terminal domain and the smaller C-terminal domain, contains a conserved aspartate residue that allows ACS to adopt two different conformations for each of the reaction steps. The consensus linker sequence is conserved in a number of AMP binding proteins that include acyl-CoA synthetase (Gulick *et al.,* 2003; Hisanaga *et al.,* 2004; Jogl and Tong, 2004; Brasen *et al.,* 2005), CBAL (Gulick *et al.,* 2004; Reger *et al.,* 2008) and Phe (Conti *et al.,* 1997).

Crystallised structures of ACS have been characterised as a monomer (Gulick *et al.,* 2003) and a trimer (Jogl and Tong, 2004), though ACS has also been identified as a

dimer (Preston *et al.*, 1990; Martinez-Blanco *et al.*, 1992) and an octomer (Brasen *et al.*, 2005). Oligomerisation can result in both structural and functional advantages such as providing protection from degradation (Goodsell and Olson, 2000), for example by reducing the surface area exposed to solvent compared to a monomeric structure (Jones and Thornton, 1995; Goodsell and Olson, 2000). Other advantages include reducing translational errors by providing an extra step for proofreading that allows defective subunits to be discarded (Goodsell and Olson, 2000), adding an extra level of control through allosteric regulation (Marianayagam *et al.*, 2004) and increasing the concentration of active sites in a particular region (Marianayagam *et al.*, 2004). The multimeric state of AAE7/ACN1 is currently unknown and will be investigated within this chapter.

Currently, no structural information is known about the AAE superfamily of *Arabidopsis*. Whilst metabolic analysis of a limited number of family members has helped to determine the function of the superfamily members (Shockey *et al.*, 2003; Turner *et al.*, 2005; Kim and Guerinot, 2007; Lin and Oliver, 2008; Wiszniewski *et al.*, 2009), analysis of their structural properties will characterise their mechanism of action.

Structural analysis of AAE7/ACN1 is important in order to determine if the enzyme's mode of action is similar to that of atACS, even though both enzymes are involved in different, spatially separated pathways. Work presented within this chapter aims to determine the multimeric state of AAE7/ACN1 and determine if it undergoes a conformational change during catalysis.

3.2 Crystallisation trials of AAE7/ACN1

Determining the crystallographic structure of AAE7/ACN1 would be beneficial in order to understand its method of activating both acetate and butyrate (Shockey *et al.*, 2003; Turner *et al.*, 2005). Characterisation of the enzyme's active site and residues involved in catalysis will help to identify the true substrate of AAE7/ACN1.

Crystallisation trials were set up with the Structure Screen 1 kit from Molecular Dimensions Limited (Cambridgeshire, UK) with purified AAE7/ACN1 (Turner *et al.*, 2005) at 4°C and 18°C. Putative crystals (Figure 3.2C and D) were identified

after 3 days with the Structure Screen 1 kit reservoir solution 33. To determine if the observed crystals were comprised of purified protein or salt their strength and ability to take up the dye bromophenol blue was tested (Bergfors, 1999).

The methods used to distinguish between protein and salt crystals were tested with crystallised Lysozyme (Figure 3.2A and B). The protein crystals were easily re-dissolved with a needle and turned blue after incubation with 1 µl of 2.5 % bromophenol blue (Figure 3.2B). Unfortunately the putative crystals, observed with AAE7/ACN1, could not be broken with a needle and did not turn blue after incubation with bromophenol blue (Figure 3.2D) suggesting that they were comprised of salt rather than protein.

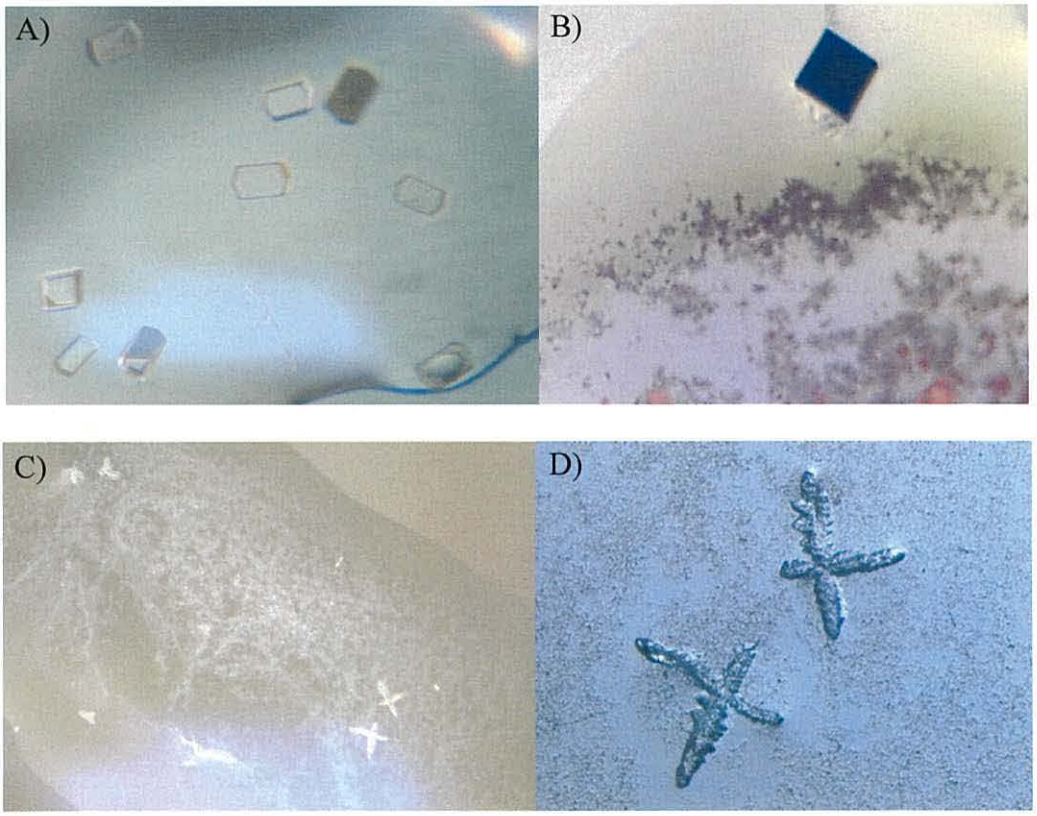


Figure 3.2: A) Lysozyme crystals B) Lysozyme crystals that have taken up the bromophenol blue dye. C) Salt crystals surrounded by protein precipitate identified with 0.2 M Magnesium chloride hexahydrate, 0.1 M Tris HCl, pH 8.5 and 30 % w/v PEG 4000 (number 33) of Structure Screen 1 (Molecular Dimensions Ltd, Cambridgeshire, UK). D) Salt crystals obtained under the same conditions as C in the presence of bromophenol blue.

In a second experiment, putative crystals were also observed using the modified reservoir solution of Gulick *et al.* (2003) (section 2.2.5). Despite repeating these

conditions with a range of ammonium sulphate concentrations this result was never re-produced, indicating that it was not a protein crystal but was in fact contamination. Common impurities within the droplet include dust particles and fibres (Bergfors, 1999; Thakur *et al.*, 2007).

3.3 Small angle x-ray scattering

Structural analysis of AAE7/ACN1 was implemented in the form of solution scattering to establish a model of the protein's overall conformation.

3.3.1 Sample preparation and data collection

AAE7/ACN1 was over-expressed, purified and concentrated to 10 mg ml^{-1} as detailed in section 2.2.1 (Figure 3.3). The specific activity was determined prior to small angle x-ray scattering (SAXS) data collection as 2390 nmol min^{-1} mg^{-1} protein. Previous specific activity measurements of AAE7/ACN1 have ranged from 1700 – 3160 nmol min^{-1} mg^{-1} protein (Turner, 2003).

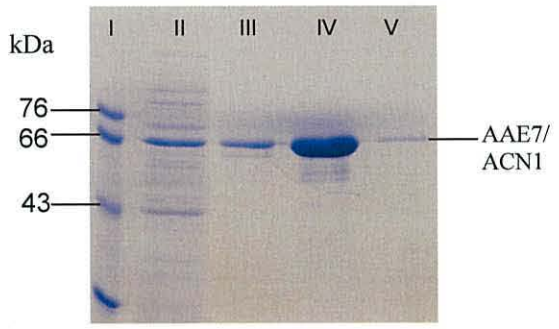


Figure 3.3: 12 % SDS-PAGE analysis of the purification of AAE7/ACN1 in 0.02 M KH_2PO_4, 0.05 M KCl and 10 % glycerol, pH 7.4 from *E. coli* using a HiTrap Ni chelating column. Lane I; Molecular weight marker, 76, 66, 43 kDa. II; Cleared cell lysate loaded onto the HiTrap column prior to purification. III, IV, V; 5 ml elution with 250 mM imidazole buffer collected in 1 ml fractions, the first three fractions of purified AAE7/ACN1 are shown.

Solution scattering data was collected at Beamline 2.1, SRS Daresbury Laboratory as outlined in section 2.2.7 with 1 mg ml^{-1}, 2 mg ml^{-1} and 5 mg ml^{-1} purified AAE7/ACN1 at 5.5 m camera length and 10 mg ml^{-1} AAE7/ACN1 at 1 m camera length. Data was normalised and processed to produce a merged 1-dimensional scattering profile (Figure 3.4).

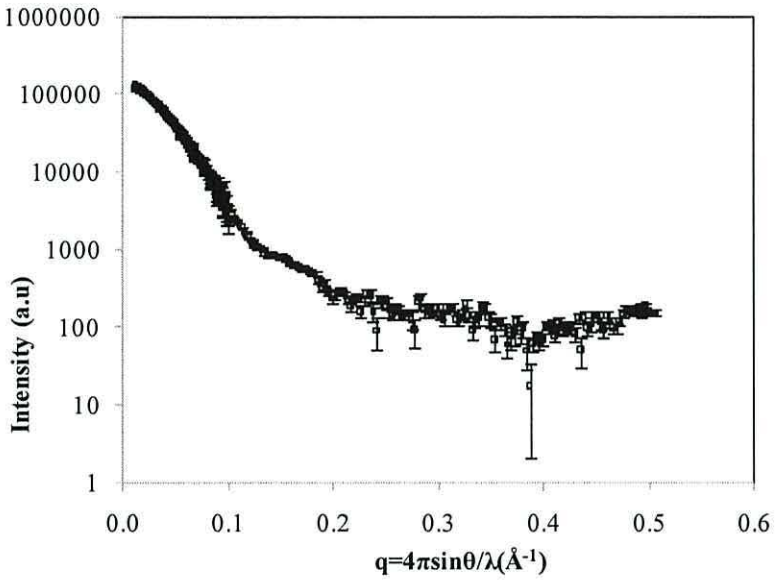


Figure 3.4: Solution scattering data collected at Beamline 2.1, SRS Daresbury. Data collected with 1 m SAXS camera (1 mg ml^{-1} AAE7/ACN1) and 5.5 m camera (10 mg ml^{-1} AAE7/ACN1 in 0.02 M KH_2PO_4, 0.05 M KCl and 10 % glycerol, pH 7.4) was merged to create the solution scattering profile over the entire available scattering range. Note that the x-axis represents the magnitude of the scattering vector q (in inverse Å) and the y-axis shows the scattering intensity (using logarithmic scale).

Scattering data was collected at increasing protein concentrations to check that AAE7/ACN1 remained monodisperse in solution. The Guinier region (Figure 3.5) showed a linear relationship between the natural log of the scattering intensity and the square of the scattering vector q between 0.019 Å^{-1} < q < 0.044 Å^{-1} indicating the protein solution to be monodisperse (Figure 3.5). The radii of gyration (R_g) were calculated from the Guinier region with 1 mg ml^{-1}, 2 mg ml^{-1} and 5 mg ml^{-1} AAE7/ACN1 and found to be 38.9 Å, 40.8 Å and 42.7 Å, respectively.

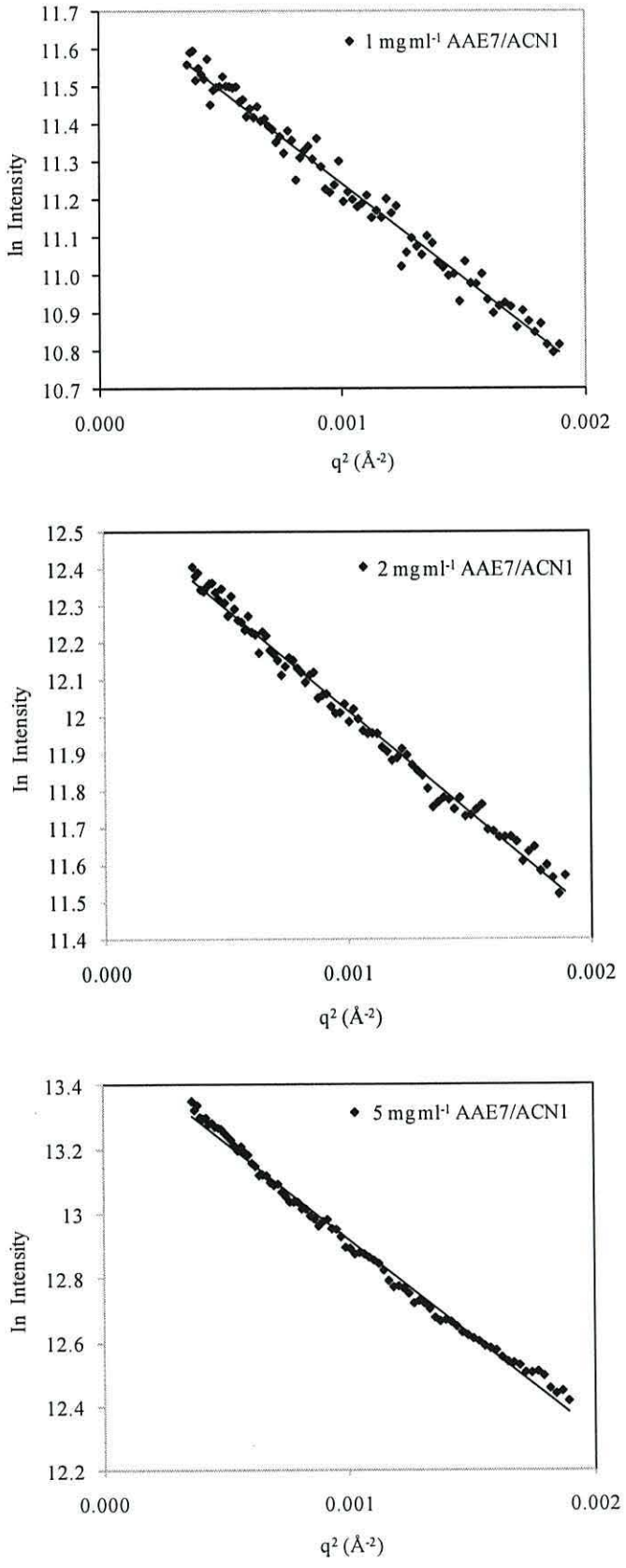


Figure 3.5: Guinier region of small angle x-ray scattering data collected at Beamline 2.1, SRS Daresbury of 1 mg ml^{-1}, 2 mg ml^{-1} and 5 mg ml^{-1} purified AAE7/ACN1 in 0.02 M KH_2PO_4, 0.05 M KCl and 10 % glycerol, pH 7.4.

As discussed in section 2.2.8 the indirect transform program, GNOM (Svergun, 1992), was utilised to determine the radius of gyration and the maximum particle dimension (*Dmax*) of AAE7/ACN1 in solution. GNOM reads the merged one dimensional scattering profile (Figure 3.4) and determines the particle distance distribution function p(r) and subsequently D*max* (Figure 3.6). GNOM estimates the R_g of AAE7/ACN1 as 41 Å and the D*max*, obtained when p(r) drops to zero as 140 Å. The R_g estimated by GNOM using the whole scattering profile was similar to the average R_g of 40.8 ± 1.9 Å predicted from the Guinier region of increasing concentrations of AAE7/ACN1. The distance distribution function suggests that AAE7/ACN1 is an elongated protein as the maximum p(r) is smaller than D*max* / 2 with a skewed distribution (section 2.2.8).

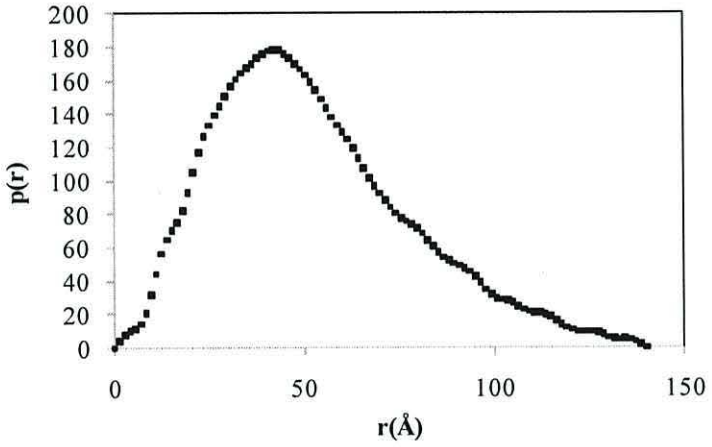


Figure 3.6: Pair distribution function, p(r) of AAE7/ACN1 with a D*max* of 140 Å was determined with the program GNOM (Svergun, 1992).

3.4 Determining the multimeric state of AAE7/ACN1

Prior to the creation of a topology model from the solution scattering profile the multimeric state of AAE7/ACN1 needs to be established for insertion of a symmetry restraint.

3.4.1 CRYSOL comparison of the AAE7/ACN1 solution scattering profile to known crystallographic structures

CRYSOL (Svergun *et al.,* 1995; as described section 2.2.10.1) was used to generate theoretical solution scattering profiles from the crystal structures of a variety of acyl-CoA synthetases to predict the multimeric state of AAE7/ACN1 (Figure 3.7).

Namely, these were monomeric acetyl-CoA synthetase (1PG4.pdb) from *Salmonella enterica* (seACS) (Gulick *et al.,* 2003), trimeric *Saccharomyces cerevisiae* acetyl-CoA synthetase (scACS) (1RY2.pdb) (Jogl and Tong, 2004) and dimeric long chain fatty acyl-CoA synthetase from *Thermus thermophilus* (ttLC-FACS) (1ULT.pdb) (Hisanaga *et al.,* 2004). Profile comparisons show that AAE7/ACN1 is not a monomeric protein in solution as the SAXS profile differs greatly in both the high and low angle regions to the generated theoretical profile of the monomeric seACS (Figure 3.7A). These observed differences are reflected in the large χ value of 14.84 obtained by CRYSOL that evaluates how well the structure fits the SAXS data. The predicted profile of the trimeric scACS is similar to AAE7/ACN1 in the low angle region but differs in the high angle region suggesting differences in overall conformation, indicating that AAE7/ACN1 is not a trimeric protein in solution (Figure 3.7C). These differences are reflected in the large χ value of 14.66. A high degree of similarity between AAE7/ACN1 and ttLC-FACS in both the high and low angle regions strongly suggests that AAE7/ACN1 is a dimeric protein in solution (Figure 3.7B). This is reflected in the lower χ value of 3.03, and indicates that the crystal structure of ttLC-FACS is similar to the structure of AAE7/ACN1.

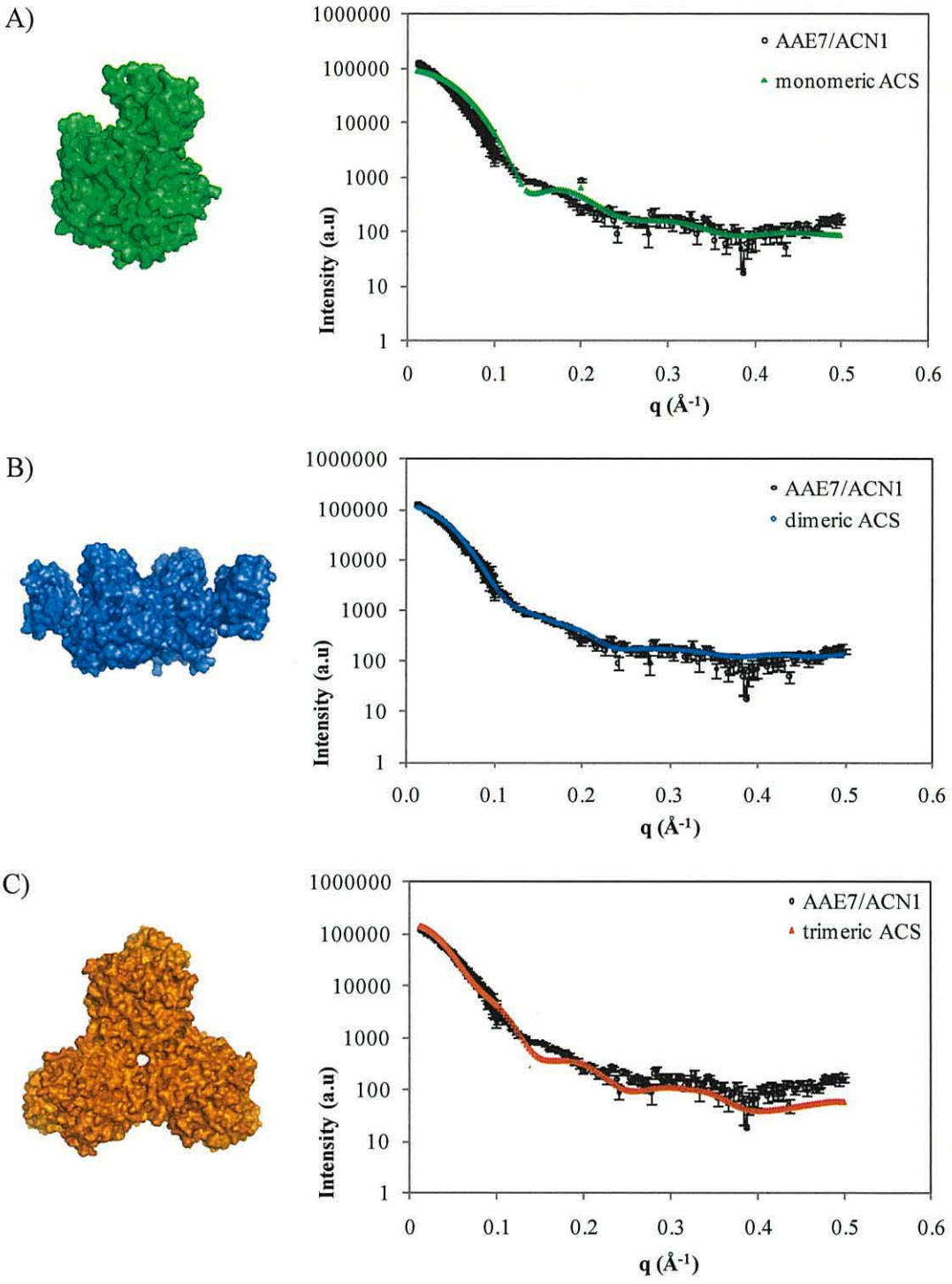


Figure 3.7: Comparison of AAE7/ACN1 small angle x-ray scattering data (shown in black), collected at Daresbury laboratory Beamline 2.1, with predicted small angle x-ray scattering data from (A) a monomeric acetyl-CoA synthetase, protein data bank 1PG4 (χ=14.84) (B) a dimeric long chain acyl-CoA synthetase, protein data bank 1ULT (χ=3.03) and (C) a trimeric acetyl-CoA synthetase, protein data bank 1YR2 (χ=14.66). Small angle x-ray scattering data was generated from crystallographic data using the default settings of CRYSOL (Svergun *et al.*, 1995). Low resolution structural images were created from crystallographic structures with Pymol (DeLano, 2002).

3.4.2 Native polyacrylamide gel electrophoresis of AAE7/ACN1

In order to experimentally confirm the multimeric state of AAE7/ACN1 predicted by CRYSOL, purified protein was subjected to native PAGE analysis. Despite altering the pH of the running buffer, sample buffer and the gel, together with the analysis being repeated on several occasions, the results were inconclusive. Initially the gels were run at pH 6.8 and 8.8 (PhastSystem, 1986a; PhastSystem, 1986b), however when protein remained in the gel wells, pH 10 was investigated, though without success.

3.4.3 Gel filtration of AAE7/ACN1

Purified AAE7/ACN1 was also run through a Superose 12 HR 12/30 column on a FPLC to elucidate its multimeric structure. Gel filtration analysis suggested that AAE7/ACN1 was monomeric (Figure 3.8 and 3.9) though further investigations showed that only around 30 % of the applied protein ran through the column with the remainder staying on its surface. During a typical experiment 0.26 mg of AAE7/ACN1 was loaded onto the column in a volume of 100 µl but only 0.08 mg eluted from the column, after 13 ml, in a similar position as BSA. The eluted protein was assayed (section 2.2.3) but no activity was recorded. This implies that AAE7/ACN1 is not monomeric and the protein is dissociating either before or during transport through the column.

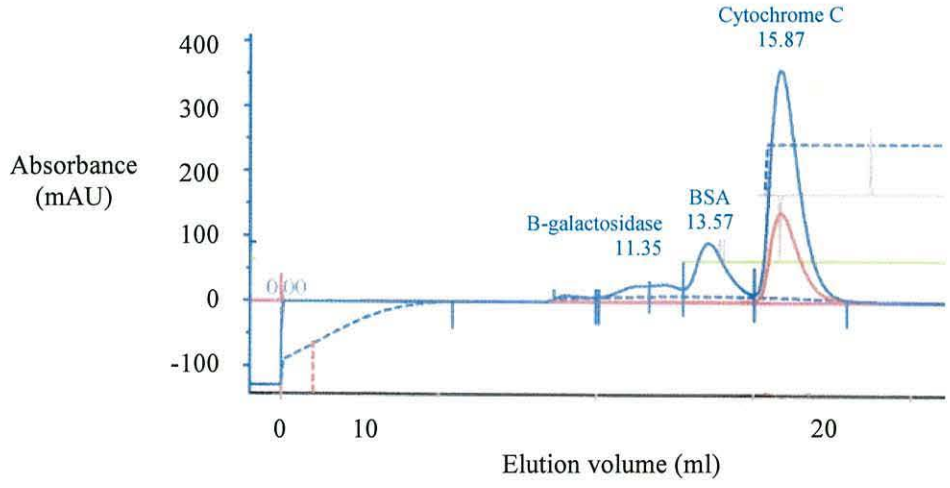


Figure 3.8: Gel filtration standards eluted from a Superose 12 HR 12/30 column. Beta galactosidase eluted at 11.35 ml, BSA at 13.57 and cytochrome C at 15.87 ml. The x-axis shows the volume eluted from the column in ml and the y-axis shows that absorbance at 280 nm (solid blue line) and 550nm (red line).

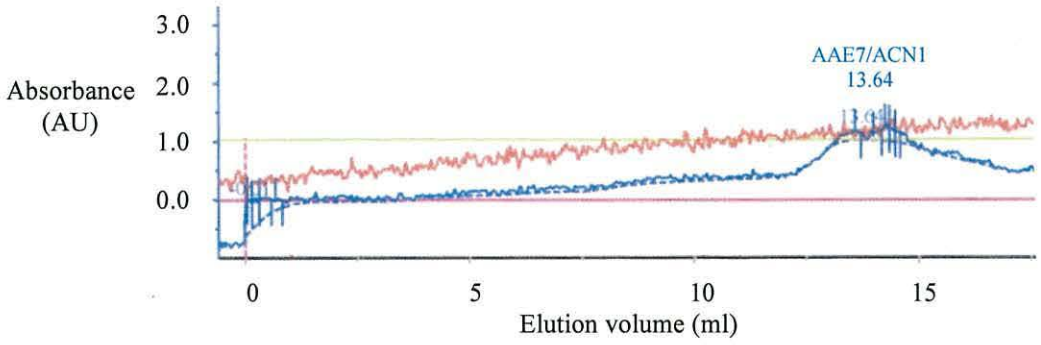


Figure 3.9: Gel filtration analysis with a Superose 12 HR 12/30 column of AAE7/ACN1. AAE7/ACN1 eluted at 13.64 ml. The x-axis shows the volume eluted from the column in ml and the y-axis shows that absorbance at 280 nm (solid blue line) and 550nm (red line).

3.4.4 Electrospray ionisation mass spectrometry of AAE7/ACN1

Mass spectroscopy analysis was conducted in order to determine the multimeric state of AAE7/ACN1.

Monomeric and dimeric forms of AAE7/ACN1 were detected in solution by electrospray ionisation (ESI) mass spectrometry (Figure 3.10) as described in section 2.2.10.5. Mass spectrometers measure mass-to-charge ratio (m/z) and whilst a monomer with n charges and a dimer with 2n charges will have the same m/z ratio and be indistinguishable in a spectrum, the presence of a dimer can be identified by peaks representing odd numbers of charges.

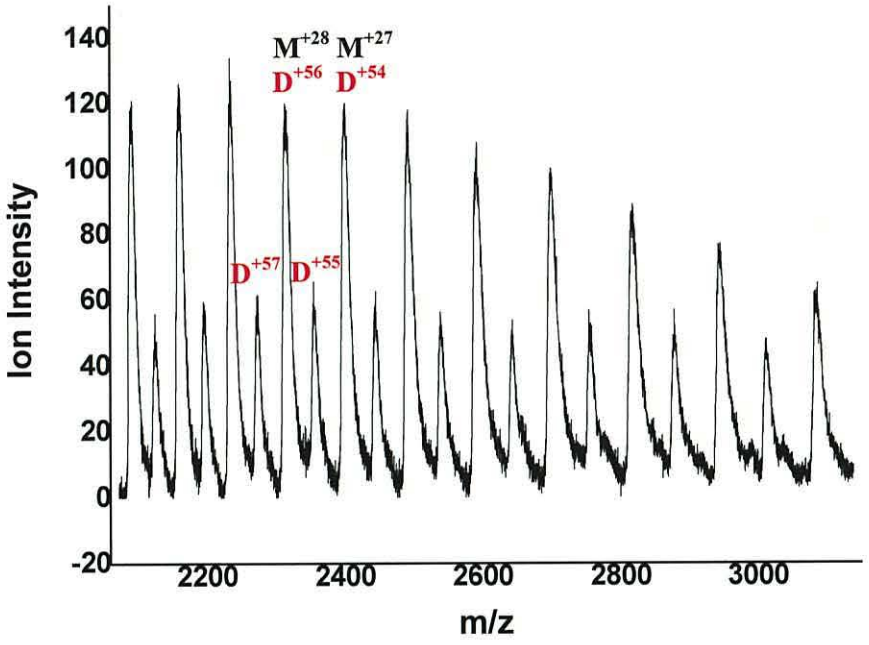


Figure 3.10: ESI mass spectrum of AAE7/ACN1 dissolved in 10 mM ammonium formate. The main figure shows the raw mass-to-charge (m/z) data. Peaks labelled with D and M correspond to m/z of both monomeric and dimeric form while the peaks labelled D only correspond to m/z of dimeric species.

The deconvoluted mass spectrum (Figure 3.11) shows that the dimeric structure is the dominant species in solution and the ratio of dimer to monomer is estimated to be 2:1. The presence of monomer in the spectrum is the result of the dissociation of dimer into monomers caused by the ionisation process (electric field, desolvation, heat).

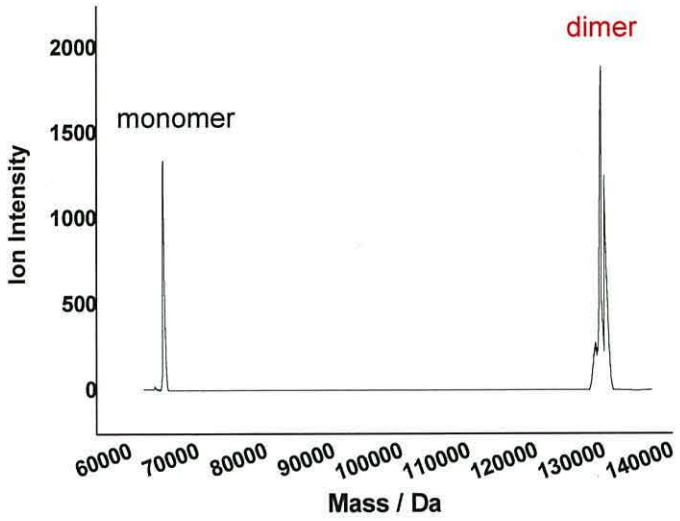


Figure 3.11: The deconvoluted mass spectrum (i.e. data presented in mass scale only) showing that the dominant form is dimeric AAE7/ACN1.

3.5 Shape reconstruction of AAE7/ACN1

A dummy residue model of AAE7/ACN1 was constructed with GASBOR (Svergun *et al.*, 2001) from the solution scattering profile and the size dimensions produced by GNOM. Sixteen independent shape reconstructions were performed with GASBOR using a two-fold symmetry constraint to ensure a homodimeric structure. The χ values produced by GASBOR demonstrate how close the model fits to the experimental scattering data of AAE7/ACN1. The χ values range from 1.86 to 1.97 (Table 3.1) suggesting that the predicted models are all very similar. The fit of the sixteen models (Models A to P) produced by GASBOR were compared to the experimental solution scattering data of AAE7/ACN1 (Figure 3.12).

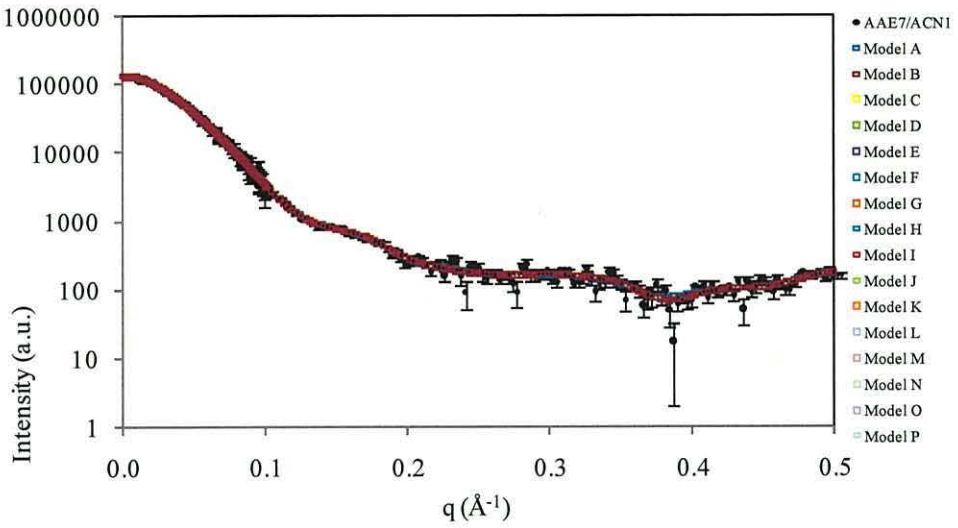


Figure 3.12: Comparison of the sixteen models created by GASBOR to the solution scattering profile of AAE7/ACN1.

In-turn the similar χ values from GASBOR models implies a close relationship between the different models (Table 3.1).

Table 3.1: GASBOR χ value for the predicted models of AAE7/ACN1

GASBOR Model	χ value from GASBOR showing the experimental fit
Model A	1.89
Model B	1.86
Model C	1.88
Model D	1.90
Model E	1.95
Model F	1.89
Model G	1.87
Model H	1.92
Model I	1.88
Model J	1.88
Model K	1.88
Model L	1.97
Model M	1.89
Model N	1.90
Model O	1.93
Model P	1.94

The scattering profile of the model with the lowest χ value (Model B) of 1.86 and therefore the 'best fit' model was compared to the experimental scattering data of AAE7/ACN1 (Figure 3.13A). Comparison of the profiles shows that they are very similar in both the low and high angle regions (Figure 3.13A) and a side view of the

dummy residue model suggests that AAE7/ACN1 is comprised of two domains (Figure 3.13B).

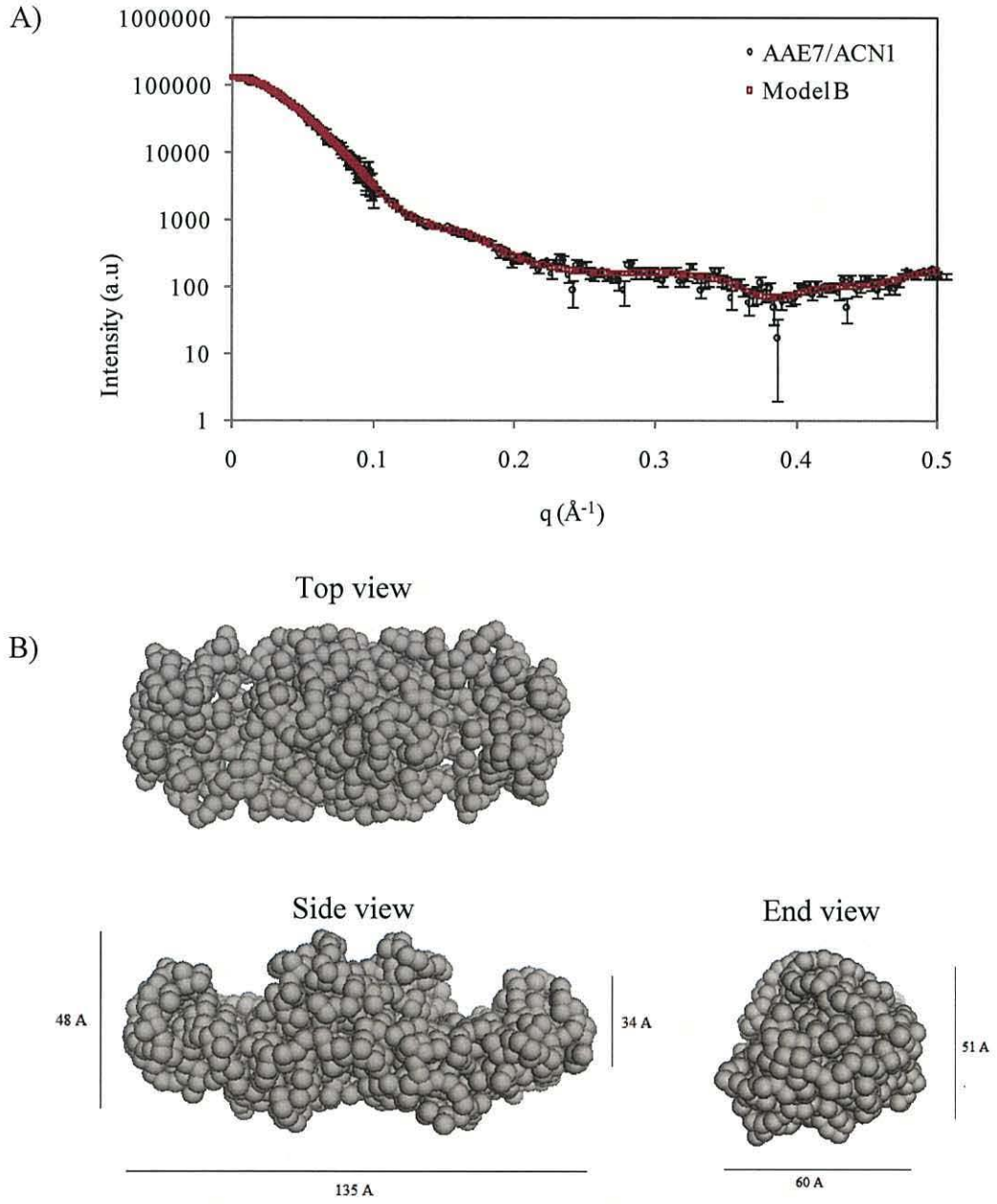


Figure 3.13: Shape reconstruction of AAE7/ACN1 implemented with GASBOR (Svergun *et al.*, 2001). A) Experimental scattering profile of AAE7/ACN1 in 0.02 M KH_2PO_4, 0.05 M KCl and 10 % glycerol, pH 7.4 compared to the scattering profile calculated from the 'best fit' GASBOR model with a χ value of 1.86. B) Top, side and end view of AAE7/ACN1 dummy residue model B created with the two-fold symmetry constraint of GASBOR. Sizes were determined and structural images created with Pymol (DeLano, 2002).

The model with the highest χ value (Model L) of 1.97 and therefore the 'worst fit' model was compared to the experimental scattering data of AAE7/ACN1 (Figure 3.14A). The profiles were found to be very similar in the low angle region, but less so in the high angle region (Figure 3.14A). A side view of the dummy residue model L shows a similar structure to model B but with less detail, for example separate domains for the C- and N- terminal are not visible (Figure 3.14B). The top view of model L also shows a different two fold symmetry compared to model B.

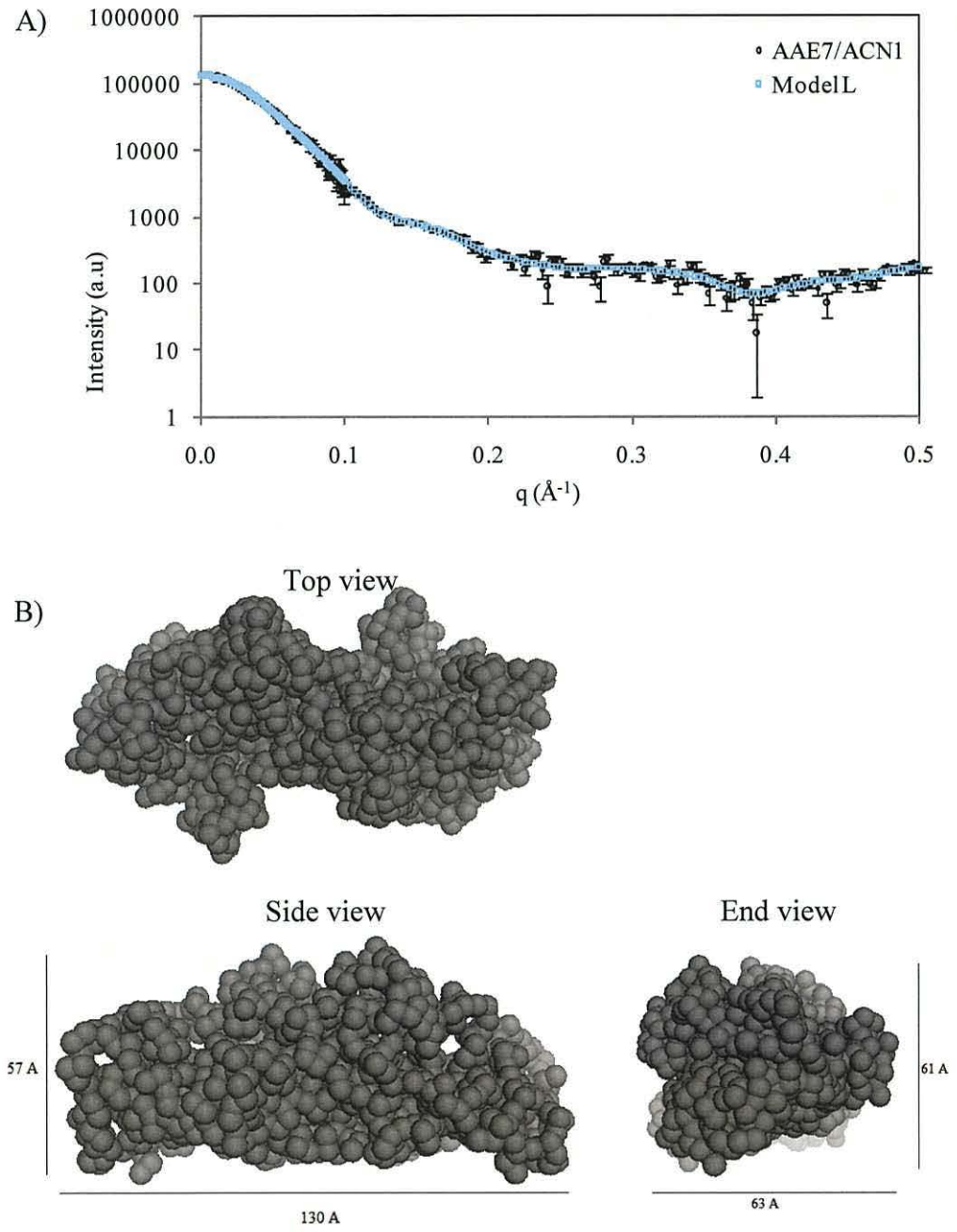


Figure 3.14: Shape reconstruction of AAE7/ACN1 implemented with GASBOR (Svergun *et al.*, 2001). A) Experimental scattering profile of AAE7/ACN1 in 0.02 M KH_2PO_4, 0.05 M KCl and 10 % glycerol, pH 7.4 compared to the scattering profile calculated from the 'worst fit' GASBOR model with a χ value of 1.97. B) Top, side and end view of AAE7/ACN1 dummy residue model L created with the two-fold symmetry constraint of GASBOR. Sizes were determined and structural images created with Pymol (DeLano, 2002).

The sixteen dummy residue models produced by GASBOR were aligned and averaged with the DAMAVER program suite (Volkov and Svergun, 2003). Prediction of the average models scattering profile and comparison to the solution

scattering profile of AAE7/ACN1 shows similarities in the very low angle region but large differences at high angles (Figure 3.15; α section). These differences are reflected in the high χ value of 11. Images of the model (Figure 3.15B) show an averaged structure with little detail.

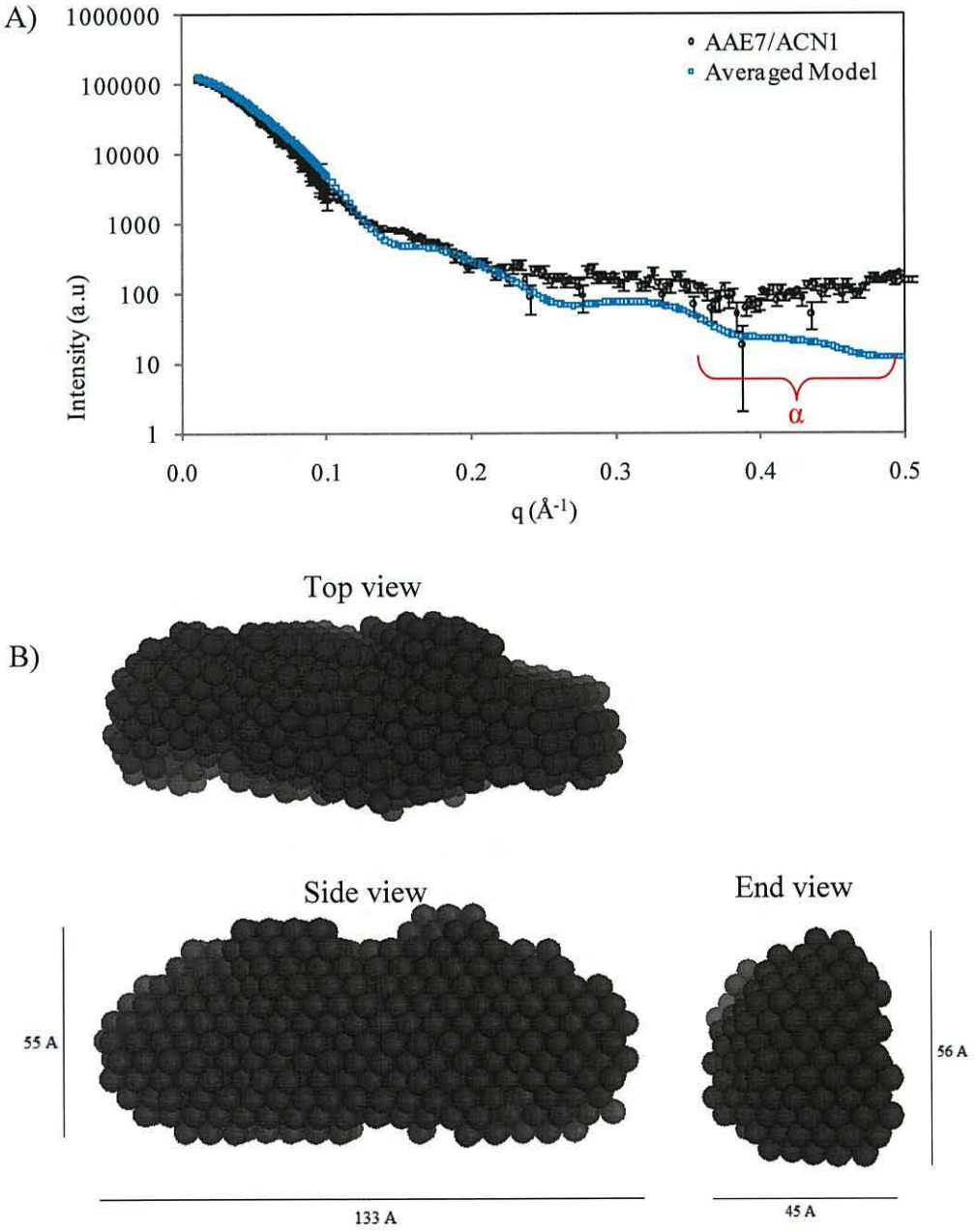


Figure 3.15: Shape reconstruction of AAE7/ACN1 implemented with DAMAVER. A) Experimental scattering profile of AAE7/ACN1 in 0.02 M KH_2PO_4, 0.05 M KCl and 10 % glycerol, pH 7.4 compared to the scattering profile produced by DAMAVER averaging of the sixteen GASBOR models with a χ value of 11. The section highlighted with a red α represents the high angle region B) Top, side and end view of AAE7/ACN1 averaged model created by DAMAVER. Sizes were determined and structural images created with Pymol (DeLano, 2002).

In view of these three comparisons, the model with the lowest χ value of 1.86 is believed to be the structure closest to that of AAE7/ACN1.

3.6 Identification of AAE7/ACN1 inhibitors

To investigate the type of inhibition kinetic analysis was conducted with two AMP analogues AMPE and AMPP (Figure 3.16). The apparent K_m and V_{max} values were estimated by varying the concentration of substrate, MgATP (section 2.2.4). These kinetic studies were conducted in preparation for future solution scattering experiments (section 3.7).

These kinetic studies reveal that the product of AAE7/ACN1 catalysis, AMP, is a mixed inhibitor of AAE7/ACN1 against ATP. A decrease in the maximum velocity (V_{max}) and an increase in the K_m (the substrate concentration at half the maximum velocity) was observed (Table 3.2; Lineweaver Burk plots shown in the Appendix). The equation of the straight line, produced by the Lineweaver Burk plot of AAE7/ACN1 without the presence of an inhibitor, was calculated to be $y= 0.0358x + 0.3015$. Using this equation the kinetic parameters, K_m and V_{max}, were calculated as 0.12 mM and 3.32 mM min^{-1} respectively.

Inhibitors classified as mixed inhibitors act both competitively and uncompetitively altering the K_m and V_{max}. However they can be further categorised into:

i) predominantly competitive inhibition where the uncompetitive inhibition constant (K_{iu}) is greater than the competitive inhibition constant (K_{ic})
ii) true mixed inhibition where K_{ic} equals K_{iu}
iii) predominantly uncompetitive inhibition where K_{ic} is greater than K_{iu}

AMP, AMPE and AMPP are all mixed inhibitors that show a decrease in V_{max} and an increase in K_m, though compared to AMP the latter two analogues have a greater inhibitory effect on AAE7/ACN1. AMP can be categorised as a predominantly competitive mixed inhibitor (group i) whereas AMPE and AMPP can be categorised as true mixed inhibitors (group ii).

Table 3.2: Kinetic studies of AAE7/ACN1 conducted at room temperature in a total volume of 1 ml with a coupled assay using citrate synthase and malate dehydrogenase varying the concentration of ATP and subsequently MgATP. A) Kinetic parametics, K_m and V_{max} of native AAE7/ACN1. B) Kinetic parameters of AAE7/ACN1 with AMP, AMPP or AMPE. K_m, $K_{m\ apparent}$, V_{max} and $V_{max\ apparent}$ were determined from a Lineweaver burk plot.

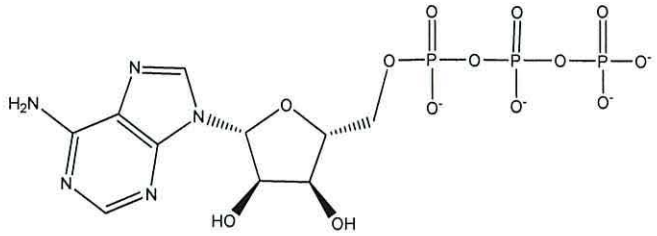
A)

		Km (mM)	Vmax (mM/min)		
native AAE7/ACN1		0.12	3.32		

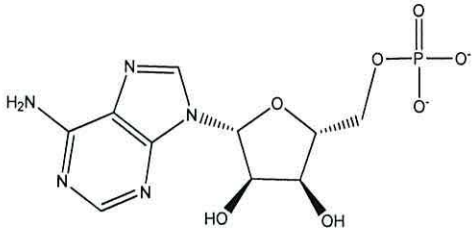
B)

Inhibitor	Inhibitor concentration (mM)	Km apparent (mM)	Vmax apparent (mM/min)	Kic (mM)	Kiu (mM/min)
AMP	50.00	3.050	1.000	2.040	21.550
	25.00	1.900	1.140	1.690	13.070
	5.00	2.487	2.700	0.250	21.770
AMPE	0.10	26.500	6.710	0.000	-0.198
	0.05	3.900	1.960	0.002	0.072
	0.01	5.380	5.090	0.000	-0.029
AMPP	0.05	2.580	0.497	0.002	0.009
	0.01	2.180	1.510	0.001	0.008

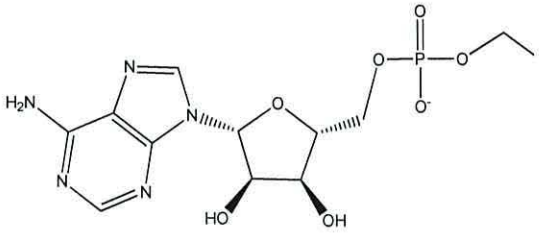
In light of these results, the strongest mixed inhibitor AMPP was added to the following AAE7/ACN1 solution scattering experiments (section 3.7) as it has the greatest potential to alter protein conformation upon binding.



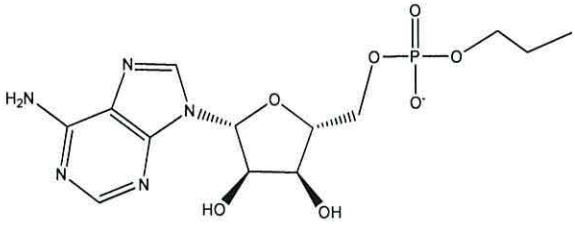
Adenosine triphosphate (ATP)



Adenosine monophosphate (AMP)



Adenosine monophosphate ethyl ester (AMPE)



Adenosine monophosphate propyl ester (AMPP)

Figure 3.16: Chemical structure of adenosine triphosphate, adenosine monophosphate and synthesised inhibitors adenosine monophosphate ethyl ester and adenosine monophosphate propyl ester created with ChemDraw.

3.7 Solution scattering of AAE7/ACN1 in the presence of adenosine monophosphate propyl ester (AMPP)

Small angle x-ray scattering data was collected at beamline 2.1, SRS Daresbury for 1 mg ml^{-1} and 2 mg ml^{-1} AAE7/ACN1 at 5.5 m camera length and 10 mg ml^{-1} at 1 m camera length in the presence of adenosine monophosphate propyl ester (AMPP). Examination of the Guinier region showed a linear relationship between the natural log of the scattering intensity and the square of the scattering vector q between 0.019 $Å^{-1} < q < 0.044$ $Å^{-1}$ indicating the solution to be monodisperse (Figure 3.17). The radii of gyration (Rg), estimated from the Guinier region was, 38.9 Å and 40.2 Å for 1 mg ml^{-1} and 2 mg ml^{-1} respectively. The R_g values for native AAE7/ACN1 were found to be 38.9 Å and 40.8 Å at 1 mg ml^{-1} and 2 mg ml^{-1} respectively (Table 3.3). These comparable radii of gyration values indicate that the overall conformation of AAE7/ACN1 does not change significantly in the presence of AMPP.

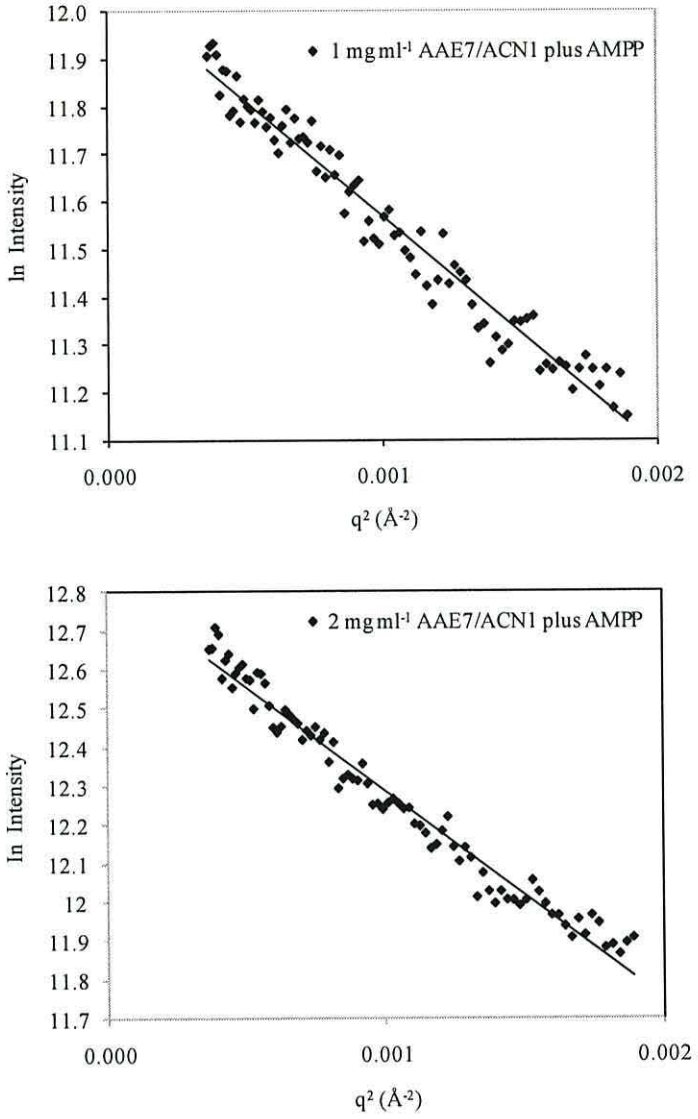


Figure 3.17: Guinier region of the small angle x-ray scattering data collected at Beamline 2.1, SRS Daresbury of 1 mg ml^{-1} and 2 mg ml^{-1} purified AAE7/ACN1 incubated with adenosine monophosphate propyl ester in 0.02 M KH_2PO_4, 0.05 M KCl and 10 % glycerol, pH 7.4.

Table 3.3: Radius of gyration calculated from the small angle region of the scattering profile of AAE7/ACN1 and AAE7/ACN1 incubated with adenosine monophosphate propyl ester (AMPP) in 0.02 M KH_2PO_4, 0.05 M KCl and 10 % glycerol, pH 7.4.

Sample concentration	Radius of gyration (Å)	
	native AAE7/ACN1	AAE7/ACN1 plus AMPP
1 mg ml^{-1}	38.9	38.9
2 mg ml^{-1}	40.8	40.2

The solution scattering profiles obtained for AAE7/ACN1 in both the absence and presence of the mixed inhibitor AMPP are very similar (Figure 3.18). The two profiles, even though slightly displaced, have the same scattering profile pattern (Figure 3.18). The similarity between the two profiles further suggests that the protein does not undergo a conformational change upon addition of AMPP under the experimental conditions implemented.

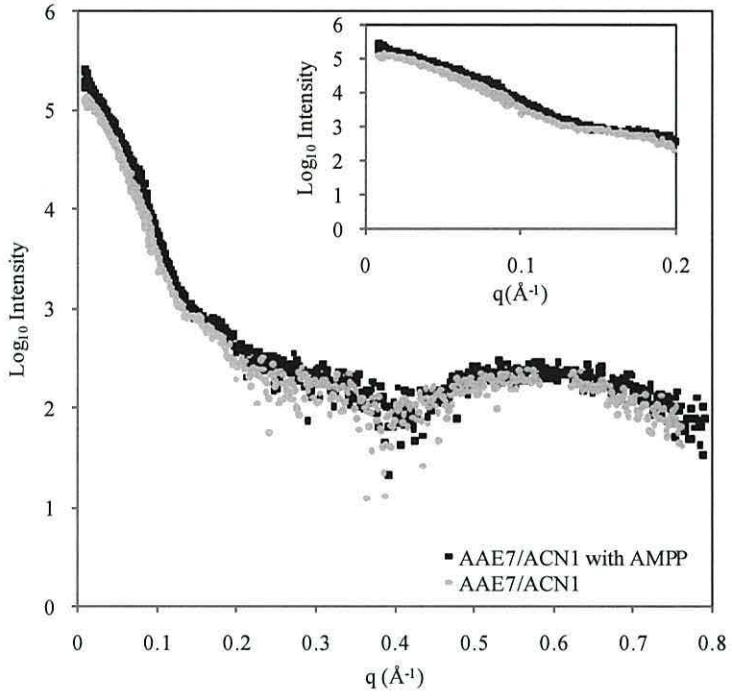


Figure 3.18: Small angle scattering of AAE7/ACN1 in the presence and absence of the mixed inhibitor adenosine monophosphate propyl ester (AMPP). The scattering profiles were constructed from scattering data collected at Daresbury, Beamline 2.1 in 0.02 M KH_2PO_4, 0.05 M KCl and 10 % glycerol, pH 7.4. The insert is an enlarged image of the small angle region of the scattering profile of AAE7/ACN1 in the presence and absence of the mixed inhibitor AMPP.

Supporting this, the small angle region of the solution scattering profile of AAE7/ACN1 was compared to the predicted scattering profile of the crystallised monomeric, dimeric and trimeric acyl-CoA synthetase (Figure 3.19). Visual comparisons clearly show that the scattering profiles of both AAE7/ACN1 with and without AMPP are very similar to each other. However they do not mimic the predicted scattering profiles as closely as they do each other (Figure 3.19).

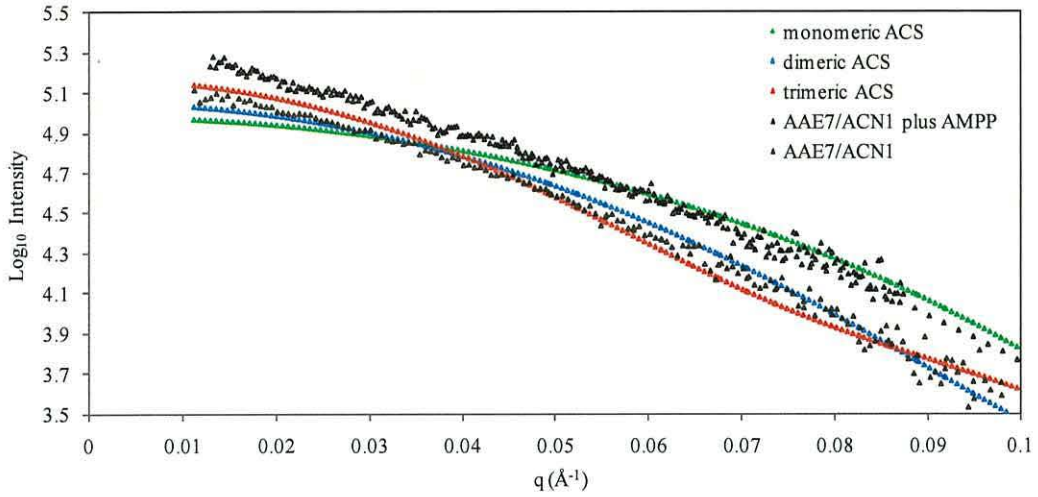
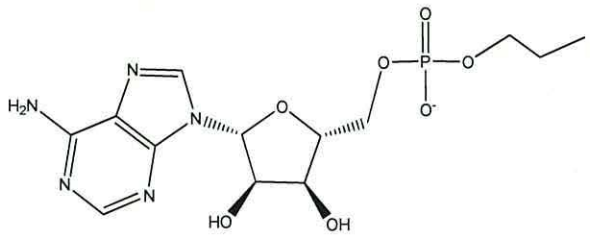
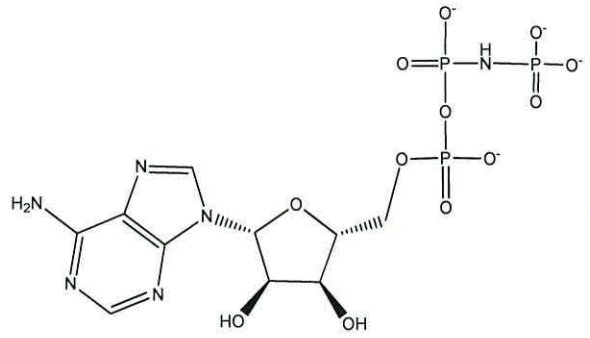


Figure 3.19: Solution scattering profile of AAE7/ACN1 both in the presence and absence of AMPP compared to the predicted solution scattering profile of the monomeric (Gulick *et al.*, 2003) and trimeric (Jogl and Tong, 2004) acetyl-CoA synthetase and the dimeric long chain fatty acyl-CoA synthetase (Hisanaga *et al.*, 2004) predicted from their atomic structure by CRYSOL (Svergun *et al.*, 1995).

Previous studies have determined the three dimensional crystallographic structure for the dimeric long chain fatty acyl-CoA synthetase (LC-FACS) from *Thermus thermophilus* (Hisanaga *et al.*, 2004). The protein's structure was determined in the absence of ligand (protein data bank 1ULT) and in the presence of the ligand analogue phosphoaminophosphonic acid adenylate ester (AMP-PNP) (Figure 3.20) (protein data bank 1V25).



Adenosine monophosphate propyl ester (AMPP)



Phosphoaminophosphonic acid adenylate ester (AMP-PNP)

Figure 3.20: Chemical structure of the AAE7/ACN1 inhibitor AMPP and the ligand, phosphoaminophosphonic acid adenylate ester of ttLC-FACS crystal structure (protein data bank 1V25).

The unbound crystallographic structure of ttLC-FACS shows an open conformation whilst the structure ligated to AMP-PNP shows a closed conformation (Figure 3.21A and B). A comparison between the predicted scattering profiles of the bound and unbound ttLC-FACS to the solution scattering data of AAE7/ACN1 shows greater similarity to the unbound predicted profile (Figure 3.21C). This similarity, reflected in the χ values of 3.03 and 10.00 (comparison to protein data bank files 1ULT and 1V25 respectively), suggests that the structure of AAE7/ACN1 may resemble an open conformation.

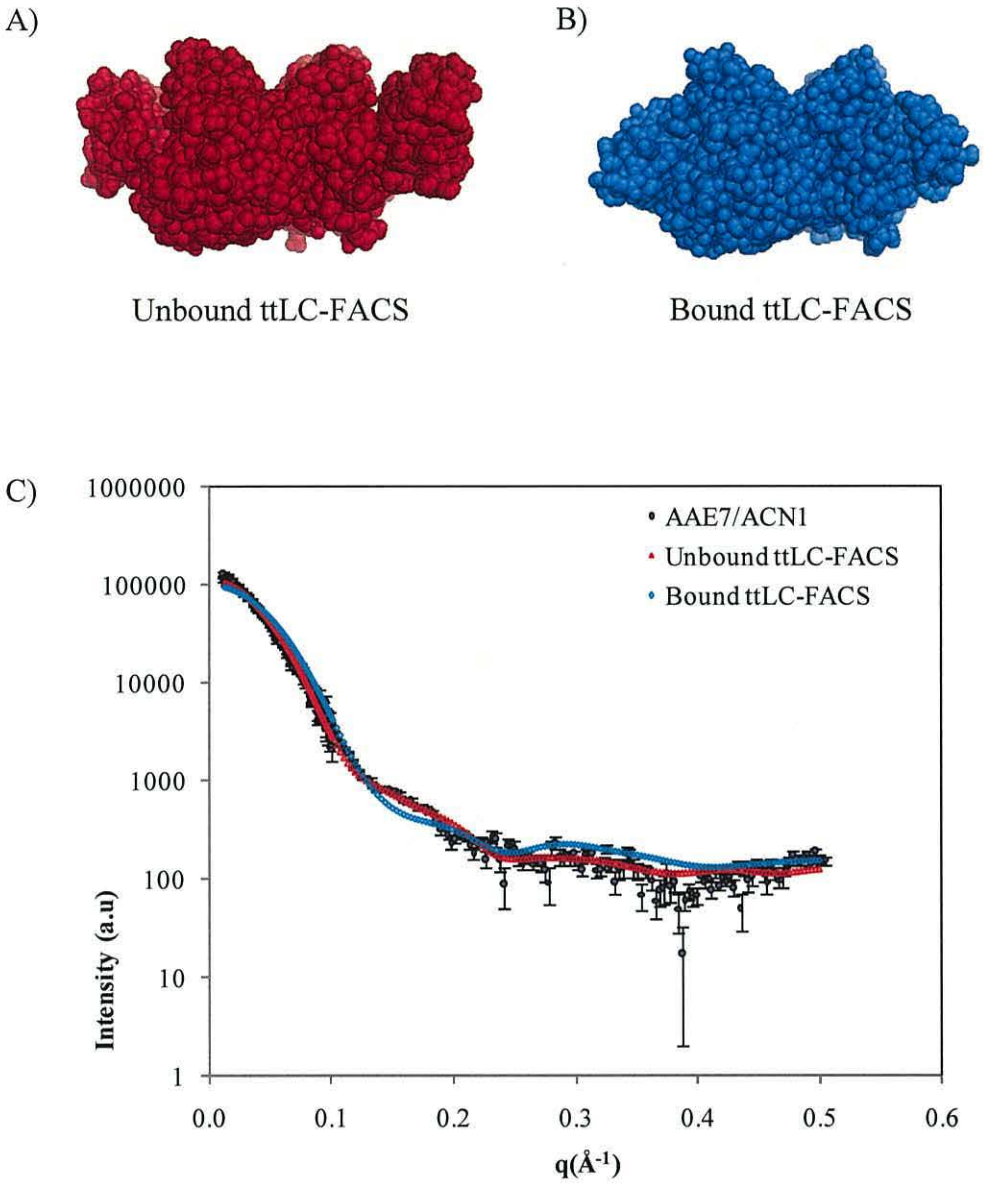


Figure 3.21: A) Sphere image of the unligated (1ULT.pdb) B) Sphere image of the ligated crystallographic structure (1V25.pdb) C) SAXS analysis predicted by CRYSOL (Svergun *et al.*, 1995) of ttLC-FACS crystallographic structures with citric acid bound (1ULT) and phosphoaminophosphonic acid-adenylate ester and magnesium ion (1V25) compared with the SAXS profile of AAE7/ACN1 with χ values of 3.03 and 10.00.

3.8 Discussion

3.8.1 AAE7/ACN1 crystallisation trials

Crystallisation trials of AAE7/ACN1 conducted with the Structure screen kit 1 were unsuccessful, yielding only precipitated protein and crystallised salt. Obtained crystals were confirmed as salt crystals by their inability to take up dye (Figure 3.2) and break with a needle (Bergfors, 1999). The presence of precipitate demonstrates that AAE7/ACN1 is precipitating out of solution at a high rate, whereas slow rates of precipitation are believed to result in the formation of crystals held together by non-covalent interactions (Bergfors, 1999).

In order to try and correct this situation a modification of the methodology was made (section 2.2.5) by microseeding crystals using cat whisker (Gulick *et al.,* 2003) as a nuclei onto which crystallisation commences more readily. Initial results, prior to microseeding, revealed a putative crystal. Despite repeating these conditions with a range of ammonium sulphate concentrations and microseeding this result was never re-produced. This could indicate that impurities, such as dust particles and fibres were present in the initial droplet (Bergfors, 1999; Thakur *et al.,* 2007).

AAE7/ACN1 crystallisation trials outlined in this chapter were based on the crystallisation conditions of the monomeric, bacterial acetyl-CoA synthetase, seACS (Gulick *et al.,* 2003). These conditions were mimicked as, at this time, AAE7/ACN1 was thought to be a monomeric protein. Subsequent work conducted within this chapter showed that AAE7/ACN1 is a dimeric protein. The crystallisation conditions of ttLC-FACS could be replicated in future studies as work presented within chapter 4 has shown greater sequence identity between AAE7/ACN1 and ttLC-FACS, compared to seACS.

The crystallographic conditions of other acetyl-CoA synthetases (Jogl and Tong 2004; Reger *et al.,* 2007) could also be mimicked in future trials to obtain protein crystals of AAE7/ACN1. However there are no common elements to the crystallisation conditions of adenylate forming enzymes, with aryl acid 2,3-dihydroxy-benzoate (DhbE) and acyl-CoA synthetase both requiring different conditions in the presence and absence of ligands (May *et al.,* 2002; Kochan *et al.,* 2009). In order to obtain crystals of AAE7/ACN1 a much larger screening process is

required covering a variety of reservoir solutions. Reservoir solution factors that could be investigated in future studies include pH, differing buffers, ionic strength and ligands to increase the probability of the protein precipitating out of solution (McPherson, 2004).

Crystallisation requires a supersaturated solution for crystals to form as the protein slowly drops out of solution (Bergfors, 1999). Trials are typically set up with a saturated solution and a supersaturated solution created over time (Bergfors, 1999). Supersaturation is obtained when the amount of water available no longer supports the high concentration of protein, commonly as a result of evaporation (Bergfors, 1999; McPherson, 2004). These conditions can also be obtained by altering the ionic strength of the reservoir solution leading to nucleation. The addition of salt reduces the amount of water available to maintain the high protein solution.

Another variable that can be altered to obtain protein crystals is pH (McPherson, 2004). Different pH levels change the structural attractions between the protein's amino acids altering the stability and packing orientation of the structure, which in turn affects the quality of the crystals produced (Bergfors, 1999). This can be achieved by either altering or omitting a buffer from the reservoir solution. The Structure Screen kit implemented within these trials covered a large pH range from 4.6 to 8.5. Further studies could be conducted at pH levels above and below this range in an attempt to obtain crystals. Whilst the Structure screen kit covered a large pH range of 4.6 to 8.5, the range was quite sparse, including pH 4.6, 5.6, 6.5, 7.5 and 8.5. Future trials could concentrate on pH values not tested within this kit. Therefore a pH gradient could also be trialled using evaporation of volatile compounds such as HCl to alter the pH of the droplet over the course of the experiment (Bergfors, 1999).

3.8.1.1 The addition of substrates

Trace elements, such as magnesium, may be essential to stabilise the protein enabling crystals to form. An example of this is the crystallisation of insulin, which is only possible in the presence of zinc (McRee, 1993). Interestingly, acyl-CoA synthetases have been crystallised both in the presence (Hisanaga *et al.*, 2004; Kochan *et al.*, 2009) and absence (Gulick *et al.*, 2003; Jogl and Tong, 2004; Kochan *et al.*, 2009) of magnesium implying that it is not an essential element for

crystallisation. However repeating and expanding the crystallisation conditions in the presence of magnesium ions could be conducted in order to investigate this factor further.

Whilst within these experiments AAE7/ACN1 was incubated prior to crystallisation with either 1) CoA and AMPP 2) CoA or 3) AMPP within this chapter other combinations of ligands have been successfully used to crystallise ACS (Gulick *et al.*, 2003; Jogl and Tong, 2004; Reger *et al.*, 2007). The variety of ligand combinations used to crystallise acyl-CoA synthetase (Table 3.4) highlights the individually of each protein. The unique conditions required for successful crystallisation of each acyl-CoA synthetase reveals the extensive amount of work still required to determine the exact conditions for successful AAE7/ACN1 crystallisation.

Table 3.4: Ligands implemented in the successful crystallisation of acyl-CoA synthetase.

Protein	Ligand	Reference
seACS	AMPP + CoA	Gulick *et al.*, 2003
scACS	AMP	Jogl and Tong, 2004
MACS2	ATP AMPCPP CoA butyryl-CoA + AMP + Mg AMP	Kochan *et al.*, 2009
ttLC-FACS	AMP-PNP + Mg	Hisanaga *et al.*, 2004
	Myristoyl-AMP + Mg	

A combination of ligands (ATP, CoA, AMP, AMP analogue and Magnesium) could be incubated with purified AAE7/ACN1 prior to future crystallisation trials in an attempt to stabilise the proteins structure.

3.8.1.2 Temperature

Another variable that can be altered to identify successful crystallisation conditions for AAE7/ACN1 is temperature. The temperature at which crystallisation occurs is very important as it determines the speed at which crystallisation occurs from a supersaturated solution. Crystallisation trials conducted within this chapter were set up at both 4°C and 18°C. Trials were conducted at 4°C as AAE7/ACN1 is stable and

can be stored at this temperature for a number of weeks prior to complete loss of activity (James Turner, personal communication). Trials were performed at 18°C as this is only slightly lower than the room temperature at which enzyme activity was measured. This slightly lower temperature was used to maintain activity for as long as possible during the lengthy period that may be required to obtain protein crystals. Future crystallisation trials could be incubated at a variety of different temperatures as differing ACS molecules have reported to successfully crystallise at 4°C, 14°C, 20°C and 22°C (Gulick *et al.,* 2003; Hisanaga *et al.,* 2004; Jogl and Tong, 2004; Kochan *et al.,* 2009). A temperature gradient could be a further development of this, where the experiment is set up at a high temperature and gradually cooled over time (Bergfors, 1999).

3.8.1.3 Tags

There are a variety of different affinity tags available to aid over-expression and purification of proteins. These include both large tags such as glutathione-s-transferase (GST) and maltose binding protein (MBP), and small tags such as histidine and arginine. The removal of any affinity tag is desirable before crystallisation trials as the presence of a tag could introduce a flexible surface region that may inhibit the formation of crystals (Stevens, 2000). However, the cleavage of such tags presents additional difficulties with extra purification steps to remove the protease and the cleaved tag (Waugh, 2005). Unfortunately, even if the affinity tag is removed it is not usually possible to remove all the residues that attach the tag to the expressed protein (Waugh, 2005).

Cloned AAE7/ACN1 (Turner, 2003; Turner *et al.,* 2005) contains an N-terminal six residue histidine tag that cannot be removed prior to expression without cloning the gene into another vector as a protease cleavage site is not present on the pQE-31 vector used in this work (section 2.1). The histidine tag is the smallest tag available, and furthermore small tags have been frequently used successfully by many research groups without the need for removal prior to crystallisation (Jogl and Tong, 2004; Kochan *et al.,* 2009). Bucher *et al.* (2002) compared the His tag, Arg tag, FLAG tag, Strep tag II and the biotin acceptor peptide (BAP) to the native, untagged protein and concluded that the sequence of certain tags (BAP, Strep, Arg) can have negative effects on both the formation of crystals and their ability to diffract, whereas His tags

did not. Another advantage of the histidine tag is its ability to purify proteins under denaturing conditions (Waugh, 2005), although this is not desirable as the protein needs to be re-folded, which in turn creates additional steps and potential problems. The successful purification of AAE7/ACN1 in an active form (section 3.3.1) would suggest that the affinity tag is surface exposed and therefore does not interfere with the protein's internal structure. It may however, introduce instability into the protein's external structure, stopping it from forming ordered crystals (Bergfors, 1999). For this reason, one possibility for further work would be to attempt to remove the histidine affinity tag, by cloning the gene into another vector, prior to crystallisation trials of AAE7/ACN1.

Large affinity tags are generally considered more problematic than smaller ones introducing greater heterogenerity to the protein (Bergfors, 1999). However, large tags such as GST have been used to aid crystallisation of small peptides in a technique called GST-driven crystallisation (Zhan *et al.,* 2001). Whilst AAE7/ACN1 is not a small peptide, large affinity tags such as GST, MBP and thioredoxin can be used for proteins, notably membrane proteins (section 3.8.1.7), that are difficult to crystallise (Smyth *et al.,* 2003). As discussed in section 3.8.1.7 and 3.8.8, if AAE7/ACN1 is a membrane associated protein then the use of these larger tags could provide another avenue for further work.

Finally, other studies have shown the possibility of protein crystallisation after purification without the use of an affinity tag. This may prevent introducing instability into AAE7/ACN1's external structure. However, as discussed in section 3.8.1.4, these non-tag methods of purification may not be suitable for AAE7/ACN1.

3.8.1.4 Purification

The purity of AAE7/ACN1 after IMAC was visualised by gel electrophoresis (Figure 3.3). Along with a large band at 66 kDa smaller bands were also present showing that AAE7/ACN1 was not 100 % pure. The purity of a protein sample used in crystallisation trials is important to achieve a well ordered crystal that is comprised of only the desired protein. If the protein sample contains contaminants, not only may this decrease the likelihood of crystal formation (Bergfors, 1999) but furthermore any crystals obtained may not form a well-ordered pattern and

consequently not diffract x-rays during crystallography. Even the presence of small quantities of contaminants could compromise confidence in the final protein structure.

An additional purification step could be implemented to purify ACS after IMAC (Stevens, 2000). These include size exclusion chromatography (Kochan *et al.*, 2009) and anion exchange chromatography (Hisanaga *et al.*, 2004; Kochan *et al.*, 2009).

3.8.1.5 Protein aggregation

Whilst it is important to optimise over-expression conditions in order to obtain the largest quantity of soluble, purified protein as possible. It is also important to ensure that the protein has not aggregated as this could also hinder crystal formation (Schlieben *et al.*, 2004; Yoshida *et al.*, 2005). Whilst it is possible to observe aggregated protein during solution scattering analysis it would have been preferable to test for protein aggregation prior to structural analysis. This would have ensured that native protein was being used for structural analysis. A variety of techniques could have been implemented to identify aggregation, and include analytical ultracentrifugation, quasielastic light scattering (Asherie, 2004), size exclusion chromatography multiangle laser light scattering (SEC-MALLS) (Ye, 2006), dynamic light scattering (Chayen, 2008) or electrospray differential mobility analysis (ES-DMA) (Pease, 2008).

3.8.1.6 Protein concentration

The concentration of protein included in the described crystallisation trials may not have been high enough to obtain a supersaturated solution before protein activity was lost. Whilst the concentration of AAE7/ACN1 utilised within these trials fell within the Structure Screen 1 kits' recommended protein range of 5 - 25 mg ml^{-1}, other studies have obtained crystals using levels up to 60 mg ml^{-1} (Stevens, 2000), which therefore could be investigated in further work in order to maximise the potential for crystal formation.

3.8.1.7 Properties of AAE7/ACN1

The identification of a glyoxysomal targeting sequence, SRL, at the C-terminal of AAE7/ACN1 together with enzymatic analysis of organelle fractions, indicates that

the protein is located within the glyoxysome of *Arabidopsis* (Turner *et al.*, 2005). AAE7/ACN1 activity measurements prove that it is able to activate both acetate and butyrate but there is some discrepancy over which substrate has the greater activity level (Shockey *et al.*, 2003; Turner *et al.*, 2005). This discrepancy arises because it is unclear whether or not AAE7/ACN1 is a membrane protein. Shockey *et al.* (2003) identified AAE7/ACN1 during amino acid comparison searches based on the membrane bound LACS family (Shockey *et al.*, 2002) and subsequently believed that it was a membrane bound protein. However, Turner *et al.* (2005) identified AAE7/ACN1 during screening studies with fluoroacetate, the toxic analogue of acetate, indicating that it is involved in acetate metabolism. This agreed with activity measurements that showed greater activity with acetate in the absence of an artificial membrane (Turner *et al.*, 2005). During screening experiments an ABC transporter was also identified that was believed to transport acetate into the glyoxysome prior to activation by AAE7/ACN1 (Hooks *et al.*, 2007). Even though a transporter has been identified it does not rule out the association of AAE7/ACN1 with the plasma membrane. Fatty acids may be transported in a similar manor to that observed in *E. coli*, where fadD encodes an inner membrane associated acyl-CoA synthetase and fadL encodes an outer membrane protein that binds fatty acids (Black *et al.*, 1991; Black *et al.*, 1992; Schaffer *et al.*, 1995).

Previously characterised membrane proteins, *E. coli* FACS (Black *et al.*, 2000), mammalian FACS (Soupene and Kuypers, 2006) and ttLC-FACS (Hisanaga *et al.*, 2004), are only active in the presence of a detergent. The solubility of AAE7/ACN1 in the absence of a detergent suggests that the protein is not associated with the membrane. However, this evidence alone is not enough to rule out the possibility of membrane association as other proteins have been previously characterised as soluble and membrane bound (Delgado *et al.*, 1998). Further work in the form of enzymatic activity measurements in the presence of a detergent, to identify any potential effect on AAE7/ACN1, and glyoxysomal membrane isolation from *Arabidopsis* seedlings, followed by enzymatic activity measurements, are required to determine if AAE7/ACN1 is associated with the plasma membrane (section 3.8.8).

If AAE7/ACN1 is identified as a membrane bound protein then this may explain why the formation of crystals was unsuccessful as membrane proteins are

particularly difficult to crystallise (Smyth *et al.,* 2003; Asherie, 2004). Detergents have been shown to aid in the crystallisation of membrane proteins by binding to hydrophobic surfaces (Garavito *et al.,* 1996) and forming a micelle around the protein (O'Stermeirer and Michel, 1997; Garavito and Ferguson-Miller, 2001), mimicking the lipid bilayer (Bergfors, 1999). Crystallisation trials of AAE7/ACN1 could be repeated with a variety of detergents in an attempt to obtain protein crystals. Unfortunately the identification of a successful detergent is usually obtained by trial and error (Bergfors, 1999). However, a promising detergent to trial would be triton-x as it was successfully used to solubilise and crystallise the membrane associated protein, ttLC-FACS (Hisanaga *et al.,* 2004). Once a suitable detergent has been identified that successfully solubilises AAE7/ACN1, then conventional crystallisation trials could commence.

If the entire structure of AAE7/ACN1 cannot be crystallised as one then one option would be to structuralise the protein in sections. The N-terminus or C-terminus of AAE7/ACN1 could be truncated to prevent the formation of a dimer or remove the glyoxysomal targeting sequence respectively (section 3.8.3.4 and 3.8.3.5).

If a high quality crystal of AAE7/ACN1 still remains elusive, then another alternative could be to structuralise a protein that shares a large degree of amino acid homology with AAE7/ACN1. This structure could then be compared to AAE7/ACN1 and used as a template for homology modelling in an attempt to indirectly obtain structural information about AAE7/ACN1. A member of the AAE superfamily of *Arabidopsis*, in particular clade VI, would be an obvious candidate for this. Unfortunately, attempts to over-express AAE2 (a member of clade VI) in a native state have proved unsuccessful as the protein is produced only in inclusion bodies that require purification under denaturing conditions (Dr Barry Grail, personal communication). However, the structuralisation of AAE2 may prove to yield un-transferable information as a function or substrate range has not been identified. Enzymatic analysis of AAE2 conducted by Shockey *et al.* (2003) showed that the protein was inactive against C2 to C14 straight chain acid substrates, unlike AAE7/ACN1. The only member of clade VI that showed any activity against straight chain acid substrates, C6 and C8, was AAE11 (Shockey *et al.,* 2003). Despite the

observed differences between substrate chain length, structural characterisation of AAE11 could aid in the characterisation of AAE7/ACN1.

Another alternative to determine the mechanism of AAE7/ACN1 action could be to over-express and crystallise an acyl-CoA synthetase from another organism. Phylogenetic analysis placed AAEL14 and AAEL15 both from *Oryza sativa* and AAEL8 from *Populus trichocarpa* within the same group as AAE7/ACN1 (Azevedo-souza *et al.*, 2008).

3.8.1.8 Nucleation

As described previously, microseeding was conducted with a cat whisker in an attempt to produce protein crystals of AAE7/ACN1. Unfortunately, nucleation with a cat whisker was unsuccessful. This could have been caused by incompatibility of the nucleating agent and a component of the reservoir solution. For example, the success of horse hair as a nucleate was increased in the absence of PEG (Thakur *et al.*, 2007). Variability could also be introduced via operator inconsistencies. As microseeding using animal fibres is conducted by the researcher, variability will occur between replicates and researchers in factors such as fibre length in contact with the droplet and duration/speed of transit. It is plausible that these could therefore be responsible for the difference in the reported success of this technique (Gulick *et al.*, 2003) and the results reported in section 3.2. Other methods of nucleation that have been reported to be successful include dried seaweed, cellulose and hydroxyapatite (Thakur *et al.*, 2007), and thus could provide avenues for further exploration.

3.8.1.9 Crystallisation methods

The hanging drop, vapour diffusion method was employed in crystallisation trials of AAE7/ACN1 section 3.2, where the solvent evaporates from the droplet creating a supersaturated protein solution that is more likely to crystallise. Other methods such as the sandwich drop (Bergfors, 1999) or sitting drop (Bergfors, 1999) could be investigated in future work. Whilst financial constraints meant that the latter technique was not investigated in this work, future experiments could be conducted using the sitting drop methodology. This has been reported to give more favourable conditions for crystal formation as the drop position results in less disruption during

nucleation and protects the drop from the outside environment (Bergfors, 1999). As the rate of diffusion is slower with the sitting drop technique it requires greater than twenty microlitres of sample. By contrast, the hanging drop method used during these trials requires only a few microlitres of sample (Bergfors, 1999) and was thus used in these experiments.

3.8.1.10 Crystallisation trial summary

As discussed above every protein is unique with a tendency to aggregate and precipitate from different solutions at different rates. This variety makes it impossible to predict the optimal protein concentration or buffer solution required for successful crystallisation (Bergfors, 1999). In the quest for the successful crystallisation of AAE7/ACN1 a large range of conditions and variables still remain unexplored. These variables include solvent conditions, temperature and the presence/combination of cofactors and affinity tags. A pre-crystallisation screen should be set up to analyse the monodispersity and stability of protein-detergent complexes in an attempt to identify successful conditions (Kawate and Gouaux, 2006). This mammoth task, of identifying optimal conditions for successful crystallisation of AAE7/ACN1, could be implemented with an automated robotic system allowing a large range of conditions to be screened at once (Heinemann *et al.*, 2003; Thakur *et al.*, 2007).

3.8.2 Structural analysis of AAE7/ACN1

Solution scattering analysis, resulting in the creation of the 'best fit' GASBOR model, and CRYSOL analysis both suggests that AAE7/ACN1 to be a dimeric protein (Figure 3.7 and 3.13). The Guinier region of the solution scattering data, collected at increasing protein concentrations, demonstrated that the protein remained monodisperse and did not aggregate at higher concentrations. Both graphs (Figure 3.5) show no difference between the radii of gyration for AAE7/ACN1. However, the observed differences in the Guinier region intensity between the graphs are a result of the increase in protein concentration. At higher concentrations there is more protein present to attenuate the x-ray intensity.

The solution scattering profile of AAE7/ACN1 was used to generate sixteen GASBOR models (A to P) with very similar profiles to the raw data (χ values of 1.86

to 1.97). Visualisation of the 'best fit' model clearly shows a dimeric structure with each monomer comprising of two domains (Figure 3.13). The 'best fit' model with the lowest χ value of 1.86 is believed to be the structure closest to that of AAE7/ACN1. This dimeric model shows a two domain structure similar to that of ttLC-FACS (Hisanaga *et al.,* 2004). A two domain structure has also been observed in the crystallographic structure of other acyl-CoA synthetases (Gulick *et al.,* 2003; Jogl and Tong, 2004; Hisanaga *et al.,* 2004; Kochan *et al.,* 2009).

Subsequently the sixteen GASBOR models (A to P) were aligned and averaged, using the program DAMAVER, to generate an averaged model. The averaging of more than ten models by DAMAVER to produce an averaged model that aims to represent the general structural features of each reconstruction is an established technique (Costenaro *et al.,* 2005; Hammel *et al.,* 2007; Edwards *et al.,* 2008; Jain *et al.,* 2009). Unfortunately, the alignment and averaging of the sixteen GASBOR models (A to P) resulted in the creation of a representative model that did not accurately represent the solution scattering data of AAE7/ACN1 (Figure 3.15). It is believed that this loss of detail was caused by model averaging and the variation in the rotation of the two monomers. Others have reported similar issues regarding DAMAVER averaging of GASBOR models, explained by the presence of hollow cavities within the cores and peripheral regions of the molecule (Vijayakrishnan *et al.,* 2010).

3.8.3 The multimeric state of AAE7/ACN1

A suite of methods were implemented to determine the multimeric state of AAE7/ACN1 with varying levels of success, namely native PAGE, gel filtration and mass spectroscopy.

3.8.3.1 Native PAGE analysis of AAE7/ACN1

The first method attempted to determine the multimeric state of AAE7/ACN1 was native PAGE, which proved inconclusive. Native gels were run in the absence of sodium dodecyl sulphate (SDS) with firstly 0.4 M Tris at pH 6.8 followed by 0.4 M Tris at pH 8.8 (section 2.2.10.2). The pH of the running buffer, loading buffer and gel were all altered with no effect. Subsequently the protein was either dialysed into the same buffer used for SAXS analysis or buffer exchanged with a pd10 column

into Tris. Unfortunately this also had no effect as the protein appeared to drop out of solution during electrophoresis and remained on the gel surface within the well.

The isoelectric focusing point (pI) of AAE7/ACN1 was determined a s 8.6 (see appendix) in an attempt to establish a suitable pH for native PAGE analysis. Due to the identification of a relatively high pI the pH of the running buffer, loading buffer and the gel was increased to pH 10.0. The buffer was altered to a higher pH as it has been reported that proteins with a pI at or above the pH of the buffer do not migrate through the gel towards the anode (Su *et al.* 1994). Unfortunately, this also proved unsuccessful and the protein remained on the gel surface. One explanation may be the stability of AAE7/ACN1 at pH 10.0. If AAE7/ACN1 is not active at this high pH then it is plausible that the protein might be unstable and dissociate. In order to determine whether this is the case, the activity of AAE7/ACN1 at pH 10.0 could be investigated in order to determine if the protein is active and in its native multimeric form.

AAE7/ACN1 appeared to remain on the surface of the gel within the well during all of the native PAGE conditions implemented. Future native PAGE analysis could be implemented with horizontal gel apparatus such as the PhastSystem (GE Healthcare, Buckinghamshire) as smaller quantities of soluble protein are required for detection. However this further study using the PhastSystem was not conducted due to financial resource constraints precluding the purchase of the necessary gel and buffer strips.

A native gradient gel could also be implemented before narrowing the acrylamide concentration in a homogenous gel. A gradient gel contains a lower acrylamide concentration near the well. This lower concentration will present less resistance and increases the likelihood that AAE7/ACN1 will enter the gel. The concentration of acrylamide is then increased through the gel to allow separation of mixed samples.

Another alternative is the addition of the PAGE compatible detergent perfluoro-octanoicacid (PFO) (Ramjeesingh *et al.,* 1999). PFO could be added to AAE7/ACN1 prior to electrophoresis to increase solubility which may prevent it from remaining

within the well. However, prior to this it is essential to conduct trials in order to ensure that this detergent does not denature AAE7/ACN1.

3.8.3.2 Gel filtration of AAE7/ACN1

The second method used to try and determine the multimeric state of AAE7/ACN1 was gel filtration. AAE7/ACN1, dissolved in the same buffer used for solution scattering analysis, eluted in the same fraction as BSA suggesting a monomeric structure. However, only about 30 % of the loaded protein ran through the column with the remainder staying on the gel surface. It is plausible that the difference in temperature observed between the solution scattering analysis, conducted at 4°C, and gel filtration, at room temperature, explains why the protein appeared to drop out of solution and remained on the columns surface. Further experiments could be conducted at 4°C in the presence of ligands to increase the stability of AAE7/ACN1. However these were not conducted due to lack of resources. Decreasing the temperature of the column to 4°C would cause the gel to contract and thus make it unsuitable for further work at room temperature (Loretta Murphy, personal communication).

3.8.3.3 Mass spectrometry of AAE7/ACN1

The third experimental method implemented to determine the multimeric state of AAE7/ACN1 was mass spectrometry. Analysis revealed both monomeric and dimeric forms of AAE7/ACN1 within the spectrum (Figure 3.10). The measured mass of the monomeric AAE7/ACN1 was 64823 Da (theoretical mass = 64480) and the measured mass of dimeric AAE7/ACN1 was 129513 Da (theoretical mass = 128900). The discrepancy between measured and theoretical mass was due to the retention of water and the buffer molecules by the protein. The deconvoluted spectrum showed that the dimeric form of AAE7/ACN1 was more prominent in solution. Furthermore the presence of monomeric AAE7/ACN1 in the spectrum was believed to be due to the harsh buffers used in this technique disassociating dimeric molecules to form monomers.

3.8.3.4 Confirming the multimeric state of AAE7/ACN1

Other methods that could have been used to determine the multimeric state of AAE7/ACN1 include analytical ultracentrifugation (Schuck, 2003) or dynamic light scattering (Phillies, 1990).

3.8.3.5 Sequence comparison of residues involved in the dimerisation of acyl-CoA synthetase

Experimental evidence in sections 3.4.1 and 3.4.4 suggests that AAE7/ACN1 is a dimeric protein with each monomer comprising of two domains. Currently, the only structuralised dimeric acyl-CoA synthetase is the ttLC-FACS that catalyses long chain substrates (Hisanaga *et al.,* 2004). Despite the difference in substrate specificity, AAE7/ACN1 and ttLC-FACS share 30 % amino acid sequence identity (see chapter 4). The dimeric structure of ttLC-FACS is created by swapping N-terminal domain residues. The dimeric structure of AAE7/ACN1 thus may be created in a similar manner as three residues are conserved out of the five (Table 3.5) that Hisanaga *et al.* (2004) state as important in domain swapping. However, it could be argued that four out of these five important residues are in effect conserved and able to contribute to the dimerisation of AAE7/ACN1 as the effect of the substitution of a glutamic acid to an aspartic acid in AAE7/ACN1 (Table 3.5) is minimal.

The major difference observed between these important residues is the substitution of Arg^{176} for tyrosine in AAE7/ACN1. However, aromatic residues have been reported to be effective at attaching subunits together (Jones and Thornton, 1995) and it is therefore not surprising to find an aromatic residue at the domain interface helping to maintain the tertiary structure.

The ttLC-FACS Asp^{15} and Arg^{199} residues, believed to be abundant at protein-protein interfaces (Jones and Thornton, 1995; Crowley and Golovin, 2005), are both conserved in AAE7/ACN1 but not in the monomeric seACS adding further credibility to the dimeric structure of AAE7/ACN1.

Table 3.5: Important residues involved in the dimerisation of ttLC-FACS (Hisanaga *et al.*, 2004) along with the corresponding residues in acyl-CoA synthetase (Gulick *et al.*, 2003).

	ttLC-FACS position				
	15	16	175	176	199
seACS	Tyr	Lys	Ala	Glu	Gly
ttLC-FACS	Asp	Glu	Glu	Arg	Arg
AAE7/ACN1	Asp	Asp	Glu	Tyr	Arg

3.8.4 The advantages of an oligomeric AAE7/ACN1

Previously crystallised ACS structures have been characterised as monomeric (Gulick *et al.*, 2003) and trimeric (Jogl and Tong, 2004). Whilst the dimeric structure of AAE7/ACN1 is not unique for an ACS (Jetten *et al.*, 1989; Martinez-Blanco *et al.*, 1991) little attempt has been made to identify the advantages of an oligomeric ACS over a monomeric ACS.

Protein oligomers are believed to have many advantages over their monomeric equivalents. Some of these advantages include allosteric control (Marianayagam *et al.*, 2004; Ali and Imperiali, 2005), a higher local concentration of active sites (Marianayagam *et al.*, 2004) and the creation of new active sites at subunit interfaces (Jones and Thorton, 1995; Goodsell and Olson, 2000).

3.8.4.1 Increased enzymatic control by acetylation?

The activity of seACS can be regulated by the acetylation of Lys^{609}, with acetylation rendering the enzyme inactive (Starai *et al.*, 2002; Gulick *et al.*, 2003). The conservation of the corresponding residue in the trimeric scACS (Jogl and Tong, 2004) may result in increased acetylation sites and level of enzymatic control. The corresponding lysine residue (Lys^{548}) is also conserved in the dimeric AAE7/ACN1, which may result in increased control over enzyme activation in response to rapidly changing physiological conditions (Gardener *et al.*, 2006) leading to increased control of acetate levels within the glyoxysome. However this lysine residue is also important in the adenylation step of the reaction where it is believed to interact with the oxygen of AMP (Gulick *et al.*, 2003; Hisanaga *et al.*, 2004). Analysis is required to determine what, if any, affect acetylation has on the enzymatic activity of

AAE7/ACN1 in order to establish if dimerisation increases the level of enzymatic control.

3.8.4.2 Increased local concentration of AAE7/ACN1 active sites?

AAE7/ACN1 is involved in the utilisation and detoxification of the weak acid, acetate. The toxic effects of a build up of acetate caused cotyledon bleaching in the AAE7/ACN1 mutant (Hooks *et al.,* 2004). Long term exposure to reduced cytosolic pH, produced by an increase in acetate levels, is believed to be detrimental to seedling development (Hooks *et al.,* 2004). Therefore the control of acetate levels, to reduce cytosolic acidification, is important during development. The dimerisation of AAE7/ACN1 increases the local concentration of active sites, enabling the fast elimination of acetate before levels become toxic. Acetate feeding studies of AAE7/ACN1 mutants show that approximately 80 % of the external supply enters the glyoxysome, and of this roughly 75 % is metabolised through AAE7/ACN1 (Turner *et al.,* 2005). To date, the fate of the remaining acetate remains unclear. However a plastidal atACS, active during germination, has been previously identified by Ke *et al.* (2000). The identification of a second ACS located within a different subcellular compartment is not surprising as acetyl-CoA, an important branch point in metabolism, and other CoA derivatives are impermeable to membranes (Fatland *et al.,* 2002). Acetyl-CoA is an important molecule involved in the TCA cycle, the glyoxylate cycle, fatty acid biosynthesis and biosynthesis of isoprenoids, flavonoids and malonated derivatives. This makes ACS and other enzymes involved in the activation of acetate very important. During conditions of low acetate, seACS is responsible for acetyl-CoA synthesis, while at higher concentrations, acetyl-CoA production is carried out by acetate kinase and phosphotransferase (Reger *et al.,* 2007). Homologous proteins of either acetate kinase or phosphotransferase have not been identified in *Arabidopsis* (Tair, 2009) increasing the importance of ACS.

3.8.4.3 The creation of an additional allosteric site?

Previous structural analysis of ACS has shown that the active site is situated at the interface between the large and small domain (Gulick *et al.,* 2003; Jogl and Tong, 2004). The competitive inhibition of AMPP against ATP for propionyl-CoA synthetase (Gulick *et al.,* 2003) suggests that the inhibitor binds only at the active

site. The mixed inhibition of AMP analogues for AAE7/ACN1 suggests that they may bind allosterically to a second position on the protein, created by dimerisation (Ali and Imperiali, 2005), as well as competing with ATP at the active site. The presence of a second binding site may be a result of dimerisation as a second site has not been identified in previous studies or any crystallised structures (Black *et al.*, 1997; Gulick *et al.*, 2003; Jogl and Tong, 2004). Further structural analysis of AAE7/ACN1, in the form of crystallisation, is required to determine if this second allosteric site could be located at, or near the dimer interface.

The mechanism of AAE7/ACN1 action should also be structurally and enzymatically investigated to determine if, like ttLC-FACS, there are separate entrance sites for the fatty acid and ATP. If AAE7/ACN1 functions in a similar manner to ttLC-FACS then there are two locations for AMPP to compete or bind to the protein and thus inhibit catalysis.

3.8.4.4 Improved stability?

The site of AAE7/ACN1 activity, the glyoxysome, has been previously referred to as the "garbage pail of the cell" (Schrader and Yoon, 2007). However, within seedlings it is the site of aerobic respiration and, as a result, ROS production (Apel and Hirt, 2004). These ROS (hydrogen peroxide, hydroxyl radicals and superoxide radicals) can cause oxidative damage to proteins, DNA and lipids (Scandalios, 1993; Apel and Hirt, 2004; Schrader and Yoon, 2007) which has already resulted in a glyoxylate cycle detour via the cytosol (Courtois-Verniquet and Douce, 1993). Whilst plant cells have evolved mechanisms to detoxify ROS (Apel and Hirt, 2004) the formation of an oligomer is believed to protect against degradation by reducing the exposed surface area (Jones and Thornton, 1995; Goodsell and Olson, 2000). Therefore the dimerisation of AAE7/ACN1 may have a protective function against degradation and thus an advantage over a monomeric form.

The peroxisomal scACS has also been characterised as an oligomeric protein (Jogl and Tong, 2004). This unique protein, characterised as a trimer, is also required to activate acetate as the peroxisomal membrane is impermeable to acetyl-CoA (van Roermund *et al.*, 1995). However, the advantages of a trimeric protein were reported as being unknown (Jogl and Tong, 2004) and no subsequent publications have shed

light on this matter. An octameric ACS (paACS) has been characterised from the hyperthermophilic organism *Pyrobactulum aerophilum* (Brasen *et al.,* 2005). Unsurprisingly paACS is thermostable and thermoactive at temperatures around 100°C. The ability of paACS to activate acetate under these conditions could be due to its homooctameric structure. The higher oligomeric state could be a significant stabilizing mechanism for hyperthermophilic proteins. Other ACS proteins that are active at high temperatures, 65-70°C, include the dimeric and trimeric proteins from *Methanothermobacter thermautotrophicus* and *Archaeoglobus* respectively (Ingram-Smith and Smith, 2006; Ingram-Smith *et al.,* 2006). *Thermus thermophilus* is also a thermophilic organism that may increase the oligomeric state of its proteins as a stabilisation strategy (Ali and Imperiali, 2005). Unfortunately this may not explain why ttLC-FACS is a dimeric protein as a large proportion of LC-FACS function as homodimers (Stinnett *et al.,* 2007). Perhaps then, AAE7/ACN1 evolved from a LC-FACS as it shares greater sequence homology with the MC-ACS, AAE11, than atACS (see chapter 4).

3.8.5 Kinetic analysis of AAE7/ACN1

Kinetic studies with the AMP analogues, AMPP and AMPE, were conducted to identify the degree and type of inhibition for future solution scattering experiments with AAE7/ACN1. Results showed that AMP, AMPP and AMPE are all mixed inhibitors, with AMPP and AMPE reducing the specific activity to a greater extent. The mixed inhibitors, AMPP and AMPE, are believed to form a larger number of interactions with the protein and are therefore more likely to cause a conformational change in shape than AMP. Greater inhibition also indicates that ATP cannot displace AMPP or AMPE as easily as AMP, which may be due to tighter binding of these inhibitors at or near the active site. With this in mind, solution scattering data was collected both in the presence and absence of the mixed inhibitor, AMPP, to determine if AAE7/ACN1 undergoes a conformational change in shape during catalysis (section 3.8.6). AMPP has also been identified as an inhibitor of the related propionyl-CoA synthetase (Gulick *et al.,* 2003) but inhibition analysis has not been previously conducted on ACS.

The kinetic analysis of AAE7/ACN1 with AMP analogues should be refined and repeated with lower concentrations of MgATP to confirm their mixed inhibition

nature as previous kinetic analysis of AMP inhibition against ATP has demonstrated competitive inhibition (Behal *et al.,* 2002). A reduction in the substrate concentration will result in increased accuracy of the calculated kinetic parameters and reduce the variability observed within the data (see appendix). The enzymatic assays implemented by Reger *et al.* (2007) could also be used to study the inhibitor's kinetic parameters of AAE7/ACN1. Implementing the same method of analysis will allow a direct comparison between the two enzymes.

3.8.6 Does AAE7/ACN1 undergo a conformational change during catalysis?

Under the experimental conditions implemented within this chapter results suggest that AAE7/ACN1 does not undergo a conformational change during catalysis. Comparison of AAE7/ACN1 solution scattering profiles in the absence and presence of AMPP showed that the profiles both follow the same pattern, despite a slight displacement (Figure 3.18). The displacement observed between the two profiles is likely to be due to a slight variation in protein concentration, resulting from the addition of AMPP.

The solution scattering profile was predicted, from the bound (1V25.pdb) and unbound (1ULT.pdb) crystallographic structure of ttLC-FACS, to determine if a change in conformation can be detected with this technique. The difference observed between the predicted profiles of the crystallographic structures (Figure 3.21) clearly demonstrates that a conformational change in shape is detectable by solution scattering. Therefore if a conformational change in shape is induced by the addition of AMPP it should be detectable.

The crystallographic structure of seACS was determined bound to both AMPP and CoA revealing a new conformation of the enzyme family for the thioester forming reaction (Gulick *et al.,* 2003). Perhaps then, the addition of CoA is also required to create a domain rotation into the thioester forming conformation and explains why a conformational change was not observed for AAE7/ACN1. A CoA induced structural change was also suggested by Bar-Tana and Rose (1968) for butyryl-CoA synthetase and by Bar-Tana and colleagues (1973) for palmitoyl-CoA synthetase. However, crystal structures have also been obtained in the thioester conformation (Table 3.6) bound only to AMP or CoA (Kochan *et al.,* 2009) suggesting that the

presence of AMPP alone may be enough to generate a conformational change in shape. Despite this observation, structural analysis of AAE7/ACN1 should be conducted with the addition of CoA to determine if a conformational change can be induced.

Table 3.6: Conformation undertaken by acyl-CoA synthetase crystallised in the presence of a range of ligands. Acyl-CoA synthetase catalyses a two-step reaction: a) the adenylation reaction followed by b) the thioester forming step.

Protein	Ligand	Suggested reaction step	Suggested conformation	Reference
seACS	AMPP + CoA	Thioester	Start of the second reaction	Gulick *et al.*, 2003
scACS	AMP	Adenylation	First conformation	Jogl and Tong, 2004
MACS2	i) ATP ii) AMPCPP	Adenylation	First conformation	Kochan *et al.*, 2009
	i) CoA	Adenylation	Close to adenylation conformation	
	i) Butyryl-CoA + AMP + Mg ii) AMP	Thioester	Second conformation	
	i) Unligated	Thioester	End of the second conformation	
AAE	Unligated	Thioester	Second	Shah *et al.*, 2009
ttLC-FACS	Unligated	Adenylation	Open	Hisanaga *et al.*, 2004
	AMP-PNP + Mg	Adenylation	Closed	
	Myristoyl-AMP + Mg	Thioester	Closed	

The binding of AMPP alone to AAE7/ACN1 may not create the necessary interactions required to generate a closed structure and perhaps a second ligand is required to induce a conformational change in shape. The addition of magnesium along with AMPP may increase the number of interactions between the ligand and AAE7/ACN1 enabling it to undergo a conformational change in shape. Structural analysis of AAE7/ACN1 should be repeated in the presence of magnesium to try to identify a conformational change as ttLC-FACS and MACS2 (crystallised in the closed conformation) both contain a magnesium ion (Hisanaga *et al.*, 2004; Kochan *et al.*, 2009). However, magnesium may prove not to be critical as the scACS, MACS2 and AAE structures that also represent the thioester conformation, were

obtained in the ion's absence (Gulick *et al.,* 2003; Kochan *et al.,* 2009; Shah *et al.,* 2009).

The crystallographic structure of ttLC-FACS bound to myristoyl-AMP shows that Thr^{184} and Glu^{328} are bound to the magnesium ion. Sequence alignment analysis shows that both of these residues are conserved in AAE7/ACN1 (Thr^{205} and Glu^{347}) indicting the importance of magnesium. Whilst ttLC-FACS is inactive in the absence of both detergent and magnesium, detergent is not required to activate AAE7/ACN1. Whilst the necessity of magnesium ions for AAE7/ACN1 activity has not been explored, it is widely believed to be important for ATP binding (Behal *et al.,* 2002; Schmelz and Naismith, 2009). Magnesium binds tightly to ATP with a high dissociation constant of 20000 M^{-1} this is in contrast to AMP which has a much lower dissociation constant of 100 M^{-1} and weaker binding (Purich and Fromm, 1972). Webster (1963; 1967) showed that magnesium stimulates the first half reaction but is not required for the formation of acetyl adenylate from acetyl-CoA and AMP. The omission of magnesium has been shown to reduce activity, but does not result in the complete inactivation of adenylate forming enzymes (Webster and Campagnari, 1962; Reger *et al.,* 2008; Kochan *et al.,* 2009).

AMP, like its analogue AMPP, is a mixed inhibitor of AAE7/ACN1. The crystal structure of scACS with AMP alone represents the adenylation conformation (Jogl and Tong, 2004). Perhaps a conformational change was not detected as the binding of AMPP simply mimics the actions of AMP and the solution scattering models represent the same catalytic step. In order to investigate this further structural analysis could be repeated with AMP-PNP and magnesium as ttLC-FACS has been crystallised in the second step conformation in their presence (Hisanaga *et al.,* 2004). The addition of magnesium to MACS2 along with butyryl-CoA and AMP also lead to the creation of the second step conformation and therefore could be replicated for AAE7/ACN1 analysis. Surprisingly the addition of an ATP analogue (AMPCPP) without magnesium to MACS2 resulted in an adenylate conformation (Kochan *et al.,* 2009) highlighting the potential importance of magnesium in the formation of the second step conformation.

Crystal structures of the adenylate forming family show that the unligated enzyme is able to form both the adenylate (May *et al.,* 2002; Reger *et al.,* 2008; Hisanaga *et al.,* 2004) and thioester (Kochan *et al.,* 2009; Shah *et al.,* 2009) conformation making it difficult to predict which conformation unligated AAE7/ACN1 has adopted during structural analysis. Comparison between the predicted scattering profiles of the bound and unbound ttLC-FACS to AAE7/ACN1 showed that AAE7/ACN1 shares greater similarity to the unbound profile (Figure 3.21). This suggests that, under these experimental conditions, AAE7/ACN1 adopts an open conformation.

Further structural analysis of AAE7/ACN1 is required to determine if the protein is able to undergo a conformational change during catalysis both in the absence and presence of ligands. Structural techniques that could be implemented include x-ray crystallography and circular dichroism (CD). T he far-UV region of the circular dichroism (CD) spectrum is especially useful since it provides a quantitative measure of the average secondary structure content (Pain, 1994).

3.8.7 Conservation of important hinge residues in AAE7/ACN1

The hinge region, also known as the A8 sequence, is important in catalysis for the movement of the small domain towards the larger domain in ACS. Sequence comparison of AAE7/ACN1 to the structurally characterised seACS, scACS and ttLC-FACS (Table 3.7) suggests that the protein may be capable of undergoing a conformational change as important residues within the linker region are conserved.

The A8 sequence contains a conserved Dx_6GxR motif with an invariant glycine (Gly^{524}) residue (Gulick *et al.,* 2003) that is also conserved in AAE7/ACN1 (Table 3.7).

Table 3.7: Hinge sequence of AMP binding proteins. Italicised resides show the glycine residue that is believed to add flexibility to the linker allowing two conformations to be formed. Underlined residues indicate the conserved aspartate hinge residue and the highly conserved glycine residue is depicted in bold. Sequence alignment was conducted with Multalin (Corpet, 1988).

Enzyme	Multimeric state	A8 sequence	Reference
seACS	Monomer	^{514}*G*RV<u>D</u>DVLNVS**G**HR^{526}	Gulick *et al.*, 2003
scACS	Trimer	*G*RV<u>D</u>DVVNVS**G**HR	Jogl and Tong, 2004
ttLC-FACS	Dimer	*D*RL<u>K</u>DLIKSG**G**EW	Hisanaga *et al.*, 2004
AAE7/ACN1	Dimer	*D*RS<u>K</u>DVIISG**G**EN	This study

Crystallographic analysis conducted by Gulick *et al.* (2003) showed that Asp^{517} (D) is the hinge residue responsible for the two structural conformations of seACS (Gulick, 2009). Whist this residue appears to be involved in domain movement Reger *et al.* (2007) have shown, by mutagenesis, that an aspartate residue at this position is not essential for catalysis. However, substitution to a proline residue reduced enzymatic activity and conformational flexibility, making it difficult for the protein to enter a stable adenylate conformation. The corresponding residue in luciferase, like AAE7/ACN1, has been identified as lysine. Despite this residue change, luciferase is still predicted to undergo a conformational change as the active site residues of the crystal structure are not in direct contact, requiring domain movement prior to catalysis (Conti *et al.*, 1996; Reger *et al.*, 2007; Gulick, 2009). The presence of a lysine residue at this position is also true for ttLC-FACS, which has been shown to undergo a conformational change prior to catalysis (Hisanaga *et al.*, 2004). The conservation of a lysine residue in AAE7/ACN1 suggests that a conformational change in shape should be possible.

The Arg^{526} residue of seACS is not conserved in either ttLC-FACS or AAE7/ACN1. This residue is part of the conserved Dx_6GxR motif and forms a salt bridge across the active site of seACS with Glu^{417}. However, the interaction between these two residues is not essential for catalysis as 4-chlorobenzoyl-CoA ligase (CBAL) contains an asparagine residue at this position (Gulick *et al.*, 2004; Wu *et al.*, 2007). The corresponding residue in AAE7/ACN1 is also an asparagine residue. The variability observed at the Arg^{526} site of seACS suggests that the formation of a salt bridge is not essential for catalysis as it is not conserved in CBAL, ttLC-FACS or

AAE7/ACN1. However, Glu^{417}, the residue that Arg^{526} forms a salt bridge with, is conserved in AAE7/ACN1. This glutamate residue, within the A5 or GxTE motif (Gulick *et al.,* 2003; Shah *et al.,* 2009), plays an important part in the positioning of the magnesium ion. The conservation of this residue suggests that the magnesium ion, which binds with ATP, can be correctly coordinated during catalysis.

The substitution of Gly^{514} in seACS to an aspartate residue in ttLC-FACS is believed to reduce the linkers' flexibility and perhaps explain why seACS and scACS both adopt two different conformations for the adenylation and thioester reactions (Gulick *et al.,* 2003; Jogl and Tong, 2004). The presence of an aspartate residue at this position in AAE7/ACN1 may suggest that, similar to ttLC-FACS, the same conformation is adopted for both the adenylation and thioester half reactions. This may explain why a conformational change was not observed after the addition of AMPP.

3.8.8 Membrane association

As discussed earlier (sections 3.8.1.3 and 3.8.1.7), it has been suggested that AAE7/ACN1 may be a membrane associated protein. Techniques that could be implemented to determine if AAE7/ACN1 is a membrane protein include glyoxysomal fractionation followed membrane purification (Leighton *et al.,* 1969; Schwitzguebel and Siegenthaler, 1984) and enzymatic analysis (Millerd and Bonner, 1954) or construction of a GFP protein fusion followed by confocal microscopy (Curtis and Grossniklaus, 2003; Koroleva *et al.,* 2005). Sucrose density gradients have been previously used to isolate membranes from other organelles prior to disruption and separation into matrix, core and membrane fractions (Stumpf and Conn, 1980).

3.8.9 Summary

Work conducted within this chapter clearly demonstrates that AAE7/ACN1 is a dimeric protein in solution. The construction of a structural model suggests that each monomer of AAE7/ACN1, like other acyl-CoA synthetases, is comprised of two domains. The solution scattering profiles of AAE7/ACN1 in the absence and presence of AMPP along with very similar Rg values at increasing protein concentrations strongly suggest that, under these experimental conditions,

AAE7/ACN1 does not undergo a conformational change in shape. The conservation of important residues within the hinge sequence between AAE7/ACN1 and ttLC-FACS (section 3.8.7) implies that movement is possible. However, further analysis is required in the presence of additional ligands and a metal ion to fully determine if a conformational change in shape occurs during catalysis.

Homology modelling of AAE7/ACN1

4.1 Introduction

AAE7/ACN1 is a member of the adenylate forming superfamily of *Arabidopsis thaliana* responsible for the conversion of acetate to acetyl-CoA for entry into the glyoxylate cycle (Turner *et al.,* 2005). The glyoxysomal acetyl-CoA synthetase, AAE7/ACN1 was placed in clade VI by Shockey *et al.* (2003) along with a medium chain fatty acyl-CoA synthetase (AAE11), though in a different group to the previously characterised *Arabidopsis* short chain acetyl-CoA synthetase (Ke *et al.,* 2000) which was placed in clade II. The search, based on the presence of an AMP signature motif, showed that AAE7/ACN1 shares greater similarity to a medium chain fatty acyl-CoA synthetase compared to a short chain fatty acyl-CoA synthetase but is only able to catalyse substrates shorter than four carbons. In chapter 3 of this thesis, AAE7/ACN1 was characterised as a dimeric protein that contains the two domain structure observed in previously characterised members of the adenylate forming family (Gulick *et al.,* 2003; Hisanaga *et al.,* 2004; Jogl and Tong, 2004).

Members of the adenylate forming family catalyse the ATP dependent acetylation of fatty acid in a two step reaction by activating the unreactive carboxylic acid (COOH) and transforming the hydroxyl (OH) leaving group into AMP (Schmelz and Naismith, 2009). The activation of various carboxylic acid substrates through the formation of a thioester bond proceeds via a bi uni uni bi ping-pong mechanism (Bar-Tana and Rose, 1968; 1973). The uni and bi terms refer to the number of substrates that enter the enzyme and the number of products that leave the enzyme (Bisswanger, 2008). Uni describes a situation where a single substrate leaves or enters the enzyme, pyrophosphate (PPi) and coenzyme A (CoA) respectively, and bi describes a situation where two substrates enter or leave the enzyme at the same time, namely ATP and a fatty acid or AMP and fatty acyl-CoA. Ping-pong signifies that a product is released (pyrophosphate) before another substrate (CoA) can bind to the enzyme (Sachs, 1980; Frey and Hegeman, 2007) (Figure 4.1 and 4.2).

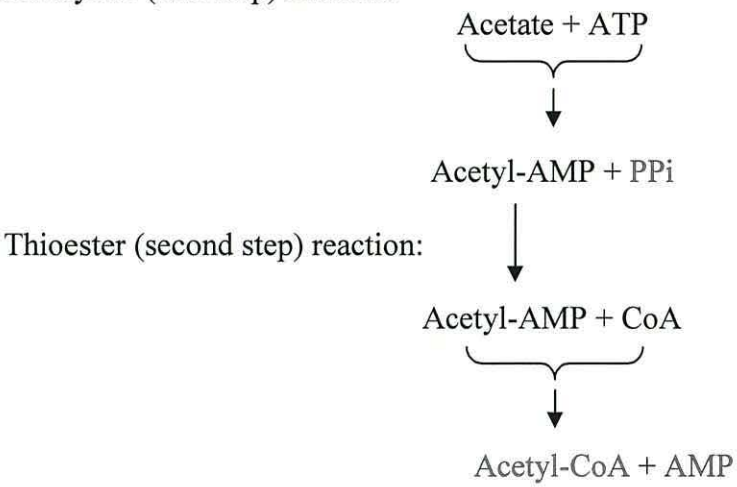


Figure 4.1: Two step reaction mechanism of acyl-CoA synthetase. Molecules highlighted in grey represent the three products of the reaction.

In step one (Figure 4.1 and 4.2), ATP and a fatty acid combine with the enzyme to yield an enzyme bound intermediate, acyl-AMP, before the release of PPi (Bar-Tana and Rose, 1973; Bar-Tana *et al.,* 1973; Young and Anderson, 1974). The acyl-AMP intermediate becomes the substrate for the second reaction (Gulick *et al.,* 2003).

In the second step (Figure 4.1 and 4.2), CoA enters the enzyme's active site and reacts with acyl-AMP to produce acyl-CoA (Bar-Tana and Rose, 1973). CoA binding causes a rotation of the C-terminal domain in seACS positioning the CoA for nucleophilic attack on the acyl group (Gulick *et al.,* 2003). Two products, acyl-CoA and AMP, are released at the end of the reaction (Starai and Senerena, 2004).

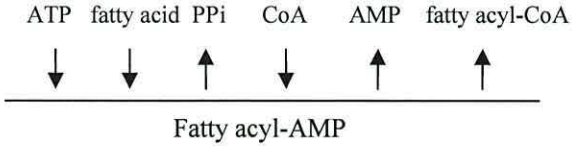


Figure 4.2: Mechanism of acyl-CoA synthetase demonstrating the order molecules are utilised during the production of fatty acyl-CoA.

Attempts to obtain the crystallographic structure of AAE7/ACN1 have so far proved unsuccessful (section 3.2). Comparative modelling techniques are employed within this chapter to obtain information about the structure of AAE7/ACN1. Comparative modelling of AAE7/ACN1 is important in order to identify potentially important residues involved in catalysis. This, along with further experiments, may help to establish if AAE7/ACN1 is in fact associated with the plasma membrane and able to

catalyse butyrate at a greater rate than acetate (Shockey *et al.*, 2003) or is free in the glyoxysome where it catalyses acetate at a greater rate than butyrate (Turner *et al.*, 2005). It may also help to determine why AAE7/ACN1 is not active with substrates that are greater than four carbon atoms in length (Turner *et al.*, 2005).

4.2 AAE7/ACN1 sequence comparison to the protein data bank

The amino acid sequence of AAE7/ACN1 was compared to known crystallographic structures with the threading program Phyre (www.sbg.bio.ic.ac.uk/phyre/) to identify potential template structures for homology modelling. The crystallographic structure of the LC-FACS from *Thermus thermophilus* (ttLC-FACS) (Hisanaga *et al.*, 2004) was identified with the largest amino acid sequence identity to AAE7/ACN1 of 30 % (Table 4.1). Other crystal structures identified (sharing between 20 and 25 % amino acid sequence identity to AAE7/ACN1) include 4-chlorobenzoyl-CoA ligase / synthetase (CBAL), benzoate coenzyme A ligase, 2, 3-dihydroxybenzoate AMP ligase (DhbE), acetyl coenzyme A synthetase from *Saccharomyces cerevisiae* (scACS) and *Salmonella enteric* (seACS) (Table 4.1). All of these enzymes are members of the adenylate forming enzyme family and are comprised of two domains with the active site at the interface between them.

Table 4.1: Amino acid sequence comparison between AAE7/ACN1 and known crystallographic structures implemented with the online server Phyre (www.sbg.bio.ic.ac.uk/phyre/).

Protein	Protein Data Bank	Sequence identity to AAE7/ACN1 (%)
Long chain fatty acyl-CoA synthetase	1ULT	30
4-Chlorobenzoyl-CoA ligase/synthetase	1T5H	26
Benzoate coenzyme A ligase	2V7B	23
2,3-Dihydroxybenzoate AMP ligase	1MDB	22
Medium chain acyl-CoA synthetase (*H. sapiens*)	3C5E	20
Acetyl coenzyme A synthetase (*Saccharomyces cerevisiae*)	1RY2	19
Acetyl coenzyme A synthetase (*Salmonella enteric*)	1PG4	19

Whilst AAE7/ACN1 activates short chain fatty acid substrates, sequence analysis shows that it shares greater sequence identity to ttLC-FACS that utilises long chain

fatty acid substrates, such as myristate (Figure 4.3). Sequence alignment analysis of scACS, seACS, ttLC-FACS and AAE7/ACN1 (Figure 4.4) shows that whilst all four proteins have conserved regions (shown in red on Figure 4.4) there does not appear to be a particular section that is conserved between scACS, seACS and AAE7/ACN1 that is not present in ttLC-FACS.

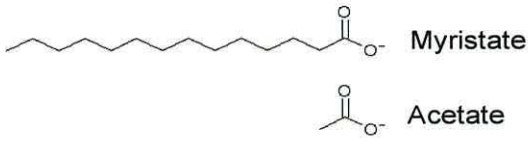


Figure 4.3: Skeletal structure created with ChemDraw of the preferred ttLC-FACS substrate, myristate ($C_{14}H_{28}O_2$) and AAE7/ACN1 substrate, acetate ($C_2H_3O_2$).

Figure 4.4: Sequence alignment of seACS (1PG4), scACS (1RY2), ttLC-FACS (1ULT) and AAE7/ACN1 conducted by multalin (Corpet, 1988). Red represents highly conserved regions between the four enzymes and blue represents regions that are moderately conserved between the four enzymes. All other residues are represented in black.

Sequence comparison studies implemented with the default settings of emboss (Rice *et al.,* 2000) reveal that AAE7/ACN1 shares greater sequence identity with the medium chain acyl-activating enzyme (AAE) 11 (At1g66120) than the monomeric acetyl-CoA synthetase (At5g36880) characterised by ke *et al.* (2000). Whilst AAE7/ACN1 shares 45.0 % identity and 59.2 % similarity at the amino acid level

with AAE11 (Figure 4.5), it only shares 18.4 % identity and 30.6 % similarity with the monomeric acetyl-CoA synthetase (Figure 4.6). Shockey *et al.* (2003) suggests that there are two other acetyl-CoA synthetases' in *Arabidopsis*, AAE17 (At5g23050) that shares 16.1 % identity and 27.5 % similarity and AAE18 (At1g55320) that shares 15.1 % identity and 26.0 % similarity with AAE7/ACN1.

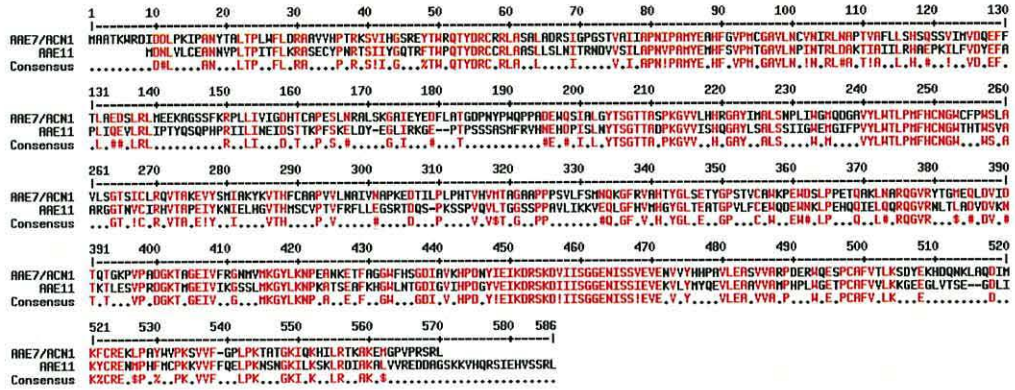


Figure 4.5: Sequence alignment of AAE11 and AAE7/ACN1 conducted with Multalin (Corpet, 1998). Red represents highly conserved regions between the two enzymes. All other residues are represented in black.

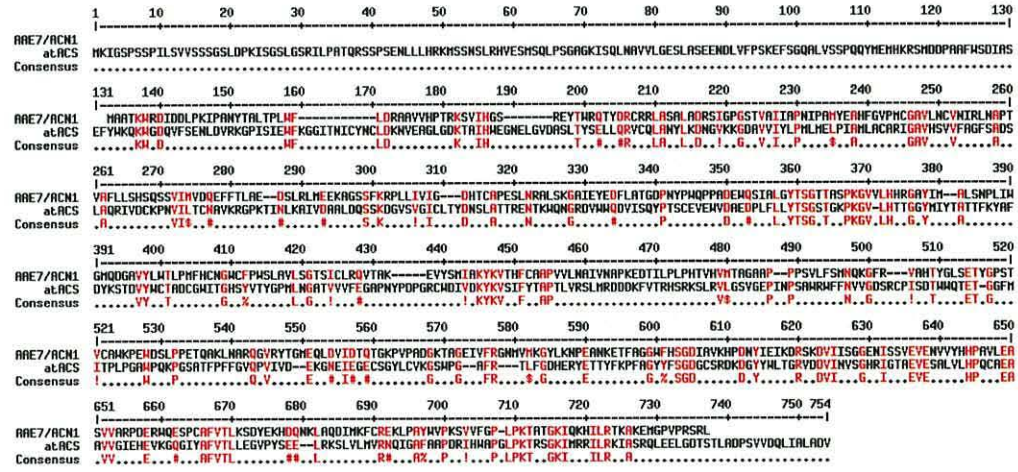


Figure 4.6: Sequence alignment of atACS and AAE7/ACN1 conducted with Multalin (Corpet, 1998). Red represents highly conserved regions between the two enzymes. All other residues are represented in black.

4.3 Comparative modelling of AAE7/ACN1

The dimeric structure of ttLC-FACS was chosen as the most suitable template structure for homology modelling of AAE7/ACN1. Prior to the construction of a

comparative model, the suitability of the potential template structures were evaluated using the default settings of Swiss-pdbViewer (Guex and Peitsch, 1997) to identify residues that fall within the disallowed region of the ramachandran plot.

4.3.1 Evaluation of the potential template structure

The dimeric structure of ttLC-FACS has been crystallised in both an unbound (1ULT.pdb) and bound state (1V25.pdb and 1V26.pdb) by Hisanaga *et al.* (2004). The native structure of ttLC-FACS was crystallised with a resolution of 2.55 Å (1ULT). The crystallographic structure of ttLC-FACS bound with adenosine monophosphate, magnesium ion and myristic acid was structuralised at a similar resolution of 2.50 Å (1V26). The third crystal structure of ttLC-FACS bound with phosphoaminophosphonic acid adenylate ester (AMP-PNP) and magnesium ion was crystallised at a resolution of 2.30 Å (1V25).

4.3.1.1 Evaluation of the crystallographic structure of ttLC-FACS, 1ULT

The crystallographic structure of ttLC-FACS (1ULT) was analysed and no bad contacts or side chain angles were identified. Residues shown in red (chain A: Thr^{187}, His^{230} and Glu^{328} and chain B: Lys^{431} and Arg^{519}) on the ramachandran plot were identified as unfavourable backbone angles and residues shown in blue represent unfavourable glycine residues (Figure 4.7). This analysis showed that 3.5 % of the total residues fell within the unfavourable region of the ramachandran plot, with 86.1 % of these being glycine residues.

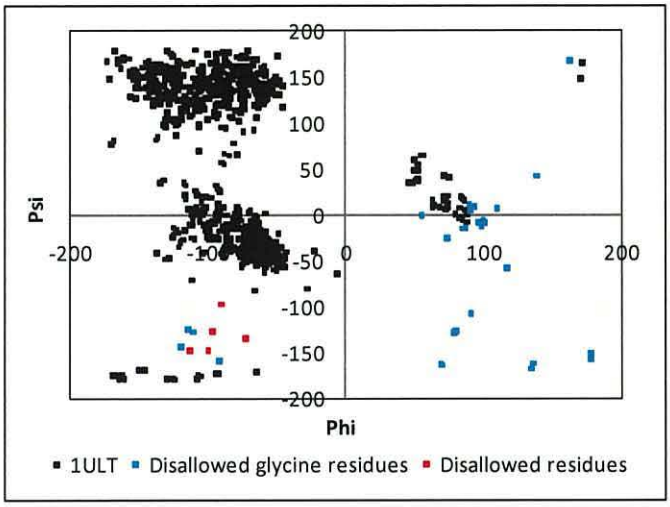


Figure 4.7: Ramachandran plot of the template structure of LC-FACS from *Thermus thermophilus* (Protein data bank file 1ULT). Normal residues are shown in black, problem glycine residues in blue and other problem residues identified by Swiss-pdbViewer (Guex and Peitsch, 1997) in red.

4.3.1.2 Evaluation of the crystallographic structure of ttLC-FACS, 1V25

The crystallographic structure of the bound ttLC-FACS (1V25) was analysed to identify residues with unfavourable backbone angles. The ramachandran plot revealed that 3.7 % of the total residues fell within the unfavourable region with 94.4 % of these being glycine residues (Figure 4.8). Residues shown in red (chain A: Trp^{473} and chain B: Ala^{541}) represent residues with unfavourable backbone angles and residues shown in blue represent unfavourable glycine residues (Figure 4.8).

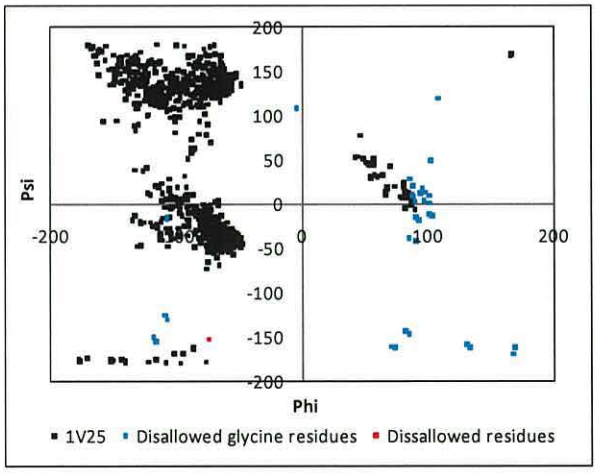


Figure 4.8: Ramachandran plot of the template structure of LC-FACS from *Thermus thermophilus* (Protein data bank file 1V25). Normal residues are shown in black, problem glycine residues in blue and other problem residues identified by Swiss-pdbViewer (Guex and Peitsch, 1997) in red.

4.3.1.3 Evaluation of the crystallographic structure of ttLC-FACS, 1V26

The bound crystallographic structure of ttLC-FACS (1V26) was analysed with Swiss-pdbViewer to identity unfavourable residues. Residues shown in red (chain A: Ser^{330} and chain B: Ser^{330}, Ala^{514}) represent residues with unfavourable backbone angles and residues shown in blue represent unfavourable glycine residues (Figure 4.9). The unfavourable region of the ramachandran plot contains 3.8 % of the residues, 86.5 % of these are glycine residues.

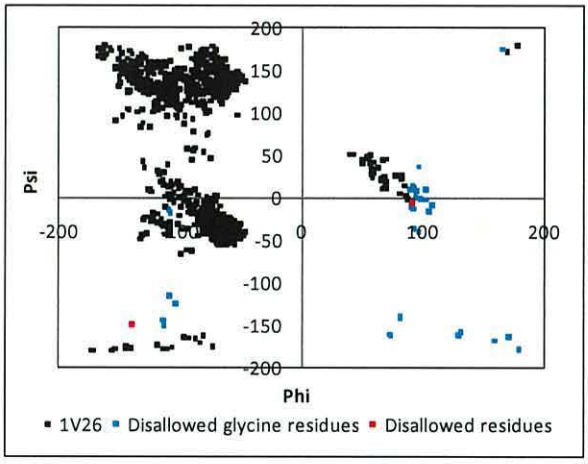


Figure 4.9: Ramachandran plot of the template structure of LC-FACS from *Thermus thermophilus* (Protein data bank file 1V26). Normal residues are shown in black, problem glycine residues in blue and other problem residues identified by Swiss-pdbViewer (Guex and Peitsch, 1997) in red.

Analysis of the protein data bank file 1ULT of ttLC-FACS (Hisanaga *et al.,* 2004) in the open conformation revealed that three residues from chain A and two residues from chain B fell within the disallowed region of the ramachandran plot. All three disallowed residues from chain A are considered important. Residue Thr^{187} forms part of the P-loop, His^{230} forms part of the G-motif and Glu^{328} forms part of the A-motif. Therefore, as a result of these being in the disallowed region, chain B of protein data bank file 1ULT was used as a template structure for homology modelling comparisons of AAE7/ACN1.

Other work has crystallised ttLC-FACS in the closed conformation in the presence of either phosphoaminophosphonic acid adenylate ester (1V25) or AMP and myristic acid (1V26) (Hisanaga *et al.,* 2004). The two crystallographic structures of ttLC-

FACS in the closed conformation are very similar with few disallowed residues. Therefore, the decision to use 1V26 as a template structure was based on the presence of AMP and myristic acid and the structural similarities (Figure 4.10) between the ligands of 1V26 and AMPP, used during experimental analysis in chapter 3 of this thesis. Chains A and B of 1V26 both contained very few disallowed residues. Therefore, chain B of 1V26 was also used as a template structure for homology modelling comparisons.

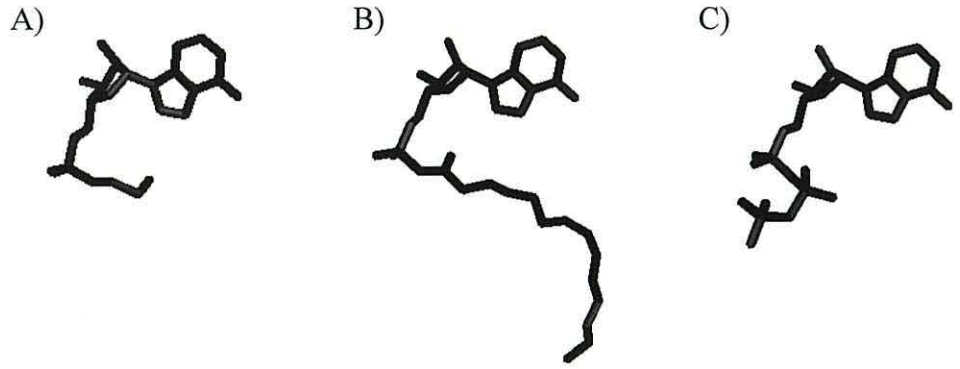


Figure 4.10: Ligand comparison. A) The structure of AMPP. B) The ligand structure (AMP and myristic acid) of 1V26.pdb (Hisanaga *et al.*, 2004). C) The ligand structure (ANP) of 1V25.pdb (Hisanaga *et al.*, 2004).

4.3.2 Homology modelling of AAE7/ACN1 with I-TASSER

The amino acid sequence of AAE7/ACN1 was submitted to the online modelling server I-TASSER (Zhang, 2008; 2009; Roy *et al.*, 2010). This automated server showed that the long chain fatty acyl-CoA synthetase, ttLC-FACS shares the greatest sequence identity (33 %) with AAE7/ACN1 but interestingly predicted the enzymes function as a short chain acyl-CoA synthetase which supports experimental enzyme assays (Shockey *et al.*, 2003; Turner *et al.*, 2005; chapter 3). Analysis of the ramachandran plot of the AAE7/ACN1-TASSER model showed that 53 residues (9.3 %) fell within the disallowed region (Figure 4.11 and Table 4.2), and of these 23 (43.4 %) were glycine or proline residues.

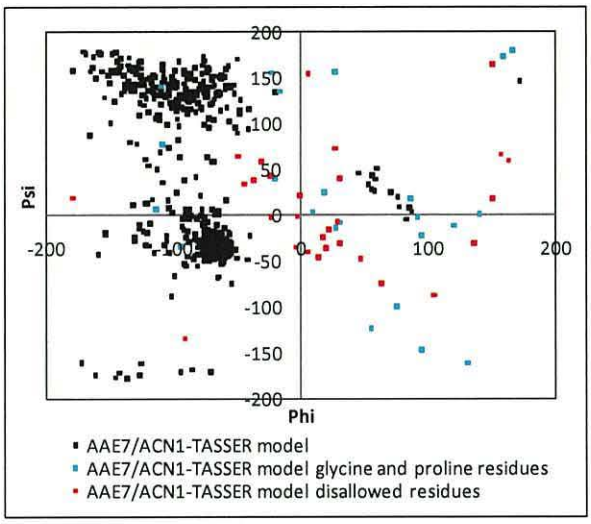


Figure 4.11: Ramachandran plot analysis of AAE7/ACN1-TASSER model. Residues that fall within the disallowed region of the ramachandran plot of AAE7/ACN1-TASSER model are represented in red and disallowed glycine and proline residues are present in blue.

The location of the disallowed residues ofAAE7/ACN1-TASSER model can be seen in Figure 4.12.

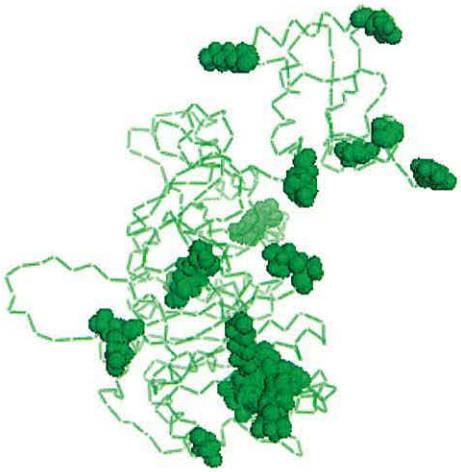


Figure 4.12: Disallowed residues of AAE7/ACN1-TASSER model are shown as spheres (minus glycine and proline residues). All other residues are shown as ribbon.

Table 4.2: All AAE7/ACN1-TASSER model residues that fall within the disallowed region of the ramachandran plot.

Disallowed residue in AAE7/ACN1-TASSER model	Aligned residues in ttLC-FACS
Pro^{13}	Pro^{10}
Pro^{24}	Leu^{20}
Ser^{68}	Gly^{68} in disallowed region
Gly^{96}	Gly^{96}
Asp^{126}	Asp^{126}
Gln^{127}	Pro^{127}
Gly^{145}	No alignment
Ser^{146}	No alignment
Phe^{148}	Val^{143}
Pro^{151}	Phe^{146}
Ile^{154}	Met^{149}
Val^{155}	No alignment
Ile^{156}	No alignment
Asp^{158}	No alignment
His^{159}	No alignment
Ser^{165}	No alignment
Asn^{167}	Asp^{150}
Gly^{184}	Glu^{167}
Pro^{192}	Val^{171}
Pro^{193}	Arg^{172}
Asp^{195}	Pro^{174}
Glu^{196}	Glu^{175}
Gly^{207}	Gly^{186} in disallowed region
Thr^{209}	Thr^{188}
Ala^{210}	Gly^{189} in disallowed region
Gly^{234}	Ala^{215}
Phe^{255}	Leu^{236}
Lys^{285}	Gly^{267}
Thr^{307}	Gly^{289}
Ile^{308}	His^{290}
Pro^{310}	No alignment
Gly^{321}	Gly^{302} in disallowed region
Gly^{336}	Gly^{317}
Thr^{348}	Thr^{329}
Pro^{351}	Val^{332}
Gly^{377}	Gly^{358} in disallowed region
Gly^{382}	Leu^{363}
Gly^{411}	Gly^{390} in disallowed region
Lys^{420}	Gly^{399}
Gly^{432}	Gly^{412} in disallowed region
Gly^{437}	Gly^{417}
Ser^{454}	Leu^{434}
Lys^{455}	Lys^{435}
Gly^{461}	Gly^{441}
Gly^{462}	Gly^{442}
Gln^{512}	No alignment
Lys^{526}	Gly^{501} in disallowed region
Gly^{539}	Ala^{514}
Gln^{550}	Leu^{526}
Ala^{558}	Tyr^{534}
Pro^{563}	Gly^{539}
Arg^{566}	No alignment
Arg^{568}	No alignment

4.3.3 Homology modelling of AAE7/ACN1 with MODELLER

The amino acid sequence of AAE7/ACN1 was aligned to ttLC-FACS using the default settings of MultAlin (Corpet, 1988) for sequence comparison and homology modelling (Figure 4.13). The resultant homology model of AAE7/ACN1 was based on the crystallographic structures of ttLC-FACS in both the unbound (1ULT) and bound (1V26) form.

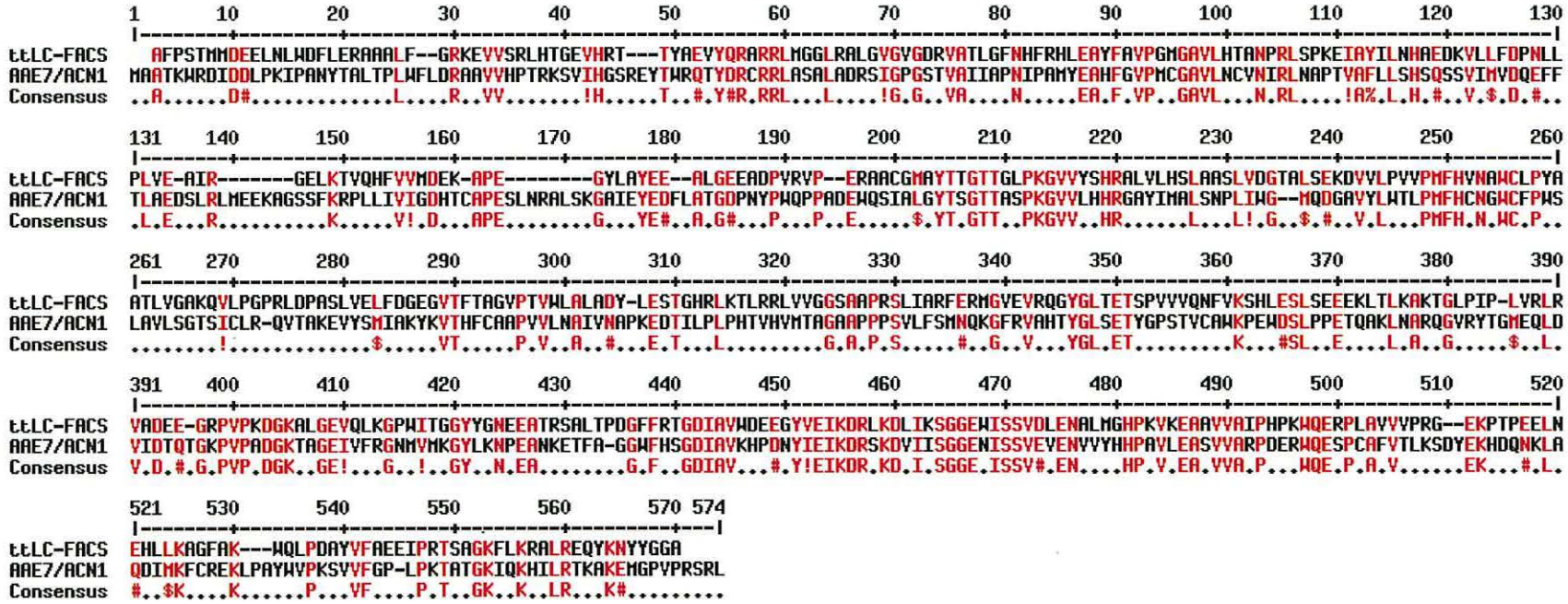


Figure 4.13: Sequence alignment of ACN1 and the dimeric long-chain fatty acyl-CoA synthetase (LC-FACS) from *Thermus thermophilus*. Sequences were aligned with the default settings of MultAlin (http://bioinfo.genopole-toulouse.prd.fr/multalin/multalin.html). Conserved residues are visible in the consensus region in red.

4.3.3.1 AAE7/ACN1 modelling to the template structure 1ULT in the open conformation

The resulting model of AAE7/ACN1 (AAE7/ACN1-A), based on the unbound ttLC-FACS (1ULT.pdb), was evaluated with the ramachandran plot function of Swiss-pdbViewer (Guex and Peitsch, 1997). The unfavourable region of the ramachandran plot contained 6.1 % of AAE7/ACN1-A residues (Figure 4.14; shown in red and blue) with 52.9 % of these being glycine or proline residues (Figure 4.14; shown in blue).

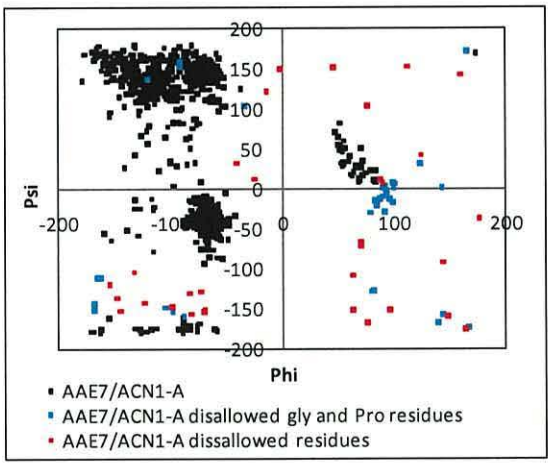


Figure 4.14: Ramachandran plot of the homology model AAE7/ACN1-A generated with Modeller 9v1 based on unbound ttLC-FACS (1ULT.pdb). Residues in blue and red represent residues that fall within the disallowed region of the ramachandran plot. Blue residues represent glycine and proline residues and red represents all other disallowed residues identified by Swiss-pdbViewer (Guex and Peitsch, 1997).

Energy minimisation was conducted with Swiss-pdbViewer (Guex and Peitsch, 1997) to improve the structure of AAE7/ACN1-A (to give AAE7/ACN1-B) and reduce the number of bonds within the disallowed region (Figure 4.15). Once this was conducted 5.3 % of the residues fell within the unfavourable region of the ramachandran plot (Figure 4.15; residues shown in blue and red) and of these 45.8 % were found to be glycine or proline residues (Figure 4.15; residues shown in blue).

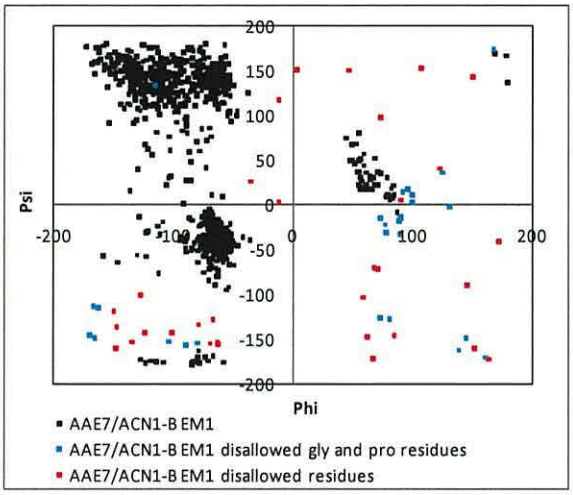


Figure 4.15: Comparative model AAE7/ACN1-B after one energy minimisation (EM1) process with the default settings of Swiss-pdbViewer (Guex and Peitsch, 1997). All disallowed residues are present in either red or blue. Blue represents disallowed glycine and proline residues.

Subsequently, a second energy minimisation process was conducted to further refine the model using the default settings of Swiss-pdbViewer (to give AAE7/ACN1-modelI). The resulting torsion angles were plotted on a ramachandran plot in figure 4.16. This second energy minimisation step resulted in 5.3 % of the residues within the unfavourable region (Figure 4.16; shown in red and blue). Of these 45.8 % were either glycine or proline residues (Figure 4.16; shown in blue).

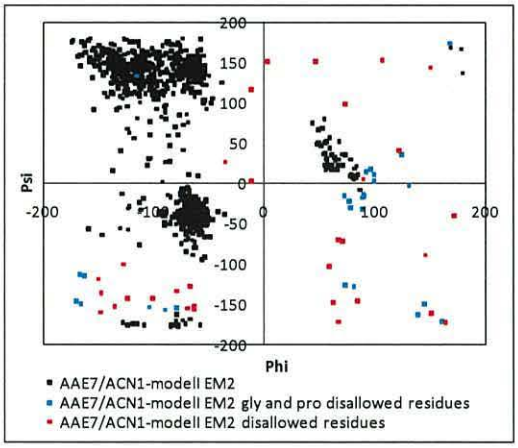


Figure 4.16: Ramachandran plot of the comparative model of AAE7/ACN1-modelI generated with Modeller 9v1 based on unbound ttLC-FACS (1ULT.pdb) crystallographic structure. The ramachandran plot was created with Swiss-pdbViewer 4.0 (Guex and Peitsch, 1997) after two energy minimisation processes (EM2) with default settings showing all residues with blue representing the disallowed glycine and proline residues and red representing all other disallowed residues.

The homology model constructed of AAE7/ACN1 is a dimeric structure comprised of two monomers. Each monomeric structure contains the same sequence of amino acid residues that fold to create a three dimensional protein structure. The amino acid chains that form the two monomeric structures are denoted chain A and chain B. Using the ramachandran plot of AAE7/ACN1-modelI (Figure 4.16) the two chains (A and B) were differentiated and the respective disallowed residues identified (Table 4.3 and 4.4).

Table 4.3: Disallowed residues (minus glycine and proline) of AAE7/ACN1-modelI, chain A, and the corresponding residues in the template structure ttLC-FACS (1ULT.pdb).

AAE7/ACN1-modelI chain A residue	ttLC-FACS (1ULT.pdb) residue
Leu^{28}	No alignment
Glu^{141}	No alignment
Ala^{194}	No alignment
Trp^{197}	Arg^{176}
Ala^{210}	Gly^{189} in disallowed region
His^{249}	His^{230} in disallowed region
Arg^{270}	Gly^{251} in disallowed region
Glu^{347}	Glu^{328} in disallowed region
Thr^{381}	Pro^{362}
Asn^{446}	Gly^{426} in disallowed region
Ser^{460}	Ser^{440}
Asp^{506}	No alignment
Glu^{508}	Glu^{486}
Cys^{523}	Gly^{501} in disallowed region
Leu^{527}	No alignment
Ile^{541}	Ile^{517}

Table 4.4: Disallowed residues (minus glycine and proline) of AAE7/ACN1-modelI, chain B, and the corresponding residues in the template structure ttLC-FACS (1ULT.pdb).

AAE7/ACN1-modelI chain B residue	ttLC-FACS (1ULT.pdb) residue
Asp^{580}	Glu^{17}
Glu^{616}	No alignment
Thr^{618}	Thr^{49}
Thr^{729}	Lys^{152}
Trp^{766}	Arg^{176}
Arg^{839}	Gly^{251} in disallowed region
Thr^{950}	Pro^{362}
Asn^{1015}	Gly^{426} in disallowed region
Lys^{1020}	Lys^{431} in disallowed region
Glu^{1060}	Arg^{476}
Arg^{1061}	No alignment
Gln^{1063}	No alignment
Ser^{1065}	No alignment
Leu^{1072}	Glu^{483}
Asp^{1087}	Ala^{500}
Met^{1089}	Ala^{503}

The location of the disallowed residues of AAE7/ACN1-modelI are shown in Figure 4.17. Chain A is represented in red with disallowed residues shown as spheres and chain B is shown in orange with disallowed residues also represented by spheres (Figure 4.17).

Figure 4.17: Ribbon structure of AAE7/ACN1-modelI. Disallowed residues (minus glycine and proline residues) in chain A are represented as red spheres and those in chain B are represented as orange spheres.

This final structure, known as AAE7/ACN1-modelI, was used for the subsequent structural comparisons described in section 4.4.

4.3.3.2 AAE7/ACN1 comparative modelling to the template structure 1V26 in a closed conformation

A model of AAE7/ACN1 (AAE7/ACN1-A1) was created with Modeller using the bound protein data bank file 1V26. Six point one percent of the residues from AAE7/ACN1-A1 fell within the disallowed region of the ramachandran plot (Figure 4.18; shown in red and blue), with 50.0 % of these being glycine or proline residues (Figure 4.18; shown in blue).

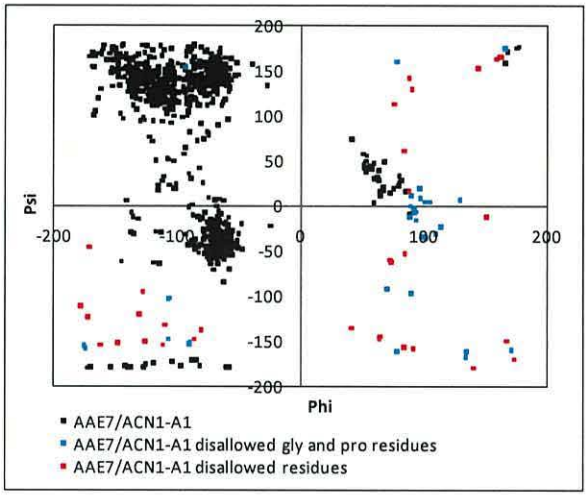


Figure 4.18: Ramachandran plot of the homology model AAE7/ACN1-A1 generated with Modeller 9v1 based on bound ttLC-FACS (1V26.pdb). Residues in blue and red represent residues that fall within the disallowed region of the ramachandran plot. Blue residues represent glycine and proline residues and red represents all other disallowed residues identified by Swiss-pdbViewer (Guex and Peitsch, 1997).

The comparative model (AAE7/ACN1-A1) was then subjected to energy minimisation by Swiss-pdbViewer, yielding AAE7/ACN1-B1. This gave a model with, 5.6 % of the residues within the disallowed region of the ramachandran plot (Figure 4.19; shown in red and blue), with 47.5 % of these being glycine or proline residues (Figure 4.19; shown in blue).

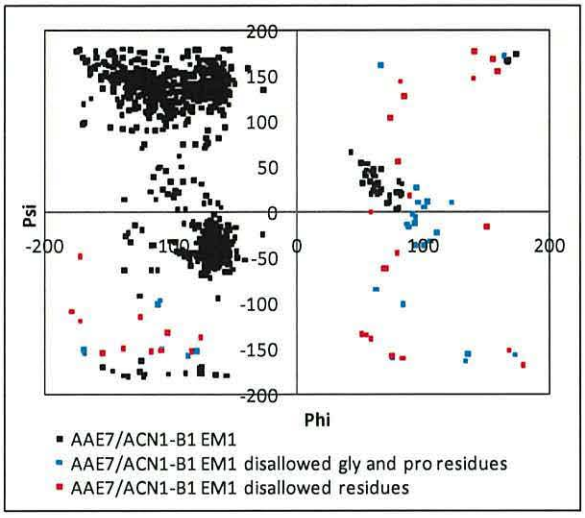


Figure 4.19: Comparative model AAE7/ACN1-B1 after one energy minimisation (EM1) process with the default settings of Swiss-pdbViewer (Guex and Peitsch, 1997). All disallowed residues are present in either red or blue. Blue represents disallowed glycine and proline residues.

The comparative model (AAE7/ACN1-B1) was subjected to a second energy minimisation step, resulting in AAE7/ACN1-modelII. Five point eight percent of the residues fell within the unfavourable region of the ramachandran plot (Figure 4.20; shown in red and blue), and of these 47.6 % were glycine or proline residues (Figure 4.20; shown in blue).

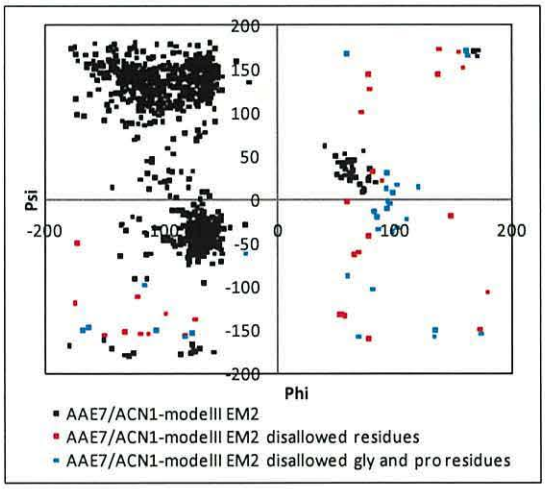


Figure 4.20: Ramachandran plot of the comparative model of AAE7/ACN1-modelII generated with Modeller 9v1 based on bound ttLC-FACS (1V26.pdb) crystallographic structure. The ramachandran plot was created with Swiss-pdbViewer (Guex and Peitsch, 1997) after two energy minimisation processes (EM2) with default settings showing all residues with blue representing the disallowed glycine and proline residues and red representing all other disallowed residues.

The amino acid chains that form the two monomeric structures of AAE7/ACN1-modelII are denoted chain A and chain B. Using the ramachandran plot of AAE7/ACN1-modelII (Figure 4.20) the two chains (A and B) were differentiated and the respective disallowed residues identified (Table 4.5 and 4.6).

Table 4.5: Disallowed residues (minus glycine and proline) in AAE7/ACN1-modelII, chain A, and the corresponding residues of the template structure ttLC-FACS (1V26.pdb).

AAE7/ACN1-modelII chain A residue	ttLC-FACS (1V26.pdb) residue
Trp^{26}	Phe^{31}
Phe^{27}	No alignment
Ser^{40}	Gly^{43} in disallowed region
Glu^{47}	No alignment
Ser^{68}	Gly^{68} in disallowed region
Ser^{136}	Ala^{135}
Glu^{141}	No alignment
Lys^{143}	No alignment
Lys^{172}	No alignment
Phe^{180}	No alignment
Ala^{210}	Gly^{189} in disallowed region
Arg^{270}	Gly^{251} in disallowed region
Ala^{430}	Thr^{409}
Val^{480}	Val^{460}
Ala^{499}	Ala^{479}
Phe^{538}	No alignment
Thr^{544}	Thr^{520}
Ala^{545}	Ser^{521}

Table 4.6: Disallowed residues (minus glycine and proline) of AAE7/ACN1-modelII, chain B, and the corresponding residues in the template structure ttLC-FACS (1V26.pdb).

AAE7/ACN1-modelII chain B residue	ttLC-FACS (1V26.pdb) residue
Leu^{597}	No alignment
Arg^{707}	Arg^{137}
Lys^{712}	No alignment
Ser^{734}	No alignment
Leu^{735}	No alignment
Asp^{764}	No alignment
Ala^{779}	Gly^{189} in disallowed region
Arg^{839}	Gly^{251} in disallowed region
Thr^{950}	Pro^{362}
Lys^{964}	Arg^{374}
Asn^{1015}	Gly^{426} in disallowed region
Glu^{1051}	Glu^{462}
Ser^{1053}	Ala^{464}
Val^{1070}	Val^{481}

The location of the disallowed residues of AAE7/ACN1-modelII are shown in Figure 4.21. Chain A is represented in red with disallowed residues shown as spheres and chain B is shown in orange with disallowed residues shown as spheres (Figure 4.21).

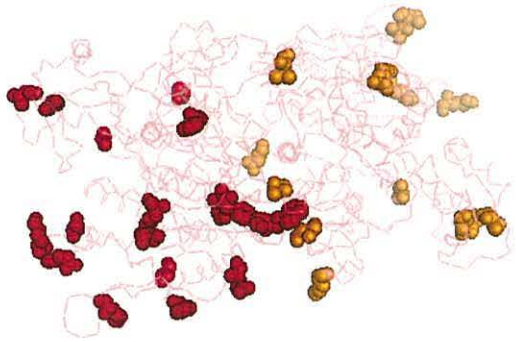


Figure 4.21: Ribbon structure of AAE7/ACN1-modelII. Disallowed residues in chain A are represented as red spheres and those in chain B are represented as orange spheres.

Thus, by producing both bound and unbound models, AAE7/ACN1-modelII and AAE7/ACN1-modelI respectively, these have been used in the following structural comparisons. Chain B of each model was chosen for structural analysis based on the number of disallowed residues and the role these residues play in catalysis of ttLC-FACS. The issues surrounding this are discussed further in section 4.3.1.

4.4 Structural comparison of AAE7/ACN1-modelI, AAE7/ACN1-modelII and AAE7/ACN1-TASSER model to acyl-CoA synthetase

Potentially important acyl-CoA synthetase residues and motifs were compared between the three AAE7/ACN1 models created within this chapter and to known acyl-CoA synthetase crystallographic structures. Sequence and structural comparison aims to further understand how AAE7/ACN1 catalyses the conversion of acetate to acetyl-CoA.

4.4.1 Adenosine binding region comparison

There are a number of important motifs and residues conserved in acyl-CoA synthetases that are considered essential for the positioning and binding of adenosine 5' monophosphate (Gulick *et al.,* 2003; Hisanaga *et al.,* 2004). Sequence and structural comparisons have been conducted to determine the level of conservation between the AAE7/ACN1-models and both the open and closed crystallographic structures of ttLC-FACS.

4.4.1.1 Comparison of the ATP/AMP signature motif

AAE7/ACN1 was identified as an adenylate forming enzyme, based on the presence of an AMP signature motif (Shockey *et al.,* 2003). This motif is common to all acyl-CoA synthetases' and contains a sequence that is believed to specify AMP binding (Babbitt *et al.,* 1992; Black *et al.,* 1997). The generic residues that make up this motif are [LIVMFY]-X-X-[STG]-[STAG]-G-[ST]-[STEI]-[SG]-X-[PASLIVM]-[KR] (Shockey *et al.,* 2003). The corresponding residues within ttLC-FACS and AAE7/ACN1 are ^{181}MAYTTGTTGLPK^{192} and ^{202}LGYTSGTTASPK^{213} respectively (Figure 4.22), denoted the P-loop by Hisanaga *et al.* (2004) (section 4.4.1.3).

Other reports have suggested that the ATP/AMP signature motif is comprised of two separated sequences, YTSGTTGxPKGV and GYGxTE (Weimar *et al.,* 2002). The first sequence is similar to the signature motif proposed by Shockey *et al.* (2003) but displaced by two residues, YTTGTTGLPKGV and YTSGTTASPKGV in ttLC-FACS and AAE7/ACN1 respectively. The second sequence corresponds to ^{323}GYGLTE^{328} and ^{342}TYGLSE^{347} in ttLC-FACS and AAE7/ACN1 respectively (Figure 4.22), denoted the A-motif by Hisanaga *et al.* (2004) (section 4.4.1.2).

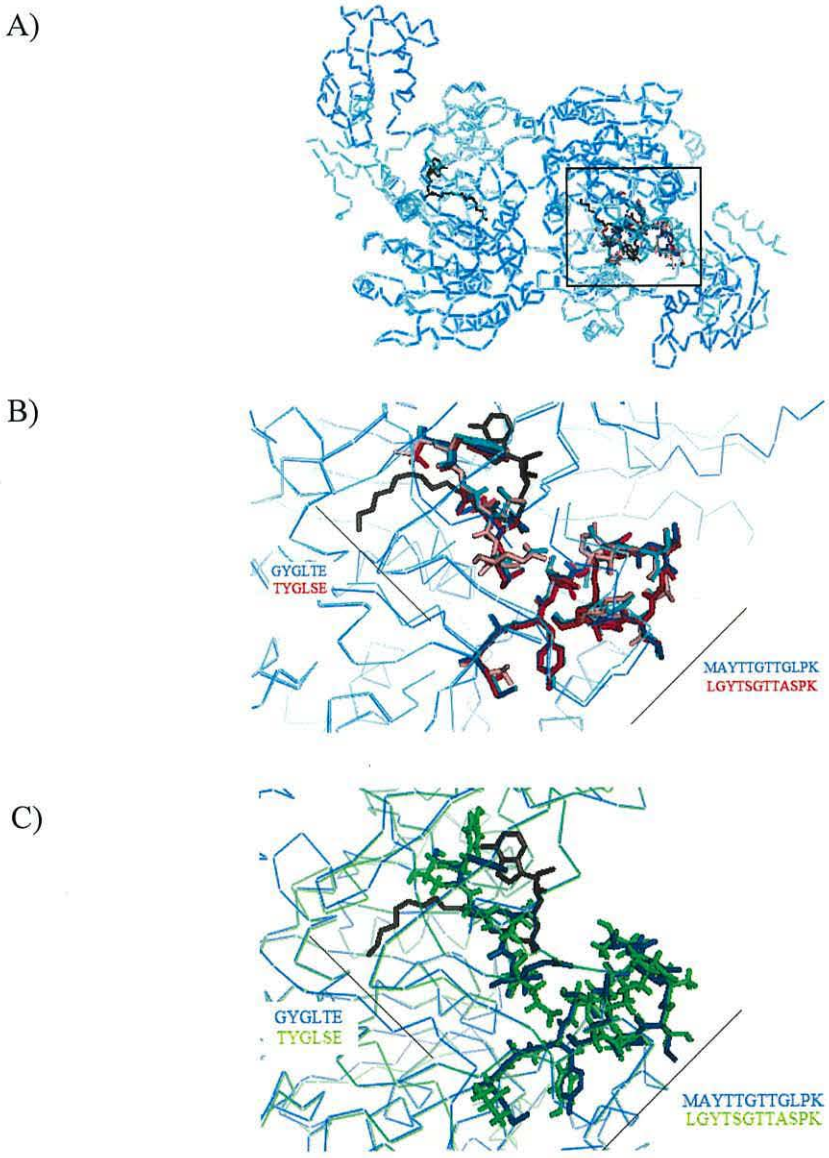


Figure 4.22: Comparison of the ATP/AMP signature motif between AAE7/ACN1-modelI, AAE7/ACN1-modelII, AAE7/ACN1-TASSER model and ttLC-FACS (1ULT.pdb and 1V26.pdb). The ttLC-FACS crystal structure 1ULT is shown in dark blue, 1V26 is shown in light blue, AAE7/ACN1-modelI is shown in red, AAE7/ACN1-modelII is shown in pink and AAE7/ACN1-TASSER model is shown in green. A) Shows the location of the AMP binding site (present within the black box) within the crystallographic structures of ttLC-FACS. B) An enlarged image of the AMP signature motif. C) An enlarged image of the AMP signature motif with AAE7/ACN1-TASSER model is shown in green and the corresponding residues of ttLC-FACS in blue.

4.4.1.2 Comparison of the adenosine (A) motif

The A-motif of ttLC-FACS, ^{323}GYGLTE^{328}, is believed to be critical for adenine binding (Hisanaga *et al.,* 2004), containing important residues that are conserved in AAE7/ACN1, ^{342}TYGLSE^{347} (Figure 4.24). The A-motif contains essential residues that interact with myristoyl-AMP of ttLC-FACS (Hisanaga *et al.,* 2004). Whilst this motif is considered important for the positioning of the substrate, two motif residues (Gly^{323} and Thr^{327}of ttLC-FACS) are not conserved in AAE7/ACN1 (Figure 4.23). Gly^{323}, of ttLC-FACS, aligns with Thr^{342} in AAE7/ACN1 and Thr^{327}, of ttLC-FACS, aligns with Ser^{346} in AAE7/ACN1.

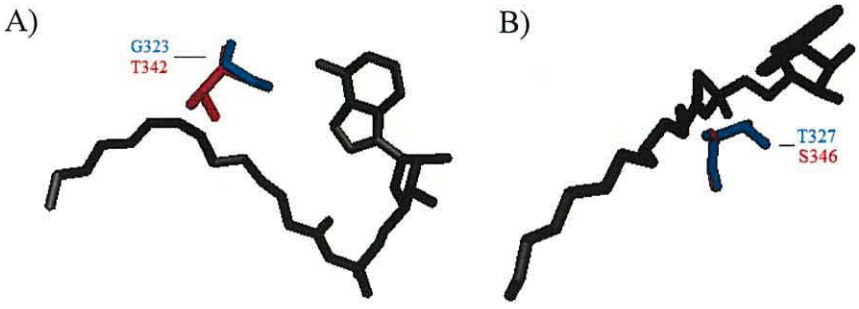


Figure 4.23: Residues that are not conserved within the A-motif of AAE7/ACN1-modelI and ttLC-FACS. A) Shows the location of Gly^{323} of ttLC-FACS (shown in blue) and the corresponding residue, Thr^{342} of AAE7/ACN1-modelI (shown in red). B) Shows the location of Thr^{327} of ttLC-FACS (shown in blue) and the corresponding residue, Ser^{346} of AAE7/ACN1-modelI (shown in red).

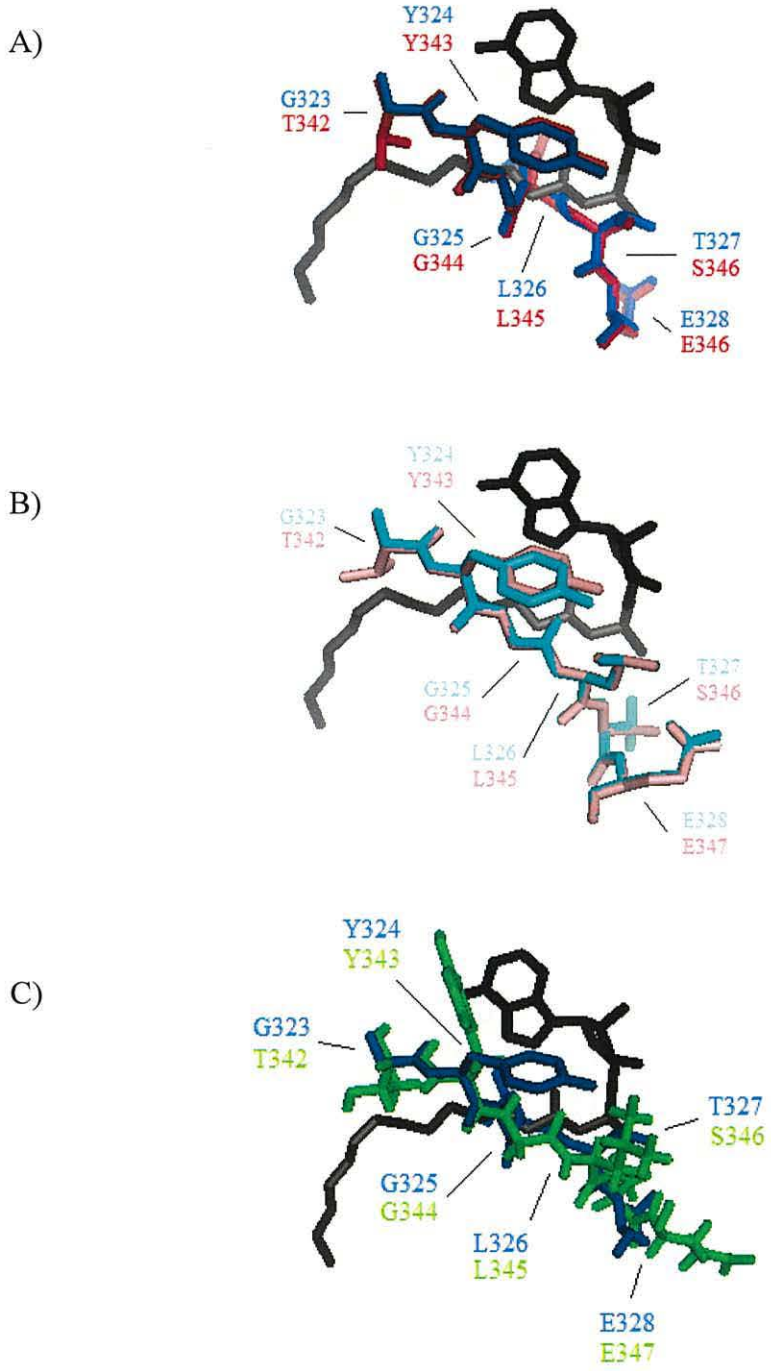


Figure 4.24: The A-motif. A) Open conformation of ttLC-FACS (blue) compared with AAE7/ACN1-modelI (red). B) Closed conformation of ttLC-FACS (light blue) aligned with AAE7/ACN1-modelII (pink). C) AAE7/ACN1-TASSER model (green) aligned with ttLC-FACS (blue).

4.4.1.3 Comparison of the phosphate (P) loop

Tyr^{184} is an important residue in ttLC-FACS forming part of the P-loop involved in phosphate binding (Hisanaga *et al.,* 2004). This loop also forms part of the

ATP/AMP signature motif reported by Weimar *et al.* (2002) and Shockey *et al.* (2003) (section 4.4.1.1). The P-loop of ttLC-FACS consists of ^{184}TTGTTGLPK^{192}, which aligns to ^{205}TSGTTASPK^{213} in AAE7/ACN1 (Figure 4.25). Tyr^{184}, which aligns to Tyr^{205} in AAE7/ACN1, is important in ttLC-FACS as the hydroxyl group interacts with magnesium during catalysis (Hisanaga *et al.*, 2004).

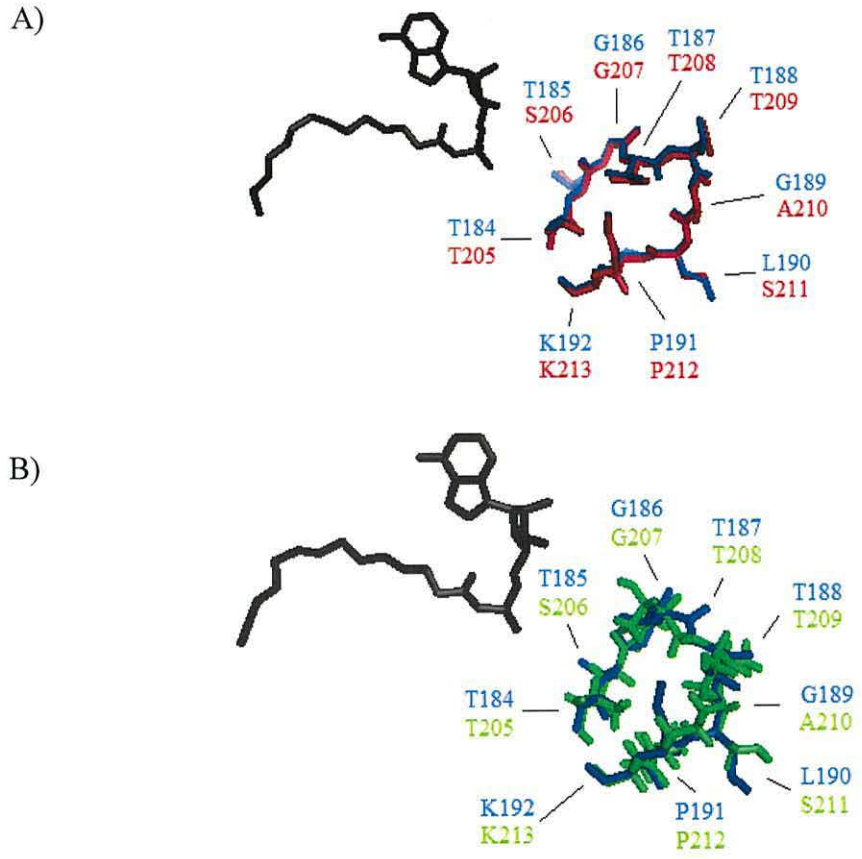


Figure 4.25: The P-loop of ttLC-FACS (shown in blue) and in A) the corresponding residues in AAE7/ACN1-modelI (shown in red). B) The corresponding residues in AAE7/ACN1-TASSER model (shown in green).

4.4.2 Comparison of the linker (L) motif

The L-motif of ttLC-FACS is also known as the linker region, consisting of residues ^{432}DRLKDL^{437}. The corresponding residues in AAE7/ACN1 are ^{452}DRSKDV^{457} (Figure 4.26). Whilst the L-motif contains the amino acid that acts as a linker, Asp^{452}, between the N- and C-terminal domains of ttLC-FACS, the motif is also involved in substrate binding. Arg^{433} and Lys^{435} of ttLC-FACS both play a role during catalysis, forming hydrogen bonds with the AMP moiety (Hisanaga *et al.*,

2004). These potentially important residues are both conserved in AAE7/ACN1 (Figure 4.26).

A)

B)

C)

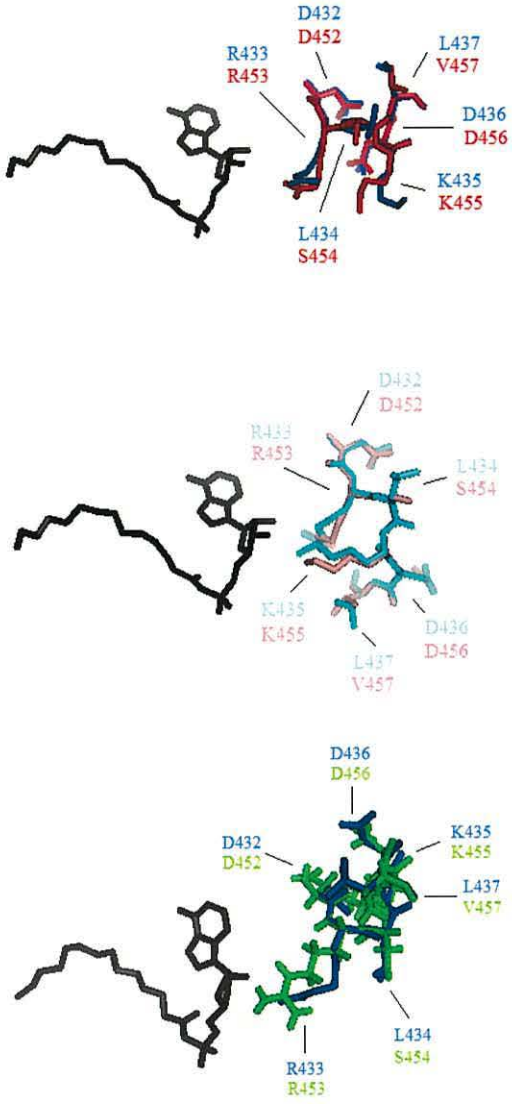


Figure 4.26: The L-motif in A) the open conformation of ttLC-FACS (1ULT) aligned with AAE7/ACN1-modelI and B) the closed conformation of ttLC-FACS (1V26) aligned with AAE7/ACN1-modelII. C) AAE7/ACN1-TASSER model aligned with ttLC-FACS (1ULT).

4.4.3 Conformational change of ttLC-FACS during catalysis

The crystallographic structure of ttLC-FACS has been solved in both an open and closed conformation (Hisanaga *et al.*, 2004). The open structure (1ULT) was crystallised in the absence of ligands and the closed structure (1V26) was crystallised in the presence of myristoyl-AMP. Hisanaga *et al.* (2004) propose that the binding of ATP triggers both a conformational change in shape and opening of the fatty acid binding tunnel, by movement of the gate residue within the G-motif (section 4.4.4.2).

The closed structure of ttLC-FACS is stabilised by three residues, Glu^{443}, Glu^{475} and Lys^{527} in the C-terminal domain (Hisanaga *et al.*, 2004) (Figure 4.27). The carboxyl oxygen of Glu^{443} forms a salt bridge with the side chain of Arg^{105} and a hydrogen bond with the side chain of His^{230}. The formation of a hydrogen bond between these two residues (Glu^{443} and His^{230}) plays an important role in the movement of the G-motif (section 4.4.4.2). During catalysis the side chains of Glu^{475} and Lys^{527} form salt bridges with the side chains of Arg^{433} and Asp^{436} in the L-motif (section 4.4.2).

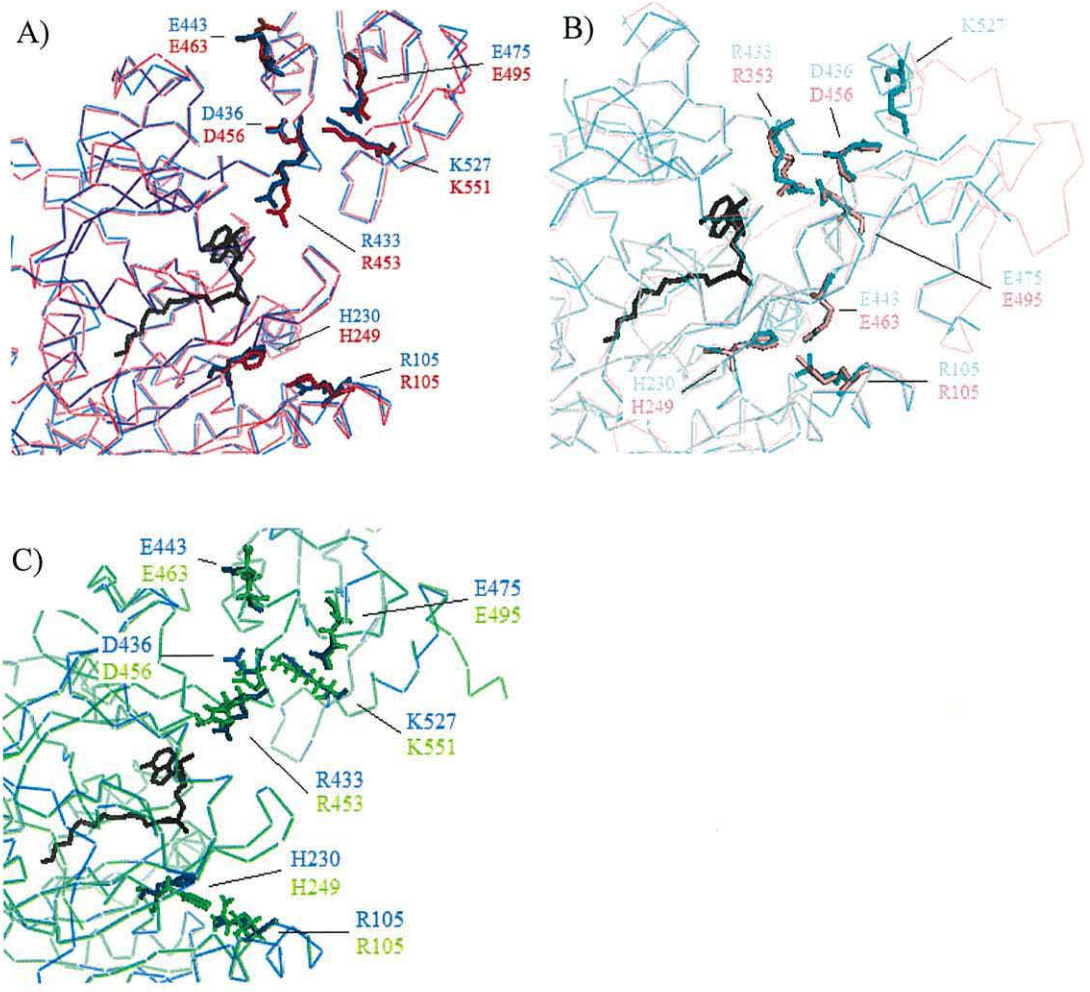


Figure 4.27: Residues involved in maintaining the closed structure of ttLC-FACS (Hisanaga *et al.*, 2004). A) Shows the open conformations of ttLC-FACS in blue (1ULT.pdb; chain A) and the corresponding residues of AAE7/ACN1-modelI in red. B) Shows the closed conformation of ttLC-FACS in light blue (1V26.pdb; chain B) with the corresponding residues of AAE7/ACN1-modelII in pink. C) Shows AAE7/ACN1-TASSER model and the corresponding residues of ttLC-FACS (1ULT.pdb; chain A).

4.4.4 Comparison of the fatty acid binding region

Residues believed to be involved in the entrance, movement and positioning of the fatty acid of ttLC-FACS (Hisanaga *et al.*, 2004) were aligned with the AAE7/ACN1 models for structural comparison.

4.4.4.1 Fatty acid entrance site

The fatty acid entrance site of ttLC-FACS is comprised of positivity charged, hydrophilic residues (Lys^{219}, Arg^{296}, Arg^{297}, Arg^{321}, Lys^{350} and Lys^{354}) that attract the negativity charged carboxyl group of the fatty acid (Figure 4.28). These residues are not conserved in AAE7/ACN1 and correspond to Gly^{238}, Val^{315}, His^{316}, Ala^{340}, Gln^{369} and Asn^{373} respectively (Table 4.7).

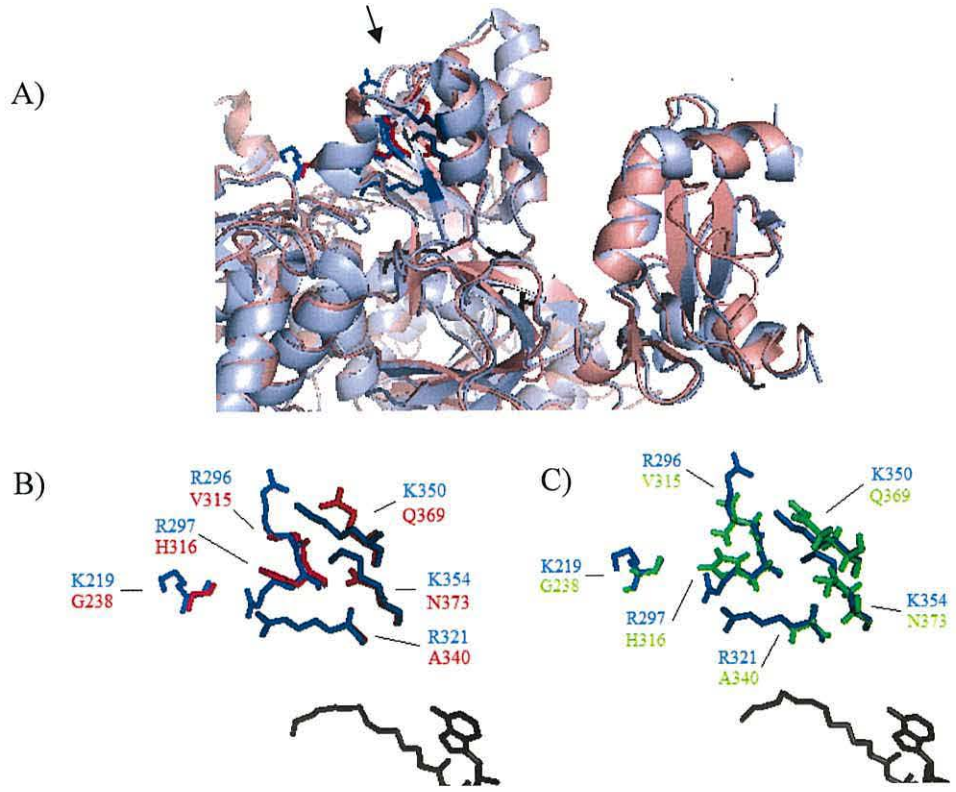


Figure 4.28: The proposed fatty acid entrance site of ttLC-FACS (Hisanaga *et al.*, 2004). A) The location of the fatty acid entrance site is indicated by an arrow on the ribbon structure of ttLC-FACS (shown in light and dark blue) aligned with the ribbon structure of AAE7/ACN1-modelI (shown in red and pink). B) Residues that are believed to be important (Hisanaga *et al.*, 2004) for fatty acid entrance of ttLC-FACS (shown in blue) and the corresponding residues of AAE7/ACN1-modelI (shown in red). C) ttLC-FACS fatty acid entrance site residues (shown in blue) are aligned with the corresponding residues of AAE7/ACN1-TASSER model (shown in green).

Table 4.7: Residues at the fatty acid entrance site of ttLC-FACS are made up of positively charged hydrophilic residues which attract the negatively charged carboxyl group of the fatty acid. These residues are not conserved in AAE7/ACN1.

AAE7/ACN1	Residues not conserved	ttLC-FACS
Gly^{238}	*	Lys^{219}
Val^{315}	*	Arg^{296}
His^{316}	*	Arg^{297}
Ala^{340}	*	Arg^{321}
Gln^{369}	*	Lys^{350}
Asn^{373}	*	Lys^{354}

4.4.4.2 The gate (G) motif

Trp^{234} is an important gate residue in ttLC-FACS that moves during catalysis, allowing the fatty acid to enter via a fatty acid tunnel (section 4.4.4.1). This gate residue forms part of the G-motif, ^{226}VPMFHVNAW^{234}, which aligns to ^{245}LPMFHCNGW^{253} in AAE7/ACN1 (Figure 4.29).

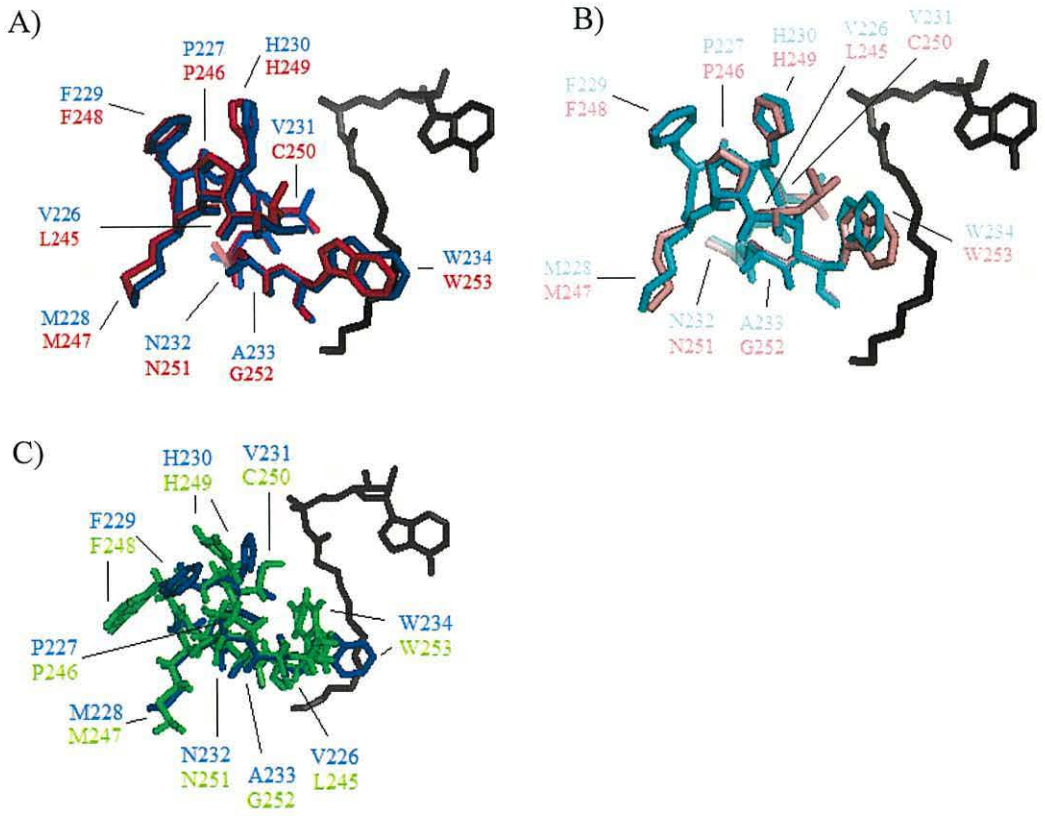


Figure 4.29: The G-motif of ttLC-FACS. A) Shows the open conformation of ttLC-FACS (1ULt) in blue and AAE7/ACN1-modelI in red. B) Shows the closed conformation of ttLC-FACS (1V26) in light blue and AAE7/ACN1-modelII in pink. C) Shows AAE7/ACN1-TASSER model in green and the corresponding residues of ttLC-FACS (1ULT) in blue.

4.4.4.3 The fatty acid binding region predicted by I-TASSER

In addition to the production of AAE7/ACN1-TASSER model the I-TASSER server (Zhang, 2008) also predicted potential binding site residues (Table 4.8 and Figure 4.30). The predicted binding site comprises of twenty three residues, of these only six are conserved between AAE7/ACN1 and ttLC-FACS. Three of these conserved residues (Tyr^{343}, Gly^{344} and Leu^{345} of AAE7/ACN1) are present in the ATP/AMP signature motif, also called the A-motif, involved in adenosine binding (section 4.4.1.1 and 4.4.1.2). A fourth residue (Gly^{321} of AAE7/ACN1) aligns with Gly^{302} of ttLC-FACS that is believed to interact with the oxygen of the sugar ribose and a nitrogen atom of the nucleobase adenosine of AMP (Hisanaga *et al.*, 2004).

The AAE7/ACN1 residue Trp^{253}, which aligns to the gate residue (section 4.4.4.2) of the ttLC-FACS G-motif, was identified by I-TASSER as a potential binding site residue. The gate residue of ttLC-FACS moves during catalysis opening the fatty acid binding tunnel, allowing entry into the active site. This residue is present in the upper section of the C4 to C9 region of the myristoyl moiety and therefore could form part of the base of the fatty acid binding pocket in AAE7/ACN1.

The AAE7/ACN1 residue Cys^{254} aligns to a cysteine residue of ttLC-FACS that surrounds the myristoyl moiety between C9 and C12 (section 4.4.4.5.3). This is the only residue conserved within this section of the myristoyl moiety. The importance of this conserved residue is currently unclear.

Table 4.8: Potential binding site residues predicted by I-TASSER during the creation of AAE7/ACN1-TASSER model.

AAE7/ACN1-TASSER model residues	Residues not conserved	Corresponding residues in ttLC-FACS
Ala^{226}	*	Ser^{205}
Asn^{229}	*	Ala^{208}
Pro^{230}	*	Ser^{209}
Trp^{233}	*	Asp^{212}
Met^{235}	*	Leu^{216}
Leu^{245}	*	Val^{226}
Cys^{250}	*	Val^{231}
Trp^{253}		Trp^{234}
Cys^{254}		Cys^{235}
Phe^{255}	*	Leu^{236}
Ser^{258}	*	Ala^{239}
Cys^{290}	*	Ala^{272}
Met^{318}	*	Val^{299}
Ala^{320}	*	Gly^{301}
Gly^{321}		Gly^{302}
Thr^{342}	*	Gly^{323}
Tyr^{343}		Tyr^{324}
Gly^{344}		Gly^{325}
Leu^{345}		Leu^{326}
Ser^{346}	*	Thr^{327}
Gly^{350}	*	Pro^{331}
Pro^{351}	*	Val^{332}
Val^{354}	*	Gln^{335}

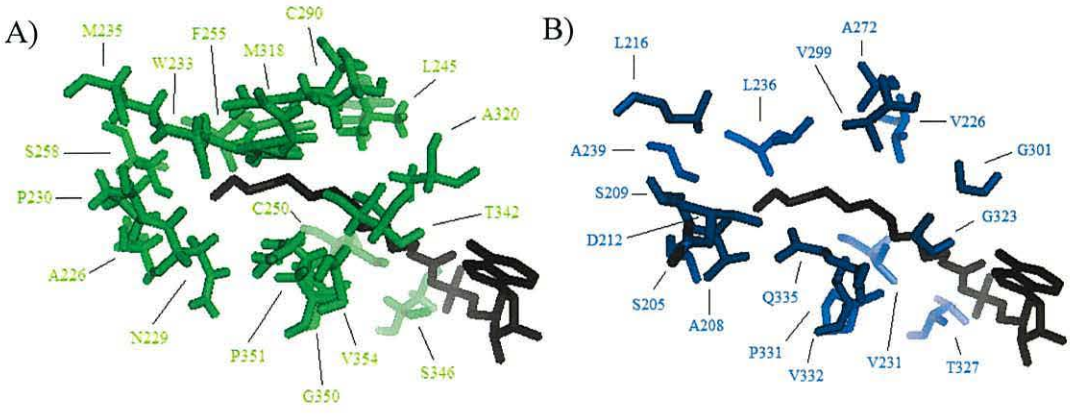
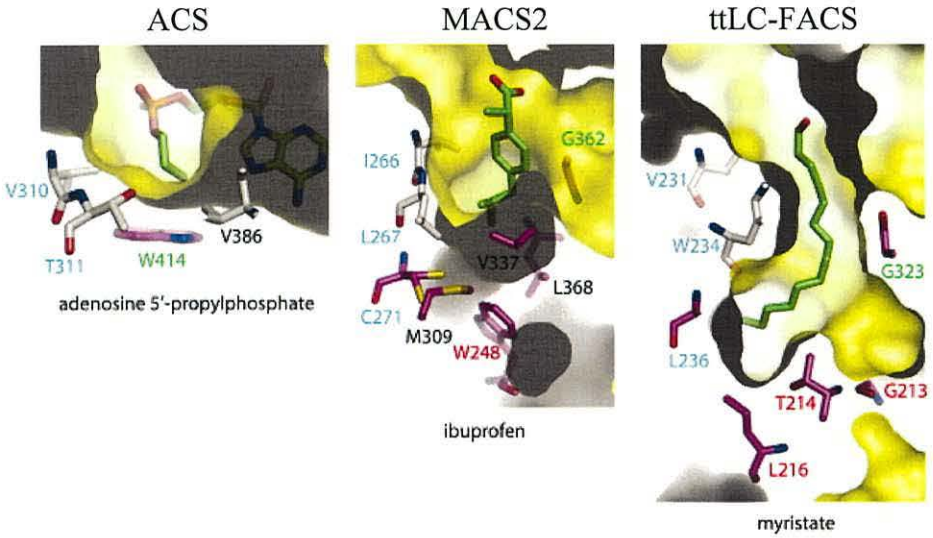


Figure 4.30: Binding site residues predicted by the online modelling server I-TASSER that are not conserved between AAE7/ACN1 and ttLC-FACS. A) AAE7/ACN1-TASSER model shown in green and myristoyl AMP ligand of ttLC-FACS shown in black. B) The corresponding residues of ttLC-FACS (1ULT) are shown in blue and the myristoyl AMP ligand of ttLC-FACS in black.

4.4.4.4 Comparison of the fatty acid binding site of acyl-CoA synthetase

The fatty acid binding region of acyl-CoA synthetase has been recently compared (Figure 4.31) between seACS, MACS2 and ttLC-FACS (Kochan *et al.*, 2009). Left wall and base residues of the active site of ACS, MACS2 and ttLC-FACS (Kochan *et al.* 2009) were compared to the active site of AAE7/ACN1 in both the open and closed conformation (Figure 4.32).

A)



B)

	Left pocket wall	seACS pocket base	MACS2 pocket base
ACS	^{308}GW**V**T**G**HTYV	^{411}DTW**W**QTE	
MACS2	^{264}GW**IL**NILCS	^{359}ESY**G**QTE	^{248}**W**TGL
ttLC-FACS	^{229}FH**V**-NAWCL	^{320}QGY**G**LTE	^{213}**G**TA**L**
AAE7/ACN1	^{248}FH**C**-NGWCF	^{341}HTY**G**LSE	^{234}**G**—**M**

Figure 4.31: A selection of residues that comprise the fatty acid binding pocket wall and base of acyl-CoA synthetase. A) Image taken from the structural comparison of acyl-CoA synthetase conducted by Kochan *et al.* 2009. B) Sequence comparison of residues that Kochan *et al.* 2009 state as involved in the left wall and base of the binding pocket of seACS, MACS2 and ttLC-FACS.

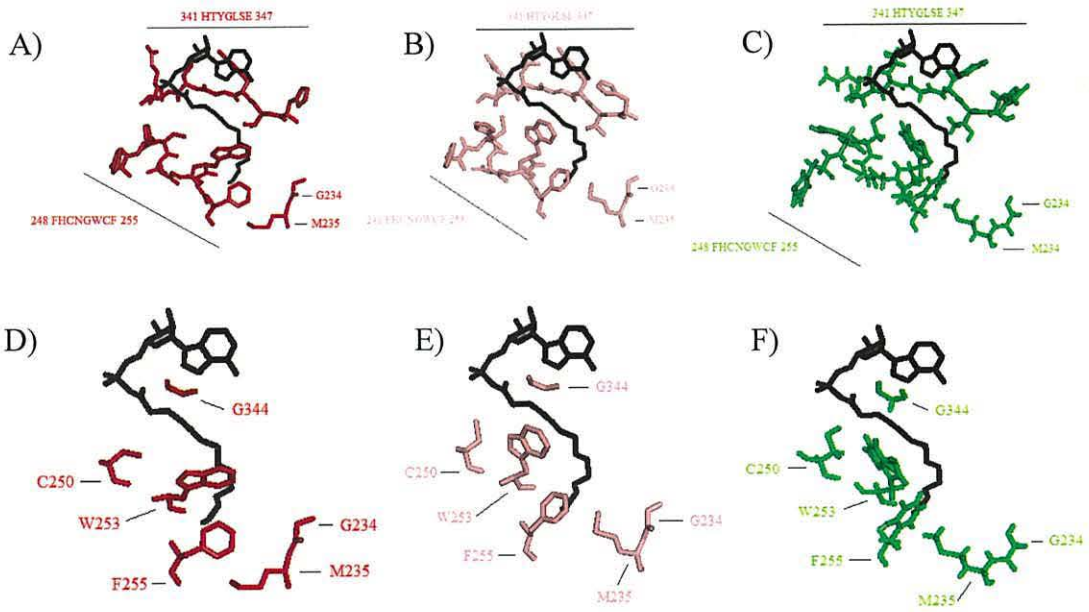


Figure 4.32: A selection of residues that comprise the fatty acid binding pocket wall and base of acyl-CoA synthetase. A) AAE7/ACN1-modelI residues that align to the left wall and base residues noted by Kochan *et al.* 2009. B) AAE7/ACN1-modelII residues that align to the left wall and base residues noted by Kochan *et al.* 2009. C) AAE7/ACN1-TASSER model residues that align to the left wall and base residues noted by Kochan *et al.* 2009. D) AAE7/ACN1-modelI residues within the left wall and base region that are not conserved between AAE7/ACN1 and ttLC-FACS. E) AAE7/ACN1-modelII residues within the left wall and base region that are not conserved between AAE7/ACN1 and ttLC-FACS. F) AAE7/ACN1-TASSER model residues within the left wall and base region that are not conserved between AAE7/ACN1 and ttLC-FACS.

4.4.4.5 ttLC-FACS and AAE7/ACN1 fatty acid binding region comparison

The crystallographic structure of ttLC-FACS and the AAE7/ACN1-models were aligned and residues surrounding the ligand of ttLC-FACS, myristoyl adenosine monophosphate (Figure 4.33), were compared.

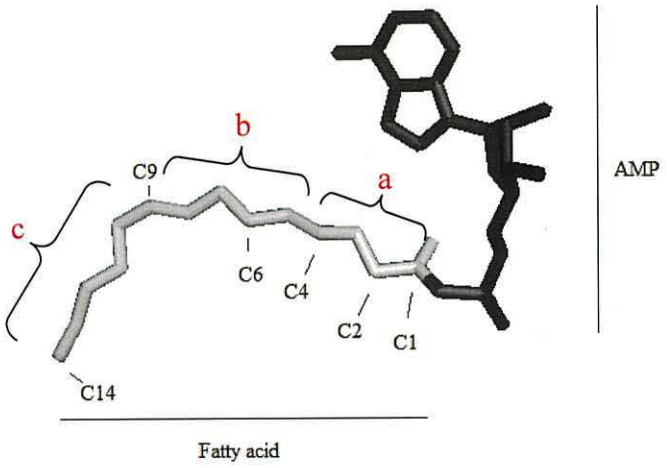


Figure 4.33: Stick model created with Pymol of the ligand myristoyl adenosine monophosphate from ttLC-FACS. Adenosine monophosphate ($C_{10}H_{14}N_5O_7P$) region is shown in black and the myristic acid ($C_{14}H_{28}O_2$) region is shown in grey. Carbons one to four are designated as section a, carbons four to nine are designated as section b and carbons nine to fourteen as designated as section c. Hydrogen atoms have been omitted for clarity.

The fatty acid binding region of ttLC-FACS has been divided into three sections (Figure 4.33). The first section contains carbons one to four, the second contains carbons four to nine and the third contains carbons nine to fourteen of the myristoyl adenosine 5' monophosphate ligand of ttLC-FACS. A summary of the residues located within each section that Hisanaga *et al.* (2004) describe as within 4.5 Å of the myristoyl moiety and the corresponding aligning residues in AAE7/ACN1 are shown in Table 4.9.

Table 4.9: Amino acid residues within a 4.5 Å of myristoyl adenosine monophosphate ligand of ttLC-FACS along with the corresponding residues in AAE7/ACN.

Carbon number	ttLC-FACS	Residues not conserved	AAE7/ACN1
C1-C4	Val^{231}	*	Cys^{250}
	Gly^{301}	*	Ala^{320}
	Tyr^{324}		Tyr^{343}
	Gly^{325}		Gly^{344}
	Leu^{326}		Leu^{345}
	Thr^{327}	*	Ser^{346}
	Pro^{331}	*	Gly^{350}
C4-C9	Trp^{234}		Trp^{253}
	Val^{299}	*	Met^{318}
	Gly^{323}	*	Thr^{342}
	Val^{332}	*	Pro^{351}
C9-C14	His^{204}	*	Met^{225}
	Ala^{208}	*	Asn^{229}
	Ser^{209}	*	Pro^{230}
	Thr^{214}	*	-
	Cys^{235}		Cys^{254}
	Leu^{236}	*	Phe^{255}
	Ala^{239}	*	Ser^{258}
	Gln^{335}	*	Val^{354}

4.4.4.5.1 Section a

The first section surrounding carbons one to four of the ligand of ttLC-FACS is the potential acetate binding site within AAE7/ACN1. It is therefore surprising that within this region only three out of the seven residues are conserved between AAE7/ACN1 and ttLC-FACS (Figure 4.34).

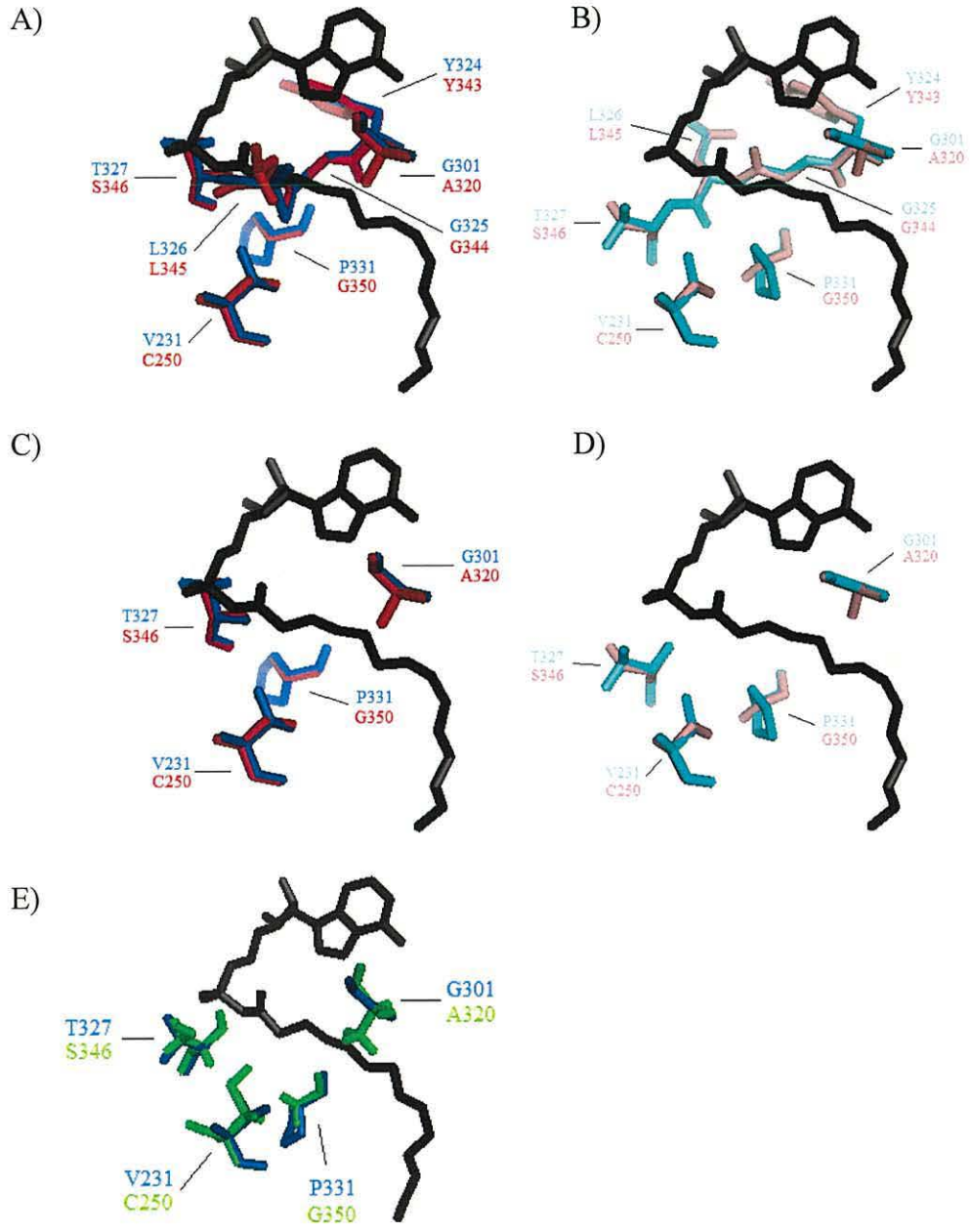


Figure 4.34: Section a. Residues surrounding carbon 1 to carbon 4 region of myristoyl adenosine monophosphate of ttLC-FACS (Hisanaga *et al.*, 2004). A) ttLC-FACS (1ULT; open conformation) residues are represented in blue and the corresponding residues of AAE7/ACN1-modelI are shown in red. B) ttLC-FACS (1V26; closed conformation) residues are represented in light blue and the corresponding residues of AAE7/ACN1-modelII are shown in pink. C) Residues that are not conserved between ttLC-FACS (1ULT; shown in blue) and AAE7/ACN1-modelI (shown in red). D) Residues that are not conserved between ttLC-FACS (1V26; shown in light blue) and AAE7/ACN1-modelII (shown in pink). E) Residues that are not conserved between ttLC-FACS (shown in blue) and AAE7/ACN1-TASSER model (shown in green).

4.4.4.5.2 Section b

Within section b of the myristoyl moiety only one residue is conserved between AAE7/ACN1 and ttLC-FACS. This conserved residue is the gate residue of the G-motif (section 4.35), tryptophan. Visual comparison of the AAE7/ACN1 models and ttLC-FACS in both the open and closed conformations shows the proximity of Trp^{234}, Val^{299}, Val^{332} and Gly^{323} of ttLC-FACS and Trp^{253}, Met^{318}, Thr^{342} and Pro^{351} of AAE7/ACN1 to the proposed active site.

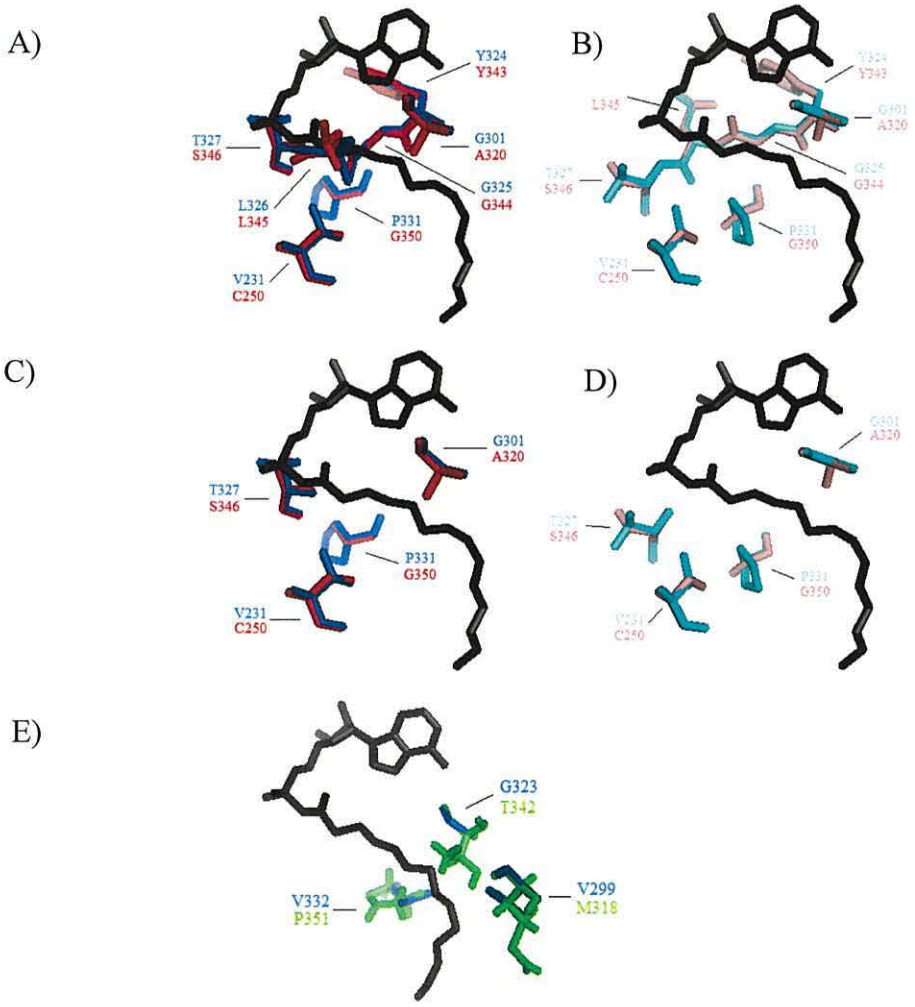


Figure 4.35: Section b. Residues surrounding carbon 4 to carbon 9 of the myristoyl adenosine monophosphate ligand (shown in grey and black) of ttLC-FACS (Hisanaga *et al.*, 2004). A) ttLC-FACS (1ULT; open conformation) residues are represented in blue and the corresponding residues of AAE7/ACN1-modelI are shown in red. B) ttLC-FACS (1V26; closed conformation) residues are represented in light blue and the corresponding residues of AAE7/ACN1-modelII are shown in pink. C) Residues that are not conserved between ttLC-FACS (1ULT; shown in blue) and AAE7/ACN1-modelI (shown in red). D) Residues that are not conserved between ttLC-FACS (1V26; shown in light blue) and AAE7/ACN1-modelII (shown in pink). E) Residues that are not conserved between ttLC-FACS (shown in blue) and AAE7/ACN1-TASSER model (shown in green).

4.4.4.5.3 Section c

Within section c of the myristoyl moiety only one residue is conserved between AAE7/ACN1 and ttLC-FACS (Figure 4.36).

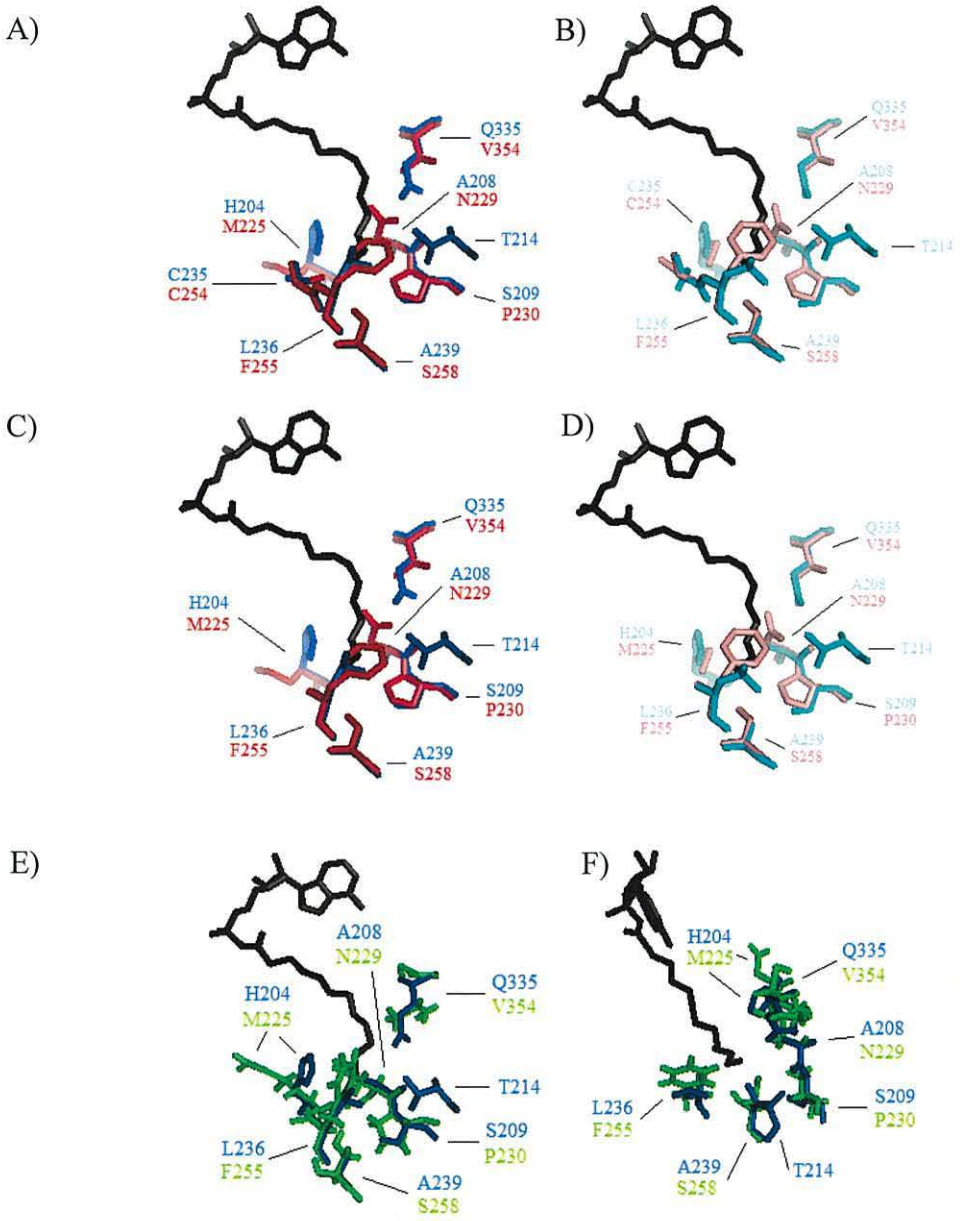


Figure 4.36: Section c. Residues surrounding carbon 9 to 14 of the myristoyl adenosine monophosphate ligand (shown in grey) of ttLC-FACS (Hisanaga *et al.*, 2004). A) ttLC-FACS residues (1ULT; open conformation) are represented in blue and the corresponding residues of AAE7/ACN1-modelI are shown in red. B) ttLC-FACS (1V26; closed conformation) residues are represented in light blue and the corresponding residues of AAE7/ACN1-modelII are shown in pink. C) Residues that are not conserved between ttLC-FACS (1ULT; shown in blue) and AAE7/ACN1-modelI (shown in red). D) Residues that are not conserved between ttLC-FACS (1V26; shown in light blue) and AAE7/ACN1-modelII (shown in pink). E) Residues that are not conserved between ttLC-FACS (shown in blue) and AAE7/ACN1-TASSER model (shown in green). F) A different orientation of (E).

4.5 Discussion

This chapter contains AAE7/ACN1 amino acid comparisons to the protein data bank (pdb) by the online server Phyre (Bennett-Lovsey *et al.,* 2008; Kelley and Sternberg, 2009). Phyre analysis revealed that the LC-FACS from *thermus thermophilus* shares greater sequence comparison with AAE7/ACN1 (30 %) than AAE7/ACN1 does to the previously characterised ACS's of *Saccharomyces cerevisiae* (19 %) and *Salmonella enteric* (19 %).

Sequence comparison of AAE7/ACN1 to the *Arabidopsis* genome showed that AAE7/ACN1 shares greater sequence identity with the medium chain acyl-activating enzyme (AAE) 11 than with the characterised monomeric *Arabidopsis* acetyl-CoA synthetase (atACS). This greater sequence identity, observed between AAE7/ACN1 and AAE11, would suggest that these two proteins are closer in evolutionary history compared to AAE7/ACN1 and the *Arabidopsis* acetyl-CoA synthetase.

There are a wide variety of homology modelling programs and servers currently used in academic research. These include Geno3D (Skouri-Gargouri *et al.,* 2009), I-TASSER (Wong *et al.,* 2010), MOE (Vlachakis, 2009), Modeller (Fazil *et al.,* 2010), Rossetta (Ridgen *et al.,* 2009) and Swiss model (Gao *et al.,* 2008). The large variety of programs available for use makes it difficult to determine which is best suited for the construction of a homology model. The introduction of critical assessment of protein structure prediction (CASP) aims to compare homology modelling servers and provide a method of identifying the most accurate method of creating a homology model. With this in mind homology models of AAE7/ACN1 were created.

4.5.1 Homology modelling of AAE7/ACN1 with I-TASSER

The amino acid sequence of AAE7/ACN1 was submitted to the online modelling server I-TASSER which was nominated best server by CASP 7 analysis (Zhang, 2009). The success of I-TASSER in both CASP 7 (Zhou, 2007) and CASP 8 (Zhang, 2009) suggests that it is one of the most accurate automated homology modelling servers currently available.

The AAE7/ACN1-TASSER model is a monomeric structure based on the threading alignments of ttLC-FACS (1ULT), human MC-ACS (3C5E), firefly luciferase

(1LCI) and DhbE (1MD9). Analysis of the ramachandran plot (Ramakrishnan and Ramachandran, 1965) shows that a small proportion (9.3 %) of the residues fell within the disallowed region, and a large percentage of these (43.4 %) are glycine or proline residues. Glycine residues are considered the simplest amino acid comprising of a hydrogen atom as its R group, resulting in a symmetrical structure. Due to the flexibility of this hydrogen atom glycine residues can form many conformations in all four regions of the ramachandran plot (Schulz and Schirmer, 1979; Buxbaum, 2007), resulting in exclusion from the ramachandran plot (Greeley, 2004). In contrast, proline is only able to form a limited number of possible angles (Hovmoller *et al.*, 2002). As a result of the steric constraints that the side chains of glycine and proline create (Choi and Mato, 2006), these residues are either excluded or analysed separately during ramachandran plot analysis (Baldwin and Rose, 1999; Bertini *et al.*, 2003; Yadav *et al.*, 2010).

Comparison of the disallowed residues of AAE7/ACN1-TASSER model to ttLC-FACS (Table 4.2) shows that three residues (Gly^{207}, Thr^{209} and Ala^{210}) form part of the ATP/AMP signature motif, also denoted the P-loop (section 4.4.1.1 and 4.4.1.3). AAE7/ACN1 residues Gly^{207}, Thr^{209} and Ala^{210} align to Gly^{186}, Thr^{188} and Gly^{189} of ttLC-FACS. The glycine residues within chain A of ttLC-FACS (1ULT) are both present in the disallowed region of the ramachandran plot. This explains why both Gly^{207} and Ala^{210} also appear in the disallowed region of the ramachandran plot as the homology model is based on the crystallographic structure of ttLC-FACS (1ULT). Unfortunately this does not explain why Thr^{209} of AAE7/ACN1-TASSER model falls within the disallowed region as Thr^{188} does not. However, both residues either side of Thr^{188} in ttLC-FACS fall within the disallowed region of the ramachandran plot and thus may affect the orientation of the side chain of Thr^{209}.

Other AAE7/ACN1-TASSER model residues of interest that fall within the disallowed region of the ramachandran plot include Ser^{454} and Lys^{455}. These residues align to Leu^{434} and Lys^{435} of ttLC-FACS and form part of the L-motif (section 4.4.2). Due to the presence of these potentially important residues within the disallowed region of the ramachandran plot care must be taken when analysing these regions as AAE7/ACN1-TASSER model may not represent the true orientation of these residues.

4.5.2 Homology modelling of AAE7/ACN1 with Modeller

Modeller is a well established comparative modelling program with a large interactive community of users who are willing to provide assistance with model creation (Eswar *et al.,* 2007; Qu *et al.,* 2009). Despite the age of Modeller (first released in 1989) the program is under constant review and regularly updated, making it a popular program that is still widely used as a research tool (Ahmed *et al.,* 2009; Dhananjeyan *et al.,* 2009; Khurana *et al.,* 2010).

The main advantage of the modelling program Modeller over online servers is the ability to specify a specific template structure. Structurally characterised acyl-CoA synthetases are believed to undergo a conformational change in shape during catalysis. It was therefore important to identify a modelling program that would allow AAE7/ACN1 to be modelled to template structures that represent both the open and closed structure of acyl-CoA synthetase (section 4.4.3). Homology modelling to specific template structures in the proposed conformations, open and closed, allowed comparisons to be made between the potential forms of AAE7/ACN1 (section 4.5.4.4).

4.5.3 Deciding upon a template structure for AAE7/ACN1 homology modelling

The template structure ttLC-FACS was chosen for homology modelling of AAE7/ACN1 for a variety of different reasons. The main reason was the percentage identity between ttLC-FACS and AAE7/ACN1, identified by Phyre as 30 %. Although AAE7/ACN1 shares the greatest sequence identity with ttLC-FACS both enzymes utilise different substrates. The ttLC-FACS catalyses long chain fatty acids and AAE7/ACN1 catalyses short chain fatty acids. Despite the difference in substrate specificity multiple sequence alignments of seACS, scACS, ttLC-FACS and AAE7/ACN1 failed to reveal conserved regions between the ACS's and AAE7/ACN1 that may indicate a specific region essential for acetate binding. The percentage identity observed between AAE7/ACN1 and ttLC-FACS along with the inability to identify a potential acetate binding region in scACS or seACS, are both reasons why ttLC-FACS was chosen as a template structure for homology modelling of AAE7/ACN1. However, further work could include the creation of an AAE7/ACN1 homology model based on all three crystallographic structures, as what

appear to be small substitutions in amino acid sequence between the enzymes could result in large changes in substrate specificity.

Another reason for choosing ttLC-FACS as the template structure for homology modelling of AAE7/ACN1 was the crystallisation of ttLC-FACS in both the presence and absence of ligands (closed and open conformation respectively). Acyl-CoA synthetases are believed to undergo a conformation change in shape during catalysis, where the small domain moves closer to the large domain (Bar-Tana and Rose, 1968; Gulick *et al.*, 2003; Jogl and Tong, 2004). Crystallisation of ttLC-FACS in both the presence and absence of ligands shows that ttLC-FACS undergoes a conformational change in shape during catalysis (Hisanaga *et al.*, 2004). Whilst different acyl-CoA synthetase enzymes have been crystallised in either the open or closed conformation, at the time of AAE7/ACN1 model creation, ttLC-FACS was the only enzyme crystallised in both conformations. Homology modelling of AAE7/ACN1 using both conformations of ttLC-FACS as template structures allowed residues involved in domain movement to be visually compared.

Enzymatic analysis of AAE7/ACN1 shows that the enzyme is a short chain fatty acyl-CoA synthetase catalysing both acetate and butyrate (Shockey *et al.*, 2003; Turner *et al.*, 2005). It is currently undetermined whether AAE7/ACN1 is associated with the plasma membrane or free in the glyoxysome. Research conducted by Shockey *et al.* (2003) and Turner *et al.* (2005) showed that the association with an artificial membrane affects which substrate, acetate or butyrate, AAE7/ACN1 shows greater activity with. The ttLC-FACS is a membrane associated protein that catalyses the unidirectional movement of fatty acids across the plasma membrane. Construction of a homology model to ttLC-FACS and structural comparison between ttLC-FACS and AAE7/ACN1 residues, identified by Hisanaga *et al.* (2004), that are believed to form the fatty acid entrance site (sections 4.4.4.1 and 4.5.4.5.1) may help to determine if AAE7/ACN1 functions in a similar way to ttLC-FACS and is therefore associated with the plasma membrane.

AAE7/ACN1 was identified by Shockey *et al.* (2003) within a different group, clade VI, to the previously characterised atACS (Ke *et al.*, 2000). The only other characterised member of clade VI is the medium chain acyl-CoA synthetase, AAE11.

The greater sequence identity observed between AAE7/ACN1 and AAE11 suggests that these two proteins are closer in evolutionary history, compared to AAE7/ACN1 and atACS. Since construction of the AAE7/ACN1-models the crystallographic structure of the human medium chain acyl-CoA synthetase (MACS2) has been characterised in both an open and closed conformation (Kochan *et al.,* 2009). Therefore, this could be a more appropriate template model for AAE7/ACN1 as the protein utilises medium chain fatty acid substrates, in contrast to the long chain fatty acid substrates utilised by ttLC-FACS. Further work could include homology modelling of AAE7/ACN1 to MACS2 in both conformations.

4.5.4 Structural analysis of AAE7/ACN1 homology models

The homology models of AAE/ACN1 created by both Modeller and I-TASSER were broken down and analysed in the following sections. The models were divided into sections to compare the adenosine binding site, the P-loop, the A-motif, the L-motif, the G motif, residues involved in domain closure, the fatty acid entrance site and the fatty acid binding region of ttLC-FACS to AAE7/ACN1. The fatty acid binding region of ttLC-FACS was further divided to compare the carbon one to carbon four, carbon four to carbon nine and carbon nine to carbon fourteen regions of myristic acid.

4.5.4.1 Comparison of the adenosine binding site of AAE/ACN1 to ttLC-FACS

The adenosine binding site of ttLC-FACS was divided into motifs (ATP/AMP signature motif, the A-motif and the P-loop) and compared to the corresponding adenosine binding site residues of AAE7/ACN1.

4.5.4.1.1 Analysis of the AAE7/ACN1 ATP/AMP signature motif

The ATP/AMP signature motif is a common sequence used to identify AMP binding proteins. The signature motif is made up of two separate sequences, YTSGTTGxPKGV and GYGxTE, which are also denoted the P-loop and A-motif respectively. The potential importance of conserved residues within these two amino acid sequences is discussed in sections 4.5.4.1.3 and 4.5.4.1.2 respectively.

4.5.4.1.2 Analysis of the AAE7/ACN1 adenosine (A) motif

The A-motif, forming part of the ATP/AMP signature motif, is believed to be important for adenine binding in AMP binding proteins. As AAE7/ACN1 is an AMP binding protein that catalyses the conversion of ATP to AMP it is not surprising that only two residues within the A-motif differ between ttLC-FACS and AAE7/ACN1.

The first substituted residue is Gly^{323} of ttLC-FACS which aligns to Thr^{342} in AAE7/ACN1. The substitution to a much larger, polar residue in AAE7/ACN1 is unlikely to affect the interaction observed in the crystal structure of ttLC-FACS between the oxygen of glycine and nitrogen of adenosine noted by Hisanaga *et al.* (2004). However, the addition of a larger R group may interfere with substrate binding and affect substrate specificity as Gly^{323} is situated within 4.5 Å radius of the myristoyl moiety in the crystallographic structure of ttLC-FACS (Hisanaga *et al.,* 2004). Homology modelling with Modeller in both the open and closed conformation shows that Thr^{342} is positioned very close to the position of the active site of ttLC-FACS (Figure 4.27). The effect that this larger residue (Thr^{342}) in AAE7/ACN1 may have on substrate specificity is unclear, as the position of Thr^{342} does not visibly restrict the fatty acid binding site. The side chain residues of AAE7/ACN1-TASSER model are positioned in a different orientation to the template structure of ttLC-FACS. Despite the observed differences and the closer position of Thr^{342} to the position of the active site of ttLC-FACS Thr^{342} does not appear to block substrate binding. The Thr^{342} residue of AAE7/ACN1 could be the target for future site directed mutagenesis as a potential candidate to increase the substrate specificity of AAE7/ACN1.

The second substituted residue is Thr^{327} of ttLC-FACS which aligns to Ser^{346} in AAE7/ACN1. The threonine residue of ttLC-FACS is believed to be important during catalysis as it interacts with the oxygen of the α-phosphate of myristoyl-AMP (Hisanaga *et al.,* 2004). However, the substitution of a threonine to a serine is only a minor exchange as both amino acids are neutral, polar residues. The substitution to a slightly smaller amino acid residue is unlikely to have a large effect on catalysis as the hydrogen bonds of ttLC-FACS are formed between the backbone nitrogen and the side chain hydroxyl group of the amino acid and the α-phosphate oxygen of myristoyl-AMP.

Previous site directed mutagenesis research conducted by Weimar *et al.* (2002) suggests that Glu^{316} of the long chain fatty acyl-CoA synthetase named fadD is important for enzymatic activity. A Glu^{316} mutation to alanine resulted in a reduction to about 10 % wild type activity. It is therefore not surprising that Glu^{316} of fadD aligns with Glu^{347} of AAE7/ACN1.

4.5.4.1.3 Analysis of the AAE7/ACN1 phosphate (P) loop

The P-loop contains residues that Hisanaga *et al.* (2004) report as important for phosphate binding. The hydroxyl group of ttLC-FACS's Tyr^{184} is an important residue binding to magnesium during catalysis. The conservation of this residue in AAE7/ACN1, Tyr^{205}, suggests that this residue is important and may also bind to magnesium during catalysis. The conservation of this residue in AAE7/ACN1 was expected as other research has shown that magnesium stimulates the first half reaction of acyl-CoA synthetase (Webster, 1963; 1967) and omission reduces enzyme activity (Behal *et al.*, 2002; Schmelz and Naismith, 2009). The stabilisation of the magnesium ion might be important during catalysis as the ion could interact with an oxygen of the phosphate group, stabilising its structure as it does in ttLC-FACS (Hisanaga *et al.*, 2004). The importance of Tyr^{205} was also demonstrated by Weimar *et al.* (2002) who mutated the equivalent residue, Tyr^{214}, in fadD. The mutation of Tyr^{214} in fadD to alanine resulted in the dramatic loss of enzyme activity to less than 10 % that of the wild type fadD. However, the importance of a magnesium ion for AAE7/ACN1 activity has not been fully determined and could be explored in future enzyme activity experiments by reducing the concentration of magnesium chloride.

Out of the nine residues that make up the P-loop of ttLC-FACS, six are conserved in AAE7/ACN1. However the substitution of Thr^{185} and Gly^{189} in ttLC-FACS for Ser^{206} and Ala^{210} in AEE7/ACN1 are both minimal substitutions that probably have little effect on catalysis. For this reason it could be argued that eight out of the nine P-loop residues are conserved.

Previous site directed mutagenesis analysis of fadD Gly^{219} (equivalent to Gly^{189} in ttLC-FACS) to alanine resulted in a reduction in activity, compared to wild type, of 25 to 45 % (Weimar *et al.*, 2002). The observed reduction in activity with oleate led

Weimar *et al.* (2002) to conclude that this glycine residue was important for catalysis rather than ATP binding. It is therefore interesting that the equivalent residue in AAE7/ACN1 is an alanine. Sequence alignment analysis shows that the corresponding residue in scACS and seACS are also glycine residues and surprisingly the corresponding residue in AAE11 is alanine. Enzymatic analysis previously conducted however demonstrates that the presence of an alanine at this position in either AAE7/ACN1 (Shockey *et al.*, 2003; Turner *et al.*, 2005; chapter 3) or AAE11 (Shockey *et al.*, 2003) does not render the enzymes inactive. Further work could include site directed mutagenesis of Ala^{210} of AAE7/ACN1 to a glycine residue and the effect on enzymatic activity observed to see if this substitution is able to increase enzymatic activity.

Weimar *et al.* (2002) also showed that the substitution of Gly^{216}, Thr^{217} and Lys^{222} in fadD to alanine residues led to a reduction in enzyme activity. These residues are all conserved in AAE7/ACN1 and correspond to Gly^{207}, Thr^{208} and Lys^{213} respectively. The conservation of these residues suggests that they too are important for catalysis.

The final P-loop residue that is not conserved between ttLC-FACS and AAE7/ACN1 is the replacement of Lys^{190} for Ser^{211} in AAE7/ACN1 (Figure 4.25). Sequence alignment analysis between scACS, seACS, ttLC-FACS and AAE7/ACN1 shows that the residue at this position is highly variable (Figure 4.4). Therefore the substitution of lysine for a smaller neutral amino acid is unlikely to affect catalysis.

4.5.4.2 Analysis of the AAE7/ACN1 linker (L) motif

The L-motif (also known as the hinge or A8 region) is believed to be important for the movement of the small domain closer to the larger domain during catalysis (section 3.8.7). Whilst it is still unclear whether AAE7/ACN1 is able to undergo a conformational change in shape during catalysis (section 3.8.6) structural comparisons will help to determine if residues that are believed to be important are conserved. Sequence comparisons reveal that only two out of the six residues of the L-motif are not conserved between ttLC-FACS and AAE7/ACN1. The ttLC-FACS Leu^{434} residue aligns with Ser^{454} of AAE7/ACN1 and Leu^{437} of ttLC-FACS aligns to Val^{457} of AAE7/ACN1.

The alignment of Leu^{434} to Val^{457} in AAE7/ACN1 is a relatively minor exchange as they are both hydrophobic amino acids. The presence of a valine residue at this position in AAE7/ACN1 is not unduly surprising as a valine residue can be found at the equivalent position in seACS and scACS (section 4.4).

Replacement of the Leu^{437} for the smaller Ser^{454} residue in AAE7/ACN1 exchanges a hydrophobic residue with a neutral residue. The reduction in hydrophobic residues at this location is unlikely to prevent domain movement in AAE7/ACN1, however this residue could be a target for future site directed mutagenesis analysis (to a hydrophobic amino acid) as a valine residue aligns with Ser^{454} in both seACS and scACS, suggesting that a hydrophobic residue at this position could be important.

Visualisation of both the open and closed conformations of ttLC-FACS and the AAE7/ACN1-models shows only minor differences between 1ULT, 1V26, AAE7/ACN1-modelI and AAE7/ACN1-modelII structures (Figure 4.26). This is not surprising as both replacements are for smaller residues that are unlikely to interfere with domain movement. Structural differences can be seen between ttLC-FACS and the AAE7/ACN1-TASSER model, however. These differences may be the result of the multiple alignments created during model creation as Asp^{452} of AAE7/ACN1 aligns to a glycine in both MACS2 (Kochan *et al.*, 2009) and DhbE (May *et al.*, 2002). The presence of a larger residue in AAE7/ACN1 at this position, compared to the template structures MACS2 and DhbE, may have lead to the differences observed for the position of Asp^{452} and Arg^{453} in AAE7/ACN1 (Figure 4.26). Another explanation for the structural differences observed could result from the presence of two residues of AAE7/ACN1-TASSSER model (Ser^{454} and Lys^{455}) that fall within the disallowed region of the ramachandran plot. The unfavourable angles of these residues could affect the position of the surrounding residues and therefore care should be taken when comparing this region as these residues may not represent their true location within AAE7/ACN1.

4.5.4.3 Comparison of residues involved in magnesium ion binding of ttLC-FACS

Hisanaga *et al.* (2004) suggests that the magnesium ion present in the crystallographic structure of ttLC-FACS interacts with two amino acid residues

(Thr^{184} and Glu^{328}) and the oxygen of AMP. The two residues of ttLC-FACS are Glu^{328} from the A-motif (Figure 4.23) and Thr^{184} from the P-loop (Figure 4.25). The conservation of both of these residues, Glu^{347} and Thr^{205} in AAE7/ACN1 suggests that magnesium could interact within the structure of AAE7/ACN1 in a similar way to ttLC-FACS during catalysis.

4.5.4.4 Comparison between the open and closed conformations of AAE7/ACN1 homology models

Prior to the identification of the crystallographic structure of acyl-CoA synthetase Bar-Tana and Rose (1968) suggested that the protein undergoes a conformational change in shape during catalysis. Whilst crystallographic structures have not been identified to date for ACS in both the open and closed conformation they have been determined for ttLC-FACS (Hisanaga *et al.,* 2004), and more recently MACS2 (Kochan *et al.,* 2009). Initial comparison of ttLC-FACS and AAE7/ACN1 models indicated that all of the residues reported to be involved in forming a stable closed conformation were conserved between the two proteins. However, closer inspection revealed that after Pro^{471}, the C-terminal template structures of ttLC-FACS differ between chains A and B. This observed difference makes it difficult to determine if residues Glu^{475} and Lys^{527} are conserved between ttLC-FACS and AAE7/ACN1 as different sequence alignment results are obtained for AAE7/ACN1 with chains A and B of ttLC-FACS (see appendix). The most complete crystallographic chain of ttLC-FACS appears to be chain A for 1ULT and chain B for 1V26. For this reason chain A of 1ULT and the corresponding section of AAE7/ACN1-modelI, modelled to chain A, was used for structural analysis in place of chain B (section 4.4.3). The results obtained (Figure 4.27) by comparison of 1ULT, chain A and 1V26 chain B to the AAE7/ACN1 models indicate that all of the residues that Hisanaga *et al.* (2004) suggest are involved in the formation of the closed conformation are conserved in AAE7/ACN1. This would suggest that AAE7/ACN1 should be able to form a closed conformation.

The credibility of the sequence alignment between Glu^{495} of AAE7/ACN1 and Glu^{475} of ttLC-FACS is verified by the multiple alignment conducted during the creation of the AAE7/ACN1-TASSER model. This analysis reveals that a glycine residue is also conserved between MACS2, firefly luciferase and DhbE. Whilst

Lys^{527} of ttLC-FACS in the protein data bank file 1ULT appears to align to Lys^{551} of AAE7/ACN1-modelI, these residues do not align between the protein data bank file 1V26 and AAE7/ACN1-modelII because of missing residues in the crystallographic structure of ttLC-FACS (see appendix). The multiple alignment constructed by I-TASSER aligns Lys^{551} of AAE7/ACN1 with other adenylate forming enzymes. The equivalent position to Lys^{551} of AAE7/ACN1 aligns to a lysine in DhbE, another basic residue (arginine) in MACS2 and alanine in firefly luciferase, providing further evidence that the sequence alignment between AAE7/ACN1 and ttLC-FACS is the most accurate. The use of multiple alignments between members of the adenylate forming family helps to support the sequence alignment between AAE7/ACN1 and ttLC-FACS as important residues involved in the creation of a closed conformation are likely to be conserved within the enzyme family.

The conservation of AAE7/ACN1 residues involved in the creation of the closed conformation of ttLC-FACS suggests that AAE7/ACN1 should be able to undergo a conformational change in shape during catalysis, although experimental analysis conducted within chapter 3 of this thesis was unable to determine if this was the case. Further work in the form of crystallisation or solution scattering in the presence of ligands would be required to clarify this situation, as discussed previously in section 3.8.

4.5.4.5 Comparison of the fatty acid binding of ttLC-FACS

For structural comparisons the fatty acid binding region of the ttLC-FACS was divided into sections; the fatty acid entrance site, the G-motif and the fatty acid binding region. The fatty acid binding region was further divided in sections a, b and c containing carbons one to four, four to nine and nine to fourteen of the myristoyl moiety respectively.

4.5.4.5.1 The fatty acid entrance site of ttLC-FACS

Hisanaga *et al.* (2004) propose that during catalysis of ttLC-FACS the fatty acid substrate enters the active site via a fatty acid tunnel that originates from the plasma membrane (Figure 4.37). The fatty acid binding site of ttLC-FACS contains six hydrophilic residues that Hisanaga *et al.* (2004) suggest are important for the attraction of the negatively charged carboxyl group of the fatty acid (Figure 4.28).

Structural analysis of ttLC-FACS and AAE7/ACN1 reveals that whilst the β-strand structure of AAE7ACN1 appears to be conserved in AAE7/ACN1-modelI (Figure 4.28), none of the residues contained within the β-strand are. The loss of the negatively charged hydrophilic residues at the fatty acid entrance site suggests that acetate does not enter AAE7/ACN1 in the same way as the ttLC-FACS substrate myristate. If the fatty acid substrate of AAE7/ACN1 does not enter the active site via a fatty acid entrance site then perhaps it enters at the interface between the large and small domains along with the other substrates required for catalysis. Further analysis could be conducted in the form of site directed mutagenesis followed by enzymatic analysis to block the proposed entrance site by addition of large amino acid residues. Further work could also include homology modelling of AAE7/ACN1 to a medium chain ACS (for example MACS2) as catalysis is conducted in the absence of a second entrance site and therefore all substrates are believed to enter and leave the active site at the interface between the large and small domains. Plausibly short chain ACS enzymes such as seACS or scACS could also be used for homology modelling, although this would have to be taken in view of their limitations, as discussed in section 4.5.3. Homology modelling and structural comparison to either of these structures would help to determine if their mechanism of action is similar to that of AAE7/ACN1.

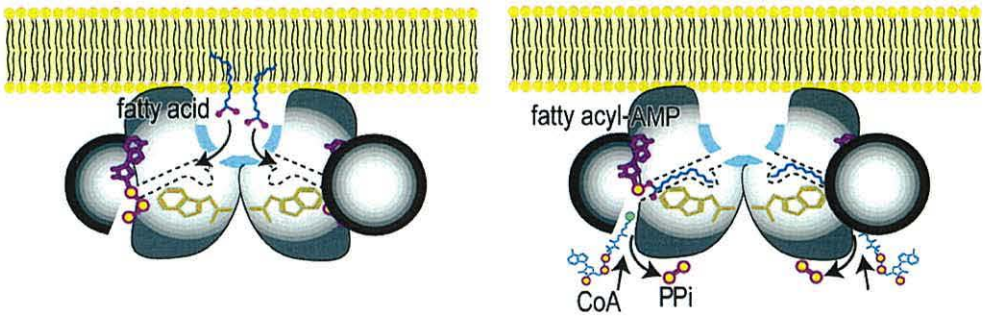


Figure 4.37: The proposed entrance site for the fatty acid substrate of ttLC-FACS (Hisanaga *et al.*, 2004).

The LC-FACS's of *E. coli* and *S. cerevisiae* are both involved in the movement of exogenous fatty acids across the plasma membrane. It is currently unknown whether AAE7/ACN1 is associated with the plasma membrane and therefore involved in the movement of acetate, or simply its activation to acetyl-CoA. The absence of a

conserved fatty acid entrance site in AAE7/ACN1 suggests that AAE7/ACN1 is not associated with the plasma membrane or involved in the transport of fatty acids as LC-FACS's are. However, this conflicts with the suggestions, derived from attempts to crystallise AAE7/ACN1 (section 3.8.1), that AAE7/ACN1 is membrane bound, and thus could suggest methodological reasons for the lack of success in crystallisation attempts. This reaction, catalysed by both short chain and long chain fatty acyl-CoA synthetases, proceeds with a bi uni uni bi ping pong mechanism through an enzyme bound intermediate, comprising of nucleotide and fatty acid in a molar ratio of 1:1 (Bar Tana *et al.,* 1971; 1973). The association of LC-FACS with the plasma membrane and the suggestion of a separate fatty acid entrance site in ttLC-FACS could explain why the transport of long chain fatty acids is considered unidirectional. The transport of long chain fatty acids is conducted undirectionally by the outer membrane protein, FadL and a cytoplasmic fatty acyl-CoA synthetase FadD in *E. coli* (DiRusso and Black, 1999; Black *et al.,* 2000; Weimar *et al.,* 2002). An ABC membrane transporter (COMATOSE) has been identified in *Arabidopsis,* which is proposed to transport acetate across the plasma membrane into the glyoxysome (Hooks *et al.,* 2007). The differences observed between the fatty acid entrance site of ttLC-FACS and AAE7/ACN1 suggests that AAE7/ACN1 is not associated with the plasma membrane. However, experimental analysis is required to determine if AAE7/ACN1 is associated with the plasma membrane, discussed further in section 3.8.8. To fully explore the possibility of a separate entrance site in AAE7/ACN1 further work could also be conducted, in the form of site directed mutagenesis, to reinstate the hydrophilic entrance site followed by enzymatic activity measurements. However substitution of the six residues required to create a hydrophilic entrance site may affect surrounding residues and indirectly affect enzyme activity.

The sequence alignment constructed by I-TASSER reveals that none of the fatty acid entrance site residues of ttLC-FACS (Gly^{238}, Val^{315}, His^{316}, Ala^{340}, Gln^{369} and Asn^{373}) are conserved between AAE7/ACN1 and MACS2, firefly luciferase or DhbE. The only conserved residue appears to be between Arg^{321} of ttLC-FACS to an arginine in MACS2 and firefly luciferase. However, this residue may be conserved between ttLC-FACS and MACS2 as Arg^{321} of ttLC-FACS appears to be in close

proximity to the fatty acid substrate and therefore could be involved in catalysis of LC-FACSs' (Figure 4.28).

4.5.4.5.2 Analysis of the gate (G) motif

The G-motif plays an important role in catalysis of ttLC-FACS as it contains the gate residue Trp^{234} that adopts two different positions in the open and closed conformations of the crystallographic structures (1ULT and 1V26). The movement of the G-motif, and more importantly the gate residue during catalysis opens the fatty acid tunnel allowing the fatty acid to enter the active site. Trp^{234} aligns with Trp^{253} of AAE7/ACN1 suggesting that a similar mechanism may occur during catalysis of AAE7/ACN1. However, under the experimental conditions implemented in chapter 3 of this thesis AAE7/ACN1 did not undergo a conformational change in shape during catalysis. Visualisation of AAE7/ACN1-modelI and ttLC-FACS suggests that the position of the tryptophan residue may restrict entry into the fatty acid binding tunnel preventing fatty acid movement (Figure 4.29). If AAE7/ACN1 is unable to change shape during catalysis, then the position of this tryptophan residue may prevent the fatty acid from entering the active site via the same fatty acid entrance site as ttLC-FACS. If the position of the tryptophan residue is not altered during catalysis as a result of domain movement, then it may also affect the substrate specificity of AAE7/ACN1 by blocking the active site. However, the position of Trp^{253} in AAE7/ACN1-TASSER model adopts a different position that does not block the proposed fatty acid tunnel (Figure 4.29). The conserved tryptophan residue could therefore be the target of future site directed mutagenesis analysis followed by enzyme activity measurements in an attempt to increase substrate specificity. Trp^{253} of AAE7/ACN1 aligned with isoleucine, methionine and serine in MACS2, firefly luciferase and DhbE respectively in I-TASSER multiple alignments. The variety of residues observed at this position would suggest that this tryptophan residue is not essential for catalysis in all acyl-CoA synthetases. Further work could be conducted in the form of crystallisation or solution scattering analysis in the presence of ligands to fully explore the possibility of a conformational change in shape of AAE7/ACN1 during catalysis.

4.5.4.5.3 Comparison of fatty acid binding site of acyl-CoA synthetase

Residues that comprise the base and left wall of the fatty acid binding pocket of acyl-CoA synthetase's have been previously compared by Ingram-Smith *et al.* (2006), Reger *et al.* (2007) and Kochan *et al.* (2009) (Figure 4.31 and 4.32). Sequence alignment analysis (Figure 4.4 and 4.31 and appendix I) reveals that none of the acetate binding pocket residues of seACS (Val^{310}, Val^{386} and Trp^{414}) are conserved in AAE7/ACN1 (Cys^{250}, Ala^{320} and Gly^{344}). Interestingly one out of these three residues is conserved between AAE7/ACN1 (Gly^{344}), MACS2 and ttLC-FACS. This is surprising as enzymatic analysis shows that AAE7/ACN1 catalyses the same reaction as scACS (Shockey *et al.,* 2003; Turner *et al.,* 2005; chapter 3 of this thesis).

The equivalent residue to Ala^{320} of AAE7/ACN1 is also a valine residue in *Methanothermobacter thermautotrophicus* (mtACS) (Ingram-Smith *et al.,* 2006). The substitution of this valine residue (Table 4.10: underlined in column B) of mtACS to an alanine residue altered substrate specificity from acetate (C2) to propionate (C3). Reger *et al.* (2007) also showed in seACS that substitution, of the equivalent Ala^{320} residue, from valine (V^{386}) to alanine increased the size of the substrate binding pocket. This substitution also altered substrate specificity from acetate to propionate (Reger *et al.,* 2007). This is not surprising as medium chain and long chain acyl-CoA synthetases typically have either a glycine or alanine residue at this position. The presence of Ala^{320} at this position in AAE7/ACN1 is surprising as the enzyme is unable to catalyse substrates longer than four carbons (Turner *et al.,* 2005).

The location of alanine at this position in AAE7/ACN1, and the effect that this residue has on substrate specificity of mtACS and seACS, may suggest that AAE7/ACN1 could have a preference for butyrate. Experimental analysis conducted by Shockey *et al.* (2003) in the presence of an artificial membrane showed that the substrate specificity of AAE7/ACN1 was butyrate (C4). Other research conducted in the absence of an artificial membrane showed that AAE7/ACN1 had greater activity with acetate, compared to butyrate (Turner *et al.,* 2005). The replacement for alanine in mtACS and seACS increased the size of the active site and the size of the preferred substrate. The presence of alanine at this position in AAE7/ACN1 suggests that AAE7/ACN1 should be able to utilise larger substrates with greater efficiency.

This along with enzymatic analysis conducted by Shockey *et al.* (2003) suggests that AAE7/ACN1 may be associated with the plasma membrane. Further work could include the site directed mutagenesis of Ala^{320} to valine followed by enzymatic analysis to determine the effect on substrate specificity.

The loss of a tryptophan residue in AAE7/ACN1 to glycine (Gly^{344}) is interesting as this residue is believed to form the base of the binding pocket in acetyl-CoA synthetase (Figure 4.31). Ingram-Smith *et al.* (2006) substituted the equivalent residue in mtACS, Try^{416}, to glycine. This substitution increased substrate specificity allowing mtACS to utilise larger substrates such as butyrate (C4). The presence of a glycine at this position may explain why AAE7/ACN1 is able to catalyse both acetate (C2) and butyrate (C4) as wild type mtACS is only able to utilise acetate (C2) and propionate (C3). It is not unusual for an acetyl-CoA synthetase to catalyse more than one substrate, for example wild type seACS utilises both acetate and propionate, with a preference for acetate (Reger *et al.,* 2007). Other examples include bovine heart mitochondrial ACS (Campagnari and Webster, 1963) and scACS (Frenkel and Kitchens, 1977).

Sequence analysis of the human acyl-CoA synthetase genes (Watkins *et al.,* 2007) reveals a tryptophan in the short chain fatty acyl-CoA synthetase's and a glycine residue in the medium and long chain fatty acyl-CoA synthetase's at the equivalent location to Gly^{344} of AAE7/ACN1. If a glycine residue at this position allows larger substrates to enter the active site, then other residues within AAE7/ACN1 must be limiting substrate specificity.

Perhaps then the presence of a cysteine residue (Cys^{250}) in AAE7/ACN1 restricts substrate specificity as this residue appears to align to a hydrophobic alanine, valine or isoleucine in other acyl-CoA synthetases (Ingram-Smith and Smith, 2006; Kochan *et al.,* 2009). Further work could include the substitution of Cys^{250} of AAE7/ACN1 to alanine, valine or isoleucine and enzymatic analysis conducted to determine the effect this amino acid has on substrate specificity.

Table 4.10: Comparison of the active site residues of acyl-CoA synthetase. Table modified from Ingram-Smith and Smith *et al.* (2006). Underlined amino acids represent residues that align to important active site residues of seACS (V^{310}, V^{386} and W^{414}).

Enzyme	A	B	C
seACS	^{305}ADVGW<u>V</u>TGHSY^{315}	^{385}S<u>V</u>GEP^{389}	^{411}DTW<u>W</u>QTETGGFMIT^{424}
Yeast ACS1	^{361}GDIGW<u>I</u>TGHTY^{371}	^{441}S<u>V</u>GEP^{445}	^{467}DTY<u>W</u>QTESGSHLVT^{480}
Yeast ACS2	^{319}GDVGW<u>I</u>TGHTY^{329}	^{399}S<u>V</u>GEP^{403}	^{425}DTM<u>W</u>QTESGSHLIA^{438}
mtACS1	^{307}ADIGW<u>I</u>TGHSY^{317}	^{387}T<u>V</u>GEP^{391}	^{413}DTW<u>W</u>QTETGMHLIA^{426}
Human SA	^{276}SDTGW<u>A</u>KSAWS^{286}	^{351}S<u>A</u>GEP^{355}	377EGY<u>G</u>QTETV-LICG389
Human MACS1	^{267}SDSGW<u>I</u>VATIW^{277}	^{342}T<u>G</u>GEV^{346}	^{368}ENY<u>G</u>QSETG-LICA^{380}
ttLC-FACS	^{226}VPMFH<u>V</u>NAWCL^{236}	^{300}V<u>G</u>GSA^{304}	^{322}QGY<u>G</u>LTETSPVVV^{334}
AAE7/ACN1	^{245}LPMFH<u>C</u>NGWCF^{255}	^{319}TAGAA^{323}	^{341}HTY<u>G</u>LSETYGPSTV^{354}

The I-TASSER server predicted 23 potential binding site residues within the structure of AAE7/ACN1-TASSER model (Table 4.2). Of these, 26 % are conserved in AAE7/ACN1. Visual comparison of the predicted binding site residues in AAE7/ACN1-TASSER model and ttLC-FACS show a large proportion of the residues in close proximity to the potential ligand position in AAE7/ACN1-TASSER model. Whilst these residues are in close proximity to the potential ligand position none appear to block the potential active site and are therefore not solely responsible for the substrate specificity of AAE7/ACN1.

4.5.4.5.4 Comparison of the fatty acid binding sections of ttLC-FACS

The fatty acid binding region of ttLC-FACS has been divided into sections for easy comparison to AAE7/ACN1. Section a comprises of residues surrounding carbons one to four of myristate (Figure 4.34). Section b includes residues surrounding carbons four to nine (Figure 4.35) and section c contains residues surrounding carbons nine to fourteen (Figure 4.36).

4.5.4.5.4.1 Section a

Comparison of the upper region of the myristoyl moiety of ttLC-FACS to AAE7/ACN1 surprisingly showed that only three out of the seven residues are conserved (Table 4.9; Figure 4.34). Unsurprisingly, two of these conserved residues (Gly^{344} and Leu^{345} of AAE7/ACN1) are situated close to carbons one and two of the myristoyl moiety. The third conserved residue (Tyr^{343} of AAE7/ACN1) is also a

member of the A-motif (section 4.4.1.2), and therefore its conservation is not unexpected. A fourth residue close to carbon one is Thr^{327} in ttLC-FACS, and whilst this residue is not conserved in AAE7/ACN1, the substitution to Ser^{346} is minimal. The substitution of Pro^{331} in ttLC-FACS for Gly^{350} in AAE7/ACN1 replaces the proline with a much smaller residue. The presence of a smaller residue at this position in AAE7/ACN1 is unlikely to block the active site and affect substrate specificity.

The residue substitution that is considered to make the largest difference to substrate specificity is the presence of a hydrophilic cysteine residue (Cys^{250}) in AAE7/ACN1 instead of a hydrophobic valine residue in ttLC-FACS. Val^{231} of ttLC-FACS is located at the end of section a close to carbon four. The change in hydrophobicity at this location may explain why AAE7/ACN1 is not able to catalyse substrates larger than four carbons (section 4.4.4.2 and 4.5.4.5.3).

Initial comparisons concerning the substitution of Gly^{301} in ttLC-FACS to Ala^{320} in AAE7/ACN1 suggested that this substitution was minor as alanine is only slightly larger than glycine. Visualisation of Ala^{320} of AAE7/ACN1 shows that alanine is situated closer to the myristoyl moiety than the Gly^{301} of ttLC-FACS in AAE7/ACN1-modelI and AAE7/ACN1-modelII (Figure 4.34). Structural comparison between AAE7/ACN1-TASSER model and ttLC-FACS also shows that Ala^{320} is very close to the myristoyl moiety (Figure 4.34 C). The location of this residue may also help to explain why AAE7/ACN1 is not able to catalyse substrates longer than four carbon atoms.

Sequence alignment analysis between AAE7/ACN1 and AAE11 reveals that only two out of the seven residues within section a are different (Figure 4.5). The first substituted residue is minimal as Thr^{327} in AAE7/ACN1 aligns with a serine residue in AAE11. The second residue is Ala^{320} of AAE7/ACN1 that aligns to a glycine amino acid in AAE11. The substitution of Ala^{320} of AAE7/ACN1 for a glycine residue in AAE11 suggests that this residue may be involved in substrate specificity as AAE11 is active against medium chain fatty acids. The location of Ala^{320} in AAE7/ACN1 may restrict access to the active site limiting substrate specificity to short chain fatty acids. Further work could include site directed mutagenesis of

Ala^{320} in AAE7/ACN1 to the slightly smaller glycine residue followed by enzymatic analysis to determine if this residue affects substrate specificity of AAE7/ACN1.

4.5.4.5.4.2 Section b

Within section b only the gate residue (tryptophan) is conserved between AAE7/ACN1 and ttLC-FACS. It is not surprising that a large proportion of the residues within this section are not conserved as this is the proposed site for medium chain fatty acids that AAE7/ACN1 is unable to catalyse (Shockey *et al.*, 2003). The addition of the larger threonine residue (Thr^{342}) in AAE7/ACN1 in place of Gly^{323} in both ttLC-FACS and AAE11 may also affect the substrate specificity as it is positioned closer to the proposed substrate binding site (Figure 4.5). This residue, in combination with Ala^{320}, of AAE7/ACN1 could be a target for site directed mutagenesis to increase the substrate range of AAE7/ACN1. Substrate specificity could be determined by enzymatic analysis with a variety of fatty acid substrates.

The substitution of Val^{299} of ttLC-FACS for the larger methionine residue (Met^{318}) in AAE7/ACN1 (Figure 4.35) may also affect substrate specificity as Met^{318} aligns to the hydrophobic residue, leucine in AAE11. Met^{318} could be a target for future site directed mutagenesis experiments restoring the hydrophobic nature of the amino acid at this position in an attempt to increase substrate specificity. However substitution of the leucine residue (Leu^{399}) of the rat acyl-CoA synthetase 4 to methionine did not significantly alter enzymatic activity (Stinnett *et al.*, 2007). This suggests that a hydrophobic residue at the position of Met^{318} in AAE7/ACN1 may not affect substrate specificity.

Interestingly, Trp^{253} and Pro^{351} are the only residues within section b that are conserved between AAE7/ACN1 and AAE11. The importance of these residues in AAE7/ACN1 however remains unclear. Comparison of the fatty acid entrance tunnel between the AAE7/ACN1 and ttLC-FACS suggests that the fatty acid does not enter in the same way, implying that a gate residue would be redundant during catalysis as there would be no need to close the entrance tunnel. The importance of Trp^{253} could be investigated by site directed mutagenesis followed by enzymatic analysis to determine if the presence of a tryptophan at this position affects enzyme activity.

4.5.4.5.4.3 Section c

The amino acid pattern has been totally disrupted in section c (Figure 4.36) with only Cys^{235} of ttLC-FACS being conserved in AAE7/ACN1 (Cys^{254}). The complete alteration in charge within the region of carbons nine to fourteen of the myristoyl moiety may help to explain why AAE7/ACN1 is unable to catalyse long chain fatty acid substrates. Despite the large number of unconserved residues in AAE7/ACN1, visualisation of AAE7/ACN1-modelI, AAE7/ACN1-modelII and AAE7/ACN1-TASSER model does not suggest a structural reason for the inability of AAE7/ACN1 to catalyse long chain fatty acids. Future work could be conducted by site directed mutagenesis to completely alter the residues within this region in an attempt to increase substrate specificity of AAE7/ACN1. However the effect on the enzymes conformation as a result of altering a large number of amino acid residues is unknown.

Residues His^{204}, Ser^{209} and Thr^{214} of ttLC-FACS are positioned at the base of myristoyl moiety surrounding carbons ten to fourteen, creating a hydrophilic environment (Figure 4.36). Comparison between AAE7/ACN1 and ttLC-FACS shows that this hydrophilic chamber is not present in AAE7/ACN1 as His^{204} and Ser^{209} of ttLC-FACS align with Met^{225} and Pro^{230} of AAE7/ACN1. The AAE7/ACN1-TASSER model shows that whilst residues in section c do not restrict entrance into the fatty acid binding pocket, the orientation of the amino acid side chains are different between AAE7/ACN1-TASSER model and ttLC-FACS, particularly for His^{204} of ttLC-FACS and Met^{225} of AAE7/ACN1.

4.5.4 A summary of the AAE7/ACN1 modelling analysis and future directions

Homology modelling of AAE7/ACN1 to the template structures of ttLC-FACS was a valuable exercise to identify potential residues for future site directed mutagenesis experiments. Visual comparisons, conducted within this chapter, of AAE7/ACN1 to ttLC-FACS have identified residues that appear to be in close proximity to the potential active site of AAE7/ACN1 and therefore may restrict access. To fully determine the position of the potential AAE7/ACN1 active site residues relative to the ligand, future work could include docking experiments. These could be conducted in a similar way to acyl-CoA synthetase with the ligand docking program AUTODOCK (Khurana *et al.*, 2010).

Due to the similarity observed between AAE7/ACN1 and ttLC-FACS it is surprising that AAE7/ACN1 is not able to catalyse substrates longer than four carbons. The high degree of sequence similarity observed between AAE7/ACN1 and AAE11 compared to atACS suggests that AAE7/ACN1 may have evolved from a LC-FACS. Further work could include the creation of a homology model of AAE7/ACN1 to the crystallographic structure of MACS2 in both the open and closed conformations.

A number of residues have been suggested for site directed mutagenesis within this chapter in an attempt to determine the role they play in catalysis. Along with enzymatic analysis to determine the effect these mutations may have on enzyme activity, structural techniques such as crystallisation, SAXS and circular dichroism (CD) spectroscopy could be implemented to determine their effect on the structure of AAE7/ACN1. Previous research has been conducted to determine the effect of site directed mutagenesis of acyl-CoA synthetase by crystallisation (Reger *et al.,* 2007) and CD spectroscopy (Lee *et al.,* 2001).

Sequence analysis, expression and purification of aconitase in *Pichia pastoris*

5.1 Introduction

Aconitase is responsible for the reversible isomerisation of citrate to isocitrate via *cis*-aconitate. Members of the aconitase family include mitochondrial aconitase, aconitase B, IRP-aconitase A, homoaconitase and isopropylmalate isomerase (Gruer *et al.,* 1997). Phylogenetic analysis groups plant aconitase with IRP1, IRP2 and bacterial aconitase A, under the umbrella of IRP-aconitase A (Gruer *et al.,* 1997). A characterised member of this group includes the bifunctional protein IRP1. In addition to aconitase function IRP1 is also able to bind to mRNA (Kaptain *et al.,* 1991; Klausner and Rouault; 1993; Pantopoulos *et al.,* 1995). This is in contrast to IRP2 which does not function as aconitase (Guo *et al.,* 1995a).

Structural analysis of aconitase reveals a four domain structure with a [4Fe-4S] cluster at its active site. This [4Fe-4S] centre is very sensitive and often requires reactivation by addition of iron and cysteine after purification (Dickman and Cloutier, 1950; Morrison, 1954). The first three irons of the [4Fe-4S] cluster are ligated by cysteine side chains. The fourth iron is available for participation in the isomerisation reaction (Volz, 2008).

Acontiase was first identified in *Arabidopsis thaliana* by Peyret *et al.* (1995). Subsequent studies have identified three aconitase isoforms in *Arabidopsis* denoted ACO1, ACO2 and ACO3 (Arnaud *et al.,* 2007; Moeder *et al.,* 2007). Research suggests that ACO1 is located within the cytosol and ACO2 and ACO3 are both located within the mitochondria (Bernaud *et al.,* 2009). In four week old leaf tissue Moeder *et al.* (2007) noted a reduction in total aconitase activity of 70 % in *aco3* and 20 % in both *aco1* and *aco2*.

Previous studies of aconitase over-expression from *Nicotiana tabacum* (ntACO1) in a prokaryotic system required reactivation after expression under denaturing conditions (Navarre *et al.,* 2000). Over-expressed ntACO1 was produced in inclusion bodies of *E. coli*, solubilised, refolded and activated. Despite this complete reactivation was not obtained. The specific activity of purified ntACO1 from four to five week old leaf tissue was determined as 700 U mg^{-1} protein. The specific activity of ntACO1 from over-expressed *E. coli* cells was measured as 2 - 4 U mg^{-1} protein. Navarre *et al.* (2000) suggested that this observed reduction in specific activity was

either due to the improper expression in *E. coli* or an inability to reactivate plant aconitase.

ACO1 from *Arabidopsis* has been over-expressed in a prokaryotic system under denaturing conditions similar to Navarre *et al.* (2000) and mRNA binding studies conducted (Moeder *et al.*, 2007). Binding studies with recombinant tobacco, *Arabidopsis* aconitase (ACO1) and mammalian IRP1 showed binding to the 5' UTR of the *Arabidopsis* chloroplastic CuZn superoxide dismutase 2 (CSD2) (Moeder *et al.*, 2007). To ensure that a correctly folded, active protein is being investigated a system is required to over-express aconitase under native conditions without the need for reactivation. A method has been developed and described within this chapter to over-express and purify active aconitase in *Pichia pastoris* without the need for reactivation. In addition homology modelling analysis has been conducted to compare the predicted structure of ACO1, ACO2 and ACO3.

5.2 Characterisation of *Arabidopsis* aconitase genes

Wu BLAST (BlastP) analysis was conducted on the *Arabidopsis* proteome with the mammalian IRP1 (GenBank® accession number NP002188) amino acid sequence to identify potential aconitase isoforms (Figure 5.1). Three genes, At4g35830 (ACO1), At4g26970 (ACO2) and At2g05710 (ACO3) were identified that are believed to have aconitase activity. This is in agreement with previous studies by Moeder *et al.* (2007) and Arnaud *et al.* (2007). BLAST analysis was conducted with the aconitase nucleotide sequences to identify the location of the three genes within the *Arabidopsis* genome. Chromosome four contains ACO1 at position 16973011 - 16977953 and reads in the reverse direction. Chromosome four also contains ACO2 at position 13543083 - 13548433 and reads in the forward direction and chromosome two contains ACO3 at position 2141588 - 2146347 and reads in the forward direction.

Sequence alignment analysis shows that the *Arabidopsis* aconitase isoforms share a large amount of amino acid identity with both the mammalian IRP1 and ntACO1 (Figure 5.1).

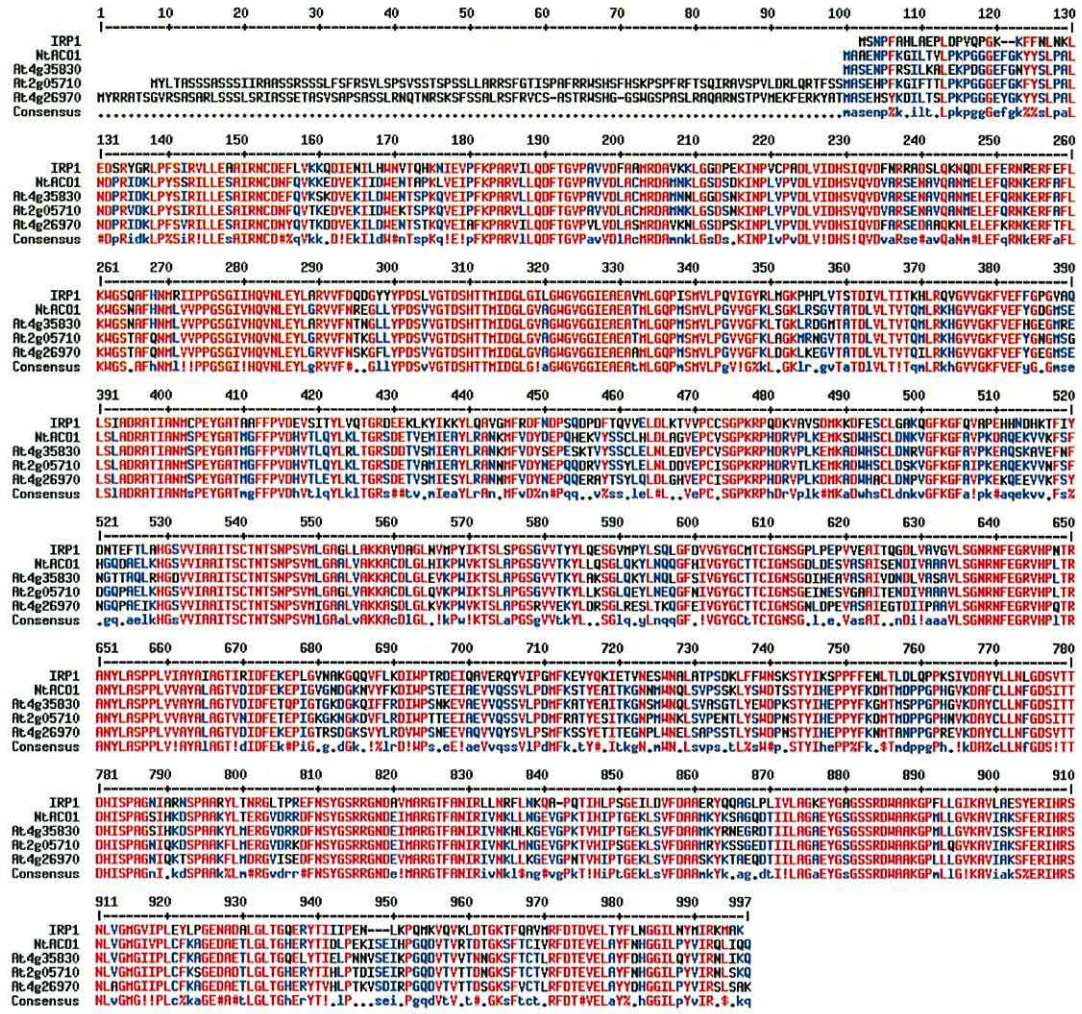


Figure 5.1: Sequence comparison of the mammalian IRP1, *Nicotiana tabacum* ACO1 (ntACO1) and *Arabidopsis thaliana* aconitase isoforms At4g35830, At4g26970 and At2g05710 performed with Multalin (Corpet, 1988) using default settings. Residues represented in red show regions of conservation between all five amino acid sequences, residues represented in blue show conservation between three or more sequences and all other residues are shown in black.

Sequence alignment of the three aconitase isoforms (ACO1, ACO2 and ACO3) conducted with the default settings of EMBOSS Pairwise alignment algorithm (Rice *et al.,* 2000) revealed a large amount of amino acid sequence conservation with the genes sharing greater than 70 % identity (Table 5.1 and Figure 5.2).

Table 5.1: Amino acid Sequence comparison between the three *Arabidopsis* aconitase isoforms performed using the default settings of EMBOSS pairwise alignment algorithm (Rice *et al.*, 2000) reveals a large percentage similarity and identity between the three aconitase isoforms.

Comparison between		% Identity	% Similarity
ACO1	ACO2	71.9	80.3
ACO1	ACO3	78.0	84.2
ACO2	ACO3	77.6	88.2

Amino acid sequence comparison (Figure 5.2) shows that the *Arabidopsis* aconitase isoforms share a large amount of sequence identity. The main difference observed between the aconitase sequences is the addition of a 100 amino acid sequence to the N-terminal of ACO2 and ACO3. The purpose of this leader sequence is currently unknown.

Figure 5.2: Sequence comparison of ACO1, ACO2 and ACO3 performed with Multalin (Corpet, 1988) using default settings. Residues represented in red show regions of conservation between all three amino acid sequences. All other resides are shown in black.

5.3 Proposed location of the *Arabidopsis* aconitase isoforms

Protein prediction analysis was conducted using the default setting of the online servers Predotar (Small *et al.*, 2004), MitoProt II (Claros and Vinces, 1996) and

iPSORT (Bannai *et al.*, 2002) (Table 5.2). MitoProt II predicts ACO2 and ACO3 as mitochondrial. Predotar and iPSORT also predict a mitochondrial location for ACO2.

Table 5.2: Molecular weight and PI were predicted with compute pI/MW from the proteomics server, ExPASy (Gasteiger *et al.*, 2003). Protein location was predicted using online databases Predotar (Small *et al.*, 2004), MitoProt II (Claros and Vinces, 1996) and iPSORT (Bannai *et al.*, 2002). MitoProt II results show the probability of transportation to the mitochondria.

	Predicted MW	Predicted PI	Predotar (Small *et al.*, 2004)	MitoProt II (Claros and Vinces, 1996)	iPSORT (Bannai *et al.*, 2002)
ACO1	98152	5.98	Other	0.2313	Not Mitochondrial
ACO2	108481	6.71	Mitochondrial	0.9986	Mitochondrial
ACO3	108201	6.72	Plastid	0.9977	Chloroplast

Amino acid sequence alignment of the aconitase isoforms revealed a large degree of similarity (Figure 5.2). The most striking difference between the amino acid sequences is the N-terminus leader sequence of ACO2 and ACO3. It is possible that this sequence is required for transport into the mitochondria as MitoProt II predicts, with a high probability of 99 %, a mitochondrial location for both ACO2 and ACO3.

The *Homo sapiens* IRP1 (hsIRP1) is a dual functioning protein with iron response binding protein and aconitase function located within the cytosol (Kaptain *et al.*, 1991; Klauser and Rouault, 1993). A second aconitase isoform (hsACO2) is located within the mitochondria and only has aconitase activity (Mirel *et al.*, 1998). These proteins along with the hsIRP2 were aligned with *Arabidopsis* ACO1, ACO2 and ACO3 (Figure 5.3) to shed light on the purpose of the N-terminal leader sequence of ACO2 and ACO3 (see section 5.6.1 and 5.14.4.1).

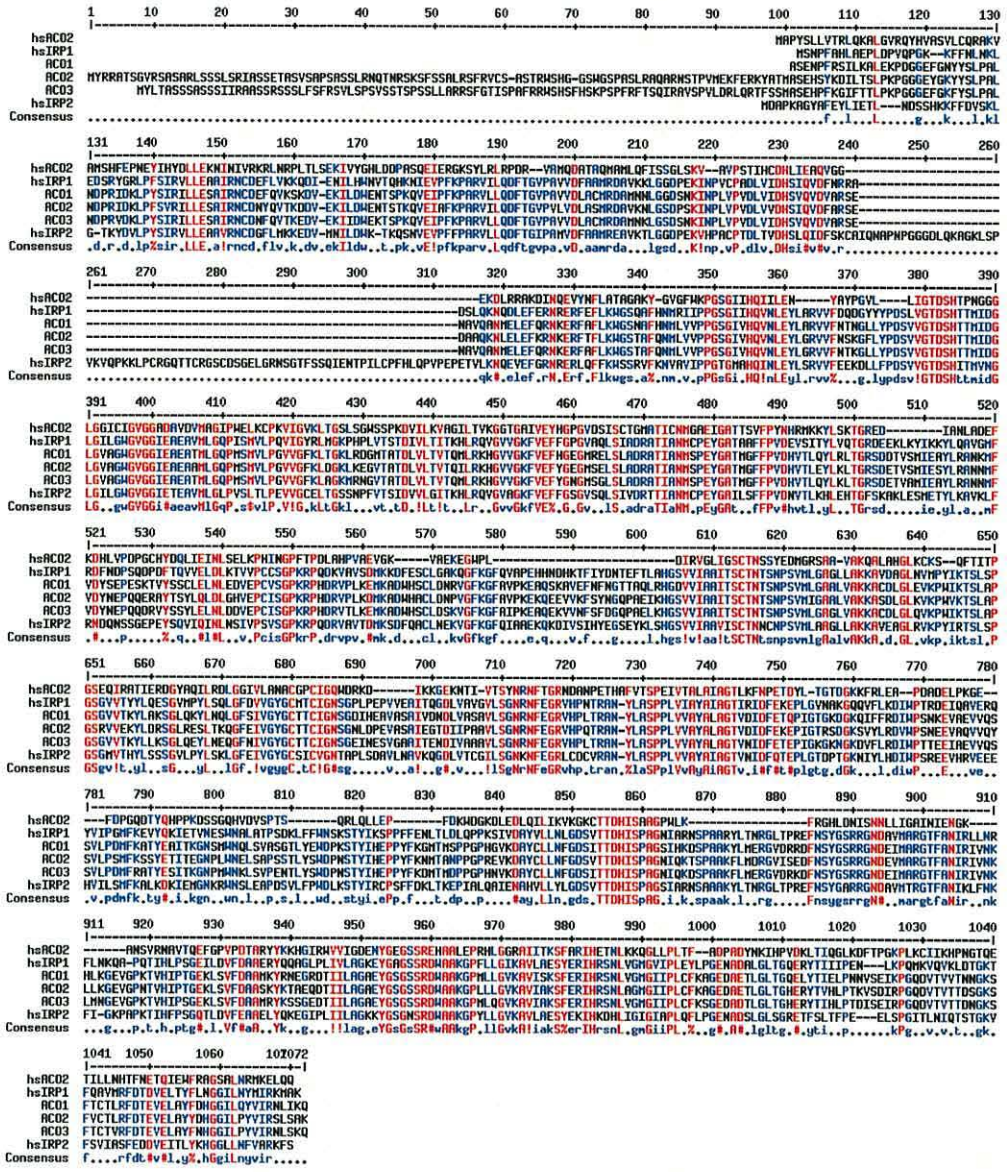


Figure 5.3: hsIRP1, hsACO2, hsIRP2, ACO1, ACO2 and ACO3 sequence alignment conducted with the default settings of Multalin (Corpet, 1988). Residues represented in red show regions of conservation between all six amino acid sequences, residues represented in blue show conservation between four or five sequences and all other residues are shown in black.

Sequence alignment analysis reveals that sections within the hsACO2, hsIRP1, hsIRP2, ACO1, ACO2 and ACO3 are all conserved (Figure 5.3; shown in red). Sequence conservation between all six sequences is surprising as they are not all responsible for the same function. The hsIRP2 does not possess aconitase function and hsACO2 does not bind to mRNA.

5.4 Phyre analysis of ACO1, ACO2 and ACO3

The amino acid sequences of ACO1, ACO2 and ACO3 were submitted to the online server Phyre. Sequence comparison of the aconitase isoforms to the protein data bank revealed that the crystallographic structure of the cytosolic hsIRP1 (2b3x.pdb) shared the greatest similarity to the *Arabidopsis* aconitase isoforms (Table 5.3).

Table 5.3: Phyre (Kelley and Sternberg, 2009) analysis of *Arabidopsis* aconitase isoforms, ACO1, ACO2 and ACO3.

ACO1		ACO2		ACO3	
Protein	% Identity	Protein	% Identity	Protein	% Identity
Homo sapiens cytosolic IRP1 (2b3x.pdb)	61	*Homo sapiens* cytosolic IRP1 (2b3x.pdb)	60	*Homo sapiens* cytosolic IRP1 (2b3x.pdb)	61
Bos taurus mitochondrial ACO (1aco.pdb)	30	*Pyrococcus horikoshii* Isopropylmalate isomerase (1v7l.pdb)	31	*Pyrococcus horikoshii* Isopropylmalate isomerase (1v7l.pdb)	29
Pyrococcus horikoshii Isopropylmalate isomerase (1v7l.pdb)	29	*Bos taurus* mitochondrial ACO (1aco.pdb)	29	*Bos Taurus* mitochondrial ACO (1aco.pdb)	28
E. coli aconitase B (1l5j.pdb)	19	*E. coli* aconitase B (1l5j.pdb)	19	*E. coli* aconitase B (1l5j.pdb)	18

5.5 Homology model construction of ACO1, ACO2 and ACO3

The amino acid sequences of ACO1, ACO2 and ACO3 were submitted to the online modelling program I-TASSER (Zhang, 2008). All three aconitase isoforms were modelled to the template structure of the *Homo sapiens* iron response protein (hsIRP1) (2B3X.pdb). Ramachandran plot analysis of the template structure revealed that 3.7 % of the residues fall within the disallowed region, and of these 72.7 % are glycine and proline residues (Figure 5.4). Glycine residues are considered the simplest amino acid comprising of a hydrogen atom as its R group, resulting in a symmetrical structure. Due to the flexibility of this hydrogen atom glycine residues can form many conformations in all four regions of the ramachandran plot (Schulz and Schirmer, 1979; Buxbaum, 2007). In contrast, proline is only able to form a

limited number of possible angles (Hovmoller *et al.,* 2002). As a result of the steric constraints that the side chains of glycine and proline create (Choi and Mato, 2006), these residues are either excluded or analysed separately during ramachandran plot analysis (Baldwin and Rose, 1999; Bertini *et al.,* 2003; Yadav *et al.,* 2010).

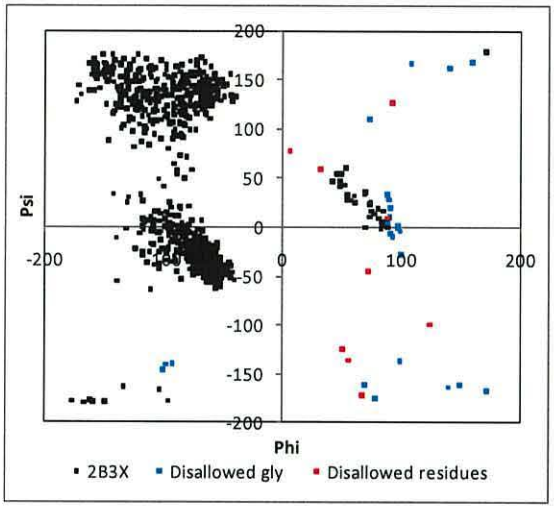


Figure 5.4: Ramachandran plot of the template structure of IRP1 from *Homo sapiens* (Protein data bank file 2B3X). Normal residues are shown in black, problem glycine residues in blue and other problem residues identified by Swiss-pdbViewer (Guex and Peitsch, 1997) in red.

Ramachandran plot analysis of the ACO1 model, created by I-TASSER, was analysed for residues with unfavourable angles using the default settings of Swiss-pdb Viewer (Guex and Peitsch, 1997). The ramachandran plot revealed that 8.4 % of the residues fall within the disallowed region, and 44.0 % of these are glycine and proline residues (Figure 5.5).

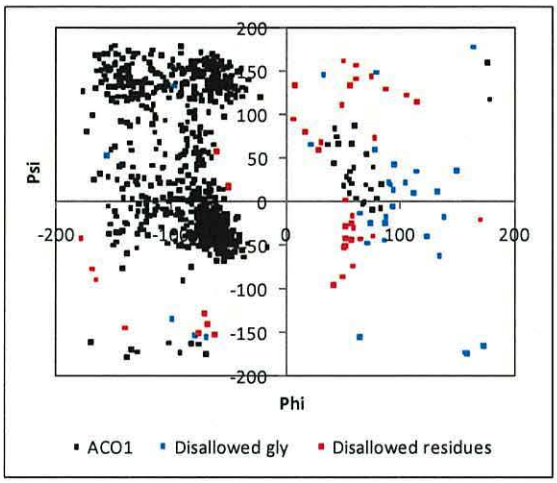


Figure 5.5: Ramachandran plot of the homology model of ACO1 generated with I-TASSER (Zhang, 2008) based on hsIRP1 (2B3X.pdb). Normal residues are shown in black, problem glycine and proline residues in blue and other problem residues identified by Swiss-pdbViewer (Guex and Peitsch, 1997) in red.

Ramachandran plot analysis of the ACO2 model revealed that 18.2 % of the residues fall within the disallowed region (Figure 5.6). Of these residues, 32.0 % are glycine and proline residues (Figure 5.6).

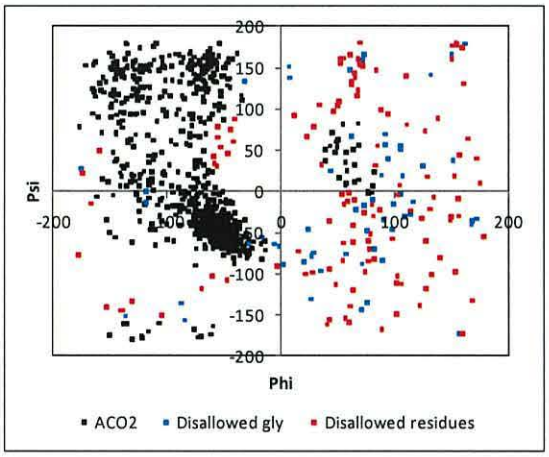


Figure 5.6: Ramachandran plot of the homology model of ACO2 generated with I-TASSER (Zhang, 2008) based on hsIRP1 (2B3X.pdb). Normal residues are shown in black, problem glycine and proline residues in blue and other problem residues identified by Swiss-pdbViewer (Guex and Peitsch, 1997) in red.

Ramachandran plot analysis of the ACO3 model revealed that 18.2 % of the amino acid residues fall within the disallowed region (Figure 5.7). Of these, 39.4 % are glycine and proline residues (Figure 5.7).

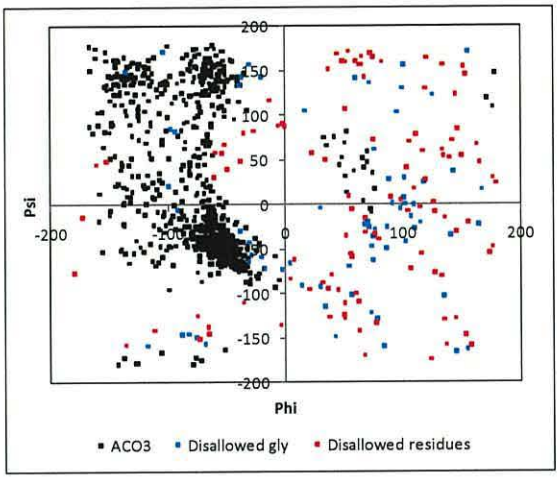


Figure 5.7: Ramachandran plot of the homology model of ACO3 generated with I-TASSER (Zhang, 2008) based on hsIRP1 (2B3X.pdb). Normal residues are shown in black, problem glycine and proline residues in blue and other problem residues identified by Swiss-pdbViewer (Guex and Peitsch, 1997) in red.

5.6 Structural comparison of the *Arabidopsis* aconitase homology models

The *Arabidopsis* aconitase models were compared to establish if all three isoforms had the potential to function as aconitase and convert citrate to isocitrate. In addition, the aconitase models were compared to the template structure of hsIRP and *Sus scrofa* mitochondrial aconitase (ssACO), 2b3x.pdb and 7acn.pdb respectively.

Visual comparison of the ribbon structure of the aconitase models revealed a similar three dimensional structure. The most obvious difference between ACO1, ACO2 and ACO3 is the presence of a flexible loop in ACO2 (Figure 5.8). This loop is comprised of the N-terminal leader sequence residues. This loop structure is also observed in ACO3 as a result of its N-terminal leader sequence. However, this loop sequence occupies a different position in ACO3 and is therefore not visible in the orientation presented in Figure 5.8.

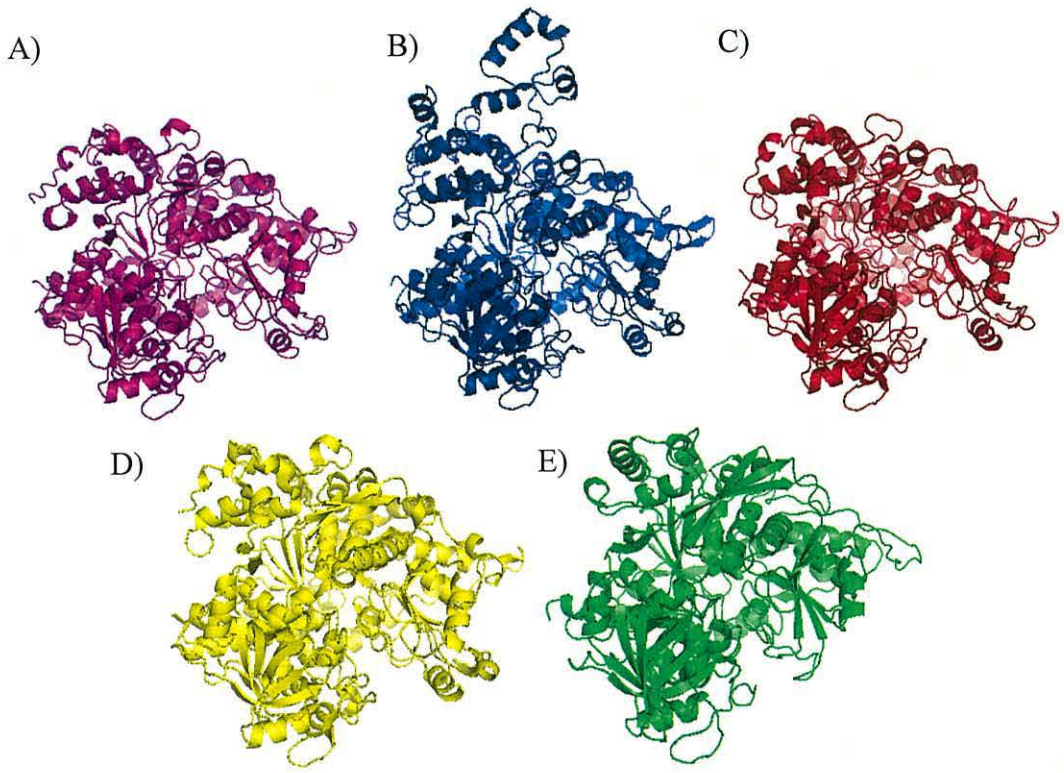


Figure 5.8: Ribbon structure of A) ACO1, B) ACO2, C) ACO3, D) hsIRP1 and E) ssACO created with Pymol.

5.6.1 The N-terminal sequence of aconitase

The N-terminal regions of ACO1, ACO2 and ACO3 are all very different. The amino acid sequences of ACO2 and ACO3 both contain an N-terminal leader sequence that is not present in other aconitases. Other work has compared the N-terminal sequence of cytosol aconitase to mitochondrial aconitase (Dupuy *et al.*, 2006). Comparisons revealed a different fold structure (Figure 5.9) for the N-terminal region of domain 1 of hsIRP1 (residues 1-33) compared to mitochondrial aconitase (Dupuy *et al.*, 2006). The corresponding residues to the N-terminal of hsIRP1 have been compared between ACO1, ACO2 and ACO3 (Figure 5.9; Table 5.4).

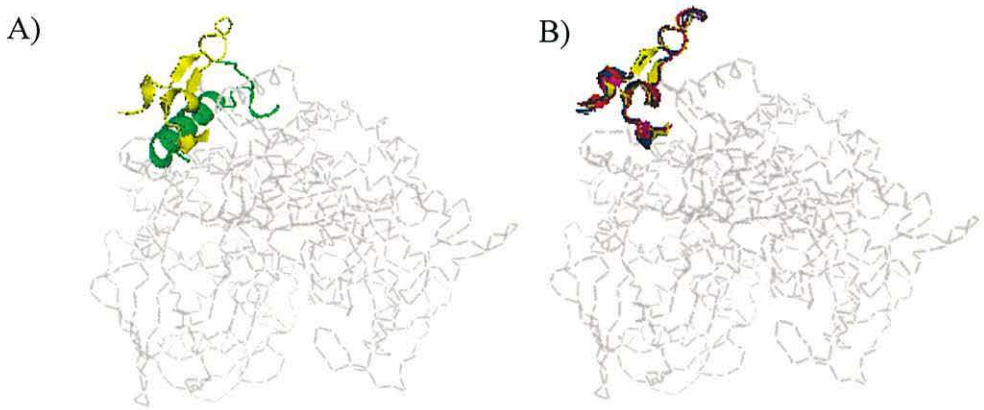


Figure 5.9: N-terminal sequence comparison. A) The first 33 residues of hsIRP1 (shown in yellow) compared to the first 33 residues of ssACO (shown in green). The remaining residues of hsIRP1 are shown in grey. B) The corresponding residues in ACO1 (shown in pink), ACO2 (shown in blue) and ACO3 (shown in red). The remaining residues of hsIRP1 are shown in grey.

Table 5.4: The N-terminal sequence of hsIRP1 and the corresponding residues on ssACO, ACO1, ACO2 and ACO3. Yellow bold letters highlight identical residues between hsIRP1 and *Arabidopsis* aconitase isoforms (ACO). Green bold letters highlight identical residues between ssACO and *Arabidopsis* aconitase isoforms (ACO). Underlined letters highlight identical residues between ACO1, ACO2 and ACO3. Italicised letters highlight identical residues between ACO2 and ACO3.

Aconitase	Sequence
hsIRP1	^{1}MSNPFAHLAEPLDPVQPGKKFFNLNKLEDSRYG^{33}
ssACO	^{2}RAKVAMSHFEPHEYIRYDLLEKNIDIVRKRLNR^{34}
ACO1	^{3}SENPFRSILKALEKPDGGEFGNYYSLPALNDPR^{35}
ACO2	^{100}*SEH*SYKD*I*L*T*S*L*P*KPGG*G*E*Y*G*K*Y*YSLPALNDPR*^{133}
ACO3	^{95}*SEH*PFKG*I*F*T*T*L*P*KPGG*G*E*F*G*K*F*YSLPALNDPR*^{127}

The N-terminal sequence of ACO2 and ACO3 contain approximately 100 extra amino acids. The proposed structure of these sequences can be observed in Figure 5.10.

Figure 5.10: The N-terminal residues 1 - 33 of hsIRP1 (shown in yellow), ACO1 residues 1 - 35 (shown in pink), ACO2 resides 1 – 133 (shown in blue) and ACO3 residues 1 - 127 (shown in red). The N-terminal sequence of each protein is represented in a ribbon structure.

5.6.2 The aconitase linker sequence

The aconitase linker sequence is an extended sequence located at the protein surface that joins domains 3 and 4 (Dupuy *et al.*, 2006). Dupuy *et al.* (2006) showed that this linker sequence in hsIRP1 (R593 to L654) is about 20 residues longer than ssACO (D501 to P543). Sequence analysis aligns the linker of hsIRP1 to S^{596} - M^{658} (ACO1), S^{693} - A^{755} (ACO2) and T^{688} - M^{750} (ACO3). A structural comparison of the linker sequences of ssACO, hsIRP1, ACO1, ACO2 and ACO3 can be seen in Figure 5.11.

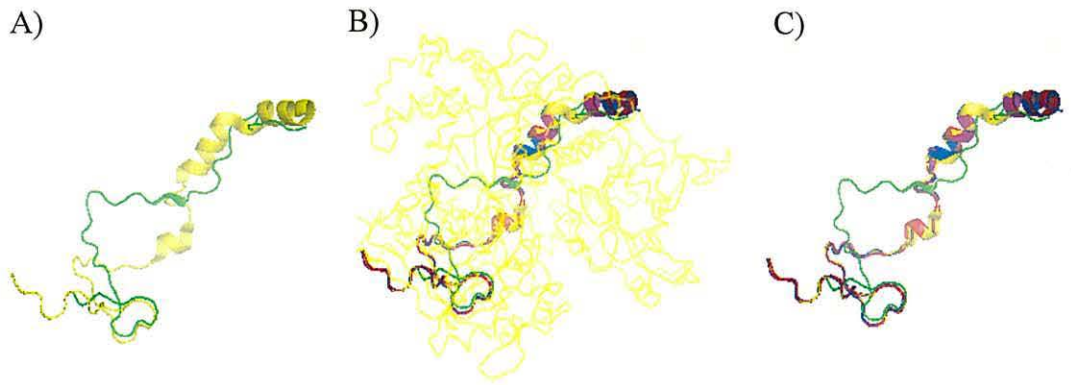


Figure 5.11: The domain 3 and 4 linker sequence involved in the movement of domain 4. A) The different structure observed between hsIRP1 (shown in yellow) and ssACO (shown in green). B) the ribbon structure of hsIRP1 (shown in yellow) demonstrating the location of the linker sequence within the cytosolic aconitase structure. The corresponding residues of the linker sequence of ACO1 (shown in purple), ACO2 (shown in blue), ACO3 (shown in red) and ssACO (shown in green) are also shown. C) The linker sequence of hsIRP1 (shown in yellow) and the corresponding residues in ACO1 (shown in purple), ACO2 (shown in blue), ACO3 (shown in red) and ssACO (shown in green).

5.6.3 Comparison of residues involved in aconitase activity

Residues believed to be involved in substrate recognition, cluster ligation and catalysis have been compared to determine the potential of ACO1, ACO2 and ACO3 to function as aconitase (Table 5.5).

Table 5.5: Important catalytic residues of mitochondrial aconitase (Lauble *et al.*, 1992) and the corresponding residues in hsIRP1, ACO1, ACO2 and ACO3.

	ssACO	hsIRP1	ACO1	ACO2	ACO3
Substrate recognition					
	Gln^{72}	Gln^{86}	Gln^{90}	Gln^{187}	Gln^{182}
	Ser^{166}	Ser^{206}	Ser^{210}	Ser^{307}	Ser^{302}
	Arg^{447}	Arg^{536}	Arg^{540}	Arg^{637}	Arg^{632}
	Arg^{452}	Arg^{541}	Arg^{545}	Arg^{642}	Arg^{637}
	Arg^{580}	Arg^{713}	Arg^{717}	Arg^{814}	Arg^{809}
	Ser^{643}	Ser^{779}	Ser^{784}	Ser^{881}	Ser^{876}
	Arg^{644}	Arg^{780}	Arg^{785}	Arg^{882}	Arg^{877}
Cluster ligation and interaction					
	Asn^{258}	Asn^{298}	Asn^{302}	Asn^{399}	Asn^{394}
	Cys^{358}	Cys^{437}	Cys^{441}	Cys^{538}	Cys^{533}
	Cys^{421}	Cys^{503}	Cys^{507}	Cys^{604}	Cys^{599}
	Cys^{424}	Cys^{506}	Cys^{510}	Cys^{607}	Cys^{602}
	Asn^{446}	Asn^{535}	Asn^{539}	Asn^{636}	Asn^{631}
Catalysis					
	Asp^{100}	Asp^{125}	Asp^{129}	Asp^{226}	Asp^{221}
	His^{101}	His^{126}	His^{130}	His^{227}	His^{222}
	His^{147}	His^{178}	His^{182}	His^{279}	His^{274}
	Asp^{165}	Asp^{205}	Asp^{209}	Asp^{306}	Asp^{301}
	His^{167}	His^{207}	His^{211}	His^{308}	His^{303}
	Glu^{262}	Glu^{302}	Glu^{306}	Glu^{403}	Glu^{398}
	Ser^{642}	Ser^{778}	Ser^{783}	Ser^{880}	Ser^{875}
Conserved residues in [Fe-S] isomerase					
	Ile^{145}	Ile^{176}	Ile^{180}	Ile^{277}	Ile^{272}
	Thr^{168}	Thr^{208}	Thr^{212}	Thr^{309}	Thr^{304}
	Ile^{355}	Ile^{434}	Ile^{438}	Ile^{535}	Ile^{530}
	Arg^{447}	Arg^{536}	Arg^{540}	Arg^{637}	Arg^{632}
	Ser^{642}	Ser^{778}	Ser^{783}	Ser^{880}	Ser^{875}
Supporting residues					
	Ala^{74}	Phe^{88}	Phe^{92}	Phe^{189}	Phe^{184}
	Thr^{567}	Thr^{677}	Thr^{681}	Thr^{778}	Thr^{773}
	Asp^{568}	Asp^{678}	Asp^{682}	Asp^{779}	Asp^{774}
	Ser^{571}	Ser^{681}	Ser^{685}	Ser^{782}	Ser^{777}

5.6.4 Aconitase residues involved in substrate recognition

The seven residues that Lauble *et al.* (1992) state as being involved in substrate recognition (Table 5.5) are all conserved in hsIRP1, ACO1, ACO2 and ACO3 (Figure 5.12).

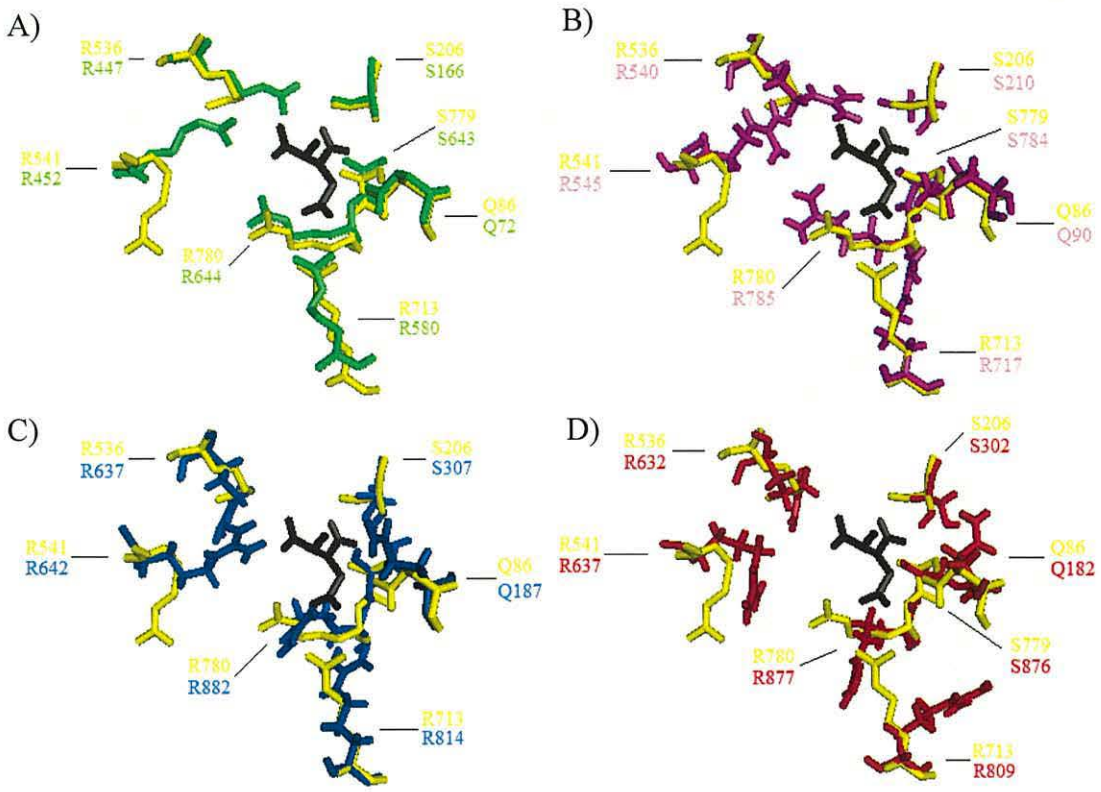


Figure 5.12: Residues of ssACO (shown in green) involved in substrate recognition. A) Residues of ssACO (shown in green) aligned to the corresponding residues of hsIRP1 (shown in yellow). B) Residues of hsIRP1 (shown in yellow) aligned to the corresponding residues of ACO1 (shown in purple). C) Residues of hsIRP1 (shown in yellow) aligned to the corresponding residues of ACO2 (shown in blue). D) Residues of hsIRP1 (shown in yellow) aligned to the corresponding residues of ACO3 (shown in red). Isocitrate is shown in black.

5.6.5 Residues involved in cluster ligation

Residues involved in the cluster ligation of ssACO were compared to hsIRP1, ACO1, ACO2 and ACO3 (Figure 5.13). Analysis revealed that the essential residues are all conserved (Table 5.5).

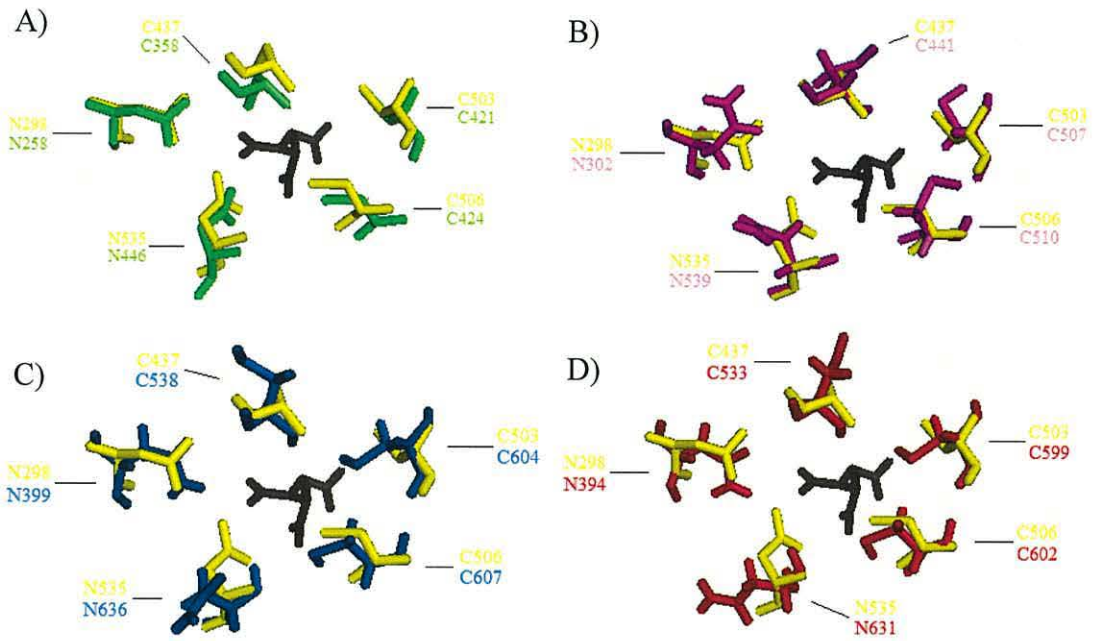


Figure 5.13: Residues of ssACO (shown in green) involved in cluster ligation. A) Residues of ssACO (shown in green) aligned to the corresponding residues of hsIRP1 (shown in yellow). B) Residues of hsIRP1 (shown in yellow) aligned to the corresponding residues of ACO1 (shown in purple). C) Residues of hsIRP1 (shown in yellow) aligned to the corresponding residues of ACO2 (shown in blue). D) Residues of hsIRP1 (shown in yellow) aligned to the corresponding residues of ACO3 (shown in red). Isocitrate is shown in black.

5.6.6 Catalytic residues of aconitase

Residues involved in the catalysis of ssACO were aligned and compared to ACO1, ACO2 and ACO3 (Figure 5.14). Analysis revealed that all of the catalytic residues are conserved in hsIRP1, ssACO, ACO1, ACO2 and ACO3.

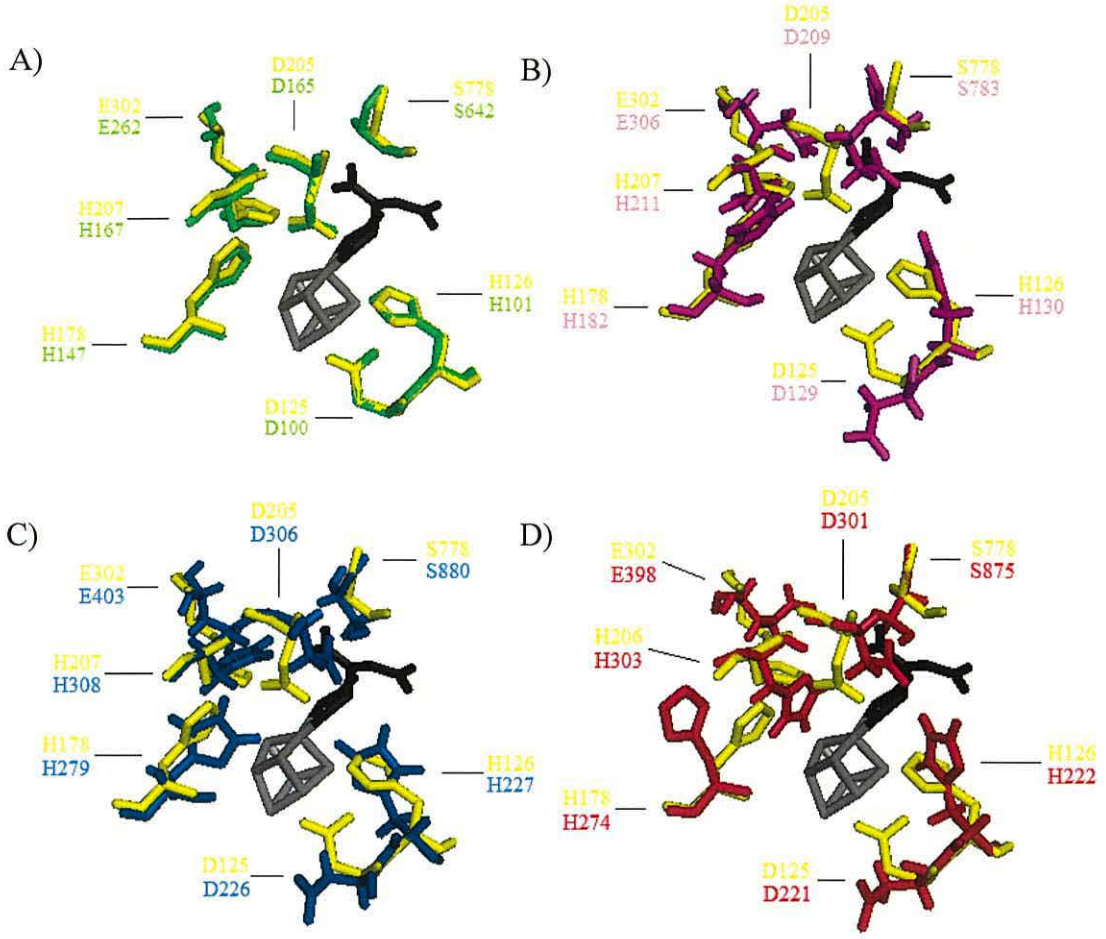


Figure 5.14: Residues of ssACO (shown in green) involved in catalysis. A) Residues of ssACO (shown in green) aligned to the corresponding residues of hsIRP1 (shown in yellow). B) Residues of hsIRP1 (shown in yellow) aligned to the corresponding residues of ACO1 (shown in purple). C) Residues of hsIRP1 (shown in yellow) aligned to the corresponding residues of ACO2 (shown in blue). D) Residues of hsIRP1 (shown in yellow) aligned to the corresponding residues of ACO3 (shown in red). Isocitrate is shown in black and the [4Fe-4S] cluster in grey.

5.6.7 Conserved residues in [Fe-S] isomerases

A number of residues are believed to be conserved in [Fe-S] isomerases (Frishman and Hentze, 1996). These residues are also conserved in ACO1, ACO2 and ACO3 (Table 5.5; Figure 5.15).

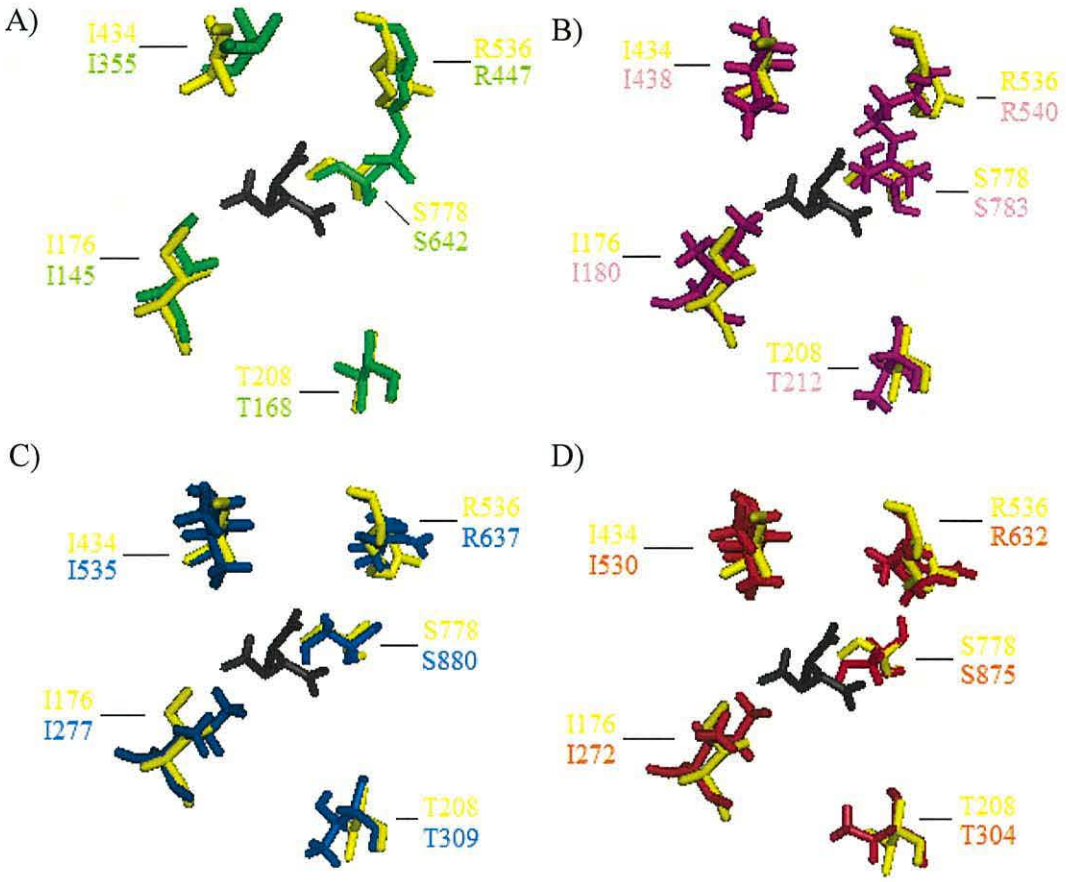


Figure 5.15: Residues of ssACO (shown in green) conserved in [Fe-S] isomerases. A) Residues of ssACO (shown in green) aligned to the corresponding residues of hsIRP1 (shown in yellow). B) Residues of hsIRP1 (shown in yellow) aligned to the corresponding residues of ACO1 (shown in purple). C) Residues of hsIRP1 (shown in yellow) aligned to the corresponding residues of ACO2 (shown in blue). D) Residues of hsIRP1 (shown in yellow) aligned to the corresponding residues of ACO3 (shown in red). Isocitrate is shown in black.

5.6.8 Supporting active site residues

The active site of aconitase is believed to be supported by a network of hydrogen bonds (Lauble *et al.*, 1992). Residues that form hydrogen bonds with the active site and residues associated with the active site were compared (Figure 5.16). Analysis revealed that all of the residues were conserved between hsIRP1, ACO1, ACO2 and ACO3. Analysis also revealed that Ala^{74} of hsIRP1 aligns to a phenylalanine residue in ssACO.

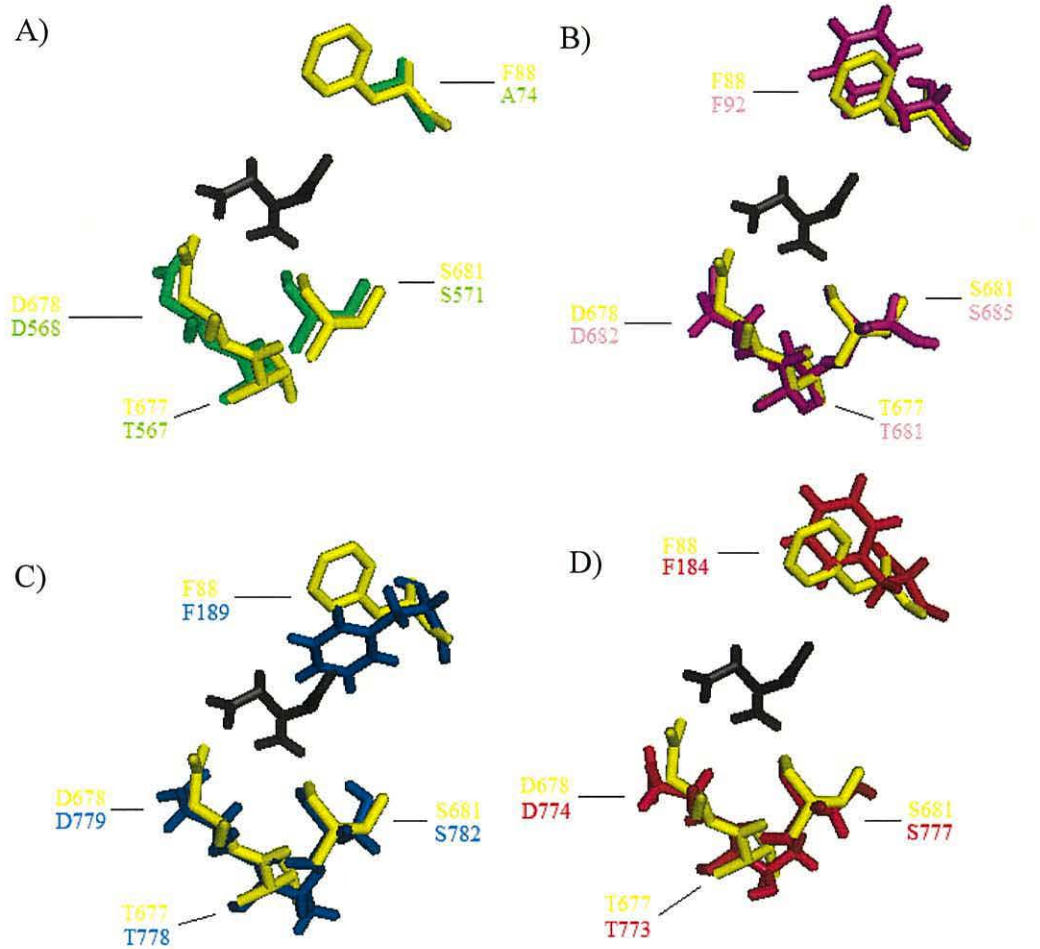


Figure 5.16: Residues of ssACO (shown in green) involved in hydrogen bonding. A) Residues of ssACO (shown in green) aligned to the corresponding residues of hsIRP1 (shown in yellow). B) Residues of hsIRP1 (shown in yellow) aligned to the corresponding residues of ACO1 (shown in purple). C) Residues of hsIRP1 (shown in yellow) aligned to the corresponding residues of ACO2 (shown in blue). D) Residues of hsIRP1 (shown in yellow) aligned to the corresponding residues of ACO3 (shown in red). Isocitrate is shown in black.

5.7 Aconitase activity from crude *Arabidopsis* extracts

The presence of three aconitase isoforms within *Arabidopsis* has been confirmed previously by Arnaud *et al.* (2007), Moeder *et al.* (2007) and within this thesis (see chapter 6). Homozygous mutants for each isoform have been identified and described in chapter 6. Aconitase activity was measured (Rose and O'Connell, 1967; section 2.3.13) in the *Arabidopsis* mutants *aco1*, *aco2* and *aco3* (see chapter 6) during germination and establishment (Figure 5.17). Total aconitase activity was measured to try to establish which isoforms were active and involved in metabolism.

Specific activity measurements of wild type seedlings showed that aconitase activity increased over the first two days. This was followed by a decline in total activity during days three and four to a constant level of activity thereafter.

The specific activity profile of *aco2* is very similar to that of wild type Col-0. The main difference observed between the profiles is the increase in activity in day 1 seedlings of *aco2* (Figure 5.17). The aconitase activity profile of *aco1* shows a reduction in specific activity for days 1, 2 and 3 compared to wild type. The largest reduction in aconitase activity compared to wild type was observed in *aco3* (Figure 5.17).

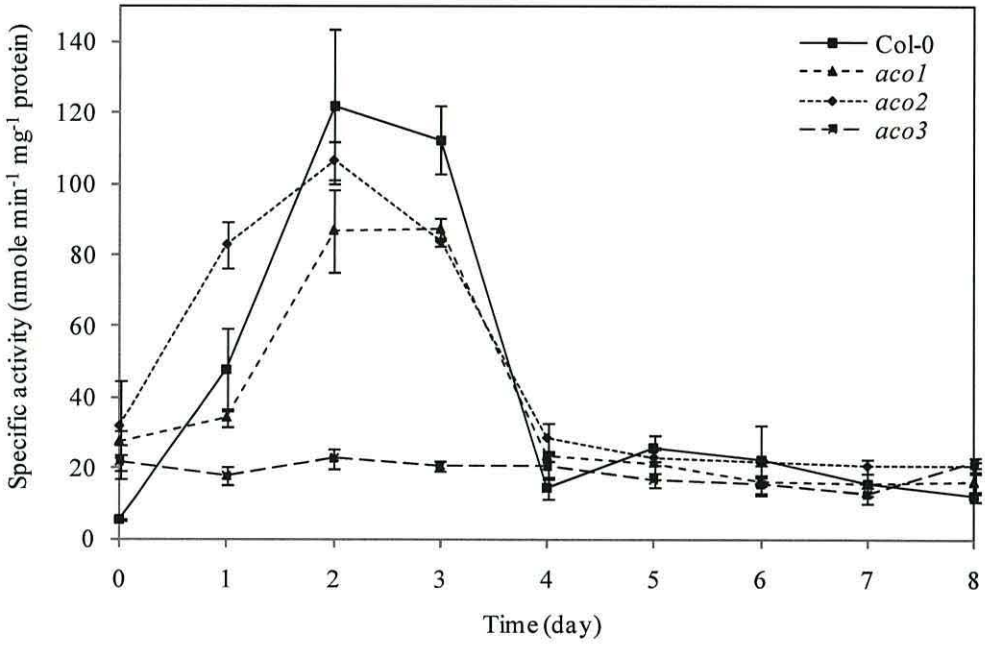


Figure 5.17: Aconitase activity in day 0 - 8 of wt Col-0, *aco1*, *aco2* and *aco3*. Aconitase activity was determined in a coupled assay modified from (Rose and O'Connell, 1967) using isocitrate dehydrogenase. The reaction was initiated by adding *cis*-aconitic acid to a final concentration of 1 mM. Error bars correspond to the standard deviation from three individual experiments.

At day 2, when aconitase activity is greatest *aco1*, *aco2* and *aco3* show an average reduction in total activity compared to Col-0 of 23, 14 and 63 % respectively (Table 5.6). The range of specific activities calculated by both *aco1* and *aco2* suggests that these enzymes have very little or no effect on overall aconitase activity. To determine if ACO1 and ACO2 are true aconitases, and are therefore capable of

contributing to total activity, the genes were successfully cloned into *Pichia pastoris* for expression analysis (see section 5.8).

The percentage reduction in aconitase activity has been determined by taking into account the variation in specific activity. The range obtained for the specific activity of Col-0 was 100 to 144 nmoles min^{-1} mg^{-1} protein. The range of specific activity calculated by both *aco1* and *aco2* suggests that these enzymes have very little effect on the overall aconitase activity. Calculations show that *aco3* has the largest effect on aconitase activity.

Table 5.6: Specific activity of aconitase at day 2 in Col-0, *aco1*, *aco2* and *aco3*.

	Specific activity range (nmole min^{-1} mg^{-1} protein)	Difference in specific activity (nmole min^{-1} mg^{-1} protein)	Average decrease in specific activity (nmole min^{-1} mg^{-1} protein)	Percentage reduction in specific activity (nmole min^{-1} mg^{-1} protein)
Col-0	100-144			
aco1	75-98	2-69	36	23
aco2	100-110	0-44	22	14
aco3	21-26	74-123	99	63

5.8 Aconitase cloning overview

In order to obtain pure protein samples of the aconitase isoforms, the genes were first sub-cloned into the *E. coli* vector pCR®2.1TOPO® (Figure 5.18) and subsequently into the *P. pastoris* vector pPiczC (Figure 5.19) (see chapter 2 for more details).

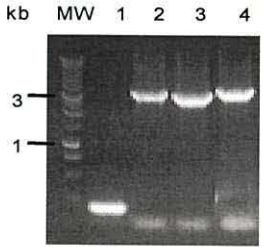


Figure 5.18: 1 % Agarose gel electrophoresis of colony PCR products with M13 forward and reverse primers of *E. coli* transformed with pCR®2.1TOPO® vector after ligation with an aconitase isoform. MW. 1 kb molecular weight marker; 1. Empty TOPO2.1 vector; 2. TOPO2.1 vector with At2g05710 insert; 3. TOPO2.1 vector with At4g35830 insert; 4. TOPO2.1 vector with At4g26970 insert.

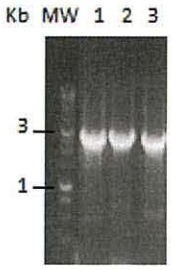


Figure 5.19: Positive identification of *E. coli* transformants with pPiczC vector containing an aconitase isoform. Gel electrophoresis after colony PCR with gene specific aconitase primers. MW. 1 kb molecular weight marker; 1. Atg05710; 2. Atg26970; 3. Atg35830.

Positive *E. coli* transformants with the pPiczC vector were linearised and integrated into *P. pastoris*. Positive *P. pastoris* lines were selected for by antibiotic resistance and PCR (Figure 5.20).

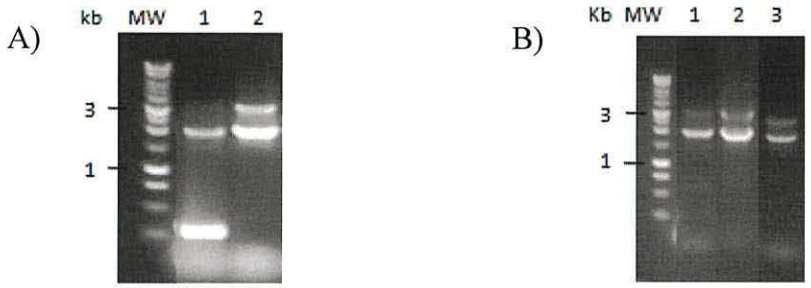


Figure 5.20: Gel electrophoresis of PCR fragments amplified from antibiotic resistant *P. pastoris* colonies. A) MW. 1 kb molecular weight marker; 1. *P. pastoris* colony transformed with pPiczC without an aconitase gene; 2. Positive transformant containing Atg05710. B) MW. molecular weight marker; 1 – 3. *P. pastoris* with Atg05710, Atg26970 and Atg35830 respectively.

5.9 Protein expression and purification of positive *P. pastoris* cells

Small scale expression and purification experiments were conducted on positively identified transformants to identify over-expressing cell lines (see materials and methods chapter 2.3.11). Due to the production of small aconitase concentrations, protein expression trials were conducted by inducing 50 ml of minimal media with 0.5 % methanol for 7 days at 28°C. Purified protein was visualised by SDS-PAGE. Visualisation of purified ACO1 and ACO2 revealed a 98 kDa protein (Figure 5.21).

The same methods were employed for ACO3 expression. However despite a number of attempts these were unsuccessful at producing the aconitase isoform ACO3, and therefore the subsequent activity measurements were unable to be conducted.

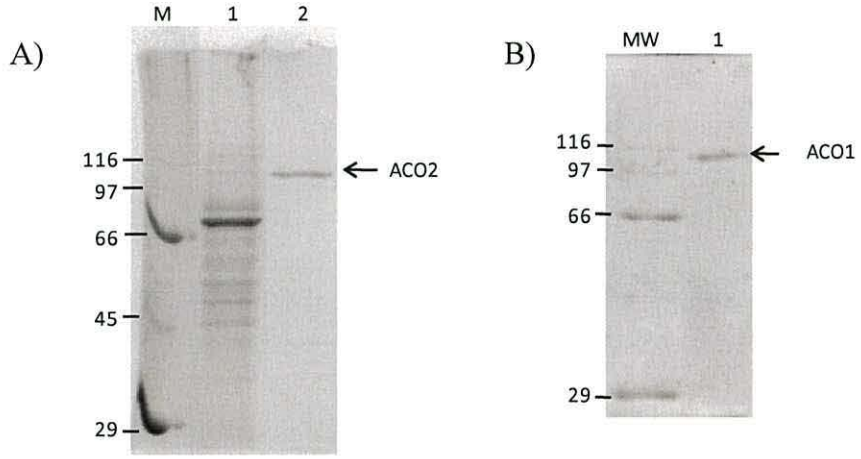


Figure 5.21: Over-expression and purification of aconitase genes ACO1 and ACO2 from *P. pastoris* x-33 with HisTrap Ni-chelating column. A) M. molecular weight marker; 1. Crude lysate loaded onto the HisTrap column; 2. purified ACO2 eluted with 100 mM imidazole buffer. B) MW. molecular weight marker; 1. purified ACO1 eluted with 100 mM imidazole buffer.

5.9.1 Confirmation of ACO1 expression in *P. pastoris*

Trypsin digests followed by MALDI MS analysis were conducted on purified protein sample of ACO1. Confirmation of ACO1 over-expression was required as activation methods had proved unsuccessful. Analysis confirmed that the over-expressed protein was aconitase (Table 5.7).

Table 5.7: Trypsin digestion followed by MALDI mass spectroscopy analysis confirmed that the over-expressed and purified protein from *P. pastoris* is aconitase from *Arabidopsis*.

	Accession	Mass	Description
1	gi |15233349	98090	Aconitate hydratase, cytoplasmic / citrate hydrolyase / aconitase (ACO) [Arabidopsis thaliana]
2	gi|599625	100755	Aconitase [Arabidopsis thaliana]
3	gi|4586021	98093	Cytoplasmic aconitate hydratase [Arabidopsis thaliana]
4	gi|30678219	108133	Aconitate hydratase, cytoplasmic, putative / citrate hydrolyase/aconitase, putative [Arabidopsis thaliana]
5	gi|125584839	115445	Hypothetical protein OsJ_008986 [Oryza sativa (japonica cultivar-group)]
6	gi|115450595	106235	Os03g0136900 [Oryza sativa (japonica cultivar-group)]
7	gi|125542317	134056	Hypothetical protein OsI_009689 [Oryza sativa (indica cultivar group)]
8	gi|34851120	98817	Putative aconitase [Prunus avium]
9	gi|162667942	94906	Predicted protein [Physcomitrella patens subsp. patens]
10	gi|1351856	97943	Aconitate hydratase, cytoplasmic (Citrate hydrolyase) (Aconitase)

5.10 Activity determination of expressed aconitases

Aconitase is responsible for the reversible reaction of citrate to isocitrate via *cis*-aconitate (Figure 5.22).

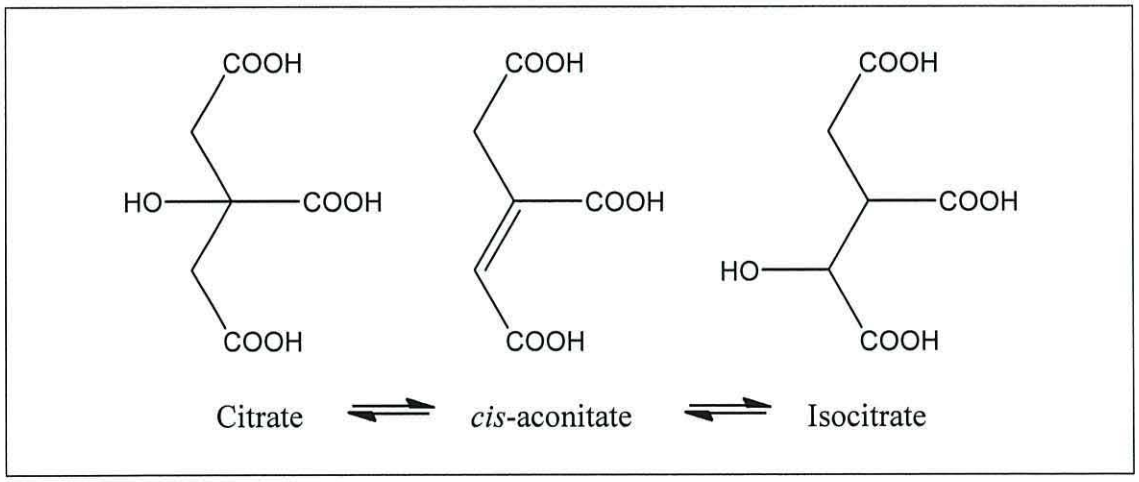


Figure 5.22: Metabolic reaction of aconitase.

Aconitase activity was measured in a coupled assay with isocitrate dehydrogenase (see materials and methods chapter 2.3.13). Previous research shows that isocitrate

dehydrogenase is inhibited by citrate (Henson and Cleland, 1967). Enzymatic analysis of isocitrate dehydrogenase both in the presence and absence of citrate was conducted to corroborate the findings of Henson and Cleland, 1967. Results clearly showed that the presence of 1 mM citrate reduces that rate of isocitrate dehydrogenase (Figure 5.23). For this reason aconitase assays were conducted with 1 mM *cis*-aconitate.

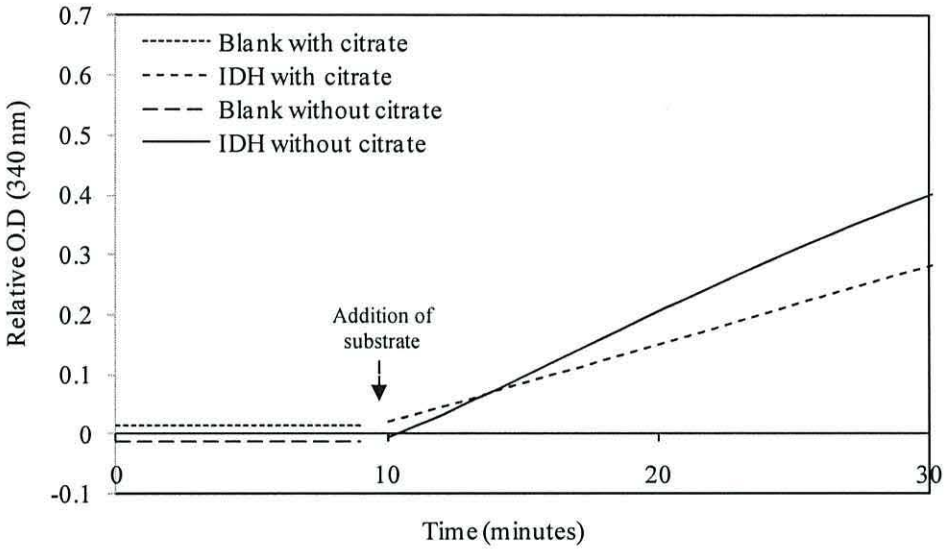


Figure 5.23: Rate of reaction trace of isocitrate dehydrogenase (IDH) in the presence and absence of 1 mM citrate.

5.11 Porcine heart aconitase activation

Porcine heart aconitase was purchased from Sigma-Aldrich to test methods of aconitase activation. Porcine heart aconitase is responsible for the conversion of citrate to isocitrate (Figure 5.22) in the mitochondria as part of the TCA cycle.

A successful activation method was identified for porcine heart aconitase (Morrison, 1954). Half a ml of a 1.4 mg ml^{-1} solution was prepared in 1 M Tris, pH 7.4 and activated for 1 hr with 0.02 mM ferrous ammonium sulphate and 2.3 mM cysteine on ice. Activated aconitase was added to a 1 ml assay and activity determined in a coupled reaction with isocitrate dehydrogenase (Rose and O'Connell, 1967). The reaction was conducted with 1 mM *cis*-aconitate rather than citrate, as the latter inhibits the activity of isocitrate dehydrogenase (Henson and Cleland, 1967; section

5.10). The assay was buffered with 0.1 M triethanolamine, pH 7.4 and the production of NADPH by isocitrate dehydrogenase was monitored at 340 nm. The specific activity was determined to be 4.8 nmoles min^{-1} mg^{-1} protein (Figure 5.24).

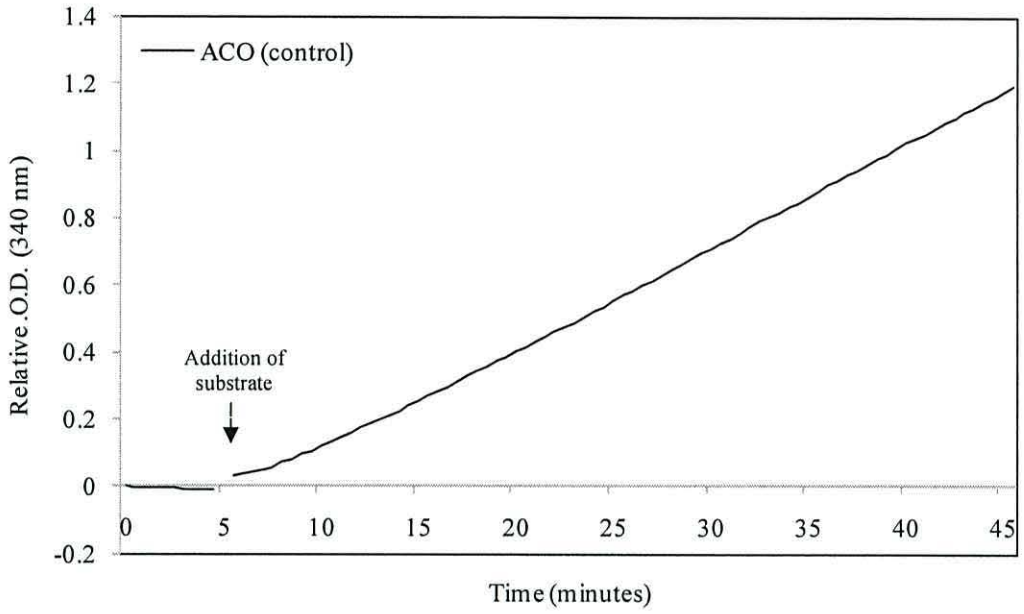


Figure 5.24: Rate of reaction trace of porcine heart aconitase used as a control to test methods of activation. A 1 ml assay was incubated at room temperature with 0.7 mg of reactivated ACO and the absorbance measured at 340 nm over the first five minutes before the addition of *cis*-aconitate which initiated the reaction.

5.12 Activation of ACO1

ACO1 was assayed directly after purification to determine if a method of activation was required. However this resulted in an inactive protein (Figure 5.25).

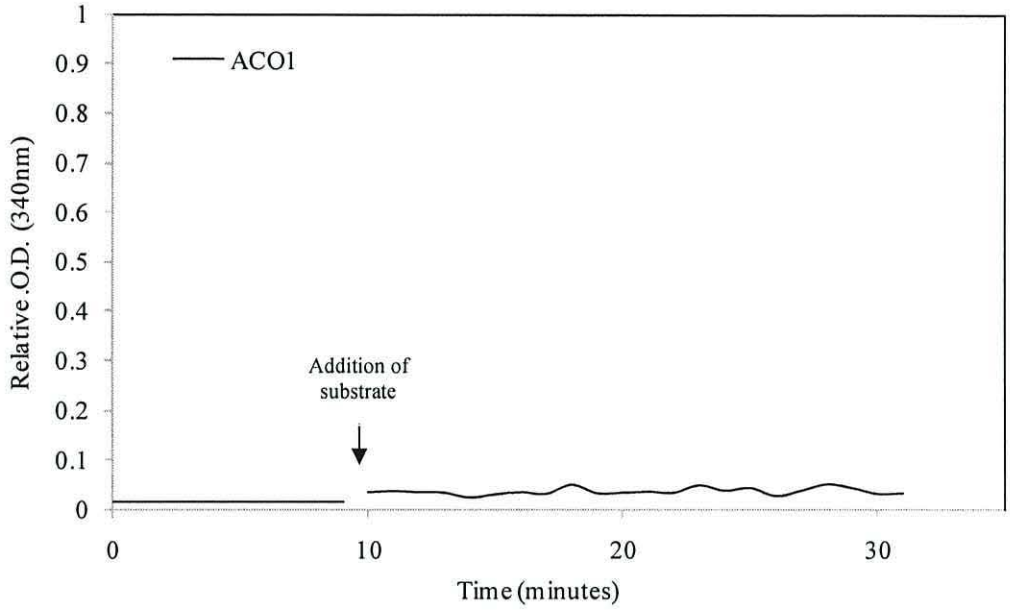


Figure 5.25: Rate of reaction trace of ACO1 expressed on minimal media and purified directly. 6 µg ACO1 were added to a 1 ml assay and absorbance at 340 nm for 10 min before the addition of *cis*-aconitate to initiate the reaction.

A successful method of ACO1 activation (Method 1) included growing cells in minimal media at 100 rpm at 28°C. The cell culture was induced with methanol every 24 hrs and harvested after 7 days. Proteins were then purified with a HisTrap column, buffer exchanged on a pd10 column with 1M Tris pH 7.4 and the resultant solution concentrated at 4°C on a viva spin column to 0.22 mg ml^{-1}. The concentrated protein was activated using the method described in section 5.11 for porcine heart aconitase at 4°C. The specific activity of ACO1 was determined as 160 nmoles min^{-1} mg^{-1} protein (Table 5.8).

Another successful method (Method 2) included growing cell cultures at 100 rpm at 28°C for 7 days with the addition of 0.06 mM ferrous ammonium sulphate at 48, 96 and 144 hr and induction with 0.5 % methanol every 24 hrs. Cells from 3.5 L of minimal media were harvested, lysed and purified as previously stated (see materials and methods chapter 2.3). The specific activity calculated using this method was more than double that obtained when the protein was purified then activated (Figure 5.26; Table 5.8).

5.13 Activation of ACO2

The same techniques used to activate ACO1 (Methods 1 and 2) were also implemented with cell cultures expressing ACO2. Attempts to activate ACO2 after purification by incubation with ferrous ammonium sulphate and cysteine were unsuccessful. Activation by addition of ferrous ammonium sulphate to the growth medium was more successful, producing specific activities of 4-5 nmoles min^{-1} mg^{-1} (Figure 5.26; Table 5.8).

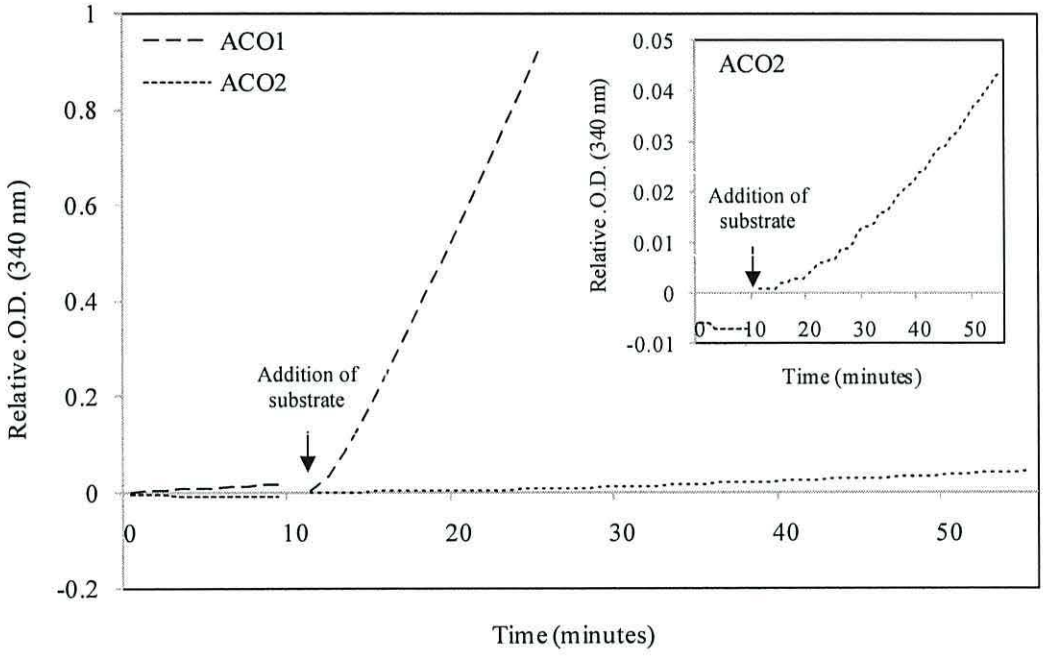


Figure 5.26: Rate of reaction trace of ACO1 and ACO2. Protein was over-expressed in minimal media using method 2. Direct purification resulted in active ACO1 and ACO2. 0.032 mg ACO1 and 0.046 mg ACO2 were added to a 1 ml assay and absorbance measured at 340 nm for 10 min before the addition of *cis*-aconitate to initiate the reaction. Insert: enlarged absorbance profile at 340 nm of ACO2.

Porcine heart aconitase was used to identify a successful method for activating aconitase. Whist method 1 successfully activated ACO1, enzymatic activity was not detected for ACO2 (Table 5.8). Further experimental analysis revealed that method 2 can successfully activate both ACO1 and ACO2. As porcine heart aconitase was purchased in a purified form, method 2 was not suitable for testing.

Table 5.8: Specific activity of Porcine heart aconitase (Sigma-Aldrich) and purified ACO2 (At4g26970) and ACO1 (At4g35830) *Arabidopsis* genes from *P. pastoris*. Method 1: Protein activated after purification with Iron and Cysteine. Method 2: Protein activated during induction with the inclusion of Iron in the growth medium.

	Specific activity (nmoles min^{-1} mg^{-1}) Method 1	Specific activity (nmoles min^{-1} mg^{-1}) Method 2
Porcine heart aconitase (Sigma)	5 3 3	Not applicable
Purified ACO2 (At4g26970)	0	4 4 5
Purified ACO1 (At4g35830)	160	572 442 331

5.14 Discussion

Three *Arabidopsis* aconitase isoforms, *aco1*, *aco2* and *aco3* have been identified and described both within this chapter and within chapter 6. These findings are in agreement with previous research conducted by Arnaud *et al.* (2007) and Moeder *et al.* (2007). Sequence alignment analysis showed that the aconitase isoforms were very similar in amino acid sequence to the mammalian IRP1 (hsIRP1) and the *Nicotiana tabacum* ACO1 (NtACO1). Initial observations regarding the sequence similarity between the *Arabidopsis* aconitase isoforms and mammalian IRP1 could suggest that the aconitase isoforms are able to bind to iron response elements (IREs) and regulate cellular iron. However, previous research has shown that the cytosolic *Arabidopsis* aconitase isoform is not converted into an IRP and does not regulate iron homeostasis (Arnaud *et al.*, 2007). Despite the conservation of IRP1 mRNA binding residues (Navarre *et al.*, 2000) it has also been suggested that NtACO1 does not bind to RNA (Gourley *et al.*, 2003).

5.14.1 Location of aconitase expression

The three *Arabidopsis* aconitase isoforms share greater than 80 % similarity and 70 % identity with each other. Sequence alignment of the three isoforms shows that the greatest difference between the sequences is the addition of an N-terminal leader

sequence to ACO2 and ACO3 (Figure 5.2). The function of the ACO2 and ACO3 leader sequence is currently unknown. One possible explanation for this could be transport into the mitochondria. Previous research with rat mitochondrial L-alanine:glyoxylate aminotransferase 1 (Takada *et al.,* 1990), human DNA glycosylase (Takao *et al.,* 1998), calpain1 (Badugu *et al.,* 2008) and mannheimia haemolytica leukotoxin (Kisiela *et al.,* 2010), to name just a few, show that the protein's N-terminal region is important for mitochondrial targeting.

Protein prediction analysis was conducted with online severs to determine the location of ACO1, ACO2 and ACO3 (Table 5.2). Analysis conducted with MitoProt II predicted that both ACO2 and ACO3 are mitochondrial proteins. A mitochondrial location was also predicted for ACO2 by the online servers Predotar and iPSORT. Previous research has suggested that both ACO2 and ACO3 are located in the mitochondria by activity analysis (Arnaud *et al.,* 2007; Bernard *et al.,* 2009) and by protein identification (Millar *et al.,* 2001; Sweetlove *et al.,* 2002). ACO1 has also been identified in the mitochondrial proteome (Kruft *et al.,* 2001), which is in contrast to other activity analysis (Arnaud *et al.,* 2007; Bernaud *et al.,* 2009) that concluded ACO1 to be located within the cytosol. If these latter studies are correct, a cytosolic location of ACO1 could explain why this isoform does not contain an N-terminal leader sequence.

Arnaud *et al.* (2007) also noticed a reduction in aconitase activity in *aco1*, *aco2* and *aco3* within the mitochondrial fraction. The reduction in aconitase activity within both the mitochondrial and cytosolic fraction could suggest a dual location (see section 6.9.2). Further work could include the use of fluorescent proteins to determine the location of aconitase expression. A fusion protein could be constructed and integrated into the *Arabidopsis* genome. The resultant plants could then be crossed to directly compare the location of expression for each isoform. Fluorescent tags can be added to either the C-terminus or N-terminus of a protein. However, previous work suggests that fluorescent protein tagging to the C-terminus would be more successful. Tian *et al.* (2004) have suggested that N-terminal fusions are likely to obscure the mitochondrial transit peptide.

The cytosolic *Homo sapiens* IRP1 (hsIRP1) is a dual functioning protein. In addition to aconitase function, hsIRP1 is also able to bind to IREs (Hentze and Argos, 1991; Kaptain *et al.,* 1991; Klausner and Rouault, 1993). A nuclear gene encoding mitochondrial aconitase 2 (hsACO2) has been identified that functions solely as aconitase (Mirel *et al.,* 1998). A second IRP2 has also been identified in *Homo sapiens* that does not have aconitase function (Guo *et al.,* 1995a). These three proteins (hsIRP1, hsACO2 and IRP2) were aligned with *Arabidopsis* ACO1, ACO2 and ACO3 to aid in determining the purpose of the N-terminal leader sequence of ACO2 and ACO3. It was hypothesised that if the leader sequence contains a mitochondrial targeting signal then hsACO2 should also contain an N-terminal leader sequence. The sequence alignment of hsIRP1, hsACO2, hsIRP2, ACO1, ACO2 and ACO3 shows that only ACO2 and ACO3 have an N-terminal leader sequence. This would suggest that ACO2 and ACO3's N-terminal leader sequences may not be involved in mitochondrial import and therefore may have a currently unknown function.

Sequence alignment analysis revealed that IRP2 contains an additional 73 amino acid insert that is not present in ACO1, ACO2 or ACO3. This additional sequence is believed to form a surface loop in IRP2 (Guo *et al.,* 1995a). Mutagenesis of this additional sequence in IRP2 suggests that this section could be involved in IRE binding (Butt *et al.,* 1996). The absence of this sequence from functioning aconitases (hsACO2, ntACO1, ssACO) could suggest that it prevents aconitase activity. The absence of this sequence in ACO1, ACO2 and ACO3 would therefore suggest that they are more likely to function as aconitase than IRP2.

Sequence alignment analysis (Figure 5.3) highlighted the insertion sequences observed by Dupuy *et al.* (2006) between *H. sapiens* cytosolic (hsIRP1) and mitochondrial aconitase (hsACO2). These additional sequences are believed to result in the different structure, and therefore different functions of these proteins (Dupuy *et al.,* 2006). Comparison of the amino acid sequences showed that these additional sequences are also present in ACO1, ACO2 and ACO3. This suggests that ACO1, ACO2 and ACO3 should be able to bind to mRNA. Previous research has shown that ACO1 is able to bind to the 5' UTR of the *Arabidopsis* CuZn superoxide

dismutase 2 (Moeder *et al.,* 2007). However, the mechanism of action and exact sequences involved are currently unknown.

5.14.2 Phyre analysis of ACO1, ACO2 and ACO3

Phyre analysis of ACO1, ACO2 and ACO3 showed that the amino acid sequences of the *Arabidopsis* aconitase isoforms shared the greatest amino acid sequence identity (>60 %) with the cytosolic hsIRP1. Additional crystallographic structures that shared between 18 and 31 % identity with ACO1, ACO2 and ACO3 included the mitochondrial aconitase from *Bos taurus* (1aco.pdb), isopropylmalate isomerase from *Pyrococcus horikoshii* (1v7l.pdb) and aconitase B from *E. coli* (115j.pdb). The slight differences observed with the percentage identity between the *Arabidopsis* aconitase isoforms and the crystallographic structures may be due to the observed amino acid differences (Table 5.3; Figure 5.2) and the absence of an N-terminal leader sequence in ACO1.

Phylogenetic analysis of the aconitase family grouped *Arabidopsis* aconitase with IRP1 (Gruer *et al.,* 1997). In addition to the plant aconitase, other members of the group included IRP2 and the bacterial aconitases. Gruer *et al.* (1997) organised thirty five members of the aconitase family into five groups based on amino acid sequence comparisons. These included IRP-aconitase A, aconitase B, mitochondrial aconitase, homoaconitase and isopropylmalate isomerase. The addition of *Arabidopsis* aconitase into the IRP-aconitase A group with IRP shows that all three isoforms resemble cytosolic aconitase more than mitochondrial aconitase. This is surprising considering that all three proteins have been identified within the mitochondria (Kruft *et al.,* 2001; Millar *et al.,* 2001; Sweetlove *et al.,* 2002).

5.14.3 Homology modelling of ACO1, ACO2 and ACO3

The amino acid sequences of ACO1, ACO2 and ACO3 were all independently submitted to the online homology modelling server I-TASSER (Zhang, 2008; Zhang, 2009; Roy *et al.,* 2010). Homology modelling was conducted with this online server due to its recent success in CASP 7 (Zhou, 2007) and CASP 8 (Zhang, 2009) analysis (see section 4.5.1). Sequence alignment (Figure 5.2) and blast analysis (Table 5.1) demonstrate that all three aconitase isoforms share a large degree of

sequence identity. For these reasons it is not surprising that I-TASSER modelled all three isoforms to the template crystallographic structure of hsIRP1 (2b3x.pdb).

The hsIRP1 was crystallised in its aconitase form by Dupuy *et al.* (2006). Under normal cellular conditions the protein is present as aconitase. The crystallised structure of aconitase is comprised of four domains. A linker segment on the proteins surface joins domains 3 and 4 together (Dupuy *et al.*, 2006). The [4Fe-4S] cluster is situated close to the centre of the protein at the interface of the four domains. The homology models of ACO1, ACO2 and ACO3 all appear to adopt a similar four domain structure.

Research conducted within this chapter has focused on sequence and structural analysis of the aconitase isoforms to determine if they have the potential to function as aconitase. Amino acid sequence analysis of the *Arabidopsis* aconitase isoforms has been previously compared by Arnaud *et al.* (2007). Where residues involved in RNA loop interactions of IRP1 were compared. Analysis revealed that only sixteen out of the twenty two residues were conserved between IRP1, IRP2, ACO1, ACO2 and ACO3. Additional research showed that aconitase does not bind to ferritin mRNA and therefore does not regulate cellular iron levels (Arnaud *et al.*, 2007). Interestingly, research conducted by Moeder *et al.* (2007) showed that the *Arabidopsis* cytosolic aconitase is able to bind to the 5' UTR of the CuZn superoxide dismutase 2 transcript (CSD2). However, the method and amino acid sequences that allow aconitase to bind to CSD2 are currently unknown. For these reasons residues involved in mRNA binding were not compared within this chapter.

5.14.4 Structural comparison of the homology models of ACO1, ACO2 and ACO3

The homology models of ACO1, ACO2 and ACO3 were compared to the template structure of hsIRP1 and the mitochondrial aconitase ssACO. Structural comparisons were made to both a cytosolic (hsIRP1) and mitochondrial (ssACO) isoform of aconitase to ensure that important residues involved in catalysis were compared in ACO1, ACO2 and ACO3.

The active site cleft of mitochondrial aconitase contains seven residues involved in substrate recognition, five residues involved in co-ordination of the [Fe-S] cluster and seven residues involved in catalysis (Lauble *et al.,* 1992; Gruer *et al.,* 1997). All of these important residues have been aligned and compared to ACO1, ACO2 and ACO3 within this chapter.

In addition, the five residues that are conserved within all of the [Fe-S] isomerases, except IRP2 (Frishman and Hentze, 1996), have been compared to ACO1, ACO2 and ACO3. These residues in ssACO are Ile^{145}, Thr^{168}, Ile^{355}, Arg^{447} and Ser^{642} and are substituted in IRP2 by methionine, isoleucine, valine, lysine and asparagine, respectively (Guo *et al.,* 1995a; Frishman and Hentze, 1996). Sequence analysis showed that all five of these residues are conserved in ACO1, ACO2 and ACO3 (Table 5.5; Figure 5.15). It has been suggested that the conservation of these residues could indicate an important role in aconitase function (Frishman and Hentze, 1996).

Visual comparison of the aconitase homology models showed that the greatest difference between them was the additional leader sequence of ACO2 and ACO3. This N-terminal leader sequence was not present in the template structure of hsIRP1. Therefore the leader sequences of both ACO2 and ACO3 are not modelled to a template structure, and consequently adopt a different conformation in the two models.

5.14.4.1 The N-terminal sequence of aconitase

The N-terminal sequence of hsIRP1 was compared to that of mitochondrial aconitase by Dupuy *et al.* (2006) who noticed a structural fold difference. Dupuy *et al.* (2006) also noted that hsIRP1 contains a number of insertion sequences when compared to ssACO. The presence of these insertions suggests an explanation for the shape differences observed and the different RNA binding properties of these two proteins (Dupuy *et al.,* 2006).

The N-terminal sequences of ACO1, ACO2 and ACO3 were compared to the N-terminal regions of hsIRP and ssACO. Visual comparison showed that the structures of ACO1, ACO2 and ACO3 are similar to that of hsIRP1. However, this is not surprising as the aconitase models are based on the crystallographic structure of

hsIRP1. Alignment of the N-terminal region of hsIRP to ACO1, ACO2 and ACO3 showed that hsIRP1 and ACO1 share the greatest number of residues. However, this includes only six out of the thirty four N-terminal region residues. Whilst this region shows greater conservation between ACO1, ACO2 and ACO3, only twenty one residues are identical. The variability observed within this region suggests that these residues are not essential for aconitase activity.

The N-terminal sequences of ACO2 and ACO3 contain approximately 100 additional amino acid residues compared to ACO1. This N-terminal leader sequence is also absent from the template structure of hsIRP1, resulting in a different orientation in the ACO2 and ACO3 models. Whilst the purpose of this N-terminal sequence is unknown, it is believed to be required for mitochondrial import. However, a leader sequence has not been observed in the mitochondrial isoform hsACO2 (Figure 5.3). Despite the presence of a leader sequence in both ACO2 and ACO3 the sequences are not identical. If the sequences were responsible for protein import then identical sequences might be expected.

5.14.4.2 The aconitase linker sequence

The flexible linker sequence of aconitase joins domains 3 and 4 together. Movement of this linker sequence results in a conformational change in shape allowing RNA to bind to the protein (Philpott *et al.,* 1994; Walden *et al.,* 2006). Comparison of the linker sequence between mitochondrial and cytosolic aconitase showed that the mitochondrial linker sequence was about twenty residues shorter (Dupuy *et al.,* 2006). The presence of a smaller linker sequence could help to explain why mitochondrial aconitase is unable to bind to RNA. Sequence and structural analysis of ACO1, ACO2 and ACO3 shows a similar, larger linker sequence to that of hsIRP1 (Figure 5.9). The presence of a longer linker sequence could suggest that these proteins are able to bind to RNA. Whilst research has shown that ACO1 is unable to bind to ferritin mRNA and regulate cellular iron, ACO1 is able to bind to CuZn superoxide dismutase 2 (Arnaud *et al.,* 2007; Moeder *et al.,* 2007). The dual functionality of aconitase is not restricted to mRNA binding. In addition to involvement in the TCA cycle, yeast mitochondrial aconitase has been shown to have an additional role in mitochondrial DNA maintenance by binding to both

dsDNA and ssDNA (Chen *et al.,* 2005; 2007). Further work could be conducted to determine whether *Arabidopsis* aconitase is also able to bind to DNA.

5.14.4.3 Residues involved in substrate recognition

Lauble *et al.* (1992) state that the seven residues Gln^{72}, Ser^{166}, Arg^{447}, Arg^{452}, Arg^{580}, Arg^{644} and Ser^{643} are all involved in substrate recognition. Sequence alignment analysis showed that all of these residues are conserved in hsIRP1, ACO1, ACO2 and ACO3. Despite the sequence similarity, structural observations show a different orientation for a number of these residues. The positioning of residues Arg^{536} and Arg^{541} of hsIRP1 are different to those observed in the crystal structure of ssACO (residues Arg^{447} and Arg^{542} respectively). The orientation of these residues also appears to be different in the aconitase models. The orientation of the corresponding residue to Arg^{536} in hsIRP1 resembles that of hsIRP1 in ACO2 and ACO3 but ssACO in ACO1. The different orientations of these residues could suggest an explanation for the differences observed in specific activity between ACO1 and ACO2. The orientation of the corresponding residue to Arg^{541} of hsIRP1 resembles that of hsIRP1 in ACO3 but ssACO in ACO1 and ACO2. The orientations of Arg^{713} and Arg^{780} of hsIRP1 are all very similar in ssACO, ACO1 and ACO2, but not ACO3. Further analysis could include determination of the specific activity of purified ACO3. For this, a positive colony of *P. pastoris* over-expressing ACO3 would need to be identified.

The Arg^{447} residue of ssACO is believed to form a salt bridge with Glu^{262} (Lauble *et al.,* 1992). This interaction contributes to a network of hydrogen bonds between side chains in the active site, and surrounding protein atoms that support substrate binding (Lauble *et al.,* 1992). This glutamate residue is also conserved in hsIRP1, ACO1, ACO2 and ACO3 (Table 5.5; Figure 5.12). The Arg^{452} residue of ssACO is also involved in supporting substrate binding through the formation of hydrogen bonds with Asp^{568} (Lauble *et al.,* 1992). In addition, Arg^{452} of ssACO also forms a hydrogen bond with the cluster ligating residue Cys^{424}. Sequence alignment analysis showed that all of these residues are conserved in hsIRP1, ACO1, ACO2 and ACO3.

Arg^{580} of ssACO is believed to form hydrogen bonds with Ala^{74}. This alanine residue aligns to a phenylalanine residue in hsIRP1, ACO1, ACO2 and ACO3. As

both ssACO and hsIRP1 are able to function as aconitase the presence of either an alanine or phenylalanine residue at this position does not appear to affect aconitase activity. However, a phenylalanine residue has also been identified at this position in hsIRP2 (Dupuy *et al.,* 2006), suggesting an involvement in RNA binding rather than aconitase activity. The position of the corresponding residue to Ala^{74} of ssACO shows a different orientation in ACO2. The difference in orientation could suggest a reason for the different specific activity observed between ACO1 and ACO2.

Additional residues believed to form hydrogen bonds to the substrate of ssACO include Ser^{166} and Ser^{643} (Lauble *et al.,* 1992). The Arg^{644} residue of ssACO is believed to form bonds with Asp^{568}, Ser^{571} and Ser^{642} (Lauble *et al.,* 1992). Contacts are also formed between the carboxylate of Asp^{165} and Gln^{72}. In addition, residues Gln^{72}, Arg^{580}, Ser^{643} and Arg^{644} are also believed to bind to sulphate in the absence of substrate (Lauble *et al.,* 1992). Sequence analysis revealed that these residues are conserved in ACO1, ACO2 and ACO3.

5.14.4.4 Aconitase cluster ligation

Three cysteine residues (437, 503 and 506) have been predicted to ligate the [Fe-S] cluster in hsIRP1 (Philpott *et al.,* 1994). Mutational analysis of these residues to serine showed that they are all required for aconitase activity (Philpott *et al.,* 1994). The corresponding residues in ssACO have also been shown to coordinate the [3Fe-4S] cluster of mitochondrial aconitase (Lauble *et al.,* 1992). Previous research suggests that Cys^{358} also forms a hydrogen bond with Asn^{258} (Lauble *et al.,* 1992). An additional asparagine residue (Asn^{446}) is also believed to interact with the cluster. Sequence and structural comparisons revealed that all of these cysteine residues are conserved between ACO1, ACO2 and ACO3 (Figure 5.13). The conservation of these residues strongly suggests that the *Arabidopsis* aconitase isoforms are able to form a [4Fe-4S] centre and therefore have the potential to function as aconitase.

5.14.4.5 Catalytic residues of aconitase

All of catalytic residues listed by Lauble *et al.* (1992) for ssACO are conserved in hsIRP1, ACO1, ACO2 and ACO3. The conservation of these residues suggests that ACO1, ACO2 and ACO3 have the potential to function as aconitase.

Residue Ser^{778} of hsIRP1 corresponds to the critical catalytic serine in mitochondrial aconitase. Mutagenesis of Ser^{778} to alanine demonstrated that this residue is important for enzymatic activity (Zheng *et al.,* 1992). Sequence alignment analysis showed that this serine residue was conserved in ssACO, ACO1, ACO2 and ACO3. Ser^{778} of ssACO was substituted with asparagine in hsIRP2 (Dupuy *et al.,* 2006). Dupuy *et al.* (2006) suggested that the replacement of Ser^{778} and the absence of a [Fe-S] cluster could be sufficient to explain the lack of aconitase activity in hsIRP2. The [Fe-S] cluster is essential for aconitase activity as citrate and isocitrate bind to the fourth iron during catalysis (Lauble *et al.,* 1992; Beinert and Kennedy, 1993). The conservation of serine in the *Arabidopsis* aconitase isoforms could suggest that the isoforms are able to function as aconitase. Further work could include investigation of the [Fe-S] cluster of ACO1, ACO2 and ACO3 by UV spectroscopy, EPR or EXAFS to confirm its presence.

The corresponding residue to Ser^{778} of hsIRP1 is Ser^{642} in ssACO. This residue is believed to interact with Thr^{567} in ssACO. Other interactions are believed to be made between Asp^{100} and His^{101}, Asp^{165} and His^{147}, and His^{167} and Glu^{262} in ssACO (Lauble *et al.,* 1992). Additional residues that are stated as essential for aconitase activity include His^{126} and His^{207} of hsIRP1 (Gegout *et al.,* 1999). Sequence analysis reveals that all of these residues are conserved in ACO1, ACO2 and ACO3. Comparison of the homology models of ACO1, ACO2 and ACO3 shows structural differences between the models. Residues that are orientated differently in ACO1 include Asp^{129}, His^{130} and Asp^{209}. Residues in ACO2 that show a different configuration include Asp^{226}, His^{279} and Asp^{306}. In the ACO3 model residues Asp^{221}, His^{222}, His^{274}, Asp^{301} and His^{303} show a different orientation compared to the crystallographic structure of hsIRP1 and ssACO.

5.14.4.6 Conserved residues in [Fe-S] isomerase

In addition to important active site residues, residues Ile^{145}, Thr^{168}, Ile^{355}, Arg^{447} and Ser^{642} of ssACO are also believed to be conserved in [Fe-S] isomerases (Frishman and Hentze, 1996). Sequence alignment analysis showed that all three of these residues are conserved in ACO1, ACO2 and ACO3. Structural analysis shows that these residues also occupy a similar orientation in ssACO, hsIRP1, ACO1, ACO2 and ACO3. Ile^{145} of ssACO is believed to form hydrophobic contacts with the [Fe-S]

cluster. Thr^{168} forms a hydrogen bond with the catalytic residue Asp^{165}. It has been hypothesised that these five residues are important for aconitase function as they are not conserved in hsIRP2 (Frishman and Hentze, 1996). The conservation of these residues in ACO1, ACO2 and ACO3 suggests that these proteins have the potential to function as aconitase.

5.14.4.7 Summary of aconitase homology modelling

Sequence analysis and homology modelling was conducted with ACO1, ACO2 and ACO3. Analysis shows that the *Arabidopsis* aconitase isoforms share a large percentage of sequence conservation with both hsIRP1 and ssACO. The conservation of essential catalytic residues suggests that ACO1, ACO2 and ACO3 are all able to function as aconitase.

5.14.5 Aconitase activity from *Arabidopsis* extracts

Aconitase activity measurements (Rose and O'Connell, 1967; section 2.3.13) were conducted with *Arabidopsis* seedling extracts during germination and establishment. Activity measurements were conducted to try to determine which isoforms are active and therefore possess aconitase activity.

Specific activity measurements revealed that aconitase activity of wild type seedlings increased over the first two days. This increase in aconitase activity follows a similar pattern observed for the glyoxylate cycle enzymes isocitrate lyase and malate synthase (Eastmond *et al.*, 2000; Cornah *et al.*, 2004). During this growth stage stored triacylglycerol is degraded through β-oxidation, the glyoxylate cycle and gluconeogenesis (Eastmond *et al.*, 2000; Cornah and Smith, 2002). Aconitase activity was subsequently observed to fall after day two. During this stage of development seedling metabolism begins to change from heterotrophic to autotrophic.

The specific activity profile of *aco2* is similar to that of Col-0. The main difference is observed during day 1, where aconitase seedling activity is visibly higher in *aco2* (Figure 5.17). The activity profile of *aco1* shows a reduction in specific activity during days 1, 2 and 3 compared to wild type. The knockout mutant *aco3* shows the largest reduction in aconitase activity over the whole 8 days. This observation

suggests that ACO3 makes the greatest contribution to total aconitase activity during germination (see chapter 6). The small decrease in *aco1* and *aco2* could be due to sample variation and therefore further analysis is required to determine if ACO1 and ACO2 have aconitase activity.

5.14.6 Protein expression of ACO1, ACO2 and ACO3 in *P. pastoris*

The *Arabidopsis* aconitase isoforms were cloned and purified from *P. pastoris* to determine which of the aconitase isoforms have aconitase activity. Purified protein was visualised on an 8 % agarose gel. Results revealed that both ACO1 and ACO2 to be approximately 98 kDa. Previous research shows that IRP from mice and *H. sapiens* (Festa *et al.*, 2000; Dupuy *et al.*, 2005) and aconitase from tomato, pumpkin and *Trypanosoma brucei* (Hayashi *et al.*, 1995; Saas *et al.*, 2000; Carrari *et al.*, 2003) are all around 98 kDa. This is in contrast to mitochondrial aconitase that has been isolated as an 83 kDa protein (Beinert and Kennedy, 1993). Klauser and Rouault (1993) suggest that the difference in size observed between mitochondrial and cytosolic aconitase resulted from multiple deletions. Klauser and Rouault (1993) also suggest that cytosolic aconitase evolved quite early on from prokaryotic precursors.

5.14.7 Conformation of ACO1 expression in *P. pastoris*

The over-expression and purification of ACO1 from *P. pastoris* was confirmed by trypsin digest and MALDI MS. Purified ACO1 was subjected to sequencing analysis to confirm the presence of aconitase. This was required as initial aconitase activation experiments proved unsuccessful. Analysis of the digested protein strongly suggested that aconitase was being purified as all the predicted proteins were aconitases (Table 5.7).

5.14.8 Activity determination of expressed aconitases

Enzymatic activity of aconitase was conducted in a coupled assay with isocitrate dehydrogenase (IDH). Prior to determining the enzymatic activity of aconitase enzymatic assays were conducted with IDH. These were conducted to determine the effect of citrate on IDH activity as previous research has shown that IDH is inhibited by citrate (Henson and Cleland, 1967). Enzymatic analysis confirmed that the presence of citrate reduced the rate of IDH (Figure 5.23). For this reason, *cis*-

aconitate was used as a substrate for enzymatic analysis of aconitase throughout this chapter.

5.14.9 Porcine heart aconitase activation

Porcine heart aconitase, purchased from Sigma-Aldrich, was used as a control to identify a successful method of aconitase activation. Aconitase was activated by incubation with cysteine and iron on ice to reconstruct the [4Fe-4S] cluster. Enzymatic analysis, conducted with *cis*-aconitate produced specific activities of between 3-5 nmoles min^{-1} mg^{-1} protein.

5.14.10 Activation of ACO1

Purified ACO1 was activated in a similar manner to porcine heart aconitase using method 1 (see section 5.11). Prior to activation the purified protein was buffer exchanged to remove phosphate buffer. Dickman and Cloutier (1950) reported that if phosphate is present before the addition of iron then activation is limited. The rapid inactivation of aconitase in the absence of substrate has also been demonstrated by Morrison (1954). Under these experimental conditions the specific activity of ACO1 was calculated as 160 nmoles min^{-1} mg^{-1} protein. This is a higher result than that obtained for mitochondrial porcine heart aconitase (Table 5.8). Though less than the activity found for ACO1 using an alternative activation method (Method 2).

Additional samples of purified ACO1 were activated using method 2 (see section 5.12) with the addition of ferrous ammonium sulphate to the cell culture. Ferrous ammonium sulphate was added to the growth medium to ensure that any aconitase produced was not iron-starved. This constant addition of iron during cellular growth allowed the [4Fe-4S] cluster of aconitase to be formed. The specific activity calculated for ACO1 using this method of activation was more than double that obtained using method 1. This large increase in specific activity shows that the method of activation during protein folding within the cell is more effective than activation afterwards in the presence of iron.

Problems with protein folding have been previously highlighted by Navarre *et al.* (2000). Cytosolic tobacco aconitase was over-expressed, solubilised and re-folded from inclusion bodies of *E. coli*. Using this method a specific activity of 2-4 nmol

min^{-1} mg^{-1} protein was obtained. Aconitase was purified from leaf extracts for comparison (Navarre *et al.,* 2000). Purified aconitase from tobacco leaves produced a specific activity of 700 nmol min^{-1} mg^{-1} (Navarre *et al.,* 2000). The difference in specific activity observed was believed to be the result of either improper expression in *E. coli* or the inability to reactivate plant aconitase. The difference observed between the specific activity of tobacco and *Arabidopsis* cytosolic aconitase suggests that re-activation during induction is a more effective method of obtaining a functioning aconitase than re-folding after His tag purification. However, further work is required to determine if ACO1 is fully active.

Further work could include the extraction and purification of the aconitase isoforms from *Arabidopsis*. The specific activity of these proteins could then be calculated and compared to those obtained for over-expressed protein from *P. pastoris*. This could determine if the over-expressed aconitase isoforms are fully active. This is important as previous research has shown that the specific activity of over-expressed aconitase and purified plant extracted aconitase can be very different (Navarre *et al.,* 2000; Described in section 5.1). Prior to determining the specific activity of mitochondrial aconitase from *Arabidopsis* a method needs to be developed to separate the two mitochondrial isoforms of aconitase. Whilst sub-cellular fractionation could be used to divide cytosolic and mitochondrial aconitase isoforms, a method of separating mitochondrial isoforms would need to be investigated, prior to specific plant aconitase activity studies. Previous research suggests that plant aconitase isoforms may be separated by hydrophobic interaction, hydroxylapatite and anion exchange chromatography (De Bellis *et al.,* 1993). These methods of purification could be investigated for the *Arabidopsis* aconitase isoforms in addition to determining their isoelectric point.

Whilst the specific activity of cytosolic aconitase over-expressed in *E. coli* was lower than that of purified ACO1 a direct comparison cannot be made. The specific activity of tobacco aconitase was determined by following the production of *cis*-aconitate from isocitrate at 240 nm (Navarre *et al.,* 2000). Attempts to measure the specific activity of aconitase using this method from either purified ACO1 and ACO2 or *Arabidopsis* extracts failed to detect any activity. As this method was unsuccessful, specific activities determined within this chapter followed the

production of isocitrate from *cis*-aconitate. Previous research conducted by Zheng *et al.* (1992) showed that the specific activity of bovine heart aconitase varies depending on the substrate used.

The limiting factor in determining the specific activity of isocitrate to *cis*-aconitate when working with seedling is the quantity of tissue required. Approximately 600 g of tobacco leaf tissue was harvested by Navarre *et al.* (2000), compared to the 8-32 mg of tissue used in the work within this chapter. Further work could investigate this in more depth though obviously the limitation is the small tissue available from *Arabidopsis* seedlings.

5.14.11 Activation of ACO2

The same experimental methods used to activate ACO1 were also implemented with purified samples of ACO2. Unfortunately, attempts to activate ACO2 after purification with method 1 were unsuccessful. One explanation for this could be the small quantities of protein purified from cell cultures potentially being below the level required for activity detection using this method. Another explanation could be that only a small, undetectable percentage of ACO2 is successfully activated. Previous research has shown that only 70-80 % of beef heart mitochondrial aconitase was successfully activated after purification (Kennedy *et al.,* 1983). Further work could include the extraction of the aconitase isoforms from *Arabidopsis* to establish if purified aconitase from *P. pastoris* is fully active.

Further work could also include the production of a higher concentration protein sample. This could be achieved by using a large scale over-expression system such as a batch fermenter. This technique is possible because of the inexpensive growth medium components *P. pastoris* requires. Yields of protein are believed to be higher in high density fermenters, as oxygen is utilised at a higher rate and levels can be tightly regulated (Cereghino and Cregg, 2000).

An alternative option could be to increase the number of stable aconitase integrates in *P. pastoris*. This could be achieved by increasing the concentration of the antibiotic zeocin used during the section process (Higgins and Cregg, 1998). Increasing the number of aconitase genes within *P. pastoris* would increase the

volume of protein over-expressed and allow further structural analysis to be conducted.

Attempts to activate purified samples of ACO2 using method 2 were successful. Using this method the specific activity of ACO2 was measured between 4-5 nmoles min^{-1} mg^{-1}. This specific activity is very similar to the specific activity obtained for porcine heart mitochondrial aconitase (Table 5.8). The similarities observed between ACO2 and mitochondrial porcine heart aconitase could suggest that ACO2 is also located within the mitochondria.

The specific activities of ACO1 observed within this chapter are considerably higher than those obtained for ACO2 (Table 5.8). These differences could reflect the different sub-cellular location of the isoforms and the role that they play in metabolism. The *Arabidopsis* aconitase isoform ACO1 is believed to function as part of glyoxylate cycle within the cytosol (Peyret *et al.,* 1995; Hayashi *et al.,* 1995), whereas ACO2 is believed to form part of the TCA cycle within the mitochondria (Arnaud *et al.,* 2007; Bernaud *et al.,* 2009).

Results presented within this chapter suggest that the specific activity of the cytosolic isoform of Arabidopsis (ACO1) is larger than that of the mitochondrial isoform (ACO2). Further work should be conducted to complete the successful identification of a *P. pastoris* clone over-expressing ACO3. Whilst previous attempts have failed to identify an expressing clone, this is believed to be the result of time constraints, rather than a problem with ACO3. Sequence analysis comparison between ACO1, ACO2 and ACO3 (Figure 5.2) show a large degree of sequence identity between the three proteins. This large degree of similarity suggests that there is no reason why ACO3 should not be successfully over-expressed in *P. pastoris*, when ACO1 and ACO2 are. Sequence alignment and homology modelling also failed to identify a reason why ACO3 should not have aconitase activity.

Further work could also include analysis of the proteins [Fe-S] cluster in order to determine its iron content. One technique that could be used to compare the [Fe-S] cluster of purified aconitase is extended x-ray absorption fine structure spectroscopy (EXAFS). Previous research has been conducted using the K edge of EXAFS to

investigate the [Fe-S] cluster of *Solanum tuberosum* aconitase in the presence and absence of citrate (Jordanov *et al.,* 1992). The stability of the aconitase [Fe-S] cluster could also be monitored by optical or EPR spectroscopy over time under a variety of conditions. Previous research with optical spectroscopy and EPR has shown that four stable species of aconitase exist. These include $[3Fe\text{-}4S]^{1+}$, $[3Fe\text{-}4S]^{0}$, $[4Fe\text{-}4S]^{2+}$ and $[4Fe\text{-}4S]^{1+}$ (Emptage *et al.,* 1983). It would be interesting to determine which state is more prominent in the aconitase isoforms and if all three aconitase isoforms are identical. Optical spectroscopy has also been used in yeast and *E. coli* to compare the state of the [Fe-S] cluster. This has been conducted previously to compare active and inactive forms of aconitase, different isoforms of aconitase and to determine the affect of pH on the [Fe-S] cluster (Gawron *et al.,* 1974; Suzuki *et al.,* 1976; Jordan *et al.,* 1999).

Further analysis could include the purification of the aconitase isoforms under anaerobic conditions. Previous research has suggested that tobacco aconitase is sensitive to aerobic conditions and once inactive it cannot be reactivated (Navarre *et al.,* 2000). In contrast, research has also shown that inactive beef heart mitochondrial aconitase can be activated anaerobically with reducing agents (Kennedy *et al.,* 1983). However, this only resulted in 70-80 % activation. Testing under anaerobic conditions was not conducted in the work in this chapter due to a lack of functioning equipment.

Previous research with over-expressed ACO1 has shown that this isoform of aconitase is unable to bind to ferritin and therefore does not regulate cellular iron levels (Arnaud *et al.,* 2007). However, research has shown that ACO1 is able to bind to CuZn superoxide dismutase 2 (Moeder *et al.,* 2007). Further work could be conducted using similar methods to determine if ACO2, and eventually ACO3 are also able to bind to the mRNA of CuZn superoxide dismutase 2. Further work could also be conducted to determine if the *Arabidopsis* aconitase isoforms are able to bind to DNA. Previous research has shown that yeast aconitase is involved in DNA maintenance, binding to both single stranded and double stranded DNA (Chen *et al.,* 2005; 2007).

5.14.12 Summary

Work presented within this chapter shows that the sequence and structure of ACO1, ACO2 and ACO3 are very similar. The similarity observed between their structure and those of ssACO and hsIRP1 would suggest that all three *Arabidopsis* isoforms have the potential to function as aconitase.

Total aconitase activity measurements of *Arabidopsis* seedlings were conducted to determine which isoforms of aconitase are active during germination and establishment. Activity measurements of *aco1*, *aco2* and *aco3* showed that ACO3 makes the greatest contribution to total aconitase activity. The activity differences observed between Col-0 and both *aco1* and *aco2* could be the result of seedling variation. However the aconitase genes were cloned and over-expressed in *P. pastoris* and it was confirmed that ACO1 and ACO2 both function as aconitases.

Within this chapter a novel method has been developed to over-express and purify aconitase under native conditions in a eukaryotic system, without the need for reactivation. This was achieved by adding iron to the growth medium during induction, ensuring that the [4Fe-4S] cluster was not iron-starved. This is in contrast to previous studies of aconitase expression in a prokaryotic system that required reactivation after expression (Navarre *et al.*, 2000). Through the application of this method, it has been demonstrated that *Arabidopsis* ACO1 and ACO2 have aconitase activity.

The development of an over-expression method that does not require protein re-folding was very important to ensure that aconitase was folded correctly. Misfolded proteins are often segregated in bacterial inclusion bodies in recombinant bacteria (Carrio and Villaverde, 2002). This is usually the result of over-expressed genes (Carrio and Villaverde, 2002). The resulting over-expressed protein then requires purification and re-folding. However, re-folding procedures are often complex and can produce a protein that is not fully functional. Whilst this method would seem to be a development over previous methods, further analysis is still required to determine if the purified aconitase is fully active.

Different metabolic roles for aconitase isoforms during establishment of *Arabidopsis thaliana* seedlings

6.1 Introduction

Early research conducted with caster bean endosperms showed that fatty acids are broken down into acetyl units by β-oxidation within glyoxysomes (Cooper and Beevers, 1969; Hutton and Stumpf, 1969). It is believed that these acetyl units are then metabolised by the glyoxylate cycle and gluconeogenesis to yield sucrose (Kornberg and Beevers, 1957). Feeding studies with ^{14}C radiolabeled acetate demonstrated that exogenously supplied acetate is also metabolised by the glyoxylate cycle and gluconeogenesis to produce sucrose (Canvin and Beevers, 1961).

Latter work conducted with lettuce embryos suggested that the acetyl units, incorporated into citrate by citrate synthase, were also metabolised in the mitochondria as part of the tricarboxylic acid (TCA) cycle (Salon *et al.,* 1988; Raymond *et al.,* 1992). Citrate was believed to be exported from the glyoxysome because of the cytosolic and mitochondrial location of aconitase. The process of β-oxidation within the glyoxysome produces hydrogen peroxide which is believed to disrupt the iron sulphur cluster of aconitase, resulting in inactivation (Verniquet *et al.,* 1991). This inactivation led to the conclusion that the glyoxylate cycle must take a detour via the cytosol (Courtois-Verniquet and Douce, 1993).

More recent research has shown that the cytosolic *Arabidopsis* aconitase (ACO1) participates in the glyoxylate cycle where it is responsible for the conversion of citrate to isocitrate (Peyret *et al.,* 1995; Hayashi *et al.,* 1995). The glyoxylate cycle contains two single copy enzymes isocitrate lyase (icl) and malate synthase (ms) that allow two decarboxylic steps of the TCA cycle to be bypassed (Figure 6.1). The *icl* mutant, which catalyses the subsequent reaction in the glyoxylate cycle to aconitase, allowed Eastmond *et al.* (2000) to demonstrate that the glyoxylate cycle is not essential for germination if adequate light is supplied. This suggests that an alternative route of metabolism is also available during germination, the TCA cycle. Research conducted by Pracharoenwattana *et al.* (2005) has proved that carbon is exported from the glyoxylate cycle as citrate. However the exact fate of this exported citrate and the mechanisms involved are currently unknown (Figure 6.1). Exported citrate may be transported to the mitochondria by a carboxylic acid transporter or metabolised by cytosolic aconitase (Pracharoenwattana *et al.,* 2005). The metabolic pathway of isocitrate produced by cytosolic aconitase is also undetermined. It may

re-enter the glyoxysome and be metabolised by the glyoxylate cycle or enter the mitochondria for metabolism by the TCA cycle.

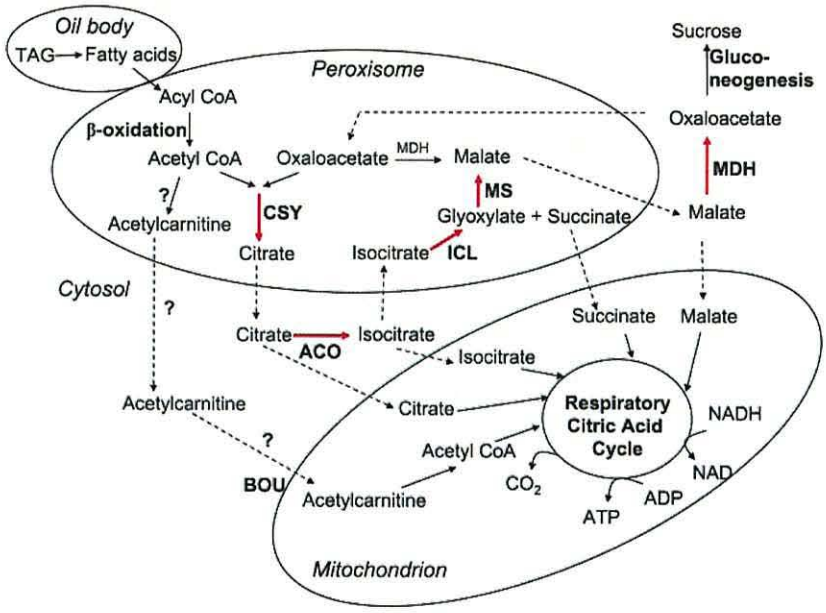


Figure 6.1: The main reactions involved in the peroxisome (also known as the glyoxysome) and routes of transport to the cytosol and mitochondria. The glyoxylate cycle reactions are shown in red. Image taken directly from Eckardt, 2005.

Two additional isoforms of aconitase have been identified in *Arabidopsis* (Arnaud *et al.,* 2007; Moeder *et al.,* 2007) and sets of homozygous knockout mutants *aco1*, *aco2* and *aco3* created. Aconitase activity measurements conducted by Arnaud *et al.* (2007) show a reduction in aconitase activity within all of the three mutants and concluded that ACO1 is located within the cytosol and ACO2 and ACO3 within the mitochondria. The location of the aconitase isoforms has subsequently been confirmed by Bernard *et al.* (2009). The *aco3* mutant showed the largest reduction in aconitase activity (Moeder *et al.,* 2007), suggesting that *aco3* is responsible for the majority of citrate to isocitrate metabolism in four week old plants. However, the specific role each aconitase isoform plays in seedling establishment is currently undetermined. Work conducted within this chapter aims to assist in determining the metabolic roles of ACO1, ACO2 and ACO3.

6.2 Identification of homozygous mutants

Homozygous knockout mutants were identified for all three aconitase isoforms in *Arabidopsis thaliana* for metabolic analysis. SALK lines from the SIGnAL project (Alonso *et al.*, 2003) and GABI-KAT (Rosso *et al.*, 2003) T-DNA insertion lines were purchased from the Nottingham *Arabidopsis* Stock Centre (NASC) and screened by PCR to identify homozygous mutants. SALK lines 008620, 026148 and 079347 were purchased and screened for At4g35830, lines 009537, 0122247, 030717, 031352, 037286, 054196, 079922, 090112, 090196, 090200 and 122245 for At4g26970, and lines 013368 and 014661 for At2g05710. Despite re-ordering the mutant batches and screening approximately fifty plants for each line a positive T-DNA mutant was only identified for line 090200. In light of this, the GABI-KAT T-DNA insertion lines 138A08 for At4g35830, 110E12 for At4g26970, and 298E10 and 423C03 At2g05710 were purchased and screened. GABI-KAT homozygous lines 138A08 and 298E10 were identified for At4g35830 (*aco1*) and At2g05710 (*aco3*) respectively and a SALK homozygous line 090200 was identified for At4g26970 (*aco2*). The 138A08 GABI-KAT line was the same mutant strain analysed by Arnaud *et al.* (2007).

T-DNA mutants were screened with both genomic and T-DNA / genomic junction primers (see materials and methods chapter for primer sequences). Heterozygous mutants were identified as they produced a PCR product with both genomic and T-DNA / genomic junction primers. Homozygous mutants were distinguished from heterozygous mutants as they only produce a PCR fragment with T-DNA / genomic specific primers. Homozygous mutants can be distinguished because the presence of the T-DNA insert prevents amplification with the genomic primers (Figure 6.2).

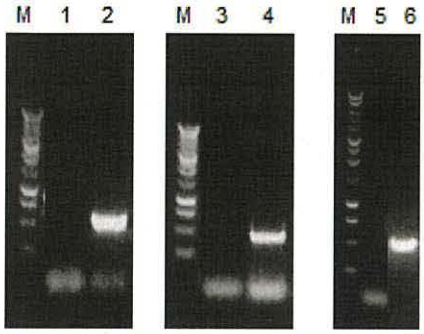


Figure 6.2: Conformation of homozygous mutants by PCR and visualisation with a 1 % agarose gel. M. 1kb molecular marker; 1. At4g35830 with 138A08 forward and reverse genomic primers; 2. At4g35830 with 138A08 and T-DNA specific primer; 3. At2g05710 with 298E10 genomic primers; 4. At2g05710 with 298E10 forward and Lb primer; 5. At4g26970 with 090200 genomic primers; 6. At4g26970 with 090200 forward and Lba1 primer set.

The T-DNA inserts are located within the thirteenth intron, the first exon and the eighth exon for homozygous mutants *aco1*, *aco2* and *aco3* respectively (Figure 6.3).

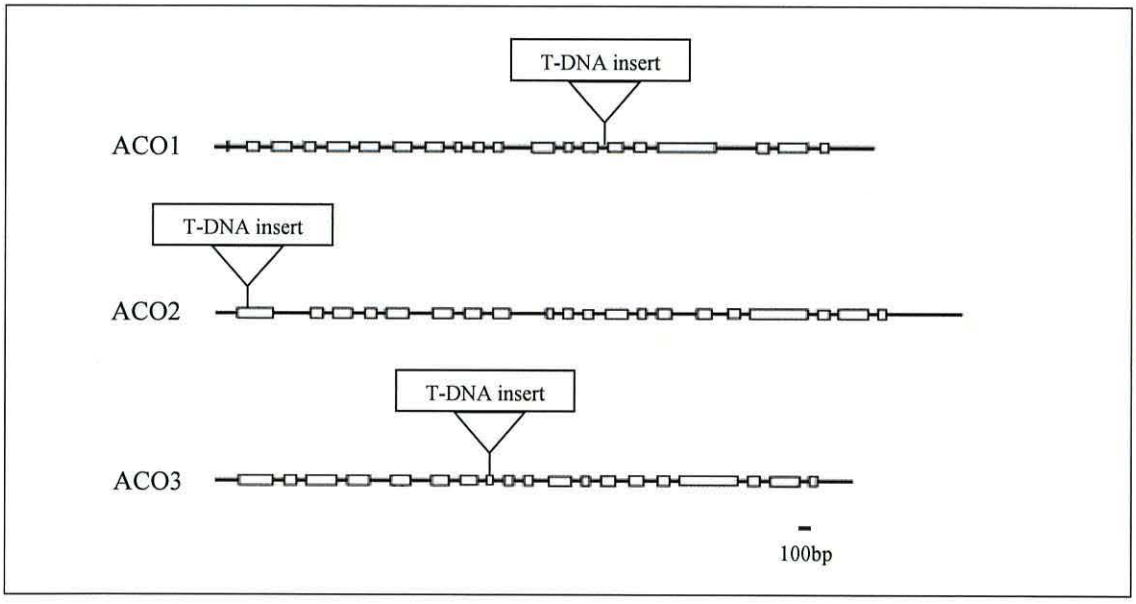


Figure 6.3: Characterisation of *Arabidopsis* knockout mutants. The diagram shows the site of T-DNA insertion within *aco1*, *aco2* and *aco3*. Boxes and lines represent exons and introns, respectively.

RT-PCR analysis was conducted to confirm that the aconitase homozygous knockout mutants were not producing mRNA. RT-PCR was conducted with gene specific oligonucleotide primers on cDNA extracted from 3 day old seedlings of *aco1*, *aco2*

and *aco3*. Analysis confirmed that the seedlings were knockout mutants and that only one gene was knocked out in each mutant (Figure 6.4). Actin was used as a control as it is constitutively expressed.

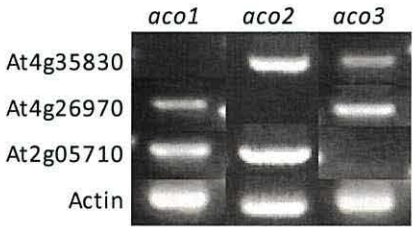


Figure 6.4: RT-PCR analysis of 3 day old *Arabidopsis* seedlings demonstrating the absence of transcript in the knockout lines *aco1*, *aco2* and *aco3*. Actin was used as a control. Data was collected from a series of electrophoresis gels during screening.

6.3 Physiological comparison between *aco1, aco2, aco3* and Wild type seedlings

Growth comparison trials of *aco1*, *aco2* and *aco3* showed that all three mutants germinated without the addition of an external carbon source. Analysis also revealed that *aco3* is delayed compared to wild type (Col-0) during early development (Figure 6.5).

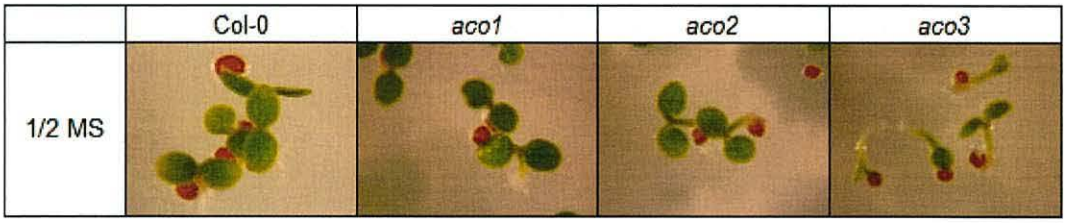


Figure 6.5: Three day old Col-0, *aco1*, *aco2* and *aco3* transferred to constant light conditions on ½ MS after four days imbibition in the dark at 4°C showing the observed delay in development of *aco3*.

The delayed germination of *aco3* was also reflected in the reduction in fresh weight observed of triplicate measurements over four days (Figure 6.6).

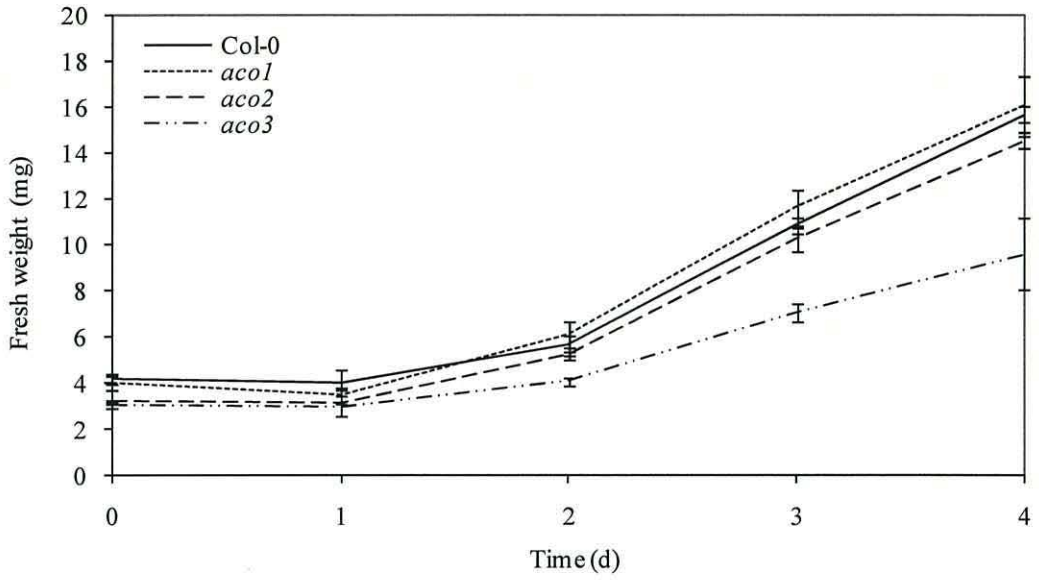


Figure 6.6: Triplicate measurements of fifty seeds were weighted every twenty four hours for four days to obtain an average weight and estimate the time required for all three mutants to reach the same growth stage. Data are the mean weights ± SD of 50 seedlings.

6.4 Acetate metabolism of *aco1*, *aco2* and *aco3* during seedling establishment

In addition to the utilisation of acetyl-CoA produced from β-oxidation during germination, seedlings also metabolise externally supplied acetate (Figure 6.7).

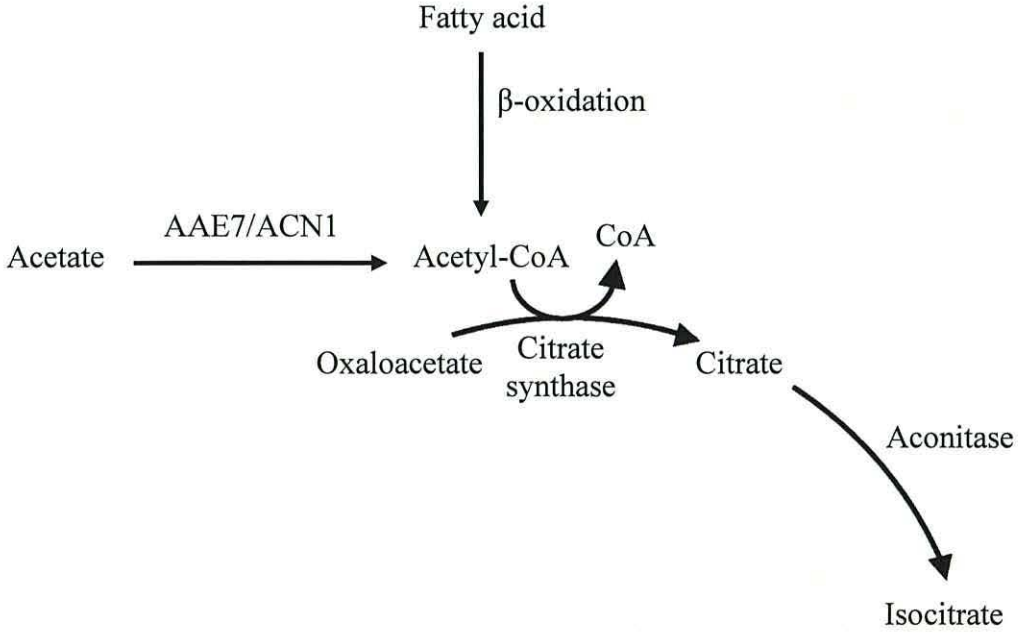


Figure 6.7: Metabolism of acetyl-CoA produced from either the breakdown of fatty acids by β-oxidation or externally supplied acetate.

Mutants *aco1*, *aco2* and *aco3* were germinated in the presence of fluoroacetate, the toxic analogue of acetate, to determine if the mutants are able to activate externally supplied acetate. Visualisation of *aco1*, *aco2* and *aco3* showed that all three mutants are sensitive to fluoroacetate (Figure 6.8) suggesting that fluoroacetate is converted to fluoroacetyl-CoA, which is subsequently metabolised by citrate synthase to 2-fluorocitrate that inhibits aconitase activity.

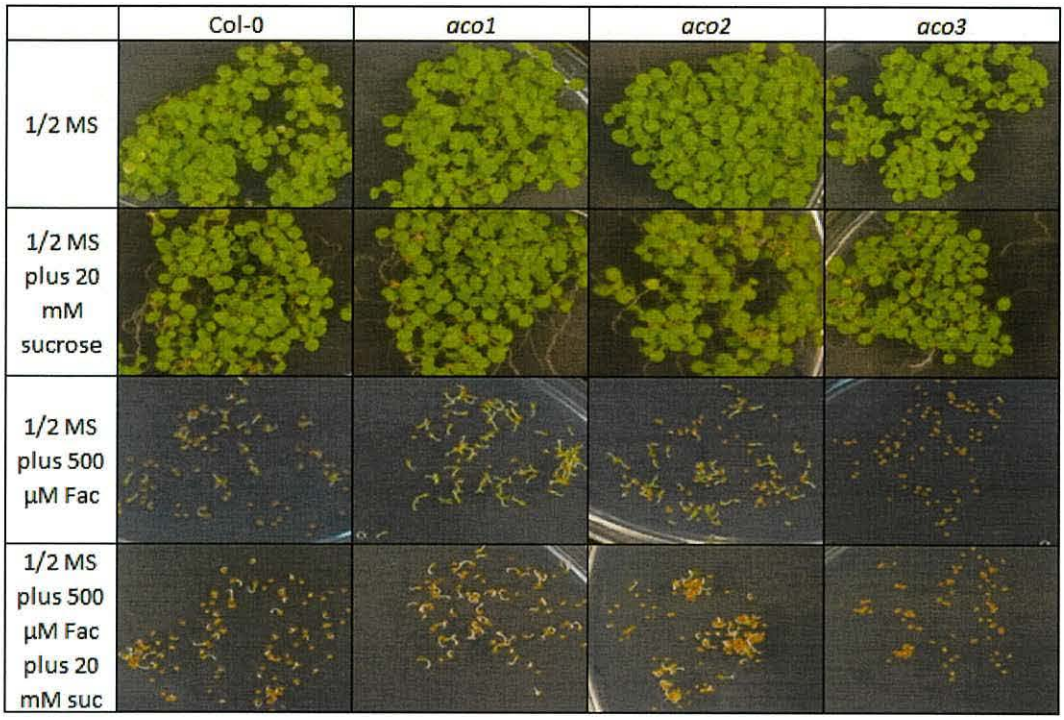


Figure 6.8: Six day old wt Col-0, *aco1*, *aco2* and *aco3* seedlings. Seeds were imbibed for four days in the dark at 4°C prior to transfer to constant light at 20°C onto ½ MS medium ± 20 mM sucrose and ± 500 µM fluoroacetate.

6.5 Quantification of citrate, oxaloacetate and pyruvate in wild type and aconitase mutants

The combined levels of citrate, oxaloacetate and pyruvate were determined within the wild type and aconitase mutants (Figure 6.9).

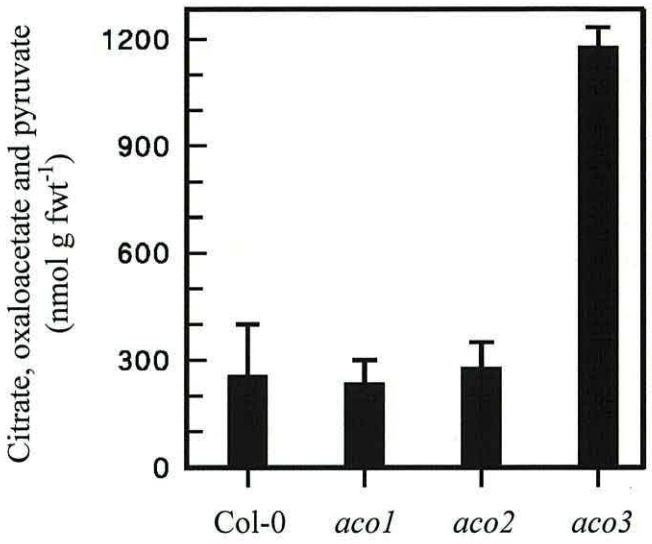


Figure 6.9: Citrate, oxaloacetate and pyruvate levels. Seedlings were grown under normal conditions until the average weight for each genotype was 0.12 mg seedling^{-1}. Approximately 0.5 g of seedlings were extracted and assayed enzymatically for citrate, oxaloacetate and pyruvate. Data are means ± SD of replicate samples.

6.6 Metabolite profiling by GC-MS of Col-0, *aco1*, *aco2* and *aco3*

Gas chromatography mass spectrometry was conducted in order to determine the metabolic profile of aconitase mutants. Metabolic profiles should assist in establishing the effect that the mutations have on metabolism and subsequently the role each isoform plays in metabolism.

Gas chromatography mass spectroscopy (GC-MS) is an established technique widely used for plant metabolite profiling (Fiehn, 2008). GC-MS analysis was conducted on three biological replicates of Col-0, *aco1, aco2* and *aco3*, to account for biological variation. Analytic error was assessed by collecting three metabolite profiles from each of the biological replicates. The quality of the metabolite extraction was assessed by addition of the internal standard ribitol (data not shown) and analytical error was assessed with myristic acid. A series of N-alkanes were also injected into the GC-MS as an external standard. Three derivatisation blanks were also produced to check that cross contamination was not occurring.

6.6.1 Analysis of the GC-MS chromatographs of Col-0, *aco1*, *aco2* and *aco3*

Visual comparison of the raw GC-MS chromatographs reveals that the Col-0 profile has its largest peak at approximately 23.8 minutes (Figure 6.10; highlighted in blue).

Whilst this peak is still present in *aco1* and *aco2* the relative percentage abundance is less and strikingly less in *aco3* (Figure 6.11, 6.12 and 6.13; all highlighted in blue). Conversely there is a peak at approximately 16.5 minutes in *aco3* (Figure 6.13; highlighted in red) that is far more pronounced than all other samples (Figure 6.10, 6.11 and 6.12; all highlighted in red).

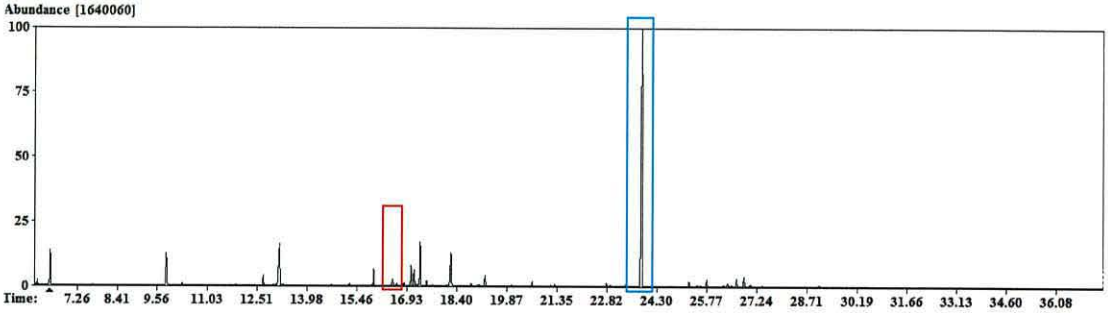


Figure 6.10: GC-MS chromatograph of wild type Col-0 seedlings with time in minutes on the x-axis and percentage abundance on the y-axis.

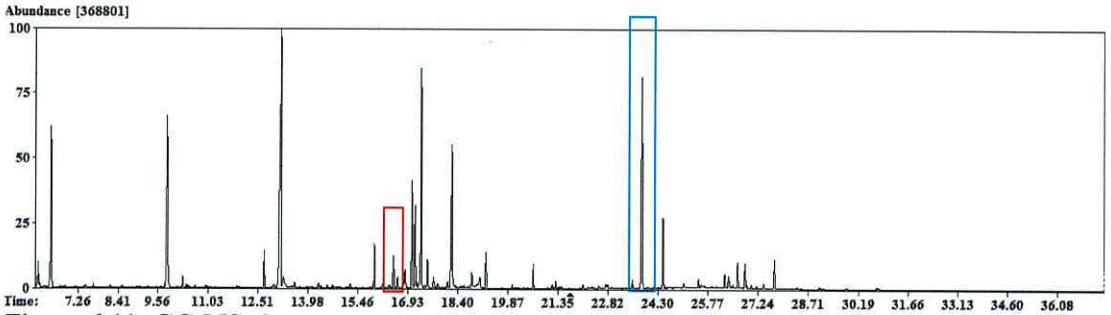


Figure 6.11: GC-MS chromatograph of *aco1* with time in minutes on the x-axis and percentage abundance on the y-axis.

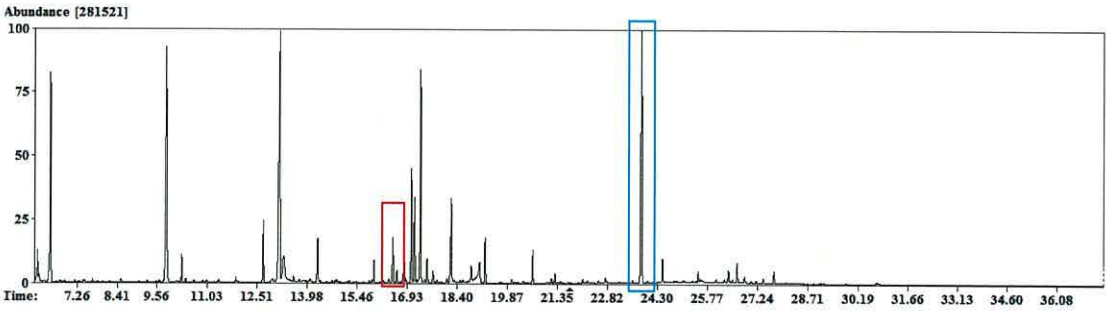


Figure 6.12: GC-MS chromatograph of *aco2* with time in minutes on the x-axis and percentage abundance on the y-axis.

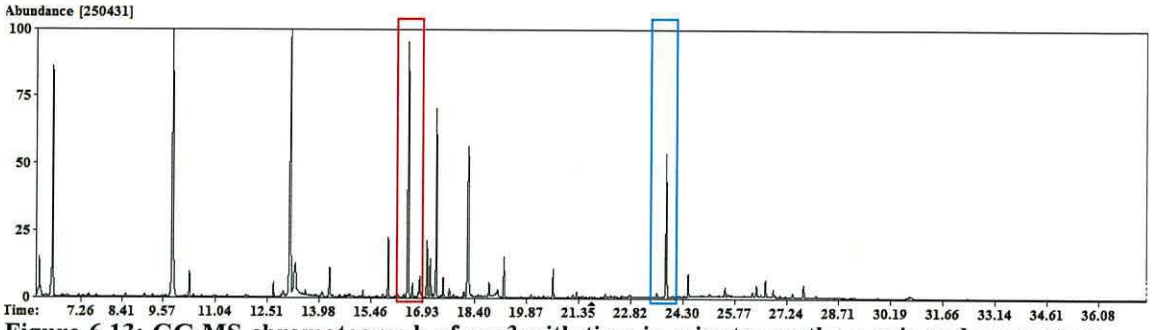


Figure 6.13: GC-MS chromatograph of *aco3* with time in minutes on the x-axis and percentage abundance on the y-axis.

6.6.2 Comparison between the processed GC-MS data of Col-0, *aco1*, *aco2* and *aco3*

The raw GC-MS chromatographs were processed and the identified peaks assigned to a specific compound by Dr J.W. Allwood. The external standard, myristic acid, added during metabolite extraction was identified with a retention time of 16.7 minutes within the processed data set. Statistical analysis was conducted with either a 1-way ANOVA or a Kruskal-Wallis, depending on the equality of variance (see appendix), to identify significant differences between the seedling groups (Col-0, *aco1*, *aco2* and *aco3*) for a specific metabolite.

The mean abundance of the control standard, myristic acid, in Col-0, *aco1*, *aco2* and *aco3* can be seen in Figure 6.14. Statistical analysis of myristic acid with a 1-way ANOVA showed no significant difference between Col-0, *aco1*, *aco2* or *aco3* (p=0.167), indicating that the analysis of the different samples was consistent.

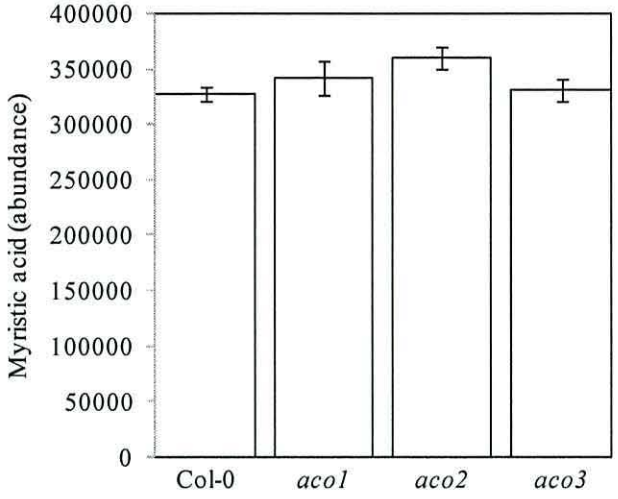


Figure 6.14: Myristic acid levels. Wild type, *aco1*, *aco2* and *aco3* seedlings were grown under normal conditions until the average weight for each genotype was 0.12 mg seedling^{-1}. Metabolites were extracted and analysed by gas chromatography mass spectrometry. Data are means ± SE of three biological replicates with typically three analytical replicates. Statistical analysis with a 1-way ANOVA showed no significant difference (p<0.05) between seedling groups.

Closer inspection of the processed GC-MS data reveals that two compounds, that show increased levels in *aco3*, elute at approximately 16.52 minutes (Figure 6.13; highlighted in red). These two compounds were identified as isocitric acid (Figure 6.15) and citric acid (Figure 6.16) with retention times of 16.58 and 16.62 respectively. A Leven's test showed that both isocitric acid and citric acid have unequal variances (p<0.001). Therefore a Kruskal-Wallis analysis revealed that *aco3* is significantly different to Col-0, *aco1* and *aco2* for both isocitric and citric acid, whereas there are no significant differences between these three groups.

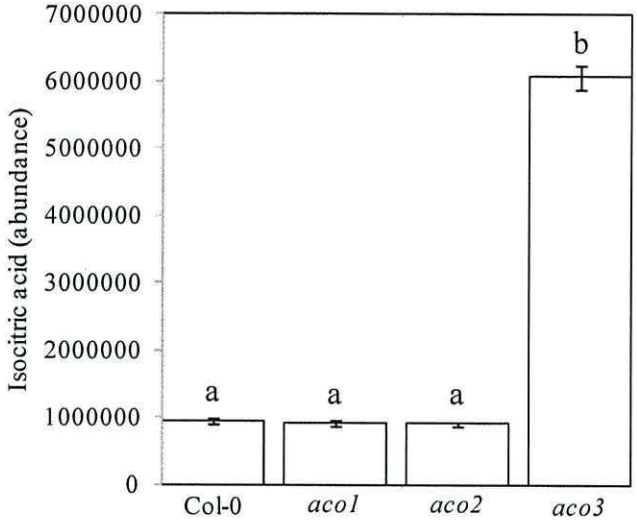


Figure 6.15: Isocitric acid levels. Wild type (Col-0), *aco1*, *aco2* and *aco3* seedlings were grown under normal conditions until the average weight for each genotype was 0.12 mg seedling^{-1}. Metabolites were extracted and analysed by gas chromatography mass spectrometry. Data are means ± SE of three biological replicates with typically three analytical replicates. Column labels with the same letter are not significantly different (using Dunnett's C tests) at the $p<0.05$ level.

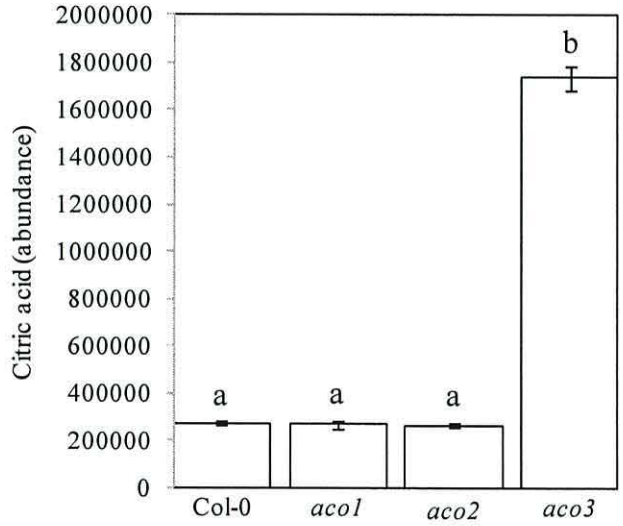


Figure 6.16: Citric acid levels. Wild type (Col-0), *aco1*, *aco2* and *aco3* seedlings were grown under normal conditions until the average weight for each genotype was 0.12 mg seedling^{-1}. Metabolites were extracted and analysed by gas chromatography mass spectrometry. Data are means ± SE of three biological replicates with typically three analytical replicates. Column labels with the same letter are not significantly different (using Dunnett's C tests) at the $p<0.05$ level.

The unprocessed GC-MS chromatographs revealed that the profile of Col-0 has its largest peak at approximately 23.84 minutes (Figure 6.10; highlighted in blue). The chromatographs also showed that this compound is present in *aco1*, *aco2* and *aco3*, but at significantly reduced levels (Figures 6.11, 6.12 and 6.13). When the data is processed, this peak corresponds to sucrose (Figure 6.17). Statistical analysis of this dataset was conducted using a Kruskal-Wallis test as a Leven's test indicated that the dataset has unequal variances ($p<0.001$). Kruskal-Wallis analysis gave a significant difference between the data ($p<0.001$) and post-hoc analysis illustrated that Col-0 is significantly different to *aco1*, *aco2* and *aco3*.

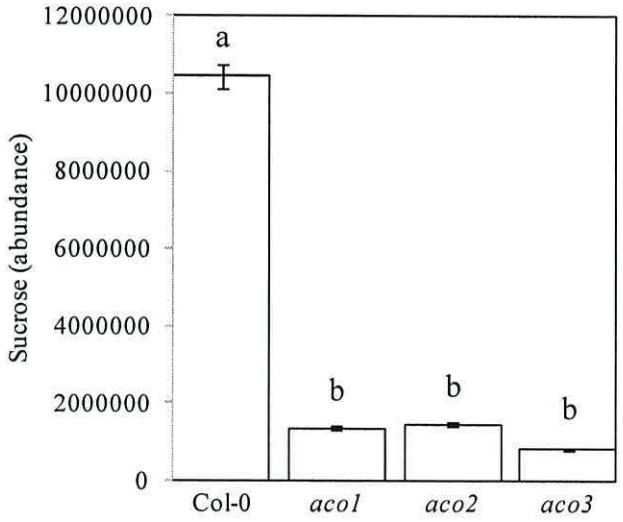


Figure 6.17: Sucrose levels. Wild type (Col-0), *aco1*, *aco2* and *aco3* seedlings were grown under normal conditions until the average weight for each genotype was 0.12 mg seedling^{-1}. Metabolites were extracted and analysed by gas chromatography mass spectrometry. Data are means ± SE of three biological replicates with typically three analytical replicates. Column labels with the same letter are not significantly different (using Dunnett's C tests) at the $p<0.05$ level.

The glyoxylate cycle functions during the germination of *Arabidopsis* seedlings with the net reaction of;

$$2 \text{ acetyl-CoA} + NAD^+ \longrightarrow \text{succinate} + 2 \text{ CoA} + NADH + H^+$$

Whilst succinate is the net product of the glyoxylate cycle it is also produced as part of the TCA cycle. It is therefore not surprising that a block in either the glyoxylate or

TCA cycle would lead to a reduction in the levels of succinic acid. GC-MS analysis shows that the levels of succinic acid were lower in *aco1* and *aco2* and markedly lower in *aco3* (Figure 6.18). This was confirmed when statistical analysis, conducted with a 1-way ANOVA, found a significant difference between the seedling groups ($p<0.001$). Post-hoc Tukey tests showed a significant difference between Col-0 and *aco1, aco2,* and *aco3*. This analysis also showed a significant difference between *aco3* and both *aco1* and *aco2*.

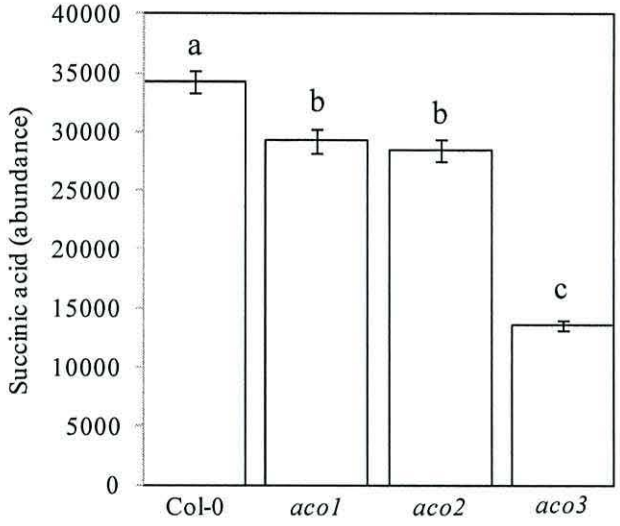


Figure 6.18: Succinic acid levels. Wild type (Col-0), *aco1*, *aco2* and *aco3* seedlings were grown under normal conditions until the average weight for each genotype was 0.12 mg seedling^{-1}. Metabolites were extracted and analysed by gas chromatography mass spectrometry. Data are means ± SE of three biological replicates with typically three analytical replicates. Column labels with the same letter are not significantly different (using Tukey tests) at the $p<0.05$ level.

Succinic acid is produced within the glyoxysome by isocitrate lyase along with glyoxylate, which is subsequently metabolised to malate. The production of succinic acid through the TCA cycle is also metabolised to malate via fumarate, making it difficult to determine the origin of identified malic acid. Nevertheless, GC-MS shows that the levels of malic acid were reduced in both *aco1* and *aco3* (Figure 6.19).

Statistical analysis conducted on the malic acid dataset with a 1-way ANOVA showed a significant difference between the seedling groups. Post-hoc Tukey tests demonstrated that this significant difference is between Col-0 and *aco1*, and *aco3*.

Interestingly, a significant difference was not observed between Col-0 and *aco2*. However, *aco1* is significantly different to *aco2*. In addition, analysis also revealed significant differences between *aco3* and both *aco1* and *aco2* (Figure 6.19).

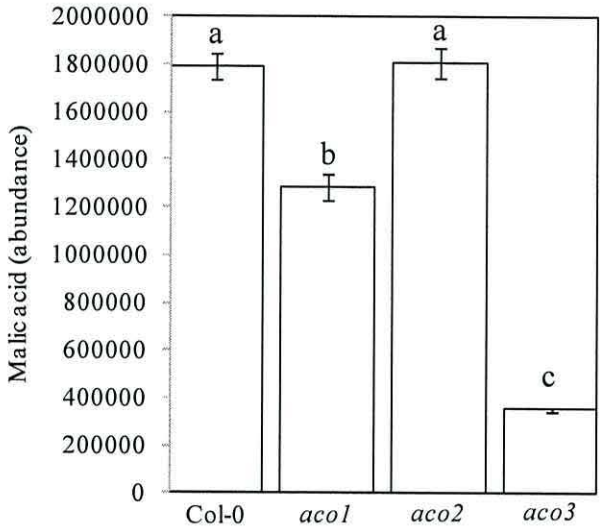


Figure 6.19: Malic acid levels. Wild type (Col-0), *aco1*, *aco2* and *aco3* seedlings were grown under normal conditions until the average weight for each genotype was 0.12 mg seedling^{-1}. Metabolites were extracted and analysed by gas chromatography mass spectrometry. Data are means ± SE of three biological replicates with typically three analytical replicates. Column labels with the same letter are not significantly different (using Tukey tests) at the $p<0.05$ level.

Whilst fumarate is not an intermediate of the glyoxylate cycle succinate, produced by the glyoxylate cycle, is channelled into the mitochondria for metabolism via the TCA cycle. The levels of fumarate detected by GC-MS were elevated in all three aconitase mutants, but dramatically so in *aco2* (Figure 6.20). Statistical analysis with a 1-way ANOVA showed a significant difference between the seedling groups. Post-hoc analysis using Tukey tests revealed a significant difference between Col-0 and both *aco1* and *aco2*. This analysis also revealed a significant difference between *aco2* and both *aco1* and *aco3*.

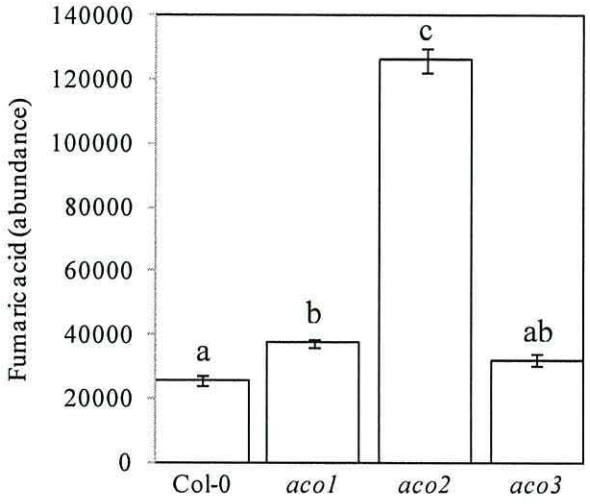


Figure 6.20: Fumaric acid levels. Wild type (Col-0), *aco1*, *aco2* and *aco3* seedlings were grown under normal conditions until the average weight for each genotype was 0.12 mg seedling^{-1}. Metabolites were extracted and analysed by gas chromatography mass spectrometry. Data are means ± SE of three biological replicates with typically three analytical replicates. Column labels with the same letter are not significantly different (using Tukey tests) at the $p<0.05$ level.

The levels of the metabolic intermediate, pyruvic acid were reduced in all three aconitase mutants (Figure 6.21). Statistical analysis of this dataset using a 1-way ANOVA showed the mean levels to be significantly different between the seedling groups. Tukey tests revealed Col-0 to be significantly different to *aco1*, *aco2* and *aco3*. This post-hoc analysis also revealed that *aco3* is significantly different to both *aco1* and *aco2*.

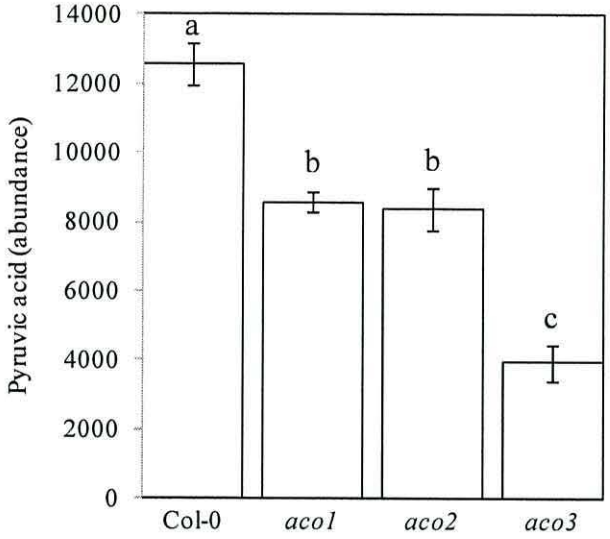


Figure 6.21: Pyruvic acid levels. Wild type (Col-0), *aco1*, *aco2* and *aco3* seedlings were grown under normal conditions until the average weight for each genotype was 0.12 mg seedling^{-1}. Metabolites were extracted and analysed by gas chromatography mass spectrometry. Data are means ± SE of three biological replicates with typically three analytical replicates. Column labels with the same letter are not significantly different (using Tukey tests) at the $p<0.05$ level.

6.7 Fatty acid metabolism during periods in the dark

During germination seedlings break down stored lipids through the processes of β-oxidation, the glyoxylate cycle and gluconeogenesis. As seedlings develop in the dark the hypocotyls elongate in search of light breaking down stored oil reserves. Mutants *aco1*, *aco2* and *aco3* were exposed to a 6 hr light period to initiate germination prior to a period of dark growth. Mutants *aco1*, *aco2* or *aco3* show no sucrose dependent phenotype or shortened hypocotyls suggesting that they are not compromised in their ability to convert fatty acids into sucrose (Figure 6.22).

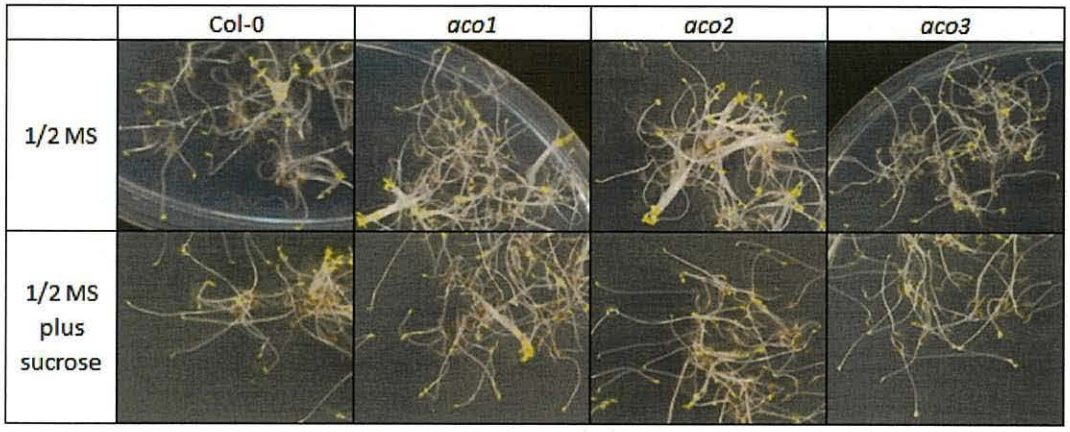


Figure 6.22: Six day old wild type (Col-0), *aco1*, *aco2* and *aco3* seedlings. Seeds were imbibed for four days prior a 6 hr light period to initiate germination and transferred to the dark at 20°C.

6.8 [2-^{14}C] Acetate metabolism in *aco1*, *aco2* and *aco3*

Feeding studies with radiolabeled acetate were conducted to compare the metabolism of acetate between Col-0, *aco1*, *aco2* and *aco3*. Fluoroacetate growth studies (see section 6.4) showed that all three aconitase mutants were sensitive to the acetate analogue (Figure 6.23). ^{14}C acetate feeding studies have been previously implemented to identify the roles of acetyl-CoA synthetase (Turner *et al.*, 2005), isocitrate lyase (Eastmond *et al.*, 2000) and malate synthetase (Cornah *et al.*, 2004) in acetate metabolism. Feeding studies with [2-^{14}C] acetate were conducted to determine acetate metabolised within the aconitase isoforms.

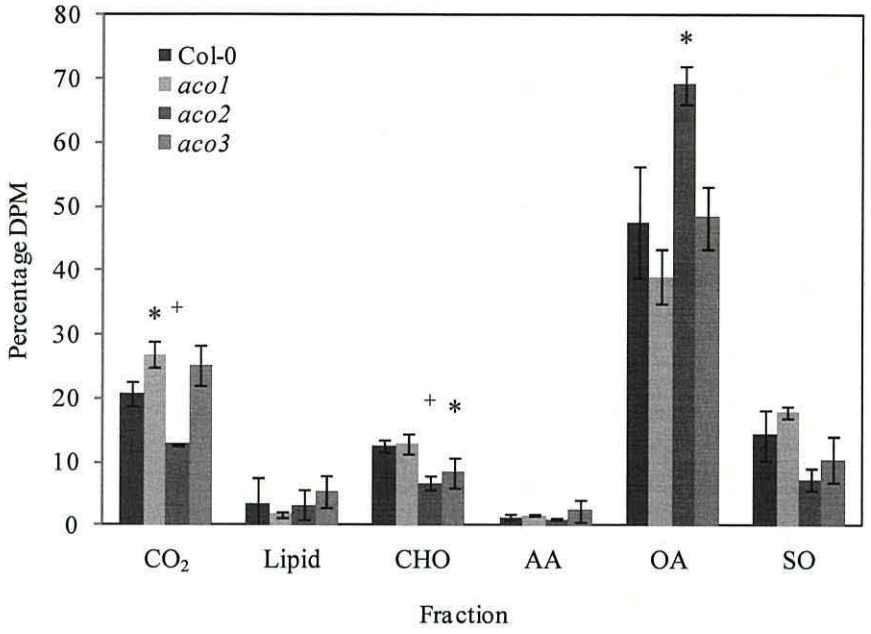


Figure 6.23: The metabolism of [2-^{14}C] acetate by emerged wild type (Col-0), *aco1*, *aco2* and *aco3* seedlings. One hundred light-grown seedlings were incubated in the presence of radiolabeled acetate in the dark for 4 h. The incorporation of ^{14}C into carbon dioxide (CO_2), lipid, carbohydrate (CHO), amino acids (AA), organic acids (OA) and ethanol insoluble fraction (SO) was determined by scintillation counting. Data are the means ± SD of three biological replicates. The symbols * and + indicate a significant difference of $p<0.05$ and $p<0.01$, respectively, between the aconitase mutant and Col-0 as determined by Student's t-test.

6.9 Discussion

Three isoforms of aconitase have been identified within *Arabidopsis* and denoted *aco1*, *aco2* and *aco3*. The initial aim was to identify multiple knockout mutants for each of the aconitase genes at different locations for metabolic comparison. Due to the difficulties experienced in obtaining either heterozygous or homozygous T-DNA insertion lines, a single set of mutants were analysed for each aconitase gene. Different sets of homozygous lines have been identified both within this chapter and previously by Arnaud *et al.* (2007) and Moeder *et al.* (2007). Within this chapter homozygous lines were determined by PCR analysis (Figure 6.2) and the knockout of a single aconitase gene was confirmed by RT-PCR (Figure 6.4). Northern blot analysis conducted by Arnaud *et al.* (2007) showed that the disruption of one aconitase isoform did not affect the expression of the other isoforms. Homozygous mutants were identified for each isoform of aconitase within this thesis to aid in determining the role each isoform plays in metabolism during seedling establishment.

6.9.1 Physiological comparison between *aco1, aco2, aco3* and wild type seedlings

Germination studies of the aconitase isoforms on standard agar (see materials and methods chapter 2.4.2) showed that all three mutants were able to germinate and establish in the absence of an external carbon source (Figure 6.8). The addition of the external carbon source sucrose has no affect on the germination of *aco1*, *aco2* or *aco3* suggesting that metabolism is not blocked in the mutants. The germination of *aco1*, *aco2* and *aco3* in the absence of sucrose also suggests that β-oxidation is not affected as the *ped1* mutant has a sucrose dependent phenotype in the light (Hayashi *et al.,* 1998). The addition of sucrose to the growth medium also had no affect on germination in the dark. This is in contrast to the glyoxylate cycle mutants of *isocitrate lyase* (Eastmond *et al.,* 2000) and *malate synthase* (Cornah *et al.,* 2004) that are sucrose dependent in the dark. *Arabidopsis* contains only a single isoform of isocitrate lyase and malate synthase. Mutation of either of these enzymes would therefore result in a block in the glyoxylate cycle. This is in contrast to aconitase, which contains three isoforms in *Arabidopsis*. The different physiological phenotypes observed between *isocitrate lyase*, *malate synthase*, *aco1*, *aco2* and *aco3* suggests that the glyoxylate cycle is not blocked in *aco1*, *aco2* or *aco3*. The presence of a functioning glyoxylate cycle suggests that multiple aconitase isoforms participate in the glyoxylate cycle. One explanation for this could be the presence of multiple aconitase isoforms within the cytosol, as a result of dual locality. An alternative explanation could be the transport of citrate to the mitochondria for conversion to isocitrate, followed by transportation back to the glyoxysome.

Trials of light conditions also revealed that *aco3,* grown under constant illumination, is delayed compared to wild type during early development (Figure 6.5 and 6.6). This suggests that ACO3 plays an important, but not essential role in development as seedlings are able to establish in its absence. Interestingly Moeder *et al.* (2007) reported no delay in germination when the aconitase isoforms were observed under a 16:8 h light:dark regime at 22°C. This may have masked the delay in germination that was observed in the study contained in this thesis, under constant light, as different metabolic pathways are utilised in the presence and absence of light.

During development glyoxysomes undergo a functional transition from β-oxidation and glyoxylate cycle activity to metabolism associated with photosynthesis

(Gonzalez and Vodkin, 2007; Pracharoenwattana and Smith, 2010). During this transition period from heterotrophic to autotrophic, glyoxysomes are converted into leaf peroxisomes (Jiang *et al.,* 1994).

6.9.2 Acetate metabolism of *aco1*, *aco2* and *aco3* during seedling establishment

During germination acetyl-CoA, produced from the breakdown of fatty acids, is believed to be metabolised through aconitase as part of the glyoxylate cycle (Hayashi *et al.,* 1995). Previous studies have shown that a proportion of exogenously supplied acetate is metabolised through the glyoxylate cycle with the use of the toxic analogue fluoroacetate (Turner *et al.,* 2005). Fluoroacetate is believed to be metabolised to fluoroacetyl-CoA and subsequently fluorocitrate within both mammals (Morrison, 1954) and plants (Lovelace *et al.,* 1968). The production of fluorocitrate is believed to be the site of toxicity, inhibiting the action of aconitase (Treble *et al.,* 1962).

Lauble *et al.* (1996) have shown that 4-hydroxy-*trans*-aconitate (HTn) binds tightly to aconitase, inhibiting enzyme activity. HTn is created from the (-)-*erythro* diastereomer of 2-fluorocitrate through fluoro-*cis*-aconitate by aconitase (Figure 6.24). A hydroxide is subsequently added to fluoro-*cis*-aconitate with the loss of fluoride to yield HTn.

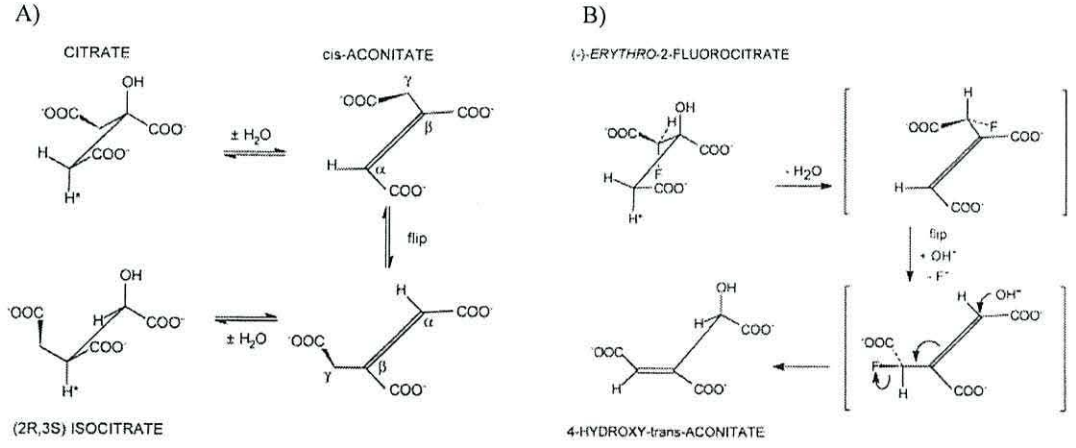


Figure 6.24: A) Mechanism of citrate to isocitrate conversion catalysed by aconitase. B) Proposed mechanism for the synthesis of 4-hydroxy-*trans*-aconitate (HTn) by aconitase. Images were taken directly from Lauble *et al.* (1996).

The toxic analogue of acetate, fluoroacetate was added to the growth medium (see section 2.4.2) to determine if a single or multiple isoforms of aconitase were important for the metabolism of citrate, or fluorocitrate during seedling establishment. Aconitase growth trials showed that all three mutants were sensitive to fluoroacetate (Figure 6.8). This suggests that multiple aconitase isoforms were active during germination as single mutations (*cts* and *aae7/acn1*) in the metabolism of fluoroacetate produced a fluoroacetate resistant phenotype (Turner *et al.,* 2005; Hooks *et al.,* 2007).

Growth trials showed that *aco1*, *aco2* and *aco3* are able to germinate and establish in the absence of an aconitase isoform. Previous research suggests that acetyl-CoA is metabolised to succinate during germination through the glyoxylate cycle and therefore ACO1, the cytosolic isoform of aconitase (Hayashi *et al.,* 1995; Peyret *et al.,* 1995). It was hypothesised that metabolism solely through the glyoxylate cycle by cytosolic aconitase would result in a *aco1* fluoroacetate resistant phenotype, in a similar way to *cts* (Hooks *et al.,* 2007) and *aae7/acn1* (Turner *et al.,* 2005). However a fluoroacetate sensitive phenotype was observed for all three aconitase mutants. This suggests that at least two aconitase isoforms are involved in the metabolism of exogenously supplied acetate.

Turner *et al.* (2005) have shown that only a proportion of the externally supplied acetate enters the glyoxysome for activation by AAE7/ACN1, suggesting that other routes of acetate metabolism are available. The fluoroacetate sensitive phenotype of *aco1*, *aco2* and *aco3* suggests that additional isoforms are involved in the metabolism of exogenous acetate. ACO1 participates in acetate metabolism as part of the glyoxylate cycle within the cytosol (Peyret *et al.,* 1995; Bernard *et al.,* 2009). The mitochondrial location of ACO2 and ACO3 (Bernard *et al.,* 2009) indicates that there must be an additional route for acetate metabolism to create an aconitase fluoroacetate sensitive phenotype. Other methods could include the direct entry of acetate into the mitochondria or the transport of a proportion of the citrate produced from the glyoxylate cycle to the mitochondria.

An alternative explanation for the fluoroacetate sensitive phenotype observed for the aconitase mutants could be the result of dual locality. ACO1 has been shown to be

cytosolic and ACO2 and ACO3 have both been shown to be mitochondrial proteins (Bernard *et al.,* 2009). However, a reduction in aconitase activity has been observed in both the cytosolic and mitochondrial fractions of *aco1* and *aco3* (Arnaud *et al.,* 2007). The dual location of aconitase within both the cytosol and the mitochondria has been previously observed in tomato and is believed to be the result of 'inefficient import' (Carrari *et al.,* 2003). The dual location of aconitase as a result of 'inefficient import' is not unique to plants. *Saccaromyces cerevisiae* contains a single isoform of aconitase that is active in both the cytosol and the mitochondria (Regev-Rudzki *et al.,* 2005). Further work could include the replication of methods implemented by Regev-Rudzki *et al.* (2005) to identify the location of aconitase activity in the *Arabidopsis* mutants.

Total aconitase activity measurements presented in chapter 5 (see summary Table 6.1) show that each isoform of aconitase contributes to the total activity. The total aconitase activity is decreased from 14 to 63 % showing that *aco1*, *aco2* and *aco3* are compromised in their ability to metabolise citrate. The addition of fluoroacetate will result in a further reduction in total aconitase activity as aconitase is inhibited. The contribution of *aco1*, *aco2* and *aco3* to total aconitase activity therefore explains why all three mutants are sensitive to fluoroacetate.

A similar pattern for the reduction in aconitase activity was observed in two day old, four week old and the mitochondrial fraction of 10 day old aconitase mutants (Table 6.1). Interestingly the percentage reduction observed in the cytosolic fraction of 10 day old seedlings was very different. An explanation for the differences observed could be due to inactivation of ACO1 during extraction. The glyoxylate cycle isoform of aconitase, ACO1, is believed to be located within the cytosol due to the presence of hydrogen peroxide in the glyoxysome (Verniquet *et al.,* 1991; Courtois-Verniquet and Douce, 1993). Disruption of cellular organelles during extraction may have resulted in the release of hydrogen peroxide, disrupting the [Fe-S] cluster of ACO1. The partial inactivation of ACO1 could have resulted in the smaller reduction in aconitase activity observed in *aco1* compared to that observed by Arnaud *et al.* (2007). Further work could include activity measurements of ACO1 and ACO2 in the presence of hydrogen peroxide to compare their sensitivity.

Table 6.1: Summary of the average percentage reduction in aconitase specific activity of *Arabidopsis* extracts from this study conducted on 2 day old seedlings (see chapter 5 for raw data), Moeder *et al.* (2007) and Arnaud *et al.* (2007).

Aconitase mutant	This study, Total extraction 2 days old	Moeder *et al.*, Total extraction 4 weeks old	Arnaud *et al.*, 10 days old	
			Cytosolic fraction	Mitochondrial fraction
aco1	23	20	75	20
aco2	14	20	No change	20
aco3	63	70	25	55

Whilst all three mutants were sensitive to fluoroacetate, *aco3* showed increased sensitivity suggesting that ACO3 is an important isoform during seedling establishment (Figure 6.8). The substantial reduction in aconitase activity observed in *aco3* (~60 %) coupled with the presence of fluoroacetate would severely reduce aconitase activity. In *aco3* this severe reduction in aconitase activity in the presence of fluoroacetate could mimic a double mutant phenotype. Previous attempts to create a double mutant with *aco3* and either *aco1* (Arnaud *et al.*, 2007) or *aco2* (Moeder *et al.*, 2007) have been unsuccessful. Arnaud *et al.* (2007) noticed only aborted seeds during crossing experiments and suggested that ACO1 and ACO3 are required for the development of the embryo, or for the proper development of the seed.

Further work could include the implementation of matrix-assisted laser desorption/ionisation (MALDI) imaging to directly map the location of metabolites within seedling tissue. Atmospheric pressure infrared MALDI mass spectrometry (AP IR-MALDI) imaging has been previously conducted on a variety of plant tissues, including *Prunus amygdalus* seeds (Li *et al.*, 2008). A wide variety of molecules including carbohydrates, oligosaccharides, amino acids, organic acids and lipids have been identified in plant tissue (Li *et al.*, 2008). This method could assist in determining the direct route of fluoroacetate metabolism during seedling establishment by identifying the location of fluorinated compounds. Previous research has suggested that only a proportion of exogenous acetate enters the glyoxysome and is metabolised to acetyl-CoA (Turner *et al.*, 2005). MALDI imaging could therefore assist in determining the route of the remaining externally supplied acetate.

6.9.3 Quantification of citrate, oxaloacetate and pyruvate levels in wild type and aconitase mutants by enzymatic analysis

Combined citrate, oxaloacetate and pyruvate levels were determined by enzymatic analysis within Col-0, *aco1*, *aco2* and *aco3*. Enzymatic analysis showed that total levels were five-fold higher in *aco3* compared to Col-0, *aco1* and *aco2*. Whilst the lack of a full GC-MS dataset for oxaloacetate precludes one hundred percent certainty, the comparison of enzymatic analysis patterns (Figure 6.9) to citrate (Figure 6.16) and pyruvate (Figure 6.21) GC-MS data shows interesting similarities. These comparisons may indicate that citrate is the most abundant compound and thus responsible for the pattern of relative results shown in the enzymatic analysis data. This pattern is given further credence by Shen *et al.* (2006) who reported levels of oxaloacetate, pyruvate and citrate to be approximately 10, 45 and 1300 nmol g^{-1} FW respectively. The levels of citrate reported by Shen *et al.* (2006) are higher than the total levels of citrate, oxaloacetate and pyruvate measured within this chapter. These differences could be due to differences in seedling age. Metabolite levels within this chapter were determined from seedlings aged between 2-3 days (see materials and methods chapter). This is in contrast to work conducted by Shen *et al.* (2006) who measured metabolite levels in two week old seedlings.

6.9.4 Quantification of metabolites during germination by GC-MS

Visual comparison of the GC-MS data clearly shows that the metabolic profiles of Col-0, *aco1, aco2* and *aco3* are not identical. These differences represent altered levels of metabolites that are likely to be related to the lack of aconitase function in the mutants, and the delay in establishment observed in *aco3* during seedling growth trials (Figure 6.5 and 6.6).

The observation that all three aconitase profiles differ from that of Col-0 suggests that the different aconitase isoforms are responsible for different functional processes and are not able to completely compensate for one another during metabolism. Closer inspection of the GC-MS data reveals that a number of metabolic intermediates of the TCA and glyoxylate cycle are affected by the knockout of an aconitase isoform in *Arabidopsis*. These alterations are highlighted in Figure 6.25. Sucrose levels are low in all three mutants due to a disruption in the TCA and

glyoxylate cycles, as this is the cycles' principle product. However within these cycles the GC-MS data paints a less clear picture (Figure 6.25).

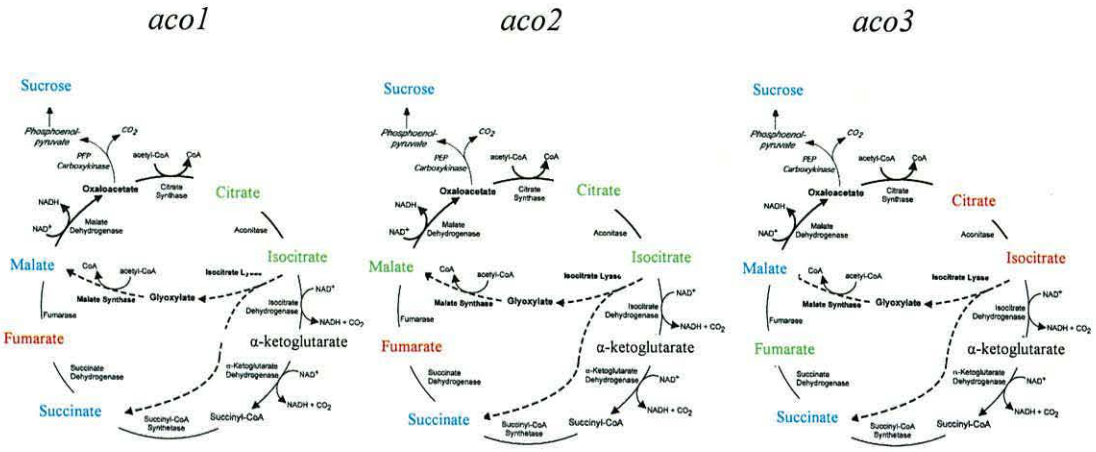


Figure 6.25: Summary of the changes in TCA and glyoxylate cycle intermediates. Image modified from Lorenz and Fink, 2002. Red indicates a statistically significant increase in metabolite levels, blue indicates a statistically significant decrease and green indicates unaltered levels, all compared to wild type (Col-0) levels.

6.9.4.1 Citric acid levels within the aconitase mutants

Aconitase is believed to be responsible for the isomerisation of citrate to isocitrate as part of both the glyoxylate and TCA cycle (Brouquisse *et al.,* 1987; Cornah and Smith, 2002). The partitioning of citrate between these cycles is believed to distribute carbon produced from fatty acid breakdown between respiration and gluconeogenesis in *Arabidopsis* seedlings (Kunze *et al.,* 2006). The identification of elevated citrate levels by GC-MS (Figure 6.16) in *aco3* was not unexpected, as the increase in citrate levels observed was believed to be the result of a block in either cycle, preventing further metabolism of citrate. Increased citrate levels due to a decrease in aconitase activity have been previously observed in *Citrus limon* (Shlizerman *et al.,* 2007) *Lupinus albus* (McCluskey *et al.,* 2004) and tomato plants (Carrari *et al.,* 2003; Lopez Millan *et al.,* 2009).

An initial hypothesis concerning the unaltered levels of citrate in *aco1* and *aco2* was that both isoforms were able to compensate for the activity of the other, resulting in a wild type phenotype. Further experimental analysis suggests that this is unlikely to be the full story however, as the concentrations of other intermediate metabolites differ between the mutants. The different spatial location of ACO1 and ACO2,

determined by Bernard *et al.* (2009), also suggests that the isoforms are unable to fully compensate for one another. In order to further investigate the roles of ACO1, ACO2 and ACO3 additional work could be conducted in the form of gene expression analysis. Work could include the quantification of the aconitase genes during establishment or microarray analysis to determine what affect the knockout mutants have on gene expression. Proteomic analysis could also be conducted in the form of two-dimensional polyacrylamide gel electrophoresis (2-D PAGE) in combination with mass spectroscopy to compare proteins present between the mutants.

Research conducted by Bernard *et al.* (2009) showed that ACO1 is located within the cytosol, whilst ACO2 and ACO3 are both located within the mitochondria of *Arabidopsis*. Experimental analysis presented in chapter 5 shows that the most dominant isoform of aconitase during germination is ACO3. Moeder *et al.* (2007) showed that ACO3 is the dominant isoform of aconitase in four week old plants which would corroborate with the GC-MS data in this study. Additional results presented in chapter 5 showed that both ACO1 and ACO2 are able to function as aconitase, converting *cis*-aconitate to isocitrate.

The unaltered levels of citrate observed in *aco1* could be due to transport of citrate to the mitochondria where functioning aconitases, ACO2 and ACO3 are present. The export of citrate from the glyoxysome is essential for continued metabolism and has been demonstrated by Pracharoenwattana *et al.* (2005). However it is unclear whether citrate or isocitrate is transported to the mitochondria (Thorneycroft *et al.*, 2001). The continued metabolism of citrate, in the absence of the cytosolic aconitase (ACO1), suggests that citrate exported from the glyoxysome can be transported to the mitochondria for further metabolism.

The unaltered levels of citrate observed in *aco2* may be due to the presence of ACO1 in the cytosol and ACO3 in the mitochondria converting citrate to isocitrate. Results presented in chapter 5 show that whilst ACO2 contributes to total aconitase activity, its contribution is small. The presence of ACO1 and ACO3 in addition to the unaltered levels of citrate, suggest that metabolism of citrate continues. Enzymatic results presented in chapter 5 also show that ACO2 is able to convert *cis*-aconitate to

isocitrate. This, in addition to differences observed within the cycles, suggests that ACO2 has an effect on metabolism. Whether this is a direct affect as an aconitase, or indirectly by binding to mRNA or DNA is still undetermined.

The increase in citrate levels observed in *aco3* may be due to the large flux of citrate through ACO3, which ACO1 and ACO2 alone cannot metabolise. The levels of citrate observed in *aco1* and *aco2* suggest that ACO1 and ACO2 isoforms do not convert a large amount of citrate to isocitrate in *vivo*, leading to the accumulation of citrate in *aco3*.

An alternative explanation for the elevated levels of citrate in *aco3* could be the result of different tissue expression of ACO1 and ACO2 compared to ACO3. Different levels of mRNA have been observed within seeds, roots, caulinary leaves and flowers between the aconitase mutants (Arnaud *et al.,* 2007). In addition the *Arabidopsis* eFP Browser shows that the three aconitase isoforms are expressed at different levels within dry seed, imbibed seed, cotyledons, roots and hypocotyl (Winter *et al.,* 2007). Further work could include the determination of mRNA levels of the aconitase isoforms during germination and establishment.

The build up of citrate observed in *aco3* could also be the result of compartmentalisation of citrate away from ACO1 and ACO2 within each cell. The vacuole has been shown to serve as a storage organelle for primary metabolites that can be retrieved for metabolism when required (Taiz, 1992; Gout *et al.,* 1993). The transport and accumulation of citrate within the vacuole has been shown by Hurth *et al.* (2005). Further work could include the isolation of the vacuole from *Arabidopsis* (Frangne *et al.,* 2002) followed by quantification and comparison of citrate levels to determine if citrate accumulates in the vacuole.

6.9.4.2 Isocitric acid levels within the aconitase mutants

GC-MS analysis revealed that the levels of isocitric acid were elevated in *aco3*. This was an unexpected result as aconitase is believed to function in the forward direction, from citrate to isocitrate during metabolism. However, the reversal of a key metabolic enzyme has been shown by Pracharoenwattana *et al.* (2007) with malate dehydrogenase.

Whilst elevated levels of isocitrate were unexpected in *aco3*, this observation is not unique to *Arabidopsis*. Elevated levels of isocitrate were also identified in the *Lycopersicon pennellii* (tobacco) aco1 (lpaco1) mutant (Carrari *et al.*, 2003). The lpaco1 is believed to function in both the cytosol and mitochondria of mature plants (Carrari *et al.*, 2003). To date, a reason for the build up of isocitrate is unknown. Yeast aconitase 1 has been shown to be controlled by phosphorylation (Pitula *et al.*, 2004). Interestingly, phosphorylation prevented the conversion of citrate to isocitrate but had little effect on the conversion of isocitrate to citrate. Further research could therefore be conducted to determine if aconitase is regulated by phosphorylation.

Isocitrate may be metabolised during germination by either isocitrate lyase or isocitrate dehydrogenase to succinate and glyoxylate or α-ketoglutarate respectively. The increase in isocitrate observed in *aco3* could therefore be the result of a reduction in enzyme activity of $NADP^+$ isocitrate dehydrogenase or isocitrate lyase, as both enzymes are inhibited by citrate (Popova *et al.*, 1998). Aconitase B of *E. coli* has been shown to form a multi-enzyme complex with isocitrate dehydrogenase (Tsuchiya *et al.*, 2008). Multi-enzyme complexes have also been suggested with TCA cycle enzymes in rat liver extracts (Robinson and Srere, 1985) and *Pseudomonas aeruginosa* (Mitchell, 1996). Further research could be conducted to determine if aconitase interacts with isocitrate dehydrogenase in *Arabidopsis* and what affect this may have on metabolism.

The levels of isocitrate are not significantly different to the levels of wild type in *aco1* and *aco2*. The unaltered levels of isocitrate could therefore be explained by the presence of ACO3. Results presented in chapter 5 show that all three isoforms contribute to the total aconitase activity. However, the majority of citrate appears to be metabolised by ACO3 during seedling development.

6.9.4.3 Sucrose levels in the aconitase mutants

The levels of sucrose are reduced in all three aconitase mutants compared to Col-0. During germination triacylglycerol (TAG) is broken down to fatty acids which are further metabolised to acetyl-CoA by β-oxidation. Acetyl-CoA then enters the glyoxylate cycle to yield succinate, which is subsequently transported to the mitochondria where it enters the TCA cycle. The product of the TCA cycle,

oxaloacetate is then metabolised by gluconeogenesis to carbohydrates such as sucrose. A block in either the glyoxylate or TCA cycle would therefore result in a decrease in the concentration of sucrose produced.

Whilst the levels of sucrose are lower in the three mutants compared to Col-0, the seedlings are able to germinate and establish in the absence of an external carbon source indicating that the cycles are not completely blocked. This is in contrast to the *isocitrate lyase* mutants that establish in the light but are defective in hypocotyl elongation in the dark, unless an external carbon source is supplied (Eastmond *et al.*, 2000). Growth trials conducted with *aco1*, *aco2* and *aco3* in the dark (Figure 6.22) showed normal hypocotyl elongation suggesting that the glyoxylate cycle is not completely blocked. This suggests that either isocitrate produced in the mitochondria is transported to the glyoxysome or that aconitase has a dual location, allowing the glyoxylate cycle to continue. Aconitase has been reported to have a dual location in *Saccharomyces cervisiae*, where it participates in both the glyoxylate and TCA cycle (Regev-Rudzki *et al.*, 2005).

The main form of carbon transported within higher plants is sucrose (Xu *et al.*, 2010). During germination lipids are converted into sucrose in the embryo and endosperm, which is then transported throughout the germinating seedling (Penfield *et al.*, 2005). In addition to the role of carbon transport, sucrose is an important signalling molecule that regulates genes involved in photosynthesis, metabolism and developmental processes (Xu *et al.*, 2010). Therefore some of the changes in metabolism observed within the mutants could also have resulted from a disruption in sucrose signalling.

6.9.4.4 Succinic acid levels within the aconitase mutants

Succinate is an intermediate of both the glyoxylate and TCA cycle during germination. Succinate produced by the glyoxylate cycle is exported from the glyoxysome and imported into the mitochondria for metabolism through the TCA cycle. The levels of succinic acid are reduced in all three aconitase mutants but are significantly lower in *aco3* compared to Col-0, suggesting a reduction in flux through metabolism. Due to the production of succinate from both the glyoxylate and TCA cycle it is not possible to determine which cycle each of the aconitase

isoforms participate in. The significantly lower levels observed in *aco3* suggest that the majority of citrate is metabolised through ACO3 in the mitochondria. The reduction in succinate observed within the mutants could potentially affect the electron transport chain as electrons are transferred during the conversion of succinate to fumarate by complex II (Brown, 2000).

6.9.4.5 Malic acid levels within the aconitase mutants

Malate is an intermediate of both the glyoxylate and TCA cycle. The precise routes for malate metabolism are currently undetermined but malate is believed to be exported from the glyoxysome in exchange for aspartate via an α-ketoglutarate driven malate-aspartate shuttle (Mettler and Beevers, 1980).

GC-MS data reveals that the levels of malic acid are lower in *aco1* and significantly lower in *aco3*. The reduction in malic acid indicates a reduction in flux through metabolism in both *aco1* and *aco3*. This suggests that the majority of citrate is metabolised through ACO1 and ACO3 during germination, resulting in a decrease in malate levels.

The levels of malic acid appear to be unaltered in *aco2*, compared to Col-0. This could suggest that ACO2 has little effect on metabolism. However, differences observed in the levels of succinate and fumarate indicates that ACO2 does affect metabolism. The unaltered levels of the glyoxylate cycle intermediates would therefore suggest that this pathway is unaffected by ACO2. The mitochondrial location of ACO2 (Bernard *et al.,* 2009) could explain the observed differences in the levels of the mitochondrial intermediates succinate and fumarate.

6.9.4.6 Fumaric acid levels within the aconitase mutants

The levels of fumaric acid measured by GC-MS were shown to be elevated in all three aconitase mutants compared to Col-0, but markedly so in *aco2*. The increased levels of fumaric acid observed in the aconitase mutants could be the result of a reduction in fumarase activity, a reduction in the production of fumarate metabolites (arginine) or an increase in the breakdown of fumarate metabolites (arginine and purine nucleotides). However, to date, no direct evidence exists to support the interaction of aconitase or citrate with fumarase or enzymes involved in the

production or breakdown of fumarate. Further work could be conducted to determine the levels of α-ketoglutarate within the mutants as α-ketoglutarate inhibits fumarase in peas (Behal and Oliver, 1997).

The accumulation of fumarate from malate has been previously observed in *Arabidopsis* leaves in the light (Pracharoenwattana *et al.,* 2010) and from a reduction in fumarase activity in response to oxidative stress (Sweetlove *et al.,* 2002). Sweetlove *et al.* (2002) also noted that aconitase activity was reduced under conditions of oxidative stress. Lehmann *et al.* (2008) showed that the transcript of ACO1 decreased and the transcript of ACO2 increased in response to an oxidative stress. The increase in ACO2 transcript could suggest that this isoform is involved in regulating reactive oxygen species as aconitase has been shown to regulate resistance to oxidative stress (Moeder *et al.,* 2007). Aconitase has been previously reported to be a sensitive redox sensor of reactive oxygen species and nitrogen species in cells (Vasquet-Vivar *et al.,* 2000). This may be important as an increase in superoxide within mitochondria can cause severe dysfunction (Vasquet-Vivar *et al.,* 2000).

A succinate / fumarate translocator has been identified within *Arabidopsis* that transfers succinate, produced by the glyoxylate cycle, into the mitochondria for conversion into fumarate by succinate dehydrogenase (Catoni *et al.,* 2003). This allows fumarate to be produced in the absence of a fully functioning TCA cycle. Sweetlove *et al.* (2010) suggest that the TCA cycle may not be fully functioning during different developmental stages, in different tissue types and under different environmental conditions in different organisms. For example, the citrate, isocitrate and α-ketoglutarate steps, together with the fumarate, malate and oxaloacetate steps, function in nitrogen assimilation in *Xanthium strumarium*. In contrast the citrate to malate steps are believed to be redundant in *Spinacia oleracea* during nitrogen assimilation.

Until recently it was believed that the only fumarase present in *Arabidopsis* was the mitochondrial isoform that functions as part of the TCA cycle. Pracharoenwattana *et al.* (2010) have shown that a second fumarase (FUM2) exists, located within the cytosol. Pracharoenwattana *et al.* (2010) concluded that FUM2 is not required for plant growth, but is necessary for the build up of fumarate in leaves, which is

subsequently used for rapid nitrogen assimilation and growth under high nitrogen conditions. The build up of fumarate has also been observed in photosynthetically active tissue of *Arabidopsis* (Chia *et al.,* 2000), specifically high levels within the vacuole. High levels of fumarate were also identified within the phloem suggesting participation in the transport of carbon throughout the plant (Chia *et al.,* 2000).

6.9.4.7 Pyruvic acid levels within the aconitase mutants

Pyruvate, the final product of glycolysis, is also involved in the metabolism of sucrose through gluconeogenesis. The production of sucrose through gluconeogenesis is essentially the reversal of glycolysis, except that phosphofructokinase is replaced with fructose-1,6-bisphosphatase. The levels of pyruvate are reduced in all three aconitase mutants, but noticeably so in *aco3*. The build up of citrate observed in *aco3* may explain why pyruvate levels are significantly reduced in *aco3*. Elevated levels of citrate negatively affect phosphofructokinase by feedback inhibition, reducing the flow of metabolites through glycolysis (Campbell and Reece, 2002).

Pyruvate is also involved in the metabolism of amino acids, serving as a precursor for the production of alanine, leucine and valine. Pyruvate is also a product during the production of methionine. The reduction in pyruvate levels observed in the mutants could be associated with the production of amino acids. Further work could include comparing the levels of alanine, leucine and valine by high performance liquid chromatography (HPLC) between the aconitase mutants.

The reduction in pyruvate levels could also be the result of reduced malate concentrations. In germinating lettuce embryos malate has been shown to be converted to pyruvate, which is metabolised to acetyl-CoA that enters the TCA cycle for metabolism by citrate synthase (Salon *et al.,* 1988).

6.9.4.8 Summary of the effects on metabolism in the aconitase mutants

A reduction in the intermediate metabolites succinic acid, malic acid and pyruvic acid were observed in both *aco1* and *aco3* suggesting a reduction in flux through metabolism. The identification of ACO1 within the cytosol (Bernard *et al.,* 2009) suggests that the glyoxylate cycle intermediate is metabolised though the TCA cycle.

However, the involvement of the intermediate metabolites in both cycles makes it difficult to determine if isocitrate is completely metabolised by the TCA cycle or is transported to the glyoxysome where metabolism continues.

The identification of ACO3 within the mitochondria (Bernard *et al.,* 2009) suggests a reduction in metabolism in *aco3* through the TCA cycle. The large build up of intermediate metabolites in combination with the production of citrate, suggests that the majority of citrate is metabolised through ACO3 during germination. Whilst the majority of citrate appears to be metabolised through ACO3, an additional aconitase must also function during germination as the *aco3* mutant is able to establish in the absence of an external carbon source. Enzymatic analysis conducted in chapter 5 shows that both ACO1 and ACO2 are able to catalyse the conversion of *cis*-aconitate to isocitrate.

The reduction in succinic acid, pyruvic acid and sucrose observed in *aco2* could suggest a reduction in flow through metabolism. However, malate levels remain unaltered and the levels of fumaric acid were elevated suggesting an additional unknown function. Aconitase has previously been shown to play an important role in cellular iron homeostasis in animal systems (Rouault, 2006) and in nitric oxide signalling pathways in tobacco (Navarre *et al.,* 2000). A reduction in aconitase activity occurred as a result of oxidative stress treatment with hydrogen peroxide (Verniquet *et al.,* 1991; Tan *et al.,* 2010) and interaction with copper in the mitochondria (Kung *et al.,* 2006; Tan *et al.,* 2010). Carrari *et al.* (2003) recorded an increase in the rate of photosynthesis in lpaco1 and suggest that the reduction in aconitase activity could disrupt a cellular signalling pathway. Unlike the mammalian aconitase, research conducted by Arnaud *et al.* (2007) concluded that the *Arabidopsis* isoforms are not able to operate as iron regulatory proteins. However, Moeder *et al.* (2007) have shown that *Arabidopsis* aconitase is able to bind to the mRNA of CuZn superoxide dismutase. Research conducted under oxidative stress shows that the levels of α-ketoglutarate (Sweetlove *et al.,* 2002; Lehmann *et al.,* 2008) and fumarate (Lehmann *et al.,* 2008) are elevated. The latter work also showed a reduction in the levels of the metabolites succinate and malate and the gene transcripts of α-ketoglutarate dehydrogenase, succinate dehydrogenase and ACO1 (Lehmann *et al.,* 2008). Studies of wild type *Arabidopsis* found that as the transcript

of ACO1 decreased, citrate levels and the transcript levels of ACO2 increased (Lehmann *et al.,* 2008). This could suggest that ACO2 plays a currently unknown role in oxidative stress. Further research could be conducted to determine if aconitase is able to bind to the UTR of other mRNAs, or if in fact it is able to bind to DNA and participate in mtDNA maintenance similar to the yeast aconitase 1 (Chen *et al.,* 2005; 2007)

6.9.5 Fatty acid metabolism during periods in the dark

Growth trials were conducted in the dark to determine if the glyoxylate cycle was blocked within the aconitase mutants. The glyoxylate cycle enzymes isocitrate lyase (icl) and malate synthase (ms) are both coded by a single gene in *Arabidopsis*. When grown in the dark the hypocotyl of both *icl* (Eastmond *et al.,* 2000) and *ms* (Cornah *et al.,* 2004) remained short. This short hypocotyl phenotype was eliminated in *icl* and *ms* by the addition of sucrose, suggesting a block in the conversion of fatty acids to sucrose. In the *icl* mutants Eastmond *et al.* (2000) suggested that fatty acids were broken down to citrate for respiration in the mitochondria. The export of carbon from the glyoxysome as citrate has since been confirmed by Pracharoenwattana *et al.* (2005). The respiration of fatty acid alone is not sufficient to maintain growth (Eastmond *et al.,* 2000). Additional carbon is required either through the glyoxylate cycle or the development of photosynthesis (Eastmond *et al.,* 2000).

When grown in the absence of light the aconitase mutants have elongated hypocotyl in both the presence and absence of sucrose (Figure 6.22). The normal phenotype observed in the dark suggests that the glyoxylate cycle is not blocked in *aco1*, *aco2* or *aco3*. The presence of a functioning glyoxylate cycle in *aco1*, *aco2* and *aco3* could be the result of a dual locality. A reduction in aconitase activity within both cytosolic and mitochondrial fractions of *aco1* and *aco3* was observed by Arnaud *et al.* (2007). A dual location would allow the glyoxylate cycle to continue in the absence of a single isoform. An alternative, more complex hypothesis could include the transport of citrate from the cytosol to the mitochondria for conversion to isocitrate. The resulting isocitrate could then be transported back to the glyoxysome for continued metabolism through the glyoxylate cycle.

6.9.6 [2-^{14}C] Acetate metabolism in *aco1*, *aco2* and *aco3*

Feeding studies were conducted with [2-^{14}C] acetate to try to determine how externally supplied acetate was metabolised by *aco1*, *aco2* and *aco3*. Growth trials in the presence of fluoroacetate have shown that all three mutants are able to metabolise externally supplied acetate. Previous studies have been conducted to compare the metabolism of acetate with [1-^{14}C] (Eastmond *et al.*, 2000) and [2-^{14}C] acetate (Cornah *et al.*, 2004; Turner *et al.*, 2005). [2-^{14}C] acetate was used within these trials rather than [1-^{14}C] acetate as less carbon is lost as CO_2 (Beevers, 1961; Canvin and Beevers, 1961). The feeding experiments were also conducted in the dark to prevent re-fixation of $^{14}CO_2$ (Cornah *et al.*, 2004).

During seedling establishment fatty acids are broken down by β-oxidation and metabolised through the glyoxylate cycle and gluconeogenesis to produce sucrose (Eastmond and Graham, 2001). Research has shown that the glyoxylate cycle is not essential for germination but is required for seedling establishment after periods of darkness (Eastmond *et al.*, 2000). Previous acetate feeding studies have shown that a block in the glyoxylate cycle reduces the conversion of acetate to sucrose. This was shown within *aae7/acn1* (Turner *et al.*, 2005), *icl* (Eastmond *et al.*, 2000; Cornah *et al.*, 2004) and *mls* (Cornah *et al.*, 2004) by a reduction in ^{14}C label in carbohydrate. A significant reduction in ^{14}C label was also identified in the carbohydrate fractions of *aco2* and *aco3* (Figure 6.23). The reduction in label, compared to Col-0, is consistent with a reduction in the production of carbohydrates from fatty acids.

The unaltered level of ^{14}C in the carbohydrate fraction of *aco1* suggests that all of the components of a functioning glyoxylate cycle are present. Unaltered levels of ^{14}C were also identified in the peroxisomal malate dehydrogenase double mutant (pmdh1pmdh2). Pracharoenwattana *et al.* (2007) showed that peroxisomal malate dehydrogenase does not function as part of the glyoxylate cycle. Instead, evidence suggests that malate dehydrogenase works in the reverse direction producing NAD for β-oxidation. This similarity between *aco1* and *pmdh1pmdh2* could suggest that ACO1 is involved in β-oxidation. However, *aco1* is able to establish in the absence of external sucrose (Figure 6.8) and therefore does not have a sucrose-dependent establishment phenotype. This is in contrast to previously characterised β-oxidation mutants (Pinfield-Wells *et al.*, 2005; Baker *et al.*, 2006; Pracharoenwattana *et al.*,

2007), suggesting that *aco1* is not involved in β-oxidation. Further analysis could be conducted to determine if the aconitase isoforms are involved in β-oxidation. The sensitivity of the aconitase mutants to 2,4-dichlorophenoxybutyric acid (2,4-DB) could be tested. 2,4-DB has been previously used to identify genes involved in β-oxidation (Hayashi *et al.,* 1998).

During seedling establishment carbon is believed to be exported from the glyoxylate cycle as either citrate or malate (Pracharoenwattana *et al.,* 2005; 2007). The direct route of metabolism for the exported citrate is currently undetermined. Citrate could be transported directly to the mitochondria or metabolised to isocitrate by cytosolic aconitase. The resulting isocitrate could then either re-enter the glyoxysome or be transported to the mitochondria. A disruption in either of the glyoxylate cycle enzymes icl or ms resulted in an increase in label in carbon dioxide (Cornah *et al.,* 2004). The differences observed in labelled carbon dioxide showed that the metabolism of acetate was less efficient in both *mls* and *icl* compared to wild type, as carbon is lost through the TCA cycle. A significant increase in radiolabeled CO_2 was observed in *aco1*, suggesting an increase in the metabolism of acetate through the TCA cycle. This observed increase in labelled CO_2 could be the result of a block or reduction in acetate metabolism through the glyoxylate cycle.

Interestingly, *aco2* showed a decrease in labelled CO_2. This reduction could suggest a reduction in the respiration of acetate through the TCA cycle in *aco2*. This is in keeping with the proposed mitochondrial location of ACO2 (Arnaud *et al.,* 2007; Bernard *et al.,* 2009).

The *aco3* mutant showed a slight but not significant increase in labelled CO_2 during incubation in the dark. This could suggest a slight reduction in metabolism through the glyoxylate cycle and an increase through ACO2. Another explanation could be that ACO3 is not involved in the respiration of externally supplied acetate in the dark, despite its obvious involvement in the light (see section 6.9.1). This observed difference in metabolism could be explained by Mohamed and Anderson (1983) who showed that purified aconitase can be light activated by washed thylakoid membranes.

The *aco2* mutant showed a significant decrease in label within the CO_2, carbohydrate and the insoluble fraction and an increase in organic acids. The increase observed in labelled organic acids may reflect the decrease in CO_2, carbohydrate and insoluble fraction. The significant differences observed between Col-0 and *aco2* could suggest that ACO2 is involved in the metabolism of exogenously supplied acetate during metabolism in seedlings. However it is not clear what function ACO2 plays in metabolism. The decrease in labelled carbohydrate observed in *aco2* suggests a block in the glyoxylate cycle. In contrast the reduction in labelled CO_2 suggests an involvement in metabolism through the TCA cycle. This conflicting evidence makes ACO2 an important isoform involved in the partitioning of citrate.

Another explanation for the labelling observed in *aco2* could result from the direct feeding of acetate into the mitochondria. This could result in the reduction of citrate metabolised through the TCA cycle. Exogenously supplied acetate could be converted to acetyl-CoA within either the plastid or cytosol of *Arabidopsis*, as acetyl-CoA synthetase activity has been detected (Behal *et al.*, 2002; Turner *et al.*, 2005). Turner *et al.* (2005) calculated that only 75 % of the externally supplied acetate that enters the glyoxysome is activated to acetyl-CoA by AAE7/ACN1 and subsequently metabolised to carbohydrates. This suggests that other routes of acetate utilisation are present. As acetyl-CoA synthetase activity has not been detected within the mitochondria, acetate may be activated within the cytosol and transported to the mitochondria. A carnitine acyl carrier like protein, present within the mitochondrial membrane (Lawland *et al.*, 2002) could transport acetyl-CoA into the mitochondria, completely bypassing the glyoxylate cycle. This could explain why the glyoxylate mutants *aae7/acn1* (Turner *et al.*, 2005) and *icl* (Eastmond *et al.*, 2000) are able to respire and assimilate exogenous acetate.

Further work could be conducted to determine the amount of radiolabelled acetate taken up by the mutants. Previous research has shown a difference in the uptake of ^{14}C acetate between glyoxylate cycle mutants (Eastmond *et al.*, 2000; Turner *et al.*, 2005). Whilst a reason for this is currently unknown, it would be useful to determine levels taken up by the aconitase mutants. This could be conducted by directly counting the initial radioactivity and a sample of the bubbling solution after 4 hours.

6.9.7 Summary

Results presented within this chapter suggest that all three aconitase isoforms are involved in acetate metabolism. Experiments showed that all three mutants are sensitive to the toxic analogue of acetate, fluoroacetate, with *aco3* showing the greatest sensitivity. Elevated levels of citrate observed in the GC-MS data from *aco3*, suggests a greater involvement in citrate metabolism. The greater sensitivity of *aco3* to fluoroacetate together with the elevated levels of citrate could be the result of the large contribution (~60 %) that ACO3 makes to total aconitase activity (see chapter 5).

Growth trials of *aco1*, *aco2* and *aco3* in the dark suggest that the mutants are not compromised in their ability to convert fatty acids to sucrose. This is in contrast to the *icl* mutant which has a sugar dependent hypocotyl elongation in the dark phenotype (Cornah *et al.*, 2004). The normal growth of the aconitase mutants suggests that they are able to produce glyoxylate for gluconeogenesis and succinate for respiration through the use of multiple aconitase isoforms.

^{14}C acetate studies in the dark show a significant difference in the metabolism of *aco2* compared to wild type and only a significant reduction in labelled soluble carbohydrates in *aco3*. This suggests that whilst ACO3 appears to be important for total aconitase activity and citrate metabolism, ACO2 is important for the metabolism of exogenously supplied acetate in the dark.

Mutant *aco1* showed little deviation from wild type in either citrate levels or radiolabelling studies. This suggests that little, if any, citrate is metabolised by the cytosolic aconitase, a step in the classic glyoxylate cycle.

Discussion, conclusion and future work

7.1 Introduction

Plants are very important organisms providing a large proportion of the world's food either directly (fruit and vegetables) or indirectly (via animals). This makes it important to study and understand how plants grow and function. Work conducted within this thesis concentrates on the process of seedling establishment in *Arabidopsis*, and in particular the proteins AAE7/ACN1 and aconitase.

During germination seedlings breakdown stored fatty acid reserves through β-oxidation, the glyoxylate cycle, the TCA cycle and gluconeogenesis in order to produce energy (Figure 7.1). Whilst a lot of research has been conducted on seedling germination, the whole process is not fully understood. Work conduced within this thesis aims to add to current knowledge by focusing on the enzymes AAE7/ACN1 and aconitase.

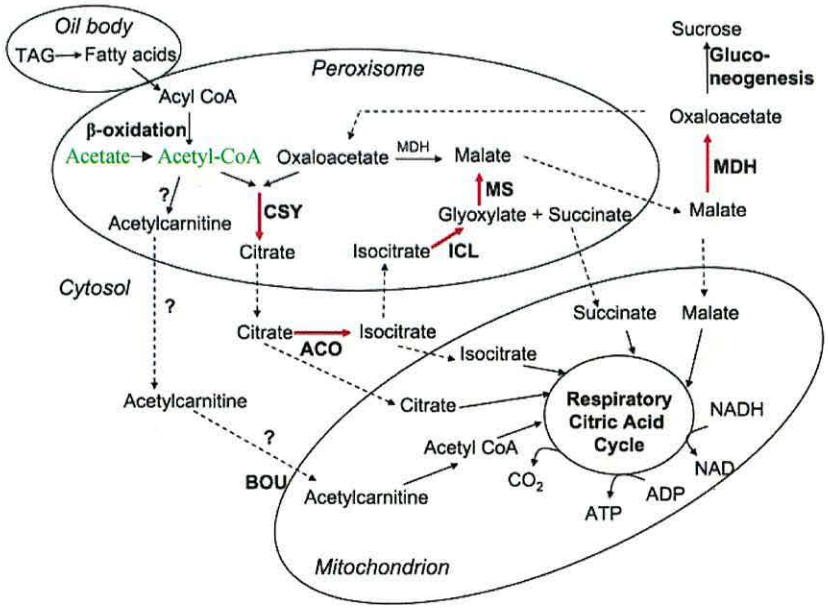


Figure 7.1: The main reactions involved in the peroxisome (also known as the glyoxysome) and routes of transport to the cytosol and mitochondria. The glyoxylate cycle reactions are shown in red. The reaction catalysed by AAE7/ACN1 is shown in green. Image modified from Eckardt, 2005.

Previous research has shown that AAE7/ACN1 is responsible for the conversion of acetate to acetyl-CoA for glyoxylate cycle metabolism (Turner *et al.*, 2005). Whilst

the metabolic reaction of AAE7/ACN1 is known, no structural analysis has been conducted to determine how this enzyme functions. The aim of work conducted within this thesis was to determine the previously unknown structure of AAE7/ACN1.

Published research shows that there are three isoforms of aconitase in *Arabidopsis* (Arnaud *et al.,* 2007; Moeder *et al.,* 2007). Current research also suggests that all three isoforms are able to function as aconitase, as a reduction in aconitase activity has been observed in the knockout mutants *aco1*, *aco2* and *aco3* (Arnaud *et al.,* 2007; Moeder *et al.,* 2007). However the location of expression and the specific cycle that each isoform participates in is currently unclear. Work conducted within this thesis aims to investigate the role each isoform plays in metabolism.

Prior to investigating the metabolic roles of the three aconitase isoforms in *Arabidopsis,* homozygous knockout mutants were identified and metabolic analysis was conducted using these mutants. In conjunction with this, aconitase genes were cloned and over-expressed in *Pichia pastoris* for structural characterisation.

7.2 Experimental structural analysis of AAE7/ACN1

AAE7/ACN1 has been identified as an acyl activating enzyme (AAE) by amino acid sequence comparisons with LACS (Shockey *et al.,* 2003). During this analysis Shockey *et al.* (2003) also identified sixty two other members of the acyl-activating enzyme family. Whilst metabolic analysis has been initiated on this large family of proteins, no structural analysis has been previously conducted. Structural analysis of a member of this family is important to understand how they function. Within this thesis structural analysis was conducted in the form of solution scattering, mass spectrometry and x-ray crystallography.

7.2.1 Solution scattering of AAE7/ACN1

Evidence from solution scattering analysis, presented in chapter 3, shows AAE7/ACN1 to be a multimeric protein comprised of two domains. Enzymatic analysis also presented in chapter 3 shows that AAE7/ACN1 is inhibited by AMP and AMP analogues. Solution scattering analysis conducted both in the presence and

absence of a propyl ester of AMP (AMPP) was unable to detect a conformational change.

The solution scattering profiles of ttLC-FACS were predicted in chapter 3 from crystallographic structures in both open and closed conformations. Comparison of these solution scattering profiles to the experimental data of AAE7/ACN1 shows that AAE7/ACN1 adopts an open conformation in solution. Previously published crystallographic analysis also shows that unligated acyl-CoA synthetase can adopt either an open or closed conformation (Hisanaga *et al.,* 2004; Kochan *et al.,* 2009; Shah *et al.,* 2009).

Previous research shows that the multimeric state of an acyl-CoA synthetase is not unique to AAE7/ACN1. Previously characterised acyl-CoA synthetases have been characterised as a dimer (Preston *et al.,* 1990; Martinez-Blanco *et al.,* 1992), trimer (Jogl and Tong, 2004) and octomer (Brasen *et al.,* 2005). Previously published crystallographic analysis also shows that other acetyl-CoA synthetase's contain two domains, with the active site located between them (Gulick *et al.,* 2003; Jogl and Tong, 2004).

Whilst current published research does not include inhibition analysis of acetyl-CoA synthetase using AMPP, as in the experiments presented in chapter 3 of this thesis, it has been shown to inhibit a related acyl-CoA synthetase, propionyl-CoA synthetase (Horswill and Escalante-Semerena, 2002). Other published work includes the crystallographic structure of seACS, which was determined in the presence of AMPP (Gulick *et al.,* 2003). Together this would indicate that AMPP is able to bind, interact and stabilise acetyl-CoA synthetase structures.

Crystallographic structures of acyl-CoA synthetase have been previously determined in an open conformation bound to AMP, ATP and AMPCPP (Jogl and Tong, 2004; Kochan *et al.,* 2009). The crystallographic structure of acyl-CoA synthetase has also been determined in a closed position in the presence of a) AMPP and CoA, b) butyryl-CoA, AMP and magnesium and c) myristoyl-AMP and magnesium (Gulick *et al.,* 2003; Hisanaga *et al.,* 2004; Kochan *et al.,* 2009). Therefore additional

substrates may be required in order to create a stable closed conformation of AAE7/ACN1.

7.2.2 Further experimental analysis for AAE7/ACN1

Further work could include determining the structure or conformation of AAE7/ACN1 in the presence of additional ligands (e.g. acetyl-CoA, acetyl-AMP, AMP, ATP, CoA, magnesium). The addition of ligands to AAE7/ACN1 may increase the number of interactions required to create a closed conformation.

7.2.3 Determining the multimeric state of AAE7/ACN1

The multimeric state of AAE7/ACN1 was investigated within chapter 3 using mass spectrometry, CRYSOL analysis, gel filtration and native PAGE.

Mass spectrometry analysis identified the presence of both monomeric and dimeric species of AAE7/ACN1 in the spectrum. The deconvoluted spectrum showed that the dimeric form of AAE7/ACN1 was more dominant in solution. However, the presence of monomeric AAE7/ACN1 in the spectrum was believed to be due to the harsh buffers that disassociate dimeric molecules into monomers.

CRYSOL analysis conducted in chapter 3 also showed that AAE7/ACN1 is a dimeric protein in solution. CRYSOL predicted the solution scattering profile of acyl-CoA synthetase from a monomeric, dimeric and trimeric crystallographic structure. These structures were then compared to the solution scattering profile of AAE7/ACN1 presented in chapter 3. Visual comparisons showed that AAE7/ACN1 is a dimeric protein.

However, native PAGE and gel filtration analysis of AAE7/ACN1 were inconclusive. During native PAGE analysis AAE7/ACN1 appeared to drop out of solution and remain on the gel surface within the well. This could to be due to the high pI of AAE7/ACN1 coupled with the protein's instability at pH 10, or the temperature at which the experiments were conducted.

Gel filtration analysis of AAE7/ACN1 was also conducted at room temperature. The majority of the protein sample loaded onto the column remained on the column's

surface. Any protein that successfully eluted from the column appeared to be monomeric. However the application of this technique at room temperature may have lead to the dissociation and precipitation of AAE7/ACN1.

Previously published work has shown that acetyl-CoA synthetase from *Bradyrhizobium japonicum* and *Penicillium chrysogenum* are dimeric proteins (Preston *et al.,* 1990; Martinez-Blanco *et al.,* 1992). The dimeric structure of acetyl-CoA synthetase from *B. japonicum* was successfully determined by gel permeation chromatography. However, Preston *et al.* (1990) did note that purified acetyl-CoA synthetase can only be stored at room temperature for 20 minutes before an activity loss of greater than 90 % was observed.

Interestingly, within *Arabidopsis* cells AAE7/ACN1 is able to function at a room temperature of 20°C. This makes the instability of AAE7/ACN1 at room temperature surprising. However, the conditions present within the glyoxysome were not mimicked during experimental analysis. Therefore there may be a currently undetermined factor within the glyoxysome that stabilises the structure of AAE7/ACN1 enabling it to function as a dimeric protein *in vivo*.

7.2.4 Further work to determine the multimeric state of AAE7/ACN1

In order to determine the dimeric state of AAE7/ACN1 gel filtration analysis could be repeated at 4°C to see if a reduction in temperature increases the protein's stability. Gel filtration analysis could also be conducted in the presence of ligands to try to stabilise the multimeric structure of AAE7/ACN1.

Additional work could include the use of analytical ultracentrifugation or dynamic light scattering. Durin g analytical ultracentrifugation the molecular mass of AAE7/ACN1 could be determined by centrifugation at a specified speed and temperature. Alternatively, dynamic light scattering could be used to determine the size of AAE7/ACN1 by measuring the amount of scattered light.

7.2.5 X-ray crystallisation of AAE7/ACN1

Attempts to obtain crystals of AAE7/ACN1 for x-ray crystallography within this thesis have been unsuccessful. A variety of different crystallisation trials have been conducted varying the reservoir solution, temperature, protein concentration and ligands present.

Previously published work shows that AAE7/ACN1 is able to use both acetate and butyrate as a substrate (Shockey *et al.,* 2003; Turner *et al.,* 2005). Research conducted by Shockey *et al.* (2003) with an artificial membrane suggests that AAE7/ACN1 prefers to utilise butyrate. In contrast, analysis conducted by Turner *et al.* (2005) suggests that AAE7/ACN1 prefers to use acetate as a substrate. The differences observed in substrate specificity are believed to be due to the presence or absence of an artificial membrane (Turner *et al.,* 2005).

The crystallisation difficulties encountered within this thesis could suggest that AAE7/ACN1 is a membrane bound protein. If AAE7/ACN1 is indeed a membrane bound protein, then the set up used is not entirely suitable for the crystallisation of AAE7/ACN1 and thus different crystallisation techniques would need to be used.

7.2.6 Further work to obtained crystals of AAE7/ACN1

In order to obtain crystals of AAE7/ACN1 a much larger screening process is required covering a variety of reservoir solutions. Factors that could be changed within the reservoir solution include pH, buffer and ionic strength. In addition to altering the reservoir solution a variety of different ligand combinations could be used to stabilise the structure of AAE7/ACN1.

Further work could also be conducted to determine if AAE7/ACN1 is a membrane bound protein. If AAE7/ACN1 is associated with the plasma membrane then future crystallisation trials could be conducted in the presence of a detergent, such as decyl-maltoside, octyltetraoxyethylene (C_8E_4), n-octyl-β-D-glucopyranoside (OG), *n*-dodecyl-β-D-maltopyranoside (DDM) or triton X-100 (Rigaud *et al.,* 2000; Hunte and Michel, 2002; Prive, 2007).

7.3 Theoretical structural analysis of AAE7/ACN1

Due to the problems encountered with x-ray crystallography, homology modelling of AAE7/ACN1 was conducted as an alternative method of structural analysis.

7.3.1 Homology modelling of AAE7/ACN1

The homology models of AAE7/ACN1 in chapter 4 were constructed with the modelling programs I-TASSER and Modeller. The online program I-TASSER was used because of its recent success in CASP analysis (Zhou *et al.*, 2007; Zhang, 2009). Alternatively, Modeller was used because of its flexibility. The users of Modeller have greater control over the model created, as they are able to choose a specific template structure.

All homology models constructed within chapter 4 were based on the crystallographic structures of the dimeric ttLC-FACS. Sequence and structural comparisons were made between the template structures of ttLC-FACS and the homology models of AAE7/ACN1.

Comparisons show that important residues involved in AMP binding are conserved in AAE7/ACN1. These sequences include the adenine motif and the phosphate loop. Other sequences that were compared include the linker and gate motifs. Important residues within the linker motif are conserved in AAE7/ACN1, suggesting that the protein is capable of undergoing a conformational change. However, residues within this sequence are also involved in catalysis.

Homology models were constructed of AAE7/ACN1 in both an open and a closed conformation. Structural comparison of the hinge region of ttLC-FACS suggests that AAE7/ACN1 is able to undergo a conformational change during catalysis. Previously published work shows that acyl-CoA synthetase undergoes a conformational change during catalysis (Bar-Tana and Rose, 1968; Gulick *et al.*, 2003; Jogl and Tong, 2004). Crystallographic structures of acyl-CoA synthetase have been determined in both an open and closed conformation (Hisanaga *et al.*, 2004; Kochan *et al.*, 2009).

Important residues surrounding the fatty acid entrance site of ttLC-FACS were also compared to AAE7/ACN1 in chapter 4. Structural comparisons show that important residues are not conserved in AAE7/ACN1. This suggests that the fatty acid substrate of AAE7/ACN1 does not enter the active site in the same way as ttLC-FACS. This could also suggest that AAE7/ACN1 is not bound to the plasma membrane, as the fatty acid entrance site of ttLC-FACS is attached to the membrane. If this is the case, then this would cast doubt on published work (Shockey *et al.*, 2003) that suggested AAE7/ACN1 to be a membrane bound protein.

Structural comparison of the fatty acid binding region of AAE7/ACN1, suggests that it is able to use short chain fatty acid substrates. Analysis shows that a large number of residues, in the region of ttLC-FACS, responsible for the conversion of long chain fatty acids, are not conserved between AAE7/ACN1 and ttLC-FACS.

The ttLC-FACS is a membrane bound protein that is able to use long chain fatty acids as substrates (Hisanaga *et al.*, 2004). The fatty acid substrate of long chain fatty acyl-CoA synthetase (LC-FACS) enters via a fatty acid binding tunnel (Hisanaga *et al.*, 2004). In contrast, characterised short chain fatty acyl-CoA synthetases (SC-FACS) are not associated with the plasma membrane and therefore all substrates enter and leave via a single opening. Homology modelling analysis conducted within this thesis suggests that the mechanism of AAE7/ACN1 catalysis is similar to that of SC-FACS. This is in-keeping with the proposed substrate specificity of AAE7/ACN1 suggested by Shockey *et al.* (2003) and Turner *et al.* (2005). This also suggests that AAE7/ACN1 is not bound to the plasma membrane.

7.3.2 Further homology modelling analysis of AAE7/ACN1

Further work could include homology modelling of AAE7/ACN1 to a SC-FACS for comparison. Despite the lower level of sequence homology between AAE7/ACN1 and crystallographic structures of SC-FACS, they utilise the same fatty acid substrates. However, the crystallographic structure of SC-FACS from a single organism has only been determined in a single conformation. Therefore, AAE7/ACN1 could also be modelled to the medium chain fatty acyl-CoA synthetase (MC-FACS) from *H. sapiens*, as its structure has been determined in both an open and closed conformation.

7.4 Experimental and theoretical structural analysis of AAE7/ACN1

Structural analysis of AAE7/ACN1 presented in chapters 3 and 4 shows that AAE7/ACN1 is a dimeric protein comprised of two domains. Whilst solution scattering was unable to detect a conformational change, homology modelling suggests that the residues involved in domain movement are conserved. The domain movement of AAE7/ACN1 could be further investigated with solution scattering analysis in the presence of additional ligands.

Crystallographic analysis of AAE7/ACN1, conducted in chapter 3, was unsuccessful. One explanation for this is that AAE7/ACN1 may be a membrane bound protein. AAE7/ACN1 was identified by Shockey *et al.* (2003) using the amino acid sequence of LACS, which are membrane bound proteins. This led to the hypothesis that AAE7/ACN1 is itself a membrane bound protein. Therefore the homology modelling analysis conducted in chapter 4 was based on the crystallographic structure of a membrane bound protein (ttLC-FACS). However, sequence and structural analysis indicates that the mechanism of action of AAE7/ACN1 is not the same as that of ttLC-FACS. This could perhaps suggest that AAE7/ACN1 is not a membrane bound protein. Therefore further experimental analysis is required to determine if AAE7/ACN1 is able to bind to the plasma membrane.

If AAE7/ACN1 is a membrane bound protein, then plausibly it has a relationship to acetate/butyrate membrane transport proteins such as comatose (Hooks *et al.,* 2007). If so, being membrane bound could make the reaction more effective at metabolising external sources of acetate/butyrate.

Previous research has shown that microorganisms in the soil can both utilise and produce organic acids (Baumann, 1968; Kusel and Drake, 1994; Limmer and Drake, 1998) such as acetic and butyric. Furthermore, radiolabeled feeding studies show that externally supplied acetate and butyrate can be taken up by plants (Tramontano and Barreiro, 1997; Cornah *et al.*, 2004; Turner *et al.*, 2005).

It has been suggested that the main role of AAE7/ACN1 is to recycle acetate to acetyl-CoA (Hooks *et al.,* 2010), conserving carbon produced during the synthesis of

cysteine and ornithine and the deacetylation of histones (Bao *et al.,* 2000; Turner *et al.,* 2005). However, *Arabidopsis* seedlings are able to take up and metabolise external acetate (Turner *et al.,* 2005). The identification of three possible acetyl-CoA synthetases in *Arabidopsis* suggests that acetate can be converted to acetyl-CoA. The uptake and conversion of acetate to acetyl-CoA removes this organic acid from the growth medium and prevents acidification of the cytoplasm, respectively. However the metabolism of external butyrate is currently unclear as a butyryl-CoA synthetase has not been identified within *Arabidopsis*, although AAE7/ACN1 has been shown to be active with butyrate (Shockey *et al.,* 2003; Turner *et al.*, 2005).

Previously published growth trials of *Arabidopsis* in the presence of elevated levels of either acetate or butyrate show a reduction in seedling weight (Turner *et al.,* 2005). A similar phenotype has been observed in the presence of 3.5 mM acetate and 1 mM butyrate, suggesting that butyrate has a greater toxic effect on seedling growth (Turner *et al.,* 2005). One explanation for why *Arabidopsis* seedlings may show greater tolerance to external acetate compared to butyrate, is that additional acetyl-CoA synthetases are thought to function within the plant's cell. For example a second acetyl-CoA synthetase has been identified within the plastid (Ke *et al.,* 2000; Turner *et al.,* 2005) and a third suggested within the cytosol of *Arabidopsis* (Turner *et al.,* 2005; Hooks *et al.,* 2010). These may help to detoxify external acetate. However to date, no other butyryl-CoA synthetases have been identified within *Arabidopsis.* Whilst published research shows that AAE7/ACN1 is involved in acetate metabolism (Hooks *et al.,* 2004; Turner *et al.,* 2005; Hooks *et al.,* 2010), only limited work has been conducted on butyrate metabolism with purified AAE7/ACN1 (Shockey *et al.*, 2003; Turner *et al.*, 2005) and seedlings (Turner *et al.*, 2005) to date. Therefore the importance of butyrate metabolism in *aae7/acn1* could be the basis of further work.

7.5 Metabolic function of aconitase

Previously published work has shown that the first exit route for carbon from the glyoxylate cycle is in the form of citrate (Pracharoenwattana *et al.,* 2005). This exported citrate is believed to be converted to isocitrate within the cytosol, and then transported back to the glyoxysome (Eastmond and Graham, 2001). However, malate synthase mutants, with a block in the glyoxylate cycle, show that carbon from

the glyoxylate cycle can be metabolised in the mitochondria by the TCA cycle (Cornah *et al.,* 2004). This therefore suggests that citrate from the glyoxysome is transported to the mitochondria, which links in to the work within this thesis that shows that all three isoforms of aconitase participate in metabolism during seedling establishment.

7.5.1 Metabolic function of ACO1

Growth trials conducted in both the light and dark show no physical differences between Col-0 and *aco1*. The normal hypocotyl growth observed in the dark of *aco1* suggests that the glyoxylate cycle is not blocked. This could suggest that either ACO1 does not participate in the glyoxylate cycle or that multiple aconitase isoforms provide isocitrate for the glyoxylate cycle. Radioactive feeding studies of *aco1* presented in chapter 6 show an increase in labelled CO_2, suggesting an involvement in the glyoxylate cycle.

The altered levels of metabolites observed in GC-MS data presented in chapter 6 show that ACO1 participates in metabolism. However, the intermediate metabolites analysed are involved in both the glyoxylate and TCA cycle making it difficult to determine which cycle ACO1 belongs too. A reduction in total aconitase activity (via enzyme assay) of approximately 23 % was observed in *aco1*, compared to Col-0. This shows that ACO1 is involved in the conversion of citrate to isocitrate during seedling establishment.

Fluoroacetate screening studies show that *aco1* is sensitive to fluoroacetate. The sensitivity of *aco1* to fluoroacetate suggests that externally supplied acetate is being metabolised by an additional aconitase isoform. This could be the result of multiple aconitase isoforms functioning as part of the glyoxylate cycle or direct feeding of fluoroacetate into the TCA cycle.

The cytosolic location of ACO1, identified by a reduction in activity within the cytosolic fraction of the cell, suggests an involvement in the glyoxylate cycle (Arnaud *et al.,* 2007; Bernard *et al.,* 2009). Enzymatic analysis of the cytosol has also shown that multiple isoforms of aconitase are active within the cytosol (Arnaud *et al.,* 2007). The identification of both ACO1 and ACO3 activity within the cytosol

and mitochondria led to the conclusion that these two isoforms are able to compensate for one another (Arnaud *et al.,* 2007). Attempts to create a double knockout mutant for both ACO1 and ACO3 have also been unsuccessful (Arnaud *et al.,* 2007). This suggests that multiple isoforms are important for successful seed development and growth (Arnaud *et al.,* 2007).

A 75 % reduction in cytosolic aconitase activity has been previously identified by enzymatic analysis (Arnaud *et al.,* 2007). The cytosolic fraction of the cell was separated from the cellular membrane and organelles by centrifugation from whole plants (Arnaud *et al.,* 2007). In contrast, only a 20 % reduction in total aconitase activity has been previously observed in leaf tissue from 4 week old plants (Moeder *et al.,* 2007). The differences observed for the reduction in aconitase activity may be due to the different types of tissue used (e.g. whole plant, leaf tissue and seedlings). Expression analysis using northern blots showed that ACO1 mRNA was more abundant in roots, stems and caulinary leaves (Arnaud *et al.,* 2007). A larger reduction in aconitase activity may have been observed in whole plant extracts as they contain these plant tissues. In comparison, the two day old seedlings analysed within this thesis contain cotyledons and a developing root system. This may explain why only a 23 % reduction in ACO1 activity was observed between *aco1* and Col-0 in chapter 5.

Published work and results presented in chapter 6 show that ACO1 is involved in the glyoxylate cycle. However, the reduction in total aconitase activity of 23 % observed within chapter 5 suggests that the majority of citrate is not being metabolised through ACO1. Published results and work presented within this thesis both suggest that multiple isoforms of aconitase are involved in citrate metabolism.

7.5.2 Metabolic function of ACO2

Metabolic analysis of the aconitase mutants with [2-^{14}C] acetate in the dark showed the greatest difference between Col-0 and *aco2*. The significant reduction in labelled CO_2 observed in *aco2* suggests that ACO2 is involved in the TCA cycle.

In the light, metabolism is believed to change from mitochondrial respiration to photorespiration (Kromer, 1995). This change in metabolism is believed to reduce

the flow of metabolites through the TCA cycle (Gemel and Randall, 1992). If activity through the TCA cycle is increased in the dark, compared to the light, then perhaps ACO2 makes a larger contribution to metabolism in the dark. This may explain why the largest difference in [2-^{14}C] acetate feeding was observed between Col-0 and *aco2*.

A link between aconitase activity and photosynthetic rate has been previously observed in plants (Carrari *et al.*, 2003). In tomatoes a reduction in aconitase activity resulted in an increase in photosynthesis (Carrari *et al.*, 2003).

Previously published work has shown that ACO2 is located within the mitochondrial proteome (Millar *et al.*, 2001). Enzymatic analysis also shows that ACO2 is located within the mitochondria, where it is believed to function as part of the TCA cycle (Arnaud *et al.*, 2007; Bernard *et al.*, 2009). In light of this evidence, ACO2 is involved in the conversion of citrate to isocitrate as part of the TCA cycle in the mitochondria.

7.5.3 Metabolic function of ACO3

A 63 % reduction in total aconitase activity, a build up of citrate and a delay in germination were all observed in *aco3*. These results suggest that ACO3 makes a large contribution to citrate metabolism during seedling establishment. The large differences observed in the GC-MS data of *aco3* also suggest that the majority of citrate is metabolised through ACO3. Feeding studies with [2-^{14}C] acetate presented in chapter 6 were inconclusive as a significant difference in labelled CO_2 between Col-0 and *aco3* was not observed.

A 70 % reduction in total aconitase activity has been previously observed in four week old leaf tissue from *aco3* (Moeder *et al.*, 2007). A 55 % reduction in aconitase activity has also been identified in the mitochondrial fraction of *aco3* (Arnaud *et al.*, 2007). The identification of ACO3 within the mitochondria led to the conclusion that ACO3 is involved in the TCA cycle (Arnaud *et al.*, 2007; Bernard *et al.*, 2009). However, a reduction in aconitase activity has also been identified within the cytosolic fraction of *aco3* (Arnaud *et al.*, 2007).

Previously published work suggests that cytosolic aconitase participates in the glyoxylate cycle during germination (Hayashi *et al.,* 1995). The mitochondrial location of ACO3 suggests that this isoform does not participate in the glyoxylate cycle. Therefore the large involvement of ACO3 in metabolism, demonstrated within this thesis, contradicts previously published work. However, a link between ACO3 and glyoxysomal enzymes has been identified by Lee *et al.* (2008). Microarray analysis showed that ACO3 is co-expressed with β-oxidation enzymes (Lee *et al.,* 2008). This suggests that ACO3 plays a role in the mitochondrial utilisation of citrate synthesised by the glyoxysome (Lee *et al.,* 2008).

The results presented within this thesis suggest that the ACO3 isoform of aconitase makes the largest contribution to citrate metabolism. Previous published work and results presented within this thesis suggest that ACO3 metabolises citrate, produced within the glyoxysome, to isocitrate.

7.5.4 Further work

In order to further investigate the function of the aconitase isoforms a transcriptional approach could be initiated. This further work could include RT-PCR or microarray analysis to quantify the mRNA levels of ACO1, ACO2 and ACO3. Gene expression analysis of all three aconitase isoforms could be compared throughout seedling establishment in a variety of tissue types. Microarray analysis could also be conducted on *aco1*, *aco2* and *aco3* to determine what effect the knockout mutants have on the expression of other genes.

7.6 Location of aconitase isoforms in *Arabidopsis*

During germination fatty acids are believed to be broken down by the glyoxylate cycle (Eastmond and Graham, 2001). However, aconitase activity has not been identified within the glyoxysome (De Bellis *et al.,* 1994). The glyoxylate cycle is therefore believed to take a detour via the cytosol, where aconitase activity has been detected in sycamore (Brouquisse *et al.,* 1987) and pumpkin (Hayashi *et al.,* 1995).

Published work on other plant species has also identified multiple isoforms of aconitase (Wendel *et al.,* 1988; De Bellis *et al.,* 1993; Cots and Widmer *et al.,* 1999). For instance, three isoforms of aconitase have been identified in pumpkin, ACO I,

ACO II and ACO III (De Bellis *et al.,* 1993). Experimental analysis suggests that ACO I and ACO III are located within the cytosol and ACO II within the mitochondria (De Bellis *et al.,* 1995). The work in this thesis also shows differences in location between the three different *Arabidopsis* aconitases.

7.6.1 Location of ACO1 in *Arabidopsis*

Feeding studies with [2-^{14}C] acetate presented in chapter 6 showed a significant increase in labelled CO_2 in *aco1*. An increase in labelled CO_2 suggests a reduction in metabolism through the glyoxylate cycle. This suggests that ACO1 functions as part of the glyoxylate cycle within the cytosol. Sequence alignment analysis shows that ACO1 lacks the N-terminal leader sequence seen in both ACO2 and ACO3. The lack of an N-terminal leader sequence indicates that ACO1 does not have a mitochondrial location. The use of online protein prediction programs in chapter 5 also suggests that ACO1 is not located within the mitochondria.

Previous research also suggests that ACO1 is located within the cytosol (Bernard *et al.,* 2009). Bernard *et al.* (2009) isolated the cytosolic fraction of from wild type and *aco1* leaf tissue by centrifugation. The resulting fractions were run on a non-denaturing gel and subjected to in-gel activity assays. Visual comparison of the activity gels suggested a cytosolic location for ACO1 (Bernard *et al.,* 2009). Enzymatic analysis of *aco1* has shown a reduction in both the cytosolic and mitochondrial fraction of 10 day old plants (Arnaud *et al.,* 2007). This could suggest dual locality, although the work showed a far greater reduction in cytosolic activity than mitochondrial activity. The cytosolic and mitochondrial fractions of *aco1* were separated by centrifugation and assayed for aconitase activity. Whilst dual locality of ACO1 may be correct, it is also interesting to note that only one study has identified ACO1 within the mitochondrial proteome (Kruft *et al.,* 2001). However this identification was believed to be caused by contamination (Kruft *et al.,* 2001). Therefore whilst it would seem from the findings of this work and published papers that ACO1 is located in the cytosol, there remains the possibility that ACO1 is also located in the mitochondria.

7.6.2 Location of ACO2 in *Arabidopsis*

Radioactive feeding studies with [2-^{14}C] acetate showed a significant decrease in label in the CO_2 fraction of *aco2*. The decrease in labelled CO_2 presented in chapter 6 suggests a reduction in acetate metabolism through the TCA cycle. A reduction in metabolism through the TCA cycle in *aco2* would therefore suggest a mitochondrial location for ACO2. Sequence alignment analysis of the aconitase isoforms conducted in chapter 5 shows that ACO2 contains an N-terminal leader sequence that may be responsible for transportation into the mitochondria. Three independent online protein prediction programs Predotar (Small *et al.*, 2004), MitoProt II (Claros and Vinces, 1996) and iPSORT (Bannai *et al.*, 2002) were used in chapter 5. All three programs predicted a mitochondrial location for ACO2.

Previously published work has also identified ACO2 within the mitochondrial proteome by mass spectrometry of mitochondrial proteins separated using two dimensional gel electrophoresis (Millar *et al.*, 2001; Lee *et al.*, 2008). A mitochondrial location for ACO2 has also been suggested by enzymatic analysis of *aco2* mutants (Arnaud *et al.*, 2007). A reduction in aconitase activity within the mitochondria of *aco2* mutants resulted in the conclusion that ACO2 is located within the mitochondria (Arnaud *et al.*, 2007). A mitochondrial location has also been shown for ACO2 by cellular fractionation (Bernard *et al.*, 2009). Thus in light of all of this corroborating evidence, it would seem that ACO2 in *Arabidopsis* has a mitochondrial location.

7.6.3 Location of ACO3 in *Arabidopsis*

Feeding studies with [2-^{14}C] acetate presented in chapter 6 were inconclusive as a significant difference was not observed between the CO_2 levels of *aco3* and Col-0. Sequence comparison analysis presented in chapter 5 shows that ACO3 has an N-terminal leader sequence. This sequence could be responsible for transport into the mitochondria, suggesting a mitochondrial location.

Published research has identified ACO3 within the mitochondrial proteome (Millar *et al.*, 2001; Lee *et al.*, 2008). In these studies mitochondrial proteins were isolated from *Arabidopsis* tissue, separated by two dimensional gel electrophoresis and identified by mass spectrometry. Cellular fractionation has also identified ACO3

within the mitochondria (Bernard *et al.,* 2009). Enzymatic analysis conducted with *aco3* showed a reduction in aconitase activity within both the mitochondria and the cytosol, suggesting dual locality (Arnaud *et al.,* 2007). It therefore would seem plausible that ACO3 has a dual location in *Arabidopsis* that functions primarily in the mitochondria.

7.6.4 Further work to determine the location of aconitase isoforms

In order to determine the locations of aconitase isoforms, expression reporter genes such as coloured fluorescence proteins (e.g. green, yellow or blue) could be used. Reporter genes can be fused to the gene of interest so that they are both under the control of the same promoter. The fused genes are then transcribed and translated into protein. The expression of the reporter gene can then be detected visually by microscopy. If different coloured reporter genes are used to determine the expression of ACO1, ACO2 and ACO3 in *Arabidopsis*, then the different genotypes could be crossed to determine if the aconitase isoforms are expressed in the same location.

7.7 Structural analysis of aconitase

Aconitase is believed to contain a [4Fe-4S] cluster that is readily converted to a [3Fe-4S] cluster, resulting in inactivation. Previous attempts to over-express and purify plant aconitase have failed to produce a fully active enzyme. A method of over-expressing pure, active aconitase without the need for re-activation after purification has been identified within this thesis.

7.7.1 ACO1

The successful over-expression, purification and activation of ACO1 from *P. pastoris* in chapter 5 shows that ACO1 functions as aconitase. Two methods of aconitase activation were used within this thesis. The first method activates aconitase after purification with addition of iron. The second method activates aconitase during induction within *P. pastoris* cells. The second method of activation resulted in a more active form of aconitase. The successful activation of aconitase required the addition of iron to either the growth medium or incubation mixture after purification. This suggests that ACO1 requires an iron sulphur cluster for activation.

Sequence analysis of ACO1 shows that the amino acid sequences of the *Arabidopsis* aconitase isoforms are very similar to hsIRP1. Homology modelling analysis shows that the structure of ACO1 is also very similar to that of hsIRP1. Homology modelling analysis also showed that important residues involved in substrate recognition, cluster ligation and catalysis of hsIRP1 are conserved in ACO1. This suggests that the structure and method of catalysis of ACO1 is similar to that of hsIRP1.

Sequence comparison and homology modelling analysis conducted in chapter 5 shows that the linker sequence of ACO1 is a similar length to that of hsIRP1. This suggests that ACO1 is able to undergo a conformational change, where domains 3 and 4 move apart. This domain movement may allow ACO1 to bind to either RNA or DNA. Previously published work shows that ACO1 is able to bind to the mRNA of CuZn superoxide dismutase (Moeder *et al.*, 2007).

A reduction in aconitase activity in *aco1* has been observed in both chapter 6 and previously published work. The reduction in aconitase activity shows that ACO1 has aconitase activity *in vivo* (Arnaud *et al.*, 2007; Moeder *et al.*, 2007). Published research shows that current methods of aconitase activation do not result in fully active aconitase (Navarre *et al.*, 2000). Over-expression, purification and activation of tobacco aconitase resulted in a lower activity than that purified from leaf tissue (Navarre *et al.*, 2000). A method of obtaining fully functioning aconitase is required for accurate structural analysis. Whilst further work is required to determine if ACO1 is fully active, the activity measurements of ACO1 and ACO2 show that this method is a development over previous methods of activation.

7.7.2 ACO2

The *Arabidopsis* aconitase isoform ACO2 has been successfully over-expressed and purified in *P. pastoris* within chapter 5. Initial activation methods of ACO2 after purification failed to produce detectable aconitase activity. However, a second method of ACO2 activation with the addition of iron during induction was successful. Activation with the addition of iron suggests that active ACO2 contains an iron sulphur cluster.

Sequence analysis of ACO2 presented in chapter 5 shows that it's amino acid sequence is similar to that of hsIRP1. Homology modelling analysis shows that the structure of ACO2 is similar to that of hsIRP1 and important residues involved in substrate recognition, cluster ligation and catalysis are conserved.

Enzymatic analysis of *aco2* has shown a reduction in aconitase activity both within chapter 5 of this thesis and previously published work (Arnaud *et al.,* 2007; Moeder *et al.,* 2007). The reduction in aconitase activity observed in *aco2* and the successful activation of ACO2 show that this isoform of aconitase is able to function as aconitase. A slight decrease in mitochondrial aconitase activity has been observed in *aco2* during conditions of iron starvation (Arnaud *et al.,* 2007). This may also suggest that ACO2 contains an iron sulphur cluster.

Tobacco aconitase has been previously over-expressed in *E. coli* inclusion bodies (Navarre *et al.,* 2000). Once purified, aconitase has been activated with reducing agents and iron (Navarre *et al.,* 2000). However, activation after purification did not result in a fully active protein (Navarre *et al.,* 2000). The activation of ACO2 presented in chapter 5 occurs during induction in *P. pastoris* cells. This *in vivo* activation removes the need for extraction from inclusion bodies and protein re-folding after purification. This may result in a correctly folded, active from of ACO2 that represents the isoform found in *Arabidopsis* cells.

7.7.3 ACO3

Sequence analysis conducted in chapter 5 shows that all three *Arabidopsis* aconitase isoforms share over 70 % identity. Whilst over-expression analysis of ACO3 in *P. pastoris* failed to identify an over-expressing line, ACO3 was successfully cloned into *P. pastoris*. The amino acid sequence similarity observed been the aconitase isoforms suggests that, with time, a colony over-expressing ACO3 should be identified.

Sequence analysis and homology modelling of ACO3 show that all of the important residues involved in substrate recognition, cluster ligation and catalysis are conserved. The conservation of these residues between hsIRP1 and ACO3 suggests that ACO3 is able to metabolise citrate.

Previously published research and work presented in chapter 5 show a reduction in aconitase activity in *aco3* (Moeder *et al.,* 2007; Arnaud *et al.,* 2007). This analysis also suggests that ACO3 is able to metabolise citrate.

7.7.4 *Arabidopsis* aconitase activity

The specific activities of ACO1 observed within this thesis are considerably higher than those obtained for ACO2. The specific activity obtained for ACO2 is very similar to the specific activity obtained for porcine heart mitochondrial aconitase. These differences could reflect the different sub-cellular location of the isoforms and the role that they play in metabolism. The *Arabidopsis* aconitase isoform ACO1 is believed to function as part of glyoxylate cycle within the cytosol (Peyret *et al.,* 1995; Hayashi *et al.,* 1995), whereas ACO2 is believed to form part of the TCA cycle within the mitochondria (Arnaud *et al.,* 2007; Bernard *et al.,* 2009).

7.7.5 Further analysis for the purified isoforms of aconitase

With the production of pure, active ACO1 and ACO2, structural analysis in the form of x-ray crystallography, UV spectroscopy, EPR spectroscopy or extended x-ray absorption fine structure spectroscopy (EXAFS) could be conducted. However, prior to structuralisation of ACO1 and ACO2 further work could include the over-expression and purification of ACO3 from *P. pastoris*.

Additional further work could include the production of a higher concentration protein sample. This could be achieved by either using a large scale over-expression system or increasing the number of stable aconitase integrates into *P. pastoris*. Once a large volume of purified protein has been obtained, structural analysis of the protein and the iron sulphur cluster may commence.

7.8 Practical applications of this research

Whilst the research in this thesis is lab based and run under tightly controlled conditions, there are potential practical implications for commercial crops.

With regards to acetyl-CoA synthetase, the necessity of a fully functioning AAE7/ACN1 can be taken a step further. Increased levels of external acetate were shown to have a toxic effect upon *Arabidopsis* seedlings by Turner *et al.* (2005).

Other work reviewing organic acid levels in a number of agricultural soils show a variety of concentrations of acetate within the soils (Jones, 1998). Highest of these are soils containing residues of *Elytrigia* (Baziramakenga *et al.,* 1995), a grass common in certain agricultural pastures (Martineau *et al.*, 1994; Walker *et al.*, 2004) and arable lands (Njøs and Børrese, 1991). Therefore it is plausible, and perhaps merits further investigation, that in situations where pasture has been ploughed and planted with brassica crops such as oilseed rape and swede, that the level of acetate in the soil could impede seedling development. This is more likely with the increased use of minimum cultivation techniques and short fallow times before sowing (Wynn, 2006), which will increase the likelihood of seedlings being in soils with levels of high acetate and other organic acids compared to traditional ploughing techniques. Therefore in this situation, a fully functioning AAE7/ACN1 protein in commercial agricultural brassicas is important for these seedlings to successfully establish in field conditions. Increases in the level of establishment would lead to more robust seedlings that are capable of withstanding greater levels of external stress such as droughts and frosts.

The latter stress, frost, has a connection to the other aspect of this PhD thesis. Frost tolerant varieties of swede are reported to have increased levels of sucrose (Forbes and Watson, 1994). The results of GC-MS data presented in chapter 6 show significantly decreased levels of sucrose in the aconitase mutants. Therefore, plausibly in a similar manner to AAE7/ACN1, a fully functioning set of aconitase proteins may be a requirement for robust seedlings and crops.

7.9 Summary

The aim of the work presented within this thesis was to add to current knowledge about seedling establishment. The processes that occur within juvenile plants affect the speed and nature of the plant's development. Enhanced levels of plant development lead to more robust seedlings that are more able to withstand external stresses. Efficient enzymatic processes are therefore critical in the swift and successful establishment of seedlings.

Results presented in chapters 3 and 4 suggest that AAE7/ACN1 is a dimeric protein that is not associated with the plasma membrane. If AAE7/ACN1 is not associated

with the plasma membrane, then previously published results suggest that it functions as an acetyl-CoA synthetase (Turner *et al.,* 2005). Therefore, during germination and seedling establishment acetyl-CoA is produced from the breakdown of fatty acids through β-oxidation and the conversion of acetate by AAE7/ACN1. The successful germination and establishment of the *aae7/acn1* mutant (Turner *et al.,* 2005) suggests that, whilst important, it is not essential for seedling establishment in the presence of a fully functioning β-oxidation pathway.

In contrast aconitase is an essential enzyme involved in both the glyoxylate and TCA cycles during seedling establishment. The importance of functioning aconitase isoforms has been previously demonstrated by the unsuccessful identification of a double knockout mutant (Arnaud *et al.,* 2007; Moeder *et al.,* 2007). Results presented within chapters 5 and 6 of this thesis also suggest that multiple isoforms of aconitase are involved in citrate metabolism.

Previously published research also suggests that ACO1 is the most important isoform during seedling growth, functioning as part of the glyoxylate cycle (Peyret *et al.,* 1995; Arnaud *et al.,* 2007). However, results presented in chapters 5 and 6 indicate that ACO3 is the main isoform involved in seedling establishment. As ACO3 is potentially located in the mitochondria this may mean that the glyoxylate cycle is not as important during seedling establishment as first thought.

Understanding how *Arabidopsis thaliana* seedlings break down stored fatty acid reserves in order to successfully complete establishment is important to increase knowledge about processes within oilseed seedlings. Related species of plants include food crops such as broccoli, cabbage and swede, together with multipurpose crops such as oilseed rape. Oilseed rape is an important industrial scale crop that provides a source of oil for both human food and biofuels.

In a time of ever increasing global population, the greater demands placed upon the World's limited resources makes successful crops increasingly vital (McKelvey and Marshall, 2007). Furthering our understanding about the intra-cellular processes during seedling development will hopefully ultimately lead to increased seedling

survival, which in turn will lead to more constant crop establishment and thus higher yielding crops.

Appendix

8.1 IUPAC codes

Ambiguity of bases in DNA sequences is represented by a code proposed by the International Union of Pure and Applied Chemistry (IUPAC):

A- adenine
C- cytosine
G- guanine
T- thymine
R- G or A
Y- T or C
K- G or T
M- A or C
S- G or C
W- A or T
B- G, C or T
D- G, A or T
H- A,C or T
V- G, C or A
N- A,C, G or T

Appendix - chapter 3

8.2 Initial AAE7/ACN1 protein expression and purification experiments

Expression trials of the PQEHPD2 clone expressing AAE7/ACN1 (Turner *et al.*, 2005) were conducted with 50 to 1000 μM isopropyl β-thiogalactopyranoside (IPTG). Expression trials were conducted to confirm that the T5 promoter was functioning correctly and was readily induced by IPTG. Analysis showed that the PQEHPD2 clone expressing AAE7/ACN1 was inducible with as little as 50 μM IPTG (Figure 8.1). However the supplier's recommended concentration of 1000 μM IPTG (Qiagen, 2003) was still used for the duration of this work in order to ensure adequate induction and subsequent production of AAE7/ACN1.

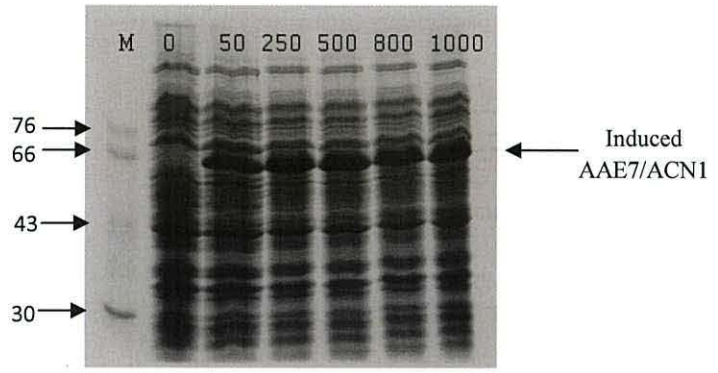


Figure 8.1: *E. coli* cells expressing AAE7/ACN1 were induced with varying concentrations of IPTG (μM). Left to right: Lane 1; molecular weight marker, 2; 0 μM, 3; 50 μM, 4; 250 μM, 5; 500 μM, 6; 800 μM, 7; 1000 μM IPTG.

AAE7/ACN1 was expressed when the PQEHPD2 clone was induced at 37°C, 28°C and 20°C. However, the majority of the protein produced at 37°C and 28°C was insoluble and remained within the *E. coli* pellet after centrifugation. Optimisation of the expression temperature was important to maximise soluble protein by minimising the insoluble protein. The formation of protein aggregates within inclusion bodies represents one of the major obstacles in the production of soluble protein (Carrio and Villaverde, 2002). AAE7/ACN1, over-expressed in *E. coli*, was produced both in a soluble and insoluble form within cells. This insoluble form, present in inclusion bodies, was extractable from the pellet with urea (Qiagen, 2003b). High concentrations of urea in the purification buffer lead to protein denaturation (Itri *et al.*, 2003), which subsequently required re-folding prior to analysis. It is therefore preferable to characterise a protein produced under native, soluble conditions.

Protein re-folding could lead to inaccurate protein characterisation. Only soluble protein produced under native conditions (in the absence of urea) was used in structural analysis of AAE7/ACN1. A reported method of increasing the concentration of soluble protein is to reduce the expression temperature (Cabantous *et al.,* 2005; Waugh, 2005). This reduces the rate of cell growth, division and over-expression leading to a greater quantity of soluble protein (Qiagen, 2003b). With this in mind, the growth and expression temperature of AAE7/ACN1 was reduced from 37°C to 20°C to increase the yield of soluble protein. Optimisation of AAE7/ACN1 expression resulted in the satisfactory yield of soluble protein and therefore eliminated the need to extract, purify and re-fold denatured protein.

Over-expressed AAE7/ACN1 contains a six residue histidine tag for purification with a HiTrap Ni-chelating column by immobilised metal affinity chromatography (IMAC). The histidine tag binds to the Ni-chelating column and protein is eluted with imidazole. The conditions required to elute pure AAE7/ACN1 were tested (Figure 8.2) with increasing concentrations of imidazole. A single purified band eluted with a molecular mass of 66 kDa with 200 mM imidazole. Purification trials were conducted using a HiTrap column to identify the optimal imidazole concentration required for elution (Figure 8.2). It was important to elute purified protein with the smallest concentration of imidazole possible in order to ensure complete removal during dialysis as imidazole has been reported to result in protein aggregation (Hefti *et al.,* 2001). A compromise between the maximum protein yield and the lowest concentration of imidazole is therefore required. Even though purified AAE7/ACN1 appears to elute at 200 mM imidazole, there could be further concentrations of protein above 200 mM but below 300 mM. For this reason AAE7/ACN1 was eluted with the recommended level of 250 mM imidazole (Qiagen, 2003), ensuring the maximum yield of purified protein was obtained.

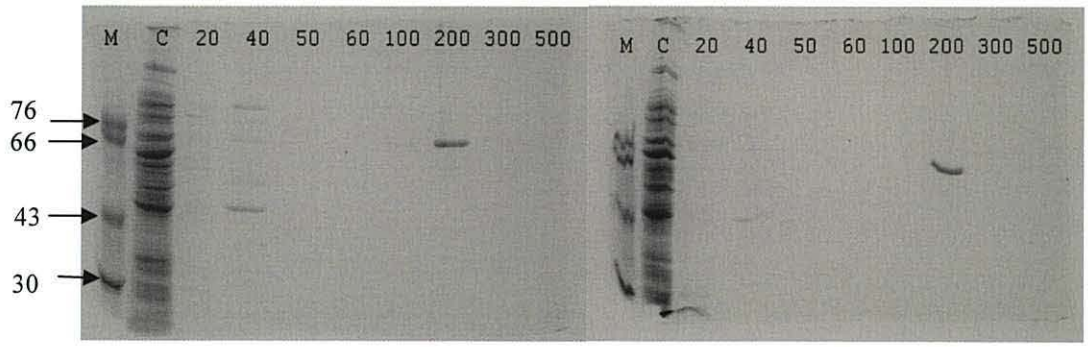


Figure 8.2: 12% Acrylamide gel electrophoresis showing the second fractions (left image) and third fractions (right image) eluted from the HiTrap column with varying concentrations of imidazole (mM). AAE7/ACN1 eluted in both images with an imidazole concentration of 200mM. M represents the molecular weight marker and C is a control sample showing the protein sample loaded onto the HiTrap column.

8.3 Isoelectric focusing (IEF) of AAE7/ACN1

A purified sample of AAE7/ACN1 was run on an IEF gel to calculate the isoelectric point (pI) of AAE7/ACN1. Analysis revealed that the pI of AAE7/ACN1 was approximately 8.6 (Figure 8.3).

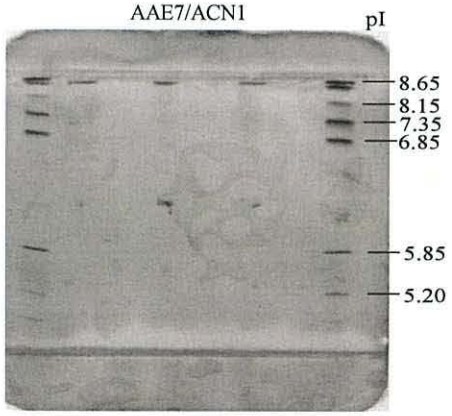


Figure 8.3: Isoelectric focusing (pI=3-9) of AAE7/ACN1 using the PhastSystem according to manufacturer's instructions (Pharmacia) demonstrates that the experimental pI of 8.5 agrees with the theoretical pI prediction conducted with ExPASY Compute pI/MW software.

8.4 AAE7/ACN1 Kinetic data

Kinetic analysis of AAE7/ACN1 was conducted with AMP, AMPE and AMPP (Figure 8.4). Analysis revealed that AMP, AMPE and AMPP are mixed inhibitors of AAE7/ACN1.

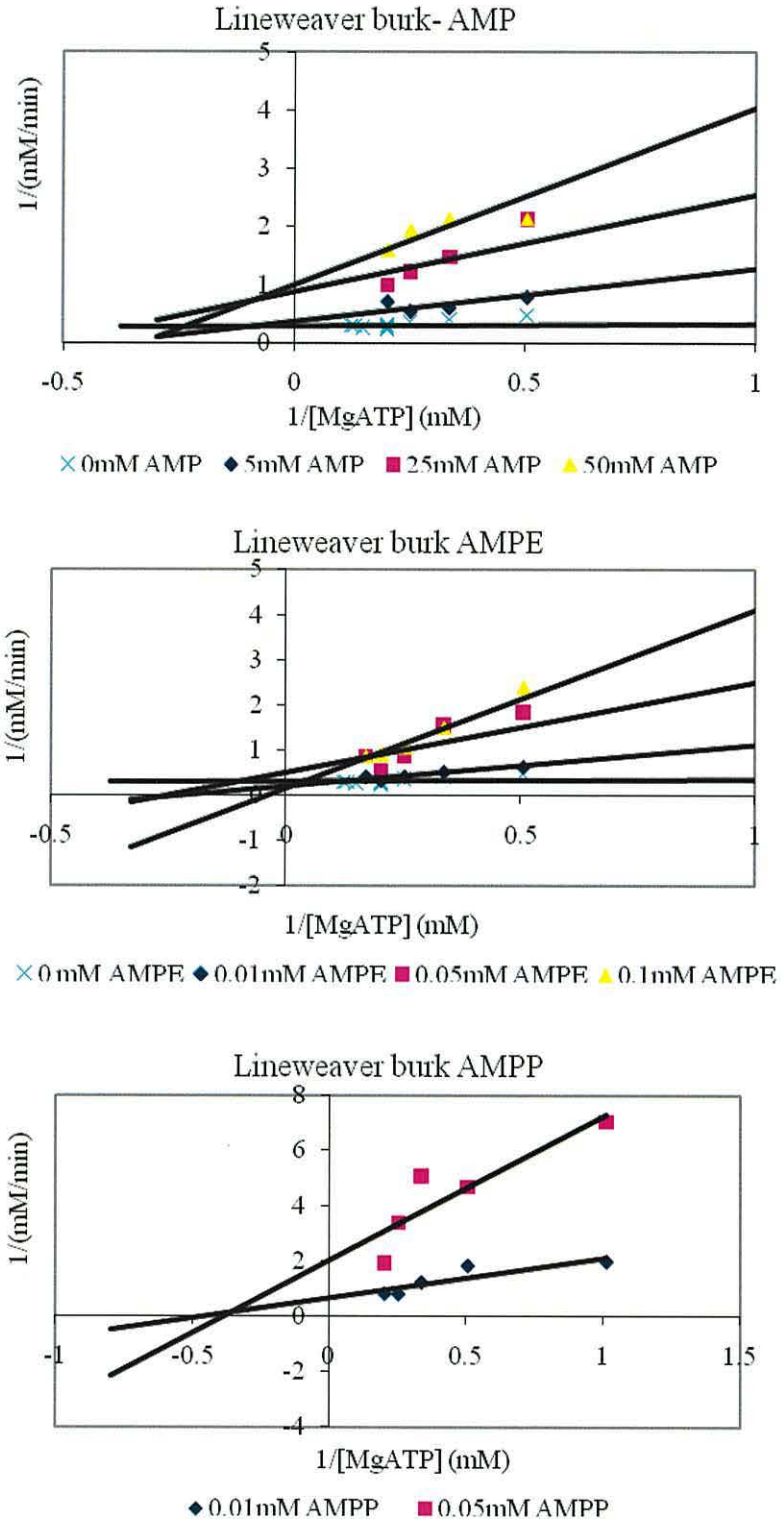


Figure 8.4: Lineweaver burk analysis of AAE7/ACN1 with AMP, AMPE and AMPP.

Appendix – chapter 4

8.5 Homology modelling of AAE7/ACN1

All residues involved in adenylate binding of ttLC-FACS were compared to AAE7/ACN1 (Table 8.1).

Table 8.1: Residues involved in adenylate binding in ttLC-FACS

Motif		ttLC-FACS	Residues that are not conserved	AAE7/ACN1
		R105		R105
	ATP/AMP signature motif (Shockey *et al.*, 2003)	M181	*	L202
		A182	*	G203
		Y183		Y205
P-loop		T184		T205
		T185	*	S206
		G186		G207
		T187		T208
		T188		T209
		G189	*	A210
		L190	*	S211
		P191		P212
		K192		K213
		F229		F248
		H230		H249
		V231	*	C250
		W234		W253
		G302		G321
		A304		A323
		Q322	*	H341
A-motif	ATP/AMP signature motif (Weimar *et al.*, 2002)	G323	*	T342
		Y324		Y343
		G325		G344
		L326		L345
		T327	*	S346
		E328		E347
		T329		T348
		T357	*	Q376
	Fatty acyl-CoA synthetase signature motif (411-435)	D418		D438
		I430		I450
L-motif		D432		D452
		R433		R453
		L434	*	S454
		K435		K455
		D436		D456
		L437	*	V457
		K439	*	I459
		E443		E463
		W444	*	N464
		E475		E495
		K527		K551

8.6 Sequence alignment of AAE7/ACN1 and ttLC-FACS for homology modelling

Sequence alignments used in the construction AAE7/ACN1 homology models with ttLC-FACS are present in Figure 8.5 and 8.6.

>1ULT
--AFPSTMMDEELNLWDFLERAAALF--GRKEVVSRLHTGEVHRT---TY
AEVYQRARRLMGGLRALGVGVGDRVATLGFNHFRHLEAYFAVPGMGAVLH
TANPRLSPKEIAYILNHAEDKVLLFDPNLLPLVE-AIR-------GELKT
VQHFVVMDEK-APE--------GYLAYEE--ALGEEADPVRVP--ERAAC
GMAYTTGTTGLPKGVVYSHRALVLHSLAASLVDGTALSEKDVVLPVVPMF
HVNAWCLPYAATLVGAKQVLPGPRLDPASLVELFDGEGVTFTAGVPTVWL
ALADY-LESTGHRLKTLRRLVVGGSAAPRSLIARFERMGVEVRQGYGLTE
TSPVVVQNFVKSHLESLSEEEKLTLKAKTGLPIP-LVRLRVADEE-GRPV
PKDGKALGEVQLKGPWITGGYYGNEEATRSALTPDGFFRTGDIAVWDEEG
YVEIKDRLKDLIKSGGEWISSVDLENALMGHPKVKEAAVVAIPHPKWQER
PLAVVVPRG--EKPTPEELNEHLLKAGFAK---WQLPDAYVFAEEIPRTS
AGKFLKRALREQYKNYYGG-----
/
--NAFPSTMMDEELNLWDFLERAAALF--GRKEVVSRLHTGEVHRT---TY
AEVYQRARRLMGGLRALGVGVGDRVATLGFNHFRHLEAYFAVPGMGAVLH
TANPRLSPKEIAYILNHAEDKVLLFDPNLLPLVE-AIR-------GELKT
VQHFVVMDEK-APE--------GYLAYEE--ALGEEADPVRVP--ERAAC
GMAYTTGTTGLPKGVVYSHRALVLHSLAASLVDGTALSEKDVVLPVVPMF
HVNAWCLPYAATLVGAKQVLPGPRLDPASLVELFDGEGVTFTAGVPTVWL
ALADY-LESTGHRLKTLRRLVVGGSAAPRSLIARFERMGVEVRQGYGLTE
TSPVVVQNFVKSHLESLSEEEKLTLKAKTGLPIP-LVRLRVADEE-GRPV
PKDGKALGEVQLKGPWITGGYYGNEEATRSALTPDGFFRTGDIAVWDEEG
YVEIKDRLKDLIKSGGEWISSVDLENALMGHPKVKEAAVVAIPHR-----
PLAVVVEKPTPEELNEHLLKAGDAYVFAEEIPRTFLKRAL--REQYKNYY
G----------
>AAE7/ACN1
MAATKWRDIDDLPKIPANYTALTPLWFLDRAAVVHPTRKSVIHGSREYTW
RQTYDRCRRLASALADRSIGPGSTVAIIAPNIPAMYEAHFGVPMCGAVLN
CVNIRLNAPTVAFLLSHSQSSVIMVDQEFFTLAEDSLRLMEEKAGSSFKR
PLLIVIGDHTCAPESLNRALSKGAIEYEDFLATGDPNYPWQPPADEWQSI
ALGYTSGTTASPKGVVLHHRGAYIMALSNPLIWG--MQDGAVYLWTLPMF
HCNGWCFPWSLAVLSGTSICLR-QVTAKEVYSMIAKYKVTHFCAAPVVLN
AIVNAPKEDTILPLPHTVHVMTAGAAPPPSVLFSMNQKGFRVAHTYGLSE
TYGPSTVCAWKPEWDSLPPETQAKLNARQGVRYTGMEQLDVIDTQTGKPV
PADGKTAGEIVFRGNMVMKGYLKNPEANKETFA-GGWFHSGDIAVKHPDN
YIEIKDRSKDVIISGGENISSVEVENVVYHHPAVLEASVVARPDERWQES
PCAFVTLKSDYEKHDQNKLAQDIMKFCREKLPAYWVPKSVVFGP-LPKTA
TGKIQKHILRTKAKEMGPVPRSRL
/
-MAATKWRDIDDLPKIPANYTALTPLWFLDRAAVVHPTRKSVIHGSREYTW
RQTYDRCRRLASALADRSIGPGSTVAIIAPNIPAMYEAHFGVPMCGAVLN
CVNIRLNAPTVAFLLSHSQSSVIMVDQEFFTLAEDSLRLMEEKAGSSFKR
PLLIVIGDHTCAPESLNRALSKGAIEYEDFLATGDPNYPWQPPADEWQSI
ALGYTSGTTASPKGVVLHHRGAYIMALSNPLIWG--MQDGAVYLWTLPMF
HCNGWCFPWSLAVLSGTSICLR-QVTAKEVYSMIAKYKVTHFCAAPVVLN
AIVNAPKEDTILPLPHTVHVMTAGAAPPPSVLFSMNQKGFRVAHTYGLSE
TYGPSTVCAWKPEWDSLPPETQAKLNARQGVRYTGMEQLDVIDTQTGKPV
PADGKTAGEIVFRGNMVMKGYLKNPEANKETFA-GGWFHSGDIAVKHPDN
YIEIKDRSKDVIISGGENISSVEVENVVYHHPAVLEASVVARPDERWQES
PCAFVTLKSDYEKHDQNKLAQDIMKFCREKLPAYWVPKSVVFGPLPKTAT
GKIQKHILRTKAKEMGPVPRSRL

Figure 8.5: Amino acid sequence alignment of AAE7/ACN1 and ttLC-FACS (1ULT).

>1V26
--AFPSTMMDEELNLWDFLERAAALF--GRKEVVSRLHTGEVHRT---TY
AEVYQRARRLMGGLRALGVGVGDRVATLGFNHFRHLEAYFAVPGMGAVLH
TANPRLSPKEIAYILNHAEDKVLLFDPNLLPLVE-AIR-------GELKT
VQHFVVMDEK-APE--------GYLAYEE--ALGEEADPVRVP--ERAAC
GMAYTTGTTGLPKGVVYSHRALVLHSLAASLVDGTALSEKDVVLPVVPMF
HVNAWCLPYAATLVGAKQVLPGPRLDPASLVELFDGEGVTFTAGVPTVWL
ALADY-LESTGHRLKTLRRLVVGGSAAPRSLIARFERMGVEVRQGYGLTE
TSPVVVQNFVKSHLESLSEEEKLTLKAKTGLPIP-LVRLRVADEE-GRPV
PKDGKALGEVQLKGPWITGGYYGNEEATRSALTPDGFFRTGDIAVWDEEG
YVEIKDRLKDLIKSGGEWISSVDLEN----------AAVVAIPHPKWQER
PLAVVGF-AKWQLPDAY-----LKRALREQYKNYYG----------

/
--AFPSTMMDEELNLWDFLERAAALF--GRKEVVSRLHTGEVHRT---TY
AEVYQRARRLMGGLRALGVGVGDRVATLGFNHFRHLEAYFAVPGMGAVLH
TANPRLSPKEIAYILNHAEDKVLLFDPNLLPLVE-AIR-------GELKT
VQHFVVMDEK-APE--------GYLAYEE--ALGEEADPVRVP--ERAAC
GMAYTTGTTGLPKGVVYSHRALVLHSLAASLVDGTALSEKDVVLPVVPMF
HVNAWCLPYAATLVGAKQVLPGPRLDPASLVELFDGEGVTFTAGVPTVWL
ALADY-LESTGHRLKTLRRLVVGGSAAPRSLIARFERMGVEVRQGYGLTE
TSPVVVQNFVKSHLESLSEEEKLTLKAKTGLPIP-LVRLRVADEE-GRPV
PKDGKALGEVQLKGPWITGGYYGNEEATRSALTPDGFFRTGDIAVWDEEG
YVEIKDRLKDLIKSGGEWISSVDLENA---------AAVVAIPHPKWQER
PLA-VTPEELNEHLLKAGFAK--WQLPDAYVFAEEIPRTRALREQYKNYY
GGA----------
>AAE7/ACN1
MAATKWRDIDDLPKIPANYTALTPLWFLDRAAVVHPTRKSVIHGSREYTW
RQTYDRCRRLASALADRSIGPGSTVAIIAPNIPAMYEAHFGVPMCGAVLN
CVNIRLNAPTVAFLLSHSQSSVIMVDQEFFTLAEDSLRLMEEKAGSSFKR
PLLIVIGDHTCAPESLNRALSKGAIEYEDFLATGDPNYPWQPPADEWQSI
ALGYTSGTTASPKGVVLHHRGAYIMALSNPLIWG--MQDGAVYLWTLPMF
HCNGWCFPWSLAVLSGTSICLR-QVTAKEVYSMIAKYKVTHFCAAPVVLN
AIVNAPKEDTILPLPHTVHVMTAGAAPPPSVLFSMNQKGFRVAHTYGLSE
TYGPSTVCAWKPEWDSLPPETQAKLNARQGVRYTGMEQLDVIDTQTGKPV
PADGKTAGEIVFRGNMVMKGYLKNPEANKETFA-GGWFHSGDIAVKHPDN
YIEIKDRSKDVIISGGENISSVEVENVVYHHPAVLEASVVARPDERWQES
PCAFVTLKSDYEKHDQNKLAQDIMKFCREKLPAYWVPKSVVFGPLPKTAT
GKIQKHILRTKAKEMGPVPRSRL
/
MAATKWRDIDDLPKIPANYTALTPLWFLDRAAVVHPTRKSVIHGSREYTW
RQTYDRCRRLASALADRSIGPGSTVAIIAPNIPAMYEAHFGVPMCGAVLN
CVNIRLNAPTVAFLLSHSQSSVIMVDQEFFTLAEDSLRLMEEKAGSSFKR
PLLIVIGDHTCAPESLNRALSKGAIEYEDFLATGDPNYPWQPPADEWQSI
ALGYTSGTTASPKGVVLHHRGAYIMALSNPLIWG--MQDGAVYLWTLPMF
HCNGWCFPWSLAVLSGTSICLR-QVTAKEVYSMIAKYKVTHFCAAPVVLN
AIVNAPKEDTILPLPHTVHVMTAGAAPPPSVLFSMNQKGFRVAHTYGLSE
TYGPSTVCAWKPEWDSLPPETQAKLNARQGVRYTGMEQLDVIDTQTGKPV
PADGKTAGEIVFRGNMVMKGYLKNPEANKETFA-GGWFHSGDIAVKHPDN
YIEIKDRSKDVIISGGENISSVEVENVVYHHPAVLEASVVARPDERWQES
PCAFVTLKSDYEKHDQNKLAQDIMKFCREKLPAYWVPKSVVFGPLPKTAT
GKIQKHILRTKAKEMGPVPRSRL

Figure 8.6: Amino acid sequence alignment of AAE7/ACN1 with ttLC-FACS (1V26).

8.7 Sequence alignment of mtACS, AAE7/ACN1 and ttLC-FACS

The amino acid sequence of AAE7/ACN1 was aligned to mtACS and ttLC-FACS to compare conserved residues (Figure 8.7). Sequence alignment analysis showed that residues are conserved between all three proteins. This is not surprising as all three proteins catalyse the same reaction.

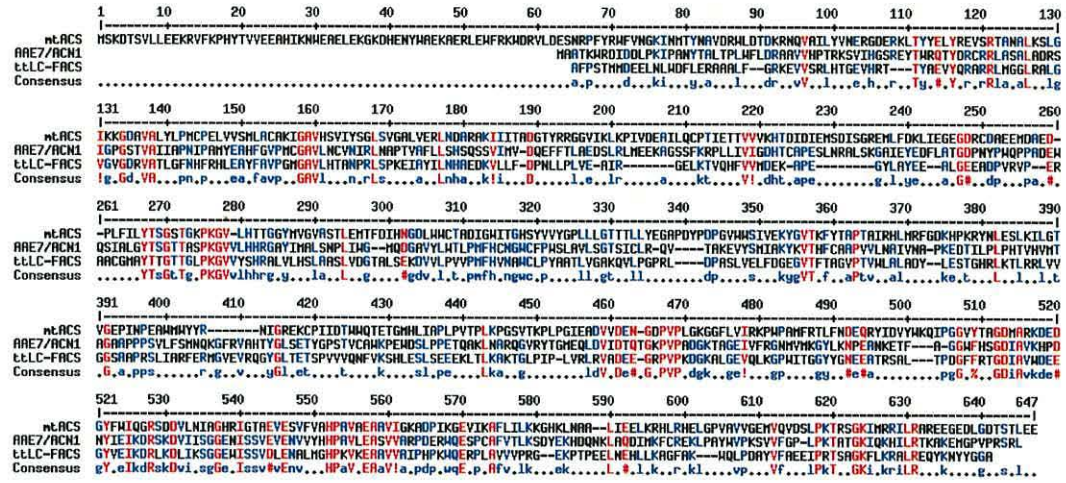


Figure 8.7: Sequence alignment of mtACS, AAE7/ACN1 and ttLC-FACS created by Multalin. Sequences conserved between two sequences are shown in blue. Sequences conserved between all three sequences are shown in red. All other residues are present in black.

Appendix – chapter 5

8.8 Cloning ACO1, ACO2 and ACO3 into *P. pastoris*

Prior to cloning into *P. pastoris* the aconitase genes were first sub-cloned into *E. coli*. The expression vector, pPiczC is an *E. coli / P. pastoris* shuttle vector with an origin of replication for *E. coli*. In addition the expression vector has a multi cloning site for the insertion of foreign genes. The expression vector was constructed in *E. coli*, linnearised and integrated into the *P. pastoris* genome to maximise the stability of expression strains (Cereghino and Cregg, 2000).

During transformation the linearised vector was inserted into the host genome at the AOX1 loci by a single crossover event (Figure 8.8). This can occur at either of the two AOX1 regions on the construct, the AOX1 promoter or the AOX1 transcription termination region (TT).

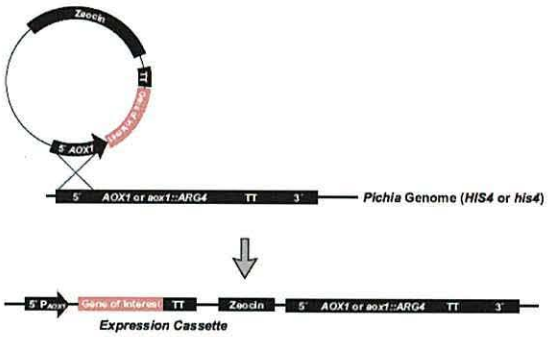


Figure 8.8: Integration of the linearised *P. pastoris* vector, PpiczC into the genomic host (Invitrogen).

Antibiotic resistance was used to select for positive transformants. Antibiotic resistance indicated recombination of the plasmid. PCR was used to identify colonies that contained a plasmid with an aconitase gene.

PCR analysis was conducted using primers specific for the alcohol oxidase (Table 8.2) gene. This resulted in the amplification of two sections of DNA and the presence of two bands after gel electrophoresis. Positive transformants will have one band corresponding to an aconitase gene, plus 324 bp. The second band will correspond to the AOX1 gene, which is approximately 2.2 kb. Transformants that have incorporated a plasmid but not an aconitase gene will produce a small 324 bp band where the insert should be. A second band corresponding to the AOX1 gene will also

be produced after PCR and gel electrophoresis. Positive transformants for all three aconitase isoforms were identified by PCR with AOX1 primers.

Table 8.2: Specific alcohol oxidase (AOX1) primers implemented in the positive identification of *P. pastoris* transformants.

	Primer Sequence (5'-3')
5' AOX1	GACTGGTTCCAATTGACAAGC
3' AOX1	GCAAATGGCATTCTGACATCC

Appendix – chapter 6

8.9 GC-MS analysis of *aco1*, *aco2* and *aco3*

Statistical analysis was conducted on the GC-MS data to identify all statistically significant differences. Analysis was conducted between the entire seedling groups (Col-0, *aco1*, *aco2* and *aco3*) for each of the metabolites compared in chapter 6. Prior to statistical analysis the data's homogeneity of variance was tested using a Levene statistic (Table 8.3).

Table 8.3: Levene Statistic. Testing of the data's homogeneity of variance using the Levene Statistic.

Metabolite	*P*
Isocitric acid	<0.001
Citric acid	<0.001
Sucrose	<0.001
Myristic acid	0.433
Succinic acid	0.225
Fumaric acid	0.006
Malic acid	0.096
Pyruvic acid	0.353

The homogeneity of variance is important in determining the type of statistical test that can be implemented to analyze the GC-MS data. The Levene statistic compares the data's variance and returns a probability (p) that the variances are significantly different ($p < 0.05$). If the variances are equal, and therefore not significantly different then a one way analysis of variance (ANOVA) can be used to determine if there is a significant difference between any of the data sets (Table 8.4). Whilst an ANOVA will identity a significant difference within the data set, it will not locate the difference. Post Hoc comparisons were conducted to determine which seedling groups were significantly different to one another.

Table 8.4: A one-way analysis of variance (ANOVA) of GC-MS data.

Metabolite	*P*
Myristic acid	0.167
Succinic acid	<0.001
Malic acid	<0.001
Pyruvic acid	<0.001

If unequal variances were identified by the Levene statistic ($p < 0.05$) then a non-parametric test must be used to identify any statistically significant differences within

the GC-MS data. The non-parametric test, Kruskal-Wallis was used to compare the GC-MS data. When $p < 0.05$ a statistical significant difference has been identified between the seedling groups for a specific metabolite (Table 8.6). However, Kruskal-Wallis does not identify which seedlings groups are significantly different. In order to determine which groups are different for a particular metabolite Dunnett's C tests were conducted.

Table 8.5: Krustal-Wallis analysis of GC-MS data.

	P
Citric acid	<0.001
Isocitric acid	<0.001
Sucrose	<0.001
Fumaric acid	<0.001

8.10 Expression levels of the *Arabidopsis* aconitase isoforms

Arabidopsis eFP Browser (www.bar.utoronto.ca/efp; Winter *et al.,* 2007) analysis. Results show the level of expression of aconitase at varying developmental stages in wild type plants (Figure 8.9, 8.10 and 8.11).

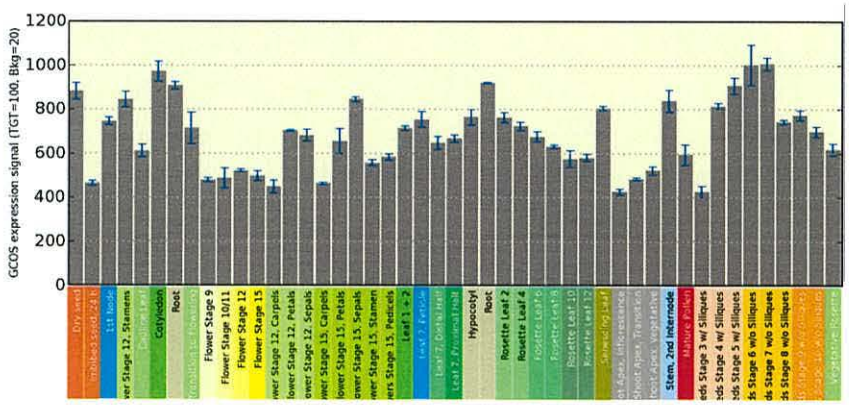


Figure 8.9: ACO1. Screenshot of the *Arabidopsis* eFP Browser results.

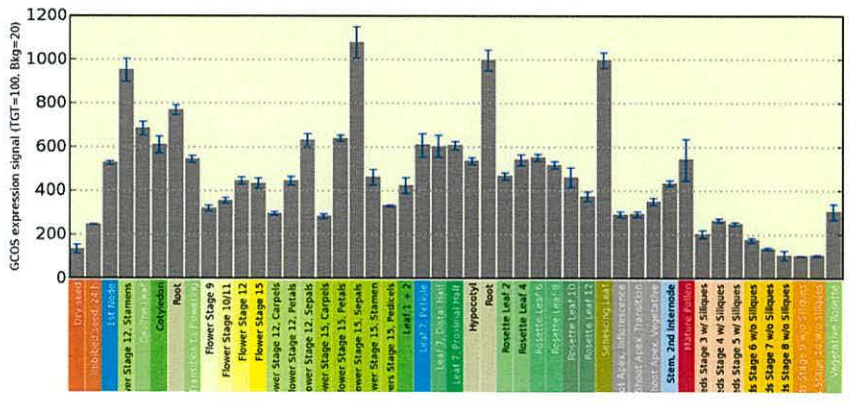


Figure 8.10: ACO2. Screenshot of the *Arabidopsis* eFP Browser results.

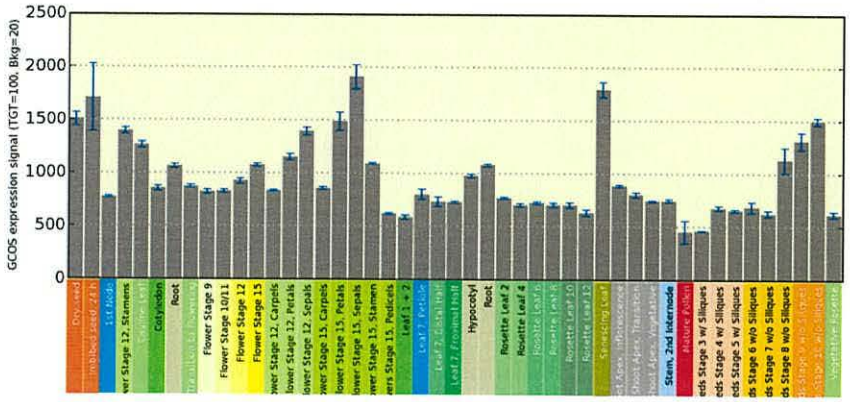


Figure 8.11: ACO3. Screenshot of the *Arabidopsis* eFP Browser results.

References

AHMED, S., SARAF, T. and GOYAL, A. 2009. Homology modeling of family 39 glycoside hydrolase from *Clostridium thermocellum*. *Current Trends in Biotechnology and Pharmacy,* 3, 210-218.

ALEN, C. and SONENSHEIN, A. L. 1999. *Bacillus subtilis* aconitase is an RNA-binding protein. *Proceedings of the National academy of Sciences of the United States of America,* 96, 10412-10417.

ALHENDAWI, R. A., ROMHELD, V., KIRKBY, E. A. and MARSCHNER, H. 1997. Influence of increasing bicarbonate concentrations on plant growth, organic acid accumulation in roots and iron uptake by barley, sorghum, and maize. *Journal of Plant Nutrition,* 20, 1731-1753.

ALI, M. H. and IMPERIALI, B. 2005. Protein oligomerization: How and why. *Bioorganic and Medicinal Chemistry,* 13, 5013-5020.

ALONSO, J. M., STEPANOVA, A. N., LEISSE, T. J., KIM, C. J., CHEN, H. M., SHINN, P., STEVENSON, D. K., ZIMMERMAN, J., BARAJAS, P., CHEUK, R., GADRINAB, C., HELLER, C., JESKE, A., KOESEMA, E., MEYERS, C. C., PARKER, H., PREDNIS, L., ANSARI, Y., CHOY, N., DEEN, H., GERALT, M., HAZARI, N., HOM, E., KARNES, M., MULHOLLAND, C., NDUBAKU, R., SCHMIDT, I., GUZMAN, P., AGUILAR-HENONIN, L., SCHMID, M., WEIGEL, D., CARTER, D. E., MARCHAND, T., RISSEEUW, E., BROGDEN, D., ZEKO, A., CROSBY, W. L., BERRY, C. C. and ECKER, J. R. 2003. Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science,* 301, 653-657.

APEL, K. and HIRT, H. 2004. Reactive oxygen species: metabolism, oxidative stress, and signal transduction. *Annual Review of Plant Biology,* 55, 373-399.

ARNAUD, N., RAVET, K., BORLOTTI, A., TOURAINE, B., BOUCHEREZ, J., FIZAMES, C., BRIAT, J. F., CELLIER, F. and GAYMARD, F. 2007. The iron-responsive element (IRE)/iron-regulatory protein 1 (IRP1)-cytosolic aconitase iron-regulatory switch does not operate in plants. *Biochemical Journal,* 405, 523-531.

ASHERIE, N. 2004. Protein crystallization and phase diagrams. *Methods,* 34, 266-272.

BABBITT, P. C., KENYON, G. L., MARTIN, B. M., CHAREST, H., SLYVESTRE, M., SCHOLTEN, J. D., CHANG, K. H., LIANG, P. H. and DUNAWAY-MARIANO, D. 1992. Ancestry of the 4-chlorobenzoate dehalogenase: analysis of amino acid sequence identities among families of acyl:adenyl ligases, enoyl-CoA hydratases/isomerases, and acyl-CoA thioesterases. *Biochemistry,* 31, 5594-5604.

BADUGU, R., GARCIA, M., BONDADA, V., JOSHI, A. and GEDDES, J. W. 2008. N-terminus of calpain 1 is a mitochondrial targeting sequence. *The Journal of Biological Chemistry,* 283, 3409-3417.

BAKER, A., GRAHAM, I. A., HOLDSWORTH, M., SMITH, S. M. and THEODOULOU, L. 2006. Chewing the fat: β-oxidation in signaling and development. *Trends in Plant Science,* 11, 124-132.

BALDWIN, R. L. and ROSE, G. D. 1999. Is protein folding hierarchic? I. Local structure and peptide folding. *Trends in Biochemical Sciences,* 24, 26-33.

BANERJEE, S., NANDYALA, A. K., RAVIPRASAS, P., AHMED, N. and HASNAIN, S. E. 2007. Iron dependent RNA binding activity of *Mycobacterium tuberculosis* aconitase. *Journal of Bacteriology,* 189, 4046-4052.

BANNAI, H., TAMADA, Y., MARUYAMA, O., NAKAI, K. and MIYANO, S. 2002. Extensive feature detection of N-terminal protein sorting signals. *Bioinformatics,* 18, 298-305.

BAO, X., FOCKE, M., POLLARD, M. and OHLROGGE, J. 2000. Understanding *in vivo* carbon precursor for fatty acid synthesis in leaf tissue. *The Plant Journal,* 22, 39-50.

BAR-TANA, J. and ROSE, G. 1968. Studies on medium-chain fatty acyl-coenzyme A synthetase. *Biochemical Journal,* 109, 283-292.

BAR-TANA, J., ROSE, G. and SHAPIRO, B. 1971. The purification and properties of microsomal palmitoyl-coenzyme A synthetase. *Biochemical Journal,* 122, 353-362.

BAR-TANA, J. and ROSE, G. 1973. Rat liver microsomal palmitoyl-coenzyme A synthetase. *Biochemical Journal,* 131, 443-449.

BAR-TANA, J., ROSE, G. and SHAPIRO, B. 1973. Palmitoyl-coenzyme A synthetase. Isolation of an enzyme bound intermediate. *Biochemical Journal,* 135, 411-416.

BASILION, J. P., KENNEDY, M. C., BEINERT, H., MASSINOPLE, C. M., KLAUSNER, R. D. and ROUAULT, T. A. 1994a. Over-expression of iron responsive element binding protein and its analytical characterization as the RNA binding form, devoid of an iron sulfur cluster. *Archives of Biochemistry and Biophysics,* 311, 517-522.

BASILION, J. P., ROUAULT, T. A., MASSINOPLE, C. M., KLAUSER, R. D. and BURGESS, W. H. 1994b. The iron-responsive element binding protein: localization of the RNA-binding site to the aconitase active site cleft. *Proceedings of the National Academy of Sciences of the United States of America,* 91, 574-578.

BAUMANN, P. 1968. Isolation of *Acinetobacter* from soil and water. *Journal of Bacteriology,* 96, 39-42.

BAZIRAMAKENGA, R., SIMARD, R. R. and LEROUX, G. D. 1995. Determination of organic acids in soil extracts by ion chromatography. *Soil Biology and Biochemistry,* 27, 349-356.

BECKWITH, J. R. 1967. Regulation of the lac operon. Recent studies on the regulation of lactose metabolism in *Escherichia coli* support the operon model. *Science,* 156, 597-604.

BEEVERS, H. 1961. Metabolic production of sucrose from fat. *Nature,* 191, 433-436.

BEHAL, R. H. and OLIVER, D. J. 1997. Biochemical and molecular characterization of Fumarase from plants: purification and characterization of the enzyme- cloning, sequencing, and expression of the gene. *Archives of Biochemistry and Biophysics,* 348, 65-74.

BEHAL, R. H., LIN, M., BACK, S. and OLIVER, D. J. 2002. Role of acetyl-coenzyme A synthetase in leaves of *Arabidopsis thaliana. Archives of Biochemistry and Biophysics,* 402, 259-267.

BEINERT, H. and KENNEDY, M. C. 1993. Aconitase, a two faced protein: enzyme and iron regulatory factor. *The Journal of the Federation of American Societies for Experimental Biology,* 7, 1442-1449.

BENNETT-LOVSEY, R. M., HERBERT, A. D., STERNBERG, M. J. E. and KELLEY, L. A. 2008. Exploring the extremes of sequence/structure space with ensemble fold recognition in the program Phyre. *Proteins,* 3, 611-625.

BERGFORS, T. M. 1999. Protein Crystallisation. Techniques, strategies, and tips. A laboratory manual. Bergfors, T. M. (ed), IUL Biotechnology Series, International University Line.

BERGMEYER, H. U., GAWEHN, K. and GRASSL, M. 1974. Methods of enzymatic analysis. Bergmeyer, H. U. (ed), volume I, second edition, Academic Press, Inc., New York, NY.

BERNARD, D. G., CHENG, Y., ZHAO, Y. and BALK, J. 2009. An allelic mutant series of *ATM3* reveals its key role in the biogenesis of cytosolic iron sulfur proteins in *Arabidopsis*. *Plant Physiology,* 151, 590-602.

BERTINI, I., CAVALLARO, G., LUCHINAT, C. and POLI, I. 2003. A use of Ramachandran potentials in protein solution structure determinations. *Journal of Biomolecular NMR,* 26, 355-366.

BISSWANGER, H. 2008. Enzyme kinetics: principles and methods. Second edition, Published by Wiley-VCH, Weinheim, Germany.

BLACK, P. N. 1991. Primary sequence of the *Escherichia coli fad*L gene encoding an outer membrane protein required for long chain fatty acid transport. *Journal of Bacteriology,* 173, 435-442.

BLACK, P. N., DIRUSSO, C. C., METZGER, A. K. and HEIMERT, T. L. 1992. Cloning, sequencing, and expression of the *fadD* gene of *Escherichia coli* encoding acyl-coenzyme A synthetase. *The Journal of Biological Chemistry,* 267, 25513-25520.

BLACK, P. N., ZHANG, Q., WEIMAR, J. D. and DIRUSSO, C. C. 1997. Mutational analysis of a fatty acyl-coenzyme A synthetase signature motif identifies seven amino acid residues that modulate fatty acid substrate specificity. *The Journal of Biological Chemistry,* 272, 4896-4903.

BLACK, P. N., DIRUSSO, C. C., SHERIN, D., MACCOLL, R., KNUDSEN, J. and WEIMAR, J. D. 2000. Affinity labeling fatty acyl-CoA synthetase with 9-*p*-azidophenoxy nonanoic acid and the identification of the fatty acid binding site. *The Journal of Biological Chemistry,* 275, 38547-38553.

BOSCOLO, P. R., MENOSSI, M. and JORGE, R. A. 2003. Aluminum induced oxidative stress in maize. *Phytochemistry,* 62, 181-189.

BOURGAUD, F., GRAVOT, A., MILESI, S. and GONTIER, E. 2001. Production of plant secondary metabolites: a historical perspective. *Plant Science, 161,* 839-851.

BRADFORD, M. M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry,* 72, 248-54.

BRASEN, C., URBANKE, C. and SCHONHEIT, P. 2005. A novel octameric AMP-forming acetyl-CoA synthetase from the hyperthermophilic crenarchaeon *Pyrobaculum aerophilum*. *FEBS Letters,* 579, 477-482.

BREWER, L. R., FRIDDLE, R., NOY, A., BALDWIN, E. and MARTIN, S. S. 2003. Packaging of single DNA molecules by the yeast mitochondrial protein Abf2p. *Biophysical Journal,* 84, 2519-2524.

BRIAT, J. F., CURIE, C. and GAYMARD, F. 2007. Iron utilization and metabolism in plants. *Current Opinion in Plant Biology, 10,* 276-282.

BROUQUISSE, R., GAILLARD, J. and DOUCE, R. 1986. Electron paramagnetic resonance characterization of membrane bound iron sulfur clusters and aconitase in plant mitochondria. *Plant Physiology, 81,* 247-252.

BROUQUISSE, R., NISHIMURA, M., GAILLARD, J. and DOUCE, R. 1987. Characterization of a cytosolic aconitase in higher plant cells. *Plant Physiology,* 84, 1402-1407.

BROWN, J. C., HOLMES, R. S. and SPECHT, A. W. 1955. Iron, the limiting element in a chlorosis: Part II. Copper phosphorus induced chlorosis dependent upon plant species and varieties. *Plant Physiology,* 30, 457-462.

BROWN, J. C. 1956. Iron chlorosis. *Annual Review in Plant Physiology,* 7, 171-190.

BROWN, S. 2000. Does succinate oxidation yield $FADH_2$ or ubiquinol? *Biochemical Education,* 28, 52-54.

BUCHER, M. H., EVDOKIMOV, A. G. and WAUGH, D. S. 2002. Differential effects of short affinity tags on the crystallization of *Pyrococcus furiosus* maldodextrin binding protein. *Acta Crystallographica,* D58, 392-397.

BUTT, J., KIN, H. Y., BASILION, J. P., COHEN, S., IWAI, K., PHILPOTT, S. A., KLAUSNER, R. D. and ROUAULT, T. A. 1996. Differences in the RNA binding sites of iron regulatory proteins and potential target diversity. *Proceedings of the National Academy of Sciences,* 93, 4345-4349.

BUXBAUM, E. 2007. Fundamentals of protein structure and function. Springer science and business media, Inc, New York.

CABISCOL, E., PIULATS, E., ECHAVE, P., HERRERO, E. and ROS, J. 2000. Oxidative stress promotes specific protein damage in *Saccharomyces cervisiae*. *The Journal of Biological Chemistry,* 275, 27393-27398.

CADENAS, E. 1989. Biochemistry of oxygen toxicity. *Annual Review of Biochemistry,* 58, 79-110.

CAMPBELL, N. A. an d REECE, J. B. 2002. Biology. Sixth edition, Benjamin Cummings, San Francisco, CA.

CAMPAGNARI, F. and WEBSTER, L. T. 1963. Purification and properties of acetyl- coenzyme A synthetase from bovine heart mitochondria. *The Journal of Biological Chemistry,* 238, 1628-1633.

CANVIN, D. T. and BEEVERS, H. 1961. Sucrose synthesis from acetate in the germinating castor bean: kinetics and pathway. *The Journal of Biological Chemistry,* 236, 988-995.

CARO, A and PUNTARULO, S. 1996. Effect of *in vivo* iron supplementation on oxygen radical production by soybean roots. *Biochimica et Biophysica Acta,* 1291, 245-251.

CARRARI, F., NUNES-NESI, A., GIBON, Y., LYTOVCHENKO, A., LOUREIRO, M. E. and FERNIE, A. R. 2003. Reduced expression of aconitase results in an enhanced rate of photosynthesis and marked shifts in carbon partitioning in illuminated leaves of wild species tomato. *Plant Physiology,* 133, 1322-1335.

CARRIO, M. M. and VILLAVERDE, A. 2002. Construction and destruction of bacterial inclusion bodies. *Journal of Biotechnology,* 96, 3-12.

CATONI, E., SCHWAB, R., HILPERT, M., DESIMONE, M., SCHWACKE, R., FLÜGGE, U. I., SCHUMACHER, K. and FROMMER, W. B. 2003. Identification of an *Arabidopsis* mitochondrial succinate-fumarate translocator. *FEBS Letters,* 534, 87-92.

CEREGHINO, J. L. and CREGG, J. M. 2000. Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. *FEMS Microbiology Reviews,* 24, 45-66.

CHAYEN, N. E. and SARIDAKIS, E. 2008. Protein crystallization: from purified protein to diffraction-quality crystal. *Nature Methods,* 5, 147-153.

CHEN, X. J., WANG, X., KAUFMAN, B. A. and BUTOW, R. A. 2005. Aconitase couples metabolic regulation to mitochondrial DNA maintenance. *Science,* 307, 714-717.

CHEN, X. J., WANG, X. and BUTOW, R. A. 2007. Yeast aconitase binds and provides metabolically coupled protection to mitochondrial DNA. *Proceedings of the National Academy of Sciences,* 104, 13738-13743.

CHIA, D. W., YODER, T. J., REITER, W. D. and GIBSON, S. I. 2000. Fumaric acid: an overlooked form of fixed carbon in *Arabidopsis* and other plant species. *Planta,* 211, 743-751.

CHOI, E. J. and MAYO, S. L. 2006. Generation and analysis of proline mutants in protein G. *Protein Engineering, Design and Selection,* 19, 285-289.

CLAROS, M. G. and VINCES, P. 1996. Computational method to predict mitochondrially imported proteins and their targeting sequences. *European Journal of Biochemistry,* 241, 779-786.

CONNOLLY, E. L. and GUERINOT, M. L. 2002. Iron stress in plants. *Genome Biology,* 3, 1024.1-1024.4.

CONTAMINE, V. and PICARD, M. 2000. Maintenance and integrity of the mitochondrial genome: a plethora of nuclear genes in the budding yeast. *Microbiology and Molecular Biology Reviews,* 64, 281-315.

CONTI, E., FRANKS, N. P. and BRICK, P. 1996. Crystal structure of firefly luciferase throws light on a superfamily of adenylate forming enzymes. *Structure,* 4, 287-298.

CONTI, E., STACHELHAUS, T., MARAHIEL, M. A. and BRICK, P. 1997. Structural basis for the activation of phenylalanine in the non-ribosomal biosynthesis of gramicidin S. *The EMBO Journal,* 16, 4174-4183.

COOPER, T. G. and BEEVERS, H. 1969. Mitochondria and glyoxysomes from castor bean endosperm. *The Journal of Biological Chemistry,* 244, 3507-3513.

CORNAH, J. E. and SMITH, S. M. 2002. Synthesis and function of glyoxylate cycle enzymes. In Baker, A. and Graham, I. A. (eds), Plant peroxisomes: biochemistry, cell biology, and biotechnological applications, pp57-101. Kluwer academic publishers, Netherlands.

CORNAH, J. E., GERMAIN, V., WARD, J. L., BEALE, M. H. and SMITH, S. M. 2004. Lipid utilization, gluconeogenesis, and seedling growth in *Arabidopsis* mutants lacking the glyoxylate cycle enzyme malate synthase. *Journal of Biological Chemistry,* 279, 42916-42923.

CORPET, F. 1988. Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Research,* 16, 10881-90.

COURTOIS-VERNIQUET, F. and DOUCE, R. 1993. Lack of aconitase in glyoxysomes and peroxisomes. *Journal of Biochemistry,* 294, 103-107.

COSTENARO, L., GROSSMANN, J. G., EBEL, C. and MAXWELL, A. 2005. Small angle x-ray scattering reveals the solution structure of the full length DNA gyrase A subunit. *Structure,* 13, 287-296.

COTS, J. and WIDMER, F. 1999. Germination, senescence and pathogenic attack in soybean (Glycine max): identification of the cytosolic aconitase participating in the glyoxylate cycle. *Plant Science,* 149, 95-104.

CROWLEY, P. B. and GOLOVIN, A. 2005. Cation-π interactions in protein-protein interfaces. *Proteins,* 59, 231-239.

CURIE, C., ALONSO, J. M., LE JEAN, M., ECKER, J. R. and BRIAT, J. F. 2000. Involvement of NRAMP1 from *Arabidopsis thaliana* in iron transport. *Journal of Biochemistry,* 347, 749-755.

CURIE, C., PANAVIENE, Z., LOULERGUE, C., DELLAPORTA, S. L., BRIAT, J. F. and WALKER, E. L. 2001. Maize yellow stripe 1 encodes a membrane protein directly involved in Fe(III) uptake. *Nature,* 409, 346-349.

CURIE, C. and BRIAT, J. F. Iron transport and signaling in plants. 2003. *Annual Review of Plant Biology,* 54, 183-206.

CURIE, C., CASSIN, G., COUCH, D., DIVOL, F., HIGUCHI, K., LE JEAN, M., MISSON, J., SCHIKORA, A., CZERNIC, P. and MARI, S. 2009. Metal movement within the plant: contribution of nicotianamine and yellow stripe 1 like transporters. *Annals of Botany,* 103, 1-11.

CURTIS, M. D. and GROSSNIKLAUS, U. 2003. A gateway cloning vector set for high-throughput functional analysis of genes in planta. *Plant Physiology,* 133, 462-469.

DAWSON, R. M. C., ELLIOTT, D. C., ELLIOTT, W. H. and JONES, K. M. 1987. Data for biochemical research. Third edition, Oxford science publications, OUP, Oxford.

DE AZEVEDO-SOUZA, C., BARBAZUK, B., RALPH, S. G., BOHLMANN, J., HAMBERGER, B. and DOUGLAS, C. J. 2008. Genome-wide analysis of a land plant specific acyl-coenzyme A synthetase (ACS) gene family in *Arabidopsis*, poplar, rice and *Physcomitrella. New Phytologist,* 179, 987-1003.

DE BELLIS, L. and NISHIMURA, M. 1991. Development of enzymes of the glyoxylate cycle during senescence of pumpkin cotyledons. *Plant Cell Physiology,* 4, 555-561.

DE BELLIS, L., TSUGEKI, R., ALPI, A. and NISHIMURA, M. 1993. Purification and characterization of aconitase isoforms from etiolated pumpkin cotyledons. *Physiologia plantarum,* 88, 485-492.

DE BELLIS, L., HAYASHI, M., BIAGI, P. P., NISHIMURA, I. H., ALPI, A. and NISHIMURA, M. 1994. Immunological analysis of aconitase in pumpkin cotyledons: the absence of aconitase in glyoxysomes. *Physiologia Plantarum,* 90, 757-762.

DE BELLIS, L., HAYASHI, M., NISHIMURA, M. and ALPI, A. 1995. Subcellular and developmental changes in distribution of aconitase isoforms in pumpkin cotyledons. *Planta,* 195, 464-468.

DEKKERS, B. J. W., SCHUURMANS, J. A. M. J. and SMEEKENS, S. C. M. 2008. Interaction between sugar and abscisic acid signaling during early seedling development in *Arabidopsis*. *Plant Molecular Biology,* 67, 151-167.

DELANO, W. L. 2002. The pymol molecular graphics system. DeLano Scientific, Palo Alto, CA, USA. http://www.pymol.org.

DELGADO, I. J., WANG, Z., DE ROCHER, A., KEEGSTRA, K. and RAIKHEL, N. V. 1998. Cloning and characterization of AtRGP1. A reversibly autoglycosylated *Arabidopsis* protein implicated in cell wall biosynthesis. *Plant Physiology,* 116, 1339-1349.

DELLEDONNE, M., XIA, Y., DIXON, R. A. and LAMB, C. 1998. Nitric oxide functions as a signal in plant disease resistance. *Nature,* 394, 585-588.

DHANANJEYAN, K. L., SIVAPERUMAL, R., PARAMASIVAN, R., THENMOZHI, V. and TYAGI, B. K. 2009. In-silico homology modeling of

three isoforms of insect defensins from the dengue vector mosquito, Aedes aegypti (Linn., 1762). *Journal of Molecular Modeling,* 15, 507-514.

DICKMAN, S. R. and CLOUTIER, A. A. 1950. Activation and stabilization of aconitase by ferrous ions. *Archives of Biochemistry,* 25, 229-231.

DIDONATO, R. J., ROBERTS, L. A., SANDERSON, T., EISLEY, R. B. and WALKER, E. L. 2004. *Arabidopsis* yellow stripe like 2 (YSL2): a metal regulated gene encoding a plasma membrane transporter of nicotianamine metal complexes. *The Plant Journal,* 39, 403-414.

DIFFLEY, J. F. and STILLMAN, B. 1991. A close relative of the nuclear chromosomal high mobility group protein HMG1 in yeast mitochondria. *Proceedings of the National Academy of Sciences of the United States of America,* 88, 7864-7868.

DIRUSSO, C. C. and BLACK, P. N. 1999. Long-chain fatty acid transport in bacteria and yeast. Paradigms for defining the mechanism underlying this protein-mediated process. *Molecular and Cellular Biochemistry,* 192, 41-52.

DUPUY, J., DARNAULT, C., BRAZZAOLOTTO, X., KUHN, L. C., MOULIS, J. M., VOLBEDA, A. and FONTECILLA-CAMPS, J. C. 2005. Crystallization and preliminary x-ray diffraction data for the aconitase form of human iron-regulatory protein 1. *Acta Crystallographica,* F61, 482-485.

DUPUY, J., VOLBEDA, A., CARPENTIER, P., DARNAULT, C., MOULIS, J. M. and FONTECILLA-CAMPS, J. C. 2006. Crystal structure of human iron regulatory protein 1 as cytosolic aconitase. *Structure,* 14, 129-139.

DURNER, J., WENDEHIENNE, D. and KLESSIG, D. F. 1998. Defense gene induction in tobacco by nitric oxide, cyclic GMP, and cyclic ADP ribose. *Proceedings of the National Academy of Sciences,* 95, 10328-10333.

EASTMOND, P. J., VERONIQUE, G., LANGE, P. R., BRYCE, J. H., SMITH, S. M. and GRAHAM, I. A. 2000. Postgerminative growth and lipid catabolism in oilseeds lacking the glyoxylate cycle. *Proceedings of the National Academy of Sciences,* 97, 5669-5674.

EASTMOND, P. J., and GRAHAM, I. A. 2001. Re-examining the role of the glyoxylate cycle in oilseeds. *Trends in Plant Science,* 6, 72-77.

ECKARDT, N. A. 2005. Peroxisomal citrate synthase provides exit route from fatty acid metabolism in oilseeds. *Plant Cell,* 17, 1863-1865.

EDWARDS, R. A., LEE, M. S., TSUTAKAWA, S. E., WILLIAMS, R. S., TAINER, J. A. and GLOVER, J. N. M. 2008. The BARD1 C-terminal domain structure and interactions with polyadenylation factor CstF-50. *Biochemistry,* 47, 11446-11456.

EIDE, D., BRODERIUS, M., FETT, J. and GUERINOT, M. L. 1996. A novel iron regulated metal transporter from plants identified by functional expression in yeast. *Proceedings of the National Academy of Sciences of the United States of America,* 93, 5624-5628.

EMPTAGE, M. H., DREYER, J. L. KENNEDY, M. C. and BEINERT, H. 1983. Optical and EPR characterization of different species of active and inactive aconitase. *The Journal of Biological Chemistry,* 258, 11106-11111.

ESWAR, N., WEBB, B., MARTI-RENOM, M. A., MADHUSUDHAN, M. S., ERAMIAN, D., SHEN, M., PIEPER, U. and SALI, A. 2007. Comparative protein structure modeling using MODELLER. *Current Protocols in Protein Science,* 2.9.1-2.9.31.

FATLAND, B. L., KE, J., ANDERSON, M. D., MENTZEN, W. I., CUI, L. W., ALLRED, C. C., JOHNSTON, J. L., NIKOLAU, B. J. and WURTELE, E. S.

2002. Molecular characterization of a heteromeric ATP-citrate lyase that generates cytosolic acetyl-coenzyme A in *Arabidopsis*. *Plant Physiology,* 130, 740-756.

FATLAND, B. L., NIKOLAU, B. J. and WURTELE, E. S. 2005. Reverse genetic characterization of cytosolic acetyl-CoA generation by ATP citrate lyase in *Arabidopsis*. *The Plant Cell,* 17, 182-203.

FAZIL, M. H., KUMAR, S., SUBBARAO, N., PANDEY, H. P. and SINGH, D.V. 2010. Homology modeling of a sensor histidine kinase from *Aeromonas hydrophila*. *Journal of Molecular Modeling,* 16, 1003-1009.

FESTA, M., COLONNA, A., PIETROPAOLO, C. and RUFFO, A. 2000. Oxalomalate, a competitive inhibitor of aconitase, modulates the RNA-binding activity of iron-regulatory protein. *Biochemical Journal,* 348, 315-320.

FIEHN, O. 2008. Extending the breadth of metabolite profiling by gas chromatography coupled to mass spectrometry. *Trends in Analytical Chemistry,* 27, 261-269.

FORBES, J. C. and WATSON, R. D. 1994. Plants in agriculture. Cambridge University Press.

FOX, T. C. and GUERINOT, M. L. 1998. Molecular biology of cation transport in plants. *Annual Review of Plant Physiology and Plant Molecular Biology,* 49, 669-696.

FRANGNE, N., EGGMANN, T., KOBLISCHKE, C., WEISSENBOCK, G., MARTINOIA, E. and KLEIN, M. 2002. Flavone glucoside uptake into barley mesophyll and *Arabidopsis* cell culture vacuoles. Energization occurs by H^+ antiport and ATP binding cassette type mechanisms. *Plant Physiology,* 128, 726-733.

FRENKEL, E. P. and KITCHENS, R. L. 1977. Purification and properties of acetyl-coenzyme A synthetase from bakers' yeast. *The Journal of Biological Chemistry,* 252, 504-507.

FREY, P. A. and HEGEMAN, A. D. 2007. Enzymatic reaction mechanisms. Oxford University Press.

FRISHMAN, D. and HENTZE, M. W. 1996. Conservation of aconitase residues revealed by multiple sequence analysis implications for structure/function relationships. *European Journal of Biochemistry,* 239, 197-200.

GANGLOFF, S. P., MARGUET, D. and LAUQUIN, G. J. 1990. Molecular cloning of the yeast mitochondrial aconitase gene (ACO1) and evidence of a synergistic regulation of expression by glucose plus glutamate. *Molecular and Cellular Biology,* 10, 3551-3561.

GAO, W., ANDERSON, P. J. and SADLER, J. E. 2008. Extensive contacts between ADAMTS13 exosites and von willebrand factor domain A2 contribute to substrate specificity. *Blood,* 112, 1713-1719.

GARAVITO, R. M., PICOT, D. and LOLL, P. J. 1996. Strategies for crystallizing membrane proteins. *Journal of Bioenergetics and Biomembranes*, 28, 13-27.

GARAVITO, R. M. and FERGUSON-MILLER, S. 2001. Detergents as tools in membrane biochemistry. *The Journal of Biological Chemistry,* 276, 32403-32406.

GARDNER, J. G., GRUNDY, F. J., HENKIN, T. M. and ESCALANTE-SEMERENA, J. C. 2006. Control of acetyl-coenzyme A synthetase (AcsA) activity by acetylation/deacetylation without NAD^+ involvement in *Bacillus subtilis*. *Journal of Bacteriology,* 188, 5460-5468.

GASTEIGER, E., GATTIKER, A., HOOGLAND, C., IVANYI, I., APPEL, R. D. and BAIROCH, A. 2003. ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic acids research,* 3784-3788.

GAWRON, O., WAHEED, A., GLAID, A. J. and JAKLITSCH, A. 1974. Iron and aconitase activity. *Biochemical Journal,* 139, 709-714.

GEGOUT, V., SCHLEGL, J., SCHLAGER, B., HENTZE, M. W., REINBOLT, J., EHRESMANN, B., EHRESMANN, C. and ROMBY, P. 1999. Ligand induced structural alterations in human iron regulatory protein 1 revealed by protein footprinting. *The Journal of Biological Chemistry,* 274, 15052-15058.

GEMEL, J. and RANDALL, D. D. 1992. Light regulation of leaf mitochondrial pyruvate dehydrogenase complex – role of photorespiratory carbon metabolism. *Plant Physiology,* 100, 908-914.

GOODSELL, D. S. and OLSON, A. J. 2000. Structural symmetry and protein function. *Annual Review of Biophysics and Biomolecular Structure,* 29, 105-153.

GONZALEZ, D. O. and VODKIN, L. O. 2007. Specific elements of the glyoxylate pathway play a significant role in the functional transition of the soybean cotyledon during seedling development. *BMC Genomics,* 8, 1-22.

GOURLEY, B. L., PARKER, S. B., JONES, B. J., ZUMBRENNEN, K. B. and LEIBOLD, E. A. 2003. Cytosolic aconitase and ferritin are regulated by iron in *Caenorhabditis elegans. The Journal of Biological Chemistry,* 278, 3227-3234.

GOUT, E., BLIGNY, R., PASCALS, N. and DOUCE, R. 1993. ^{13}C nuclear magnetic resonance studies of malate and citrate synthesis and compartmentation in higher plant cells. *The Journal of Biological Chemistry,* 268, 3986-3992.

GOUVEA, C. M. C. P., SOUZA, J. F., MAGALHAES, A. C. N. and MARTINS, I. S. 1997. NO releasing substances that induce growth elongation in maize root segments. *Plant Growth Regulation,* 21, 183-187.

GREELEY, B. H. 2004. Ramachandran plots. Amino acid configurations in proteins. http://www.greeley.org/~hod/papers/Unsorted/Ramachandran.doc.pdf.

GROTZ, N. and GUERINOT, M. L. 2006. Molecular aspects of Cu, Fe and Zn homeostasis in plants. *Biochimica et Biophysica Acta,* 1763, 595-608.

GRUER, M. J. and GUEST, J. R. 1994. Two genetically distinct and differentially regulated aconitases (acnA and acnB) in *Escherichia coli. Microbiology,* 140, 2531-2541.

GRUER, M. J., ARTYMIUK, P. J. and GUEST, J. R. 1997. The aconitase family: three structural variations on a common theme. *Trends in Biochemical Sciences,* 22, 3-6.

GUERINOT, M. L. and YI, Y. 1994. Iron: nutritious, noxious, and not readily available. *Plant Physiology,* 104, 815-820.

GUERINOT, M. L. 2000. The ZIP family of metal transporters. *Biochimica et Biophysica Acta,* 1465, 190-198.

GUEX, N. and PEITSCH, M. C. 1997. SWISS MODEL and the Swiss-PDB Viewer: An environment for comparative protein modeling. *Electrophoresis,* 18, 2714-2723.

GULICK, A. M., STARAI, V. J., HORSWILL, A. R., HOMICK, K. M. and ESCALANTE-SEMERENA, J. C. 2003. The 1.75 Å crystal structure of acetyl-CoA synthetase bound to adenosine-5'-propylphosphate and coenzyme A. *Biochemistry,* 42, 2866-2873.

GULICK, A. M., LU, X. and DUNAWAY-MARIANO, D. 2004. Crystal structure of 4-chlorobenzoate:CoA ligase/synthetase in the unligated and aryl substrate-bound states. *Biochemistry,* 43, 8670-8679.

GULICK, A. M. 2009. Conformational dynamics in the acyl-CoA synthetases, adenylation domains of non-ribosomal peptide synthetases and firefly luciferase. *ACS Chemical Biology,* 4, 811-827.

GUO, B., BROWN, F. M., PHILLIPS, J. D., YU, Y. and LEIBOLD, E. A. 1995a. Characterization and expression of iron regulatory protein 2 (IRP2) – presence of multiple IRP2 transcripts regulated by intracellular iron levels. *The Journal of Biological Chemistry,* 270, 16529-16535.

GUO, B., PHILLIPS, J. D., YU, Y. and LEIBOLD, E. A. 1995b. Iron regulates the intracellular degradation of iron regulatory protein 2 by the proteasome. *The Journal of Biological Chemistry,* 270, 21645-21651.

HAILE, D. J., ROUAULT, T. A., TANG, C. K., CHIN, J., HARFORD, J. B. and KLAUSNER, R. D. 1992. Reciprocal control of RNA binding and aconitase activity in the regulation of the iron responsive element binding protein: role of the iron sulfur cluster. *Proceedings of the National Academy of Sciences,* 89, 7536-7540.

HALLIWELL, B. and GUTTERIDGE, J. M. C. 1984. Oxygen toxicity, oxygen radicals, transition metals and disease. *Biochemical Journal,* 219, 1-14.

HALLIWELL, B. and CHIRICO, S. 1993. Lipid peroxidation: its mechanism, measurement, and significance. *The American Journal of Clinical Nutrition,* 57, 715-725.

HAMMEL, M., NEMECEK, D., KEIGHTLEY, J. A., THOMAS, G. J. and GEISBRECHT, B. V. 2007. The *Staphylococcus aureus* extracellular adherence protein (Eap) adopts an elongated but structured conformation in solution. *Protein Science,* 16, 2605-2617.

HARRISON, P. M. and AROSIO, P. 1996. The ferritins: molecular properties, iron storage function and cellular regulation. *Biochimica et Biophysica Acta,* 1275, 161-203.

HAYASHI, M., DE BELLIS, L., ALPI, A. and NISHIMURA, M. 1995. Cytosolic aconitase participates in the glyoxylate cycle in etiolated pumpkin cotyledons. *Plant and Cell Physiology,* 36, 669-680.

HAYASHI, M., TORIYAMA, K., KONDO, M. and NISHIMURA, M. 1998. 2,4-Dichlorophenoxybutyric acid resistant mutants of *Arabidopsis* have defects in glyoxysomal fatty acid β-oxidation. *The Plant Cell,* 10, 183-195.

HAYDON, M. J. and COBBETT, C. S. 2007. Transporters of ligands for essential metal ions in plants. *New Phytologist,* 174, 499-506.

HEINEMANN, U., BUSSOW, K., MUELLER, U. and UMBACH, P. 2003. Facilities and methods for the high throughput crystal structural analysis of human proteins. *Accounts of Chemical Research,* 36, 157-163.

HELL, R. and STEPHEN, U. W. 2003. Iron uptake, trafficking and homeostasis in plants. *Planta,* 216, 541-551.

HENDERSON, B. R., MENOTTI, E., BONNARD, C. and KUHN, L. C. 1994. Optimal sequence and structure of iron responsive elements. Selection of RNA stem loops with high affinity for iron regulatory factor. *Journal of Biological Chemistry,* 269, 17481-17489.

HENSON, C. P. and CLELAND, W. W. 1967. Purification and kinetic studies of beef liver cytoplasmic aconitase. *Journal of Biological Chemistry,* 242, 3833-8.

HENTZE, W. H. and ARGOS, P. 1991. Homology between IRE-BP, a regulatory RNA-binding protein, aconitase, and isopropylmalate isomerase. *Nucleic acids research,* 19, 1739-1740.

HIGGINS, D. R. and CREGG, J. M. 1998. *Pichia* protocols. *Methods in Molecular Biology,* 103.

HIRLING, H., HENDERSON, B. R. and KUHN, L. C. 1994. Mutational analysis of the [4Fe-4S] cluster converting iron regulatory factor from its RNA binding form to cytoplasmic aconitase. *EMBO Journal,* 13, 453-461.

HISANAGA, Y., AGO, H., NAKAGAWA, N., HAMADA, K., IDA, K., YAMAMOTO, M., HORI, T., ARII, Y., SUGAHARA, M., KURAMITSU, S., YOKOYAMA, S. and MIYANO, M. 2004. Structural basis of the substrate-specific two-step catalysis of long chain fatty acyl-CoA synthetase dimer. *The Journal of Biological Chemistry,* 279, 31717-26.

HOFFLAND, E., BOOGAARD, R., NELEMANS, J. and FINDENEGG, G. 1992. Biosynthesis and root exudation of citric and malic acid in phosphate starved rape plants. *New Phytologist,* 122, 675-680.

HOOKS, M. A., TURNER, J. E., MURPHY, E. C. and GRAHAM, I. A. 2004. Acetate non utilizing mutants of *Arabidopsis*: evidence that organic acids influence carbohydrate perception in germinating seedlings. *Molecular Genetics and Genomics,* 271, 249-256.

HOOKS, M. A., TURNER, J. E., MURPHY, E. C., JOHNSTON, K. A., BURR, S. and JAROSLAWSKI, S. 2007. The *Arabidopsis* ALDP protein homologue COMOTOSE is instrumental in peroxisomal acetate metabolism. *Biochemical Journal,* 406, 399-406.

HOOKS, M. A., ALLEN, E. and WATTIS, J. A. D. 2010. Modeling the peroxisomal carbon leak during lipid mobilization in *Arabidopsis*. *Biochemical Society Transactions,* 38, 1230-1234.

HORSWILL, A. R. and ESCALANTE-SEMERENA, J. C. 2002. Characterization of the propionyl-CoA synthetase (PrpE) enzyme of *Salmonella enterica*: residue Lys592 is required for propionyl-AMP synthesis. *Biochemistry,* 41, 2379-2387.

HOVMOLLER, S., ZHOU, T. and OHLSON, T. 2002. Conformations of amino acids in proteins. *Acta Crystallographica,* 58, 768-776.

HRUZ, T., LAULE, O., SZABO, G., WESSENDORP, F., BLEULER, S., OERTLE, L., WIDMAYER, P., GRUISSEM, W. and ZIMMERMANN, P. 2008. Genevestigator V3: A reference expression database for the meta analysis of transcriptomes. *Advances in Bioinformatics,* 1-5.

HUNTE, C. and MICHEL, H. 2002. Crystallization of membrane proteins mediated by antibody fragments. *Current Opinion in Structural Biology,* 12, 503-508.

HURTH, M. A., SUH, S. J., KRETZSCHMAR, T., GEIS, T., BREGANTE, M., GAMBALE, F., MARTINOIA, E. and NEUHAUS, H. E. 2005. Impaired pH homeostasis in *Arabidopsis* lacking the vacuolar dicarboxylate transporter and analysis of carboxylic acid transport across the tonoplast. *Plant Physiology,* 137, 901-910.

HUTTON, D. and STUMPF, P. K. 1969. Fat metabolism in higher plants xxxvii. Characterization of the β-oxidation system from maturing and germinating castor bean seeds. *Plant Physiology,* 44, 508-516.

IGAMBERDIEV, A. U. and KLECKOWSKI, L. A. 2001. Implications of adenylate kinase governed equilibrium of adenylates on contents of free magnesium in plant cells and compartments. *The Biochemical Journal,* 360, 225-231.

INGRAM-SMITH, C. and SMITH, K. S. 2006. AMP-forming acetyl-CoA synthetases in archaea show unexpected diversity in substrate utilization. *Archaea,* 2, 95-107.

INGRAM-SMITH, C., WOODS, B. I. and SMITH, K. S. 2006. Characterization of the acyl substrate binding pocket of acetyl-CoA synthetase. *Biochemistry,* 45, 11482-11490.

JACOB, F. and MONOD, J. 1961. Genetic regulatory mechanisms in the synthesis of proteins. *Journal of Molecular Biology,* 3, 318-356.

JAIN, R., HAMMEL, M., JOHNSON, R. E., PRAKASH, L., PRAKASH, S. and AGGARWAL, A. K. 2009. Structural insights into yeast DNA polymerase δ by small angle x-ray scattering. *Journal of Molecular Biology,* 394, 377-382.

JEONG, J. and GUERINOT M. L. 2009. Homing in on iron homeostasis in plants. *Trends in Plant Science,* 14, 280-285.

JETTEN, M. S. M., STAMS, A. J. M. and ZEHNDER, A. J. B. 1989. Isolation and characterization of acetyl-coenzyme A synthetase from *Methanothrix soehngenii*. *Journal of Bacteriology,* 171, 5430-5435.

JIANG, L. W., BUNKELMANN, J., TOWILL, L., KLEFF, S. and TRELEASE, R. N. 1994. Identification of peroxisome membrane proteins (PMPs) in sunflower (*Helianthus annuus L.)* cotyledons and influence of light on the PMP developmental pattern. *Plant Physiology,* 106, 293-302.

JOGL, G. and TONG, L. 2004. Crystal structure of yeast acetyl-coenzyme A synthetase in complex with AMP. *Biochemistry,* 43, 1425-31.

JONES, S. and THORNTON, J. M. 1995. Protein-protein interactions: A review of protein dimer structures. *Progress in Biophysics and Molecular Biology,* 63, 31-65.

JONES, D. L. 1998. Organic acids in the rhizosphere - a critical review. *Plant and Soil,* 205, 25-44.

JORDAN, P. A., TANG, Y., BRADBURY, A. J., THOMSON, A. J. and GUEST, J. R. 1999. Biochemical and spectroscopic characterization of *Escherichia coli* aconitases (acnA and acnB). *The Biochemical Journal,* 334, 739-746.

JORDANOV, J., COURTOIS-VERNIQUET, F., NEUBURGER, M. and DOUCE, R. 1992. Structural investigations by extended x-ray absorption fine structure spectroscopy of the iron center of mitochondrial aconitase in higher plant cells. *Journal of Biological Chemistry,* 267, 16775-16778.

KALDY, P., MENOTTI, E., MORET, R. and KUHN, L. C. 1999. Identification of RNA binding surfaces in iron regulatory protein 1. *The EMBO Journal,* 18, 6073-6083.

KAMPFENKEL, K., VAN MONTAGU, M. and INZE, D. 1995. Effects of iron excess on *Nicotiana plumbaginifolia* plants – implications to oxidative stress. *Plant Physiology,* 107, 725-735.

KAPTAIN, S., DOWNEY, W. E., TANG, C., PHILPOTT, C., HAILE, D., ORLOFF, D. G., HARFORD, J. B., ROUAULT, T. A. and KLAUSNER, R. D. 1991. A regulated RNA binding protein also possesses aconitase activity. *Proceedings of the National Academy of Sciences,* 88, 10109-10113.

KAWATE, T. and GOUAUX, E. 2006. Fluorescence-detection size-exclusion chromatography for precrystallization screening of integral membrane proteins. *Structure,* 14, 673-681.

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 Physiology,* 123, 497-508.

KELLEY, L. A. and STERNBERG, M. J. E. 2009. Protein structure prediction on the web: a case study using the Phyre server. *Nature Protocols,* 4, 363-371.

KENNEDY, M. C., EMPTAGE, M. H., DREYER, J. L. and BEINERT, H. 1983. The role of iron in the activation-inactivation of aconitase. *Journal of Biological Chemistry,* 258, 11098-11105.

KENNEDY, M. C., MENDE-MUELLER, L., BLONDIN, G. A. and BEINERT, H. 1992. Purification and characterization of cytosolic aconitase from beef liver and its relationship to the iron responsive element binding protein. *Proceedings of the National Academy of Sciences of the United States of America,* 89, 11730-11734.

KHURANA, P., GOKHALE, R. S. and MOHANTY, D. 2010. Genome scale prediction of substrate specificity for acyl adenylate superfamily of enzymes based on active site residue profiles. *BMC Bioinformatics,* 11, 1-17.

KILEY, P. J. and BEINERT, H. 2003. The role of Fe-S proteins in sensing and regulation in bacteria. *Current Opinion in Microbiology,* 6, 181-185.

KIM, S. A., PUNSHON, T., LANZIROTTI, A., LI, L., ALONSO, J. M., ECKER, J. R., KAPLAN, J. and GUERINOT, M. L. 2006. Seed requires the vacuolar membrane transporter VIT1. *Science,* 314, 1295-1298.

KIM, S. A. and GUERINOT, M. L. 2007. Mining iron: Iron uptake and transport in plants. *FEBS letters,* 581, 2273-2280.

KISIELA, D. I., AULIK, N. A., ATAPATTU, D. N. and CZUPRYNSKI, C. J. 2010. N-terminal region of *Mannheimia haemolytica* leukotoxin serves as a mitochondrial targeting signal in mammalian cells. *Cellular Microbiology,* 12, 976-987.

KLAUSNER, R. D. and ROUAULT, T. A. 1993. A double life: cytosolic aconitase as a regulatory RNA binding protein. *Molecular and Cellular Biology,* 4, 1-5.

KOCHAN, G., PILKA, E. S., VON DELFT, F., OPPERMANN, U. and YUE, W. W. 2009. Structural snapshots for the conformation-dependent catalysis by human medium-chain acyl-coenzyme A synthetase ACSM2A. *Journal of Molecular Biology,* 388, 997-1008.

KORNBERG, H. L. and KREBS, H. A. 1957. The glyoxylate cycle as a stage in the conversion of fat to carbohydrate in castor beans. *Biochimica et Biophysica Acta,* 26, 531-537.

KOROLEVA, O. A., TOMLINSON, M. L., LEADER, D., SHAW, P. and DOONAN, J. H. 2005. High-throughput protein localization in *Arabidopsis* using *Agrobacterium*-mediated transient expression of GFP-ORF fusions. *The Plant Journal,* 41, 162-174.

KOZIN, M. B. and SVERGUN, D. I. 2000. A software system for automated and interactive rigid body modeling of solution scattering data. *Journal of Applied Crystallography,* 36, 865-868.

KROMER, S. 1995. Respiration during photosynthesis. *Annual Review of Plant Physiology and Plant Molecular Biology,* 46, 45-70.

KRUFT, V., EUBEL, H., JANSCH, L., WERHAHN, W. and BRAUN, H. P. 2001. Proteomic approach to identify novel mitochondrial proteins in *Arabidopsis*. *Plant Physiology,* 127, 1694-1710.

KUNG, C. C., HUANG, W. N., HUANG, Y. C. and YEH, K. C. 2006. Proteomic survey of copper-binding proteins in *Arabidopsis* roots by immobilized metal affinity chromatography and mass spectrometry. *Proteomics,* 6, 2746-2758.

KUNZE, M., PRACHAROENWATTANA, I., SMITH, S. M. and HARTIG, A. 2006. A central role for the peroxisomal membrane in glyoxylate cycle function. *Biochimica et Biophysica Acta,* 1763, 1441-1452.

KUSEL, K. and DRAKE, H. L. 1994. Acetate synthesis in soil from a bavarian beech forest. *Applied and Environmental Microbiology,* 60, 1370-1373.

LAEMMLI, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature,* 227, 680-685.

LANDSBERG, E. 1981. Organic acid synthesis and release of hydrogen ions in response to Fe deficiency stress of mono- and dicotyledonous plant species. *Journal of Plant Nutrition,* 3, 579-591.

LANQUAR, V., LELIEVRE, F., BOLTE, S., HAMES, C., ALCON, C., NEUMANN, D., VANSUYT, G., CURIE, C., SCHRODER, A., KRAMER, U., BARBIER-BRYGOO, H. and THOMINE, S. 2005. Mobilization of vacuolar iron by AtNRAMP3 and AtNRAMP4 is essential for seed germination on low iron. *The EMBO Journal,* 24, 4041-4051.

LAUBLE, H., KENNEDY, M. C., BEINERT, H. and STOUT, C. D. 1992. Crystal structures of aconitase with isocitrate and nitroisocitrate bound. *Biochemistry,* 31, 2735-2748.

LAUBLE, H., KENNEDY, M. C., BEINERT, H. and STOUT, C. D. 1994. Crystal structures of aconitase with trans-aconitate and nitrocitrate bound. *Journal of Molecular Biology,* 237, 437-451.

LAUBLE, H., KENNEDY, M. C., EMPTAGE, M. H., BEINERT, H. and STOUT, C. D. 1996. The reaction of fluorocitrate with aconitase and the crystal structure of the enzyme inhibitor complex. *Proceedings of the National Academy of Sciences of the United States of America,* 93, 13699-13703.

LAWLAND, S., DORNE, A. J., LONG, D., COUPLAND, G., MARCHE, R. and CAROL, P. 2002. *Arabidopsis* A BOUT DE SOUFFLE, which is homologous with mammalian carnitine acyl carrier, is required for postembryonic growth in the light. *Plant Cell,* 14, 2161-2173.

LE JEAN, M., SCHIKORA, A., MARI, S., BRIAT, J. F. and CURIE, C. 2005. A loss of function mutation in AtYSL1 reveals its role in iron and nicotianamine seed loading. *The Plant Journal,* 44, 769-782.

LEE, H. Y., NA, K. B., KOO, H. M. and KIM, Y. S. 2001. Identification of active site residues in *Bradyrhizobium japonicum* acetyl-CoA synthetase. *Journal of Biochemistry,* 130, 807-813.

LEE, C. P., EUBEL, H., O'TOOLE, N. and MILLAR, A. H. 2008. Heterogeneity of the mitochondrial proteome for photosynthetic and non-photosynthetic *Arabidopsis* metabolism. *Molecular and Cellular Proteomics,* 7.7, 1297-1316.

LEHMANN, M., SCHWARZLANDER, M., OBATA, T., SIRIKANTARAMAS, S., BUROW, M., OLSEN, C. E., TOHGE, T., FRICKER, M. D., MOLLER, B. L., FERNIE, A. R., SWEETLOVE, L. J. and LAXA, M. 2008. The metabolic response of *Arabidopsis* roots to oxidative stress is distinct from that of heterotrophic cells in culture and highlights a complex relationship between the levels of transcripts, metabolites, and flux. *Molecular Plant,* 2, 390-406.

LEIGHTON, F., POOLE, B., LAZAROW, P. B. and DE DUVE, C. 1969. The synthesis and turnover of rat liver peroxisomes. Fractionation of peroxisome proteins. *The Journal of Cell Biology,* 41, 521-535.

LESHEM, Y. Y. 1996. Nitric oxide in biological systems. *Plant Growth Regulation,* 18, 155-159.

LI, Y., SHRESTHA, B. and VERTES, A. 2008. Atmospheric pressure infrared MALDI imaging mass spectrometry for plant metabolomics. *Analytical Chemistry,* 80, 407-420.

LIMMER, C. and DRAKE, H. L. 1998. Effects of carbon, nitrogen, and electron acceptor availability on anaerobic N_2 fixation in a beech forest soil. *Soil Biology and Biochemistry,* 30, 153-158.

LIN, M. and OLIVER, D. J. 2008. The role of acetyl-coenzyme A synthetase in *Arabidopsis*. *Plant Physiology,* 147, 1822-1829.

LLOYD, S. J., LAUBLE, H., PRASAD, G. S. and STOUT, C. D. 1999. The mechanism of aconitase: 1.8 Å resolution crystal structure of the S642A: citrate complex. *Protein Science,* 8, 2655-2662.

LODISH, H., BALTIMORE, D., BERK, A., ZIPURSKY, S. L., MATSUDAIRA, P. and DARNELL, J. 1995. Molecular Cell Biology. Third edition, W. H. Freeman and Co.

LOPEZ-MILLAN, A. F., MORALES, F., ANDALUZ, S., GOGORCENA, Y., ABADIA, A., DE LAS RIVAS, J. and ABADIA, J. 2000. Responses of sugar beet roots to iron deficiency. Changes in carbon assimilation and oxygen use. *Plant Physiology,* 124, 885-898.

LOPEZ-MILLAN, A. F., MORALES, F., GOGORCENA, Y., ABADIA, A. and ABADIA, J. 2009. Metabolic responses in iron deficient tomato plants. *Journal of Plant Physiology,* 166, 375-384.

LORENZ, M. C. and FINK, G. R. 2002. Life and death in a macrophage: role of the glyoxylate cycle in virulence. *Eukaryotic cell,* 1, 657-662.

LOVELACE, J., MILLER, G. W. and WELKIE, G. W. 1968. The accumulation of fluoroacetate and fluorocitrate in forage crops collected near a phosphate plant. *Atmospheric Environment,* 2, 187-190.

MA, J. F. and NOMOTO, K. 1993. Two related biosynthetic pathways of mugineic acids in gramineous plants. *Plant Physiology,* 102, 373-378.

MA, J. F. and NOMOTO, K. 1994. Incorporation of label from ^{13}C-, ^{2}H-, and ^{15}N-labeled methionine molecules during the biosynthesis of 2'-deoxymugineic acid in roots of wheat. *Plant Physiology,* 105, 607-610.

MARAIS, J. S. C. 1943. Isolation of the toxic principle "k cymonate" from "gifblaar" *Dichapetalum cymosum. Journal of Veterinary Science and Animal industry,* 18, 203-206.

MARIANAYAGAM, N. J., SUNDE, M. and MATTHEWS, J. M. 2004. The power of two: protein dimerisation in biology. *Trends in Biochemical Science,* 29, 618-625.

MARTI-RENOM, M. A., STUART, A. C., FISER, R. S., MELO, F. and SALI, A. 2000. Comparative protein structure modeling of genes and genomics. *Annual Review of Biophysics and Biomolecular Structure,* 29, 291-325.

MARTINEAU, Y., LEROUX, G. D. and SEOANE, J. R. 1994. Forage quality, productivity and feeding value to beef cattle of quackgrass (*Elytrigia repens* (L.) Nevski.) compared with timothy (*Phleum pretense* L.). *Animal Feed Science and Technology,* 47, 53-60.

MARTINEZ-BLANCO, H., REGLERO, A., FERNANDEZ-VALVERDE, M., FERRERO, M. A., MORENO, M. A., PENALVA, M. A. and LUENGO, J. M. 1992. Isolation and characterization of the acetyl-CoA synthetase from *Penicillium chrysogenum. The journal of Biological Chemistry,* 267, 5474-5481.

MAY, J. J., KESSLER, N., MARAHIEL, M. A. and STUBBS, M. T. 2002. Crystal structure of DhbE, an archetype for aryl acid activating domains of modular nonribosomal peptide synthetases. *Proceedings of the National Academy Sciences,* 99, 12120-12125.

McCAMMON, M. T. 1996. Mutants of *Saccharomyces cerevisiae* with defects in acetate metabolism: isolation and characterization of acn- mutants. *Genetics,* 144, 57-69.

McCLUSKEY, J., HERDMAN, L. and SKENE, K. R. 2004. Iron deficiency induces changes in metabolism of citrate in lateral and cluster roots of *Lupinus albus. Plant Physiology,* 121, 586-594.

McKELVEY, B. and MARSHALL, G. 2007. Food supply – can we meet the demand? *Journal of the Royal Agricultural Society of England,* 168, 1-6.

McPHERSON, A. 2004. Introduction to protein crystallization. *Methods,* 34, 254-265.

McREE, D. E. 1993. Practical Protein Crystallography. Second edition, Academic Press, San Diego, CA.

METTLER, I. J. and BEEVERS, H. 1980. Oxidation of NADH in glyoxysomes by a Malate-Aspartate shuttle. *Plant Physiology,* 66, 555-560.

MIDDAUGH, J., HAMEL, R., JEAN-BAPTISTE, G., BERIAULT, R., CHENIER, D. and APPANNA, V. D. 2005. Aluminum triggers decreased aconitase activity via Fe-S cluster disruption and the overexpression of isocitrate dehydrogenase and isocitrate lyase: a metabolic network mediating cellular survival. *Journal of Biological Chemistry,* 280, 159-165.

MILLAR, H. A., SWEETLOVE, L. J., GIEGE, P. and LEAVER, C. J. 2001. Analysis of the *Arabidopsis* mitochondrial proteome. *Plant Physiology,* 127, 1711-1727.

MILLERD, A. and BONNER, J. 1954. Acetate activation and acetoacetate formation in plant systems. *Archives of Biochemistry and Biophysics,* 49, 343-355.

MIREL, D. B., MARDER, K., GARZIANO, J., FREYER, G., ZHAO, Q., MAYEUZ, R. and MILHELMSEN, K. C. 1998. Characterization of the human mitochondrial aconitase gene (ACO2). *Gene,* 213, 205-218.

MISHRA, B. S., SINGH, M., AGGRAWAL, P. and LAXMI, A. 2009. Glucose and auxin signaling interaction in controlling *Arabidopsis thaliana* seedlings root growth and development. *PLoS ONE,* 4, e4502-4515.

MITCHELL, C. G. 1996. Identification of a multienzyme complex of the tricarboxylic acid cycle enzymes containing citrate synthase isoenzymes from *Pseudomonas aeruginosa. Biochemical Journal,* 313, 769-774.

MOEDER, W., DEL POZO, O., NAVARRE, D. A., MARTIN, G. B. and KLESSIG, D. F. 2007. Aconitase plays a role in regulating resistance to oxidative stress and cell death in *Arabidopsis* and *Nicotiana benthamiana. Plant Molecular Biology,* 63, 273-287.

MOHAMED, A. H. and ANDERSON, L. E. 1983. Light activation of purified aconitase by washed thylakoid membranes of pea (*Pisum satvum L.). Plant Physiology,* 71, 248-250.

MORRISON, J. F. 1954. The purification of aconitase. *Biochemical Journal,* 56, 99-105.

MURASHIGE, T. and SKOOG, F. 1962. A revised medium for rapid growth and bio assays with tobacco tissue cultures. *Plant Physiology,* 15, 473-497.

MURATA, Y., MA, J. F., YAMAJI, N., UENO, D., NOMOTO, K. and IWASHITA, T. 2006. A specific transporter for iron (III) phytosiderophores in barley roots. *The Plant Journal,* 46, 563-572.

MURGIA, I., DELLEDONNE, M. and SOAVE, C. 2002. Nitric oxide mediates iron induced ferritin accumulation in *Arabidopsis*. *The Plant Journal,* 30, 521-528.

NAVARRE, D. A., WENDEHENNE, D., DURNER, J., NOAD, R. and KLESSIG, D. F. 2000. Nitric oxide modulates the activity of tobacco aconitase. *Plant Physiology,* 122, 573-582.

NEUMANN, G. and ROMHELD, V. 1999. Root excretion of carboxylic acids and protons in phosphorus deficient plants. *Plant and Soil,* 211, 121-130.

NEWMANN, G., MASSONNEAU, A., MARTINOIA, E. and ROMHELD, V. 1999. Physiological adaptations to phosphorus deficiency during proteoid root development in white lupin. *Planta,* 208, 373-382.

NISHIMURA, M., YAMAGUCHI, J., MORI, H., AKAZAWA, T. and YOKOTA, S. 1986. Immunocytochemical analysis shows that glyoxysomes are directly transformed to leaf peroxisomes during greening of pumpkin cotyledons. *Plant Physiology,* 81, 313-316.

NJØS, A. and BØRRESE, T. 1991. Long term experiment with straw management, stubble cultivation, autumn and spring ploughing on a clay soil in S.E. Norway. *Soil and Tillage Research,* 21, 53-66.

O'ROURKE, T. W., DOUDICAN, N. A., MACKERETH, M. D., DOETSCH, P. W. and SHADEL, G. S. 2002. Mitochondrial dysfunction due to oxidative mitochondrial DNA damage is reduced through cooperative actions of diverse proteins. *Molecular and Cellular Biology,* 22, 4086-4093.

O'SULLIVAN, W. J. and SMITHERS, G. W. 1979. Stability constants for biologically important metal ligand complexes. *Methods in Enzymology,* 63, 294-336.

OELRICHS, P. B. and MCEVANS, T. 1961. Isolation of the toxic principle in *Acacia georginae. Nature,* 190, 808-809.

OSTERMEIER, C. and MICHEL, H. 1997. Crystallisation of membrane proteins. *Current Opinion in Structural Biology,* 7, 697-701.

PAIN, R. H. 1994. Mechanisms of protein folding. Oxford university press Inc., New York.

PANTOPOULOS, K., GRAY, N. K. and HENTZE, M. W. 1995. Differential regulation of two related RNA binding proteins, iron regulatory protein (IRP) and IRP_B. *RNA,* 1, 155-163.

PEASE, L. F., ELLIOTT, J. T., TSAI, D., ZACHARIAH, M. R. and TARLOY, M. J. 2008. Determination of protein aggregation with differential mobility analysis: application to IgG antibody. *Biotechnology and Bioengineering,* 101, 1214-1222.

PENFIELD, S., GRAHAM, S. and GRAHAM, I. A. 2005. Storage reserve mobilization in germinating oilseeds: *Arabidopsis* as a model system. *Biochemical Society Transactions,* 33, 380-384.

PETOUKHOV, M. V. and SVERGUN, D. I. 2003. New methods for domain structure determination of proteins from solution scattering data. *Journal of Applied Crystallography,* 36, 540-544.

PEYRET, P., PEREZ, P. and ALRIC, M. 1995. Structure, genomic organization, and expression of the *Arabidopsis thaliana* aconitase gene - plant aconitase show

significant homology with mammalian iron responsive element binding protein. *Journal of Biological Chemistry,* 270, 8131-8137.

PhastSystem Separation Technique File No. 120. 1986a. Amersham Biosciences.

PhastSystem Separation Technique File No. 121. 1986b. Amersham Biosciences.

PHILLIES, G. D. J. 1990. Quasielastic Light Scattering. *Analytical Chemistry,* 62, 1049-1057.

PHILPOTT, C. C., KLAUSNER, R. D. and ROUAULT, T. A. 1994. The bifunctional iron responsive element binding protein/cytosolic aconitase: the role of active site residues in ligand binding and regulation. *Proceedings of the National academy of Sciences of the United States of America,* 91, 7321-7325.

PINFIELD-WELLS, H., RYLOTT, E. L., GILDAY, A. D., GRAHAM, S., JOB, K., LARSON, T. R. and GRAHAM, I. A. 2005. Sucrose rescues seedling establishment but not germination of *Arabidopsis* mutants disrupted in peroxisomal fatty acid catabolism. *The Plant Journal,* 43, 861-872.

PITULA, J. S., DECK, K. M., CLARKE, S. L., ANDERSON, S. A., VASANTHAKUMAR, A. and EISENSTEIN, R. S. 2004. Selective inhibition of the citrate to isocitrate reaction of cytosolic aconitase by phosphomimetic mutation of serine 711. *Proceedings of the National Academy of Sciences,* 101, 10907-10912.

POPOVA, T. N. and PINHEIRO DE CARVALHO, M. A. A. 1998. Citrate and isocitrate in plant metabolism. *Biochimica et Biophysica Acta,* 1364, 307-325.

POVEY, S., SLAUGHTER, C. A., WILSON, D. E., GORMLEY, I. P., BUCKTON, K. E., PERRY, P. and BOBROW, M. 1976. Evidence for the assignment of the loci AK1, AK3 and ACONs to chromosome 9 in man. *Annals of Human Genetics,* 39, 413-422.

PRACHAROENWATTANA, I., CORNAH, J. E. and SMITH, S. M. 2005. *Arabidopsis* peroxisomal citrate synthase is required for fatty acid respiration and seed germination. *Plant Cell,* 17, 2037-2048.

PRACHAROENWATTANA, I., CORNAH, J. E. and SMITH, S. M. 2007. *Arabidopsis* peroxisomal malate dehydrogenase functions in beta oxidation but not in the glyoxylate cycle. *Plant Journal,* 50, 381-390.

PRACHAROENWATTANA, I. and SMITH, S. M. 2010. When is a peroxisome not a peroxisome? *Trends in Plant Science,* 13, 522-525.

PRESTON, G. G., WALL, J. D. and EMERICH, D. W. 1990. Purification and properties of acetyl-CoA synthetase from *Bradyrhizobium japonicum* bacteroids. *The Biochemical Journal,* 297, 179-183.

PRIVE, G. G. 2007. Detergents for the stabilization and crystallization of membrane proteins. *Methods,* 41, 388-397.

PROUSEK, J. 2007. Fenton chemistry in biology and medicine. *Pure and Applied Chemistry,* 79, 2325-2338.

PURICH, D. L. and FROMM, H. J. 1972. Studies on factors influencing enzyme responses to adenylate energy charge. *The Journal of Biological Chemistry,* 247, 249-255.

QU, X., SWANSON, R., DAY, R. and TSAI, J. 2009. A guide to template based structure prediction. *Current Protein and Peptide Science,* 10, 270-285.

RAMAKRISHNAN, C. and RAMACHANDRAN, G. N. 1965. Stereochemical criteria for polypeptide and protein chain conformations II. Allowed conformations for a pair of peptide units. *Biophysical Journal,* 5, 909-933.

RAMJEESINGH, M., HUAN, L., GARAMI, E. and BEAR, C. E. 1999. Novel method for evaluation of oligomeric structure of membrane proteins. *Journal of Biochemistry,* 342, 119-123.

RAYMOND, P., SPITERI, A., DIEUAIDE, M., GERHARDT, B. and PRADET, A. 1992. Peroxisomal β-oxidation of fatty acids and citrate formation by a particulate fraction from early germinating sunflower seeds. *Plant Physiology,* 30, 153-161.

REGER, A. S., CARNEY, J. M. and GULICK, A. M. 2007. Biochemical and crystallographic analysis of substrate binding and conformational changes in acetyl-CoA synthetase. *Biochemistry,* 46, 6536-6546.

REGER, A. S., WU, R., DUNAWAY-MARIANO, D. and GULICK, A. M. 2008. Structural characterization of a 140° domain movement in the two-step reaction catalysed by 4-chlorobenzonate:CoA ligase. *Biochemistry,* 47, 8016-8025.

REGEV RUDZKI, N., KARNIELY, S., BEN HAIM, N. N. and PINES, O. 2005. Yeast aconitase in two locations and two metabolic pathways: seeing small amounts is believing. *Molecular Biology of the Cell,* 16, 4163-4171.

RIC DE VOS, C. R., LUBBERDING, H. J. and BIENFAIT, H. F. 1986. Rhizosphere acidification as a response to iron deficiency in bean plants. *Plant Physiology,* 81, 842-846.

RICE, P., LONGDEN, I. and BLEASBY, A. 2000. EMBOSS: The European Molecular Biology Open Software Suite. *Trends in Genetics,* 16, 276-277.

RIGDEN, D., LIU, H., HAYES, S., URBE, S. and CLAGUE, M. 2009. Ab initio protein modeling reveals novel human MIT domains. *FEBS Letters,* 583, 872-878.

RIGAUD, J. L., CHAMI, M., LAMBERT, O., LEVY, D. and RANCK, J. L. 2000. Use of detergents in two dimensional crystallization of membrane proteins. *Biochimica et Biophysica Acta,* 1508, 112-128.

ROBBINS, A. H. and STOUT, C. D. 1989a. The structure of aconitase. *Proteins,* 5, 289-312.

ROBBINS, A. H. and STOUT, C. D. 1989b. Structure of activated aconitase: formation of the [4Fe-4S] cluster in the crystal. *Proceedings of the National academy of Sciences of the United States of America,* 86, 3639-3643.

ROBINSON, J. B. and SRERE, P. A. 1985. Organization of krebs tricarboxylic acid cycle enzymes in mitochondria. *The Journal of Biological Chemistry,* 260, 10800-10805.

ROBINSON, N. J., PROCTER, C. M., CONNOLLY, E. L. and GUERINOT, M. L. 1999. A ferric chelate reductase for iron uptake from soils. *Nature,* 397, 694-697.

ROMHELD, V. and MARSCHNER, H. 1986. Evidence for a specific uptake system for iron phytosiderophores in roots of grasses. *Plant Physiology,* 80, 175-180.

ROSE, I. A. and O'CONNELL, E. L. 1967. Mechanism of aconitase action .I. hydrogen transfer reaction. *Journal of Biological Chemistry,* 242, 1870-8.

ROSSO, M. G., LI, Y., STRIZHOV, N., REISS, B., DEKKER, K. and WEISSHAAR, B. 2003. An *Arabidopsis thaliana* T-DNA mutagenized population (GABI-Kat) for flanking sequence tag-based reverse genetics. *Plant Molecular Biology,* 53, 247-259.

ROUAULT, T. A. 2006. The role of iron regulatory proteins in mammalian iron homeostasis and disease. *Nature Chemical Biology,* 2, 406-414.

ROY, A., KUCUKURAL, A. and ZHANG, Y. 2010. I-TASSER: a unified platform for automated protein structure and function prediction. *Nature Protocols,* 5, 725-738.

RUFFER, M., STEIPE, B. and ZENK, M. H. 1995. Evidence against specific binding of salicylic acid to plant catalase. *FEBS Letters,* 377, 175-180.

SAAS, J., ZIEGELBAUER, K., VON HAESELER, A., FAST, B. and MICHAEL, B. 2000. A developmentally regulated aconitase related to iron regulatory protein 1 is localized in the cytoplasm and in the mitochondrion of *Trypanosoma brucei. The Journal of Biological Chemistry,* 275, 2745-2755.

SACHS, J. R. 1980. The order of release of sodium and addition of potassium in the sodium-potassium pump reaction mechanism. *Journal of Physiology,* 302, 219-240.

SALI, A. and BLUNDELL, T. L. 1993. Comparative protein modeling by satisfaction of spatial restraints. *Journal of Molecular Biology,* 234, 779-815.

SALI, A. and OVERINGTON, J. P. 1994. Derivation of rules for comparative protein modeling from a database of protein structure alignments. *Protein Science,* 3, 1582-1596.

SALI, A., POTTERTON, L., YUAN, F., VLIJMEN, H. and KARPLUS, M. 1995. Evaluation of comparative protein structure modeling by MODELLER. *Proteins,* 23, 318-326.

SALON, C., RAYMOND, P. and PRADET, A. 1988. Quantification of carbon fluxes through the tricarboxylic acid cycle in early germinating lettuce embryos. *Journal of Biological Chemistry,* 263, 12278-12287.

SANTI, S., CESCO, S., VARANINI, Z. and PINTON, R. 2005. Two plasma membrane H^+-ATPase genes are differentially expressed in iron deficient cucumber plants. *Plant Physiology and Biochemistry,* 43, 287-292.

SCANDALIOS, J. D. 1993. Oxygen stress and superoxide dismutases. *Plant Physiology,* 101, 7-12.

SCHAAF, G., LUDEWIG, U., ERENOGLU, B. E., MORI, S., KITAHARA, T. and VON WIREN, N. 2004. ZmYS1 functions as a proton coupled symporter for phytosiderophores and nicotianamine chelated metals. *The Journal of Biological Chemistry,* 279, 9091-9096.

SCHAAF, G., SCHIKORA, A., HABERLE, J., VERT, G., LUDEWIG, U., BRIAT, J. F., CURIE, C. and VON WIREN, N. 2005. A putative function for the *Arabidopsis* Fe phytosideropore transporter homolog AtYSL2 in Fe and Zn homeostasis. *Plant Cell Physiology,* 46, 762-774.

SCHAFFER, J. E. and LODISH, H. F. 1995. Molecular mechanism of long-chain fatty acid uptake. *TCM,* 5, 218-224.

SCHLIEBEN, N. H., NIEFIND, K. and SCHOMBURG, D. 2004. Expression, purification, and aggregation studies of His-tagged theroalkalophilic lipase from *Bacillus thermocatenulatus. Protein Expression and Purification,* 34, 103-110.

SCHMELZ, S. and NAISMITH, J. H. 2009. Adenylate forming enzymes. *Current Opinion in Structural Biology,* 19, 666-671.

SCHOPFER, P., PLACHY, C. and FRAHRY, G. 2001. Release of reactive oxygen intermediates (superoxide radicals, hydrogen peroxide, and hydroxyl radicals) and peroxidase in germinating radish seeds controlled by light, gibberellins, and abscisic acid. *Plant Physiology,* 125, 1591-1602.

SCHRADER, M. and YOON, Y. 2007. Mitochondria and peroxisomes: are the 'big brother' and the 'little sister' closer than assumed? *BioEssays,* 29, 1105-1114.

SCHUCK, P. 2003. On the analysis of protein self association by sedimentation velocity analytical ultracentrifugation. *Analytical Biochemistry,* 320, 104-124.

SCHULZ, G. E. and SCHIRMER, R. H. 1979. Principles of protein structure. Springer-Verlag, New York.

SCHWITZGUEBEL, J. P. and SIEGENTHALER, P. A. 1984. Purification of peroxisomes and mitochondria from spinach leaf by percoll gradient centrifugation. *Plant Physiology,* 75, 670-674.

SCOTLAND, R. W. and WORTLEY, A. H. 2003. How many species of seed plants are there? *Taxon,* 52, 101-104.

SHADEL, G. S. 2005. Mitochondrial DNA, aconitase 'wraps' it up. *Trends in Biochemical Sciences,* 30, 294-296.

SHAH, M. B., INGRAM-SMITH, C., COOPER, L. L., QU, J., MENG, Y., SMITH, K. S. and GULICK, A. M. 2009. The 2.1 Å crystal structure of an acyl-CoA synthetase from *Methanosarcina acetivorans* reveals an alternative acyl-binding pocket for small branched acyl substrates. *Proteins,* 77, 685-698.

SHEEN, J. 1990. Metabolic repression of transcription in higher plants. *The Plant Cell,* 2, 1027-1038.

SHEN, W., WEI, Y., DAUK, M., TAN, Y., TAYLOR, D. C., SELVARAJ, G. and ZOU, J. 2006. Involvement of a glycerol-3-phosphate dehydrogenase in modulating the NADH/NAD^+ ratio provides evidence of a mitochondrial glycerol-3-phosphate shuttle in *Arabidopsis*. *The Plant Cell,* 18, 422-441.

SHIRES, T. K. 1982. Iron induced DNA damage and synthesis in isolated rat liver nuclei. *Biochemical Journal,* 205, 321-329.

SHLIZERMAN, L., MARSH, K., BLUMWALD, E. and SADKA, A. 2007. Iron shortage induced increase in citric acid content and reduction of cytosolic aconitase activity in *Citrus* fruit vesicles and calli. *Physiologia Plantarum,* 131, 72-79.

SHOCKEY, J. M., FULDA, M. S. and BROWSE, J. A. 2002. *Arabidopsis* contains nine long chain acyl-coenzyme A synthetase genes that participate in fatty acid and glycerolipid metabolism. *Plant Physiology,* 129, 1710-1722.

SHOCKEY, J. M., FULDA, M. S. and BROWSE, J. A. 2003. *Arabidopsis* contains a large superfamily of acyl-activating enzymes. Phylogenetic and biochemical analysis reveals a new class of acyl-coenzyme A synthetases. *Plant Physiology,* 132, 1065-76.

SHOJIMA, S., NISHIZAWA, N. K., FUSHIYA, S., NOZOE, S., IRIFUNE, T. and MORI, S. 1990. Biosynthesis of phytosiderophores. *In vitro* biosynthesis of 2'-deoxymugineic acid from L-methionine and nicotianamine. *Plant Physiology,* 93, 1497-1503.

SKOURI-GARGOURI, H., ALI, M. B. and GARGOURI, A. 2009. Molecular cloning, structural analysis and modeling of the AcAFP antifungal peptide from *Aspergillus clavatus. Peptides,* 30, 1798-1804.

SLAUGHTER, C. A., POVEY, S., CARRITT, B., SOLOMON, E. and BOBROW, M. 1978. Assignment of the locus ACONM to chromosome 22. *Cytogenetics and Cell Genetics,* 22, 223-225.

SMALL, I., PEETERS, N., LEGEAI, F. and LURIN, C. 2004. Predotar: a tool for rapidly screening proteomes for N-terminal targeting sequences. *Proteomics,* 4, 1581-1590.

SMITH, C. M., BRYLA, J. and WILLIAMSON, J. R. 1973. Regulation of mitochondrial α-ketoglutarate metabolism by product inhibition at α-ketoglutarate dehydrogenase. *The Journal of Biological Chemistry,* 249, 1497-1505.

SMITH, S. M. 2002. Does the glyoxylate cycle have an anaplerotic function in plants? *Trends in Plant Science,* 7, 12-13.

SMYTH, D. R., MROZKIEWICZ, M. K., MCGARATH, W. J., LISTWAN, P. and KOBE, B. 2003. Crystal structures of fusion proteins with large affinity tags. *Protein Science,* 12, 1313-1322.

SOUPENE, E. and KUYPERS, F. A. 2006. Multiple erythroid isoforms of human long-chain acyl-CoA synthetases are produced by switch of the fatty acid gate domains. *BMA Molecular Biology,* 7, 1-12.

STARAI, V. J., TAKAHASHI, H., BOEKE, J. D. and ESCALANTE-SEMERENA, J. C. 2002. Short chain fatty acid activation by acyl-coenzyme A synthetases requires SIR2 protein function in *Salmonella enterica* and *Saccharomyces cerevisiae. Genetics,* 163, 545-555.

STARAI, V. J. and ESCALANTE-SENERENA, J. C. 2004. Review. Acetyl-coenzyme A synthetase (AMP forming). *Cellular and Molecular Life Sciences,* 61, 2020-2030.

STEVENS, R. C. 2000. Design of high throughput methods of protein production for structural biology. *Structure,* 8, 177-185.

STINNETT, L., LEWIN, T. M. and COLEMAN, R. A. 2007. Mutagenesis of rat acyl-CoA synthetase 4 indicates amino acids that contribute to fatty acid binding. *Biochimica et Biophysica Acta,* 1771, 119-125.

STUIBLE, H., BUTTNER, D., EHLTING, J., HAHLBROCK, K. and KOMBRINK, E. 2000. Mutational analysis of 4-coumarate:CoA ligase identifies functionally important amino acids and verifies its close relationship to other adenylate forming enzymes. *FEBS letters,* 467, 117-122.

STUMPF, P. K. and CONN, E. E. 1980. The biochemistry of plants. *The Plant Cell,* volume 1. Academic Press Inc., New York.

SU, C., WANG, F., CIOLEK, D. and PAN, Y. C. E. 1994. Electrophoresis of proteins and protein protein complexes in native polyacrylamide gels using a horizontal gel apparatus. *Analytical Biochemistry,* 223, 93-98.

SUZUKI, T., AKIYAMA, S., FUJIMOTO, S., ISHIKAWA, M., NAKAO, Y. and FUKUDA, H. 1976. The aconitase of yeast. IV. Studies on iron and sulfur in yeast aconitase. *Journal of Biochemistry,* 80, 799-804.

SVERGUN, D. I. 1992. Determination of the Regularization Parameter in Indirect-Transform Methods Using Perceptual Criteria. *Journal of Applied Crystallography,* 25, 495-503.

SVERGUN, D. I., BARBERATO, C. and KOCH, M. H. J. 1995. CRYSOL- a program to evaluate x-ray solution scattering of biological macromolecules from atomic coordinates. *Journal of Applied Crystallography,* 28, 768-773.

SVERGUN, D. I., PETOUKHOV, M. V. and KOCH, M. H. J. 2001. Determination of domain structure of proteins from x-ray solution scattering. *Biophysical Journal,* 80, 2946-53.

SVERGUN, D. I. and KOCH, M. H. J. 2003. Small angle scattering studies of biological macromolecules in solution. *Reports on Progress in Physics,* 66, 1735-1782.

SWEETLOVE, L. J., HEAZLEWOOD, J. L., HERALD, V., HOLTZAPFFEL, R., DAY, D. A., LEAVER, C. L. and MILLAR, A. H. 2002. The impact of oxidative stress on *Arabidopsis* mitochondria. *Plant Journal,* 32, 891-904.

SWEETLOVE, L. J., BEARD, K. F. M., NUNES-NESI, A., FERNIE, A. R. and RATCLIFFE, R. G. 2010. Not just a circle: flux modes in the plant TCA cycle. *Trends in Plant Science,* 15, 462-470.

TAIR. 2009. www.arabidopsis.org.

TAIZ, L. 1992. The plant vacuole. *Journal of Experimental Biology,* 172, 113-122.

TAKADA, Y., KANEKO, N., ESUMI, H., PURDUE, P. E. and DANPURE, C. J. 1990. Human peroxisomal L-alanine: glyoxylate aminotransferase. Evolutionary loss of a mitochondrial targeting signal by point mutation of the initiation codon. *Biochemical Journal,* 268, 517-520.

TAKAO, M., ABURATANI, H., KOBAYASHI, K. and YASUI, A. 1998. Mitochondrial targeting of human DNA glycosylases for repair of oxidative DNA damage. *Nucleic Acids Research,* 26, 2917-2922.

TAN, Y. F., O'TOOLE, N., TAYLOR, N. L. and MILLAR, A. H. 2010. Role in interactions with proteins and oxidative stress induced damage to respiratory function. *Plant Physiology,* 152, 747-761.

TANG, Y. and GUEST, J. R. 1999. Direct evidence for mRNA binding and post transcriptional regulation by *Escherichia coli* aconitases. *Microbiology,* 145, 3069-3079.

TANG, Y., QUAIL, M. A., ARTYMIUK, P. J., GUEST, J. R. and GREEN, J. 2002. *Escherichia coli* aconitases and oxidative stress: post transcriptional regulation of soda expression. *Microbiology,* 148, 1027-1037.

TANG, Y., GUEST, J. R., ARTYMIUK, P. J. and GREEN, J. 2005. Switching aconitase B between catalytic and regulatory modes involved iron dependent dimer formation. *Molecular Microbiology,* 56, 1149-1158.

THAKUR, A. S., ROBIN, G., GUNCAR, G., SAUNDERS, N. F. W., NEWMAN, J., MARTIN, J. L. and KOBE, B. 2007. Improved success of sparse matrix protein crystallization screening with heterogeneous nucleating agents. *PLoS ONE,* 10, e1091-e1097.

THEIL, E. C. 1987. Ferritin: structure, gene regulation, and cellular function in animals, plants, and microbiology. *Annual Review of Biochemistry,* 56, 289-315.

THOMINE, S., WANG, R., WARD, J. M., CRAWFORD, N. M. and SCHROEDER, J. L. 2000. Cadmium and iron transport by members of a plant metal transporter family in *Arabidopsis* with homology to Nramp genes. *Proceedings of the National Academy of Sciences of the United States of America,* 97, 4991-4996.

THOMINE, S., LELIEVRE, F., DEBARBIEUX, E., SCHROEDER, J. I. and BARBIER-BRYGOO, H. 2003. AtNRAMP3, a multispecific vacuolar metal transporter involved in plant responses to iron deficiency. *The Plant Journal,* 34, 685-695.

THORNEYCROFT, D., SHERSON, S. M. and SMITH, S. M. 2001. Using gene knockouts to investigate plant metabolism. *Journal of Experimental Botany,* 52, 1593-1601.

TIAN, G. W., MOHANTY, A., CHARY, S. N., LI, S., PAAP, B., DRAKAKAKI, G., KOPEC, C. D., LI, J., EHRHARDT, D., JACKSON, D., RHEE, S. Y., RAIKHEL, N. V. and CITOVSKY, V. 2004. High throughput fluorescent tagging of full length *Arabidopsis* gene products in planta. *Plant Physiology,* 135, 25-38.

TIFFIN, L. O. 1966a. Iron translocation. Plant culture, exudates sampling, iron citrate analysis. *Plant Physiology,* 41, 510-514.

TIFFIN, L. O. 1966b. Iron translocation. Citrate/iron ratios in plant stem exudates. *Plant Physiology,* 41, 515-518.

TONG, W. H. and ROUAULT, T. A. 2007. Metabolic regulation of citrate and iron by aconitases: role of iron sulfur cluster biogenesis. *Biometals,* 20, 549-564.

TOWNS-ANDREW, E., BERRY, A., BORDAS, J., MANT, G. R., MURRAY, P. K., ROBERTS, K., SUMNER, I., WORGAN, J. S., LEWIS, R. and GABRIEL, A. 1989. Time resolved x-ray diffraction station - x-ray optics, detectors, and data acquisition. *Review of Scientific Instruments,* 60, 2346-2349.

TRAMONTANO, W. A. and BARREIRO, C. J. 1997. Alterations in pea root meristem proteins after butyrate exposure. *Phytochemistry,* 46, 991-995.

TREBLE, D. H., LAMPORT, D. T. and PETERS, R. A. 1962. The inhibition of plant aconitase hydratase (aconitase) by fluorocitrate. *Biochemical Journal,* 85, 113-115.

TSUCHIYA, D., SHIMIZU, N. and TOMITA, M. 2008. Versatile architecture of a bacterial aconitase B and its catalytic performance in the sequential reaction coupled with isocitrate dehydrogenase. *Biochinica et Biophysica Acta,* 1784, 1847-1856.

TURNER, J. E. 2003. Characterisation of the *ACETATE NON UTILISING MUTANT 1* of *Arabidopsis.* Ph.D, University of Wales, Bangor.

TURNER, J. E., GREVILLE, K., MURPHY, E. C. and HOOKS, M. A. 2005. Characterization of *Arabidopsis* fluoroacetate-resistant mutants reveals the principal mechanism of acetate activation for entry into the glyoxylate cycle. *Journal of Biological Chemistry,* 280, 2780-7.

VALKO, M., IZAKOVIC, M., MAZUR, M., RHODES, C. J. and TELSER, J. 2004. Role of oxygen radicals in DNA damage and cancer incidence. *Molecular and Cellular Biochemistry,* 266, 37-56.

VAN ROERMUND, C. W., ELGERSMA, Y., SINGH, N., WANDERS, R. J. and TABAK, H. F. 1995. The membrane of peroxisomes in *Saccharomyces cervisiae* is impermeable to NAD(H) and acetyl-CoA under in vivo conditions. *EMBO Journal,* 14, 3480-3486.

VANLERBERGHE, G. C. and McINTOSH, L. 1996. Signals regulating the expression of the nuclear gene encoding alternative oxidase of plant mitochondria. *Plant Physiology,* 111, 589-595.

VARGHESE, S., TANG, Y. and IMLAY, J. A. 2003. Contrasting sensitivities of *Escherichia coli* aconitases A and B to oxidative and iron depletion. *Journal of Bacteriology,* 185, 221-230.

VASQUET-VIVAR, J., KALYANARAMAN, B. and KENNEDY, M. C. 2000. Mitochondrial aconitase is a source of hydroxyl radical. *The Journal of Biological Chemistry,* 275, 14064-14069.

VERNIQUET, F., GAILLARD, J., NEUBURGER, M. and DOUCE, R. 1991. Rapid inactivation of plant aconitase by hydrogen peroxide. *Biochemical Journal,* 276, 643-648.

VERT, G., GROTZ, N., DEDALDECHAMP, F., GAYMARD, F., GUERINOT, M. L., BRAIT, J. F. and CURIE, C. 2002. IRT1, an *Arabidopsis* transporter essential for iron uptake from the soil and for plant growth. *Plant Cell,* 14, 1223-1233.

VIJAYAKRISHNAN, S., KELLY, S. M., GILBERT, R. J. C., CALLOW, P., BHELLA, D., FORSYTH, T., LINDSAY, J. G. and BYRON, O. 2010. Solution structure and characterization of the human pyruvate dehydrogenase complex core assembly. *Journal of Molecular Biology,* 28, 71-93.

VOLKOV, V. V. and SVERGUN, D. I. 2003. Uniqueness of ab initio shape determination in small-angle scattering. *Journal of Applied Crystallography,* 2003, 860-864.

VOLZ, K. 2008. The functional duality of iron regulatory protein 1. *Current Opinion in Structural Biology,* 18, 106-111.

VON WIREN, N., KLAIR, S., BANSAL, S., BRIAT, J. F., KHODR, H., SHIOIRI, T., LEIGH, R. A. and HIDER, R. C. 1999. Nicotianamine chelates both Fe^{III} and Fe^{II}. Implications for metal transport in plants. *Plant Physiology,* 119, 1107-1114.

VLACHAKIS, D. 2009. Theoretical study of the usutu virus helicase 3D structure, by means of computer aided homology modeling. *Theoretical Biology and Medical Modelling,* 6, 1-9.

WALDEN, W. E., SELEZNEVA, A. I., DUPUY, J., VOLBEDA, A., FONTECILLA-CAMPS, J. C., THEIL, E. C. and VOLZ, K. 2006. Structure of dual function iron regulatory protein 1 complexed with ferritin IRE-RNA. *Science,* 314, 1903-1908.

WALKER, K. J., STEVENS, P. A., STEVENS, D. P., MOUNTFORD, J. O., MANCHESTER, S. J. and PYWELL, R. F. 2004. The restoration and re-creation of species rich lowland grassland on land formerly managed for intensive agriculture in the UK. *Biological Conservation,* 119, 1-18.

WATKINS, P. A., MAIGUEL, D., JIA, Z. and PEVSNER, J. 2007. Evidence for 26 distinct acyl-coenzyme A synthetase genes in the human genome. *Journal of Lipid Research,* 48, 2736-2750.

WATERS, B. M., CHU, H. H., DIDONATO, R. J., ROBERTS, L. A., EISLEY, R. B., LAHNER, B., SALT, D. E. and WALKER, E. L. 2006. Mutations in *Arabidopsis* yellow stripe like 1 and yellow stripe like 3 reveal their roles in metal ion homeostasis and loading of metal ions in seeds. *Plant Physiology,* 141, 1446-1458.

WAUGH, D. S. 2005. Making the most of affinity tags. *Trends in Biotechnology,* 23, 316-320.

WEBSTER, L. T. and CAMPAGNARI, F. 1962. The biosynthesis of acetyl and butyryl adenylates. *The Journal of Biological Chemistry,* 237, 1050-1055.

WEBSTER, L. T. 1963. Studies of the acetyl-coenzyme A synthetase reaction. I. Isolation and characterization of enzyme-bound acetyl adenylate. *The Journal of Biological Chemistry,* 238, 4010-4015.

WEBSTER, L. T. 1967. Studies of the acetyl-coenzyme A synthetase reaction. V. The requirement for monovalent and divalent cations in partial reactions involving enzyme-bound acetyl adenylate. *The Journal of Biological Chemistry,* 242, 1232-1240.

WECKWORTH, W., WENZEL, K., and FIEHN, O. 2004. Process for the integrated extraction, identification, and quantification of metabolites, proteins and

RNA to reveal their co-regulation in biochemical networks. *Proteomics,* 4, 78-83.

WEIMAR, J. D., DIRUSSO, C. C., DELIO, R. and BLACK, P. N. 2002. Functional role of fatty acyl-coenzyme A synthetase in the transmembrane movement and activation of exogenous long chain fatty acids. *The Journal of Biological Chemistry,* 277, 29369-29376.

WENDEL, J. F., GOODMAN, M. M., STUBER, C. W. and BECKETT, J. B. 1988. New isozyme systems for maize (*Zea mays)*: Aconitase hydratase, adenylate kinase, NADH dehydrogenase, and shikimate dehydrogenase. *Biochemical Genetics,* 26, 421-445.

WIEGAND, G. and REMINGTON, S. J. 1986. Citrate synthase: structure, control and mechanism. *Annual Review of Biophysics and Biophysical Chemistry,* 15, 97-117.

WILLIAMS, L. E., PITTMAN, J. K. and HALL, J. L. 2000. Emerging mechanisms for heavy metal transport in plants. *Biochimica et Biophysica Acta,* 1465, 104-126.

WILLIAMS, C. H., STILLMAN, T. J., BARYNIN, V. V., SEDELNIKOVA, S. E., TANG, Y., GREEN, J., GUEST, J. R. and ARTYMIUK, P. J. 2002. *E. coli* aconitase B structure reveals a HEAT-like domain with implications for protein-protein recognition. *Nature Structural Biology,* 9, 447-452.

WINTER, D., VINEGAR, B., NAHAL, H., AMMAR, R., WILSON, G. V. and PROVART, N. J. 2007. An "electronic fluorescent pictograph" browser for exploring and analyzing large scale biological data sets. *PLos ONE,* 2, 718-730.

WISZNIEWSKI, A. A., KHOU, W., SMITH, S. M. and BUSSELL, J. D. 2009. Identification of two *Arabidopsis* genes encoding a peroxisomal oxidoreductase like protein and an acyl-CoA synthetase like protein that are required for responses to pro-auxins. *Plant Molecular Biology,* 69, 503-515.

WONG F. T., CHEN, A. Y., CANE, D. E. and KHOSLA, C. 2010. Protein-protein recognition between acyltransferases and acyl carrier proteins in multimodular polyketide synthases. *Biochemistry,* 49, 95-102.

WU, H., LI, L., DU, J., YUAN, Y., CHENG, X. and LING, H. Q. 2005. Molecular and biochemical characterization of the Fe (III) chelate reductase gene family in *Arabidopsis thaliana. Plant Cell Physiology,* 46, 1505-1514.

WU, R., REGER, A. S., CAO, J., GULICK, A. M. and DUNAWAY-MARIANO, D. 2007. Rational redesign of the 4-chlorobenzoate binding site of 4-chlorobenzoate: coenzyme A ligase for expanded substrate range. *Biochemistry,* 46, 14487-14499.

WYNN, P. 2006. Changing face of crop production. *Royal Agricultural Society of England,* 167, 1-6.

XU, F., TAN, X. and WANG, Z. 2010. Effects of sucrose on germination and seedling development of *Brassica Napus. International Journal of Biology,* 2, 150-154.

YADAV, M., NAYARISSERI, A., RAJPUT, G. S., JAIN, A., VERMA, A. and GUPTA, P. 2010. Comparative modeling of 3-oxoacyl-acyl-carrier protein synthase I/II in *Plasmodium falciparum* – a potent target of malaria. *International Journal of Bioinformatics Research,* 2, 1-4.

YANG, Z. M., YANG, H., WANG, J. and WANG, Y. S. 2004. Aluminum regulation of citrate metabolism for Al-induced citrate efflux in roots of *Cassia tora L. Plant Science,* 166, 1589-1594.

YE, H. 2006. Simultaneous determination of protein aggregation, degradation, and absolute molecular weight by size exclusion chromatography- multiangle laser light scattering. *Analytical Biochemistry,* 356, 76-85.

YOSHIDA, H., HENSGENS, C. M. H., METSKE VAN DER LANN, J., SUTHERLAND, J. D., HART, D. J. and DIJKSTRA, B. W. 2005. An approach to prevent aggregation during the purification and crystallization of wild type acyl-coenzyme A: isopenicillin N acyltransferase from *Penicillium chrysogenum. Protein Expression and Purification,* 41, 61-67.

YOUNG, O. A. and ANDERSON, J. W. 1974. Properties and substrate specificity of some reactions catalysed by a short chain fatty acyl-coenzyme A synthetase from seeds of *Pinus radiata. Biochemistry,* 137, 423-433.

YUAN, K. and WYSOCKA-DILLER, J. 2006. Phytohormone signaling pathways interact with sugars during seed germination and seedling development. *Journal of Experimental Botany,* 57, 3359-3367.

ZHAN, Y., SONG, X. and ZHOU, G. W. 2001. Structural analysis of regulatory protein domains using GST-fusion proteins. *Gene,* 281, 1-9.

ZHANG, Y. 2008. I-TASSER server for protein 3D structure prediction. *BCM Bioinformatics,* 9, 40-48.

ZHANG, Y. 2009. I-TASSER: Fully automated protein structure prediction in CASP8. *Proteins,* 77, 100-113.

ZHENG, L., KENNEDY, M. C., BEINERT, H. and ZALKIN, H. 1992. Mutational analysis of active site residues in pig heart aconitase. *The Journal of Biological Chemistry,* 267, 7895-7903.

ZHOU, L., JANG, J. C., JONES, T. L. and SHEEN, J. 1998. Glucose and ethylene signal transduction crosstalk revealed by an *Arabidopsis* glucose insensitive mutant. *Proceedings of the National Academy of Sciences of the United States of America,* 95, 10294-10299.

ZHOU, H., PANDIT, S. B., LEE, S. Y., BORREGUERO, J., CHEN, H., WROBLEWSKA, L. and SKOLNICK, J. 2007. Analysis of TASSER-based CASP7 protein structure prediction results. *Proteins,* 69, 90-97.

World Wide Web (WWW) addresses accessed over the course of this project

The *Arabidopsis* Information Resource
www.arabidopsis.org
Reverse complement converter
www.bioinformatics.org/SMS/rev_comp.html
Multalin sequence alignment tool
http://bioinfo.genotoul.fr/multalin/multalin.html
TargetP Server
www.cbs.dtu.dk/services/TargetP/
ClustalW alignment at the European Bioinformatics Institute
www.ebi.ac.uk/
Small angle scattering data analysis software
www.embl-hamburg.de/ExternalInfo/Research/Sax/software.html
ExPASy proteomics Server
www.expasy.org/
Primer 3: Primer design website
http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi
Phyre the protein homology / analogy recognition engine
www.sbg.bio.ic.ac.uk/phyre/index.cgi
SIGNAL, the Salk Institute Genomic Analysis Laboratory
http://signal.salk.edu/