Biotransormations with plant coenzyme A-dependent acyltransferases

Oliver David Cunningham

BIOTRANSFORMATIONS WITH PLANT COENZYME A-DEPENDENT ACYLTRANSFERASES

Oliver David Cunningham

PhD Thesis

2008

School of Biological and Biomedical Sciences
Durham University

Oliver D. Cunningham

Biotransformations with Plant Coenzyme A-Dependent Acyltransferases

Abstract

The acylation of plant natural products is widely observed in nature and this modification often confers novel bioactivities. In nature, each of these acylations is selectively catalysed by a coenzyme A-dependent acyltransferase and biotransformations with these enzymes offer several potential advantages, including pre-disposed regio- and substrate- selectivity, reduced side-reactions and increased yield.

The enzymatic introduction of non-natural acyl groups into natural product biosynthesis has the potential to diversify the application of these acylating enzymes, with several examples of unnaturally acylated natural products being reported to have novel bioactivity. To explore the potential for biosynthesising such acylated-derivatives *in vivo,* feeding studies with both natural and fluorinated phenylpropanoids were carried out in *Arabidopsis thaliana* and *Petunia hybrida.* In petunia, feeding with natural phenylpropanoids caused hyper-accumulation of acylated flavonoid products. This enhancement was particularly apparent when spraying the plants with the phenylpropanoyl methyl ester. Using the methyl esters, 4- hydroxycinnamic acid and its 4- fluorocinnamic acid analogue were apparently incorporated into quercetin 3- *O*- diglycoside biosynthesis in petunia. However, only the endogenous phenylpropanoids were able to be incorporated into acylated-anthocyanin biosynthesis in arabidopsis.

In order to understand the factors that govern substrate selectivity in coenzyme A-dependent acyltransfer, the associated enzymic mechanisms were investigated *in vitro.* Plants utilise a coenzyme A acyl donor substrate formed through the action of an ATP-dependent CoA ligase, and the respective enzyme (At4CL1) from arabidopsis was cloned and expressed in *E.coli.* The purified enzyme was shown to form a diverse range of CoA esters with phenylpropanoid derivatives bearing 4- or 3- 4- hydro, hydroxyl, fluoro and / or methoxy groups, in addition to 4- azidocinnamic acid substituents. These aromatic substituents were found to have a strong influence upon acyl substrate selectivity. A CoA-dependent acyltransferase which acylates cyanidin 3- 5- *O*- diglucoside in *Gentiana triflora* with phenylpropanoids was also cloned, expressed in *E. coli* and biochemically characterised. The acyltransferase was able to transfer various 4- and 3- 4- substituted phenylpropanoids. However, in this instance substrate selectivity was not as strongly influenced by phenylpropanoid aromatic substitution, which was indicative of a 'pre-screening' role of the CoA ligase with respect to acyl substrates. The *in vitro* biosynthesis of acylated flavonoids was optimised to transfer aromatic acids onto natural product acceptors using a one-pot biosynthetic approach, combining 4CL and acyltransferase activities. Consequently, coenzyme A was able to be efficiently re-cycled and was used on a much reduced scale.

In addition, new strategies have been developed to isolate interesting acyltransferase biocatalysts based on their inhibition by a biotinylated fluorophosphonate probe. This approach showed that it was possible to identify CoA-dependent acyltransferases from crude protein preparations by recovering affinity labelled enzymes using a streptavidin sepharose solid support. The reactivity of the probe toward BAHD acyltransferases could be enhanced in the presence of UV irradiation, whilst labelling was abolished in the presence of the acyl acceptor. This afforded the ability to isolate enzyme activities according to their substrate recognition.

Table of Contents

List of Figures

Declaration

No material presented here has previously been submitted for any other degree. Except where acknowledged, all material is the work of the author.

Statement of copyright

Publications arising from work described in this thesis

Cunningham, O. D., Edwards, R., 2007. Applying Activity – based Chemical Probes to Enzyme Discovery in Plant Secondary Metabolism. In *Phytochemical Society of Europe 2007* p 157 (meeting proceeding).

Cunningham, O. D., Edwards, R., 2007. Generating Novel Phytochemicals through Biocatalysis. In *XVI International Plant Protective Congress 2007.* In press. (meeting proceeding).

Acknowledgements

First and foremost I would like to express my gratitude to Rob Edwards for his thoughts, encouragement and guidance throughout the entire course of the PhD.

I would also like to thank Patrick Steel for his expertise in synthetic chemistry and his supervision in the early phase of the PhD.

Many thanks to my colleagues in both CY 204 / Lab 2 and CY 1 (past and present) for making the last 4 years enjoyable. In particular, Ian Cummins and Lesley Edwards proved to be invaluable throughout and their endless help and support was and is very much appreciated. David Dixon, Melissa Brazier-Hicks and Mark Skipsey were equally invaluable when providing assistance on proteomics (David), cloning (Mark), mass spectrometry (Melissa, David and Ian) and general procedures (All). In recent times, the Chapter readers have been a great help and, in particular, I would like to thank Markus, Kathryn and Aggie for their time and boredom.

The PhD would not have been possible without the help and dedication of the academic services provided in both biology and chemistry departments, including services specialising in NMR, mass spectrometry, DNA sequencing and proteomics. Many thanks to all involved.

On a financial note, I'd like to acknowledge the BBSRC for their funding.

Thanks to Rob Fields from UEA for provision of azido bearing metabolites used in this thesis.

Cheers to Markus and Ed, who fall in between the colleague category above and mate category below, but have been essential in both. Thanks Markus for the enthusiasm, discussion and beers. Cheers Ed for the hill running ("we'll just walk the first mile...") and Jagermeister, Joy, Joy!

Finally, thank you to my family, Mum, Dad, Laura and Henry, who are always supportive and help me in everything that I do. Cheers Al, Rich, Matt, Barry, Andrzej, Kiran and Khairul and many thanks to Anita for everything.

Abbreviations

4CL	4- coumarate coenzyme A ligase
AAT	Aromatic acyltransferase
AEBSF	4-(2-Aminoethyl) benzenesulfonyl fluoride
At	*Arabidopsis thaliana*
AT	Acyltransferase
At4CL 1	Isoenzyme 1 of the 4 CoA ligases found in *Arabidopsis thaliana*
CADAT	Coenzyme A-dependent acyltransferase
Caff	Caffeic acid
cDNA	Complementary DNA
CHS	Chalcone synthase
CL	Coenzyme A ligase
cv.	Cultivar
Da	Dalton
DAD	Diode array detection
DEPC	diethyl phosphorocyanidate
dNTP's	deoxynucleotide triphosphates
DTT	dithiothreitol
E-64	*trans*-Epoxy succinyl-L-leucyl-amido(4-guanidino)butane
EDTA	ethylenediaminetetraacetic acid
Fer	Ferulic acid
FP	Fluorophosphonate ester
FPP	Fluorophosphonate probe
Gent5AT	5-*O*-aromatic acyltransferase from *Gentiana triflora*

HPLC	High performance liquid chromatography
HRMS	High resolution mass spectrometry
IPTG	Isopropyl β–thiogalactoside
KDG	Kaempferol 3-*O*-diglucoside
MALT	Malonyl transferase
mRNA	Messenger RNA
MS	Mass spectrometry
PAL	Phenylalanine ammonia lyase
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PMSF	Phenyl methane sulphonyl fluoride
PPT	Phenylpropanoyl transferase
QDG	Quercetin 3-*O*-diglucoside
SDS	Sodium dodecyl sulphate
TLC	Thin layer chromatography
TLCK	tosyl-L-lysine chloromethyl ketone
TPCK	tosyl-L-phenylalanine chloromethyl ketone
UV-Vis	Ultra violet-visible light spectrometry

Dedication

For Meredith

Chapter 1

Acylation of natural products

This chapter will discuss the background and work already documented on the subject of enzymatic acyl transfer. Initially, the chapter will define the process and method of acyl transfer. The metabolites capable of acyl donation will be discussed and an overview of the current application of enzymatic acyltransfer will be given.

The biochemical literature upon coenzyme A-dependent acyltransfer pathways in natural product metabolism will be reviewed, including a section on the acylation of flavonoids. The literature on acylation in this class of natural products is vast, owing to the sheer number of acylated flavonoids found.

Several coenzyme A-dependent acyltransferases from plants have previously been identified and will be reviewed. This will include aspects of their endogenous activity and their potential application as highly specific biocatalysts for the biosynthesis of synthetically-challenging natural products.

Finally, the existing application and previously identified biotechnological potential of coenzyme A-dependent acyltransfer will be discussed.

1.1 Acylation

1.1.1 Nucleophillic substitution of carboxylic acid derivatives

Acylation, in chemical terms, can be defined as the introduction of an acyl group into a compound. Acyl groups can be incorporated onto various functional groups upon a molecule. However, the focus of this thesis will be upon the acylation of alcohols, amines or thiols, to produce esters, amides and thioesters respectively.

Acylation is one of the simplest, but most common reactions in chemistry. The synthesis of esters, amides and thioesters proceeds via nucleophillic addition at a carbonyl group, to form a tetrahedral intermediate. In basic conditions, the tetrahedral intermediate is unstable and elimination of the most favourable leaving group occurs. When the leaving group is one that existed on the original molecule, nucleophillic substitution (S_N2) is completed (Figure 1.1). Carboxylic acid derivatives (RCOX) that will normally undergo nucleophillic substitution with alcohols, amines or thiols in acidic or basic conditions, include acid chlorides[1], acid anhydrides[1], *N*-hydroxysuccinimide esters[2] and imidazolium amides[3].

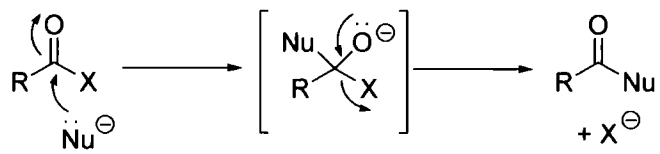


Figure 1.1 Nucleophillic substitution (S_N2) of carboxylic acid derivatives in basic conditions

Strong acid catalysts are able to promote ester formation at carboxylic acid derivatives with bad leaving groups, such as carboxylic acid, by increasing the electrophillicity of

the carbonyl group through protonation of carbonyl oxygen and lowering the pKa_H of a hydroxyl leaving group, also through protonation (Figure 1.2). However, this process is readily reversible by nucleophillic substitution with water and requires an excess of alcohol to drive the reaction toward ester formation.

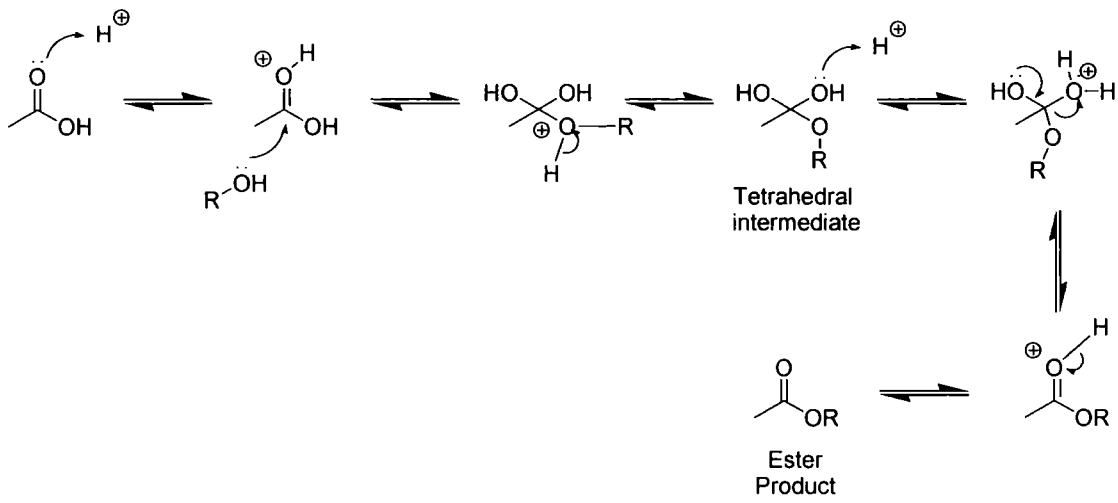


Figure 1.2 Acid catalysis of nucleophilic substitution (S_N2)

1.1.2 Enzymatic catalysis of heteroatom acylation

Throughout this thesis, acylations of heteroatoms that are catalysed by acyltransferase enzymes will be termed 'acyltransfers'. Equally, the electrophillic carboxylic acid derivatives that undergo nucleophillic substitution will normally be termed acyl donors and substrates that become acylated through nucleophillic addition to carboxyl groups will normally be termed acyl acceptors.

Similarly to chemical methodology, enzymes can catalyse acylations utilising carboxylic acid derivatives bearing both good and bad leaving groups. Various

carboxylic acid derivatives, or acyl donors, are employed by enzymes that catalyse acylations. Enzymatic acylation is observed during the synthesis of primary metabolites such as lipids[4] and proteins[5] and acyltransfer is also predominant in secondary metabolism, where enzymes catalyse the acylation of natural products[6].

Carboxylic acids[7] and carboxyl esters[8] are the simplest examples of acyl donor substrates in acyltransfer. Certain enzymes in the hydrolase family have been found to perform acylation reactions using these acyl donors and heteroatom acyl acceptors - the reverse reaction of their natural 'hydrolysis' function. Hydrolase enzymes able to catalyse acyltransfer, including lipases[9] and serine proteases[10], have been recruited for esterification in many natural and non-natural biotransformations[11].

The vast majority of hydrolases, and many acyltransferases, possess a combination of serine, histidine and aspartate residues as part of an active site catalytic triad[7,12]. Each of these residues has a role in generating the nucleophillic potential of the serine oxygen atom. The aspartate forms a hydrogen bond with the imidazole group of histidine, aligning the group with the serine alcohol, whilst increasing the pK_a and basicity of the imidazole. During catalysis, the histidine imidazole accepts a proton from serine, becomes positively charged and forms a transient salt bridge with aspartate (Figure 1.3). The oxyanion formed following nucleophillic addition of serine, is often stabilised by hydrogen bond interaction with main chain amide residues, termed the 'oxyanion hole'. Variations in this mechanism include replacement of serine for a cysteine nucleophile, orientation of the main chain amino acids surrounding the nucleophile and the proximity of aspartate to histidine[12].

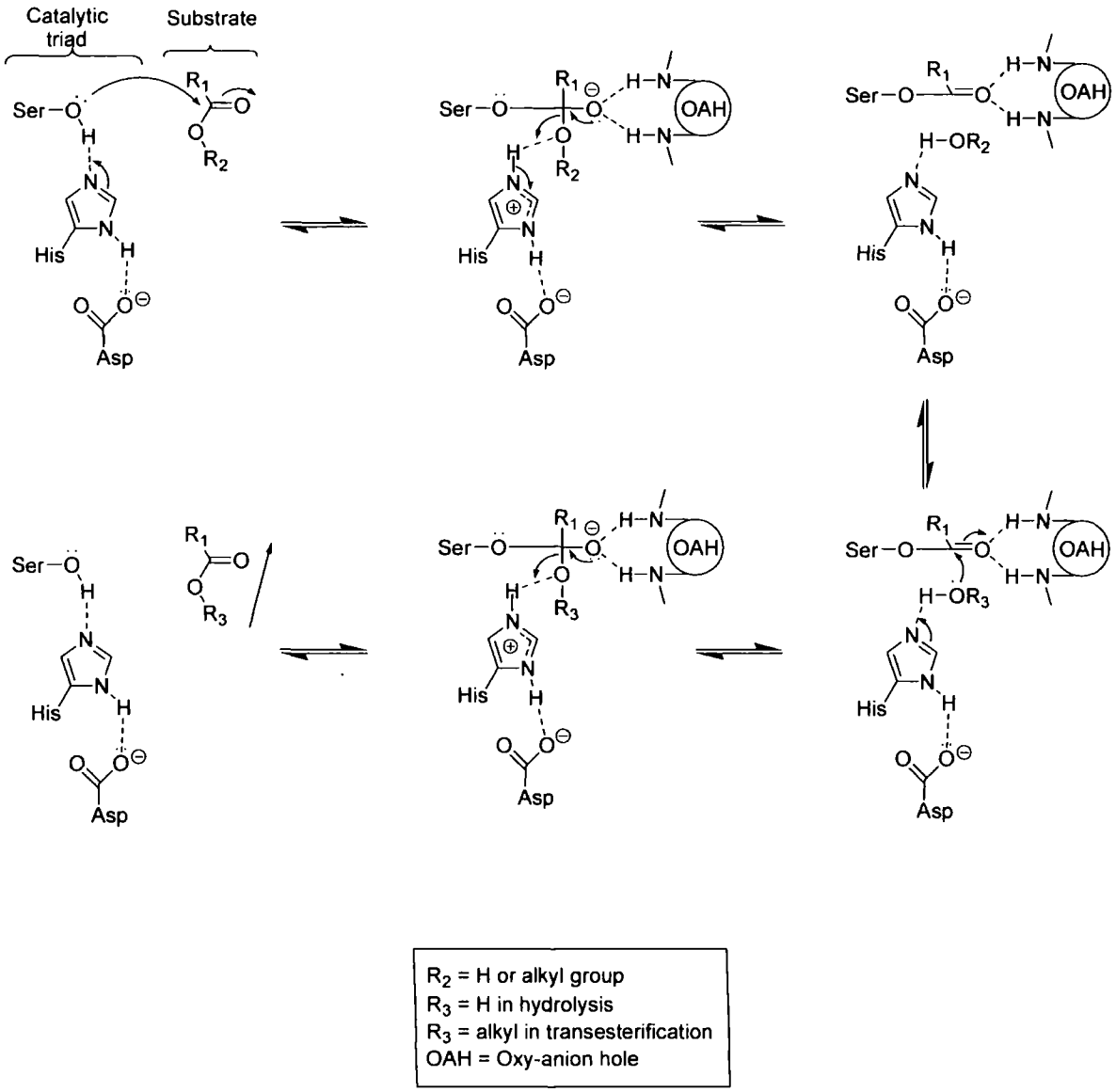


Figure 1.3 Hydrolysis (R_3 = H) or transesterification (R_3 = alkyl)) of carboxylic acid derivatives by certain Ser-His-Asp-based hydrolases

In nature, serine carboxypeptidase-like proteins catalyse the acyltransfer of acyl groups from 1-*O*-acyl-β-glucose to C-2 of a second molecule of 1-*O*-acyl-β-glucose, forming glucose polyesters in lipid biosynthesis[7]. The glucose acyl donor is initially biosynthesised from a carboxylic acid and UDP-glucose, comprising a 2-enzyme pathway of acyltransfer: initial biosynthesis of an acyl donor bearing a good leaving group prior to acyltransfer[13].

Similarly, 2-enzyme pathways are employed to biocatalyse several other acyltransfers in nature. Each of the following pathways biosynthesise an 'acyl donor' bearing a good leaving group, from a carboxylic acid and adenosine 5'-triphosphate (ATP), prior to acyltransfer. However, the strategies employed to produce the acyl donors varies remarkably.

During protein biosynthesis, amino acids are substrate to ATP-dependant amino acyl transfer tRNA synthetase enzymes. The acid group of each amino acid is converted into an acyl phosphate. This acyl donor is enzymatically transferred onto a specific tRNA alcohol acceptor, before this secondary acyl donor is bound to a ribosome for peptidyl transfer (Figure 1.4 A)[5]. Conversely, acetyl coenzyme A biosynthesis in bacteria is catalysed by an acetate kinase and a phosphotransacetylase. The kinase converts acetic acid into acetyl phosphate with ATP as cofactor. The acyltransfer is catalysed by a phosphotransacetylase, which promotes the nucleophillic substitution of acetyl phosphate with the free thiol of coenzyme A (Figure 1.4 B)[14]. In plants a single enzyme, coenzyme A ligase, uses both ATP and CoA as cofactors to convert acids into acyl-CoA thioesters, via acyl-phosphoadenylate (Figure 1.4 C)[15]. Plant coenzyme A thioesters are subsequently utilised as acyl donors by coenzyme-A dependent acyltransferases (CADATs) in fatty acid biosynthesis[16] and in secondary metabolism[6].

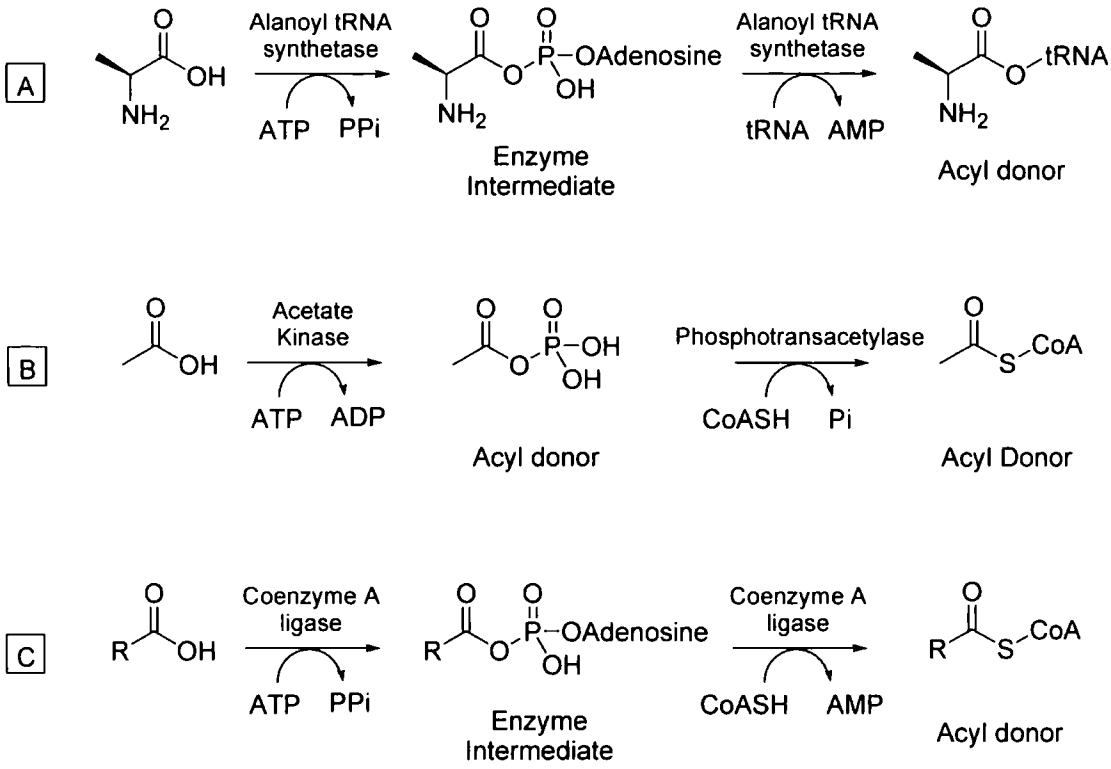


Figure 1.4 Biosynthesis of acyl donors required for multi-enzyme acyltransfer pathways

The acyltransferases which utilise biosynthesised acyl donors (with the exception of peptidyl transfer), catalyse acyltransfer by a similar nucleophile-based mechanism to a hydrolase[7,12,17,18]. Therefore the production of an acyl donor with a good leaving group is unlikely to be a *necessary* modification for enzyme catalysis as serine catalytic triads are known to use carboxyl esters or acids as 'acyl donor' substrates. More likely is that biosynthesis of the acyl donor allows for a more substrate-specific acyltransfer, by effectively pre-screening the carboxylic acids with a CoA ligase and subsequent 'recognition' of the synthesised carboxylic acid derivative by the respective acyltransferase. Thus a highly substrate-selective acyltransfer is achieved.

1.1.2.1 Biocatalysis of acyltransfer

Of the different enzymes able to catalyse acyltransfer only lipases, and to a lesser extent serine proteases, have been recruited for industrial biotransformations. Lipases have been used in organic chemistry for many years and a wealth of research has been compiled on their use[19]. Their success in this role arises from several characteristics. A high stability toward organic solvents allows metabolism of hydrophobic molecules and catalysis of condensation or transesterification with increased efficiency, in the absence of water[9]. Biochemical advantages of the use of hydrolases for the biocatalysis of acylation, include their broad substrate selectivity, no dependance on cofactors and high regio-, chemo- and stereoselectivity[19]. However, matching a lipase to an application is hard to predict and requires extensive screening, particularly with respect to large molecules that bear multiple, potentially reactive groups. Research in this area is underway and intensive[11].

In nature, one of the most diverse enzymatic acylations from species to species occurs in plant secondary metabolism. The vast majority of these acylations are catalysed by coenzyme A ligases and coenzyme A-dependent acyltransferases (CADATs), otherwise known as BAHD acyltransferases. The term BAHD was taken from the first four examples of this class of enzymes to be characterised - benzylalcohol *O*-acetyltransferase (BEAT), anthocyanin *O*- hydroxycinnamoyltransferase (AHT), *N*-hydroxycinnamoyl/benzoyltransferase (HCBT), deacetylvindoline 4- *O*-acetyltransferase (DAT). Research into the characterisation of these pathways is steadily increasing, with a view toward regiospecific acylation of several bioactive natural products found in plants[20,21]. The major advantage of the use of these enzymes

for biotransformations, is the prospect of recruiting the actual enzyme which catalyses a particular regiospecific acyltransfer onto a *natural* product, without the need for screening. Secondly, CADATs are able to catalyse an extremely diverse array of acylations and are capable of transfering a large array of carboxylic acids in a specific manner, where at least the carboxyl substrate is predictable for success [6]. However, the industrial application of these enzymes is hindered by the requirement of two enzymes, expensive cofactors, unknown substrate selectivity and the limited number of enzymes characterised to date.

The following sections review the topic of coenzyme A-dependent acyl transfer in plant secondary metabolism, in order to guage the potential of these enzymes for applied biotransformation in the future.

1.2 Biosynthesis of coenzyme A acyl donors in plant secondary metabolism

1.2.1 Carboxylic acids incorporated into coenzyme A-dependent secondary metabolism

Plants synthesise an array of simple aliphatic and aromatic carboxylic acids from primary metabolites[22,23]. The metabolites shown in figure 1.5 are common substrates in enzymatic acyltransfer onto natural product scaffolds.

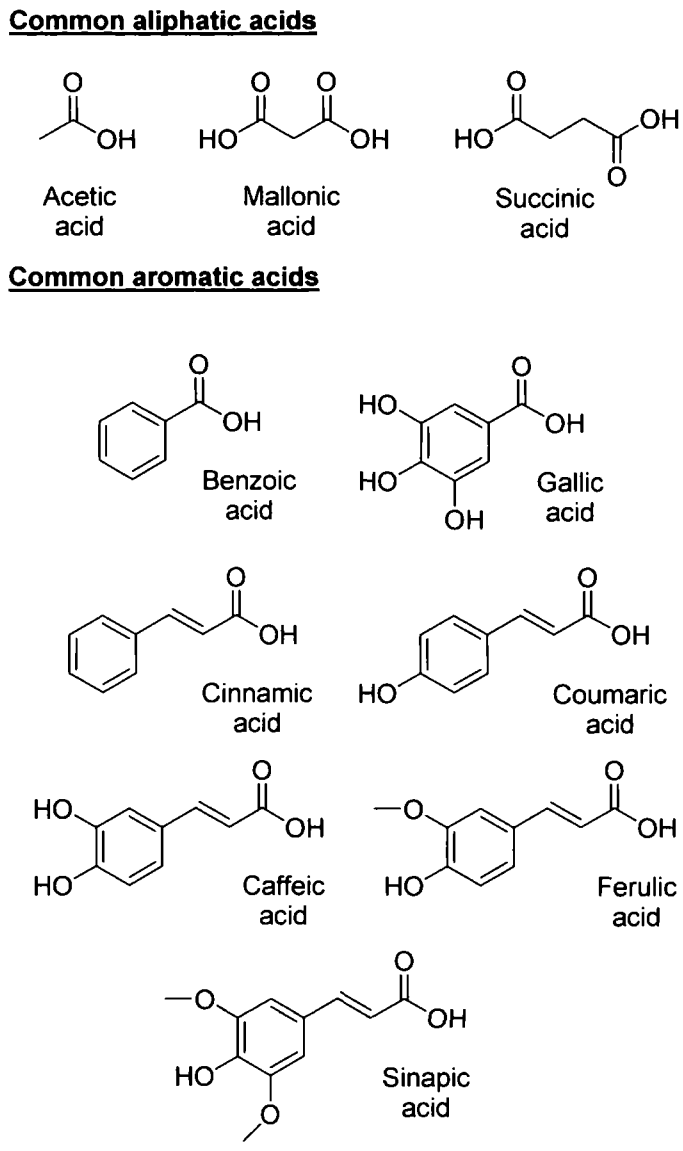


Figure 1.5 Common carboxylic acid substrates for acyltransfer in secondary metabolism

Phenylalanine is the initial substrate for the biosynthesis of most aromatic acids in plants, via numerous pathways. The phenylpropanoid pathway is one example. This pathway is known to produce the functionalised (*E*)-3-phenylprop-2-enoic acids (phenylpropanoids) shown in figure 1.5[23]. These are common substrates for acyltransfer in the flavonoid pathways[24].The aromatic acids generated via these pathways, rarely accumulatve as either the free acids or CoA esters due to the toxicity and lability of the respective metabolites[25]. Although the phenylpropanoids frequently

accumulate as esters of glycosides in many plants[7]. Aliphatic acids are largely sourced from primary metabolic pathways, such as fatty acid metabolism[16].

1.2.2 Coenzyme A ligases

Coenzyme A ligase activity has been well documented in metabolism. Kinetic studies upon a fatty acid CoA ligase from bovine liver, demonstrated coenzyme A ligase activity was based upon a double-displacement mechanism, where ATP and carboxylic acid bind primarily[26]. Competition studies with pyrophosphate and AMP respectively, revealed that pyrophosphate was eliminated from the active site first and prior to coenzyme A binding. Therefore acyl phosphoadenylate ester was concluded to be an enzyme-bound intermediate as previously described (Figure 1.4).

Plant coenzyme A ligase biosynthesis of phenylpropanoyl, benzoyl and malonyl coenzyme A-thioesters has been described in *Arabidopsis thaliana*[15], *Clarkia breweri*[27] and *Rhizobium trifolii*[28] respectively. In particular, 4-coumarate: coenzyme A ligases (4CL) in *Arabidopsis thaliana* have have been characterised with respect to several endogenous phenylpropanoids[15]. There are 4 isoenzymes of 4CL in arabidopsis, At4CL1, At4CL2, At4CL3, and At4CL4, each with different characteristics and predicted roles.

Kinetic data was acquired for each isoenzyme toward endogenous phenylpropanoid substrates. This data was sourced from multiple publications although only one quoted the activity of substrate toward a known amount of purified recombinant protein (At4CL2 toward caffeic acid in *μkats mg^{-1} pure 4CL*)[29]. The specific activities

stated below are therefore based upon comparisons made between this known activity and related data quoted for each enzyme toward acids[15].

Enzyme	*Substrate*	K_M (μM)	k_{cat} (s^{-1})	k_{cat} / K_M ($\mu M^{-1} s^{-1}$)
At4CL1	Cinnamic acid	6230	71	1.1×10^{-2}
	Coumaric acid	38	70	1.8
	Caffeic acid	11	19	1.7
	Ferulic acid	199	37	1.8×10^{-1}
	Sinapic acid	n.c.	-	-
At4CL2	Cinnamic acid	6630	4	6.0×10^{-4}
	Coumaric acid	252	18	7.1×10^{-2}
	Caffeic acid	20	13	6.5×10^{-1}
	Ferulic acid	n.c.	-	-
	Sinapic acid	n.c.	-	-
At4CL3	Cinnamic acid	2070	19	9.2×10^{-3}
	Coumaric acid	23	12	5.2×10^{-1}
	Caffeic acid	374	15	4.0×10^{-2}
	Ferulic acid	166	10	6.0×10^{-2}
	Sinapic acid	n.c.	-	-
At4CL4	Cinnamic acid	n.c.	-	-
	Coumaric acid	432	6	1.4×10^{-2}
	Caffeic acid	186	12	6.5×10^{-2}
	Ferulic acid	26	10	3.8×10^{-1}
	Sinapic acid	20	7	3.5×10^{-1}

Figure 1.6 Kinetic data for *Arabidopsis thaliana* 4CL isoenzymes toward endogenous phenylpropanoid substrates [15,30]. (n.c. = no conversion of substrate)

At4CL1 generally demonstrated a 3-4 fold higher activity when compared with isoenzymes 2, 3 and 4 and had broader tolerance toward phenylpropanoid substrates. At4CL2 and At4CL3 were efficient toward hydroxylated phenylpropanoids, coumaric

and caffeic acid, whilst At4CL4 was selective toward 3- or 3- 5- methoxylated phenylpropanoids.

Further to these kinetic studies, K_M and k_{cat} values for enzyme mutants of At4CL2 toward caffeic and ferulic acids were obtained[29]. Mutation of Met293 to proline allowed At4CL2 to bind ferulic acid. A second mutant improved the K_M of ferulic acid by replacement of Lys320 with alanine. However, no improvement on k_{cat} or K_M was achieved for several At4CL2 mutants toward caffeic acid. They concluded that the rational design of coenzyme A ligase substrate selectivity was possible and further work has been carried out on this topic[31].

1.3 Coenzyme A-dependent acyltransfer in plant secondary metabolism

1.3.1 The flavonoid pathway

1.3.1.1 Flavonol acyl acceptors

Flavonoids are the must abundant and diverse acylated metabolites in plant secondary metabolism and to date variations upon such compounds are known in over half of all angiosperm species[32]. The flavonoid pathway is first committed by a chalcone synthase-catalysed condensation of coumaroyl CoA and three molecules of malonyl CoA, yielding a tetrahydroxychalcone. Many classes of natural products are formed from this point, for example the condensed tannins, the anthocyanins, the flavones, the flavonols and the isoflavonoids[33]. Figure 1.7 shows the biosynthetic routes derived from the flavonoid pathway. However, only a few examples of the huge variety of end products are shown.

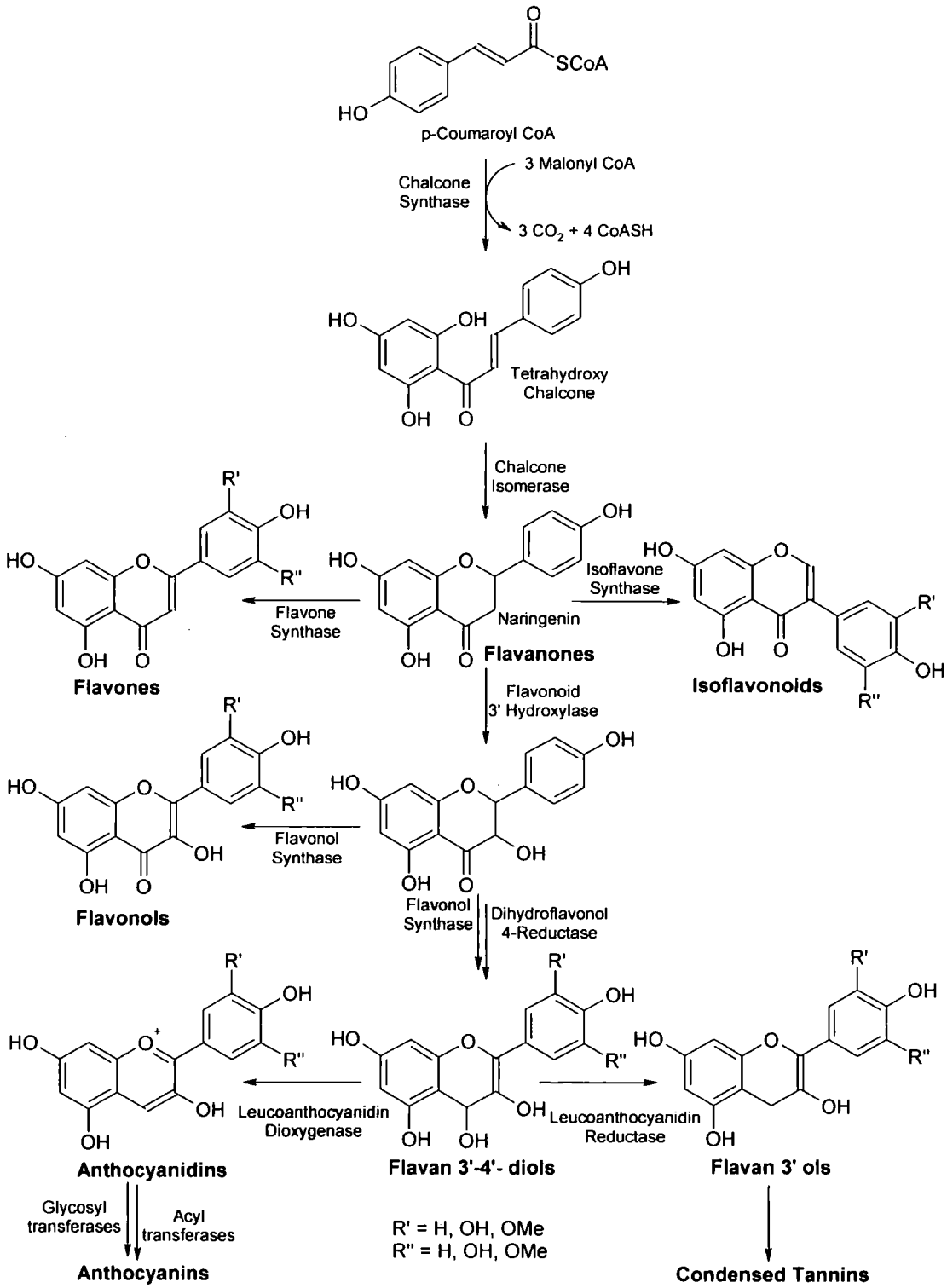


Figure 1.7 The flavonoid pathway generates a large diversity of natural products, many of which are frequently modified via glycosylation and acylation

Each of the basic flavonolic structures may be subject to hydroxylation and/or methoxylation, which allows for a diverse array of compounds. Further modification of the flavonoids is commonly observed. Most flavonoids are glycosylated[34] and many of these are found to be acylated by aliphatic and/or aromatic acids[32].

Each of the major class of flavonoid has generic UV-Vis absorption properties, which are characteristic and are often used in detection/identification[35]. Phenylpropanoids (Figure 1.8 A), flavonols (Figure 1.8 B) and anthocyanins (Figure 1.8 C) all have their own characteristic UV traces[24]. Phenylpropanoids typically absorb in the range 260-315 nm, with one absorption maximum. Flavonoids have characteristic absorbance maximum around 330 nm with the B ring absorbing in the range 260-300 nm. Upon acylation, a phenylpropanoid absorption maximum may also be observed in the same 260-300 nm region as a shoulder. Similar absorption effects are seen in the anthocyanins in the 260-300 nm region, however extended conjugation with the oxycation gives characteristic absorption in the 400-600 nm region.

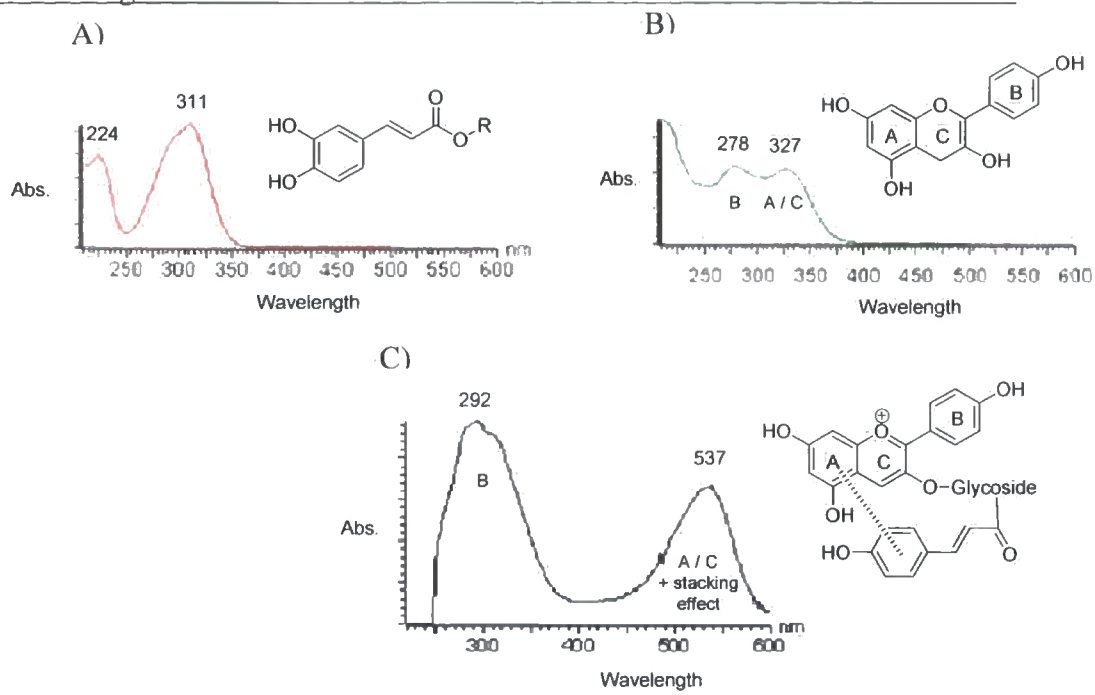


Figure 1.8 UV spectroscopy of A) phenylpropanoid, B) flavonol, and C) anthocyanin core structures.

A large proportion of acylated flavonol glycosides are based on the substitution of kaempferol and quercetin (Figure 1.9). A variety of glycosyl substitutions are seen, although the majority occur at the 3- and / or 7- carbons as indicated in figure 1.9.

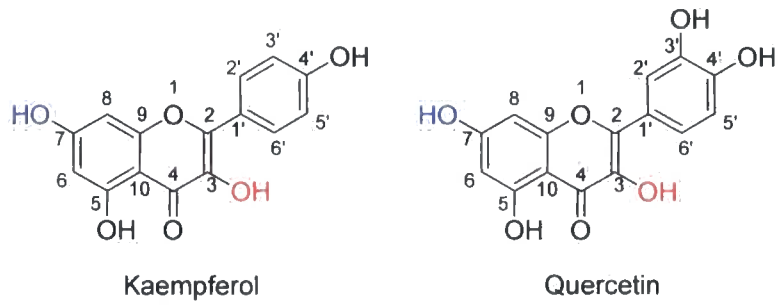


Figure 1.9 The major aglycones of acylated flavonol glycosides

Common aromatic esters of flavonoids include coumarate, ferulate and caffeate, whilst sinapate esters are much rarer. Abe *et. al.*[36] described only the second and third sinapate esters of flavonol glycosides as recent as 1996, from the leaves of *Thevetia peruviana* (Figure 1.10). The metabolites were based on 3-*O*-glycosylated kaempferol and quercetin analogues, with *mono*- and *di*- sinapoyl acylation.

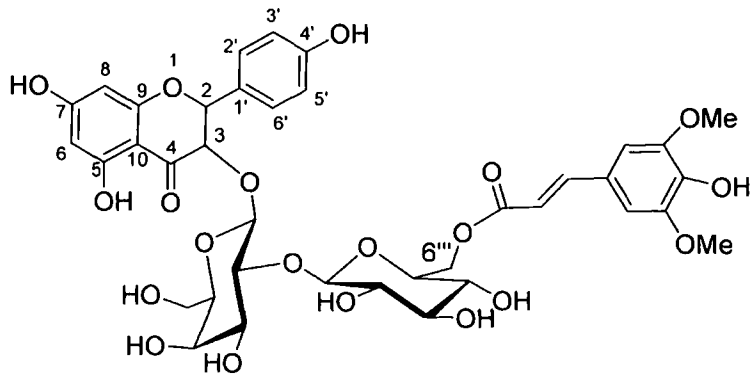
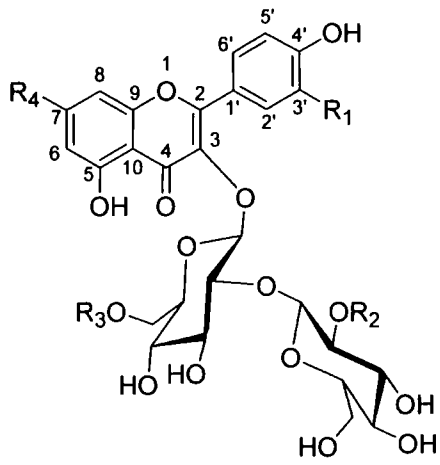


Figure 1.10 Kaempferol 3-(6'''-*O*- sinapoylglucoside)[1-2] galactoside - A rare sinapoylated flavonol glycoside

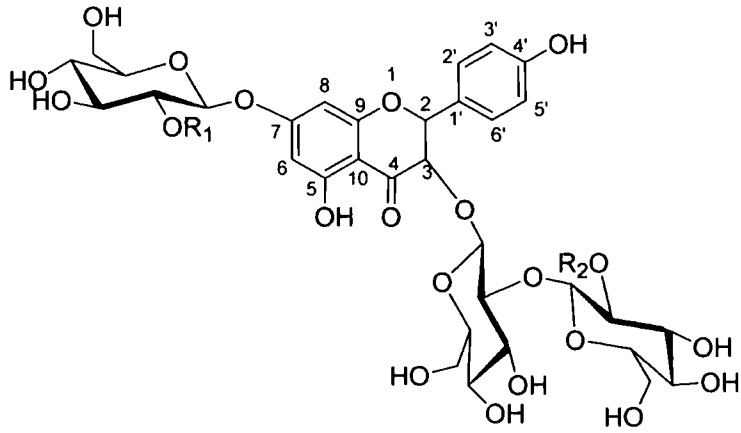
Petunia mitchell is a good example of a plant which accumulates a diverse array of flavonol glycosides. Bloor *et. al.*[37] described both kaempferol and quercetin derivatives of phenylpropanoylated 3-*O*-sophorosides and phenylpropanoylated 3-*O*-sophoroside 7-*O*-glucosides (Figure 1.11). Each of these structural variations is divergent and was exclusively located in the leaves. This suggests that these structures were produced for an evolving or diverse function.



	R1	R2	R3	R4
1	H	caffeoyl	H	OH
2	H	feruloyl	H	OH
3	OH	caffeoyl	H	OH
4	OH	feruloyl	H	OH
5	H	feruloyl	malonyl	OH
6	OH	caffeoyl	malonyl	OH
7	H	Feruloyl	H	Glucose
8	OH	Caffeoyl	H	Glucose

Figure 1.11 The array of acylated flavonol glycosides found in the leaves of *Petunia mitchell*

Cabbage also accumulates a complex mixture of over 20 flavonolic compounds with a range of acylations described[38,39]. A recent comprehensive extraction of cabbage leaves yielded four new kaempferol tetraglucoside compounds (Figure 1.12). The role of flavonoids in cabbage was hypothesised to be as subtle signalling compounds in plant-herbivore interactions and have evolved to be effective against different herbivorous species[32]. The biosynthesis of metabolites with three potential acyl moieties at the same position upon a natural product is not commonly observed. Whether there are three different enzyme activities responsible for each product, or a common enzyme pathway is unknown.



	R1	R2
1	H	caffeoyl
2	H	feruloyl
3	H	sinapoyll
4	glucoside	sinapoyl
5	glucoside	feruloyl
6	glucoside	caffeoyl

Figure 1.12 Varying acylation of *tri-* and *tetra-* glucosides observed in cabbage leaves

Schnitzler *et. al.* recently isolated 2 novel flavonol 3-*O*-(3"- 6"-dicoumaroyl) glucosides in Scots pine seedlings (*Pinus sylvestris*)[40] (Figure 1.13). These two compounds were induced by simulated solar radiation, whilst a non-acylated analogue was not induced. It was proposed that the acylation was favoured by plants under light-induced oxidative stress, to maximise UV absorbing properties.

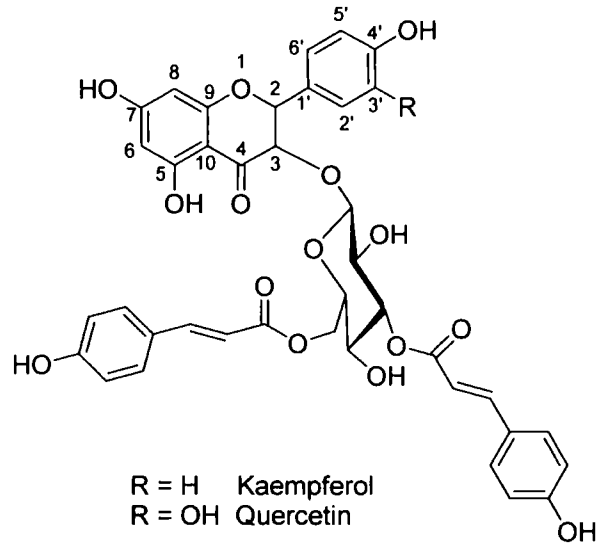


Figure 1.13 Diacylated kaempferol and quercetin 3-*O*- glucosides

The diversity of acylated flavonol glycosides implies that they have multiple functions. In the scope of this review, there are far too many anomalies and structural variations to discuss all possible acylated variants. Summarised in figure 1.14 is an overview of the acylation of the simplest flavonols reported in the recent literature. From these compounds, some generalisations can be assumed. Firstly, the most common acyl acceptors by far, have either a kaempferol or quercetin aglycone. Secondly, 3-*O*- glycosylation is predominant and 6''-*O*- acylation of 3-*O*- glycosides is also common. 7-*O*- glycosylation of flavonols is well known, but saccharides upon this carbon are frequently acylated by aliphatic as opposed to aromatic acylating moieties. Fourthly, the regiospecificity of flavonol-glycoside acylation, is dependent upon the identity of the glycoside.

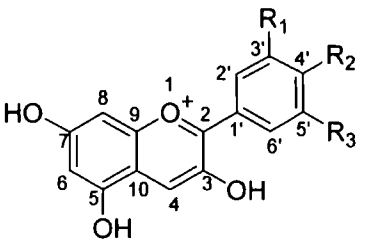
Acyl acceptor	*Acyl donor*	*Acylating position (aglycone/ saccharide)*	*Species*
Kaempferol glucoside	coumaroyl	3 / 6"	*Umbelliferae*
"	caffeoyl	3 / 6"	*Dennstaedtiaceae*
"	caffeoyl	3 / 6"	*Cruciferae*
"	sinapoyl	3 / 6"	*Apocynacea*
"	Di-coumaroyl	3 / 3" and 6"	*Pteridaceae**
"	coumaroyl	7 / 2"	*Liliaceae*
"	feruloyl	7 / 2"	*Liliaceae*
Kaempferol apioside	feruloyl	3 / 5"	*Dennstaedtiaceae*
Kaempferol rhamnoside	Di-coumaroyl	3 / 2" and 4"	*Epacridaceae*
"	Di-coumaroyl	3 / 2" and 4"	*Lauraceae*
Quercetin glucoside	sinapoyl	3 / 6"	*Apocynacea*
"	malonyl	3 / 6"	*Ranunculaceae*
"	feruloyl	3 / 6"	*Ranuncul aceae*
Quercetin gentiobioside	caffeoyl	3 / 6"	*Caprifoliaceae*
Quercetin isophorobioside	caffeoyl	3 / 6"	*Leguminoceae*

Figure 1.14 A general overview of the simplest flavonol glycosides reported from a variety of species [24,32] * *Four feruloyl, sinapoyl and caffeoyl derivatives were also reported for this species*

1.3.1.2 Anthocyanin acyl acceptors

Anthocyanins are biosynthesised from flavonols to form a similar metabolite, but with a cyclic oxonium ion/extended chromophore (Figure 1.15)[33]. Anthocyanins are also

commonly glycosylated and/or acylated or polyacylated. Glycosylation of the anthocyanins is similar in complexity and regioselectiivity to the flavonols. The most common saccharide is glucose, but galactose, sophorose, xylosylglucose, rutinose, xylosylgalactose and arabinose are also frequently observed[24]. The most common acylations are upon 3-*O*-glycosylated anthocyanins and include coumaroyl, caffeoyl, feruloyl, sinapoyl and malonyl moieties.



Anthocyanin	*R1*	*R2*	*R3*
Cyanidin	OH	OH	H
Delphinidin	OH	OH	OH
Malvadin	OMe	OH	OMe
Petunidin	OMe	OH	OH
Peonidin	OMe	OH	H
Pelargonidin	H	OH	H

Figure 1.15 Commonly acylated Anthocyanin aglycones

The most obvious function of the anthocyanins is pigmentation in flowers and fruits owing to the extended conjugation created by formation of the oxonium ion[41]. Anthocyanins, along with other flavonoids, may also be important in the resistance of plants to insects as anti-feedants[24]. For example, acylated cyanidin 3-*O*-glucosides

were shown to protect cotton leaves against the tobacco budworm[42]. Acylation is frequently essential for the activity of anthocyanins in these roles.

Delphinidins are thought to be the major anthocyanidin base of blue flower colour. Bloor *et. al.* initiated a search for stable pigments with vibrant blue colouration for the dye industry[34]. For the first time, delphinidin 3- 7- 3'- 5'-*O*-tetraglucosides were reported which were coumaroylated upon all four glycosides, as shown in figure 1.16.

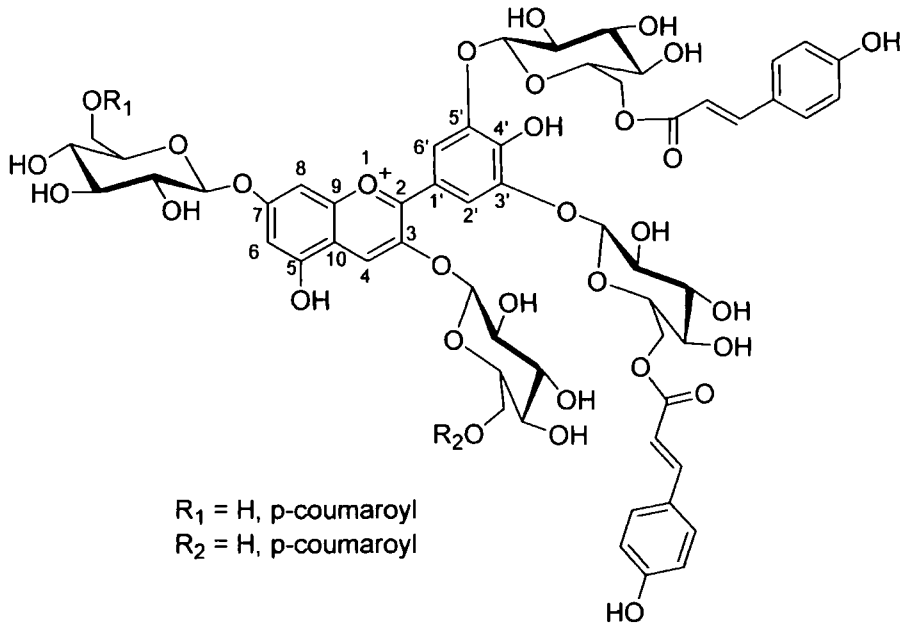


Figure 1.16 Polyacylated delphinidin

Along with the array of acylated flavonol glycosides reported in the previous section, *Petunia mitchell* was also shown to accumulate acylated anthocyanins in the petals. Bloor *et. al.* observed the accumulation of four acylated anthocyanins by HPLC-MS[43], including the major compound petunidin 3-*O*-(coumaroylrutinoside) 5-*O*-glucoside. Interestingly, they successfully transfected a foreign gene, responsible for regulation of anthocyanin biosynthesis, into *Petunia mitchell*. This transfection caused petunidin 3-*O*-(coumaroylrutinoside) 5-*O*-glucoside to accumulate in the *leaf* tissue where previously it was only found in the petals. The coumaroylated petunidin structure was

preserved in the presence of caffeoylated and ferruloylated kaempferol and quercetin glycosides. This demonstrated high specificity of anthocyanin acyltransferases toward acyl donor and acceptor substrates.

Arabidopsis thaliana represents an ideal organism for biochemical characterisation due to the availability of its genome sequence, allowing convenient identification of proteins and their respective genes. In arabidopsis, the major flavonoid end product is cyanidin 3-*O*-[2-*O*-(2-*O*-(sinapoyl)-xylosyl) 6-*O*-(4-*O*-(glucosyl) coumaroyl-glucoside] 5-*O*- 6-*O*-(malonyl)glucoside], as shown in figure 1.17[35]. The aromatic acyltransferases responsible for the incorporation of the coumaroyl and sinapoyl acyl moieties into the parent anthocyanin are as yet uncharacterised, although the malonyl transferase which effects malonylation of 5-*O*-glucose has recently been identified[44].

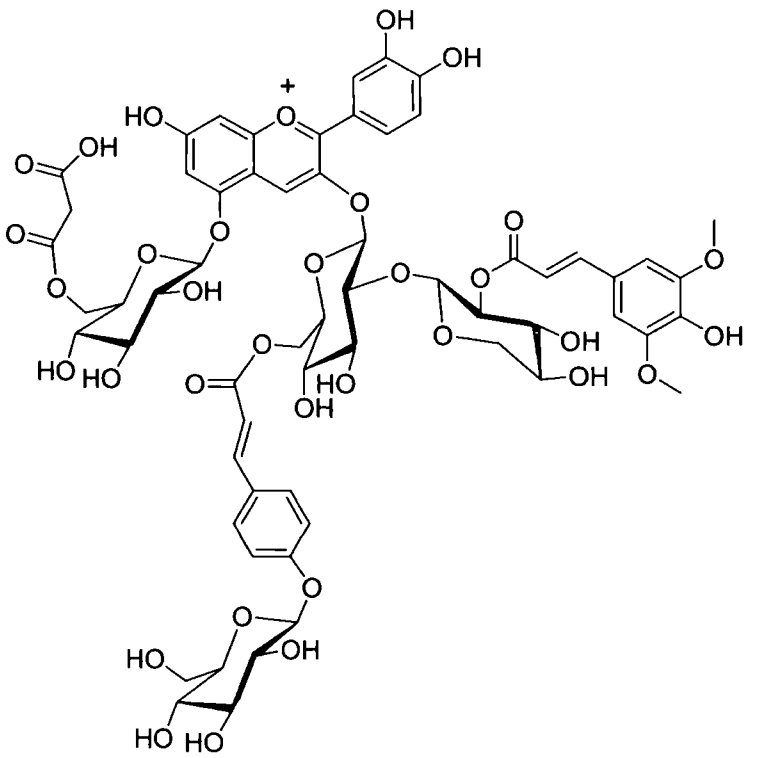


Figure 1.17 The major anthocyanin found in *Arabidopsis thaliana*

1.3.1.3 Regulation of flavonoid biosynthesis

Flavonoids have also been reported to have antibacterial and antifungal properties in plants[45]. Application of the fungus *Botrytis cinerea* to grape vine (*Vitis vinifera*) resulted in up-regulation of phenylpropanoid pathway enzymes[46]. However, fungal treatment of cell cultures of carrot (*Daucus carrota*) failed to increase anthocyanin content despite a significant increase in the activity of the key flavonoid biosynthetic enzymes phenylalanine ammonia lyase and 4-coumarate coenzyme A ligase (Figure 1.18)[47].

UV irradiation is another method that has been used to artificially regulate flavonoid biosyntheses in various plant systems[48]. Li *et. al.* used *Arabidopsis thaliana* flavonoid mutants, deficient in chalcone synthase and chalcone isomerase, to portray the regulation of flavonoids with UV treatment[49]. The results clearly showed accumulation of UV absorbing compounds in wild-type plants upon UV treatment, whereas the mutants were completely lacking the same compounds. Although up-regulation of existing metabolites was achieved in the wild-type plants, no new compounds accumulated and UV absorbing compounds present normally were simply proportionally increased. Therefore flavonoid biosynthetic enzymes: PAL (phenyl ammonia lyase), 4CL (4- coumarate CoA ligase), and CHS (chalcone synthase); can be regulated by stress, including UV-induced oxidative damage and microbial infection. Whereas latter phase biosynthetic enzymes: FVS (flavone synthase), FLS (flavonol synthase), LAD (leucoanthocyanidin dioxygenase) and AATs (aromatic anthocyanin acyltransferases); appear to be more selective in their response (Figure 1.18).

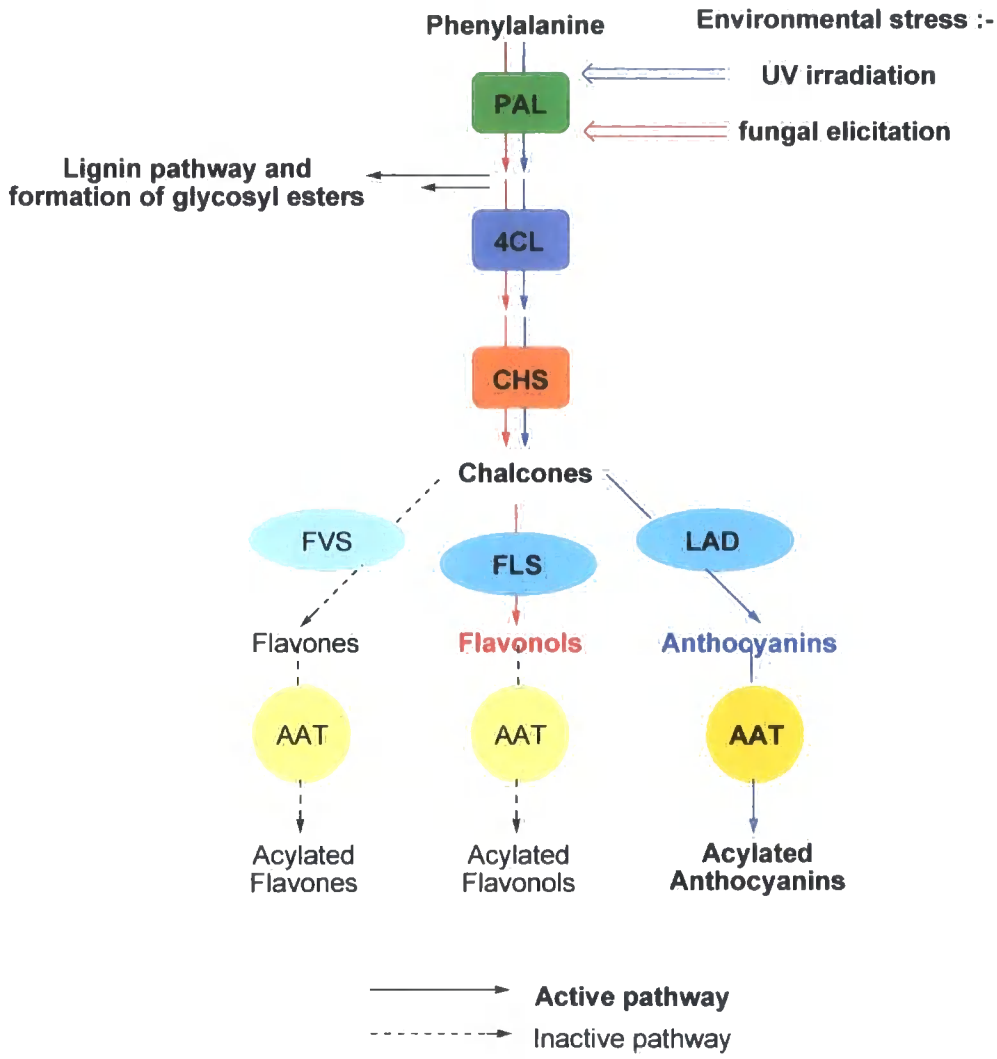


Figure 1.18 A schematic of flavonoid biosynthesis depicting metabolic responses of grape and Arabidopsis plants toward 'stress inducing' environmental conditions.

A less utilised approach toward upregulation of flavonoids in plants has been to feed metabolic precursors to a plant. The intention being to create a 'flood effect' whereby all avenues of related pathways are forced into production of terminal metabolites in an attempt to 'soak up' the excess of substrates. To this effect, Harborne *et. al.* fed the petioles of leaves with aqueous solutions of phenylpropanoid compounds[50]. In his findings, he reported several 1-*O*-acylated saccharides as the major accumulated

products, speculating that they could be activated precursors to more complex metabolites.

Utilising a different methodology, Dougall *et. al.* successfully perturbed the accumulation of acylated anthocyanin natural products in cell cultures of wild carrot (*Daucus carota*)[51]. This was achieved by addition of carboxylic acids direct to cell culture media. Whether the acyl donor was endogenous to the species or not, the fed metabolites were observed in the accumulated acylated anthocyanin (Figure 1.19). An array of endogenous and exogenous phenylpropanoid precursors were ably incorporated into anthocyanin acylation. However, the enzymes responsible for chalcone formation could not be shown to overproduce flavonoids in the presence of excess phenylpropanoid substrate.

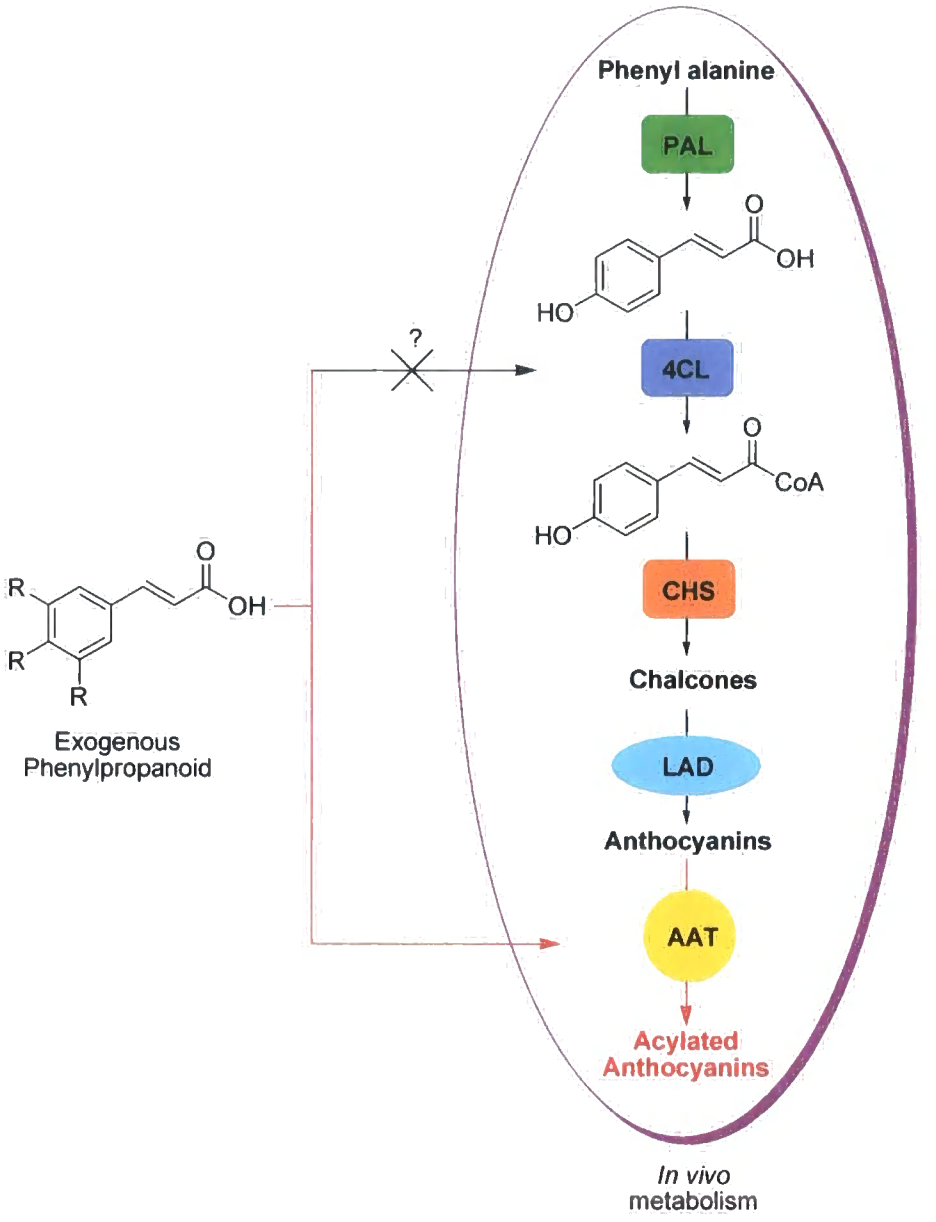


Figure 1.19 Incorporation of phenylpropanoid acylation precursors into flavonoid metabolism in cell cultures of wild carrot (*Daucus carota*)

Thus manipulation and regulation of the phenylpropanoid and flavonoid pathways in plants has been well documented and many of the results have hinted towards the endogenous purpose of the terminal products of these pathways. These studies have demonstrated a number of approaches toward attaining plant material, which is actively expressing flavonoid biosynthetic proteins such as the acyltransferases.

However, none of the described approaches were able to upregulate the whole pathway.

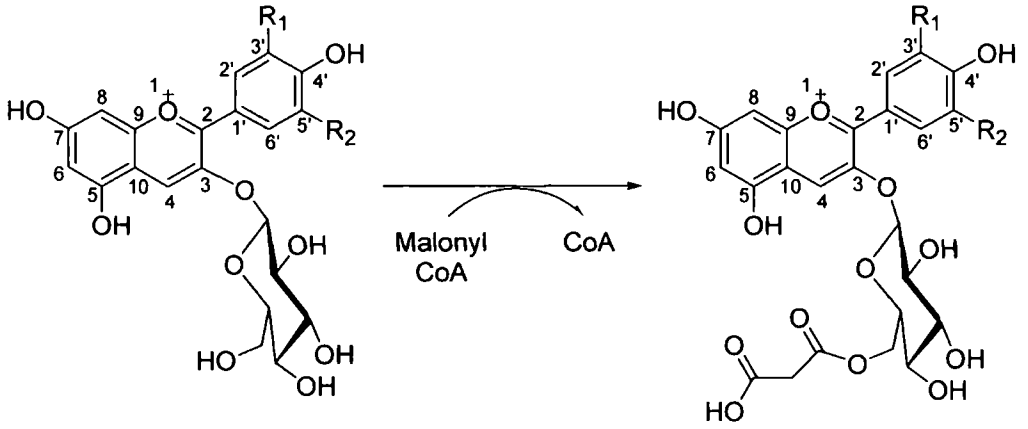
1.3.2 Coenzyme A-dependent acyltransferases isolated from the flavonoid pathway

1.3.2.1 Aliphatic acyltransferases

The commercial availability of malonyl coenzyme A has facilitated the purification of malonyl CoA: anthocyanin transferases[52]. Malonylation occurs as the penultimate step of anthocyanin biosynthesis, conferring solubility, stability and pH balance to the metabolite[18], therefore the non-acylated anthocyanin malonyl acceptor can be identified and readily isolated from plant tissue sources in relatively good yield, due to the lability of the malonic ester bond.

Suzuki *et. al.* purified the first example of a 3-*O*- glucoside specific aliphatic transferase, malonyl CoA: anthocyanidin 3-*O*- glucoside 6''-*O*- malonyltransferase (Dm3MAT1) from Dahlia flowers, using an HPLC assay based on malonyl CoA transfer to pelargonidin 3-*O*- glucoside to monitor the activity of the enzyme. The specificity of the enzyme toward acyl donors was determined, with only malonyl CoA and close homologues such as succinyl CoA being used as substrates (Figure 1.20). Of particular note, caffeoyl CoA was not transferred onto pelargonidin by the action of Dm3MAT1. Glycosyl moieties upon the aglycone structure were also found to affect acyltransferase activity toward acyl acceptors. Only metabolites bearing the exact combination and positioning of 3-*O*- glycoside residues were MAT substrates. In contrast conservation of the aglycone structure itself was found to be less critical.

An array of anthocyanins, and the flavonol quercetin, all conferred reasonable levels of transferase activity.



Substrate	*R1*	*R2*	*Relative activity %*
Acyl donor			
Malonyl CoA	H	H	100
Acetyl CoA	H	H	<0.1
Methylmalonyl CoA	H	H	11
Succinyl CoA	H	H	16
Caffeoyl CoA	H	H	<0.1
Acyl acceptor			
Pelargonidin 3- O- glucoside	H	H	100
Cyanidin 3- O- glucoside	OH	H	111
Delphinidin 3- O- glucoside	OH	OH	117
Quercetin 3- O- glucoside	(See section 1.3)		15
Cyanidin 3- O- 6''- O- malonylglucoside	OH	H	0.25
Pelargonidin 3-, 5-, O- glucoside	H	H	1.9
Cyanidin 3- O- 6''- O- malonylglucoside, 5- O- glucoside	OH	H	<0.1

Figure 1.20 Activity of Dm3MAT1 toward various acyl acceptors and donors. Acyl donors are determined versus pelargonidin 3-*O*- glucoside and acyl acceptors versus malonyl CoA

Studies based upon anthocyanin 5-*O*- malonyl transferases in *Salvia splendins (SsMAT1),* subsequently explored the catalytic properties of anthocyanin transferases[53]. Application of substrates in assays of SsMAT1 (Figure 1.31) revealed coenzyme A was a competitive inhibitor with respect to malonyl CoA. The deduction made from this result was the ordered binding of the substrate malonyl CoA and eviction of coenzyme A from the active site is the last event in the catalytic cycle. Acyl donor substrates are therefore bound furthest into the enzyme core, or possibly 'activate' sequential binding of the acyl acceptor.

Utilising antisera to SsMAT1, Suzuki *et. al.* immunoprecipitated studies a further MAT, SsMAT2, which carried out a successive malonylation shown in figure 1.21.

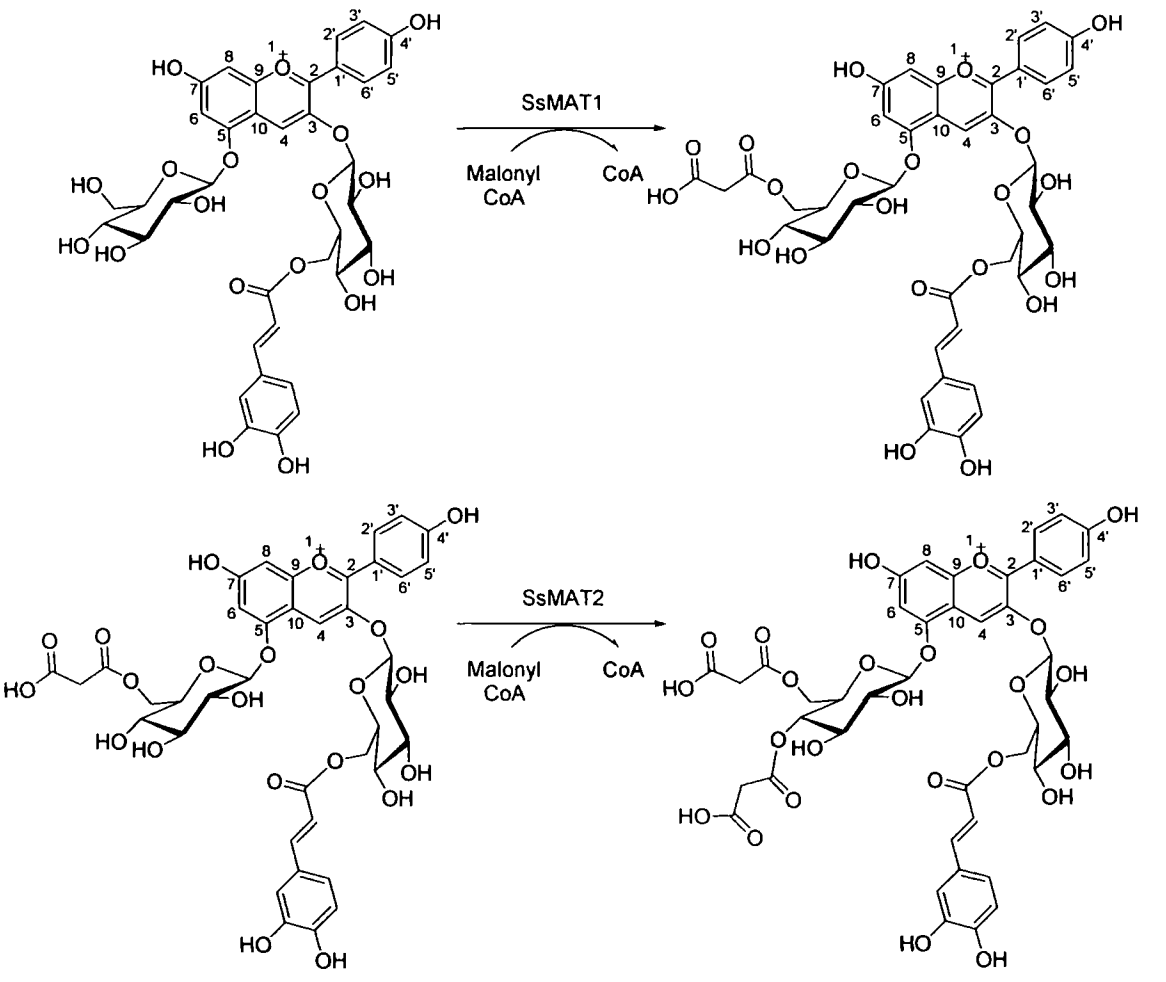


Figure 1.21 Malonyl transferases present in Salvia splendins. The primary malonylation of the 5- *O*-glucoside must be present in order for the second malonylation to take place

The sequence of SsMAT2 was aligned with those of Dm3MAT1 and Ss5MAT1. However, the alignment showed unexpected results with Dm3MAT1 and Ss5MAT1 bearing close homology, but SsMAT2 being only distantly related. It was proposed that the difference arose from the site of acylation upon the aglycone. Ss5MAT1 and Dm3MAT1 acylate at 6''- OH of the saccharide, whereas SsMat2 transfers malonyl onto the 4''- OH of a saccharide.

1.3.2.2 Aromatic acyltransferases

As reflected in the comparative number of aliphatic anthocyanin acyltransferases versus aromatic anthocyanin acyltransferases identified in plant secondary metabolism, developing activity-based purification strategies of the latter has proved more difficult. As a result, only three aromatic acyltransferases associated with flavonoid metabolism have been isolated, despite considerable interest in other aspects of the pathways.

Fujiwara *et. al.* (1997) purified to homogeneity caffeoyl CoA: delphinidin 3- 5- *O*-diglucoside 5-*O*- glucoside 6''-*O*- caffeoyltransferase (Gent5AT) from blue petals of *Gentiana triflora* with a single protein with a mass of 52 kDa being identified. Gent5AT catalysed the transfer of caffeic acid onto delphinidin 3- 5- *O*- diglucoside to form Gentiodelphinidin, the major anthocyanin accumulated in Gentian flowers (Figure 1.32)[54]. Owing to the low levels of acylated delphinidin 3-, 5- *O*- diglucoside compared to its 3'-*O*- acylated / glucosylated counterpart, Fujiwara was able to predict the sequence of biosynthesis of gentiodelphinidin (Figure 1.22).

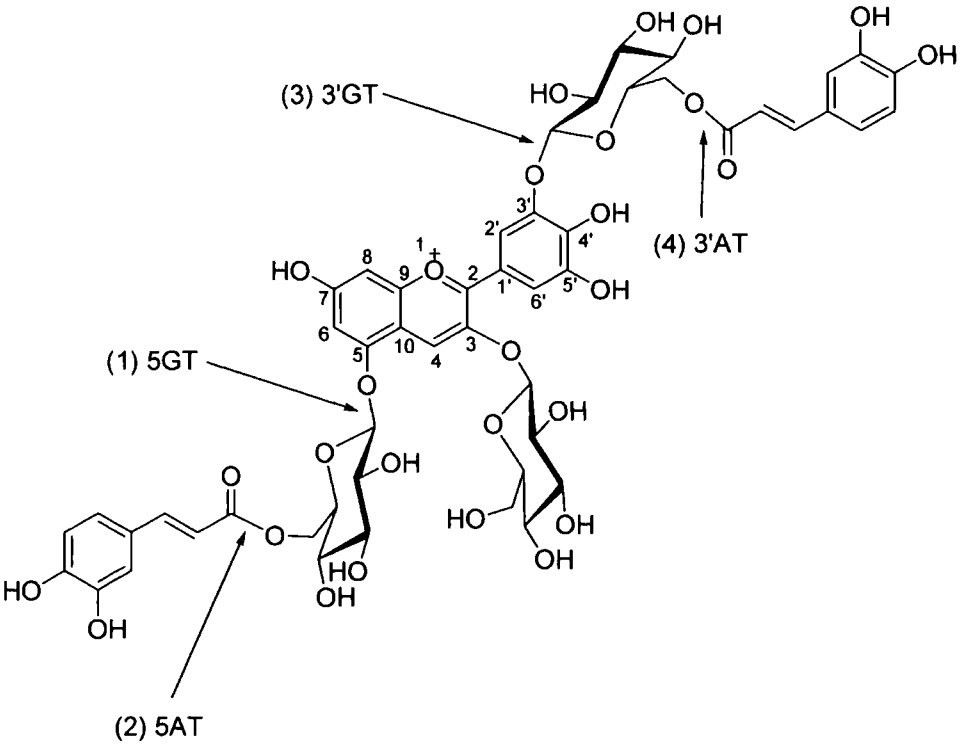


Figure 1.22 Order of biosynthesis of gentiodelphinidin

Utilizing synthesised caffeoyl and coumaroyl CoA donors and delphinidin and cyanidin 3- 5- *O*- diglucoside acceptors isolated from extracts of *Verbena hybrida*, acyltransferase activity toward all four of the identified enzyme activities was determined. Isolation of a range of other acyl acceptors from various plant sources, which were not present in *Gentiana triflora*, including 5-*O*- monoglucosides and 3-, 5- 3'- *O*- triglucosides, demonstrated the importance of conservation of 3- 5- *O*- glucosylation of delphinidin or cyanidin in Gent5AT activity. In addition a range of metallic cofactors (1 mM) were applied to the acyltransferase assay with little effect except manganese [55], which conveyed a 151 % increase in activity of Gent5AT. However, EDTA had minimal detrimental effect on activity, which called into question whether Mn^{2+} was essential.

In a related study, Fujiwara *et. al.*[56] isolated a caffeoyl CoA: 3-*O*- glucoside delphinidin 6''-*O*- caffeoyltransferase (Per3AT), which carries out the transfer of caffeoyl to a 3-*O*- glucoside of delphinidin in *Perilla frutescens* (Figure 1.23).

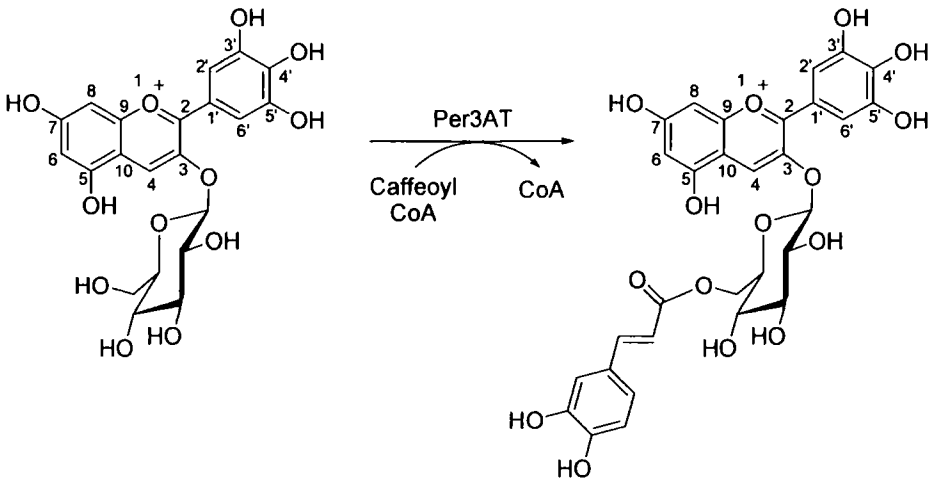


Figure 1.23 Per3AT activity in *Perilla frutescens*

A similar set of anthocyanin acceptors was screened against Per3AT activity with respect to substrate preference. Surprisingly, some activity was observed toward 3- 5- *O*- glucosylated delphinidin. A positive effect of manganese [55] was also noted of Per3AT transferase activity, whilst a series of other metallic cofactors all had detrimental effects, particularly zinc [55] which reduced activity to 61 %. Interestingly diethylpyrocarbonate and *N*-ethylmaleimide had inhibitory effects on Per3AT, suggesting important roles of histidine and cysteine amino acid residues, although no other inhibitors were screened.

1.3.3 Coenzyme A-dependent acyltransferases isolated from other secondary metabolic pathways

1.3.3.1 Vinorine synthase

A considerable amount of research has been invested in the acetyltransferase Vinorine synthase, present in the Indian medicinal plant *Rauvolfia serpentine* and represents the most thoroughly characterised member of the coenzyme A-dependant BAHD acyltransferase family. Vinorine synthase catalyses the transfer of an acetyl group to an aldehyde in a concerted ring - closure mechanism (Figure 1.24). The acetyl group enhances the stability of vinorine, preventing spontaneous epimerisation at 16- *C* causing breakdown of the product to vellosamine, a product not incorporated into the same pathway. Similarly to the interest in the morphinal and taxol acetyltransferases, vinorine synthase was initially isolated in order to provide a route to the biosynthesis of the pharmaceutically important natural product, the antiarythmic ajmaline.

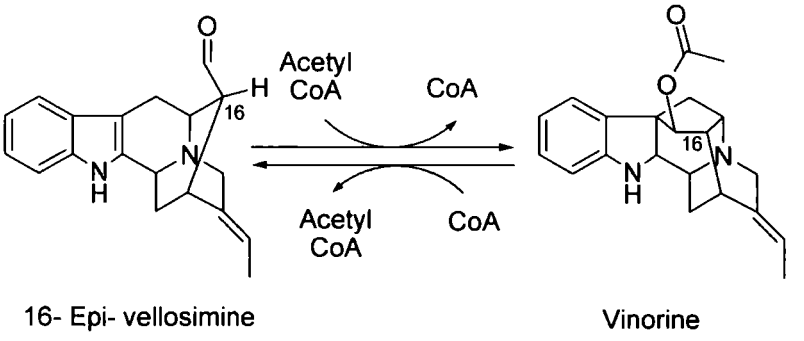


Figure 1.24 Acetyl transfer to 16- epi- vellosimine traps the ring-closed metabolite vinorine. *In vitro* enzymatic activity is found to be reversible although in the absence of enzyme vinorine is stable

Gerasimenko *et. al.* (2004) initially identified vinorine synthase activity in hybrid cell suspension cultures in *Rauvolfia serpentina* by feeding the cell cultures with deacetylvinorine[57]. Having identified the acetyltransfer reaction, they subsequently

developed an HPLC-based assay to isolate and characterise the acetyltransferase activity. Proteolytic cleavage of the purified protein with the endoproteinase LysC delivered four peptide fragments, which were subsequently sequenced. Alignment of the four peptides showed high homologies to other acetyltransferases (30-70 %), such as deacetylvindoline acetyltransferase[58] and salutiridinol acetyltransferase[59].

Bayer *et. al.* observed that the 160HxxxD164 and a 360DFGWG364 residues in the translated sequence of vinorine synthase, were highly conserved in all BAHD acyltransferases, in addition to a series of serine and cysteine residues which were semi-conserved[60]. A series of classic inhibitors were screened against vinorine synthase in order to ascertain their importance (Figure 1.25).

Inhibitor	**Type of inhibitor**	**Relative inhibition (%)**
10 mM AEBSF	*Serine specific*	*100*
0.023 mM E-64	*Cysteine specific*	*50*
0.12 mM TLCK	*Serine / Cysteine*	*50*
0.2 mM TPCK	*Serine / Cysteine*	*100*
1mM PMSF	*Serine / Cysteine*	*58*
0.2 mM Hg^{2+}	*SH group modifier*	*100*
10 mM DEPC	*Unselective Histidine*	*100*

Figure 1.25 Effects of various catalytic residue targeting inhibitors on vinorine synthase activity

The residues serine, histidine and, to a lesser extent, cysteine were all targeted by their respective inhibitors. In particular the selective serine inhibitor 4-(2-Aminoethyl) benzenesulphonyl fluoride gave complete inhibition as well as the Ser-Cys modifying agent *N*-tosyl L-phenylalanine chloromethylketone. Neither selective cysteine inhibitor *N*-(*N*-(L-3- *trans*- carboxirane 2-carbonyl)-L-leucyl)-agmantine nor the series of Ser-Cys modifiers caused total loss of activity. Whereas the His targeting diethylpyrocarbonate inhibitor caused total inhibition of vinorine synthase.

Based on these findings, Bayer *et. al.* concluded that histidine, serine and cysteine are likely to have some catalytic role, in addition to the highly conserved aspartate residues (D164 and D362). In order to further enhance their predictions, a series of site-specific alanine mutations were carried out upon the vinorine synthase residues. The totally conserved ^{160}HxxxD^{164} motif was initially targeted and muteins His160Ala and Asp164Ala were produced. Both of these mutations caused a total loss of activity, implying that these residues are part of the catalytic process. A series of cysteine and serine residues were also replaced by alanine, causing significant loss of relative activity - less than 25 % - at Ser29Ala, Ser243Ala and Cys149Ala indicating a potential role in catalysis. Interestingly, all of these mutations retained low K_M values for substrates, indicating a tight binding of the mutant enzymes to their substrate, indicating a product / byproduct release role of nucleophillic residues.

Ma *et. al.* crystallised vinorine synthase in order to gain further molecular insight into the activity of BAHD acyltransferases[61]. Vinorine synthase is comprised of two domains, connected by a loop (residues 201 – 213), which creates a solvent channel between the two which has been shown to bind both CoA and acetyl CoA from one

face and the acceptor substrate from an alternative face. The catalytic residues 160HxxxD164, present in domain 1, are found in the centre of the solvent channel. However, the Asp164 residue is directed away from the catalytic site, which implies that it does not form a catalytic dyad with His160 despite the total loss of activity upon mutation. Asp164 is alternatively involved in the formation of a salt bridge with the conserved Arg279, suggesting the importance of a highly orientated active site. Based on these results, Ma *et. al.* proposed a catalytic mechanism of vinorine synthase (Figure 1.26).

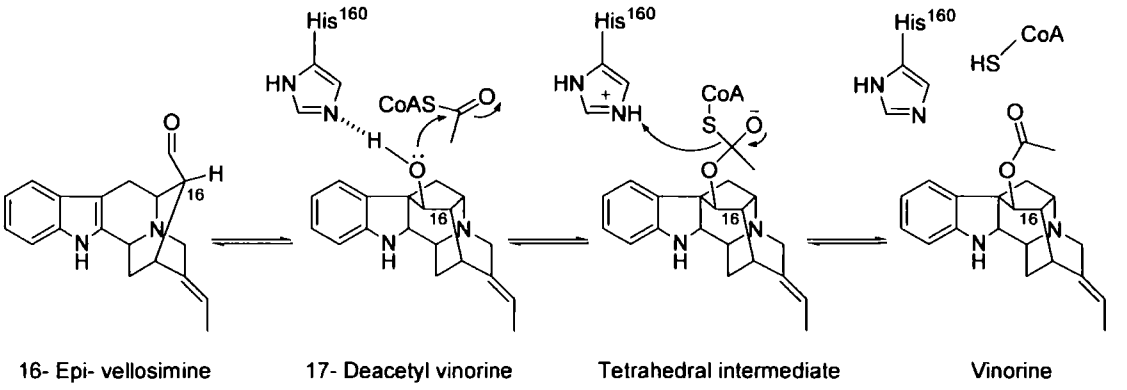


Figure 1.26 A proposed catalytic mechanism for Vinorine synthase

The highly conserved DFGWG motif present in all coenzyme A-dependent transferases identified to date appears to have no direct role in the catalytic binding of coenzyme A in the vinorine synthase crystal structure, the site being remote to the catalytic centre (Figure 1.27). Instead these residues are believed to play a structural role, with the sequence being located at a tight turn between β-11 and β-12. Hydrogen bonding occurs between Asp362, Try365 and Gly366 therefore aspartate seems to be important in maintaining the turn. To this effect, mutation of the aspartate to alanine results in a 65 % loss of activity in vinorine synthase activity[62]. The conservation of

this motif throughout all BAHD transferases suggests its role is related to a common denominator and is involved in selective binding of coenzyme A.

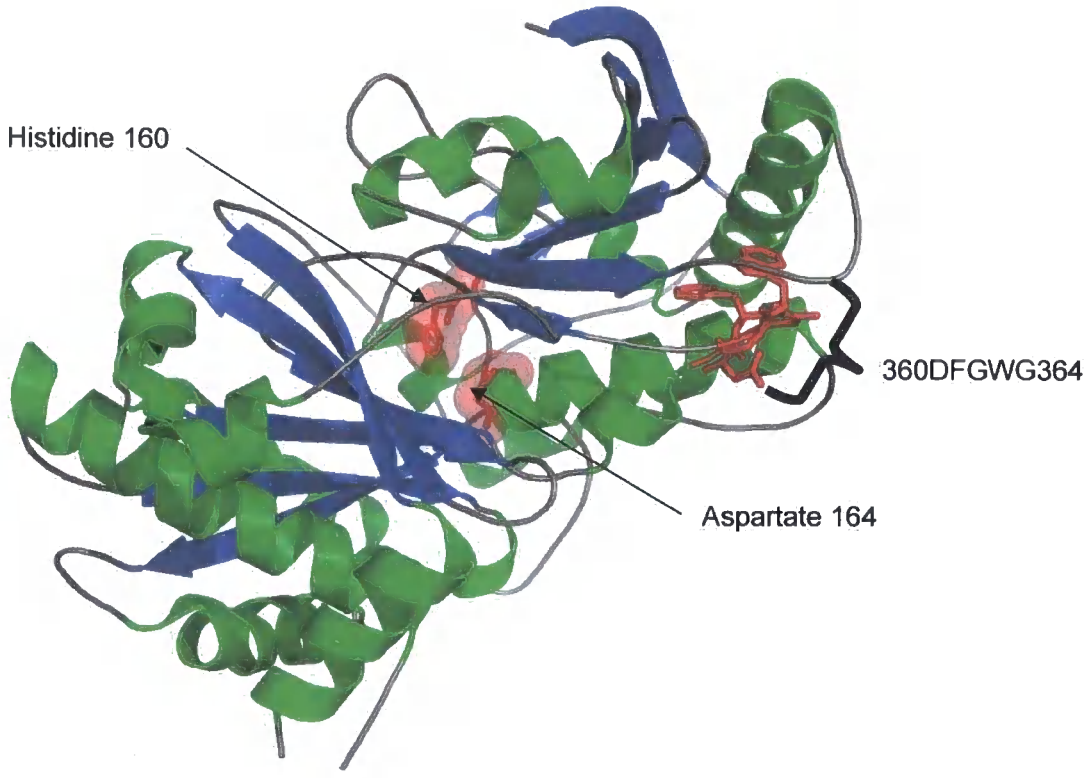


Figure 1.27 A crystal structure of vinorine synthase. The amino acid residues Histidine160 and aspartate 164 are shown, although aspartate is not proximate to histidine. Additionally the conserved DFGWG sequence is removed from the catalytic centre

Ma *et. al.* proposed that several other residues have contact with the catalytic centre, although there is less conservation of these residues across the BAHD spectrum. The large range of acyl acceptors and acyl donors, aside from the common CoA factor, all with different chemical natures, could conceivably cause the opposite effect of disarray of location of active residues in transferases to cope with differing catalytic loci, sterics of substrate and relative activation energy of reaction.

1.3.3.2 Taxol Biosynthesis

The natural product paclitaxel (Taxol) can be isolated in small quantities from pacific yew species *(Taxus brevifolia)* and has proved to be an important drug for the treatment of cancer[63]. In mitotic cells, taxol promotes the binding of tubulin hetero-dimers, stabilises microtubule assembly, disrupts cellular division and ultimately causes apoptosis[64]. Commercially, 10- deacetyl baccatin III (Figure 1.28) is isolated from natural sources (European yew) and synthetic approaches toward paclitaxel are undertaken, including 10-*O*-acetylation, use of a β- lactam synthon to incorporate a 13-*O*-phenylpropanoyl side chain and further benzoylation of the phenylpropanoid amine following deprotection[65]. However, new routes toward production of acylated Taxol are continually being explored[66]. This has included successful biosynthesis in cell cultures, derived from the bark of Yew trees[67]. To enhance the productivity of this process a series of feeding experiments with elicitors, such as methyl jasmonate, and the taxol pathway precursors phenylalanine and benzoic acid[68]. Phenylalanine and benzoic acid were found to be effective in increasing Taxol production in Taxus cell cultures.

Within the biosynthesis of taxol, five of twelve enzymatic reactions are catalysed by an acyltransferase. Recently, all of these steps have been characterised and the respective transferases cloned[64,69-72]. Acylation of taxol initiates with 2-*O*-benzoylation of 2- debenzoyltaxane, the final intermediate in the synthesis of the core diterpenoid structure. From this point, 4 distinct acylation steps are carried out in succession to yield the active product taxol (Figure 1.28).

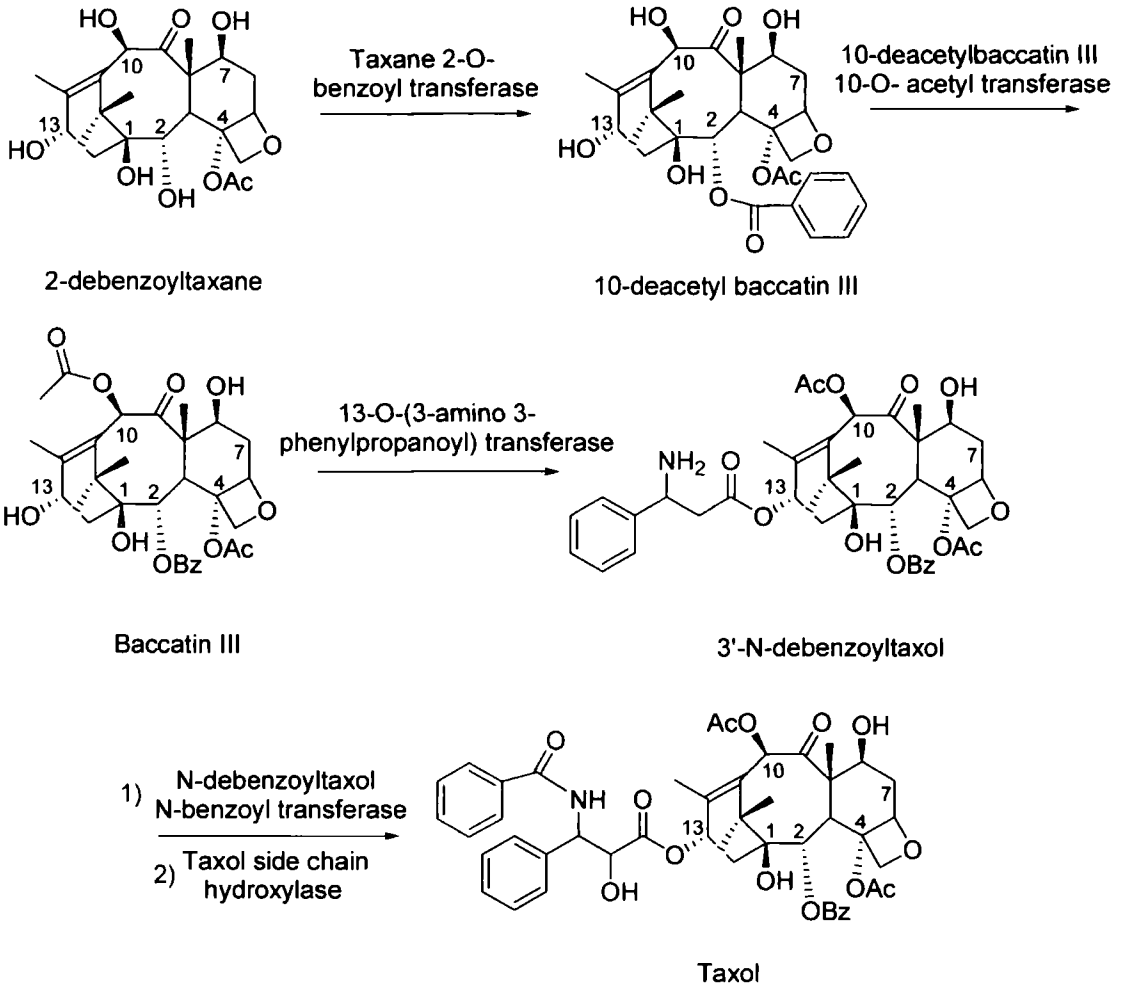


Figure 1.28 Taxol biosynthesis

The order of biosynthesis concerning 13-*O*-phenylpropanoyl side chain and its modifications in the latter stages was determined using ring per-deuterated amino acids. The order shown in figure 1.28 was concluded when studies showed per-deuterated phenylalanine and per-deuterated *N*-benzoyl phenylalanine moieties were not incorporated into the pathway, whereas per-deuterated 3- aminocinnamic acid was successfully incorporated. Essentially, this was the step allowing the crucial 13-*O*-phenylpropanoyl transferase to be isolated, as the known substrates could now be screened against protein extracts[73].

Importance of the acylation of taxol, was studied with respect to its ability to promote tubulin binding in mitotic cells. It was shown that the C-4 oxetane bridge, the 2-*O*-benzoyl group and the benzoylated 13-*O*-phenylpropanoid side chain were all vital for the molecule's pharmaceutical activity[69].

1.3.3.3 Morphine biosynthesis

The Opium poppy *Papaver somniferun* produces some of the world's most heavily used pharmaceuticals, the most desirable being the narcotic analgesic morphine, but also codeine is produced as a close morphine homologue. To date, numerous total syntheses of morphine have been reported[74], although none surpass the commercial viability of extraction of the pharmacon direct from the plant. This fact has driven interest toward the biosynthesis of morphine. Biosynthetic methods could also allow facile inter-conversion of extracted morphine intermediates to the desired pharmacon. Zenk *et. al.*[75] originally elucidated the majority of the initial pathway (Figure 1.28) and more recently T. Kutchan *et. al.*[59] characterised the four penultimate enzymatic steps, including a transfer of acetyl to the 7-*O*- of salaturidinol. In this case the transfer of an acyl group promotes nucleophillic attack of 4-OH, whereby enhanced leaving group ability of the acetyl favours the reaction (Figure 1.29).

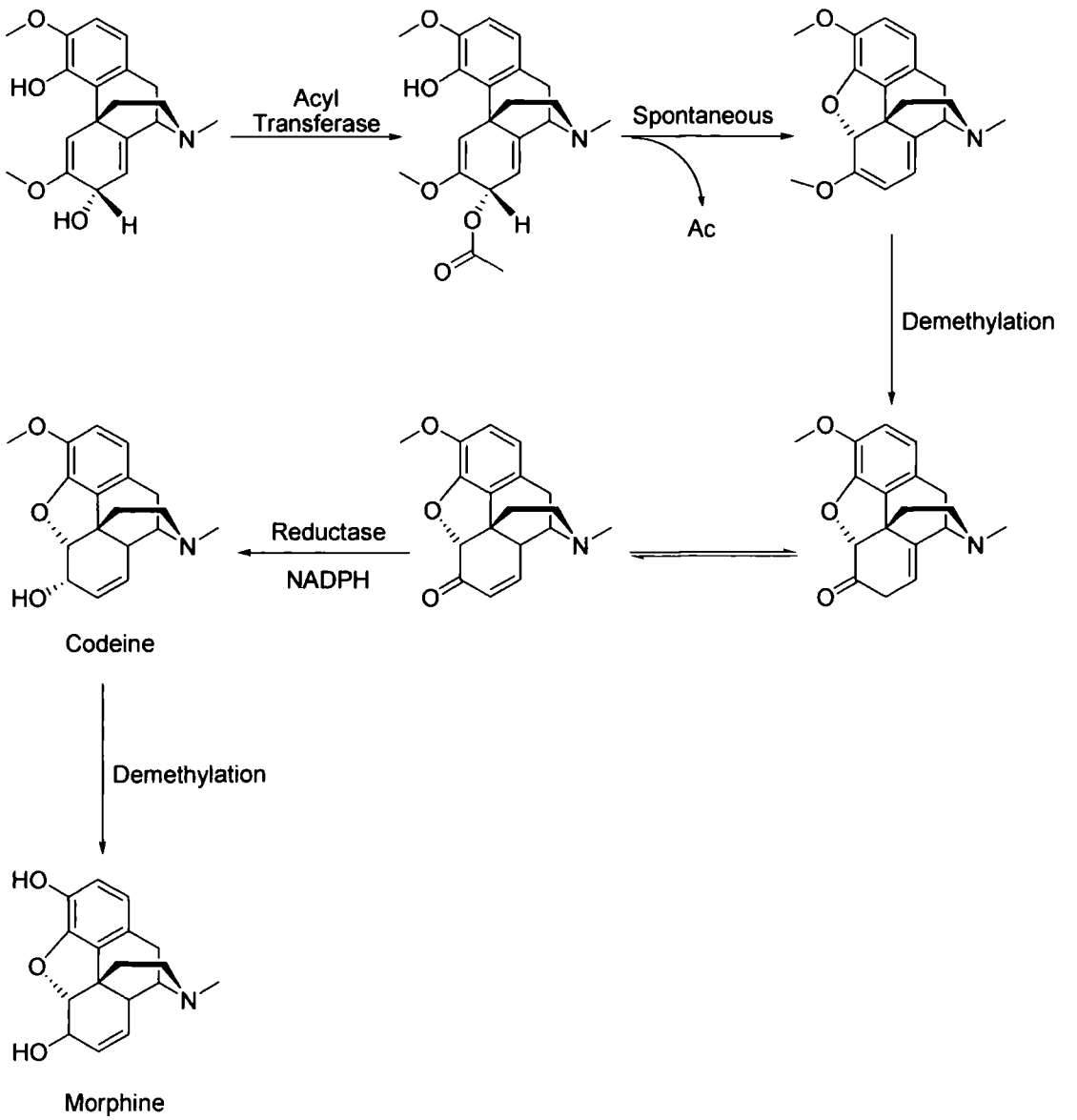


Figure 1.29 The latter stages of Morphine biosynthesis

1.3.3.4 Vindoline biosynthesis

Vindoline is a monoterpenoid indole alkaloid found in *Catharanthus roseus* and has been found to possess anti-cancer properties after coupling to Catharanthine, the product of a separate biosynthetic pathway. Vindoline possesses a terminal

acetylation step (Figure 1.30), which has received attention in both synthetic[76] and biosynthetic[58] literature as a prerequisite for synthesis of the Catharanthine dimer.

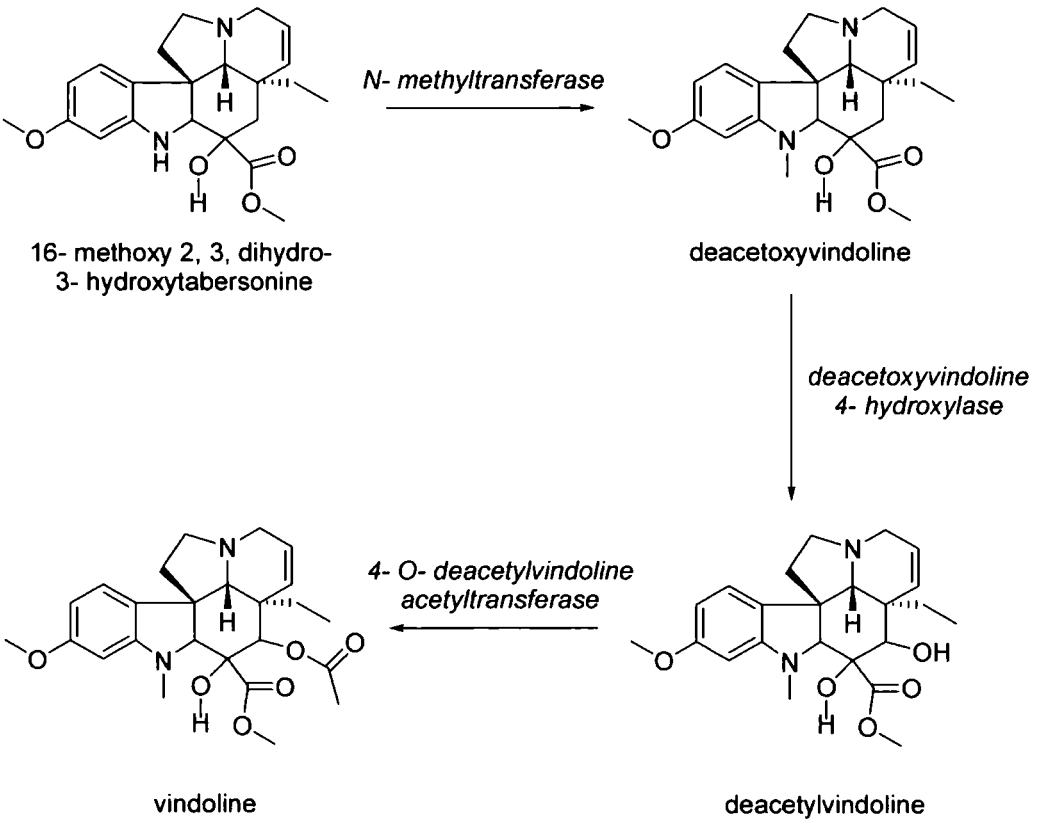


Figure 1.30 The latter stages of Vindolene biosynthesis

Studies carried out by Van der Heijden *et. al.* revealed only a small amount of the bioactive dimer was present in *Catharanthus roseus*, thus hindering the commercially viable extraction of the pharmaceutical[77]. They consequently produced several cell lines one of which produced high levels of the alkaloids catharanthine and tabersomine (a vindoline precursor), although the final acetylated vindoline metabolite was not produced. St-Pierre and De Luca established that tabersomine is converted into vindoline over 6 enzymatic steps and as such revealed the most promising route toward vindoline synthesis[58]. The terminal *O*- acetyltransferase step was elucidated by De Luca *et. al.* in 1998[58]. In contrast to the other 5 enzyme

activities, metabolite studies showed the acetyltransferase activity was not present in either the cell culture lines or wild type mature plants. The activity was subsequently noted in light-treated developing seedlings and further refinement revealed the latter stages of vindoline biosynthesis are highly light dependent and developmental stage - specific.

1.3.3.5 Other acylated plant natural products with bioactive properties

An increasing number of acylated plant natural products are being identified as having pharmaceutical properties where the acylating moiety is identified as being crucial for the bioactivity of these natural products. To this effect, three new dammar-type triterpenoid caffeates were isolated from *Celastrus rosthornianus*[78]. Of the three compounds, one of which is shown in figure 31, each showed strong antitumor activity against the human cervical squamous carcinoma cell line. Hydrolysis of the caffeoyl moiety in each case considerably reduced bioactivity.

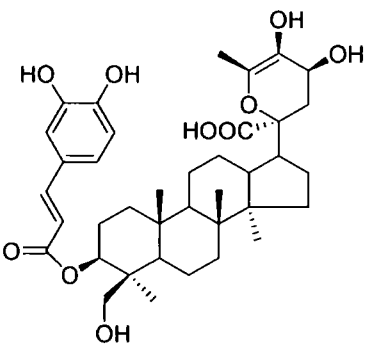
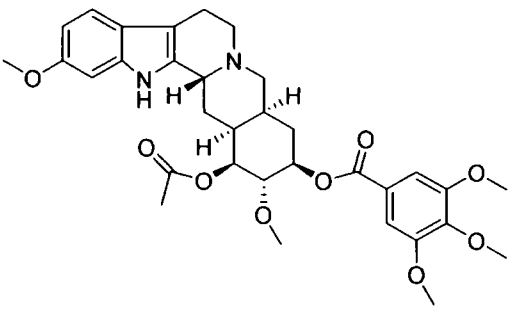


Figure 1.31 Trihydroxyl dammarene caffeate from *Celasrtus rosthornianus*

Reserpine is one of several natural products which has been isolated from *Rauvolfia serpentina* (Figure 32). The properties of this natural product have been intensively

researched owing to its anti-hypertensive properties[79]. Reserpine bears an unusual 3- 4- 5- methoxygallate acyl moiety.



Reserpine
Rauvolfia serpentina
Decreases blood pressure

Figure 1.32 Reserpine from *Rauvolfia serpentina*

1.4 Potential applications of coenzyme A-dependent acyltransferases

1.4.1 Heterologous expression of coenzyme A-dependent acyltransferases

cDNAs of several CADATs have been expressed as soluble catalytically-active enzyme in heterologous expression systems using yeast[80,81] and *Eschericia coli*[80-83] as host strains. In particular, the 10-*O*-acetyltransferase from taxol biosynthesis (section 1.3.3.2), having been overexpressed in unlysed *E.coli* cells, was able to acylate exogenously-fed 10-deacetyl baccatin III within the bacteria[84]. It was postulated that endogenous acetic acid was able to be converted into acetyl coenzyme A acyl donor by bacterial acetyl coenzyme A ligases, before acyltransfer. Furthermore a CADAT cDNA has successfully been transfected into a petunia plant, where the expressed protein allowed for the malonylation of anthocyanins *in vivo*[18]. Thus successful expressions of CADATs should be able to form the basis of successful methods in biotransformation.

1.4.2 Biosynthesis of natural products

Whilst there are currently no commercial biotransformations making use of CADATs, several scenarios where they might be applied have been suggested in the literature. The most common of these is in the biosynthesis of acylated natural products that exhibit desirable properties, such as those reviewed earlier in this chapter (Section 1.3)[66]. CADATs have been suggested as biocatalyst candidates in this role owing to their inherently high substrate-, regio- and stereoselectivity, overcoming the requirement for screening of non-endogenous enzymes. Therefore their proposed use is as replacement of chemical syntheses where yield of reaction is low due to non-specific reactions and unwanted consumption of what is often a valuable acyl acceptor[85]. To this effect, CADATs have been characterised in several instances in light of emerging pharmaceutical properties of acylated natural products that are complex and possess multiple reactable groups. This is mostly with a specifc view to future biosynthesis of the natural product[20,58,61,86].

Thus, with regard to commercial application of CAD biotransformations there remains aspects of their activity that remain sufficiently detrimental to commercial viability in the present. Such aspects are likely to include reliance upon expensive cofactors (ATP and/or CoA), required synthesis of an acyl donor and low number of previously characterised CADATs. However, relatively little research has been carried out upon optimisation or improvements in these areas, in comparison to acyltransfer processes involving hydrolase biocatalysts.

1.4.3 Modification of the functional properties of compounds by unnatural acylation

In terms of potential application of CADATs to either biocatalyse acyl transfer to non-endogenous acceptors or to transfer non-natural acyl moieties, there is again little precedent or research. Although Nakayama *et. al.* suggested that CADATs would potentially be suitable for the former role above, owing to the fact that an acyl substrate is known for a given enzyme and could therefore be screened for activity toward one substrate, the acyl acceptor, as opposed to both[21]. Furthermore, exogenously-fed, non-natural acids bearing halogens have previously been successfully incorporated into anthocyanin biosynthesis in carrot cell suspension cultures and therefore there is potential to further increase the range of carboxylic acids able to be transferred using CADAT biocatalysts[51]. Dougall *et. al* suggested that CADATs may have application in production of second generations of natural products.

To this effect, several bioactive natural product scaffolds have successfully been acylated with non-natural acyl moieties, by synthetic means, in order to regulate and/or alter the compounds pharmaceutical effect[87-89]. For example, second-generations of the pharmaceutical taxol have been synthesised[85]. Novel fluorinated acylating moieties were reacted with the core acyl acceptor structure, Baccatin III, although regiochemical control was limited resulting in reduced yield and loss of starting material (Figure 1.33). Cytotoxity studies of the two novel compounds were evaluated against 5 human cancer cell lines and were found to be significantly more potent than their paclitaxel counterpart[85].

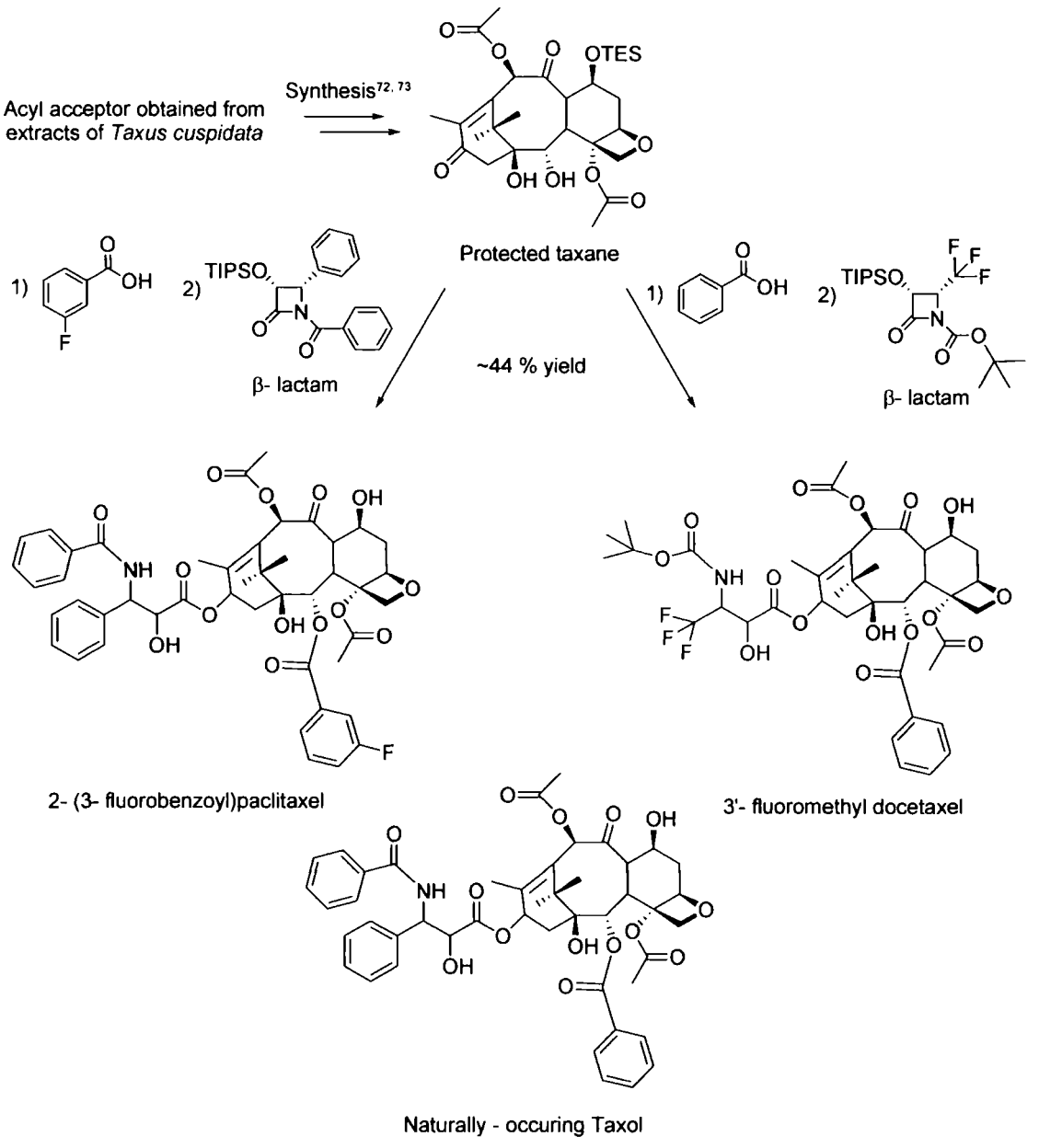


Figure 1.33 Synthesis of the second generation taxanes 2- (3- fluorobenzoyl) paclitaxel and 3'- fluoromethyl docetaxel increased the cytotoxicity of taxanes toward cancer cell lines and diversified the range of cancers treatable by these natural product homologues

The biochemical importance of halogenation has been likened to activation of non-reactive carbon centres by halogens, an effect commonly exploited in chemical synthesis[90]. Introduction of a halogen can activate a molecule toward bioreceptors, making the metabolite bioactive by the chemical activation of a carbon centre. Walsh

et. al.[91] characterised the halogenase 'CmaB'. They likened the role of halogenation to the effect of hydroxylation in most biological systems, whereby hydroxylation increases the reactivity of a substrate within a particular active site. Therefore, the reactivity / binding of a substrate toward catalytic residues within an active site can be altered by activation of key carbon centres i.e. by hydroxylation, bromination, chlorination or fluorination. Equally, the strategic positioning of a halogen atom upon a pharmacon or metabolite can prevent the action of site-specific enzymes which are found to modify or degrade the compound, thus intercepting the compound before it can reach its biological target[65].

Fluorine, in particular, has been found to impart desirable pharmokinetic and pharmacodynamic characteristics upon many bioactive agrochemical and pharmaceutical compounds[92]. As the most electronegative atom in the periodic table, fluorine simultaneously modulates electronic[93], lipophillic[94] and steric[95] parameters of compounds, all of which can critically influence the properties of bioactive molecules. As the smallest atom in the periodic table aside from hydrogen, fluorine (1.47 angstrom van der waals radii) can be utilised as a bioisosteric substitution, for hydrogen (1.20 angstrom) or oxygen (1.52 angstrom)[95]. This has been an important advance in modifying pharmaceuticals as portrayed by the range of fluorine substitutions for hydrogen (Figure 1.34A)[92,96] and hydroxyl (Figure 1.34**B**)[92,97] groups on pharmaceuticals currently synthesised for experimental or clinical use.

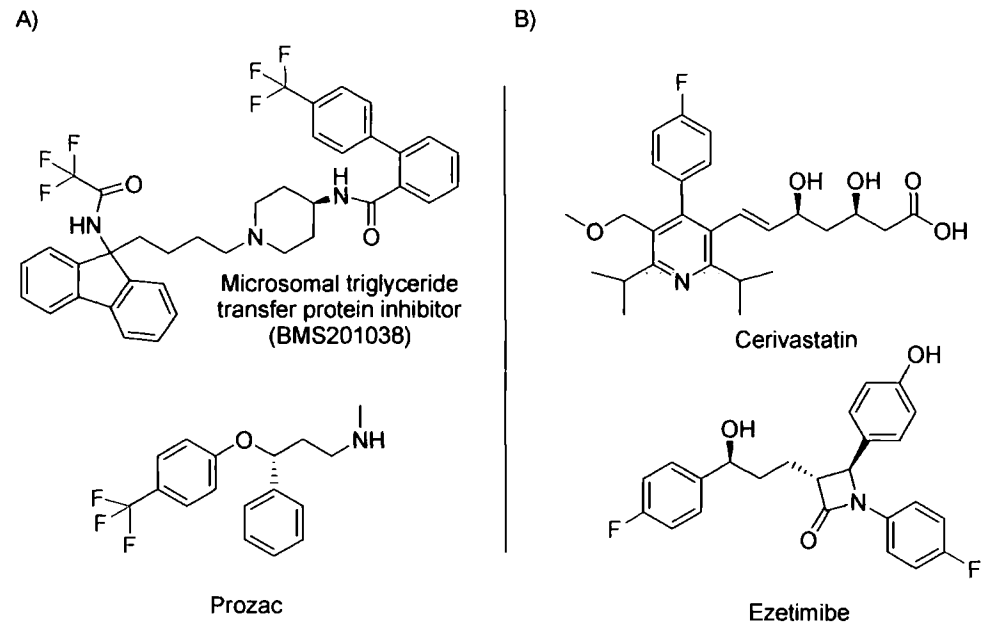


Figure 1.34 Effective fluorine substitution of (A) methylated and (B) hydroxylated pharmaceuticals to produce successful second generations of pharmaceuticals

Therefore the production of second generations of acylated natural products by biocatalysis is of interest and potentially useful. Many natural products have been identified as having a potential pharmaceutical or agrochemical role, although relatively few of these metabolites have been elaborated upon to produce refined analogues which are potentially more suited for a desired application.

Furthermore, radiolabelling of various drugs with ^{18}F has been an important diagnostic tool, allowing of cell, nerve tissue and tumour imaging by positron emmision tomography[98]. Katznellenbogen *et. al.*[99] utilized this technique for imaging of the absorption, distribution, site of action and excretion of anti-tumour drugs bearing ^{18}F labels. It was noted that the incorporation of fluorine into pharmaceuticals promotes the treatment of drugs as xenobiotics by biological systems, presumably due to the lack of fluorine containing endogenous metabolites.

Combination of these factors makes fluorine a valuable modification of pharmaceuticals and agrochemicals. However, chemical incorporation of fluorine into molecules post-synthesis is generally non-specific and often necessitates harsh conditions[93]. Both factors are undesirable in the synthesis of complex bioactive molecules and which have generated interest in bioactive fluorination methods, similar to those being researched for chlorination[90].

1.4.4 *In vivo* and *in vitro* biosynthesis of acylated flavonoids

There has been intensive research that indicates that flavonoids are beneficial to human health[100]. One of the most commonly cited reasons for this is the excellent antioxidant capacity of flavonoids[101,102]. Increased flavonoid content has also been cited to be a critical factor that contributes to enhancing the length of time plant-based food stuffs can be stored post-harvest, owing to their antioxidant capacity[33]. Aromatic acylation of flavonoids has been shown to enhance the antioxidant capacity of flavonoids in several studies[103], in addition to being shown to be critical for exhibition of several other beneficial properties, such as antimicrobial properties[104].

Such studies have led to increased research into acylated flavonoid production via biocatalysis of flavonoid acylation *in vitro*[104] and upregulation of endogenous flavonoid biosynthesis *in vivo*[105]. *In vivo* methodologies include genetic modification of plants to affect the regulation of key flavonoid biosynthetic genes such as isoflavone synthase[106] and chalcone isomerase[107]. External regulatory factors are also employed to upregulate flavonoid biosynthesis *in vivo*, such as gamma irradiation of fruits to boost flavonoid content during storage[108].

Mellou *et. al.* produced acylated derivatives of a monoglycosylated flavonoid using an immobilised *Candida antarctica* lipase and vinyl laureate as acyl donor[104]. The acylated flavonoids were found to have increased antioxidative activity toward low-density lipoprotein and serum model, respectively through increased lipophilicity.

There has been growing interest in the modification of the functional properties of anthocyanins and other flavonoids, by means of specific acylations[21]. Several flavonoid metabolites have been shown to possess beneficial health properties, such as antioxidant[109] and antimutagenic[110] properties. Acylation is thought to increase such health benefits of flavonoids by enhancing the chemical stability, solubility, membrane permeability and intestinal absorption of a metabolite[51,111]. For example, phenylpropanoyl esters of 3-sophorose 5-glucose cyanidin and peonidin from sweet potato (*Ipomaeo batatas*) were shown to have antimutagenic properties, where hydrolysis of the caffeoyl ester significantly decreased these properties[111]. A further study suggested selective absorption of the caffeoyl esters occured in rats and humans post-ingestion[112]. To this effect, recent studies have been undertaken to develop methods of production of supplementary acylated anthocyanins, by means of cultured cells of sweet potato and heterologous expression systems[113].

1.5 Modern proteomic probe approaches toward the isolation of specific enzyme activities

Recently, chemical probes have been developed that covalently modify the active site of specific proteins within a complex proteome, effectively labelling particular enzymes based upon their activities[114]. Traditionally, incorporation of either a fluorescent moiety or a biotin affinity moiety into the probe allowed rapid detection and purification of labelled proteins. Recently, it was demonstrated that both of these functional aspects of chemical probes could be effectively incorporated into a trifunctional probe based upon a linker consisting of the amino acid lysine, in order to accelerate protein discovery and characterisation[115] (Figure 1.35).

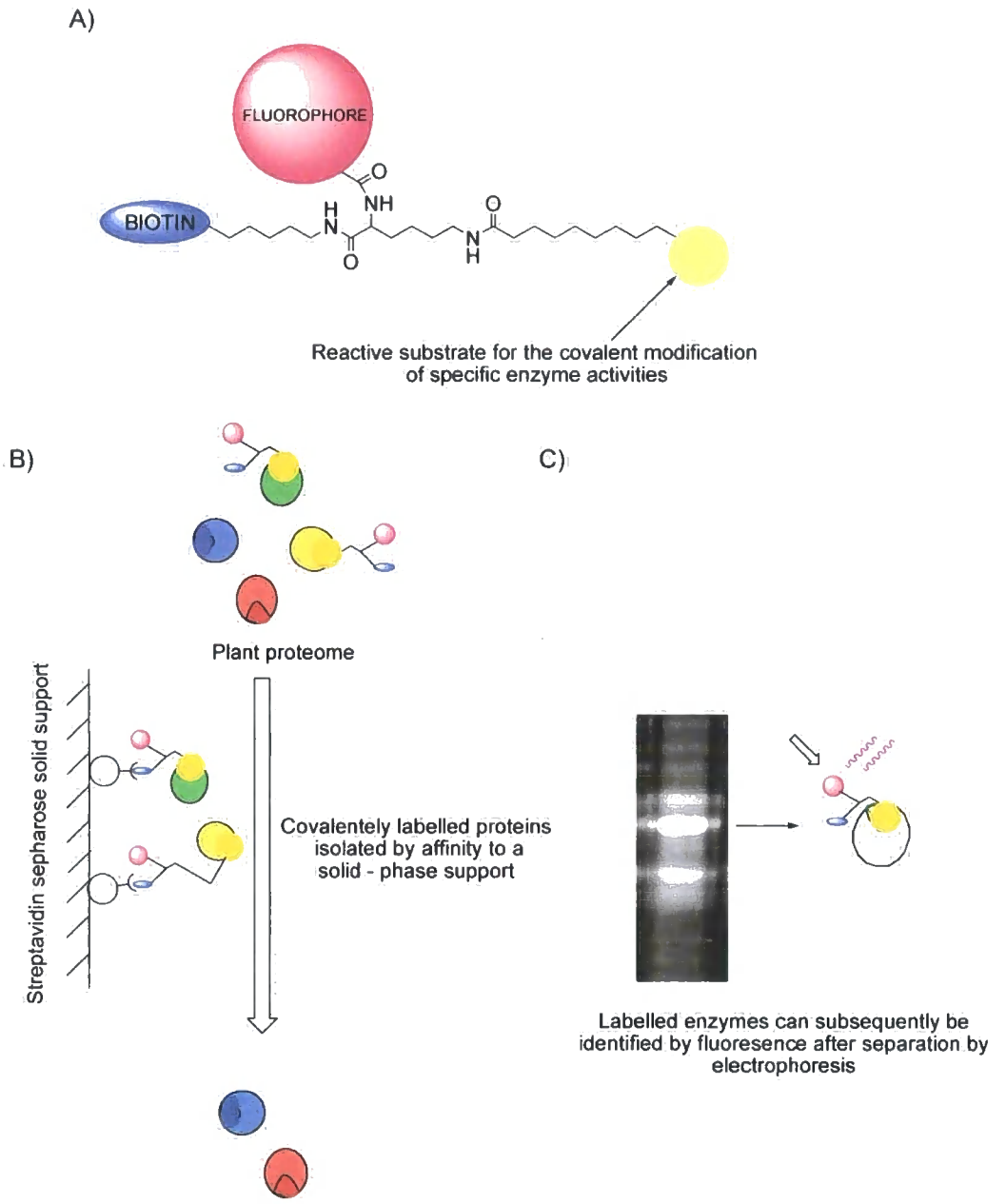


Figure 1.35 (A) A trifunctional chemical probe designed to detect and isolate specific proteins from within a complex proteome comprising a reactive moiety capable of covalent modification of specific enzyme activities, (B) a biotin residue, routinely used for isolation of labelled biological molecules from mixtures and (C) a fluorophore which allows identification of proteins that are specifically labelled by the probe

Several covalent inhibitors of enzymes have been coupled to chemical probes based upon this structure[115]. There are differing methods of covalent modification of

specific enzyme activities, which can be classed into two subgroups, mechanism specific inhibitors and affinity-substrate inhibitors. The former targets specific catalytic amino acid residues involved in the mechanism employed by a particular class of enzyme. This probe is known as a chemotype probe as it mimics the chemistry of the enzyme substrates allowing initiation of catalytic binding of the probe, but subtle differences in chemical nature means the protein is unable to eliminate the bound substrate[116]. In the case of fluorophosphonate inhibition of serine hydrolases (Figure 1.36 A), nucleophillic attack of water results in the elimination of ethanol from the phosphonate as opposed to elimination of the serine alcohol. As a result the phosphonate remains bound to the serine and a second water molecule cannot access the active site to undergo a second nucleophillic substitution.

The latter class of chemical probe mimics the structure of an enzyme substrate and is therefore preferentially bound by a protein target, whereupon a non-specific reactive moiety forms a covalent bond with the protein's infrastructure[117]. The example shown in Figure 1.36B, is a substrate bearing an azido moiety. The azide is photoreactive and forms a nitrene under UV irradiation, following elimination of nitrogen. The nitrene reacts rapidly with proximate covalent bonds (in this case a carbon-hydrogen bond) by a free radical mechanism. The probe is finally covalently bound to the proteins active site.

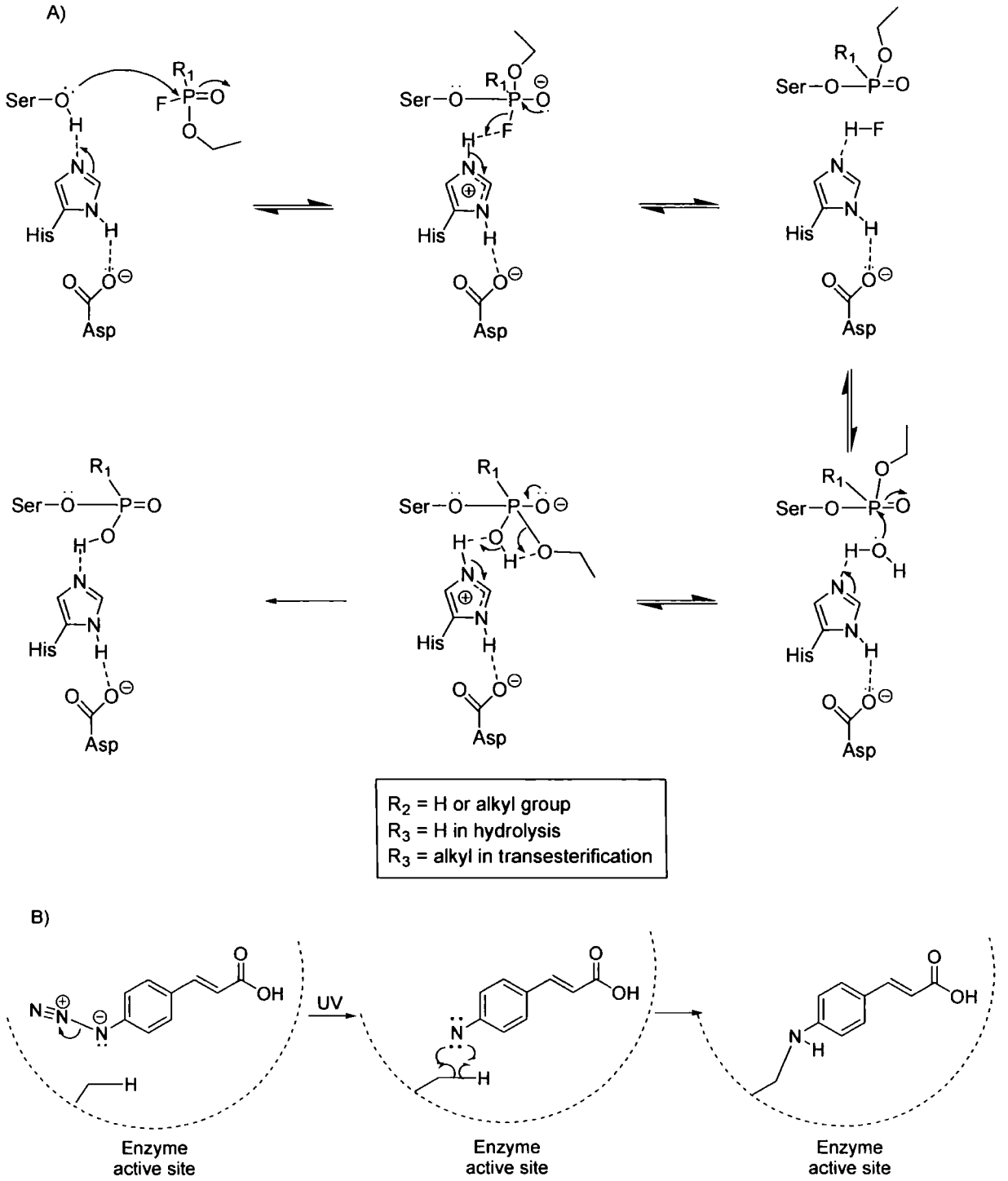


Figure 1.36 Two different mechanisms of covalent modification used in proteomic approaches. A) A mechanism-based (chemotype) fluorophosphonate inhibition B) An azido-bearing substrate-affinity probe that reacts with the proteins infrastructure following UV irradiation

Once a protein has been labelled and subsequently isolated from a proteome, identification is necessary. This is most commonly achieved by proteomics using either matrix-assisted laser desorption (MALDI) mass spectrometry or MS-MS based sequencing. In MALDI-based proteomics the protein is identified from established data based upon its unique pattern of peptide fragments, derived from predictable tryptic digestion[118]. In MS-MS based proteomics the tryptic peptides are individually sequenced by fragmentation to derive their sequential amino acid content[119].

1.6 Conclusion

In conclusion, the literature review highlights the diversity of acylation in plant secondary metabolism. Identification of novel acylated natural products is occurring at a steady rate and many of these compounds have been found to be bioactive. In contrast the isolation and characterisation of the acyltransferases responsible for these biotransformations has not been so readily achieved. It would therefore be of interest to determine whether or not chemical probes could therefore be useful tools in the facile and generic discovery of CADATs, which are otherwise challenging biochemical targets. Increasing the speed of discovery, and therefore characterisation, of these enzymes could considerably increase their potential as biocatalysts.

Development of acyltransferases for commercially-viable biotransformation has only been achieved with hydrolases as biocatalyst. However, coenzyme A-dependent enzymes in plant secondary metabolism can catalyse the transfer of a diverse array of carboxylic acids to many types of natural product acceptors. In particular, a vast number of CADATs in flavonoid biosynthesis are able to perform specific acylations

with both aliphatic and aromatic acids, although a much greater diversity of aromatic acids are known substrates. Thus this class of enzymes could add a whole new array of potential substrate-biocatalyst combinations to the existing library of enzymes able to perform acylation. Furthermore, CADATs could be recuited directly to biocatalyse their respective natural reactions, where they acylate many desirable natural products.

Taxol is an example of an acylated plant natural product with pharmaceutical properties and great commercial value, which could be semi-synthesised almost exclusivly by CADATs from 10-deacetylbaccatin III isolated from European yew trees. However, despite each of the CADATs required for the biosynthesis of Taxol being cloned, solubly expressed and shown to be active, Taxol remains to be produced by chemical methods. However, relatively little research has been undertaken into applied biocatalysis with CADATs and the reasoning behind this appears to be a deference toward chemical acylation. This is presumed to be associated with the cost of enzymatic production, for example with respect to requirement of cofactors, in comparison to a synthetic or lipase-mediated biosynthetic approach.

One advantage of harnessing the potential of CADATs as biocatalysts would be the ability to predict the nature of the acid each enzyme is able to transfer and the acyl acceptor, which is not always the case with unnaturally-recruited biocatalysts such as lipases. Furthermore plant CADATs have been proposed to have potential for application in biotransformations *in vivo*, where endogenous acyltransferases, expressed in cell cultures, have been manipulated to either boost the existing production of valuable natural products (Taxol) or perform non-natural

biotransformations with exogenously-fed carboxylic acids (Flavonoid metabolism). To this end, substitution of halogens and in particular fluorine, has been particularly effective in the modification of the properties of natural products and carboxylic acids bearing halogens have been successfully introduced into CADAT metabolism in plant-derived cell cultures. However, of the CADATs that have been isolated, most have only been characterised with respect to their endogenous substrates and close analogues.

1.7 Aims and objectives

- To characterise aromatic coenzyme A-dependent acyltransferases and coenzyme A ligases of flavonoid metabolism with regard to performing biotransformations both *in vivo* and *in vitro*

- Specifically, acyltransfer of phenylpropanoids will be characterised, as the most diverse and abundant substrates of acyltransfer in flavonoid metabolism

- To ascertain the ability of both CoA-dependent enzymes to turnover non-natural, fluorinated aromatic acids and to establish the effect substitution has upon activity

- To synthesise the required coenzyme A conjugates of both endogenous and non-natural acylating groups for use in enzyme studies

- To develop facile routes toward the isolation and identification of coenzyme A-dependent acyltransferases from plants, in order to accelerate the production of a

library of biocatalysts to add to the existing library of hydrolase-based acyltransferases

Chapter 2

Materials and Methods

2.1 Materials

Materials for the synthesis of trifunctional probes were obtained from Pierce biotechnologies, Rockford IL. 4-Azidocinnamic acid was obtained from Professor Rob Field, Dept. of Biological Chemistry, John Innes Centre, Norwich.. Cyanidin 3-5- *O*- diglucoside was obtained from Apin chemicals. DNA primers were synthesised by MWG-Biotech, Ebersberg, Germany. Other reagents and enzymes for molecular biology were obtained from Boehringer-Mannheim (now part of Roche, Basel, Switzerland) and all other chemicals and materials were obtained from Sigma - Aldrich, except where stated otherwise. All chemicals were of analytical grade unless specified.

2.2 Methods

2.2.1 Plants

2.2.1.1 Whole Plants

One tray of moist potting compost for each species was sown with seed prepared as detailed in figure 2.1. These trays were kept in a growth room for a total growing time specified below.

Species	*Seed preparation*	*Total growth time (days)*
Arabidopsis thaliana	none	21
Triticum aestivum	Soaked in water over night	21
Petunia hybrida	none	28
Gentiana triflora	none	50

Figure 2.1 Preparation of seed and total time employed for growth of each of the species in this study

Growth time depended on germination efficiency and growth rate, as plants were only harvested when they had produced sufficient material for extraction (~10 g or more). Growth room conditions were 16 hours light, 8 hours dark. In the light periods the temperature was maintained at 24 °C, which was dropped to 22 °C for the dark periods. The light intensity was 80 μEinstein m^{-2} s^{-2} in the photosynthetically active range. The compost was kept moist with regular watering. At harvest, the plants were flash frozen in liquid nitrogen and then stored at - 80 °C until extracted.

2.2.1.2 Arabidopsis Suspension Culture

Growth medium for *Arabidopsis* suspension culture contained 30 g sucrose and 4.43 g Murashige and Skoog basal salts with minimum organics (MSMO), dissolved in distilled water. Then 0.5 mL 1 mg/mL NAA and 50 µl 1 mg/mL kinetin were added and the pH adjusted to pH 5.7 using potassium hydroxide. The volume was made up to 1 l using distilled water. Conical flasks to be used for tissue culture were prepared by filling with water and autoclaving for 15 minutes at 121 °C to remove any residual detergent. Medium (80 mL) was decanted into each 250 mL flask. The flasks were stoppered with cotton wool, covered with foil, then autoclaved as above. Seven days after sub culture, a 10 mL inoculation of *A. thaliana* (Columbia) suspension culture was decanted into each flask under sterile conditions. The cultures were then incubated at 25 °C in the dark with shaking at 120 rpm. After seven days, the cultures were filtered to remove the medium and the cells flash-frozen in liquid nitrogen before being stored at -80 °C.

2.2.1.3 Regulation of plant metabolism

2.2.1.3.1 Elicitation of *Arabidopsis thaliana* cell cultures with yeast elicitor

Yeast elicitor was prepared by the method outlined by Kessman and Edwards[120]. Yeast elicitor (10 µl) was added to cell cultures grown to mid – log phase as previously described[121].

2.2.1.3.2 Regulation of flavonoid metabolism in plants with UV-B irradiation

A seed tray containing 24 plants was transferred to an enclosed UV lamp visualisation box. A CAMAG UV lamp with two 8 W light tubes was used to treat plants with UV – B (254 nm) irradiation for 10 minutes. After treatment the plants were removed and reintroduced into normal growth conditions for time periods in between 0 and 76 hours, as required, before harvesting of the leaf and stem tissue.

2.2.1.3.3 Regulation of flavonoid metabolism in plants with continuous high - light treatment

Seed trays, each containing 24 plants, were transferred into a Sanyo versatile environmental test chamber during the morning period of a given day. Conditions within the growth chamber were set to 24 hours of light (160 μ einstein m^{-2} s^{-2}) per day, at 17 ^{o}C. Seed trays were removed from the growth chamber at given time intervals, whereupon their leaf and stem tissue was immediately harvested.

2.2.1.4 Phenylpropanoid feeding studies

2.2.1.4.1 Phenylpropanoid feeding studies in *Arabidopsis thaliana* cell suspension cultures

Methanol solutions of each phenylpropanoid were prepared to a concentration of 60 mM. These solutions were added to 250 mL conical flasks containing 80 mL of mid – log phase cell cultures[121], as prescribed in figure 2.2.

Volume of 60 mM phenylpropanoid added (µl)	*Final concentration of cell culture (µM)*
67	50
134	100
201	150

Figure 2.2 Addition of phenylpropanoid solution to cell cultures of Arabidopsis

Cells were harvested by vacuum filtration upon fine gauze at predetermined time points after addition of phenylpropanoid. Control experiments consisted of addition of 200 µl of methanol.

2.2.1.4.2 Phenylpropanoid feeding studies in plants

Methyl esters of each phenylpropanoid were prepared (section 3.2.1.14) for this study. Each phenylpropanoid ester was solubilised in 50 mL 0.1 % Triton X detergent at 1 mM concentration. Phenylpropanoid – containing detergent was applied to plants as a fine spray at a rate of 25 mL per seed tray (800 cm^2) containing 24 plants. Plants were harvested at predetermined time points after application.

2.2.2 Protein extraction

Frozen plant material was ground to a powder with a pestle and mortar under liquid nitrogen and was then extracted into 5 volumes of chilled 0.1 M Tris-HCl buffer pH 8, containing 1 mM DTT, 2 mM EDTA and 5% w/v polyvinylpolypyrrolidone

(PVPP). The extract was centrifuged at 10000 g for 15 minutes at 4 °C, and the supernatant was decanted and retained while the pellet was discarded. Ammonium sulphate was added to the supernatant to 40 % saturation, and the resulting suspension was recentrifuged (10000 g, 15 minutes). The pellet was discarded and ammonium sulphate was added to the supernatant to a saturation of 80%, followed by a final centrifugation. The pellet from this final centrifugation was stored at –20 °C. Prior to use, protein was dissolved into 0.1 M potassium phosphate or 0.1 M Tris - HCl at the desired pH, then desalted using a HiTrap™ desalting column (Amersham Biosciences). This was carried out on a Bio-Rad HRLC system, with a Bio-Rad model 1740 UV/Vis monitor used to observe when the protein eluted from the column. The concentrations of the desalted protein solutions were determined using Bio-Rad Protein Assay according to the manufacturer's instructions.

2.2.2.1 SDS-PAGE

SDS-PAGE (sodium dodecyl sulphate- polyacrylamide gel electrophoresis) gels were prepared using the mini-PROTEAN II kit from Biorad according to the method of Laemmili (1970). Resolving gels were polymerised from 10 % or 12.5 % acrylamide/ *bis*-acrylamide (Sigma) in 375 mM Tris/ HCl, pH 9.0 containing 0.1 % (v/v) TEMED (Promega, Madison, WI, USA), 0.1 % (w/v) SDS and 0.1 % (w/v) ammonium persulphate. The stacking gel was polymerised from 4 % acrylamide/ *bis*-acrylamide, 126 mM Tris/ HCl, pH 6.8, 0.1% (v/v) TEMED, 0.1 % (w/v) SDS and 0.05 % (w/v) ammonium persulphate. Protein samples were diluted with an equal volume of 2 x SDS loading buffer (100 mM Tris-HCl, pH 6.8, 20 % glycerol, 200 mM DTT, 4% w/v SDS, 0.2 % w/v bromophenol blue) and incubated at 95 °C for 5 min prior to

loading on to gel. The samples were loaded into the wells and electrophoresed in SDS-PAGE running buffer (25 mM Tris, 192 mM glycine, 0.1 % SDS, pH 8.3) at 100 V for the first 10 min and then at 200 V until the dye front reached the bottom of the gel. Gels were washed thoroughly with water to remove the SDS and then if required, stained with Coomassie blue reagent (0.01 % w/v coomassie brillant blue, 5 % v/v 95 % ethanol:water (95:5 v/v) and 10 % v/v phosphoric acid:water 85:15 v/v).

2.2.2.2 Western blotting and detection of polypeptides bearing *strep* – tag II

Strep - tagged polypeptides were separated by SDS - PAGE (see above) and then electroblotted onto a PVDF membrane (Hybond-P, Amersham Biotech) using a mini Trans-Blot cell (Biorad) according to manufacturers instructions. After blotting the membrane was blocked with 3 % skimmed milk powder in Tris-buffered saline (TBS; 1.93 % w/v Tris, 9 % w/v glycine) for 1 hr. *Strep*-tactin – alkaline phosphatase conjugate (IBA technologies) was then added at a 1: 2000 to 1: 5000 dilution and incubated at 1- 2 hr at room temperature or overnight at 4 °C. The membrane was then washed twice for 10 min with TBST (TBS plus 0.1 % v/v Triton X-100) and then for a further 10 min with TBS. The membrane was rinsed in 100 mM Tris - HCl, pH 8.5 and then developed in 5 mL 100 mM Tris - HCl, pH 8.5 containing 15 µl 5-Bromo 4- chloro 3' indolyphosphate p- toluidine salt (1 mM) and 15 µl nitro – blue tetrazolium chloride (1 mM). When optimal colouration of strep – tagged proteins was achieved, the blot was rinsed thoroughly with distilled water.

2.2.3 Plant metabolites

2.2.3.1 Metabolite extraction

2.2.3.1.1 Extraction of methanol – soluble metabolites

Plant samples were extracted with ice-cold methanol (10 v/w) using a pestle and mortar with acid-washed sand as an abrasive. After centrifugation (16 000 g / 5 minutes) the extract was concentrated to ~1 mL under reduced pressure. Chloroform (1 mL) and high – purity water (200 μl) were added and the resulting emulsion was centrifuged (16 000 g / 1 minute) or until a biphasic solution was achieved. The upper layer (methanolic) was removed and retained for further analysis.

2.2.3.1.2 Extraction of anthocyanins

Anthocyanins were extracted with 5 % metaphosphoric acid (1 mL per gram plant tissue) using a pestle and mortar with acid-washed sand as an abrasive. After centrifugation (16 000 g / 5 minutes) the extract was used without further preparation.

2.2.3.2 Metabolite profiling

HPLC - MS (high – performance liquid chromatography mass spectrometry) analysis of natural products was performed with a Waters 2790 HPLC chromatography system. A pre - column and a Synergi Polar-RP (250 x2 mm) column (Phenomenex, Macclesfield, Cheshire) were used to separate metabolites. The eluents were solvent A, 0.5 % formic acid and solvent B, acetonitrile containing 0.5 % formic acid. Metabolites were eluted by a linear increase of 5 % to 100 % solvent B from 0 to 45

min, followed by isocratic elution with 100 % B for a further 6 min. A volume of 20 µl per sample was injected, unless otherwise stated. The flow rate was 0.2 mL min^{-1} and the eluant was monitored by diode array detection for UV - Vis absorbing metabolites between 200 nm and 600 nm, using a Waters 996 photodiode array detector. The eluant was subsequently analysed by ESI-TOF MS on a Micromass LCT mass spectrometer operating in positive or negative ion mode as required, with the cone voltage set at 20 V, desolvation temperature at 200 °C and sample cone heated to 120 °C.

2.2.4 Assays

2.2.4.1 Assay of phenylammonia lyase by UV spectrometry

Crude protein (0.1 mL) was dissolved in 990 µL of 50 mM Tris - HCl buffer pH 8.5 and 12.1 mM L- phenylalanine. The rate of formation of cinnamic acid was determined at 37 °C, through measurement of the change in absorbance at 290 nm in quartz cuvettes and over a time period of 2 hours. D- phenylalanine was not used as substrate and therefore was used as control.

2.2.4.2 Assay of 4- coumarate coenzyme A ligase by UV spectrometry

Protein (2.5 µg) was dissolved in 50 µL of 100 mM Tris - HCl buffer pH 8.0 containing 10 mM dithiothreitol. To this was added 550 µL of 100 mM Tris - HCl buffer pH 8.0 containing 1.6 mM coenzyme A, 8.3 mM ATP and 8.3 mM $MgCl_2$. The rate of formation of coenzyme A esters was determined at 30 °C, through measurement of the change in absorbance as specified below in quartz cuvettes.

Measurements were taken over a time period of 30 minutes. The extinction coefficients of coenzyme A esters used in this project are also specified below (Table C-Figure 2.3). Raw data is shown for novel compounds 4- fluorocinnamoyl coenzyme A (Table A) and 3- 4- fluorocinnamoyl coenzyme A (Table B). Extinction coefficients of natural coenzyme A esters were comparable to those identified in recent literature[15] (Appendix C).

A)

Reaction product	*Conc. (mM)*	*Absorbance*	*Absorbance*
4- Fluorocinnamoyl coenzyme A	0.001	0.015	0.015
4- Fluorocinnamoyl coenzyme A	0.01	0.151	0.153
4- Fluorocinnamoyl coenzyme A	0.05	0.760	0.762
4- Fluorocinnamoyl coenzyme A	0.1	1.535	1.539

B)

Reaction product	*Conc. (mM)*	*Absorbance*	*Absorbance*
3- 4- Difluorocinnamoyl coenzyme A	0.001	0.015	0.014
3- 4- Difluorocinnamoyl coenzyme A	0.01	0.149	0.144
3- 4- Difluorocinnamoyl coenzyme A	0.05	0.729	0.715
3- 4- Difluorocinnamoyl coenzyme A	0.1	1.430	1.420

C)

Reaction product	*Wavelength (nm)*	*Extinction coefficient* ($mM^{-1}cm^{-1}$)
Coumaroyl coenzyme A	333	20.3 +/- 0.08
Caffeoyl coenzyme A	347	18.1 +/- 0.11
Feruloyl Coenzyme A	350	19.0 +/- 0.07
4- Fluorocinnamoyl coenzyme A	308	15.3 +/- 0.02
3- 4- Difluorocinnamoyl coenzyme A	310	14.3 +/- 0.09

Figure 2.3 A) Raw absorption data for 4- fluorocinnamic acid B) Raw absorption data for 3- 4- fluorocinnamic acid C) Absorbance maxima and extinction coefficients of the phenylpropanoid coenzyme A thioesters by UV spectrophotometry

2.2.4.3 Assay of 4- coumarate coenzyme A ligase by High-Performance Liquid Chromatography (HPLC)

Reactions, as detailed above, were stopped by addition of 100 µL methanol and the samples were centrifuged (16000 g, 5 minutes), to pellet precipitated protein. 150 µL of supernatant was decanted for analysis by HPLC:

System:	Dionex GP50 gradient pump with a Dionex UVD 340U UV / Vis detector and AS 50 autosampler
Mobile phase:	A: 0.5 % formic acid
	B: HPLC grade acetonitrile containing 0.5 % formic acid
Stationary phase:	Synergi Polar-RP (250 x 2 mm) column (Phenomenex, Macclesfield, Cheshire)
Flow rate:	0.8 mL / min

The column was equilibrated with 5 % B. Two minutes after injection of the sample (30 µL), a linear gradient increasing solvent B up to 100 % was carried out over twelve minutes. The column was then washed with 100 % B for two minutes before returning to the start conditions in preparation for the next sample. Eluate from the column was analysed for UV absorbance by a diode array and detected at the wavelengths specifically detailed in figure 2.3. Standards were used to confirm the retention time of the substrates. Product formation was quantified by calibrating the HPLC with known amounts of product.

2.2.4.2 Assay of chalcone synthase by HPLC chromatography

Crude protein (50 µL) in 100 mM Tris - HCl buffer pH 8.0 containing 20 mM ascorbic acid, was added to 350 µL of 100 mM Tris - HCl buffer pH 8.0 containing 1 mM malonyl coenzyme A and 1 mM coumarol coenzyme A. The rate of formation of tetrahydroxychalcone at 30 °C was determined after 30 minutes by HPLC chromatography. The reaction was stopped by addition of 350 µL methanol and the samples were centrifuged (16000 g, 5 minutes), to pellet precipitated protein. 150 µL of supernatant was decanted for analysis by HPLC:

System:	Beckman System Gold 125P Solvent Module and 166 Detector with Gilson 234 Autoinjector
Mobile phase:	A: 0.5 % formic acid
	B: HPLC grade acetonitrile containing 0.5 % formic acid
Stationary phase:	4.6 x 45 mm Beckman Ultrasphere ODS column
Flow rate:	1 mL / min

The column was equilibrated with 5 % B. Two minutes after injection of the sample (30 µL), a linear gradient increasing solvent B up to 100 % was carried out over eight minutes. The column was then washed with 100 % B for two minutes before returning to the start conditions in preparation for the next sample. Eluate from the column was analysed for UV absorbance at 287 nm. Tetrahydroxychalcone (Apin chemicals) was used as authentic standard for analysis of product formation.

2.2.4.2 Assay of BAHD acyltransferases by HPLC

Crude protein (50 µL) or recombinant protein (2.5 µg) in 100 mM Tris - HCl buffer pH 8.0 containing 2 mM dithiothreitol, was added to 50 µL of 100 mM Tris - HCl buffer pH 8.0 containing 0.5 mM acyl acceptor and 0.5 mM coenzyme A acyl donor. The reaction was incubated at 37 °C and the formation of acylated product was determined after 10 minutes by HPLC chromatography except where stated otherwise. The reaction was stopped by addition of 50 µL acetonitrile and the samples were centrifuged (16000 g, 1 minute), to pellet precipitated protein. 100 µL of supernatant was decanted for analysis by HPLC:

System:	Dionex GP50 gradient pump with a Dionex UVD 340U UV / Vis detector and AS 50 autosampler
Mobile phase:	A: 0.1 % trifluoroacetic acid
	B: HPLC grade acetonitrile containing 0.1 % trifluoroacetic acid
Stationary phase:	Synergi Polar-RP (250 x 2 mm) column (Phenomenex, Macclesfield, Cheshire)
Flow rate:	0.8 mL / min

The column was equilibrated with 5% B. Two minutes after injection of the sample (30 µL), a linear gradient increasing solvent B up to 100% was carried out over 16 minutes. The column was then washed with 100% B for two minutes before returning to the start conditions in preparation for the next sample. Eluate from the column was

analysed for UV absorbance at 520 nm and enzymatic products were analysed by mass spectrometry as detailed in section 2.2.3.2.

2.2.5 Gene cloning and expression

Oligonucleotide primers were designed to amplify At4CL 1 and Gent5AT as described in the relevant chapters 5 and 6 and obtained from MWG – Biotech Ebersberg, Germany.

2.2.5.1 mRNA extraction and reverse transcription

Plant material (100 mg) was ground in liquid nitrogen and extracted with 1 mL tri-reagent (Sigma). The extract was spun at 12 000 g for 15 minutes at 4 °C. The upper phase was retained and 0.5 mL isopropanol added. After five minutes incubation at room temperature, the mixture was spun again at 12 000 g. The pellet was washed with 75 % ethanol, then dried and dissolved in 25 µL sterile water. The RNA solution was then transferred to ice for five minutes. 4 µL 5 × incubation buffer (250 mM Tris-HCl, 200 mM KCl, 30 mM $MgCl_2$, 50mM dithiothreitol pH 8.3), 1 µL RNase block, 2 µL dNTPs (mixture of dATP, dTTP, dGTP and dCTP each at 10 mM) and 1 µL M-MuLV reverse transcriptase was then added, and the reaction incubated at 37 °C for 90 minutes. The cDNA produced was kept at –20 °C.

2.2.5.2 Polymerase chain reaction

The reaction mix listed below was made up in 0.5 mL PCR tubes.

5 µL 10× buffer (Supplied with KOD HiFi polymerase)

5 µL 0.2 mM dNTP mix (dATP, dCTP, dTTP and dGTP)

5 µL $MgCl_2$ 20 mM

2.5 µL 20 mM forward primer

2.5 µL 20 mM reverse primer

2 µL template (cDNA)

30.5 µL dH_2O

0.5 µL KOD polymerase (Added during step 1 in figure 2.4)

The mixture was then subjected to the program detailed in Figure 2.4 in an Eppendorf Mastercycler Gradient PCR machine

Step	*Temperature (°C)*	*Duration (min:sec)*
1	94	2:00
Repeat steps two to four inclusive 30 times		
2	94	0:30
3	55 °C (at4cl 1) 62 °C (gent5at)	0:45
4	72	1:30
After repeats, proceed to step five		
5	72	10:00
	END	

Figure 2.4 Temperature program used for PCR of At4CL 1 and Gent5AT respectively

PCR products were analysed on a 0.8 % agarose gel, and if required, purified from the gel using a Promega Wizard® SV Gel and PCR Clean-Up kit (Promega UK, Southampton, UK).

2.2.5.3 Restriction

Restriction digests were carried out on 0.5 – 1 μg DNA using the restriction enzymes detailed in Chapters 5 and 6. 2 U of each restriction enzyme (where 1 U of enzyme will digest 1 μg of substrate DNA in a total reaction volume of 50 μl using the optimal buffer) was added to the appropriate buffer and the solution was made up to a total volume of 20 μl. Digests were incubated at 37 °C for 1-16 hr before analysis of the reaction products by agarose gel electrophoresis.

2.2.5.4 Ligation

1 μL 10 × ligation buffer (as supplied with enzyme) and 1 μL T4 ligase (Promega) were incubated with a 3: 1 ratio of gene insert to plasmid for ligation (total reaction volume 10 μL) at 15 °C overnight.

2.2.5.5 Transformation

Transformation of competent cells was carried out following the manufacturer's instructions. In brief, 20 μL competent cells were incubated on ice with 1 μL plasmid for 5 minutes. The cells were then heat-shocked at 42 °C for exactly 30 seconds before being returned to ice for a further minute. 80 μL LB medium (10 g/L NaCl, 10

g/L bacteriological peptone, 5 g/L yeast extract) was added and the cells were then incubated at 37 °C for one hour. The cells were then plated on to LB agar (as for LB, with 15 g / L agar) with appropriate antibiotic selection.

2.2.5.6 Minipreps

10 mL cultures of transformed bacteria were grown and the cells were pelleted. Plasmids were purified from the cells using a Promega Wizard® Plus Minipreps DNA Purification System.

2.2.5.7 Expression conditions and recombinant His_6 - At4CL1 purification

A single transformed *Escherichia coli* (*E. coli*) colony was used to inoculate a 10 mL LB starter culture, containing 100 μg / mL kanamycin and 35 μg / mL chloramphenicol. The culture was incubated overnight at 37 °C with shaking at 200 rpm. A 500 mL culture was then inoculated with 5 mL of the starter culture, and the cells grown under the same conditions until they achieved an optical density at 600 nm of approximately 0.7. The cells were then induced with 1 mM IPTG (unless stated otherwise) and incubated for a further three hours. The bacteria were then centrifuged at 10 000 g for 10 minutes. The pelleted bacteria were resuspended in 5 mL 0.1 M Tris-HCl pH 7.8 with 20 mM imidazole (buffer A) and sonicated whilst on ice three times for 30 seconds with 10 seconds between each burst. After the bacterial lysate had been centrifuged (16 000 g, 5 minutes), it was loaded on to a HisTrap HP column (Amersham Biosciences). After washing with buffer A the retained protein was eluted by raising the imidizole concentration to 300 mM. Further purification

was achieved by reapplying the His-tagged protein to a fresh metal chelate affinity column and following the above procedure. The protein was desalted as described in section 2.2.2.

2.2.5.8 Expression conditions and recombinant *Strep*-Gent5AT purification

A single colony of transformed Arcticexpress™ *Escherichia coli* (Stratagene) harbouring Cpn 60 and Cpn 10 chaperonins from *O. antarctica* was used to inoculate a 10 mL LB starter culture, containing 100 μg/mL kanamycin and 35 μg/mL chloramphenicol. The culture was incubated overnight at 37 °C with shaking at 200 rpm. A 500 mL culture was then inoculated with 5 mL of the starter culture, and the cells grown under the same conditions until they achieved an optical density at 600 nm of approximately 0.6. The cells were then cooled to 4 °C and induced with 0.1 mM IPTG and incubated for a further 12 hours. The bacteria were centrifuged at 5000 g for five minutes and resuspended in 10 mL 0.1 M Tris-HCl pH 8 with 150 mM sodium chloride and 1 mM EDTA (buffer A) and sonicated three times for 20 seconds with 10 seconds between each burst. After the bacterial lysate had been centrifuged (4200 g, 15 minutes), it was loaded on to a Strep-tactin column (IBA technologies). After washing with buffer A, retained protein was eluted by addition of 5 mL buffer B (0.1 M Tris-HCl pH 8 with 150 mM sodium chloride, 1 mM EDTA and 2.5 mM desthiobiotin). The protein was desalted as described in section 2.2.2.

2.2.5.9 Agarose gel electrophoresis

Agarose gels were prepared by microwaving TAE (4.84 % w/v Tris base, 1.14 v/v glacial acetic acid) and 0.8 - 1.5 % molecular biology grade agarose (Helena BioSciences) until the agarose had melted. The gel mix was cooled to approximately 60 °C and ethidium bromide added at a concentration of 10 μl L^{-1}. The gel was poured into the casting tray of HORIZON® 58 horizontal gel electrophoresis apparatus (GibcoBRL®, Paisley, UK). Samples were prepared by diluting with 6x loading buffer (0.25 % w/v bromophenol blue, 0.25 % v/v xylene, 15 % w/v Ficoll) and the gel loaded with a lane of markers (1 Kb DNA ladder, GibcoBRL®) prior to electrophoresis TAE buffer at 120 V for 20 min.

2.2.5.10 DNA sequencing and analysis

Double stranded cDNAs were sequenced using automated florescent sequencers Applied Biosystems 377 DNA sequencer XL or the 373 DNA sequencer using the University of Durham sequencing service.

DNA sequences were edited, translated into three frames and restriction sites determined using the DNA sequence editing and analysis program DNA for Windows 2.4.0 (software written by Dr D. P. Dixon, School of Biological and Biomedical Sciences, University of Durham). DNA and protein sequences were aligned using CLUSTALW[122] and sequence similarities were determined using BLAST (Basic Local Alignment Search Tool)[123].

2.2.6 Proteomic probes

2.2.6.1 Inhibition of proteins with 4- azidocinnamic acid and 4- azidocinnamoyl coenzyme A substrate-affinity probes

Aliquots (100 µL) of the protein (10 µg) to be labelled were placed into a microtitre plate. 10 µL of a methanol solution of 4- azido cinnamic acid (1 mg / mL) was added to the aliquots of protein. The microtitre plate was put on ice and placed inside a Bio-link crosslinker (BLX), where it was irradiated with 0.12 Joules / cm^2 of ultraviolet light (254 nm).

2.2.6.2 Proteomic profiling with fluorophosphonate trifunctional probe (FPP)

Frozen plant material, pulverised in a chilled mortar and pestle was extracted with 0.1 M Tris / HCl pH 8 with 1 mM DTT and 5% cross-linked polyvinylpolypyrrolidone at 4 °C. The mixture was centrifuged at 15 000 g for 15 minutes, then FPP (Details of preparation are given in Chapter 3) was added to the supernatant to a final concentration of 5 µM. The resulting solution was incubated at room temperature for one hour. The solution was then passed through a PD-10 desalting column (Amersham Biosciences) to remove excess probe. Subsequently, SDS was added to a final concentration of 0.2 % (w/v) and the resulting solution heated to 95 °C for 5 minutes. After cooling, 20 µL of a 50 % streptavidin sepharose bead slurry was added and the mixture was incubated for one hour at 4 °C with end-over-end mixing. The beads were then washed twice with 0.2 % SDS, then twice with distilled water, and the bound proteins were eluted by heating the beads to 95 °C for five minutes in 1x SDS-PAGE loading buffer. Eluted bands were then run on an SDS - PAGE gel as

described in section 2.2.6. Use of UV treatment during FPP labelling of proteins was carried out as described in section 2.2.5.1.

2.2.6.2.1 Visualisation, quantification and identification of labelled proteins

The fluorescently labelled proteins in the gel were visualised using a Fujifilm FLA-3000, exciting at 532 nm and reading the emission at 580 nm. Quantification of fluorescence was carried out using Aida Image Analyser version 3.11.002. Bands were excised and submitted for MALDI proteomics as described in section 2.2.5.2.2.

2.2.6.2.2 Identification of Purified Proteins by MALDI-TOF Mass Spectrometry

Polypeptide-bands were excised from the gel and digested with trypsin. The samples were then analysed by the Durham University proteomic service.

Method

1D band transferred to a well of a 96-well microtitre plate (Genomic Solutions). The 96-Well plate is transferred to the ProGest Workstation (Genomic Solutions Ltd.) where the samples are digested with trypsin using the ProGest long trypsin digestion protocol :-

- Each gel plug was equilibrated in 50µl of 50mM ammonium bicarbonate and subsequently reductively alkylated with 10mM DTT and 100mM iodoacetamide, followed by destaining and desiccation with acetonitrile. Gel plugs were re-hydrated with 50mM ammonium bicarbonate containing 6.6% (w/v) trypsin (Promega) and digested overnight. Peptides were extracted using 50% (v/v) acetonitrile, 0.1% (v/v) TFA into a final volume of 50µl (2 x 25µl extractions)

The resulting extracts were then freeze-dried and re-suspended in 10 µl of 0.1% formic acid. This is done using the 'thin-film' method:-

- Approx. 0.2 µl of matrix is spotted to the target plate. Matrix is α-cyano-4-hydroxy-cinnamic acid in nitrocellulose/acetone. 1 µl of digested sample is applied to the thin film & allowed to dry. The samples are washed in-situ with 0.1% TFA and left to dry again

MALDI-ToF PMF was then performed using a Voyager-DE™ STR BioSpectrometry™ Workstation (Applied Biosystems, Warrington, UK). De-isotoped and calibrated spectra are then used to generate peak lists, these are searched using MASCOT (www.matrixscience.com) mass spectrometry database search software.

2.2.7 Bioinformatics

2.2.7.1 Identification of previously characterised BAHD acyltransferases

Position specific iterative BLAST (PSI-BLAST, http://www.ncbi.nlm.nih.gov/BLAST/) was run through four iterations with an inclusion threshold of 10^{-10}, and the results were examined to identify all complete, non-redundant *Arabidopsis thaliana* sequences.

Chapter 3

Synthesis of metabolites and biological probes of acylating enzymes

3.1 Introduction

Acyltransferases have previously been identified, isolated and characterised via the synthesis and/or purification of their acyl donors[124] and acyl acceptors respectively[56]. Aside from benzoyl coenzyme A, aromatic coenzyme A thioester donors are not commercially available. This has been overcome in the past by chemical synthesis of the CoA thioester substrates[81 56 69]. In most cases, phenylpropanoid CoA donors were synthesised from the respective imidazolium [124] or *N*-hydroxysuccinimide ester [125] of the phenylpropanoid.

Acyl acceptors cannot be synthesised easily, due to their structural complexity and are normally isolated from the plant of interest[54,82]. However, this is dependent upon the presence of the precursors to the terminal natural product and is therefore a major challenge to the characterisation of acyltransferases requiring acyl acceptors that are not present within the plant. For example, the flavonoid acyl acceptors utilized by Fujiwara *et. al.*[56], Kaffarnik *et. al.*[126] and Yonekura-Sakakibara *et. al.*[80] were isolated from a wide array of plant sources through necessity, as they were not accumulated in the plants of interest.

This chapter explores synthetic routes toward (1) the acyl donors required for the identification and characterisation of aromatic BAHD acyltransferases; (2) phenylpropenoyl esters for the identification of the acyl donors and acyl acceptors of flavonoid metabolism by feeding and incorporation of metabolites; and (3) chemical

probes with potential activity toward CADATs in order to facilitate the isolation of these enzymes in future studies.

3.2 Experimental

Unless otherwise noted, starting materials and reagents were obtained from commercial suppliers and were used without further purification.

Chemicals and solvents were purchased from Fisher Scientific or Sigma-Aldrich and were used as received unless otherwise stated. All reactions were performed in flame or oven-dried glassware under a positive pressure of nitrogen or argon unless otherwise stated. Diethyl ether (Et_2O) and tetrahydrofuran (THF) were dried by refluxing with sodium-benzophenone under an atmosphere of N_2 and collected by distillation. Dichloromethane (DCM) was dried by heating under reflux over calcium hydride and distilled under an atmosphere of nitrogen. Triethylamine was dried over KOH pellets.

^1H and ^{13}C NMR spectra were measured on a Brüker Advance 400 instrument. *J* values are quoted in Hz. Chemical shifts are calibrated with reference to the residual proton and carbon resonances of the solvent ($CDCl_3$: δH = 7.26, δC = 77.0 ppm). Assignment of spectra was carried out using DEPT, COSY, HSQC and HMBC experiments.

Low and high resolution mass spectrometry was carried out as follows. 10 μl of a sample was injected onto an UPLC™ BEH with C18 column (1.7 μm, 2.1 x 100 mm;

Waters Acuity). The column was then eluted at 0.2 mL min^{-1} with a mixture of 0.5 % formic acid (solvent A) and acetonitrile containing 0.5 % formic acid (solvent B) using a linear gradient of 5 % to 95 % solvent B over 9 min, followed by isocratic elution with 95 % B for a further 2 min. The eluant was then analysed by MS after electrospray ionization using a Micromass QTof spectrometer operating in positive or negative ion mode as required. Settings were sample cone voltage = 41 kV, capillary voltage = 2.55 kV, extraction cone voltage = 5 kV, source temperature = 100 °C, desolvation temperature =180 °C and for high resolution mass spectrometry injection of leucine enkephalin at 5 second intervals was used as lock mass for real-time recalibration.

Fourier Transform Infra-red (FT IR) spectra were recorded on a Perkin Elmer Paragon 1000 FT spectrometer. Absorption maxima are reported in wavenumbers (cm^{-1}).

Analytical thin-layer chromatography (TLC) was performed on pre-coated TLC plates SIL G-25 UV_{254} (layer 0.25 mm silica gel with fluorescent indicator UV_{254}) [3]. Developed plates were air dried and analysed under a UV lamp, Model UVGL-58 (Mineralight LAMP, Multiband $UV_{254/365}$ nm) and where necessary stained with iodine on silica or a solution of ceric ammonium molybdate (CAM). Flash column chromatography was performed using silica gel (40-63 μm, Fluorochem).

Where compounds had been synthesised previously and a literature preparation was followed the literature reference(s) are given in the sub-title to the relevent protocol. Unless otherwise stated all characterisations were in accordance with values found in the literature.

3.2.1 Synthesis of metabolites

The *trans*- 3-phenylprop-2-enoic acid derivatives used in this thesis are shown in figure 3.1. They will be generically termed 'phenylpropanoids' or by their non-systematic name (also given in figure 3.1), in line with current biochemical literature.

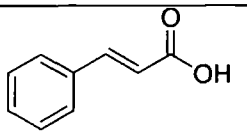
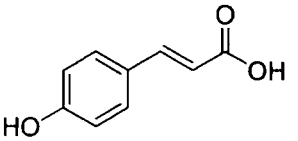
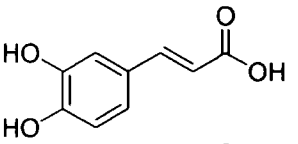
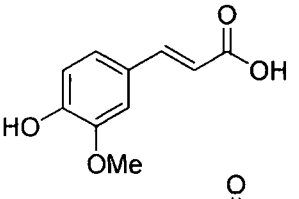
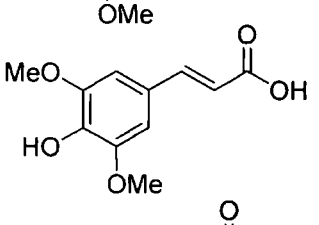
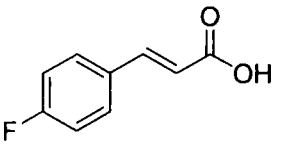
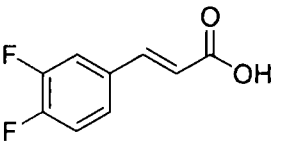
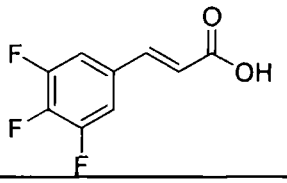
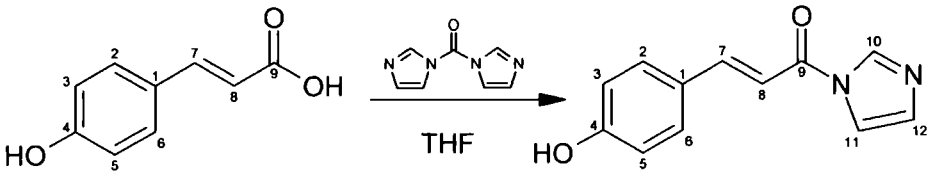
Structure	*Non-systematic name*	*Compound number*
	Cinnamic acid	**1**
	Coumaric acid	**2**
	Caffeic acid	**3**
	Ferulic acid	**4**
	Sinapic acid	**5**
	4- Fluorocinnamic acid	**6**
	3- 4- Difluorocinnamic acid	**7**
	3- 4- 5- Trifluorocinnamic acid	**8**

Figure 3.1 *trans*- 3-Phenylprop-2-enoic acid structures and terminology used

3.2.1.1 Synthesis of coumaroyl imidazole

This compound has previously been synthesised and similar methodology was used in this instance[124,125].



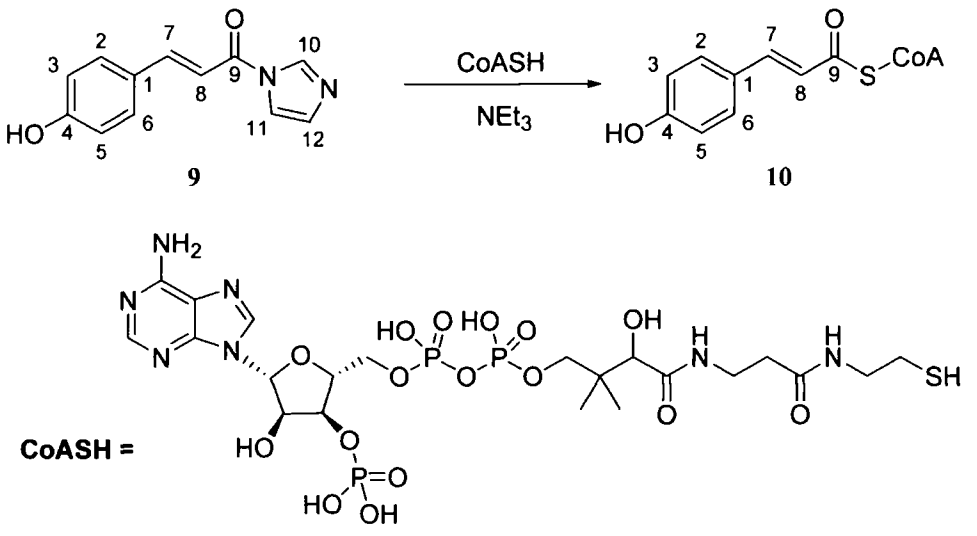
2 **9**

To a solution of coumaric acid (**2**)(0.5 g, 3.05 mmol) in tetrahydrofuran (5 mL) was added carbonyl diimidazole (2.47 g, 15.25 mmol). The reaction was stirred for 30 minutes at room temperature, before the solvent was removed *in vacuo* to afford a pale yellow powder. Immediately, diethyl ether (20 mL) was added to the solid and then decanted. This was repeated until clear ether was removed. Coumaroyl imidazole (**9**) was obtained as a pale yellow solid (0.50 g, 2.38 mmol, 78 %).

Characterisation: FTIR (NaCl disc) ν/cm^{-1}:1480, 1616 (C=C), 1667 (CON), 2590 (C=N), 3014 (sp^2 carbon un-saturation), 3450 (phenol); ^1H NMR (400 MHz, $CDCl_3$): 7.10 (1H, d, $J_{7,8}$ 16.5 Hz, 1xH8), 7.41 (2H, m, 1xH3, 1xH5), 7.59 (1H, d, $J_{11,12}$ 1.5 Hz, 1xH12), 7.65 (1H, d, $J_{12,11}$ 1.5 Hz, 1xH11), 7.79 (2H, m, 1xH2, 1xH6), 8.05 (1H, d, $J_{8,7}$ 16.5 Hz 1xH7), 8.32 (1H, s, 1xH10); ^{13}C NMR (125.7 MHz, $CDCl_3$): 117.6 (C8), 122.3 (C2,6), 130.5 (C3,5), 131.6 (C7), 132.8 (C1), 136.5 (C12), 137.7 (C11), 148.3 (C10), 152.1 (C4), 161.7 (C=O); MS *m/z* (ES^+):215 ($[M + H]^+ = 215$).

3.2.1.2 Synthesis of coumaroyl coenzyme A

This compound has previously been synthesised and similar methodology was used in this instance[124,125].

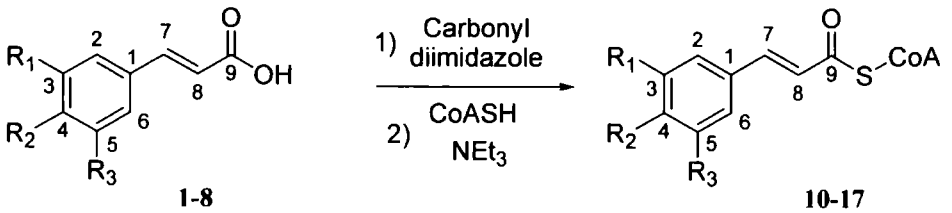


A solution of triethylamine, acetonitrile and distilled water (1: 1: 3) and a 1 mL portion of acetonitrile were thoroughly degassed. To a solution of coenzyme A (2 mg, 2.61 μmol) in triethylamine, acetonitrile and distilled water (1: 1: 3) was added coumaroyl imidazole (**2**) (1 mg, 5.22 μmol) in acetonitrile (0.1 mL). The reaction mixture was stirred under argon and on ice for 30 minutes. The resulting solution was purified by size exclusion chromatography (mobile phase: dH_2O, stationary phase: Biorad P2 SEC gel, wavelength detection at 310 and 340 nm). Fractions containing the first eluted compound (retention time = 1.2 min) were collated and lyophilized to yield coumaroyl coenzyme A (**3**).

Characterisation: See following section

3.2.1.3 Syntheses of phenylpropanoyl coenzyme A esters

The indicated compounds (*) have previously been synthesised and similar methodology was used to synthesise each of the compounds in this instance[124,125].



1-8 10-17

The synthesis of cinnamoyl, coumaroyl, ferruloyl, caffeoyl, sinapoyl, 4-fluorocinnamoyl, 3- 4- difluorocinnamoyl and 3- 4- 5- trifluorocinnamoyl coenzyme A was carried out in parallel following the method detailed in section 3.2.1.1 and section 3.2.1.2.

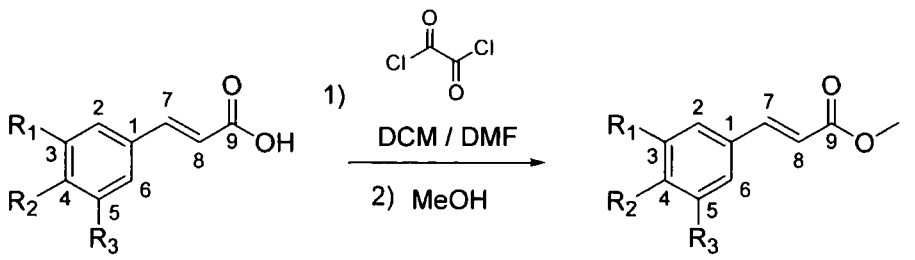
Characterisation:

Compound number	R_1	R_2	R_3	$[M-H]^-$ Da	*m/z* (ES^-) Da	Δ ppm	UV-Vis λ_{max} (nm)
10*	H	H	H	896.1498	896.1495	0.33	315
11*	H	OH	H	912.1447	912.1451	0.44	333
12*	OH	OH	H	928.1396	928.1402	0.64	347
13*	OMe	OH	H	942.1553	942.1556	0.32	353
14*	OMe	OH	OMe	972.1658	972.1653	0.51	350
15	H	F	H	914.1404	914.1409	0.55	308
16	F	F	H	932.1310	932.1306	0.43	312
17	F	F	F	950.1215	950.1221	0.63	313

* = Coenzyme A ester previously synthesised[124,125].

3.2.1.4 Synthesis of methyl phenylpropanoate esters

A standard ester synthesis was used to produce each of the compounds in this section, as detailed by numerous sources.



1-8 18-24

Phenylpropanoid derivative (0.91 mmol) was charged to a flask in dichloromethane (5 mL) and dimethylformamide (4 μL) was added. Oxalyl chloride (1.82 mmol) was subsequently added dropwise and the solution was stirred for 30 minutes. The reaction mixture was washed with water (3 x 10 mL) and dried over magnesium sulphate. Concentration *in vacuo* afforded the corresponding methyl cinnamate derivative as pale yellow - brown oil (86 – 96 % yield). Methyl phenylpropanoate esters were purified by reverse-phase preparatory HPLC.

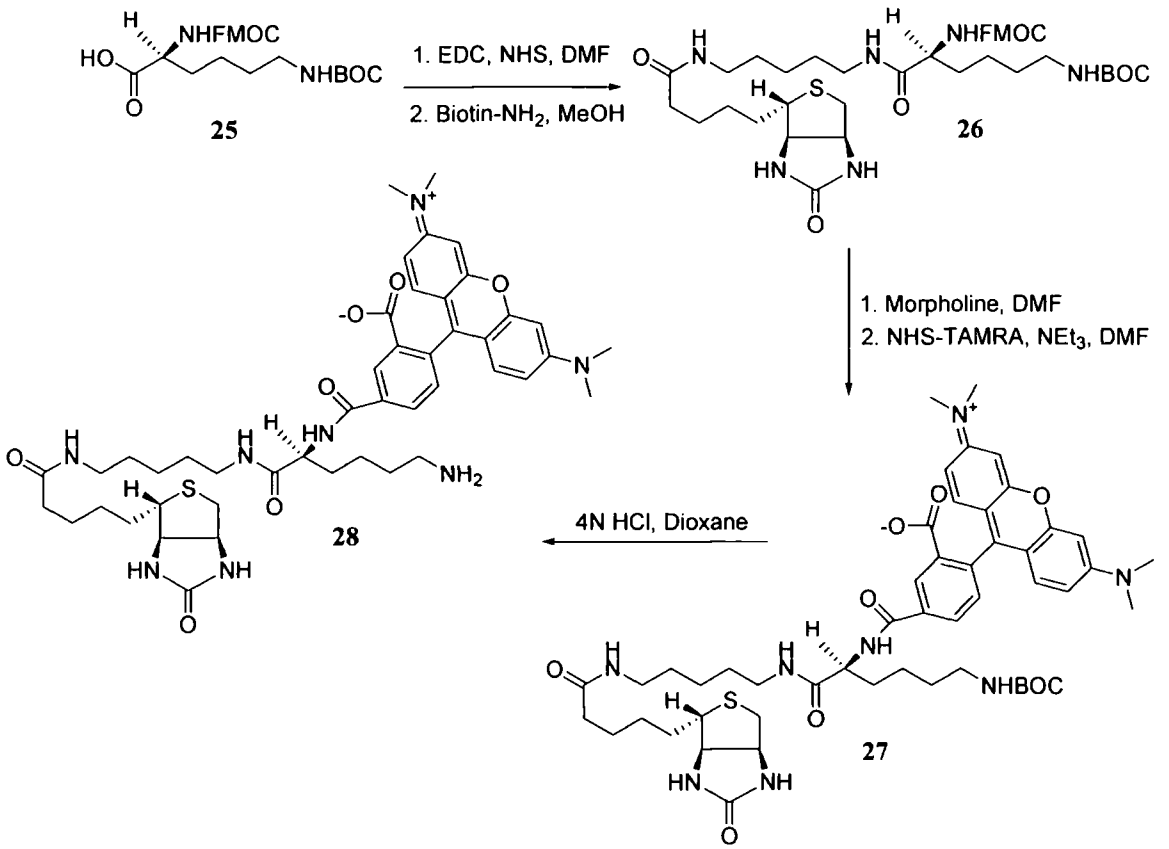
Characterisation:

Methyl ester	R_1	R_2	R_3	$[M + H]^+$ Da	*m/z* (ES^+) Da	Yield %
18	H	OH	H	179	179	92
19	OH	OH	H	209	209	89
20	OMe	OH	H	209	209	92
21	OMe	OH	OMe	239	239	91
22	H	F	H	181	181	86
23	F	F	H	199	199	95
24	F	F	F	217	217	86

3.2.2 The synthesis of trifunctional proteomic probes

All compounds within section 3.2.2 were designed and synthesised by Adam *et. al.*[115]. The methods employed in this section were directly derived from the method described by Adam *et. al.*[115]. Observations at each stage were found to be in accordance with those previously found and mass spectrometry was comparable to the characterisation achieved by Adam *et. al.*[115].

3.2.2.1 Synthesis of a rhodamine and biotin bearing bifunctional lysine backbone



Fmoc- *N*(ε)- Boc- L- lysine (**25**)(0.24 g, 0.5 mmol) was dissolved in dimethyl formamide (12 mL) and 1- ethyl- 3- (3- dimethylaminopropyl) carbodiimide (EDC) (0.128 g, 0.66 mmol) and *N*- hydroxysuccinimide (NHS) (0.128 g, 1.1 mmol) were

added. The solution was stirred over night at 25 °C, before addition of saturated $NaHCO_3$ (30 mL). The resulting suspension was extracted with ethyl acetate (3 x 30 mL). The organic fractions were washed with water (30 mL) and brine (30 mL), before being dried over magnesium sulphate and concentrated *in vacuo*. The resulting white powder was redissolved in methanol containing 1 % chloroform (10 mL) and biotin amine (60 mg) was added. After stirring at room temperature for 1 hour, the solution dried *in vacuo*. The resulting white solid was washed with ethyl acetate (2 x 20 mL) and chloroform was added (20 mL) and the suspension was stirred for 3 hours at room temperature. The resulting white solid was isolated by filtration affording compound **26** (76 mg, 0.1 mmol, 20 %). MS *m/z* (ES^+):779 ($[M+ H]^+ = 779$).

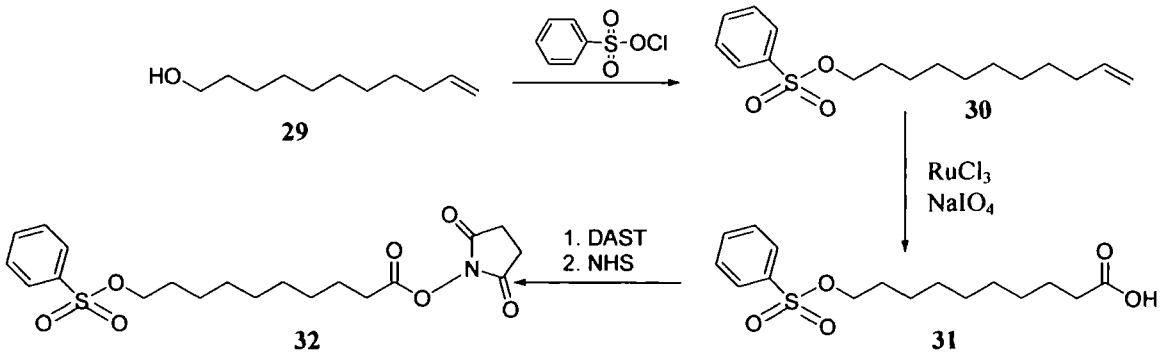
To a solution of biotinylated lysine intermediate (**26**) (30 mg, 39 μmol) dimethyl formamide (2 mL) and morpholine (0.3 mL) were added. The solution was stirred at 25 °C for one hour, before the solution was concentrated *in vacuo*. A solution of anhydrous dimethylformamide (2 mL), containing NHS-TAMRA (18 mg, 34 μmol) and triethylamine (50 μl) was used to re-dissolve the resulting white solid. The solution was stirred for two hours at room temperature, before all volatiles were removed *in vacuo*. Compound **27** (30 mg, 30 μmol, 77 %) was isolated by HPLC (Phenomenex, RP - C_{18} preparatory column, 5 - 100 % acetonitrile gradient, 30 minutes). MS *m/z* (ES^+):969 ($[M + H]^+ = 969$).

To a solution of compound **27** (15 mg, 15 μmol) in anhydrous methanol (0.3 mL) was added 4N HCl in dioxane (3 mL) and the solution was stirred for 1 hour at room temperature. The volatiles were removed under a stream of nitrogen. The de-protected compound **28** was solubilised and subjected to HPLC for analysis. The de-protection

was found to be 100 % successful and therefore, compound **28** was utilised in future steps with no further purification. MS *m/z* (ES^+):899 ($[M + H]^+ = 899$).

3.2.2.2 Synthesis of *N*- hydroxysuccinimide esters of chemotype labelling moieties

3.2.2.2.1 Synthesis of benzene sulphonoyl decanoic acid



To a solution of benzene sulphonyl chloride (1.04 g, 5.40 mmol) in pyridine (4 mL) was added ω - undecylenyl alcohol (**29**)(0.5 g, 3.00 mmol) in pyridine (4 mL). The resulting solution was stirred at 0 °C for 6 hours, before being taken up into ethyl acetate (50 mL) and washed with water (3 x 40 mL), 10 % HCl (2 x 50 mL) and brine (50 mL). The organic phase was dried over magnesium sulphate and concentrated *in vacuo*. Purification by column chromatography on silica (hexane: EtOAc, 98: 2) gave benzyl sulphonyl- 10- undecene (**30**)(0.90 g, 2.88 mmol, 96 %) as a colourless oil.

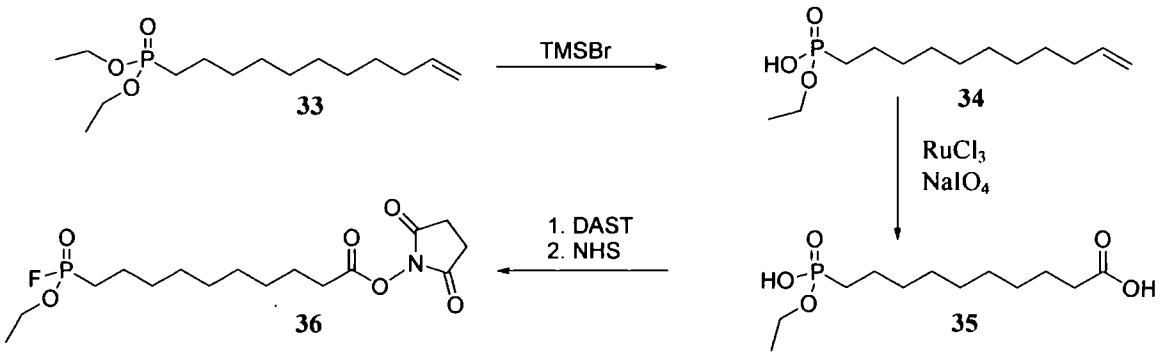
Compound **30** (0.90 g, 2.88 mmol) was dissolved in a biphasic solution composed of CCl_4: CH_3CN: H_2O (10 mL: 10 mL: 15 mL) and treated sequentially with sodium periodate (2.53 g, 11.80 mmol) and ruthenium trichloride hydrate (0.005 g, 0.02 mmol). The reaction was stirred at 25 °C overnight before partitioning with dichloromethane (100 mL) and 1 N aq. HCl (2 x 100 mL). The organic layer was washed with saturated aq. NaCl (100 mL), dried over magnesium sulphate and concentrated under reduced pressure. Column chromatography (40% EtOAc/Hex)

afforded benzene sulphonyl decanoic acid (**31**) as a white solid (0.65 g, 1.99 mmol, 69 %).

Characterisation: HRMS ES^+ *m/z:* 352.1202 ($[M + H]^+$ = 352.1189).

To a solution of compound **31** (0.009 g, 0.027 mmol) in anhydrous dichloromethane (1.5 mL) at -78 °C, was added DAST ((diethylamino)sulphur trifluoride) (81 µl, 0.3 mmol). The solution was warmed to room temperature over 10 minutes and *N*-hydroxysuccinimide (0.15 g, 1.3 mmol) was added in dimethylformamide (0.5 mL). After 15 minutes, EtOAc (50 mL) was added and the solution was washed with brine (4 x 50 mL) and dried over magnesium sulphate. The solvent was removed *in vacuo* to afford benzene sulphonyl- *N*- (hydroxysuccinyl)decanamide (**32**) as a yellow oil which was used without further purification (0.011 g, 0.026 mmol, 98 %).

3.2.2.2.2 Synthesis of 10-(fluoroethoxyphosphinyl)-*N*-(hydroxysuccinyl) decanamide



A solution of **33** (0.31 g, 1.07 mmol) in dichloromethane (4 mL) was treated dropwise with trimethylsilyl bromide (0.17 mL, 1.28 mmol). The reaction was stirred at 25 °C for 1 h, quenched with aq. $KHSO_4$ solution (5 mL, 5 % w/v) and stirred vigorously for 5 minutes. The reaction mixture was partitioned between ethyl acetate (100 mL) and

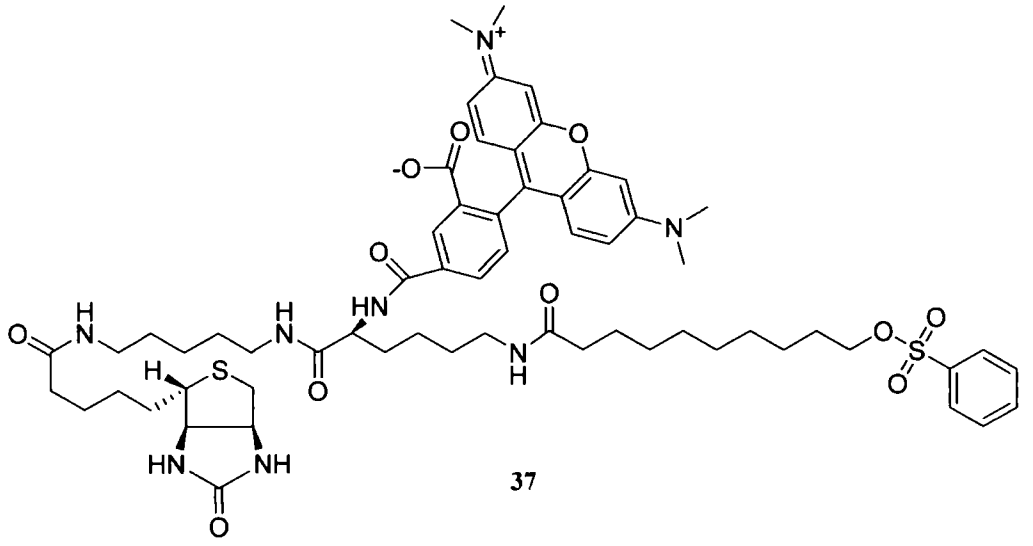
water (100 mL), and the organic layer was washed with saturated aq. NaCl (200 mL), dried over magnesium sulphate and concentrated under reduced pressure to give **34** as a colourless oil (0.22 g, 0.86 mmol, 80 %).

The synthesis of compound **35** (0.17 g, 0.62 mmol, 72 %) was carried out as described for compound **31** in section 3.2.2.2.1.

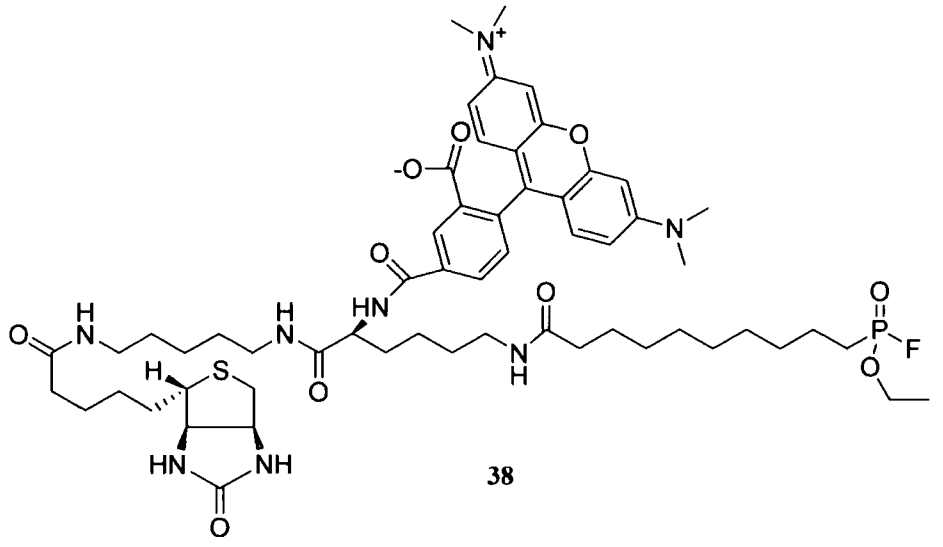
Characterisation: MS *m/z* (ES^+): 262 ($[M + H]^+ = 262$).

The synthesis of compound **36** was carried out as described for compound **32** in section 3.2.2.2.1 and was utilised without further purification.

3.2.2.3 Synthesis of trifunctional probes bearing biotin, rhodamine and chemotyping-reactive groups



37



38

To a solution of the deprotected compound **28** (5 mg, 5.5 μmol) in methanol (3 mL), was added compound **32** (9.4 mg, 22 μmol) and $NaHCO_3$ (5 mg). After stirring for 4 hours at room temperature, the reaction mixture was filtered and concentrated *in vacuo*. The trifunctional phenyl sulphonate probe (**37**) was purified by HPLC (Phenomenex, preparatory RP-C_{18} column: 5-100 % acetonitrile in 0.1 %

trifluoroacetic acid). The same method was used to synthesise **38** from compound **28** and **36**.

Characterisation **37**: MALDI - MS m/z: 1180.50 ($[M + H]^+$ requires 1180.52).

Characterisation **38**: MALDI - MS m/z: 1134.38 ($[M + H]^+$ requires 1134.38).

Chapter 4

Regulation and modification of natural product acylation *in vivo* with exogenously-fed carboxylic acids

4.1 Introduction

Characterisation of biochemical pathways requires the activity of interest to be identifiable at the point of experimentation. Secondary metabolic enzymes are expressed when a metabolic response is required to counteract an external environmental change. CADATs and their associated coenzyme A ligases have been shown to be regulated under UV-irradiating conditions and post-fungal infection (Section 1.3.1.3). Having identified active acyltransferase activity, the identity of the acyl acceptor(s) will be sought by LC-MS metabolic profiling as a prerequisite for their isolation and biochemical characterisation in future chapters.

Upon establishing acyltransferase activity, it was of interest to attempt to incorporate exogenous carboxylic acids (phenylpropanoids) into flavonoid metabolism. Firstly, this would enable assessment of the potential for CADATs and CLs to use unnatural substrates, such as fluorinated aromatic acids, in biosynthesis of unnaturally-modified natural products (Section 1.4.3). Secondly, the effect of feeding phenylpropanoids to plants will be guaged with respect to regulation of natural product biosynthesis and in particular flavonoid production. To this effect, exogenously-fed phenylpropanoids have previously been incorporated into, and caused upregulation of, flavonoid metabolism in plant-derived cell cultures in order to produce acylated flavonoids, for use as potential health products (Section 1.4.4). It was envisaged that transfer of this technology into plants could have application in enhanced production of beneficial-

natural products and/or unnaturally-acylated analogues directly in harvestable plants. To this effect upregulation of acylated flavonoid biosynthesis in to-be-harvested plant material is a topic of great interest to crop producers (Section 1.4.4). The antioxidant properties of flavonoids, and in particular acylated flavonoids, are observed to enhance shelf life in addition to increasing the potential health benefits of the foodstuff (Section 1.4.4).

Thus the aims and objectives which are associated with this chapter are as follows:

- To characterise aromatic coenzyme A-dependent acyltransferases and coenzyme A ligases of flavonoid metabolism with regard to biosynthesis of natural products both *in vivo* and *in vitro*

- To ascertain the ability of both CoA-dependent enzymes to turnover non-natural, fluorinated aromatic acids and to establish the effect substitution has upon enzyme activity

Specifically the objectives of experiments in this chapter were three-fold: To identify phenylpropanoyl coenzyme A-dependent acyltransfer in plants and elucidate the nature of the acyl acceptor by metabolic profiling; To screen fluorinated phenylpropanoids against CL and CADAT activities *in vivo*; and finally to establish the effect of feeding aromatic acids to flavonoid metabolism in plants.

The model plant Arabidopsis thaliana will be characterised with a view to identification of two putative CADATs, a coumaroyl and a sinapoyl transferase[35], which are to date unknown. Whilst feeding studies will be undertaken in two separate

plant species, with contrasting flavonoid acylation. *Arabidopsis thaliana*, which accumulates phenylpropanoylated anthocyanins, and *Petunia hybrida*, which accumulates phenylpropanoylated flavonols, will be used for this purpose.

4.2 Results

4.2.1 Identification of anthocyanin acyl acceptors in *Arabidopsis thaliana*

Several anthocyanin metabolites have been identified in the foliage of *Arabidopsis thaliana*[35], the majority of which are acylated with esters of malonic acid, coumaric acid and sinapic acid. An anthocyanin-malonyl transferase has recently been isolated and characterised[44], but the enzymes which catalyse the acyltransfer of coumaric acid and sinapic acid are currently unidentified. Equally, the identity of the acyl acceptor required for each aromatic acyltransfer is to date unknown.

4.2.1.1.1 *Arabidopsis thaliana* cell suspension cultures

Methanolic extraction of metabolites from arabidopsis cultured cells, concentration of the sample and analysis by HPLC-diode array, failed to identify any metabolites which absorb in the range 300-600 nm. This is despite known solubility of the arabidopsis anthocyanins in methanol[35]. Thus, the cells were concluded to be dormant with regard to this particular secondary metabolic pathway and measures to instigate its activity were attempted.

4.2.1.1.2 Elicitation of *Arabidopsis thaliana* cell suspension cultures

Three gateway enzymes of flavonoid biosynthesis, phenylalanine ammonia lyase (PAL), 4- coumarate Coenzyme A ligase (4CL) and chalcone synthase (CHS) are all known to be enhanced by biotic or abiotic stress treatments[47]. For example, elicitation with yeast cell wall fragments has been shown to regulate enzymes involved in secondary metabolism in *Arabidopsis thaliana*[127].

Attempts were made to instigate flavonoid biosynthesis by regulation of the external environment of the cells. Cell cultures were treated with a yeast elicitor and PAL[23], 4CL (Section 1.3) and CHS activity (Figure 4.1) were monitored and compared to control (untreated) cell cultures.

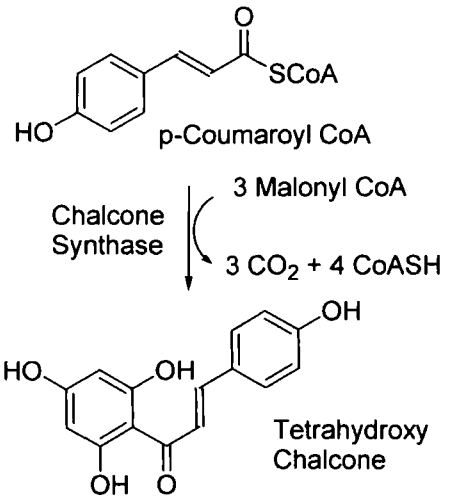


Figure 4.1 Conversion of phenylpropanoids to chalcones by chalcone synthase

PAL activity was assayed by monitoring the conversion of phenylalanine to cinnamic acid by increase in absorbance at 290 nm due to cinnamic acid. Successful induction

of PAL activity was achieved following elicitation and activity peaked after eight hours, with a 14-fold increase in activity (Figure 4.2).

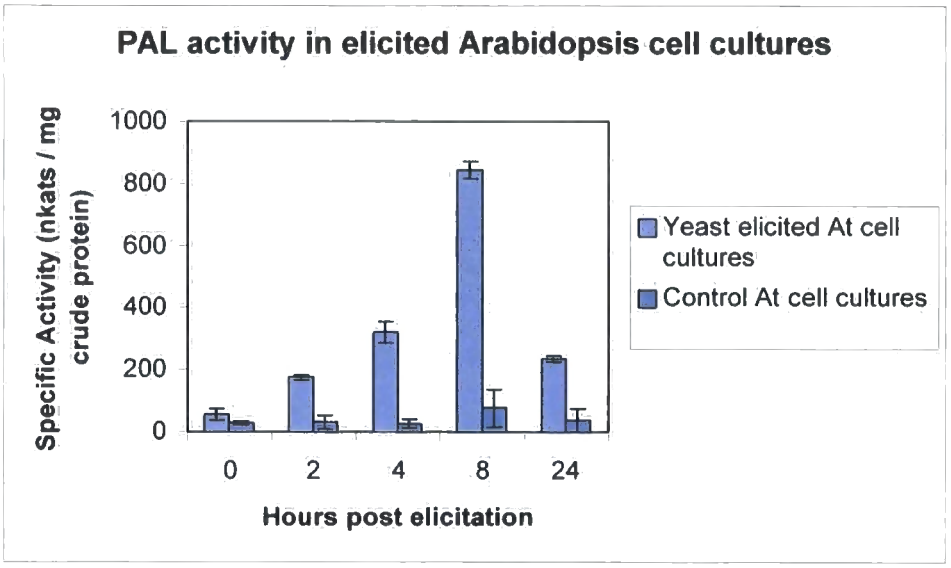


Figure 4.2 Phenylalanine ammonia lyase activity in yeast-elicited arabidopsis cell cultures. Values shown are means of two replicates, with error bars showing the range

Similarly, 4-CL activity also increased in activity 4-fold after 8 hours of elicitor treatment (Figure 4.3).

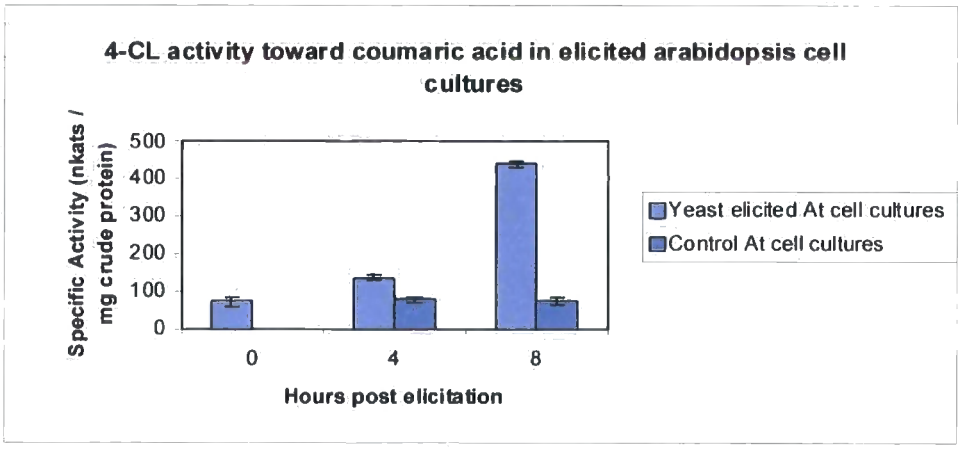


Figure 4.3 4-Coumarate coenzyme A ligase activity in yeast-elicited arabidopsis cell cultures. Values shown are means of two replicates, with error bars showing the range

PAL and 4CL activities were successfully up-regulated by elicitor treatment above the basal levels of enzyme activity required to synthesise lignin and sinapoyl malate. However, the chalcone synthase assay, monitored by HPLC and mass spectrometry, was unsuccessful and no chalcone synthase activity was shown in protein extracts from either control or elicitor-treated cells.

Equally, no UV-Vis absorbant metabolites could be observed in methanolic extracts as a consequence of elicitation. Thus all percieved avenues of ascertaining active flavonoid metabolism in arabidopsis cell cultures were exhausted.

4.2.1.2 *Arabidopsis thaliana* plants

An initial metabolic profile (methanol soluble) was obtained from wild type *Arabidopsis thaliana* plants to establish endogenous levels of flavonoid metabolism. However, no characteristic UV-Vis or mass spectra could be observed across the metabolic profile, despite the use of similar growth conditions to those described for the previous characterisation of the major anthocyanin metabolites[35].

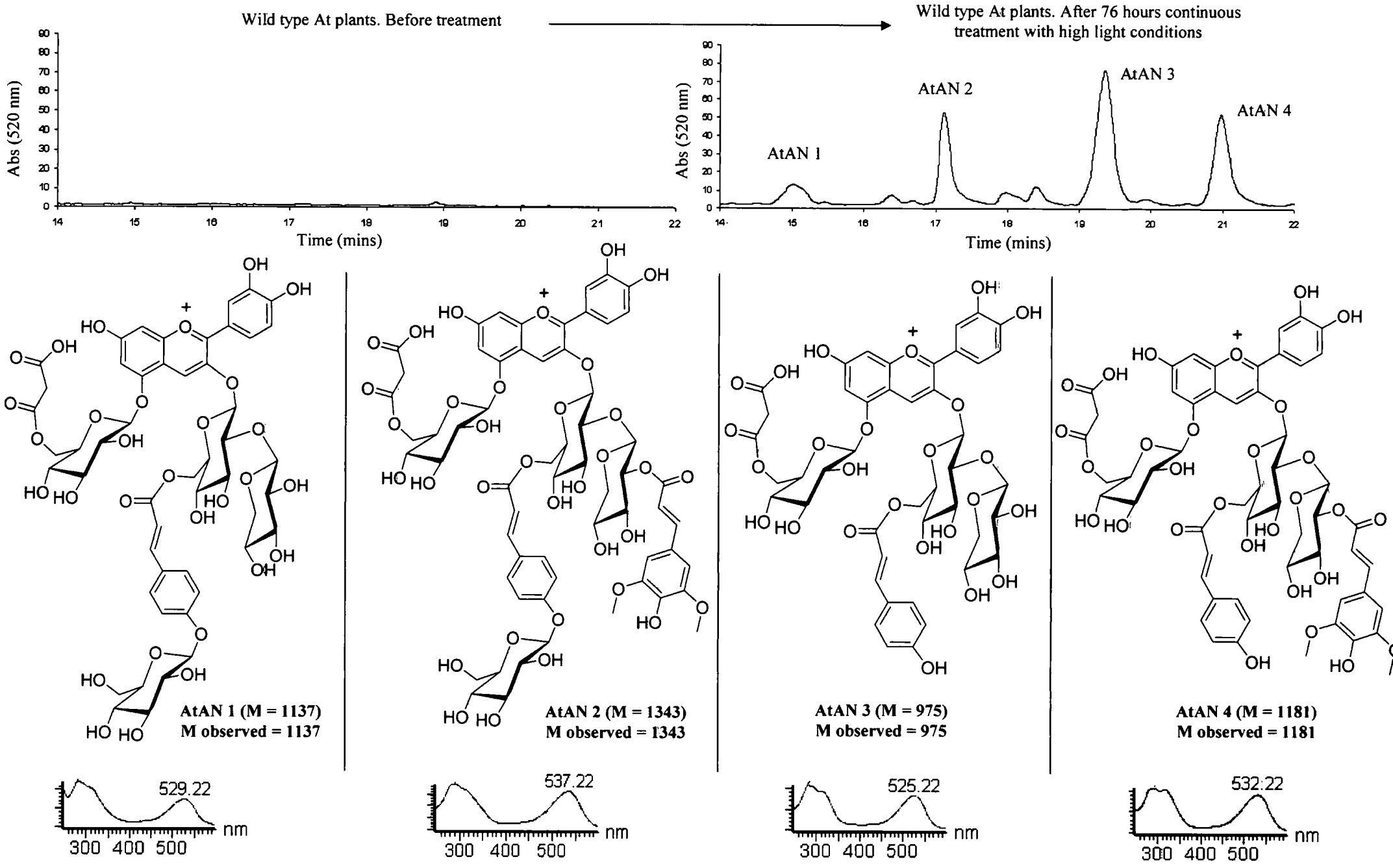
High intensity light sources have been effective in producing an increased level of flavonoid biosynthesis[48], due to the role of the metabolites as photoprotectants and antioxidants (Section 1.3.1.3). A treatment of prolonged (> 16 hours) exposure to the highest available levels of artificial visible light (160 μ einstein m^{-2} s^{-2}) was devised. After 76 hours of exposure, a red tinge was observed in vascular regions of the plants and 4 metabolites were observed to have accumulated, detected by HPLC with UV-Vis analysis at 520 nm (Figure 4.4). All four of these compounds displayed UV traces

characteristic of polyphenolic anthocyanins (Section 1.3.1.1), with absorption in the visible region arising from the extended conjugation of the oxycation[34]. Mass spectrometry (ES^+), mass fragmentation patterns and UV-Vis spectrophotometry allowed tentative identification of the four metabolites as the previously described major anthocyanin in *Arabidopsis thaliana* (Section 1.3.1.2) and its precursors[35] (Figure 4.4 and Appendix A).

Structural information for each of the anthocyanins was derived from MS/MS (ES^+) fragmentation (Appendix A). Primarily, a prevalent mass ion of 287 Da corresponded to cyanidin $(M)^+$ for each metabolite. Fragmentation of 3-*O*-glucose and it's substituents from cyanidin gave 535 Da corresponding to cyanidin 5-*O*-(6-*O*-malonyl)glucose $(M)^+$ or 449 Da when malonyl was not present originally (cyanidin 5-*O*-glucose $(M)^+$). Mass ions were also observed for sinapoyl fragments at 207 Da (M^+) when present. For each compound, the mass of the 3-*O*-glucose and any substituents was observed from cleavage of 5-*O*-glucose.

Each parent mass ion and UV-Vis spectrum were comparable to previous characterisation carried out by Bloor *et. al.*[35]. Whilst no NMR characterisation of these metabolites was performed in this study, the regiochemistry of the metabolites shown in figure 4.4 was based upon NMR evidence obtained by Bloor *et. al.* for AtAN 2[35].

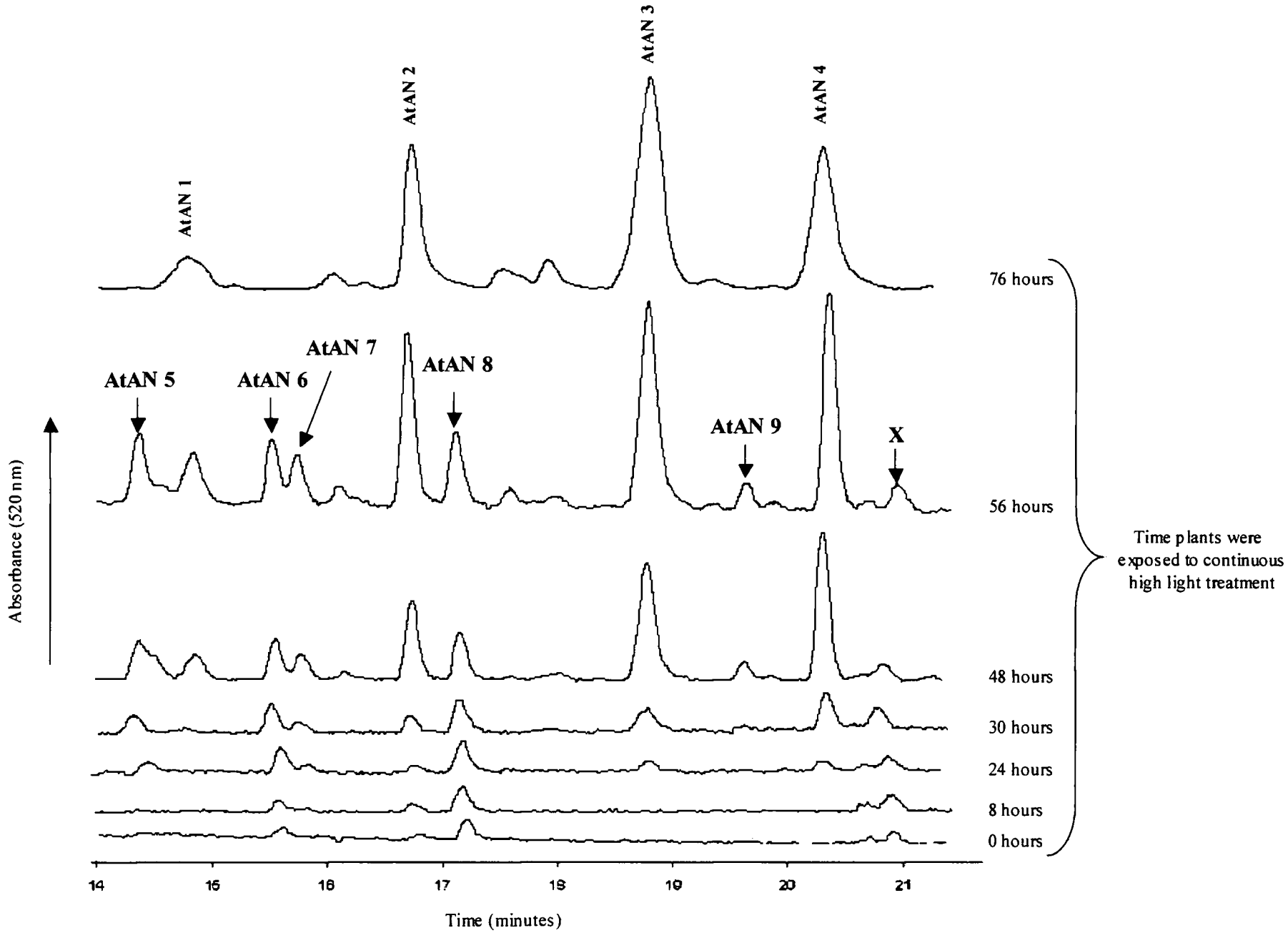
Figure 4.4 (overleaf) The four major anthocyanins accumulated in *Arabidopsis thaliana* after a 76 hour treatment of high-light conditions


Wild type At plants. Before treatment
Wild type At plants. After 76 hours continuous treatment with high light conditions
Abs (520 nm)
Time (mins)
AtAN 1
AtAN 2
AtAN 3
AtAN 4
AtAN 1 (M = 1137)
M observed = 1137
AtAN 2 (M = 1343)
M observed = 1343
AtAN 3 (M = 975)
M observed = 975
AtAN 4 (M = 1181)
M observed = 1181
529.22
537.22
525.22
532.22
nm
300 400 500

AtAN 2 was originally thought to be the terminal anthocyanin product, in accordance with the literature[35], and the other metabolites were presumed to be precursors to the most substituted metabolite. However, the four metabolites could not be placed in a sequential order of biosynthesis, with AtAN 3 the apparent precursor to both AtAN 1 and AtAN 4. This divergence was therefore indicative of the four metabolites being end-products in their own right and this assumption was further backed up by the malonylation of each metabolite, which is often seen to effectively 'cap' flavonoids in the final step of their biosynthesis[18,82,128].

If this was the case, then the acyl acceptors upon which the aromatic acylating groups were imparted were still unknown. Although, there would potentially be four anthocyanin malonyl transferases present in *Arabidopsis thaliana*, of which one has previously been characterised[44]. Therefore an attempt to identify the precursors of the four anthocyanins was required and a more detailed time course of high-level light treatment was undertaken in an attempt to establish if any precursors to aromatic acylation could be observed (Figure 4.5).

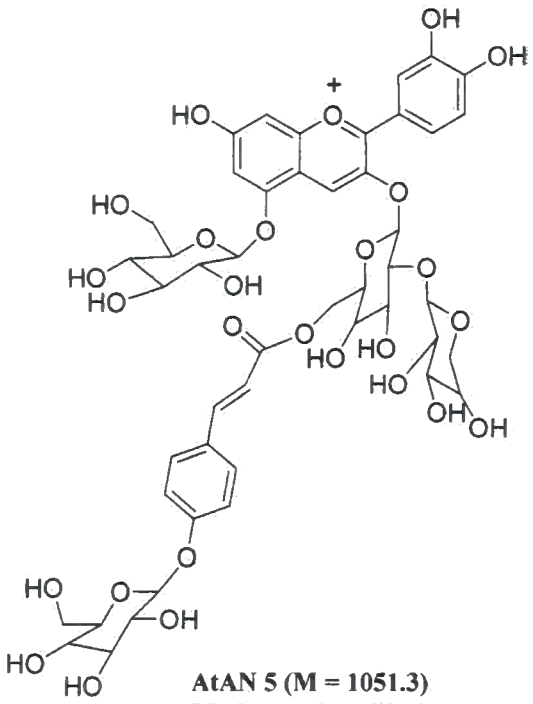
Figure 4.5 (overleaf) Metabolite profiles (520 nm) of anthocyanins in *Arabidopsis thaliana* taken at intervals over a 76-hour period of high-light treatment. Plant extracts were normalised to one gram of plant tissue per mL of extract to allow comparisons to be made. AtAN 1-4 were accumulated as before, whereas AtAN 5-9 were new metabolites to this study and metabolite X was found to not be an anthocyanin product


AtAN 1
AtAN 2
AtAN 3
AtAN 4
AtAN 5
AtAN 6
AtAN 7
AtAN 8
AtAN 9
X
Absorbance (520 nm)
76 hours
56 hours
48 hours
30 hours
24 hours
8 hours
0 hours
Time plants were exposed to continuous high light treatment
14
15
16
17
18
19
20
21
Time (minutes)

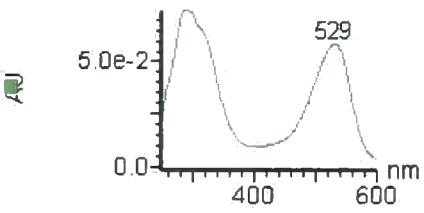
Several previously unseen metabolites were seen to temporarily accumulate over a 76-hour period, either prior to, or during, the biosynthesis of AtAN 1-4. At time points after 76 hours no further accumulation of anthocyanins was observed, although solely AtAN 1-4 remained present. However, plants exposed to continuous high-light treatment for periods longer than 76 hours became increasingly wilted and it was postulated that the cessation of anthocyanin biosynthesis was a side-effect of the declining condition of the plants. Alternatively, anthocyanin concentrations may have reached capacity, where further accumulation would be toxic toward the plants.

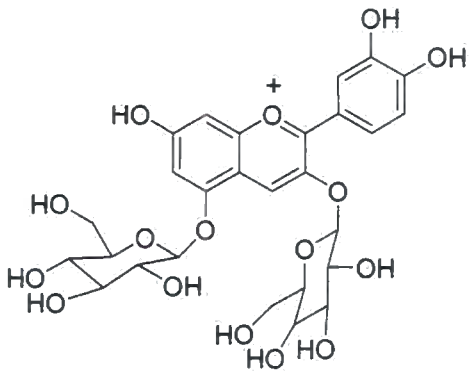
The metabolites denoted AtAN 5-9 in figure 4.5 were found to be structurally-related to AtAN 1-4 based upon MS/MS (ES^+) fragmentation (Appendix A) and characteristic UV-Vis absorption spectra (Figure 4.6). Parent mass ions of 1051, 611, 1257, 889 and 1095 Da were found for AtAN 5, AtAN 6, AtAN 7, AtAN 8 and AtAN 9 respectively (Figure 4.6). The identities of the anthocyanins were deduced from these values, mass fragmentation data, UV-Vis spectra and comparison of this data to prior characterisation of arabidopsis anthocyanins [34,44] (Figure 4.19 and Appendix A). Whilst no NMR characterisation of these metabolites was performed in this study, the regiochemistry of the metabolites shown in figure 4.6 is based upon the structural characterisation of AtAN 2, for which ^1H NMR evidence was obtained by Bloor *et. al.*[34]. The metabolite labelled X in figure 4.5 did not possess the same mass fragmentation pattern as the anthocyanins and did not bear a characteristic UV-Vis absorption spectrum and was therefore not characterised further.

Figure 4.6 (overleaf) Proposed structures and UV-Vis spectra of AtAN 5, AtAN 6, AtAN 7, AtAN 8 and AtAN 9

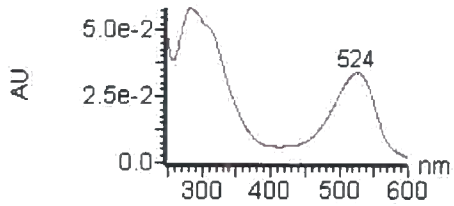


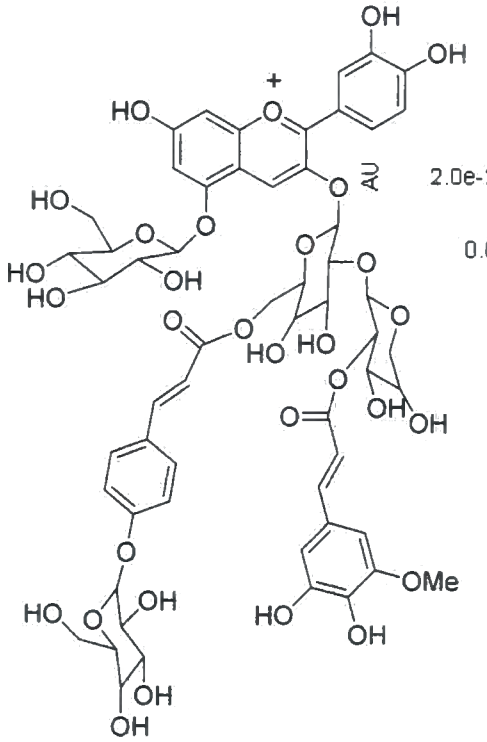
AtAN 5 (M = 1051.3)
M observed = 1051.3



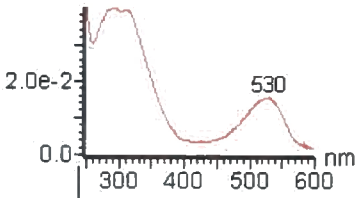


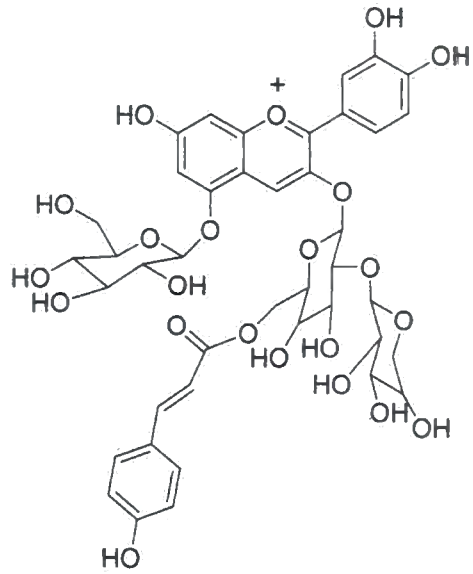
AtAN 6 (M = 611.2)
M observed = 611.2



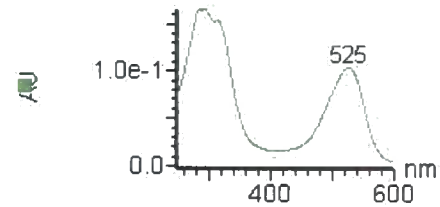


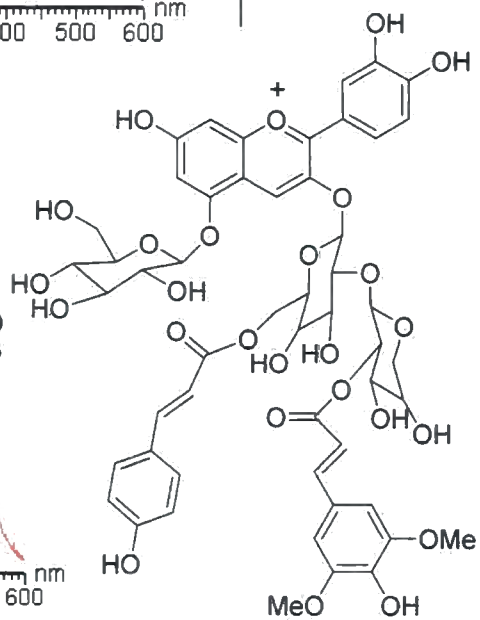
AtAN 7 (M = 1257.4)
M observed = 1257.4



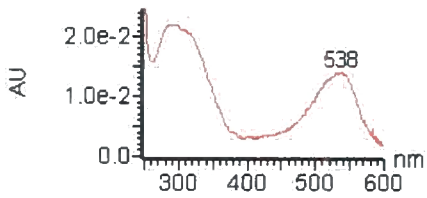


AtAN 8 (M = 889.2)
M observed = 889.2





AtAN 9 (M = 1095.3)
M observed = 1095.3



To date, AtAN 1, AtAN 2, AtAN 4 (Bloor *et. al.*), AtAN 5 and AtAN 7 (D'Auria *et. al.*) have been reported in the literature. Each of the anthocyanins AtAN 5-9 were concluded not to be degradation products of AtAN 1-4 as they could not be identified by mass spectrometry in duplicate experiments of plants treated with 76 hours of high-light, despite these plants having the highest quantity of metabolites.

Thus AtAN 3, 6, 8 and 9 were proposed to be previously unknown metabolites, which accumulated under the unnatural-upregulation of the anthocyanin pathway. Their accumulation was analysed alongside that of AtAN 1, 2, 4, 5 and 7 to draw *tentative* conclusions concerning the order of anthocyanin biosynthesis.

Firstly, AtAN 6 and AtAN 8 were unique, accumulating early in anthocyanin biosynthesis and approximately doubled in quantity within 24 hours of high-light treatment. Therefore these metabolites were thought to be the precursors to the latter stages of anthocyanin biosynthesis. However, AtAN 6 was not the precursor of AtAN 8, as the former anthocyanin would have required simultaneous modification by transfer of both xylosyl and coumaroyl moieties, in order to produce the latter. Thus an intermediary anthocyanin, with a mass of 757 Da if coumaroyl transfer was the primary modification or 743 Da if xylosylation occurred first, was expected to accumulate within the same time frame as AtAN 6 and 8. However, single-ion monitoring of mass spectral data did not identify anthocyanins of either mass. Therefore the intermediary compound was concluded to be present only transiently during biosynthesis, which suggests purposeful accumulation of its precursor, AtAN 6. Contradictory to this however, was that AtAN 6 was only observed whilst

biosynthesis of anthocyanins was active, suggesting that a relatively slow rate of coumaroylation and / or xylosylation caused accumulation of AtAN 6.

AtAN 8 appeared to be modified by three different enzymes, a glucosyl transferase in biosynthesis of AtAN 5, a sinapoyl acyltransferase in biosynthesis of AtAN 9 and a malonyl acyltransferase in biosynthesis of AtAN 3. Equally, AtAN 5, AtAN 7 and AtAN 9 were also thought to be acyl acceptors for malonyl transferase(s) during biosynthesis of AtAN 1, AtAN 2 and AtAN 4 respectively.

However, there was ambiguity over the biosynthesis of AtAN 7, which could have been committed by either a sinapoyl transferase (metabolic route 1) or a glucosyl transferase (metabolic route 2)(Figure 4.7). It was observed that AtAN 9 was absent after 30 hours of high-light treatment, when all other non-terminal anthocyanins were present, and was only observed to accumulate after 48 hours. This implies that this particular anthocyanin was not a precursor to further modification, malonylation aside, and therefore metabolic route 1 (Figure 4.7) was thought to be the active pathway.

These tentative conclusions were collated as a proposed order of anthocyanin biosynthesis (Figure 4.7).

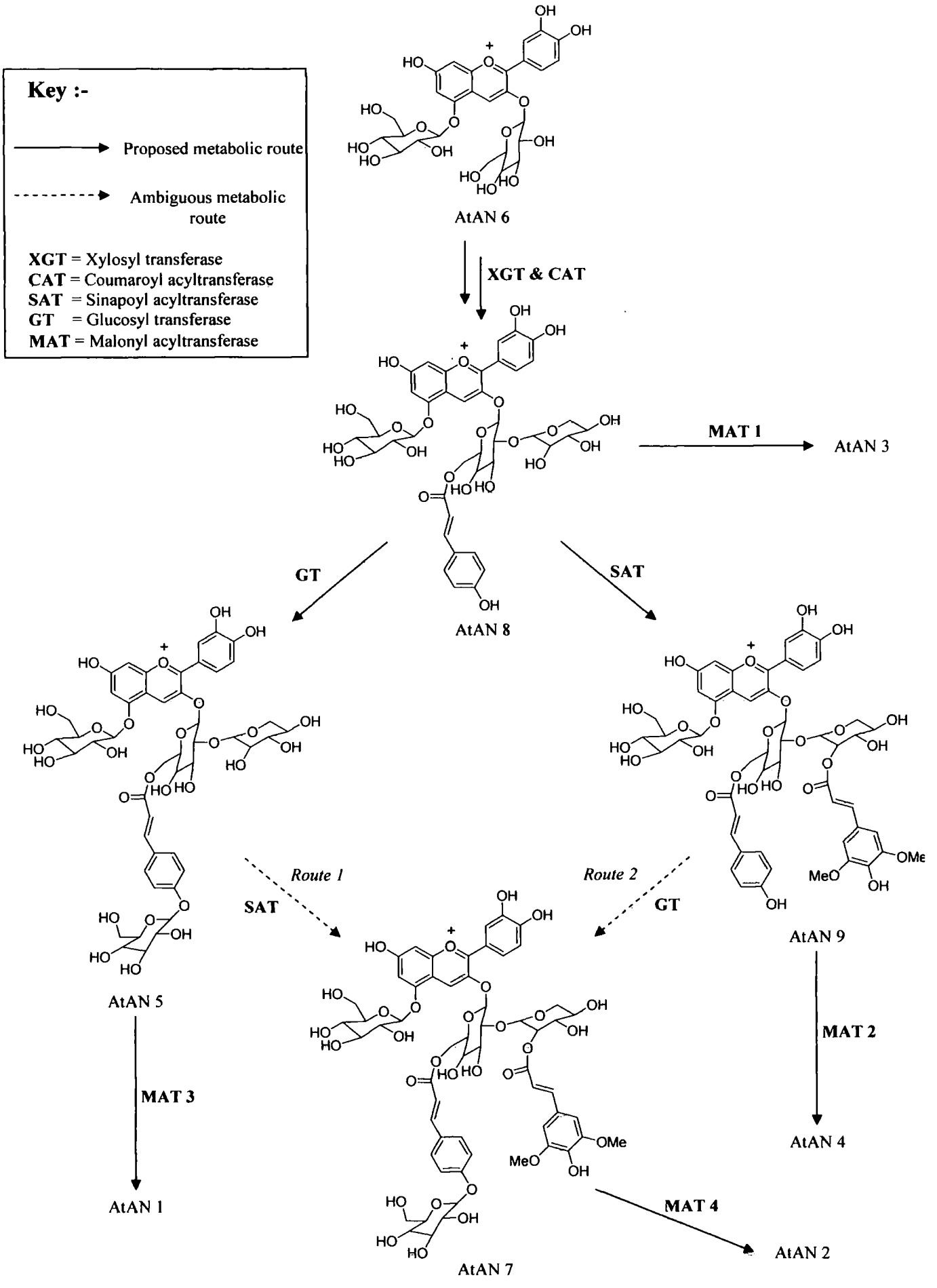


Figure 4.19 A proposed order of anthocyanin biosynthesis in *Arabidopsis thaliana*

Thus AtAN 8 and AtAN 5 were proposed to be acyl acceptors for a sinapoyl acyltransferase(s) in *Arabidopsis thaliana*. Although the identity of the acyl acceptor for the coumaroyl acyltransferase could not be determined it was proposed to be either AtAN 6 or a xylosylated derivative of AtAN 6. Additionally, the time period during which both of these aromatic acyltransferases were expressed and found to be active following initiation of continuous high-light treatment, was inferred from the rates of accumulation of the associated metabolites in figure 4.5. This suggested that a coumaroyl transferase was active between 8 and 56 hours of continuous high-light treatment, whereas sinapoyl transfer was only observed between 24 and 56 hours.

The characterisation of anthocyanin biosynthesis in *Arabidopsis thaliana* undertaken in this study, has led to identification of potential acyl acceptors for both sinapoyl and malonyl acyltransferases and the identity of the acyl acceptor for coumaroyl acyltransfer is tentatively proposed to be one of two possible structures. The conditions required for up-regulation of anthocyanin biosynthesis in arabidopsis plants were also determined, which will allow for future biochemical characterisation of the respective acyltransferase enzymes.

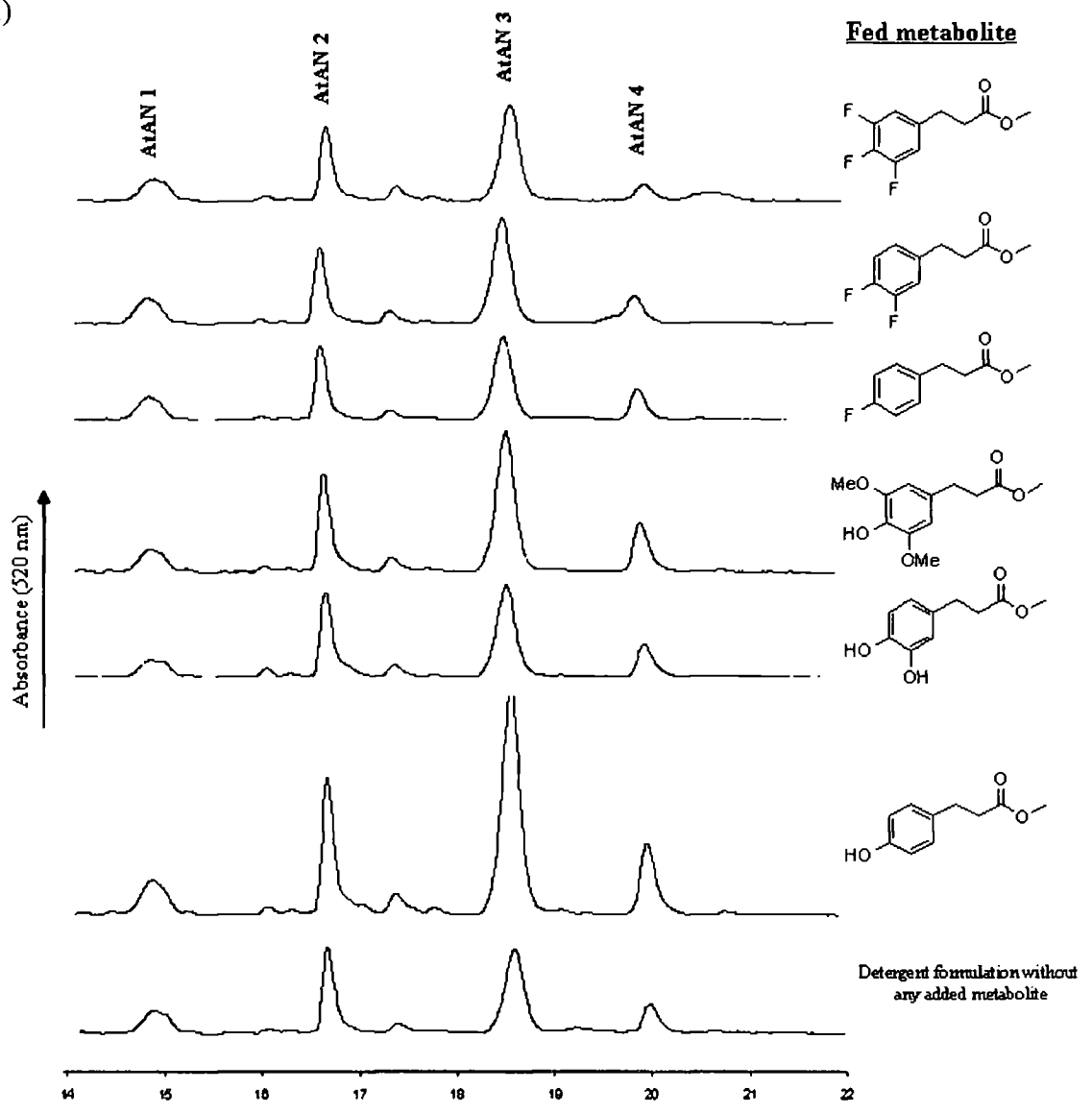
4.2.2 Feeding studies with exogenously-supplied phenylpropanoids

4.2.2.1 Feeding studies in *Arabidopsis thaliana*

Having established conditions under which acylation of flavonoids was active we were able to ascertain the potential of 4CLs and CADATs in arabidopsis plants to incorporate exogenously-fed phenylpropanoids into the biosynthesis of acylated anthocyanins. To this end, arabidopsis plants were sprayed with detergent

formulations containing phenylpropanoids (1 mM)(Compounds 2-7) after 4 hours of high-light treatment; prior to up-regulation of anthocyanin biosynthesis. The solution was applied at a rate of 25 mL per seed tray, with each tray containing 24 plants. The treated plants were harvested and weighed after 76 hours of high-light treatment and their metabolic extracts analysed by HPLC-MS to ascertain any effect of the exogenous phenylpropanoids (Figure 4.8). The concentrations of anthocyanin metabolites in each study, was calculated using the extinction coefficient of cyanidin 3-5-*O*-diglucoside at 520 nm as a close approximation to the cyanidin-based metabolites accumulated (Appendix C).

A)



B)

	Concentration of anthocyanin (nmoles/gram plant tissue)						
Anthocyanin	*Control*	*Coumaric**	*Caffeic*	*Sinapic**	*4, Fluoro cinnamic*	*3, 4, Difluoro cinnamic*	*3, 4, 5, Trifluoro cinnamic*
AtAN 1	**4.3** +/- 0.5	**6.1** +/- 0.1	**4.2** +/- 0.3	**5.8** +/- 0.4	**4.3** +/- 0.2	**4.7** +/- 0.8	**4.1** +/- 0.6
AtAN 2	**8.3** +/- 0.6	**13.4** +/- 0.7	**8.4** +/- 1.0	**11.0** +/- 0.5	**7.9** +/- 1.2	**7.6** +/- 1.1	**8.9** +/- 0.9
AtAN 3	**15.6** +/- 0.8	**36.1** +/- 1.7	**13.7** +/- 1.3	**25.0** +/- 1.8	**12.9** +/- 1.4	**14.9** +/- 0.9	**13.9** +/- 1.9
AtAN 4	**4.0** +/- 0.6	**9.3** +/- 1.1	**4.3** +/- 0.7	**7.6** +/- 0.9	**3.9** +/- 0.4	**4.4** +/- 0.5	**3.9** +/- 0.7

Figure 4.8 (A) Metabolite profiles of high-light induced *Arabidopsis thaliana* plants fed with various phenylpropanoids. (B) Analysis of the peak area at 520 nm of endogenous anthocyanins and conversion into nmoles/gram of tissue (mean of three biological replicates + / - standard deviation) (* = values significantly greater than control ($P < 0.05$)

Significant differences to observations under control conditions were observed in the accumulation of anthocyanins in plants treated with continuous light for 76 hours, when they were fed with coumaric and sinapic acids respectively. Only feeding with these phenylpropanoids was seen to have an effect upon endogenous anthocyanin content and no anthocyanin structures were found to be acylated with any other phenylpropanoid variant.

The effect of elevated coumaric and sinapic acid concentrations upon the accumulation of the terminal anthocyanins in arabidopsis was calculated as a percentage relative to the control (Figure 4.9).

	Increase in anthocyanin content when fed with phenylpropanoid (%)	
Anthocyanin	*Coumaric acid*	*Sinapic acid*
AtAN 1	41	34
AtAN 2	61	33
AtAN 3	132	60
AtAN 4	131	90

Figure 4.9 The effect of elevated coumaric acid and sinapic acid concentration upon the relative concentration of AtAN 1, AtAN 2, AtAN 3 and AtAN 4

Both sinapic acid and coumaric acid were found to cause the accumulation of each of the major anthocyanins, which indicated that the plants' response to the exogenous phenylpropanoids was to increase anthocyanin accumulation to conjugate the excess metabolites. However, blanket up-regulation of terminal anthocyanin products was not observed and malonylated metabolites arising from anthocyanin products

immediately downstream of the proposed coumaroyl (AtAN 3) and sinapoyl (AtAN 4) acyltransfer reactions (Figure 4.7) were the most influenced metabolites.

Feeding studies with methyl coumarate and methyl sinapate resulted in a modest increase in each of AtAN 1-4, but had particular influence upon the malonylated products of their respective proposed acyltransfers. Coumaric and sinapic acids were fed at an approximate rate of 350 nmoles/g (based upon the average weight of harvested plants being 1 g +/- 0.3). Thus approximately 6.7 % of fed coumaric acid was incorporated into anthocyanin biosynthesis and approximately 2.6 % of sinapic acid.

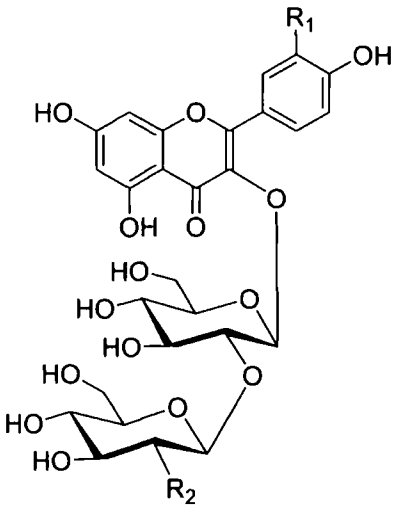
Modest incorporation of the endogenous phenylpropanoids coumaric and sinapic acid was achieved, whilst non-natural phenylpropanoids were not incorporated into flavonoid biosynthesis when applied to Arabidopsis leaves. The reason(s) for lack of incorporation of non-natural phenylpropanoids was unclear. Further studies were undertaken in *Petunia hybrida*, which produces a diverse array of aromatically acylated flavonoids (Section 1.3.2).

4.2.2.2 Feeding studies in *Petunia hybrida*

4.2.2.2.1 Identification of flavonol acyl acceptors in *Petunia hybrida*

Aromatic acylation with phenylpropanoids is the terminal modification of many flavonoids found in *Petunia hybrida*. Petunia has previously been shown to possess an array of aromatically acylated quercetin and kaempferol diglucosides, which are found in the leaves[37].

Metabolite profiling (HPLC-DAD-MS) of petunia leaf tissue led to the observation of each of the compounds previously characterised by Bloor *et. al.*[37], through UV-Vis detection, mass spectrometry (ES^+) and the corresponding fragmentation (Appendix B). The metabolites kaempferol 3-*O*-(6''-*O*-caffeoyl)diglucoside (KDG-Caf), kaempferol 3-*O*- (6''-*O*-feruloyl)diglucoside (KDG-Fer), quercetin 3-*O*-(6''-*O*-caffeoyl)diglucoside (QDG-Caf), kaempferol 3-*O*- diglucoside (KDG), quercetin 3-*O*-diglucoside (QDG) and rosmarinic acid (R) were found (Figure 4.10)[37]. The regiochemistry of the compounds shown in figure 4.10 is based upon the characterisations carried out by Bloor *et. al.*[37].



Metabolite	R_1	R_2
QDG (**1**)	OH	OH
KDG (**2**)	H	OH
QDG-Caff (**3**)	OH	OCaf
KDG-Caff (**4**)	H	OCaf
KDG-Fer (**5**)	H	OFer

R =

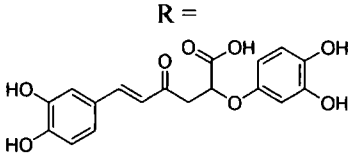


Figure 4.10 The UV-Vis absorbing metabolites identified in *Petunia hybrida*

4.2.2.2.2 Light-mediated regulation of flavonoid acylation in petunia

The relative amounts of the metabolites previously identified (Section 4.2.2.1) were compared in plants treated with continuous high-level illumination for 24 hours (160 μ einstein m^{-2} s^{-2}) with plants grown under a conventional (80 μ einstein m^{-2} s^{-2}) lighting regime for 12 hours in 24 (Both were treated with blank detergent). The relative abundance of the individual compounds were similar in each of the extracts, although the overall flavonolic content was approximately tripled by the high light treatment (Figure 4.11). Spraying with detergent had no significant effect on the accumulation of the metabolites, when compared with unsprayed controls, so the results shown are derived from plants sprayed with formulation.

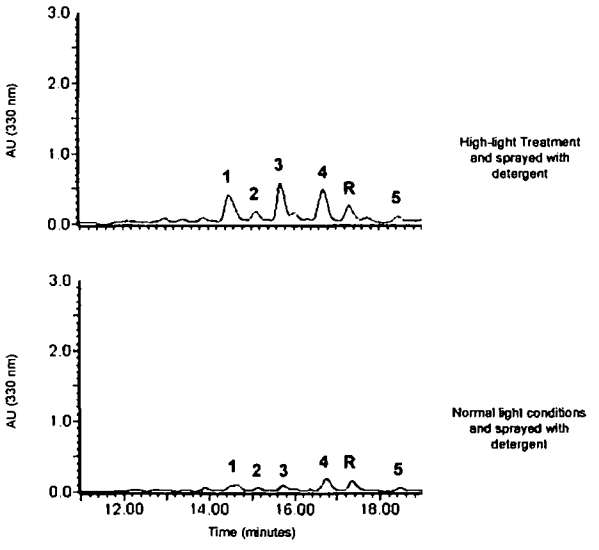


Figure 4.11 Comparison of normalised metabolite extracts (1 g/mL) from petunia plants exposed to high-light conditions for 24 hours and from plants grown under normal conditions, with UV-Vis detection at 330 nm

4.2.2.2.3 Feeding with methyl esters of phenylpropanoids in *Petunia hybrida*

Following only modest incorporation of fed-phenylpropanoids into flavonoid biosynthesis in arabidopsis, an attempt to increase the efficiency of uptake of phenylpropanoids, and therefore increase substrate concentration relative to existing CL / CADAT metabolism, was made in petunia. To this effect, synthesis of carboxyl esters is one method that has previously been employed to enhance translocation of herbicides bearing polar groups, such as carboxylic acids, into plant cells[129]. Once internalised proherbicide carboxyl esters are substrate to hydrolytic enzymes, which yield the active herbicide[129]. Such hydrolases are more often found to have activity toward small alkyl esters, such as methyl, than bulky alkyl chains[130]. Therefore the effect of feeding methyl esters of phenylpropanoids, as opposed to carboxylic acids, was assessed in petunia with respect to flavonol acylation.

Methyl esters of phenylpropanoids (compounds 2-8) were synthesised (Section 3.2.1.4). Petunia plants were sprayed with two separate detergent formulations in duplicate, one containing caffeic acid - an endogenous acylation precursor in flavonoid metabolism - (compound 3)(1mM) and the other containing methyl caffeate (compound 19)(1 mM). Each plant was harvested after a further 8 hours. An equivalent mass of leaf tissue was taken from each study, rinsed in methanol (2 x 5 mL) to remove extraneous metabolite and was finally extracted in methanol. The methanolic extracts were concentrated to 100 μL and subjected to analysis by HPLC with detection at 287 nm (Figure 4.12).

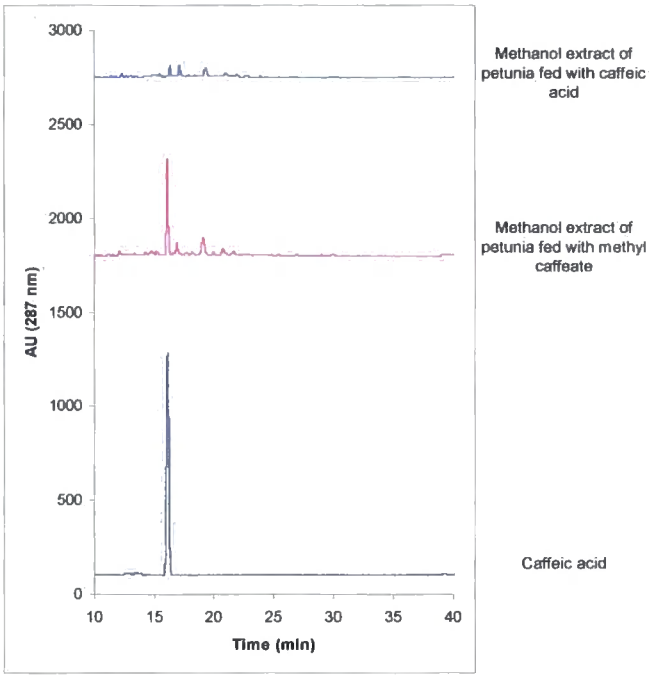


Figure 4.12 Metabolite profiles of petunia leaves fed with methyl caffeate, as opposed to caffeic acid, and with comparison to a caffeic acid standard

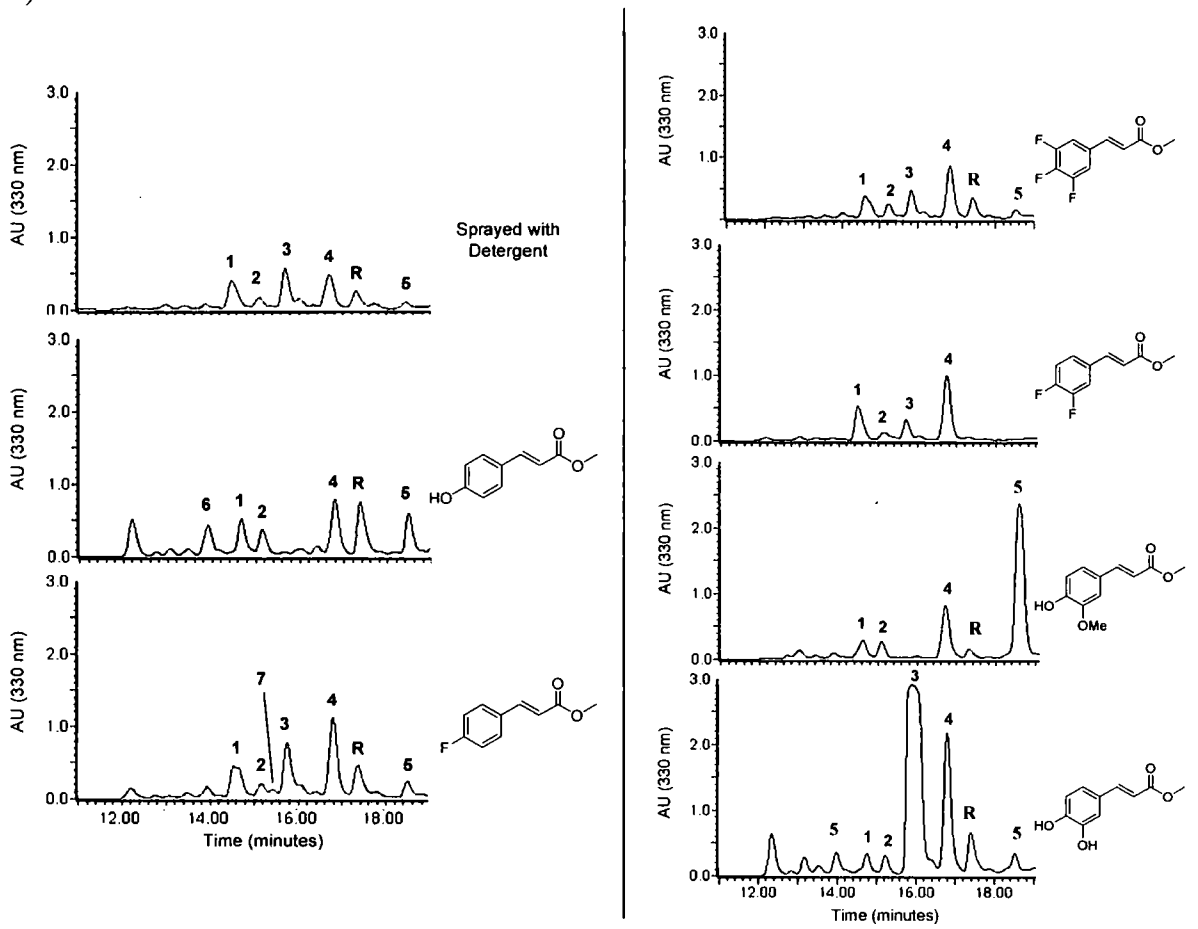
Identical studies were undertaken in petunia for each of the other phenylpropanoids and their respective methyl esters (Appendix D). In each case the carboxylic acid and not the methyl ester was found to be present, indicating that the methyl ester was efficiently hydrolysed by endogenous esterase activity. Phenylpropanoids were accumulated in far greater quantity when fed with the methyl ester in comparison to the carboxylic acid. The effect of increased phenylpropanoid concentration upon incorporation of phenylpropanoids into flavonol acylation will be assessed in the following section.

4.2.2.2.4 Effect of feeding methyl phenylpropanoid esters upon flavonolic acylation in *Petunia hybrida*

Methyl esters of known flavonol acyl donor precursors coumaric, caffeic and ferulic acid (Compounds 18-20), and non-natural methyl fluorocinnamate compounds (Compounds 22-24), were applied to leaves of petunia in the same manner as described for the feeding studies in *Arabidopsis thaliana* (Section 4.2.2.1). The plants were extracted into methanol and analysed by HPLC-UV-Vis with detection at 330 nm (Figure 4.13). The individual metabolites were then quantified based on their absorbance at 330 nm, using a standard curve prepared from quercetin-3-*O*-diglucoside (Appendix C). By quantifying at 330 nm the concentration of the parent flavonoids was determined, with minimal interference from the aromatic acylating groups. Two new metabolites were accumulated in plants sprayed with methyl coumarate and methyl 4-fluorocinnamate, metabolite 6 and metabolite 7 respectively. The characterisation of these compounds will be discussed in section 4.2.2.4.1.

Figure 4.13 (overleaf) The effect of the indicated exogenously-fed methyl esters of phenylpropanoids upon the accumulation of acylated metabolites in *Petunia hybrida*. (A) HPLC profiles with detection at 330 nm of treated plants and (B) quantification of significantly different peak areas for comparison. Plants sprayed with detergent were grown under ambient lighting (=Con LL) or were sprayed with formulated treatments then exposed to high level lighting for 24 h. (=Con HL). Values are the mean of two biological replicates +/- standard deviation (* = values significantly greater than control HL ($P < 0.05$))(Δ = occured at the same retention time as fed caffeic acid).

A)



B)

	Concentration (nmol g^{-1} FW)						
			Phenylpropanoid ester fed				
Compound	*Control LL*	*Control HL*	*Coumarate*	*Fluoro cinnamate*	*Caffeate*	*Ferulate*	*Difluoro cinnamate*
QDG (1)	72 +/- 16.8	200 +/- 21.0	191 +/- 20.7	257 +/- 34.3	**127 *** +/- 12.5	148 +/- 9.4	263 +/- 14..8
KDG (2)	25 +/- 5.7	62 +/- 8.4	**124 *** +/- 12.1	59 +/- 10.3	**118 *** +/- 10.3	**110 *** +/- 6.9	64 +/- 4.8
QDG-Caff (3)	49 +/- 10.9	235 +/- 19.9	**46 *** +/- 11.9	326 +/- 39.4	**2891 *Δ** +/- 143.2	**16 *** +/- 1.2	**140 *** +/- 7.2
KDG-Caff (4)	96 +/- 5.0	338 +/- 64.0	366 +/- 11.6	511 +/- 29.0	**1018 *** +/- 24.2	469 +/- 32.2	510 +/- 13.9
KDG-Fer (5)	22 +/- 6.8	37 +/- 1.4	**239 *** +/- 9.7	81 +/- 18.7	**121 *** +/- 15.6	**1509 *** +/- 48.7	**0 *** +/- 0.0
QGGA-Coum (6)	0 +/- 0.0	0 +/- 0.0	**220 *** +/- 18.8	**59 *** +/- 7.7	**144 *** +/- 10.2	0 +/- 0.0	0 +/- 0.0
QGGA-Fca (7)	0 +/- 0.0	0 +/- 0.0	0 +/- 0.0	**15 *** +/- 0.2	0 +/- 0.0	0 +/- 0.0	0 +/- 0.0
R	79 +/- 10.8	140 +/- 28.6	**368 *** +/- 9.6	217 +/- 2.8	**328 *** +/- 23.1	**53 *** +/- 6.2	**15 *** +/- 4.5

The application of a methyl caffeate derivative led to over a 10-fold increase in the pool size of QDG-Caf (3) and a 3-fold enhancement in KDG-Caf (4). Although caffeic acid was found to co-elute with QDG-Caf (3), analysis of the UV-Vis spectrum found for this metabolite suggested that the majority of metabolite 3 was flavonolic and not caffeic acid (Figure 4.14).

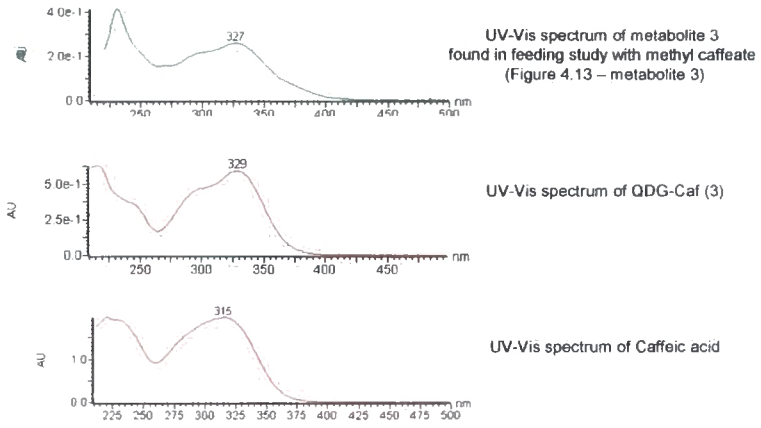


Figure 4.14 UV-Vis absorption spectrum for metabolite 3 found in feeding study with methyl caffeate in petunia and comparison to spectra of QDG-Caf (Appendix B) and caffeic acid

Feeding with methyl caffeate also increased the pool sizes of all the other UV-absorbing metabolites except for QDG. In addition to the originally identified compounds, feeding with methyl caffeate led to the appearance of a series of unidentified polar UV-absorbing metabolites. Concentrating on the major effects determined on the HPLC profiles of the flavonols, treatment with the other natural acyl donor, ferulic acid, resulted in a selective 40-fold enhancement in KDG-Fer (Figure 4.13). Feeding with methyl coumarate resulted in a large enhancement in the levels of KDG and its feruloyl ester (Figure 4.13).

Having established substantial perturbation in flavonol acylation following feeding with naturally occurring phenylpropanoids, analysis of the study was extended to

feeding of methyl esters of 4- fluoro-, 3- 4- difluoro- and 3- 4- 5- trifluorocinnamate methyl esters. The mono- and tri-fluorinated compounds did not significantly perturb the accumulation of the endogenous metabolites, while feeding with methyl 3- 4- difluorocinnamate led to a selective decline in QDG-Caf (3), KDG-Fer (5) and rosmarinic acid (R). To look for incorporation of the fluoroacyl donors into novel flavonols, the eluting metabolites were fragmented by MS and monitored for evidence of fluorcinnamoyl cationic fragments, as found for other acyl fragments (Appendix B). Using this approach we were unable to find any evidence for incorporation of 3- 4-difluorocinnamic acid (167 Da) or 3- 4- 5- trifluorocinnamic acid (185 Da). However, a compound fragmenting to release a 149 Da fragment characteristic of 4-fluorocinnamoyl cation was identified in a minor peak eluting just after KDG (Metabolite 7).

4.2.2.2.4.1 Characterisation of novel metabolites identified in petunia feeding study

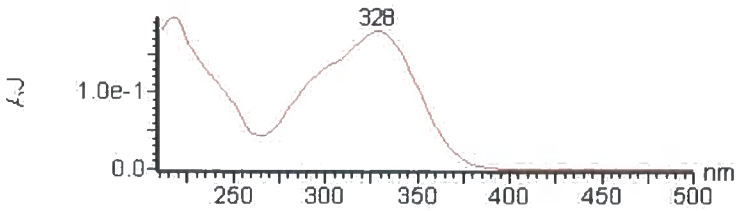
Firstly, each of the metabolites 6 and 7 had distinctive UV-Vis absorption spectra (Figure 4.15 A), which was consistent with other acylated flavonols (Appendix B). Interestingly, absorbance normally associated with aromatic acylation was at 305 nm and 291 nm for metabolite 6 and 7 respectively, indicating a different aromatic acylating group in each instance. Mass spectrometry of the metabolites (Figure 4.15 B) indicated parent masses of 787 Da and 789 Da, for metabolite 6 and 7 respectively. A difference of 2 Da is consistant with substitution of a hydroxyl for a fluorine group. Equally, positive ion (ES^+) fragments with masses with 2 Da difference were observed in positive fragments 147 Da (Metabolite 6) and 149 Da (Metabolite 7),

which is consistant with fragmentation to produce coumaroyl and fluorocinnamoyl cations respectively.

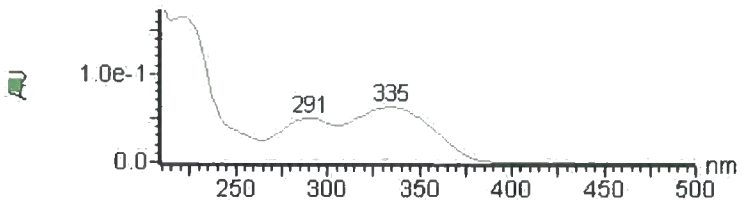
Figure 4.15 (overleaf) (A) UV-Vis spectra found for metabolite 6 and 7. (B) Mass spectra of metabolites 6 and 7 and (C) analysis of fragmentations

A) UV-Vis spectra of metabolites 6 and 7

Metabolite 6

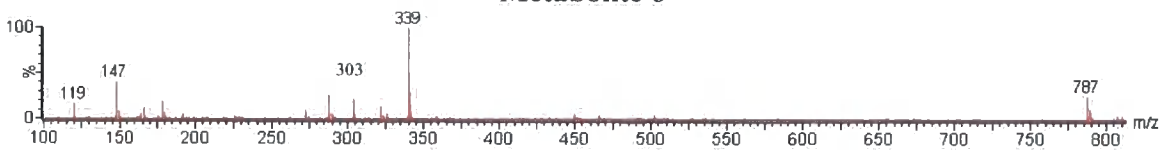


Metabolite 7

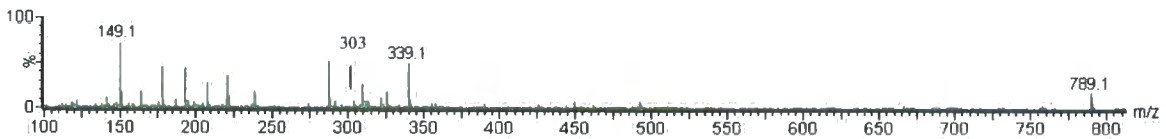


B) Mass spectra of metabolites 6 and 7

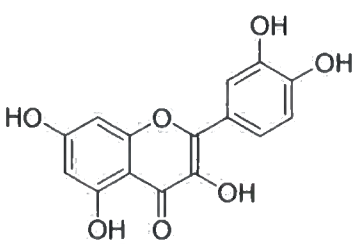
Metabolite 6



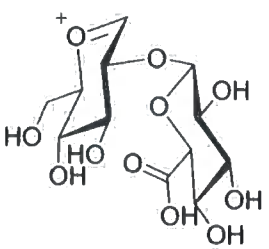
Metabolite 7



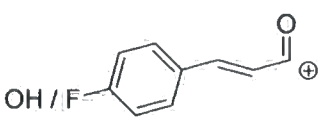
C) Proposed mass fragments



$(M + H)^+ = 303$



$(M)^+ = 339$



$(M)^+ = 147$ (OH) / 149 (F)

However, further fragments were generated that corresponded to both quercetin (M+H = 303 Da) and kaempferol (M+H = 287 Da) in each study[131] (Appendix B). In addition, only a fragment corresponding to glucosyl-glucuronide (M^+ = 339 Da) was identified, whilst the diglucoside fragment (325 Da) found in studies on QDG (1) and KDG (2) was absent(Appendix B). Therefore no firm conclusion concerning the identification of the structures was reached in either instance. Although it was thought possible that a coumaroyl (metabolite 6) or a fluorocinnamoyl (metabolite 7) group had been transferred onto a quercetin-diglycoside (glucose-glucuronyl) structure, according to the parent mass ion observed and the limited data shown.

In light of this tentative observation and the unexplained interference of 3-4-difluorocinnamic acid upon endogenous acyltransfer, it was decided that a biochemical study *in vitro* would afford the most detailed insight into enzymatic acyltransfer of fluorinated phenylpropanoids within a controlled environment. Equally, the tolerance of both 4CLs and CADATs could be assessed on an individual basis.

4.3 Discussion

4.3.1 Identification of metabolites

Metabolites in this chapter have been tentatively identified by comparison of obtained UV-Vis absorption and mass spectrometry data, to that available for similar metabolites with resolved structures - quercetin 3-*O*-diglucoside[131] and cyanidin 3-5-*O*-diglucoside[35]. Prior structural resolution of flavonoids has been achieved by performing 1 and 2D NMR experiments upon metabolites isolated in sufficient

quantity[35]. Such characterisation has afforded resolution of regioisomerism of enzymatic modifications such as glycosylation and/or acylation, geometric chemistry of *cis/trans* isomers and confirmed the stereochemistry of glycosyl residues[109]. Owing to the relatively large number of structurally-related examples of flavonoids isolated and fully characterised to date, predictability in their UV-VIS absorption and mass fragmentation has been observed. These patterns occur due to common chromophores present or predictable fragmentation events and this has recently allowed for the rapid profiling of these metabolites in plants by comparison of data generated by these techniques to that obtained for fully characterised metabolites[109,131].

To this end, the products of predictable flavonoid fragmentations in positive ion mass spectrometry have previously been proposed[109,131], which were consistent with those found in this study. Flavonoid-glycoside ether bonds were observed to fragment to produce the glycoside cation $(M)^+$ and the oxygen-retaining protonated aglycone $(M+H)^+$ respectively (Figure 4.16). Equally, phenylpropanoid-glycoside esters were frequently observed to fragment to produce the phenylpropanoyl carbocation $(M)^+$ and the oxygen-retaining protonated glycone $(M+H)^+$ respectively (Figure 4.16). Combination of these fragmentations were also observed, which each molecule to be assembled sequentially. Each of the suggested carbocations were proposed to be stabilised by a mesomeric effect and therefore favourable. Characteristic UV-Vis absorption of flavonoid chromophores was previously discussed in Chapter 1 (Section 1.3.1.1).

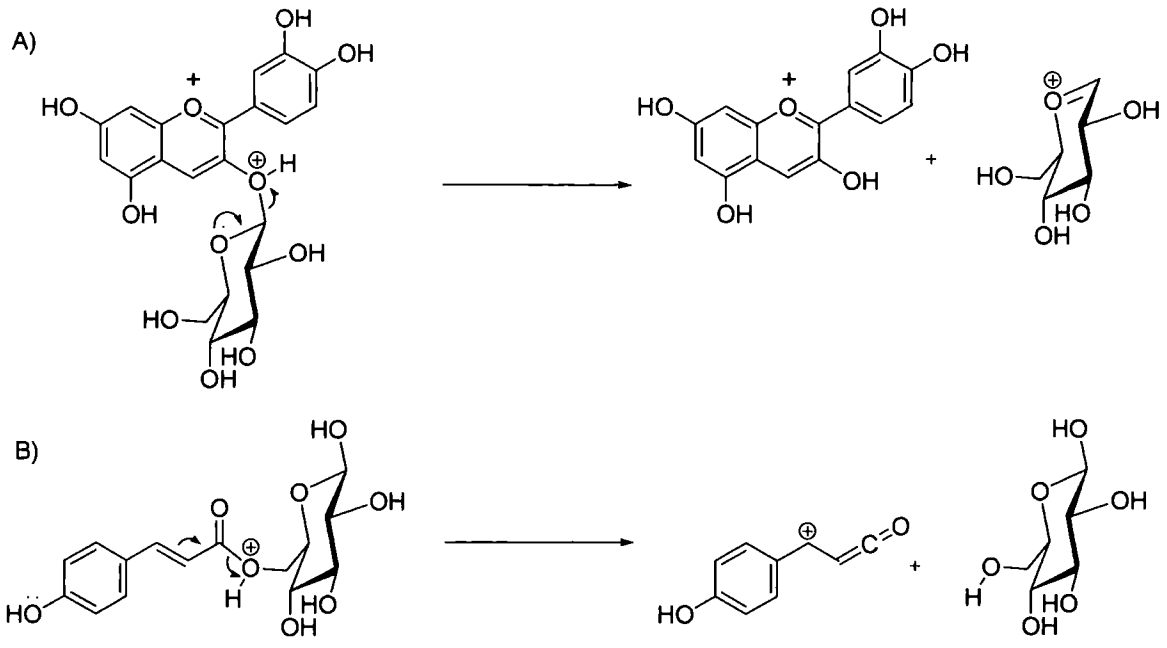


Figure 4.16 Proposed origin of fragments in positive ion electrospray mass spectrometry arising from fragmentation of A) flavonoid-glycoside ether bonds and B) phenylpropanoyl-glycoside ester bonds

Limitations of metabolite identification by mass fragmentation and UV-Vis spectroscopy are therefore untold regio- and stereochemistry. However, these variables are relatively conserved in flavonoid biosynthesis (Section 1.3). Whilst it is recognised that the region of acyltransfer would be a key factor in any potential biosynthesis of acylated products, in the scope of biocatalysis in this thesis the focus will remain upon the acyl donor in the absence of a suitably valuable product.

Equally, the flavonol, glycoside and/or acyl components of the proposed acyl acceptors for aromatic acyltransfer in arabidopsis (AtAN 5, AtAN 8 and AtAN 6 or AtAN 6-Xyl) have been identified, whilst the regio- and stereochemistry of the structures are based upon previous characterisation of the parent anthocyanins and were not known for each individual compound[35]. Precedent in accumulation of

multiple anthocyanins comprising the same aglycone and glycoside residues suggests that the region of glycosylation and the isomeric form of the glycoside are conserved[32]. This scenario is assumed in this study, particularly in light of the novel metabolites accumulating immediately prior to the observed accumulation of the structurally-resolved metabolite AtAN 2 and not being present as degradation products in extracts of plants that had accumulated maximal levels of anthocyanin (Section 4.2.1.2).

Other, more conclusive, methods have been employed to establish metabolic precursors and pathways and could be utilised to conclusively establish the course of the arabidopsis anthocyanin pathway in further studies. Of relevance, an anthocyanin malonyltransferase from *Arabidopsis thaliana* was characterised by a reverse-genetics approach by D'Auria *et. al.*[44]. They observed that this particular transferase knockout line prohibited the accumulation of both AtAN 1 and AtAN 2, although the effect upon levels of AtAN 3 and AtAN 4 was not reported. They concluded that the malonyl transferase was responsible for malonylation of all of the anthocyanins found in arabidopsis and therefore had a large tolerance of vastly differing acyl acceptors and was therefore highly unusual. However, biochemical characterisation of the enzyme toward malonyl CoA and AtAN 5 and AtAN 7 respectively, demonstrated that the enzyme was significantly biased toward AtAN 7. Therefore it would be of interest to establish whether there is a second or third malonyl transferase.

4.3.2 Use of methyl phenylpropanoid esters in feeding studies

Methyl esters of phenylpropanoids were found to be incorporated with far greater efficiency than their carboxylic acid analogues. After incorporation of the methyl ester, only the carboxylic acid could be observed in the plant extract in each study. Hydrolysis of the ester after internalisation can be likened to the mechanism exploited to allow internalisation of proherbicide esters, before hydrolysis yields the active herbicide[129]. However, the lack of incorporation of the underivitised phenylpropanoids was unexpected. Particularly, in light of the modest effect fed coumaric and sinapic acids had upon anthocyanin metabolism in arabidopsis. One possible explanation for this could be heightened-channeling of these metabolites into lignin biosynthesis[132] (Figure 4.17).

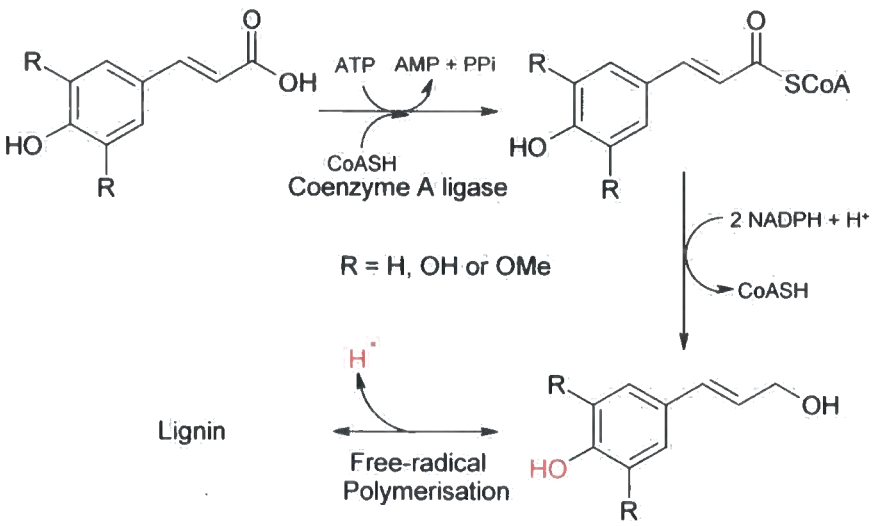


Figure 4.17 Lignification - a metabolic pathway that is able to incorporate phenylpropanoids

With respect to the levels of incorporation of the applied methylated-acyl donors in contrast to applied carboxylic acids in petunia; these could not be calculated with any confidence for the naturally occurring phenylpropanoids in the absence of isotopic

labeling. For example, whereas only approximately 200 nmol/g of methyl caffeate was sprayed onto the plants, the large increase in the levels of caffeoylated flavonols could account for over four times the applied dose. The scale of this enhancement is suggestive of metabolic regulation by positive feedback from the end products of the pathway. This conclusion is enhanced by the scale of the accumulation of the acylated flavonoids greatly exceeding the depletion in the levels of QDG and KDG intermediates, implying that the pool was being replenished by *de novo* flavonoid biosynthesis.

These results may be of great utility in selectively regulating acylated natural products in plants. As alluded to in section 4.1, the accumulation of flavonoid antioxidants has potential advantages in prolonging the shelf life of harvested plants, as well as boosting the levels of those flavonoids proposed to have beneficial health properties. Alternatively, this methodology could be transferred into medicinal plants, to boost the naturally-occuring levels of desirable acylated natural products.

4.3.3 Novel metabolites in petunia feeding study

Metabolite 6 and metabolite 7 accumulated in petunia when fed with methyl coumarate and methyl 4-fluorocinnamate respectively. MS analysis of each compound was similar, although they differed by fragments corresponding to a coumaroyl cation (metabolite 6 -147 Da) in contrast to a 4-fluorocinnamoyl cation (metabolite 7 - 149 Da) and a respective increase of 2 Da in the parent mass ion of metabolite 7 (789 Da) in contrast to metabolite 6 (787 Da). Each metabolite did not fragment to produce an observable diglucoside cation (325 Da), as found for each of

the other flavonols found in petunia (Appendix B), but did produce a 339 Da fragment, which could be attributed to a glucosyl-glucuronide fragment. This glycosylation is known in plant flavonoid metabolism[133] although not in species of petunia. However, identification of fragments typical of a flavonol aglycone is confused by the presence of ions corresponding to kaempferol (287 Da) and quercetin (303 Da) in both studies. Although assimilation of the masses corresponding to phenylpropanoid, glucoside-glucoronide and quercetin components give the observed parent mass ions, 787 and 789 Da. It was tentatively proposed that metabolites 6 and 7 were based upon a quercetin (phenylpropanoyl-)glucosyl-glucuronide structure of unknown isomeric form.

Chapter 5

Biosynthesis of phenylpropanoyl-coenzyme A acyl donors

5.1 Introduction

One of the potential advantages of recruiting characterised CADATs for biocatalyses is the ability to predict successful acyl substrates, eliminating the requirement for screening against this particular variable. However, within this scope it would be useful for a biocatalyst to have facility to utilise as large a range of foreign substrates as possible. To this end, previous literature has suggested that fluorinated phenylpropanoids are able to undergo acyltransfer to flavonoids via coenzyme A-dependent pathways. Results shown in chapter 4 also suggested that this may be possible and further characterisation *in vitro* was concluded to be the optimal way forward for elucidating the true effect of phenylpropanoid fluorination upon CAD biocatalysis, as opposed to functional characterisation of enzymes *in planta*. It was also envisaged that information about the effect of varying aromatic substituents upon coenzyme A-dependent acylation pathways would be gained.

Coenzyme A ligases (Section 1.2.2) have previously been thoroughly characterised with respect to endogenous phenylpropanoid substrates bearing oxygen containing substituents, although little is known about their tolerance to non-natural aromatic compounds[15], particularly those with electron-withdrawing groups such as fluorine. This chapter will therefore aim to elucidate the tolerance of coenzyme A ligases toward fluorination of phenylpropanoid derivatives.

Of the CLs previously reviewed in Section 1.2.2, the enzymes from *Arabidopsis thaliana* represented a well-characterised set of potential biocatalysts with respect to natural phenylpropanoid substrates. With regard to biocatalysis of unnatural carboxylic acid-coenzyme A ligations, At4CL1 was the stand-out candidate owing to its broader substrate tolerance and superior turnover number with respect to every phenylpropanoid variant. Thus this chapter will document the characterisation of At4CL1 activity toward fluorinated phenylpropanoids, as per the following aim:-

- To ascertain the ability of both CoA-dependent enzymes to turnover non-natural, fluorinated aromatic acids and to establish the effect fluorine substitution has upon activity

5.2 Results

5.2.1 Cloning and expression of a 4-coumarate coenzyme A ligase (At4CL1) from *Arabidopsis thaliana*

Complementary primers were designed for the 5' and 3' termini of At4CL1. The 5' primer included an NcoI (CCATGG) restriction site immediately upstream of the start codon. The 3' primer included a codon corresponding to a histidine residue in place of the stop codon, which was followed by a further 5 identical codons to encode for a six-residue histidine tag for purification of the recombinant protein. An XhoI (CTCGAG) restriction site was added to the 3' primer sequence immediately after the histidine sequence.

Preparation of a cDNA template from mRNA was necessary to exclude introns present in the At4CL1 gene. RNA was extracted from *Arabidopsis thaliana* plants, which had been grown under high-light conditions for 48 hours to stimulate flavonoid production and 4CL expression[47]. Reverse transcription of the RNA produced cDNA, which was used as template for PCR of *at4cl1*. PCR in the presence of both primer oligonucleotides gave rise to two amplification products, one minor (approximately 1100 bp) and a major product (approximately 1700 bp). The latter corresponded to the 1781 bp PCR product expected for *at4cl1* (Figure 5.1).

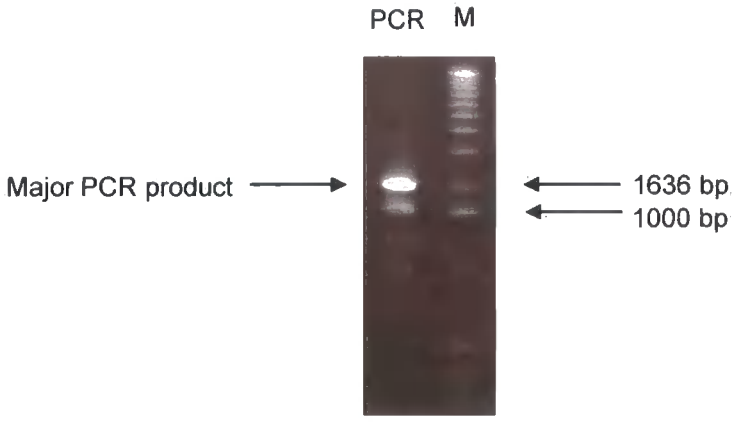


Figure 5.1 PCR of *at4cl1* from *Arabidopsis thaliana* mRNA. M denotes size markers of varying double stranded DNA fragments and PCR denotes the PCR from the cDNA template to derive a ~1700 bp amplification product

After confirming the identity of the PCR product by restriction analysis, it was ligated into a complementary pET 24a (Promega) kanamycin-resistant expression vector. The recombinant plasmid was then electroporated into XL10-Gold ultra-competent cells. Four kanamycin-resistant colonies were achieved, two of which were prepared and sequenced to confirm that they harboured authentic sequence.

IPTG was added to the bacterial cultures to stimulate recombinant protein expression. A soluble recombinant protein of approximately 62 kDa was identified in the pET 24a-*at4cl1* cultures, when analysed by SDS-PAGE following purification of His-tagged proteins by Ni chelate affinity chromatography. This protein was absent in a similarly treated empty vector control. The expected mass for recombinant At4CL1 was 61280 Da. Histidine fusion-tagged recombinant At4CL1 was bound selectively to a Ni ^{55} column in low concentration of imidazole and eluted with 5 mL of buffer at high imidazole concentration. Partially-purified recombinant At4CL1 was recovered by elution in 4 x 1 mL fractions and the polypeptides present analysed by SDS-PAGE (Figure 5.2).

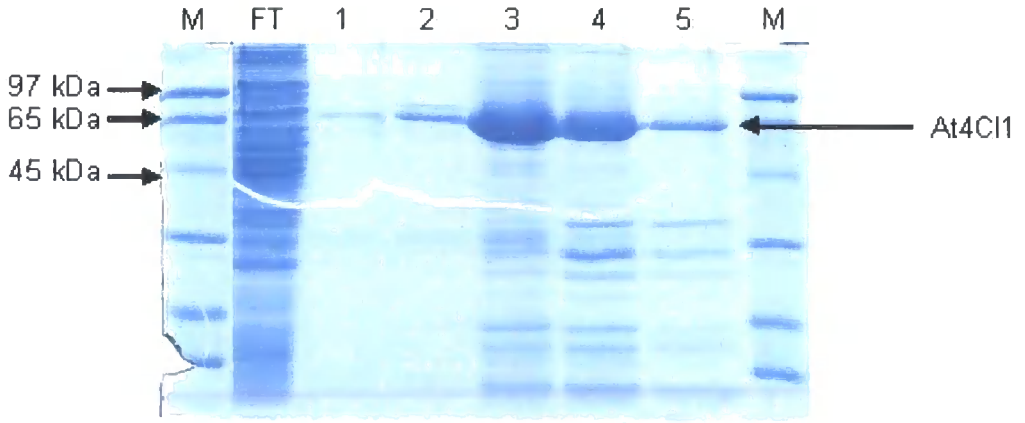


Figure 5.2 SDS-PAGE gel of Ni chelate affinity purified His-tagged recombinant At4CL1. M denotes pre-calibrated molecular mass markers, FT denotes flow-through protein which would not bind to the column and 1-5 denote fractions collected with increasing concentration of imidazole.

Further purification of the semi-pure protein was achieved by application to a second Ni chelation column and the recovered fractions were dialysed to remove

imidazolium salt. The two-stage purification yielded highly enriched At4CL1 (Figure 5.3).

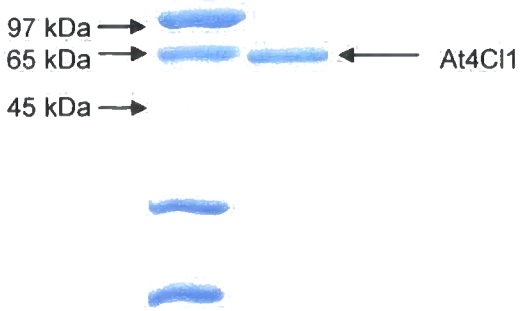


Figure 5.3 SDS-PAGE gel of highly enriched At4CL1

5.2.2 Biochemical characterisation of recombinant At4CL1

Initial activity of purified At4CL1 was determined toward coumaric acid using a colorimetric assay[120]. The activity of At4CL1 toward coumaric acid was also monitored by HPLC using diode array detection (λ max. 333 nm). Product formation was compared to a control experiment, which utilised denatured protein. A reaction product with a UV-Vis absorption maximum of 333 nm, which accumulated over 5 minutes in the presence of At4CL1 was confirmed (Figure 5.4). This product co-chromatographed with synthetic coumaroyl coenzyme A (Section 3.2.1.3). High resolution (ES^+) mass spectrometry of biosynthesised coumaroyl coenzyme A gave a parent *m/z* ion of 914.1605 Da ($M + H^+$ = 914.1598 Da) and UV-Vis absorption confirmed that biosynthesised coumaroyl coenzyme A had an identical absorption maximum to previously synthesised coumaroyl coenzyme A (Section 3.2.1.3).

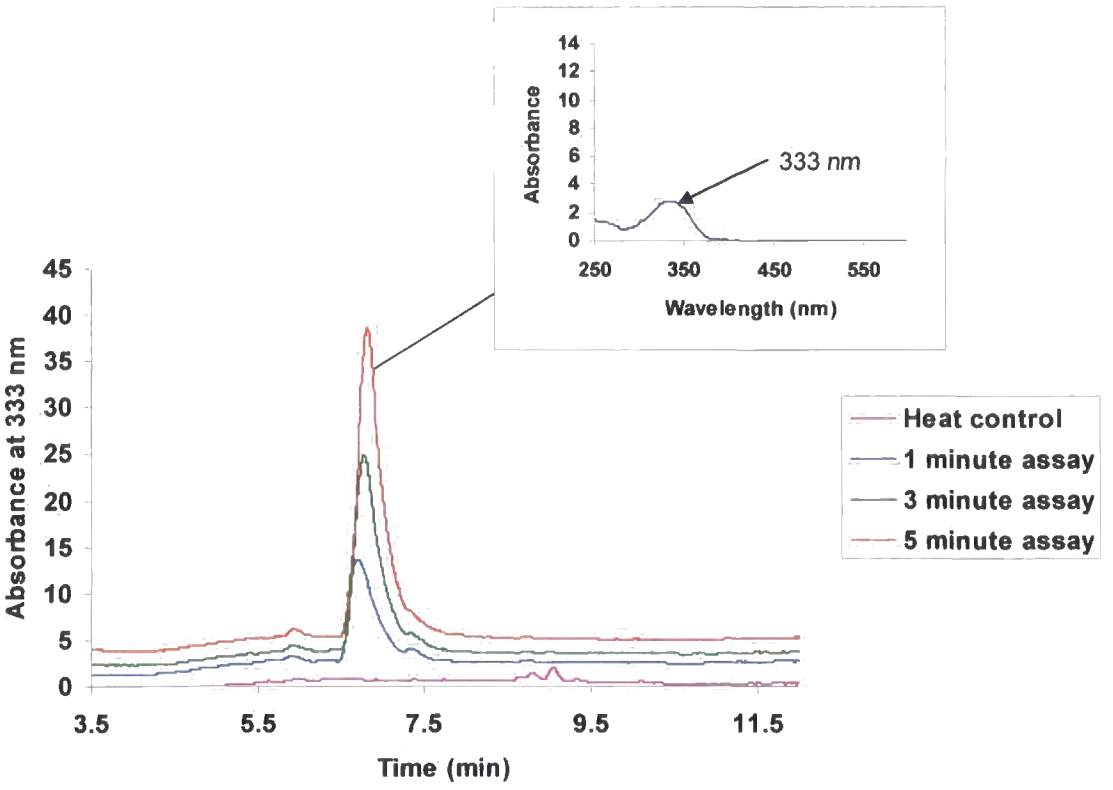


Figure 5.4 HPLC assay of At4CL1 activity toward coumaric acid including diode array analysis of the accumulated product

Using the colourimetric assay, the activity of At4CL1 was determined toward coumaric acid by monitoring the biosynthesis of coumaroyl CoA in duplicate over a time course of 5 minutes. It was found that a linear rate of product formation was achieved over this period in the presence of 21 μg At4CL1. No non-enzymatic accumulation of coumaroyl coenzyme A occurred in the presence of denatured protein, nor in the absence of coumaric acid or native At4CL1.

5.2.2.1 Biosynthesis of phenylpropanoyl coenzyme A acyl donors

Having optimised the assay conditions under which coumaroyl coenzyme A could be biosynthesised, the tolerance of At4CL1 toward fluorinated cinnamic acids was determined. Each of the previously described fluorinated phenylpropanoids was fed in

turn to assays containing 21 μg of At4CL1, 20 mM CoA, 20 mM ATP and incubated for 30 minutes, so as to enhance biosynthesis of products. HPLC-DAD-MS (ES^-) (Figure 5.5) and subsequent HRMS (ES^+) (Appendix E) confirmed successful biosynthesis of coumaroyl, caffeoyl, feruloyl, 4- fluorocinnamoyl and 3-4-difluorocinnamoyl coenzyme A esters, although no activity could be observed toward trifunctionalised cinnamic acids. Each of the successfully biosynthesised coenzyme A esters co-chromatographed with the respective synthetic coenzyme A ester and had identical UV-Vis absorption maximum (Section 3.2.1.3).

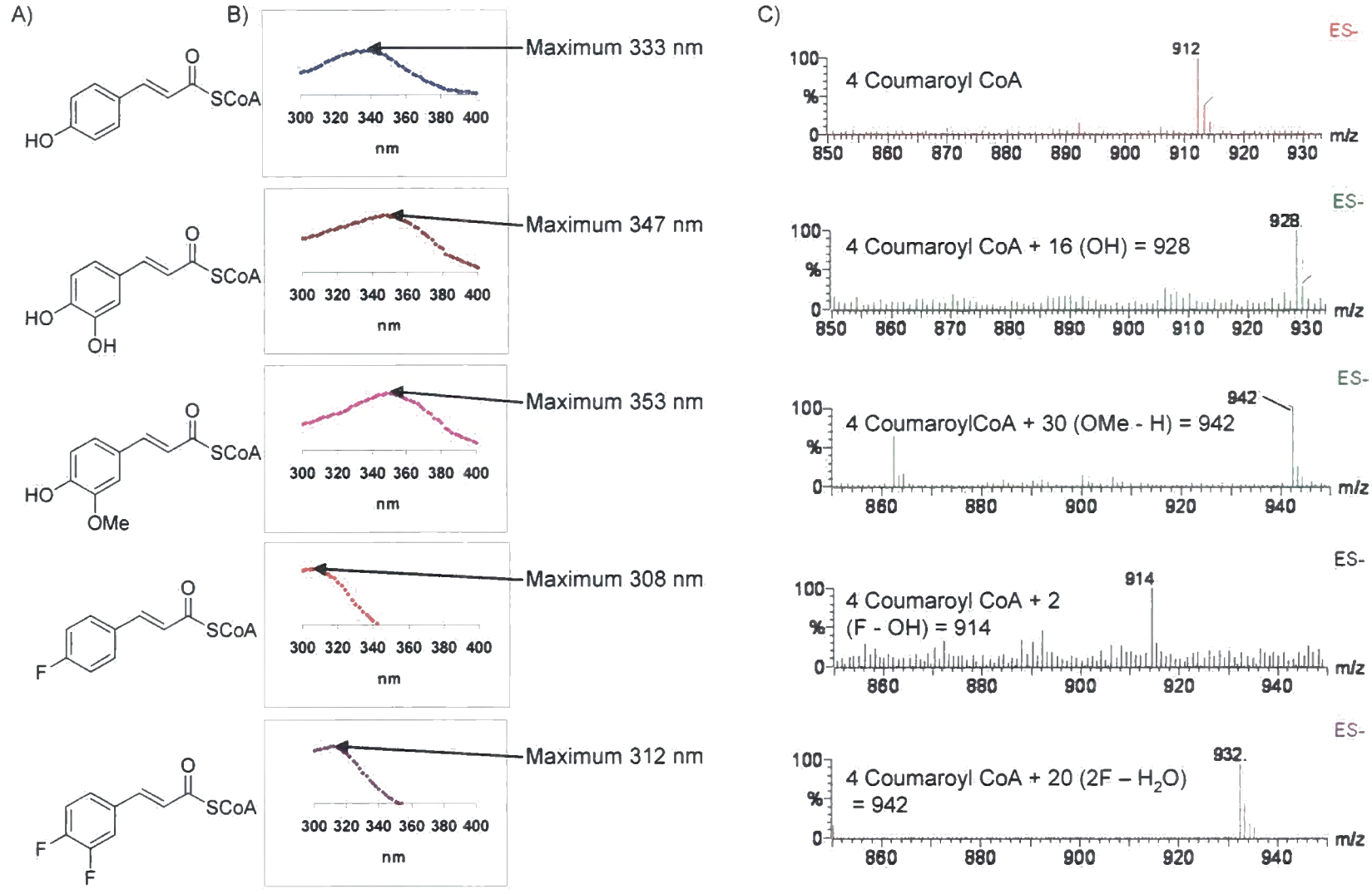


Figure 5.5 Biosynthesis of coenzyme A thioesters (A). The formation of the coenzyme A thioesters were monitored by (B) UV – VIS spectroscopy and (C) electrospray (ES^+)

The purified synthetic coenzyme A esters were utilised to quantify the rate of biosynthesis of coumaroyl, caffeoyl, ferruloyl, 4- fluorocinnamoyl and 3-4-difluorocinnamoyl coenzyme A esters via calculation of molar absorption extinction coefficients (Section 2.2.4.2, Appendix C). Activity of At4CL1 (2.5 μg) toward coumaric acid (1 mM) with CoA (20 mM) and ATP (20 mM) at 37 °C was found to produce 0.162 +/- 0.009 μmoles of coumaroyl coenzyme A in 1 minute. Whereas activity of twice the amount of At4CL1 (5 μg) toward coumaric acid (1 mM) with CoA (20 mM) and ATP (20 mM) at 37 °C was found to produce 0.330 +/- 0.011 μmoles of coumaroyl coenzyme A in 1 minute. Thus enzyme added was the limiting factor upon rate of product formation under these conditions. Under these conditions, rate of product formation, v (turnover number, s^{-1}), was measured for assays comprising decreasing concentrations of each phenylpropanoid from a starting phenylpropanoid concentration where the maximum observable rate was achieved. The kinetic constants k_{cat} and K_M were determined for each phenylpropanoid against At4CL1 (Figure 5.6) by non-linear regression analysis of a fitted hyperbola (Appendix F).

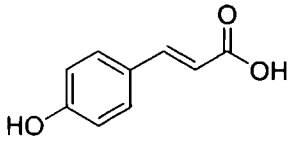
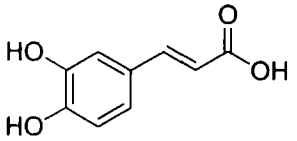
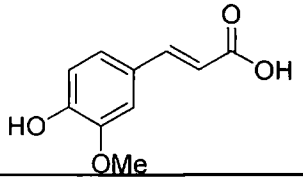
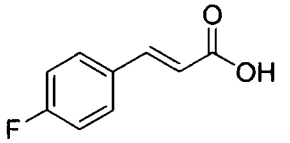
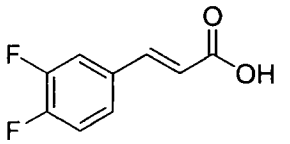
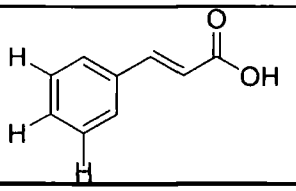
Phenylpropanoid	K_M (μM)	k_{cat} (s^{-1})	k_{cat} / K_M
	40.4 +/- 3.7	68.3 +/- 1.37	1.7
	12.0 +/- 1.2	21.5 +/- 0.4	1.8
	182.2 +/- 12.9	45.5 +/- 1.0	0.25
	856.0 +/- 70.7	62.2 +/- 1.16	0.07
	1610 +/- 177.2	7.7 +/- 0.25	0.005
	6230*	71*	0.011

Figure 5.6 k_{cat} and K_M values determined for each phenylpropanoid against At4CL1 (Values shown are +/- standard deviation; * = literature value for comparison[15])

In conclusion fluorinated phenylpropanoids were able to undergo biotransformation to produce phenylpropanoyl coenzyme A thioesters. The effect of 4- and 3-4-fluorination upon phenylpropanoid binding and rate of catalysis was ascertained. In comparison to hydroxylated analogues, each fluorocinnamic acid exhibited a far greater K_M value. Whilse only a small decline in rate of catalysis was observed for 4-fluorocinnamic acid substrate in comparison to coumaric acid, rate of catalysis of 3-4-fluorocinnamic acid substrate was significantly reduced in comparison to caffeic acid. Thus the lack of metabolism of 3-4-fluorinated cinnamic acids observed in feeding studies in petunia and arabidopsis (Chapter 4) may have been due to a low turnover

number. Whilst the binding of this substrate was weak, the observed decrease in accumulation of QDG-Caf, KDG-Fer and rosmarinic acid (Section 4.2.3) when treated with 3-4-difluorocinnamic acid derivatives, could be attributed to competitive inhibition when present in sufficient quantity.

5.3 Discussion

With regard to the UV-Vis spectra obtained for each of the phenylpropanoyl CoA thioesters, differing aromatic substitution caused differentiation in the λ_{max} observed. Hydroxyl and methoxy substituents are able to withdraw electrons through induction, but also donate electrons through resonance. However, electron donation is the dominant effect and therefore these substituents produce electron-rich aromatic systems. Whilst fluorine substituents also exhibit both of these effects, the dominant effect upon an aromatic system is strong electron withdrawal through induction. Electron-rich aromatic systems exhibit a bathochromic shift of UV-Vis absorption in comparison to electron-poor aromatics and this is in line with the UV-Vis absorbance data obtained in this chapter..

Equally, differing aromatic substitution caused variation in substrate affinity[134]. This decrease in affinity could be attributed to steric hinderance when considering At4CL1's relatively high affinity for coumaric acid, which bears a small hydrogen *meta-* substituent. However, the presence of a slightly charged hydrogen in both the *para-* position (contrast with 4-fluorocinnamic acid) and *meta-* position (contrast with 3-4-difluorocinnamic acid) appeared to be key aspect of substrate affinity in At4CL1.

This suggested the formation of hydrogen bonds to substituents in these positions was important.

To this effect, several amino acid residues have been found to be important in the selective binding of phenylpropanoyl substrates in Arabidopsis 4CL activity[29,31]. Site-specific mutagenesis of the Arabidopsis 4CL isoenzymes found two active site amino acid residues that interacted with phenylpropanoid substituents in the *meta-* position[31]. Lysine320 was proposed to be a hydrogen bond acceptor for *meta-* hydroxy substituents in At4CL2, although hydrogen bonds could also be formed with main chain carbonyl groups.

With regard to the turnover number, substrates bearing substituents that donated electrons through resonance to the site of reaction - the carbonyl carbon - were favourable. *para-* Hydroxyl residues (σ_p = -0.38) are highly electron donating, whilst *para-* fluorine (σ_p = 0.06) and hydrogen (σ_p = 0) are relatively neutral with respect to electron density at the carbonyl carbon. However, *meta-* substituents are unable to donate electrons through resonance to the carboxylic acid and therefore fluoro (σ_m = 0.34), methoxy (σ_m = 0.10) and hydroxyl (σ_m = 0.13) substituents all withdraw electrons from the carbonyl when in this position. On this basis substrates bearing electron withdrawing substituents were observed to exhibit lower rates of turnover. Although electron donation does not necessarily promote rate of catalysis.

As reviewed in Chapter 1, Coenzyme A ligases operate by a double-displacement mechanism, whereby phenylpropanoids react with ATP to form an adenylate intermediate and release PPi before coenzyme A binds and substitutes for the leaving

adenylate group. If nucleophillic substitution of CoA was the rate-determining step (RDS) then electron withdrawing groups would be likely to enhance rate of turnover due to increased electrophilicity of the carbonyl carbon. However, optimal rate of turnover where electron donating or neutral groups (EDG) are present suggests nucleophillicity of the carboxylate group (at pH 8) is important in the RDS and this could be explained through formation of the adenylate intermediate (Figure 5.7).

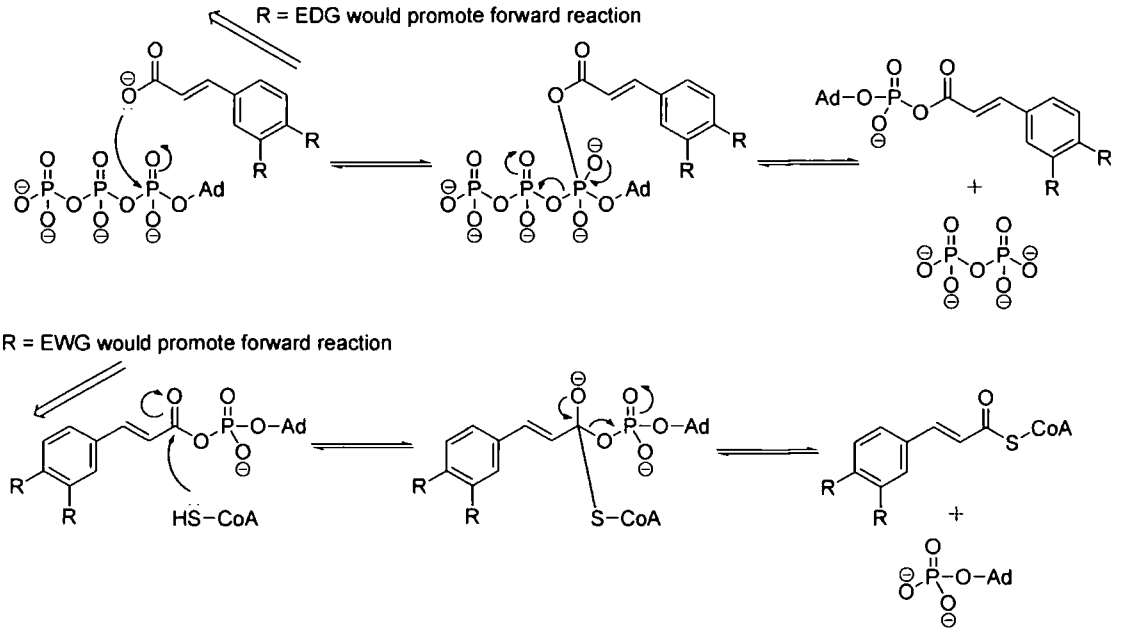


Figure 5.7 Speculation on the effect of aromatic substituents on turnover and the nature of the rate determining step

Thus the formation of the phenylpropanoyl-adenylate intermediate is tentatively proposed to be the rate determining step. In this scenario electron withdrawing groups would decrease the rate of reaction by stabilising a carboxylate anion. Kinetic analysis of the substrates detailed in figure 5.8 and construction of a Hammett plot would be useful further analysis of the effect of *meta-* and *para-* substituents in this mechanism[155].

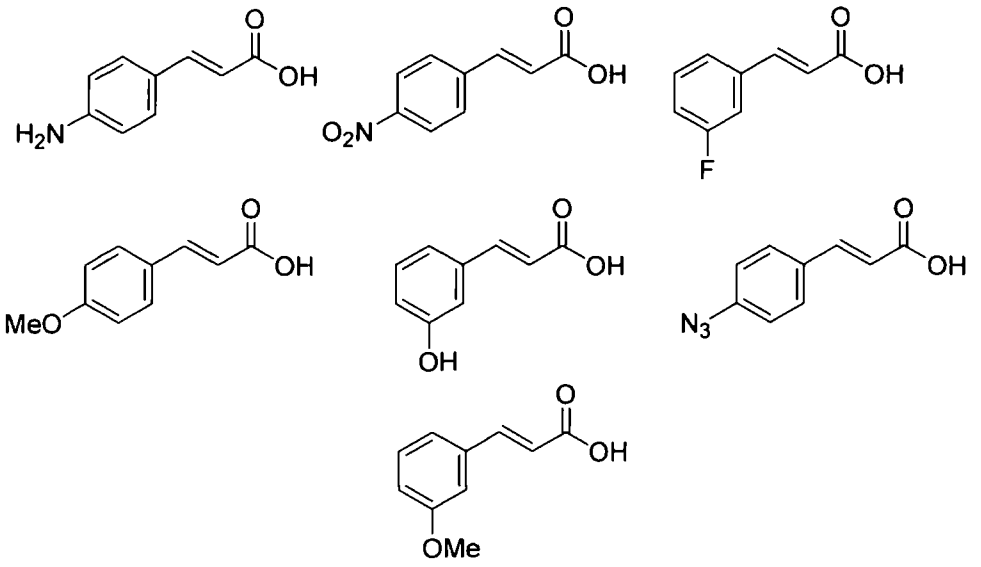


Figure 5.8 Proposed diagnostic substrates for a linear free energy relationship

Activity of At4CL1 toward 4-azidocinnamic was detected in studies in Chapter 7, however the rate of catalysis was too low to obtain kinetic data in saturating conditions. *p*-Azido groups are highly electron withdrawing and this adds weight to the structure-activity relationship observed for substituents in the *para-* position.

The most efficient catalysis (k_{cat}/K_M) was observed for coumaric and caffeic acid substrates. This is indicative of At4CL1 having a role in lignin biosynthesis in arabidopsis as lignin biosynthesis requires both coumaroyl and caffeoyl coenzyme A as substrate[132].

Chapter 6

Coenzyme A-dependent biocatalysis of flavonoid fluoroacylation

6.1 Introduction

Results from previous chapters suggested that fluorinated phenylpropanoids were able to be enzymatically-transferred onto flavonols in petunia *in vivo* (Chapter 4), with results from Chapter 5 showing that it was possible to generate the respective CoA thioester acyl donors *in vitro* utilising a CoA ligase. It was therefore aimed to determine the nature of the successive acyltransfer of the fluorinated acyl moieties from the donor substrates to natural product acceptors.

Having cloned and expressed both a CL and CADAT, a further aim of this thesis was to characterise the application of CAD enzymes with regard to performing viable biotransformations. One area where efficiency of CAD acyltransfer can be improved is dependence upon the cofactors coenzyme A and ATP.

Thus the aims relevant to this chapter are as follows:

- To characterise aromatic coenzyme A-dependent acyltransferases and coenzyme A ligases of flavonoid metabolism with regard to performing biotransformations both *in vivo* and *in vitro*
- Specifically, acyltransfer of phenylpropanoids will be characterised, as the most diverse and abundant substrates of acyltransfer in flavonoid metabolism
- To ascertain the ability of both CoA-dependent enzymes to turnover non-natural, fluorinated aromatic acids and to establish the effect of substituents upon activity

6.2 Results

6.2.1 Cloning of a coenzyme A-dependent phenylpropanoyltransferase

At present, nine BAHD acyltransferases have been shown to catalyse transfer from aromatic acyl donors with three enzymes specific for flavonoid acceptors, as reviewed in chapter 1. A summary of the characterised phenylpropanoyltransferases is shown in Figure 6.1.

Enzyme name	Originating plant species	Coenzyme A acyl donor	Acyl acceptor (**X**- = denotes the region of acyltransfer)
SagPPT1[135]	*Salvia splendins*	Hydroxycinnamoyl	Pelargonidin **3**- *O* - 5- *O* - diglucoside
PerPPT1[56]	*Perilla frutescens*	Hydroxycinnamoyl	Delphinidin **3**- *O* - glucoside
Gent5AT[56]	*Gentiana triflora*	Hydroxycinnamoyl	Delphinidin 3- *O* -, **5**- *O* - diglucoside

Figure 6.1 Previously characterised coenzyme A-dependent hydroxycinnamoyltransferases in plant secondary metabolism and their substrates

The characterised aromatic: 5- anthocyanin acyltransferase (Gent5AT) from *Gentiana triflora* catalyses the acyl transfer of caffeoyl acyl groups onto delphinidin 3-5-*O*-diglucoside (Figure 6.2). This enzyme had previously been found to have activity toward both caffeoyl and coumaroyl coenzyme A acyl donors and was able to be expressed heterogenously in soluble form. Thus this enzyme was selected for characterisation as biocatalyst of fluoracyltransfer.

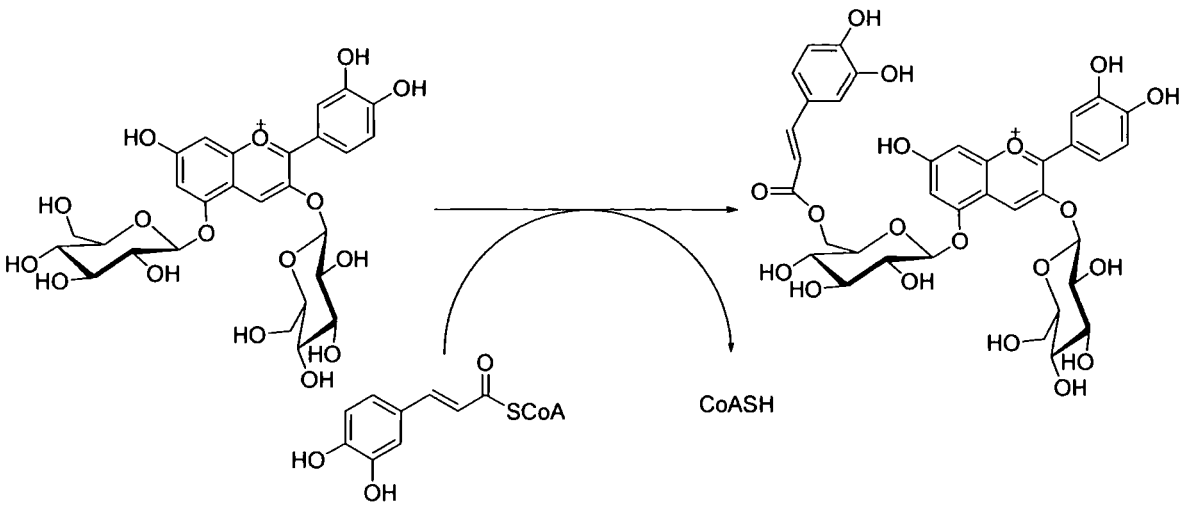


Figure 6.2 Endogenous activity of Gentiana aromatic 5- *O*- anthocyanin acyltransferase

mRNA was prepared from pale blue flower buds from *Gentiana triflora* expressing Gent5AT as described by Fujiwara *et. al.*[81]. cDNA was then prepared as a template for PCR amplification of gent5at. Primers were designed to incorporate a strep - tag within the *N*- terminus allowing for the affinity purification of Gent5AT by affinity to strep-tactin. The 5' primer also incorporated a *Pac1* restriction site (TTAATTAA) 4 base pairs (CCAT) upstream of the start codon, where the 3' primer incorporated a *Sal1* restriction site immediately downstream of the stop codon. PCR reactions were then performed in the presence of the reverse transcribed cDNA template at an annealing temperature of 55 °C (Figure 6.3).

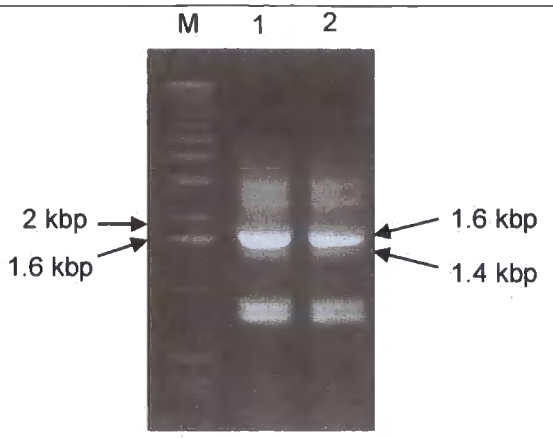


Figure 6.3 PCR of Gent5AT from pale blue petals, M = graduated double stranded DNA markers, 1 = PCR at 55 °C annealing temperature and 2 = 60 °C annealing temperature. Two PCR products of approximately the correct size were produced, in addition to other anomalous bands, due to non – specific annealing of the primers. No PCR could be achieved at 65 °C annealing temperature

The amplification product of 1.4 kbp was of the expected size for gent5at (1410 bp) and was restricted and ligated into a complementary Petstrep3 plasmid (D. Dixon *et. al.* unpublished data). The plasmid was then transformed into XL10-Gold ultracompetent cells by electroporation. Upon sequencing, the recombinant plasmid was found to contain two non - silent mutations as compared with the previous reported sequence[81]. A second cloning process was undertaken which yielded gent5at containing identical anomalies. Thus it was concluded that the amplified product corresponded to a natural variation in gent5at probably due to varietal variation in the biological sources. The location of the mutations at either termini, suggested that they were unlikely to influence catalytic activity (Figure 6.4).

Figure 6.4 (overleaf) Recombinant gent5at contained two non-silent mutations (marked *) as compared with the literature sequence

```
1          11         21         31         41         51         61         71
ATGGAGCAAATCCAAATGGTGAAGGTTCTTGAAAAATGCCAAGTTGCACCACCATCTGACACAACAGATGTCGAGTTATC
 M  E  Q  I  Q  M  V  K  V  L  E  K  C  Q  V  A  P  P  S  D  T  T  D  V  E  L  S
                                              T*

81         91         101        111        121        131        141        151
GCTACCGGTAACATTCTTTGATATCCCCTGGTTGCACTTGAATAAGATGCAGTCCCTTCTGTTTTACGACTTTCCGTACC
  L  P  V  T  F  F  D  I  P  W  L  H  L  N  K  M  Q  S  L  L  F  Y  D  F  P  Y

161        171        181        191        201        211        221        231
CAAGAACACATTTCTTGGACACTGTCATCCCTAATCTTAAGGCCTCTTTGTCTCTCACTCTAAAACACTACGTTCCGCTT
P  R  T  H  F  L  D  T  V  I  P  N  L  K  A  S  L  S  L  T  L  K  H  Y  V  P  L

241        251        261        271        281        291        301        311
AGCGGAAATTTGTTAATGCCGATCAAATCGGGCGAAATGCCAAAGTTTCAGTACTCCCGTGATGAGGGCGACTCGATAAC
 S  G  N  L  L  M  P  I  K  S  G  E  M  P  K  F  Q  Y  S  R  D  E  G  D  S  I  T

321        331        341        351        361        371        381        391
TTTGATCGTTGCGGAGTCTGACCAGGATTTCGACTACCTTAAAGGTCATCAACTGGTAGATTCCAATGATTTGCATGGCC
  L  I  V  A  E  S  D  Q  D  F  D  Y  L  K  G  H  Q  L  V  D  S  N  D  L  H  G

401        411        421        431        441        451        461        471
TTTTTTATGTTATGCCACGGGTTATAAGGACCATGCAAGACTATAAAGTGATCCCGCTTGTAGCCGTGCAAGTAACCGTT
L  F  Y  V  M  P  R  V  I  R  T  M  Q  D  Y  K  V  I  P  L  V  A  V  Q  V  T  V

481        491        501        511        521        531        541        551
TTTCCTAACCGTGGCATAGCCGTGGCTCTGACGGCACATCATTCAATTGCAGATGCTAAAAGTTTTGTAATGTTCATCAA
 F  P  N  R  G  I  A  V  A  L  T  A  H  H  S  I  A  D  A  K  S  F  V  M  F  I  N

561        571        581        591        601        611        621        631
TGCTTGGGCCTATATTAACAAATTTGGGAAAGACGCGGACTTGTTGTCCGCGAATCTTCTCCCATCTTTTGATAGATCGA
  A  W  A  Y  I  N  K  F  G  K  D  A  D  L  L  S  A  N  L  L  P  S  F  D  R  S

641        651        661        671        681        691        701        711
TAATCAAAGATTTGTATGGCCTAGAGGAAACGTTTTGGAACGAAATGCAAGATGTTCTTGAAATGTTCTCTAGATTTGGA
I  I  K  D  L  Y  G  L  E  E  T  F  W  N  E  M  Q  D  V  L  E  M  F  S  R  F  G

721        731        741        751        761        771        781        791
AGCAAACCCCCTCGATTCAACAAGGTACGAGCTACATATGTCCTCTCCCTTGCTGAAATCCAGAAGCTAAAGAACAAAGT
 S  K  P  P  R  F  N  K  V  R  A  T  Y  V  L  S  L  A  E  I  Q  K  L  K  N  K  V

801        811        821        831        841        851        861        871
ACTGAATCTCAGAGGATCCGAACCGACAATACGTGTAACGACGTTCACAATGACGTGTGGATACGTATGGACATGCATGG
  L  N  L  R  G  S  E  P  T  I  R  V  T  T  F  T  M  T  C  G  Y  V  W  T  C  M

881        891        901        911        921        931        941        951
TCAAATCAAAAGATGACGTCGTATCAGAGGAATCATCGAACGACGAAAATGAGCTCGAGTACTTCAGTTTTACAGCGGAT
V  K  S  K  D  D  V  V  S  E  E  S  S  N  D  E  N  E  L  E  Y  F  S  F  T  A  D

961        971        981        991        1001       1011       1021       1031
TGCCGAGGACTTCTGACGCCCCCGTGTCCGCCTAACTACTTTGGCAACTGTCTTGCGTCATGCGTTGCAAAAGCAACACA
 C  R  G  L  L  T  P  P  C  P  P  N  Y  F  G  N  C  L  A  S  C  V  A  K  A  T  H

1041       1051       1061       1071       1081       1091       1101       1111
TAAAGAGTTAGTTGGGGATAAAGGGCTTCTTGTTGCAGTTGCAGCTATTGGAGAAGCCATTGAAAAGAGGTTGCACAACG
  K  E  L  V  G  D  K  G  L  L  V  A  V  A  A  I  G  E  A  I  E  K  R  L  H  N

1121       1131       1141       1151       1161       1171       1181       1191
AAAAAGGCGTTCTTGCAGATGCAAAAACTTGGTTATCGGAATCTAATGGAATCCCTTCAAAAAGATTTCTCGGGATTACT
E  K  G  V  L  A  D  A  K  T  W  L  S  E  S  N  G  I  P  S  K  R  F  L  G  I  T

1281       1291       1301       1311       1321       1331       1341       1351
TGCAGAATTGATTTATGTGATTCAGTCCAGGGATTTCGAAAAAGGTGTGGAGATTGGAGTATCATTGCCTAAGATTCATA
  A  E  L  I  Y  V  I  Q  S  R  D  F  E  K  G  V  E  I  G  V  S  L  P  K  I  H

1361       1371       1381       1391       1401       1411
TGGATGCATTTGCAAAAATCTTTGAAGAAGGTTTTTGCTTTTTGTCATAG
M  D  A  F  A  K  I  F  E  E  G  F  C  F  L  S  *
                                      S*
```

6.2.2 Expression and purification of Gent5AT

The sequenced plasmid was initially transformed for expression into Rosetta 2 competent cells. However, standard expression conditions of 1 mM IPTG at 37 °C failed to produce soluble Gent5AT. Instead it appeared that the recombinant protein was expressed in the inclusion bodies. To overcome the insolubility of Gent5AT, petstrep3 – gent5at was transformed into ArcticExpress™ RIL competent expression cells, which are utilised to assemble polypeptides at low temperatures and therefore reduce the rate of protein synthesis. The slow synthesis of proteins has previously been shown to reduce protein mis - folding and hence increase the amount of soluble protein in a given lysate[136]. Transformed *E. coli* was then subjected to mild expression conditions (0.1 mM IPTG at 4 °C for 12 hours) and the soluble and insoluble proteins in the respective lysate were analysed by SDS-PAGE (Figure 6.5). The petstrep3-gent5at transformed bacteria accumulated an insoluble ~54 kDa polypeptide, which was absent in similarly treated cells containing an empty petstrep3 plasmid. This corresponded to the expected molecular mass of recombinant Gent5AT (expected 53592 Da). To determine whether any of the recombinant protein was present in the soluble fraction, the lysate supernatant was applied onto a Strep - tactin affinity column and affinity retained proteins were recovered (Figure 6.5). Whereas no recombinant Gent5AT (54 kDa) was observed a 60 kDa polypeptide could be recovered which was absent in the lysate from similarly processed control bacteria.

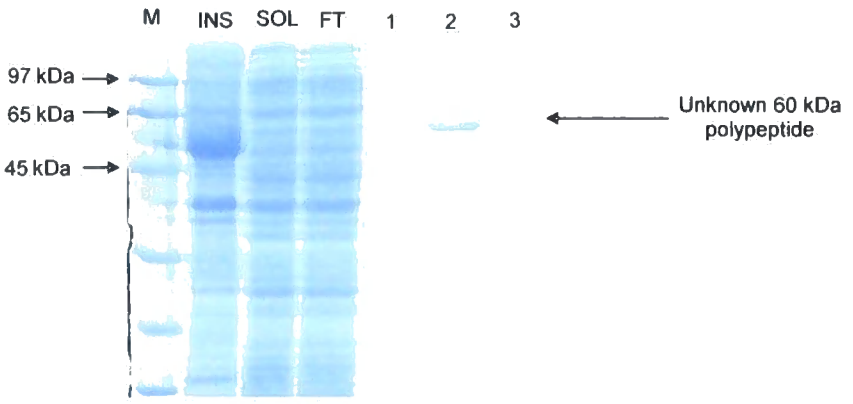


Figure 6.5 SDS – PAGE analysis of insoluble (INS) and soluble [137] proteins in the lysate of IPTG - induced *E. coli* harbouring petstrep3 – *gent5at*. The soluble lysate was subjected to strep-tactin affinity purification and non – binding polypeptides (INS) were removed from affinity - retained polypeptides, which were recovered in three fractions (1 – 3). A 60 kDa protein was recovered in the soluble fraction which did not correspond to the molecular weight of Gent5AT

The unknown 60 kDa protein was not observed in fractions taken from affinity purification of lysate from control cells, therefore the protein was concluded to be associated with the expression of Gent5AT and, to this effect, the *O. antarctica* chaperonin Cpn60 (60 kDa) which is employed in the arctic expression host cells, was thought to be a potential candidate. However, non – tagged proteins such as a chaperonins would not have affinity toward strep-tactin and as such would not be expected to be purified to homogeneity by this method, as was observed. The isolation of the chaperonin polypeptide was postulated to be due to an association with small amounts of strep-tactin – bound Gent5AT. Based upon this conclusion, a large – scale purification of lysates from the Gent5AT expression was undertaken and the *strep* - tagged polypeptides were analysed by SDS-PAGE. A protein of the correct molecular weight (55 kDa) for Gent5AT was observed (Figure 6.6 A – lane 1). To

investigate the reason for the recovery of the 60 kDa protein its association with strep-tactin bound recombinant Gent5AT was disrupted in the presence of ATP (20 mM). If the 60 kDa protein was the chaperonin Cpn60 such a treatment should cause its selective release from the strep – bound Gent5AT[138]. Successive washing steps released the 60 kDa protein and yielded partially - pure Gent5AT (Figure 6.6 A – lanes 2 - 4). The identity of the 55 kDa protein was confirmed to be Gent5AT by screening a western blot with strep-tactin – alkaline phosphatase probe, which selectively binds to *strep* - tagged proteins. Resulting *strep* - *strep*-tactin complexes were then identified by application of BCIP (5- Bromo 4- chloro 3' indolyphosphate *p*- toluidine salt), a colourless substrate of alkaline phosphatase, which forms an insoluble purple dye with nitro-blue tetrazolium chloride upon enzymatic conversion (Figure 6.6 B). The identity of the 54 kDa polypeptide was thus confirmed to have been produced utilising the harboured *gent5at* sequence.

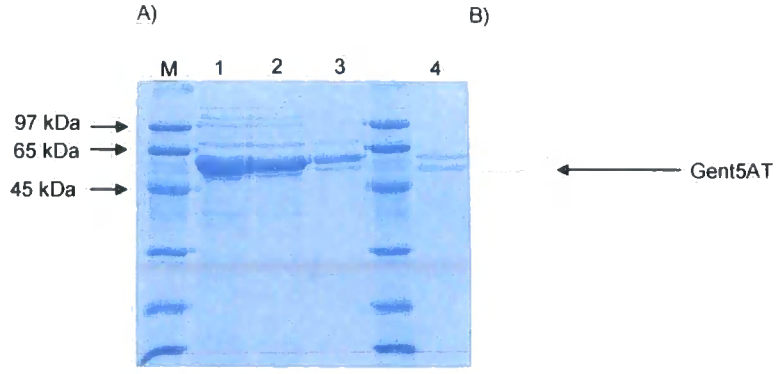


Figure 6.6 (A) Dissociation of chaperonin from strep-tactin bound - Gent5AT by washing with 20 mM ATP (lanes 1-4) and (B) identification of Gent5AT by screening for strep-tagged proteins upon a western blot with a strep-tactin – alkaline phosphatase probe (M = pre- calibrated protein size markers)

6.2.3 Biochemical characterisation of Gent5AT

In order to quantify Gent5AT activity it was necessary to ascertain the concentration of the recombinant protein. Due to residual chaperonin contamination, determination of the concentration of Gent5AT was made more difficult. However, quantification was achieved by running the purified protein extracts on an SDS – PAGE gel, eluting the coomassie blue-stained Gent5AT in 50 % isopropanol containing 0.2 % SDS and comparing the absorbance at 595 nm to that obtained with known concentrations of equally stained BSA, isolated from the same gel (Figure 6.7 A)[139]. Based upon this quantification, aliquots of Gent5AT were determined to be 0.24 μg in 5 μL and 0.49 μg in 10 μL (Figure 6.7 B).

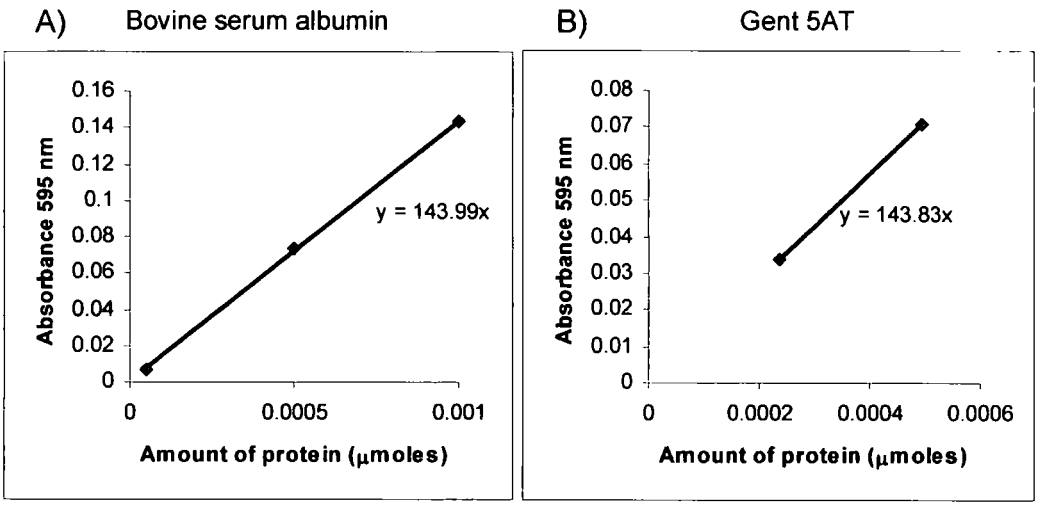


Figure 6.7 (A) Quantification by coomassie blue calibration after SDS – PAGE separation of known amounts of BSA and (B) determination of the concentration of similarly treated Gent5AT

Initial activity of recombinant Gent5AT was tested towards caffeoyl CoA as acyl donor and cyanidin 3-5-*O*- diglucoside as acyl acceptor. The reaction was followed by

HPLC-UV-Vis with detection at 520 nm and a single reaction product was observed, which was absent in controls and which was analysed by ESI MS (ES^+) (Figure 6.8). Diode array and MS analysis tentatively identified the product as the caffeoylated anthocyanin found in previous studies[81].

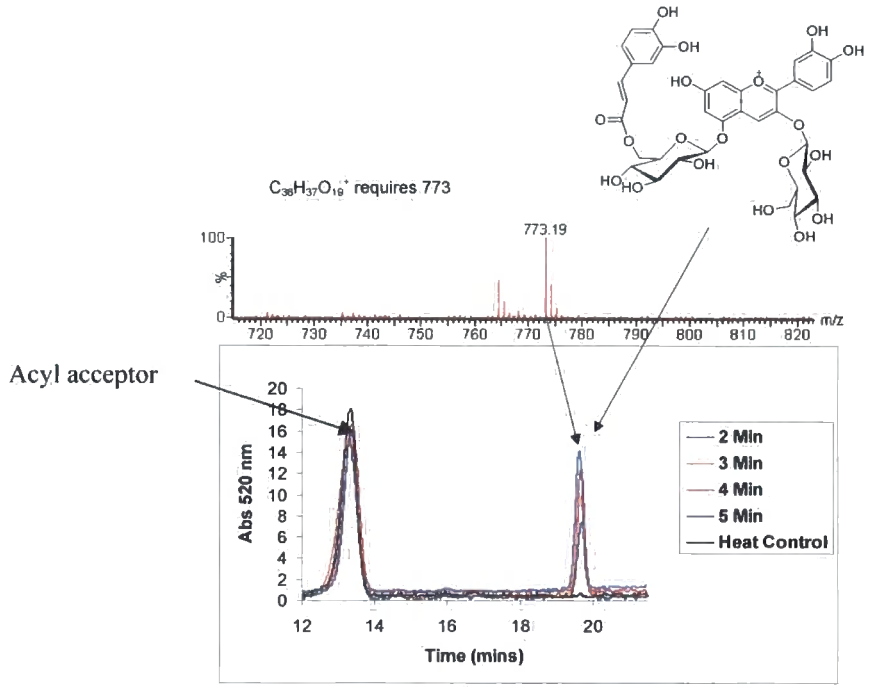


Figure 6.8 Activity of Gent5AT toward caffeoyl CoA and cyanidin 3-5-*O*- diglucoside lead to accumulation of a reaction product at RT 19.8 min. This compound had the correct mass for cyanidin 3-*O*- (5-*O*-caffeoyl)diglucoside by ESI MS: 773 ($[M]^+ = 773$)

In order to allow for the *initial* quantification of enzyme activity, the rate of disappearance of the acyl acceptor was determined using a calculated extinction coefficient (Appendix D), with the total concentration of acyl acceptor in the assay lowered to 0.15 mM, from the 0.4 mM utilised by Fujiwara *et. al.*[56]. Consumption of the acyl acceptor was converted into units of product formed (based upon a 1: 1 ratio of product formed) and the assay was followed over time in order to quantify the activity of Gent5AT toward caffeoyl CoA and cyanidin 3-5-*O*- diglucoside (Figure 6.9).

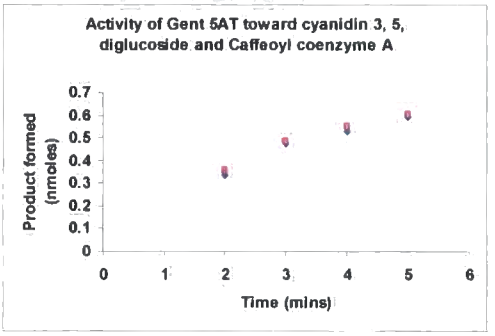


Figure 6.9 Product formation with reduced anthocyanin substrate concentration

As observed in figure 6.9, formation of the acylated anthocyanin product could be actively monitored after 2 minutes of incubation, though the slow accumulation of product was indicative of a low rate of product formation. To this effect, coenzyme A dependent acyltransfer has previously been demonstrated to be subject to inhibition by the respective product[140]. However, this was thought unlikely in this case due to the relatively low level of conversion achieved after 5 minutes. Following on from this, the consumption of the acyl donor caffeoyl coenzyme A was monitored at 347 nm during the same time period of the assay. It was found that 85 % of caffeoyl coenzyme A was consumed within 5 minutes of the assay, as compared to only 8 % of consumption of the acyl acceptor and conversely 8 % product formation (Figure 6.10). A parallel study with denatured Gent5AT was undertaken, which demonstrated that caffeoyl coenzyme A was not consumed in the absence of native Gent5AT. Equally, HPLC - diode array and ESI MS analysis confirmed that caffeic acid was the product of caffeoyl CoA degradation. Thus degradation of caffeoyl coenzyme A was occurring via thioesterase activity, which has been previously reported to be a secondary enzymatic function of BAHD acyltransferases[60].

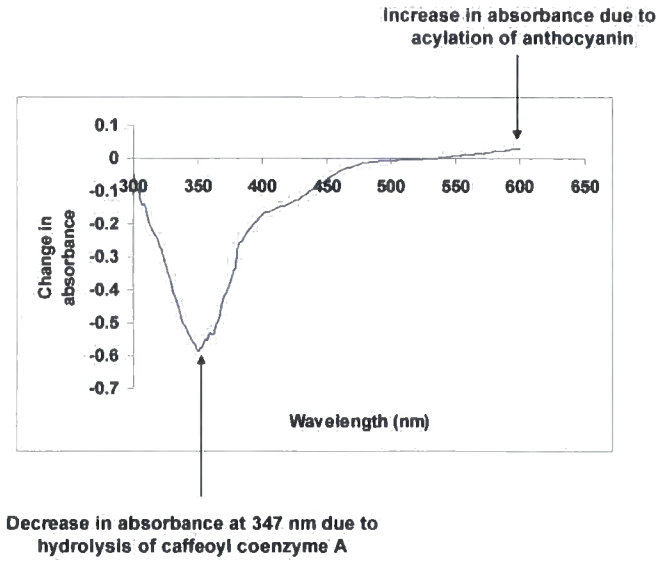


Figure 6.10 Acyltransferase and thioesterase activity toward caffeoyl coenzyme A in the presence of native Gent5AT

The extinction coefficient for cyanidin 3-O-glucoside 5-O- caffeoylglucoside at 520 nm was calculated. A duplicated assay was undertaken with both acyl acceptor and acyl donor at 0.5 mM and quantified by product formation. A linear rate of acyltransfer was achieved (Figure 6.11).

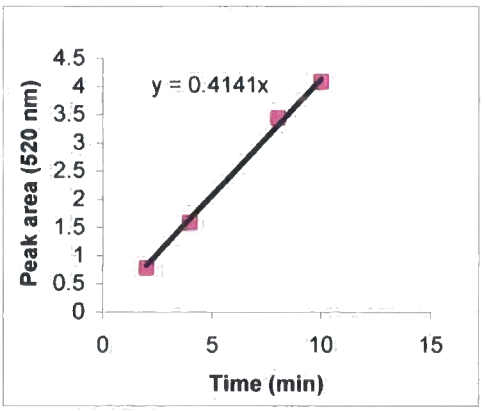


Figure 6.11 Rate of formation of cyanidin 3-(5- (6''- *O*- caffeoyl)) *O*- diglucoside in the presence of Gent5AT

The observed product formation could be solely attributed to enzymatic activity. This was confirmed by the lack of accumulation of cyanidin 3-5-(6''- *O*- caffeoyl) *O*-diglucoside in the presence of denatured protein, or in the absence of caffeoyl CoA.

6.2.4 Biochemical characterisation of phenylpropanoyl acyltransfer

A lot of attention has previously been attributed to understanding the tolerance of secondary metabolic acyltransferases to acyl acceptors[6,83], whereas little attention has been paid to the acyl donors. Having standardised the assay conditions under which acylated cyanidin diglucoside could be biosynthesised efficiently, it was desirable to ascertain the tolerance of Gent5AT toward other phenylpropanoid coenzyme A donors. Coumaroyl, feruloyl, sinapoyl, 4-fluorocinnamoyl, 3-4- difluorocinnamoyl and 3-4-5-trifluorocinnamoyl coenzyme A esters were fed in turn to assays containing 14 μg of Gent5AT and cyanidin 3-5-*O*- diglucoside and incubated for 10 minutes. HPLC-UV-Vis detection showed accumulation of metabolites absorbing at 520 nm in the presence of coumaroyl, ferruloyl, fluorocinnamoyl and difluorocinnamoyl coenzyme A donors (Figure 6.12). However, product formation was not observed for trifunctionalised phenylpropanoid coenzyme A donors at 520 nm or with diode array detection.

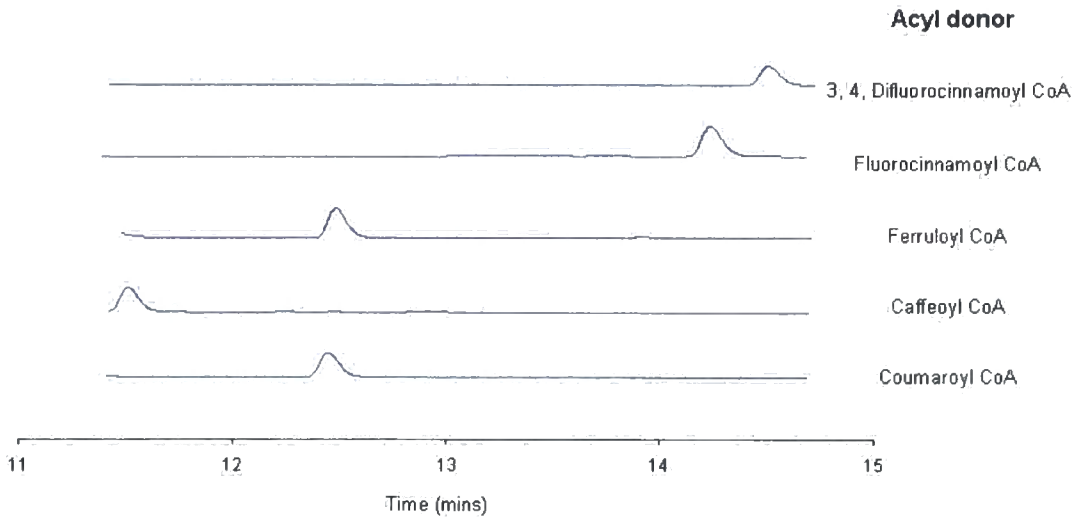


Figure 6.12 Products of enzymatic reaction with Gent5AT in the presence of cyanidin 3-5-*O*-diglucoside and an array of phenylpropanoid coenzyme A donors

Diode array detection allowed for the characterisation of the absorption properties of the products with their molecular weight determined by ESI MS (ES^+) (Figure 6.13 B). The anthocyanin products were found to be acylated from the respective acyl donor in each case (R in figure 6.13 A).

(A)

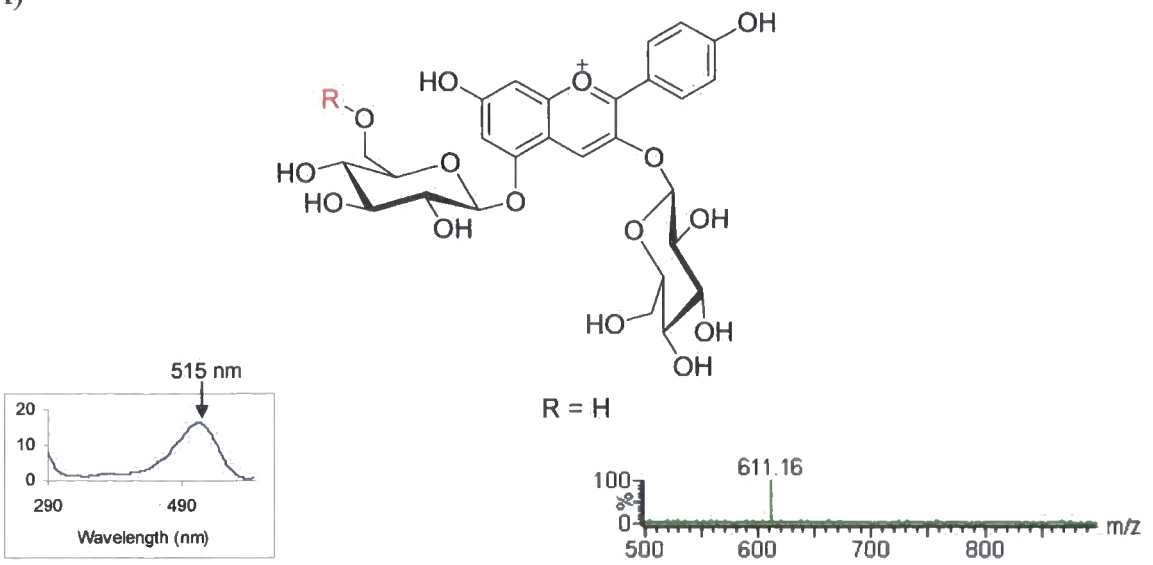


Figure 6.13 (and overleaf) (A) Biosynthesis of acylated products of cyanidin 3-5-*O*- diglucoside. (B) The respective acylations were characterised by (C) UV-VIS and (D) electrospray (ES^+) mass spectrometry

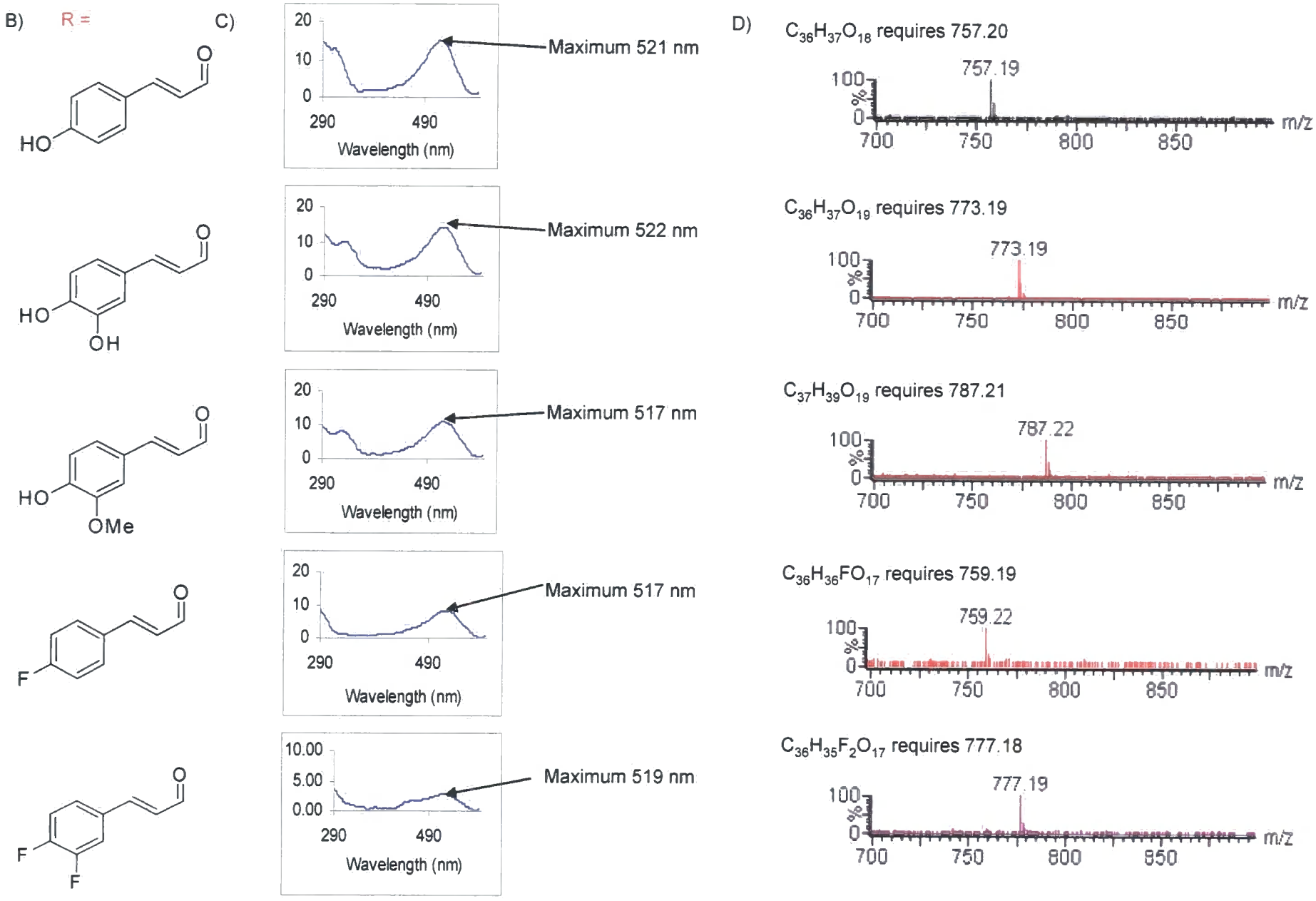


Figure 6.13 Biosynthesis of acylated versions of cyanidin 3-, 5-, diglucoside (A). The respective acylations (B) were characterised by UV – VIS (C) and electrospray (ES+) mass spectrometry (D)

The activity of Gent5AT toward each acyl donor was characterised with cyanidin 3-5-*O*- diglucoside as a constant. Kinetic data was generated for each acyl donor (Figure 6.14) following quantification of product formation, in each instance by calculation of the extinction coefficient at 520 nm (Appendix C).

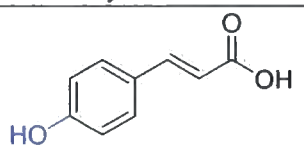
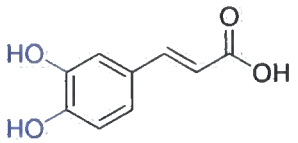
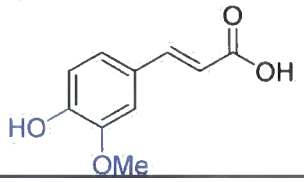
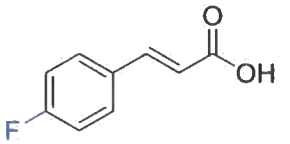
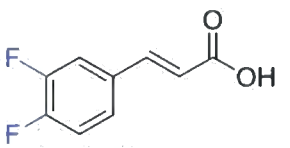
Acyl Donor	K_M (μM)	k_{cat} (s^{-1})	k_{cat} / K_M
	198.2 +/- 17.4	2.75 +/- 0.07	0.013
	68.3 +/- 9.7	4.21 +/- 0.13	0.061
	162.9 +/- 13.7	3.85 +/- 0.10	0.023
	239.0 +/- 35.1	3.20 +/- 0.13.	0.013
	224.3 +/- 19.2	4.5 +/- 0.18	0.020

Figure 6.14 Kinetic data for Gent5AT catalysis of acyltransfer of phenylpropanoids from coenzyme A to acyl acceptor. Values shown are the mean of two duplicates +/- standard deviation.

Acyl donors with varying 3- and 3-4- aromatic substituents were well tolerated by Gent 5AT. However, the magnitude of the effect of aromatic substituents upon the kinetic constants k_{cat} and K_M was much reduced in comparison to data found for At4CL1 activity. With respect to K_M values, this was attributed to binding of the CoA moiety to Gent5AT, which is consistent throughout the substrate series. Electron withdrawing aromatic substituents that increased the electrophillicity of the carbonyl

were found to have only a mild effect upon k_{cat}. Overall each phenylpropanoid was used as acyl donor with relative efficiency.

6.2.5 Sequential biosynthesis of acylated anthocyanins from phenylpropanoic acids

A one-pot biosynthetic approach toward the biosynthesis of acylated natural products from carboxylic acid substrates has many potential benefits in applied enzymatic acylation, as previously discussed (Chapter 5). In order to ascertain its feasibility, the effect of combining At4CL1 and Gent5AT activities and their various substrates / cofactors was ascertained. A variety of experiments were devised and specific activity of the final enzymatic step, Gent5AT activity toward caffeoyl CoA (Acyl donor) and cyanidin 3-5-*O*- diglucoside (acyl acceptor), was monitored (Figure 6.15).

	Assay conditions	*Specific activity nkat / mg protein*
A	Gent5AT + acyl donor + acyl acceptor	48.9 +/- 0.89
B	Gent5AT + At4CL1 + acyl donor + acyl acceptor + ATP	66.0 +/- 0.67
C	Gent5AT + At4CL1 + acyl donor + acyl acceptor + ATP + CoA	58.3 +/- 2.09
D	Gent5AT + At4CL1 + caffeic acid + acyl acceptor + ATP + CoA	51.6 +/- 2.01
E	Gent5AT + denatured At4CL1 + acyl donor + acyl acceptor + ATP	47.7 +/- 1.49

Figure 6.15 *In vitro* biosynthesis of acylated anthocyanins by a one – pot approach incorporating both coenzyme A ligase and acyltransferase activities. Acyl donor (caffeoyl coenzyme A) and acyl acceptor (cyanidin 3-5-*O*- diglucoside) were added to equimolar concentration (0.5mM) when present. Additional coenzyme A in assays C and D was added to 0.05 mM and caffeic acid was added to 0.5 mM in assay D.

In vitro acyltransferase activity was found to be significantly enhanced by the presence of native At4CL1 (Assay B as compared with Assay A) as opposed to other components of 4CL activity, such as ATP or $MgCl_2$ (Assay E). Potentially there were two explanations for this. The first of which was that the 4CL enzyme had an indirect effect upon acyltransferase activity, by maintaining the levels of acyl donor available to the acyltransferase by restoring acyl donor degraded by thioesterase activity (Figure 6.16). Interestingly, there appeared to be evidence of a more complex interaction between the two activities in Assay D, where acyltransfer occurred at a reasonable rate without the prior accumulation of acyl donor. It was postulated that

the acyl donor was being biosynthesised transiently before delivery to the acyltransferase, although the optimal enzyme activity for this assay would be expected to be achieved if this was the case. However, further conclusions on this matter were difficult to ascertain by comparison, due to the associated inhibition of acyltransfer by the presence of coenzyme A, as observed in comparison of conditions B and C. Therefore the positive effect of At4CL1 upon acyltransferase activity is, at least in part, likely to be due to its ability to bind coenzyme A and reduce inhibition of the acyltransferase.

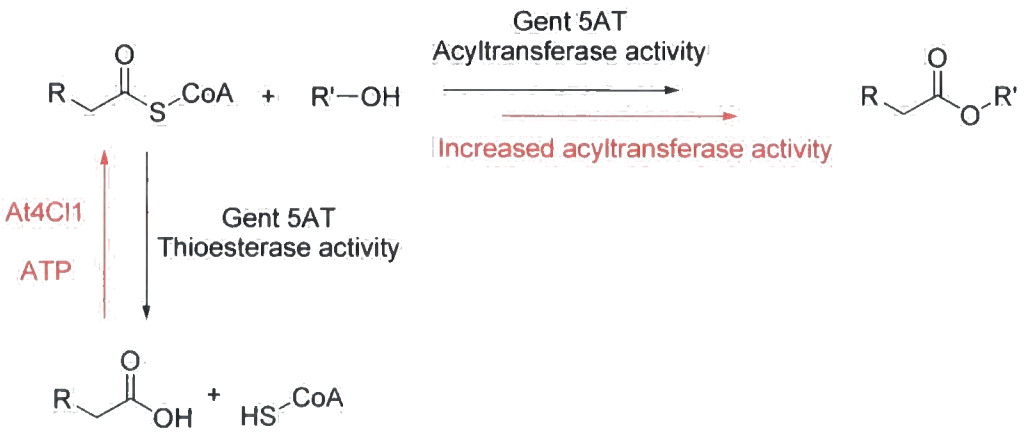


Figure 6.16 Effect of combined At4CL1 and Gent5AT activity upon the effective biosynthesis of acylated anthocyanins

The use of the coupled 4CL and CADAT to drive the *in vitro* biosynthesis of acylated anthocyanins increased the specific activity of acyltransferase activity by 5.8 % (mean of all values). This was proposed to occur due to the role of At4CL1 in counteracting the enzymatic hydrolysis of acyl donors.

6.3 Discussion

6.3.1 Biochemical characterisations

The biosynthesis of fluoracylated flavonoids was therefore achieved *in vitro* via a coenzyme A dependent acyltransfer pathway. This conclusion adds weight to the observations made in relation to the *in vivo* feeding study in Petunia (Chapter 4), where a mono-fluorinated phenylpropanoid appeared to be incorporated into flavonol biosynthesis.

The effect of aromatic fluorination upon CADAT substrate selectivity was ascertained in this chapter. Comparison of Gent5AT activity toward natural and non-natural acyl donors revealed that substrate specificity was only mildly contrasting. In terms of K_M Gent5AT had highest affinity for caffeoyl CoA. In comparison, reduced K_M values for coumaroyl and feruloyl CoA substrates suggested that a *meta-* hydroxyl substituent was important for substrate affinity, possibly due to hydrogen bonding, although the correlation is relatively weak. In general terms, Gent5AT had greatest affinity for electron-rich aromatics - coumaroyl, caffeoyl and feruloyl CoA - in contrast to the electron-poor fluorinated substrates and therefore π-π interactions were thought to potentially be involved in binding of acyl donors.

The effect of varying aromatic substituents upon the turnover number was again relatively weak. However a positive correlation between an aromatic substituents ability to withdraw electrons from the carbonyl carbon[155] and the rate of catalysis was observed. Considering the mechanism of enzymatic acyltransfer (Chapter 1), electron withdrawal would be expected to increase the electrophilicity of a carboxyl group and

therefore increase the rate of catalysis where nucleophillic addition of serine to the electrophillic carbonyl was the rate determining step (Figure 6.17).

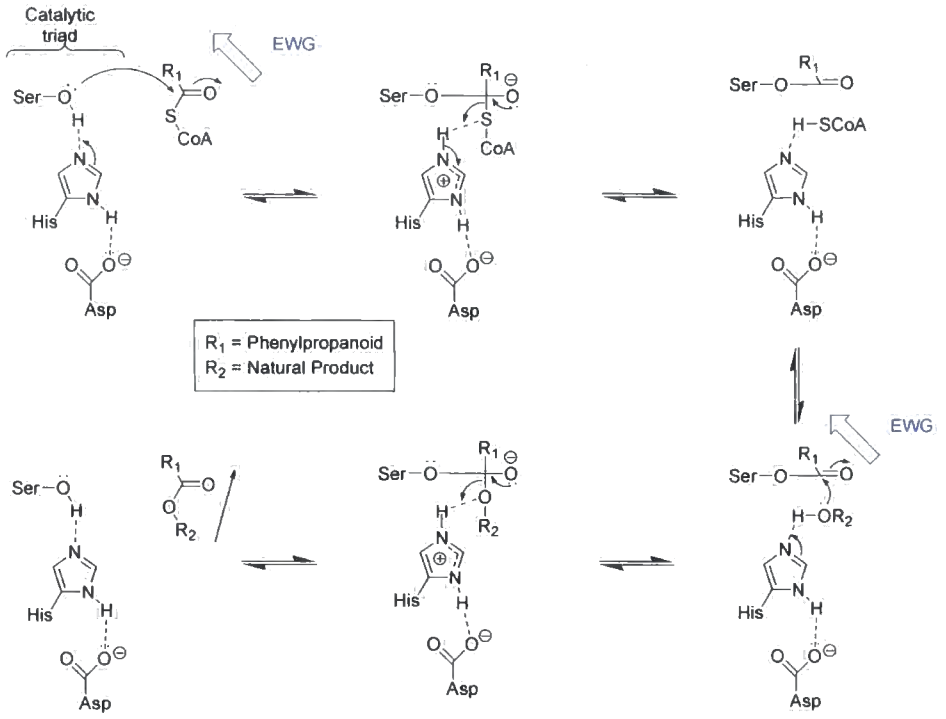


Figure 6.17 The steps in enzymatic acyltransfer mechanism which could be promoted by electron withdrawal from the carbonyl carbon

However, if nucleophillic addition was the sole rate determining step, then a greater influence of aromatic substitution would possibly be expected[155]. Therefore it appeared possible that the rate determining step was a combination of both a nucleophillic addition (aromatic substituent influenced) and elimination of a consistent leaving group i.e. coenzyme A (Figure 6.18).

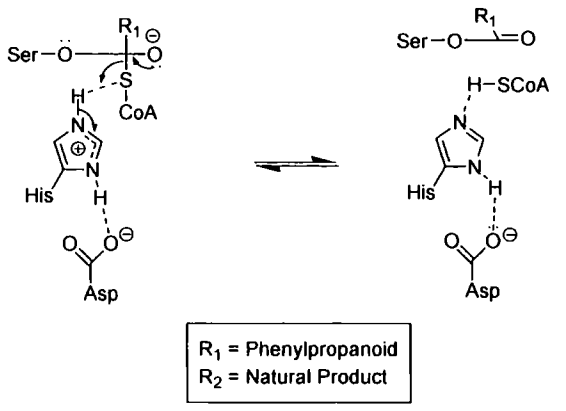


Figure 6.18 Elimination of coenzyme A from serine bound acyl substrate

Therefore the coenzyme A moiety of an acyl donor was thought to contribute to both substrate binding and turnover. This is indicative of the role of coenzyme A ligases in ensuring the correct acyl moiety is transferred by the relatively less substrate-specific acyltransferase. This would require association of a CoA ligase with a particular acyltransferase, as has been previously described in complexation of other flavonoid biosynthetic enzymes[23]. The use of acyl donor-specific CLs associated with natural product-specific CADATs is a possible method for performing natural product acylations with an array of acyl groups (phenylpropanoyl, benzoyl etc) using coenzyme A as mediator. Further experimentation upon the effect of different aromatic substituents (NO_2, NH_3 etc) to produce a Hammett plot was discussed in section 5.3 and would be similarly applicable in this study.

6.3.2 Product characterisation

Limitations concerning the regioselectivity of the acyltransfers described were present in this case-study, as the described region of acyltransfer was based upon previous

characterisation of Gent5AT activity[54]. Without detailed structural characterisations, the region of acyltransfer cannot be presumed. This would have to be ascertained when applying CADATs to a biocatalytic process to ensure the desired regioselectivity was achieved for each acyl donor used.

Chapter 7

Identification and isolation of BAHD acyltransferases

7.1 Introduction

To date, the identification of BAHD acyltransferases involved in the biosynthesis of several highly important natural products has been achieved. Thus acyltransferases involved in the biosynthesis of morphine[59], taxol[70], vindoline[58] and vinorine[62] have all been identified. However, the elucidation of the pathways involved in the acylation of these natural products took many years to establish and the biochemical characterisation of the respective acyltransferases proved very problematic. In order to harness any potential of CADATs in applied biocatalysis, a greater number of potential biocatalysts must be cloned and characterised than are currently known. To this effect it would be useful to develop facile routes toward the identification of novel acyltransferases. Thus this chapter concerns the aim:

- To develop facile routes toward the isolation and identification of coenzyme A-dependent acyltransferases from plants, in order to accelerate the production of a library of biocatalysts to add to existing library of hydrolase-based acyltransferases

Having achieved a toolkit of the necessary substrates and plant materials for the identification and *in vivo* regulation of acyltransferases (Chapters 3 and 4), this chapter will ultimately explore novel routes toward the directed isolation of coenzyme A dependent acyltransferases and coenzyme A ligases from a given proteome, by use of proteomic probes (Section 1.5). As a test for their activity such probes will be first

tested against the previously characterised 4- coumarate coenzyme A ligase (At4CL 1) and BAHD acyltransferase (Gent5AT). The use of these approaches toward the isolation of coenzyme A dependent enzymes from *Arabidopsis thaliana* will then be undertaken (section 4.2.1).

7.2 Results

7.2.1 Identification of target coenzyme A-dependent enzymes

Biosynthesis of acylated flavonoids was observed in *Arabidopsis thaliana* and *Petunia hybrida* by metabolomic studies (Chapter 4). However, assays toward determining enzymatic phenylpropanoid acyltransfer in the presence of crude protein extract from these plants, together with phenylpropanoyl CoA donor and individual acyl acceptors (purified by preparative HPLC) were not successful when carried out according to the method of Fujiwara *et. al.*[56]. As such there appeared to be no possibility of purifying these enzymes by conventional biochemistry. Instead the respective enzyme activities were thought to be suitable targets for discovery by a chemical probe approach (Figure 7.1).

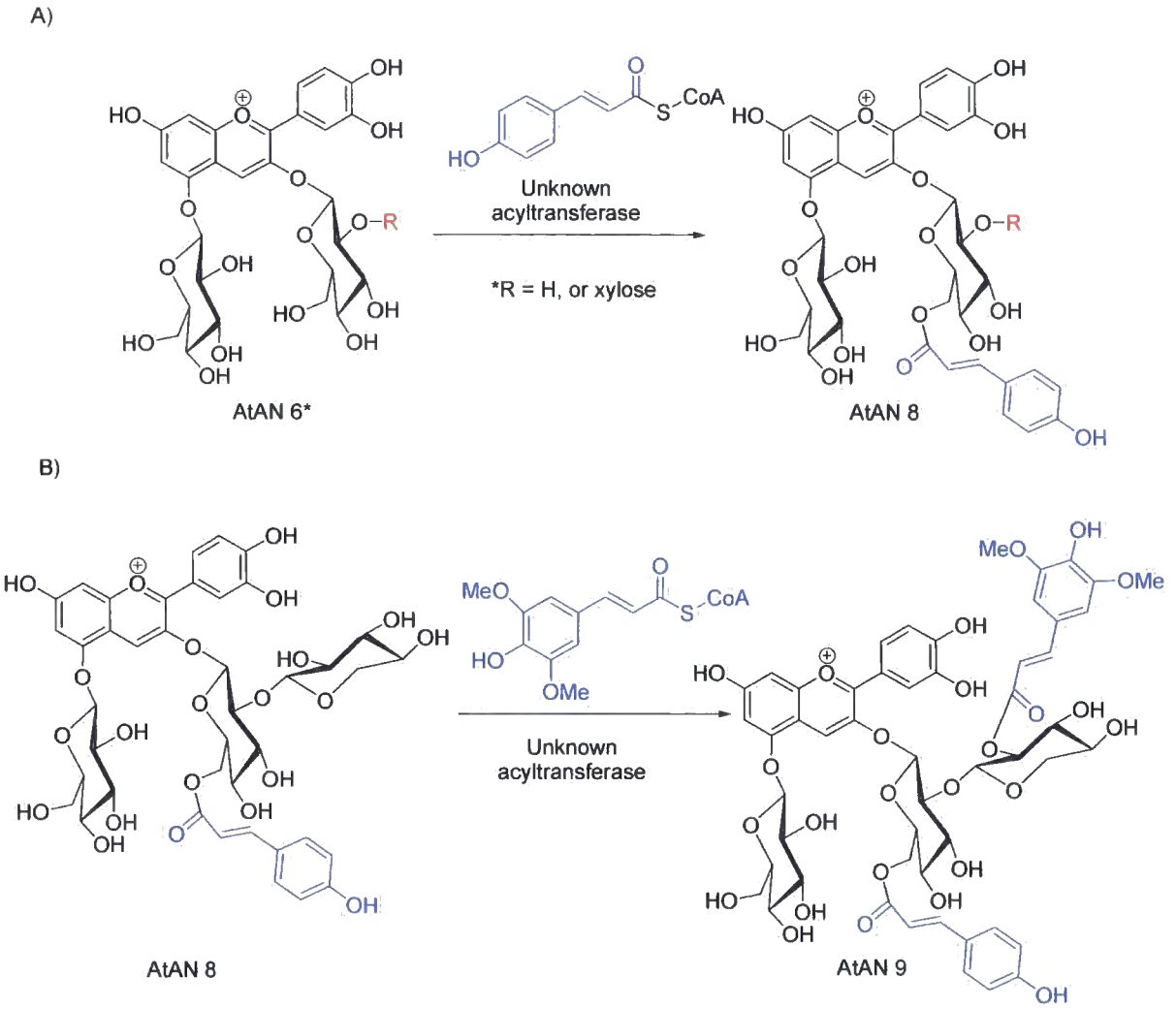


Figure 7.1 Potential target biotransformations for identification by a chemical probe approach. (A) Coumaroyl acylation of cyanidin 3- *O*- (R-glucoside), 5- *O*- glucoside, where R = either H or xylose in *Arabidopsis thaliana* and (B) sinapoyl acylation of cyanidin 3- *O*- ((xylosyl)- coumaroyl glucose), 5- *O*- glucoside in *Arabidopsis thaliana*

Anthocyanin biosynthesis in *Arabidopsis thaliana* was targeted for characterisation as sequence data was available allowing for sequencing of isolated proteins by MALDI proteomics. Whereas proteins isolated from *Petunia hybrida* would require identification by MS-MS proteomics, which is more difficult.

7.2.2 Use of substrate-affinity chemical probes toward the covalent labelling and isolation of coenzyme A dependent enzymes

As the common substrate moiety involved in both At4CL1 and Gent5AT activity, phenylpropanoates were thought to be the most likely substrate upon which to design a substrate affinity probe. It was also envisaged that a probe based upon a generic acylating motif would facilitate the labelling of a large array of both 4CL and AT enzyme activities, whereas a probe based upon a specific acyl acceptor would limit the use of the probe to specific enzymes only (Chapter 1). To this effect, an azido bearing phenylpropanoid was obtained through collaboration with Prof. R. Field's group (University of East Anglia) and a potential strategy for its use as a photoreactive substrate affinity probe was developed (Figure 7.2).

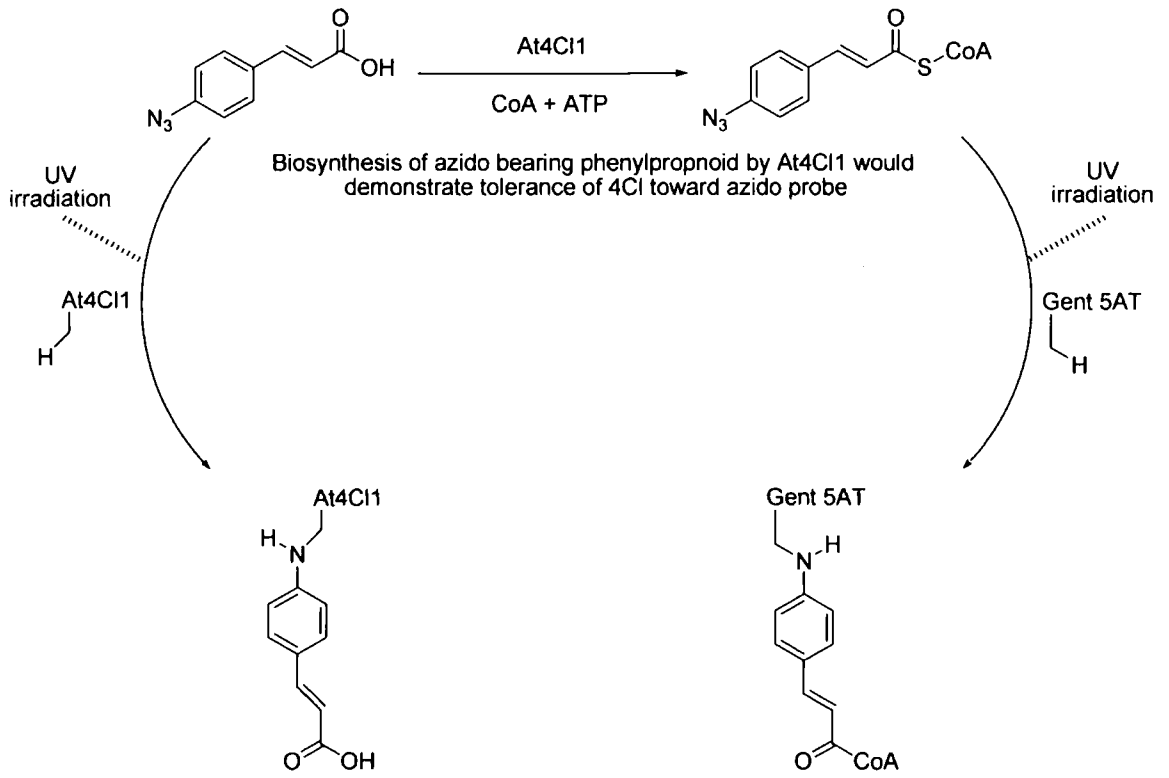


Figure 7.2 Strategy for the covalent labelling of At4CL 1 and Gent5AT with an azido bearing substrate affinity probe

It was necessary to determine that At4CL1 could use 4-azidocinnamic acid as a substrate, primarily to demonstrate binding of the probe to the active site prior to photoactivation. It was also envisaged that 4-azidocinnamoyl coenzyme A could be obtained in this manner, for use as a probe toward acyltransferases. Therefore biosynthesis of 4-azidocinnamoyl CoA was undertaken with 10 μg At4CL1. A product was observed with size-exclusion chromatography and subsequent UV-Vis characterisation demonstrated that this product had a different λ_{max} (350 nm) to those observed for 4-azidocinnamic acid (325 nm) or coenzyme A (310 nm). However, mass spectrometry characterisation of the accumulated product was not conclusive, although the 766.4 mass ion corresponded to a fragmentation of coenzyme A (Figure 7.5).

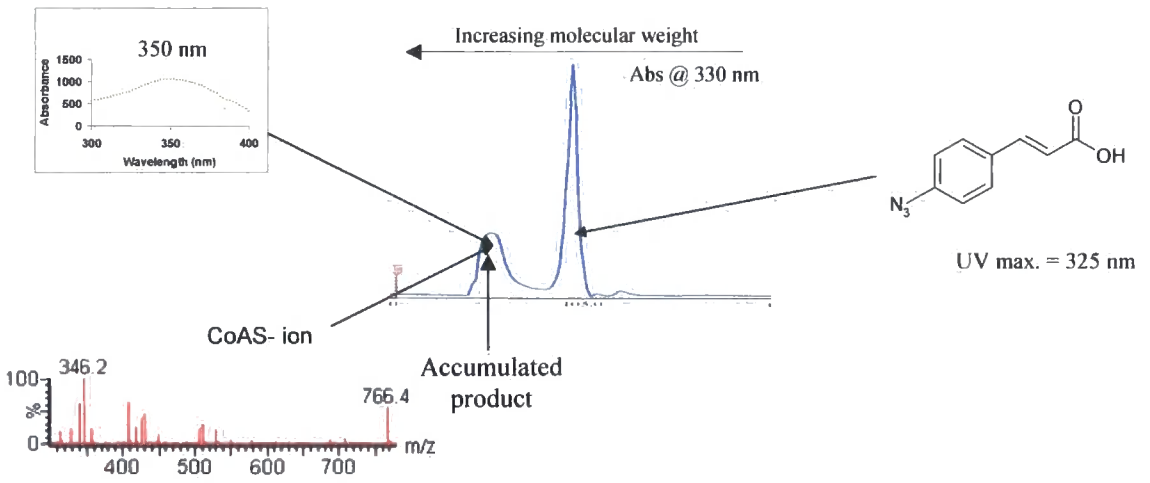
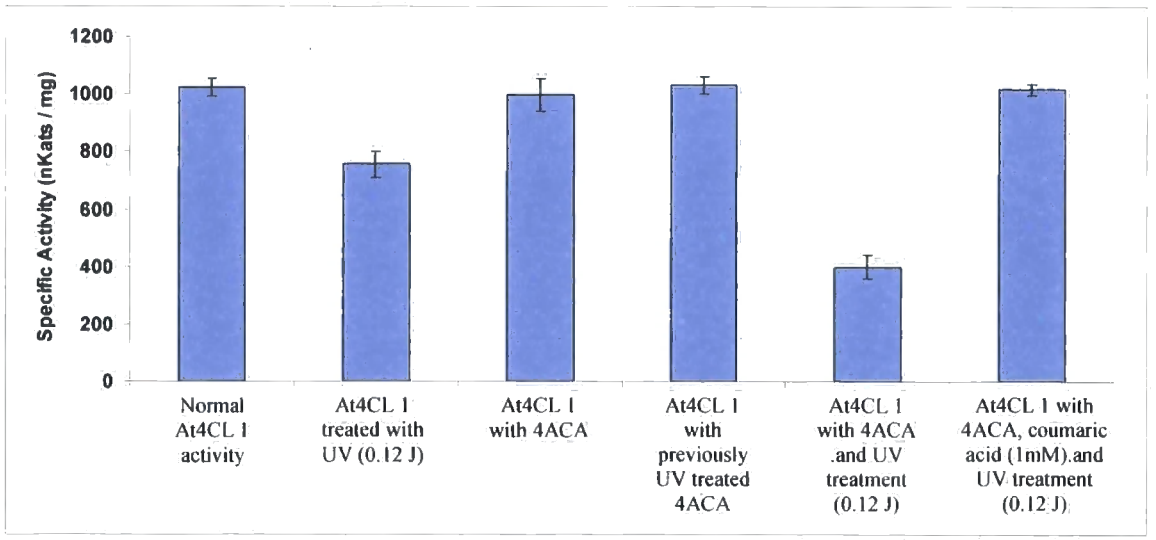


Figure 7.3 Activity of At4CL 1 toward 4- azidocinnamic acid. An enzymatic product with a UV absorption maximum of 350 nm was separated from 4- azidocinnamic acid by size exclusion chromatography. The product was shown to be a conjugate of coenzyme A by mass spectrometry (ES^-) (M^- = 766, M observed = 766)

Instead, the coenzyme A conjugate was subjected to base hydrolysis at pH 10 in order to determine the nature of the enzymatic product. HPLC analysis showed that the

conjugate had partially degraded to reform 4- azidocinnamic acid by comparison to a standard. Therefore the enzymatic product was concluded to be 4-azidocinnamoyl coenzyme A.

With the knowledge that At4CL1 could bind and turnover 4-azidocinnamic acid, the UV activation of the azido group was attempted (see Figure 7.1). To this effect, aliquots of At4CL1 were assayed toward coumaric acid after UV treatment (0.12 Joules at 254 nm) in the presence of 4-azidocinnamic acid. Comparisons were made against protein treated +/- UV, +/- 4-azidocinnamic acid and 4-azidocinnamic acid treated with UV prior to addition to At4CL1 (Figure 7.4).



4ACA = 4- azidocinnamic acid (20 μM)

Figure 7.4 4-azidocinnamic acid inhibition of At4CL 1 by UV initiated reaction of azide moiety with protein infrastructure. A series of experiments were devised to determine the specificity of 4-azidocinnamic acid as a substrate affinity probe.

A significant reduction in activity was observed, when At4CL1 was subjected to UV-irradiation in the presence of 4-azidocinnamic acid. It was also demonstrated that 4-

azidocinnamic acid had to be irradiated when in the presence of the enzyme to cause inhibition. The specific nature of the inhibition was demonstrated by the ability to prevent covalent modification of At4CL 1 by 4-azidocinnamic acid, by competition with coumaric acid (1 mM).

Inhibition of At4CL1 by 4-azidocinnamic acid could be optimised by reduction of total UV irradiation applied to 10 sequential treatments of 4.8 mJoules. This step-down minimised the loss of At4CL1 activity due to UV radiation alone, whilst achieving a greater level of inhibition when in the presence of the substrate affinity probe. However, the percentage of At4CL 1 found to be inhibited could not be greatly improved via further UV treatment nor by the addition of greater amounts of the probe (Figure 7.5).

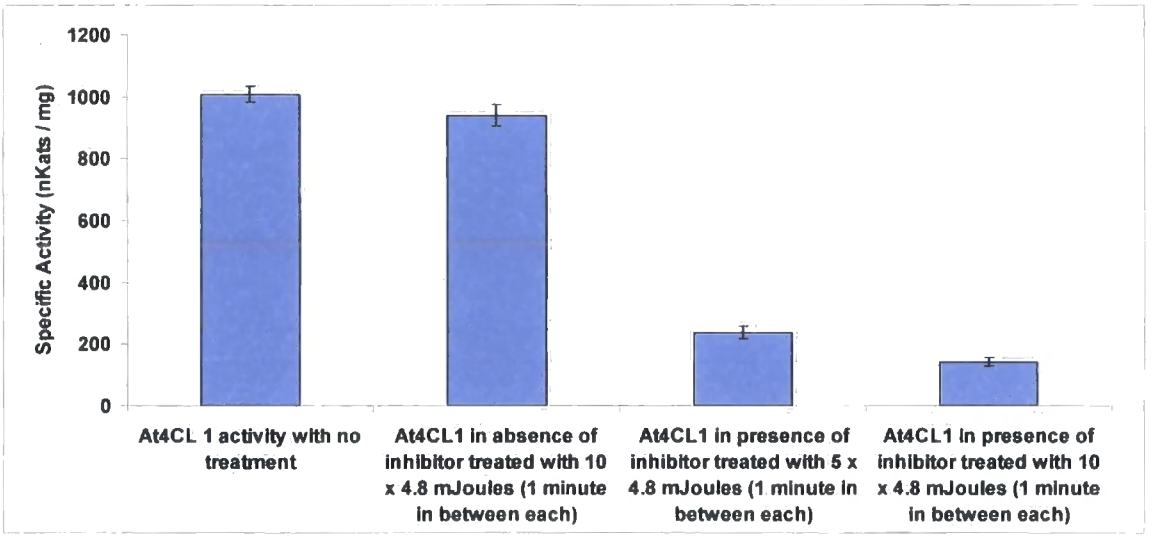


Figure 7.5 Optimisation of the inhibition of At4CL1 with 4-azidocinnamic acid.

In order to demonstrate specific labelling of At4CL1 (M = 61 815 Da), mass spectrometry of a sample of labelled protein was undertaken. This showed both

singular (M = 61 975 Da) and double addition (M = 62 135 Da) of the azido probe to At4CL1, although no further additions were observed (Figure 7.6). This observation, in context with the ability to out-compete the inhibition with coumaric acid substrate, was thought to be due to the presence of two sites capable of binding phenylpropanoids, one being the active site and the other being non-specific.

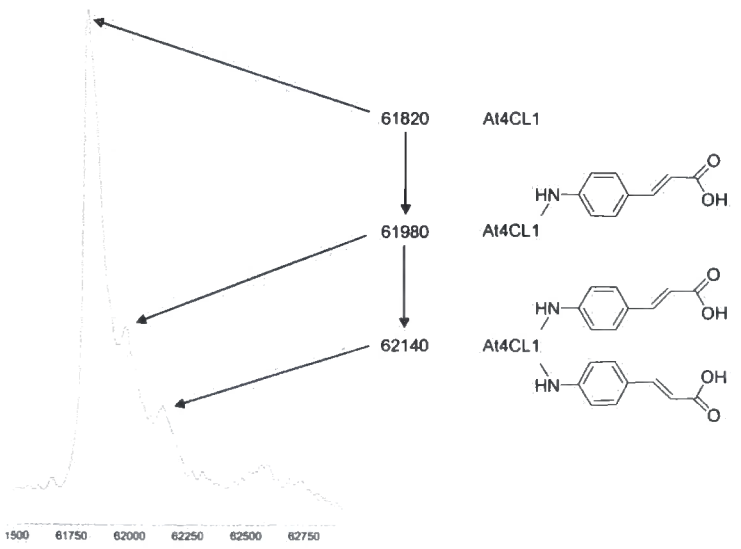


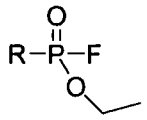
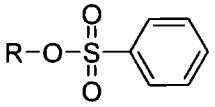
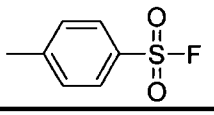
Figure 7.6 Mass spectrometry characterisation of 4- azidocinnamic acid labelled At4CL 1 Mass units are deconvoluted to 10 Da

The labelling of Gent5AT (10 µg) with 4-azidocinnamoyl coenzyme A was then undertaken. A 10-minute coupled reaction with At4CL1 (40 µg) was utilised (section 6.2.5) in order to achieve biosynthesis of the azido-labelled acyl donor *in situ*, before UV irradiation (10 sequential treatments of 4.8 mJoules) of Gent5AT-bound substrate affinity probe was applied to achieve inhibition. The activity of Gent5AT was found to be reduced by 42 +/- 3.4 % in comparison to a control comprising UV treatment, but in the absence of azido probe.

The relatively low efficiency of covalent modification of Gent5AT by 4-azidocinnamoyl coenzyme A suggested that while this reagent would be useful for characterising abundant acyltransferases *in vitro*, such a probe would have limited utility in identifying such low abundance enzymes in crude plant extracts. It was therefore concluded that 4-azidocinnamoyl CoA, in particular, was unsuitable for design of a generic chemical probe for acyltransferases. In contrast, mechanism-specific inhibitors are not recognised by target proteins by their structure and, as such, they are more readily incorporated into chemical probes, although they are less specific. A potential combined approach toward utilising the contrasting probes was therefore envisaged. This would involve the detection and isolation of a catalytically homologous set of enzymes, including BAHD acyltransferases, by chemotype proteomic profiling, whereupon a specific phenylpropanoid inhibitor would be able to identify acyltransferases within this select proteome through competitive labelling. Thus a suitable mechanism-based chemotype inhibitor was sought.

7.2.3 Use of mechanism-specific inhibitors toward the covalent labelling of BAHD acyltransferases

Although there is minimal information concerning the mechanism of coenzyme A dependent acyltransferases in the literature, specific inhibitors of catalytic serine and cysteine residues have previously been shown to be effective inhibitors of BAHD acyltransferase activity[82]. Aliquots of Gent5AT (10 μg) were incubated for one hour in the presence of various mechanistic inhibitors (5 μM) and acyltransferase activity subsequently monitored (Figure 7.7).

Inhibitor	Target amino acid	Specific activity nkats / mg	Relative inhibition %
No inhibitor	-	50.4 +/- 1.6	-
A) 	Serine	25.1 +/- 0.9	50.2
B) 	Cysteine	36.7 +/- 0.3	29.2
C) 	Serine / Cysteine	50.12 +/- 3.2	0

R = Trifunctional probe (section 3.2.2)

Figure 7.7 The relative inhibition of the mechanism – based serine and / or cysteine inhibitors (A) ethyl fluorophosphonate, (B) benzyl sulphonate and (C) *p*- toluene sulphonyl fluoride of Gent5AT activity toward cyanidin 3- 5- *O* – diglucoside and caffeoyl coenzyme A

Based upon these studies, it was decided that fluorophosphonate esters were promising inhibitors for use as chemotype probes toward BAHD acyltransferases. Ideally, total inhibition needed to be attained, in order to maximise the yield of enzyme achieved, as this would aid identification considerably. It was found that increasing the concentration of probe to 10 μM increased the proportional inhibition, with 59 % of original activity abolished. However, fluorophosphonate esters are also highly efficient inhibitors of other enzymes, in particular the carboxyl esterases[143], and previous proteomic work in this field demonstrated that greatly elevated concentrations of fluorophosphonate probes (FPPs) led to non - specific binding and complicated subsequent separation of labelled proteins by SDS-PAGE[144].

Thus a novel approach to toward the selective labelling of BAHD acyltransferases was required in order to maximise the yield of labelled target polypeptide(s) and

hence enhance the potential for successful proteomic characterisation. Comparatively, serine carboxylesterases have been found to undergo rapid inhibition by fluorophosphonates (100 % inhibition within 1 hour[111]), which is in contrast to fluorophosphonate inhibition of the acyltransferases (50.2 +/- 1.7 % after 1 hour).

However, it was observed that the inhibition of catalytic serine by fluorophosphonate could be promoted by UV-irradiation, whereby enhanced reactivity of fluorophosphonate toward serine catalytic motifs. To this effect, 100 % inhibition of Gent5AT (10 μg) was achieved utilising 5 μM fluorophosphonate probe (FPP – Figure 7.7) within one hour, when irradiated 4 times with UV light (20 mJoules - 254 nm) at 15 minute intervals. An aliquot of Gent5AT, which was treated similarly but in the absence of fluorophosphonate, was found to have its activity reduced by 5 % as a direct result of the UV irradiation. This was an unexpected result, which had no prior biochemical basis to the best of our knowledge. However, it appeared a potential solution and therefore further studies were undertaken.

Efficient inhibition of Gent5AT via this method, was interrogated by labelling Gent5AT and observing fluorescent polypeptides by electrophoresis. Therefore equivalent recombinant Gent5AT samples (10 μg) were treated with 5 μM FPP for one hour, both with and without UV treatment, and were isolated by immobilisation upon streptavidin sepharose and separated by SDS-PAGE. Only Gent5AT could be seen to fluoresce and not the chaperonin contaminant (Figure 7.8). This suggested that labelling under UV-irradiating conditions was not only able to enhance inhibition, but was also selective.

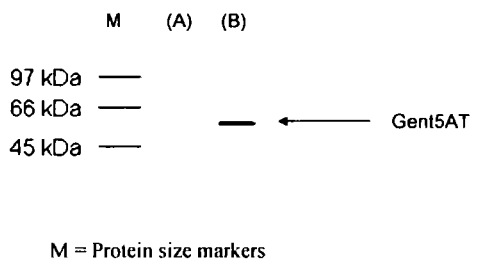
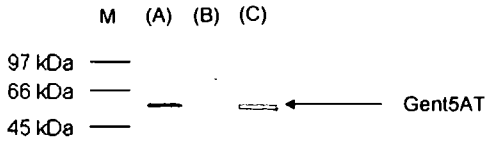


Figure 7.8 SDS-PAGE analysis of equivalent samples of Gent5AT (10 μg), visualised by fluorescence – scanning (excitation – 532 nm, emission – 580 nm) after labelling with FPP (5 μM) under A) normal conditions (1 hour) and B) with 4 x UV treatments at 15 minute intervals.

It was now possible to manipulate labelling conditions in order to achieve maximum yield of labelled acyltransferase. However, a fraction of the acyltransferase activity was still able to undergo labelling in the absence of UV irradiation. If this residual labelling could be initially prevented, activity-based separation of the multitude of enzymes possessing a catalytic serine residue from acyltransferase activities could be achieved

To this effect, the ability to protect a protein from labelling was previously achieved in section 7.2.2, where coumaric acid was successfully utilised to out-compete labelling of At4CL1 by the photoactivated inhibitor 4-azidocinnamic acid. Essentially, this was achieved via competitive-inhibition, which is of importance as it can be reversed. It was therefore of interest to see whether similar effects could be observed for the labelling of Gent5AT with FPP. Thus, a series of FPP inhibition experiments were carried out in the presence of cyanidin diglucoside (acyl acceptor - 1 mM) and caffeoyl coenzyme A (acyl donor - 1mM) respectively. The amount of enzyme labelled was quantified in-gel, via fluorescence (excitation – 532 nm, emission 580 nm, section 2.2.5.2.1) and expressed as a percentage of fluorescence

observed for an equivalent amount of Gent5AT labelled under optimal conditions (Figure 7.9).



M = Protein size markers

Lane	Labelling conditions	Relative intensity of fluorescence
(A)	Optimally labelled under UV irradiation	100 %
(B)	Labelled under UV irradiation with cyanidin diglucoside (1mM)	7 %
(C)	Labelled under UV irradiation with caffeoyl coenzyme A (1 mM)	87 %

Figure 7.9 SDS-PAGE analysis of (A) FPP-labelled Gent5AT and equivalent amounts labelled in the presence of (B) cyanidin 3- 5- O- diglucoside and (C) caffeoyl coenzyme A. The labelling efficiency of FPP in each scenario was determined by in-gel quantification of fluorescence and expressed as a percentage of Gent5AT labelled under optimal conditions (A). M = protein size - markers

The acyl acceptor, cyanidin 3- 5- *O*- diglucoside, was able to largely out-compete FPP from the active site of Gent5AT and reduce covalent inhibition to 7 % of that achieved for optimally labelled Gent5AT. In contrast, caffeoyl coenzyme A was found to have minimal effect upon the FPP-labelling of Gent5AT. This was further evidence to suggest that FPP-labelling of Gent5AT remained to be active-site specific under UV-irradiating conditions.

Having developed a good understanding of the reactivity of FPP toward recombinant Gent5AT, we applied these methods toward the targeting of BAHD acyltransferases within a complex proteome (Figure 7.10).

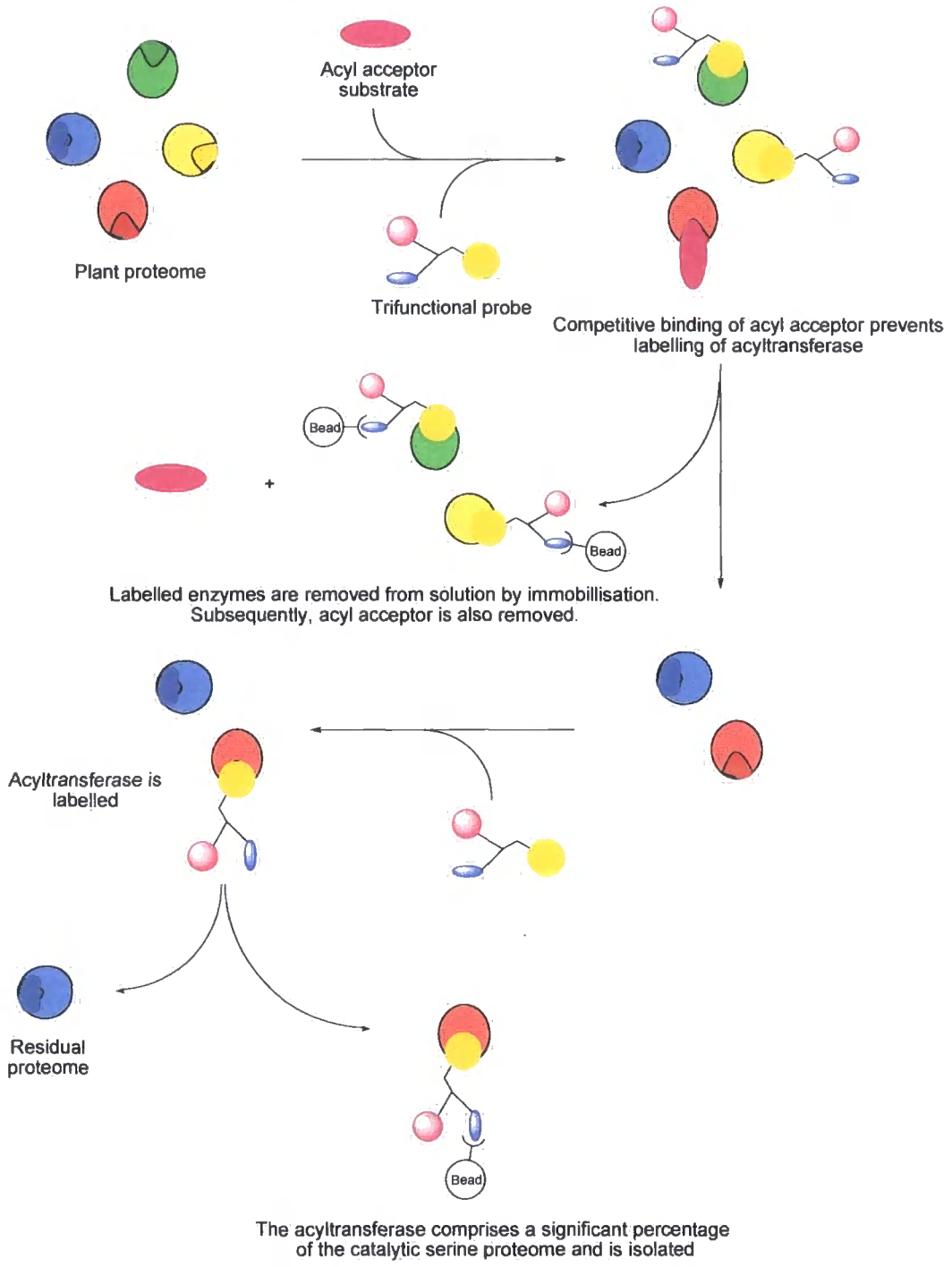


Figure 7.10 A strategy toward the isolation and identification of BAHD acyltransferases from within a complex proteome. Use of the acyl acceptor as a successful competing substrate means that acyltransferases can be prevented from labelling in a highly selective manner, allowing prior removal of other enzymes reactive toward FPP, before removal of the acyl acceptor renders the acyltransferase once more susceptible to labelling by FPP.

7.2.4 Proteomic profiling with chemotyping trifunctional probes

Primarily, it was desirable to demonstrate that Gent5AT could be isolated from a plant proteome in the same manner as the recombinant enzyme. To achieve this crude protein was extracted from flowers of *Gentiana triflora* and incubated for one hour with FPP (5 μM) and cyanidin 3- 5- *O*- diglucoside (1 mM). Proteins able to be labelled under these conditions were subsequently removed from the proteome by immobilisation upon streptavidin sepharose. The remaining soluble proteins were retrieved and extensively purified via several size-exclusion chromatography steps, in order to remove the acyl acceptor (This step was crucial – all residual acyl acceptor must be removed to enable FPP modification of the acyltransferase). The remaining protein was then subjected to a further treatment of FPP in the presence of UV and the absence of the non-covalent inhibitor cyanidin diglucoside (Figure 7.11).

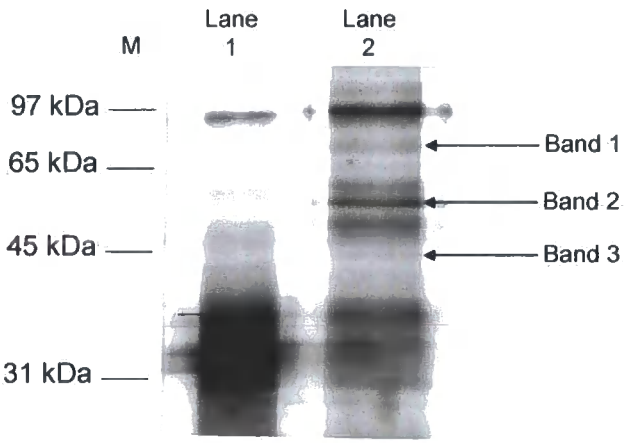


Figure 7.11 Fluorescence imaging of FPP - labelled proteins from the proteome of *Gentiana triflora* flowers, initially in the presence of 1mM cyanidin 3-, 5- *O*- diglucoside (Lane 1) and subsequently after the removal of the anthocyanin and with UV treatment (Lane 2)

Bands 1, 2 and 3 (Lane 2) was observed to be protected from FPP covalent modification in the presence of cyanidin 3-5-*O*- diglucoside, although only band 2

was found to be within the molecular weight region expected for the BAHD acyltransferases (48 – 60 kDa). Band 2 (lane 2) was shown to co-migrate with FPP labelled recombinant Gent5AT by calibration of the protein sizing markers on the respective gels. Recombinant FPP labelled-Gent5AT migrated to the equivalent of 58.0 kDa and the FPP labelled-polypeptide Band 2 to 57.8 kDa. MALDI-based proteomics confirmed the identity of the labelled protein as Gent5AT ($P = 3.4 \times 10^{-18}$, sequence coverage = 51 %, Mowse score = 229) (Appendix I).

The novel acyltransferase enzyme activities from *Arabidopsis thaliana* were targeted for isolation using chemotyping probes with and without the use of competing substrate to identify specific activity with identification by MALDI-based proteomics. Crude protein extracts were prepared from plants, which were actively accumulating both coumaroylated and sinapoylated anthocyanins, as described in Chapter 4. As the identity of the coumaroyl acyl acceptor remained unknown, the acyl acceptor utilised as competitive inhibitor in this study was AtAN 8, which is substrate for a sinapoyl transferase in *Arabidopsis thaliana* (Figure 7.12). AtAN 8 is potentially also a substrate for a malonyl transferase (At3g29590 – see section 4.3) which has previously been characterised, with the enzyme having a molecular weight of 49.5 Da (native protein)[44]. It was intended for this enzyme to be a positive control, as it was expected to be similarly inhibited when in the presence of AtAN 8. Figure 7.13 shows the sequential steps toward isolation of AtAN 8 - selected acyltransferases by chemotype probes. The conclusions drawn from this study are tabulated in figure 7.14 and the data used to draw these conclusions is presented in Appendix I.

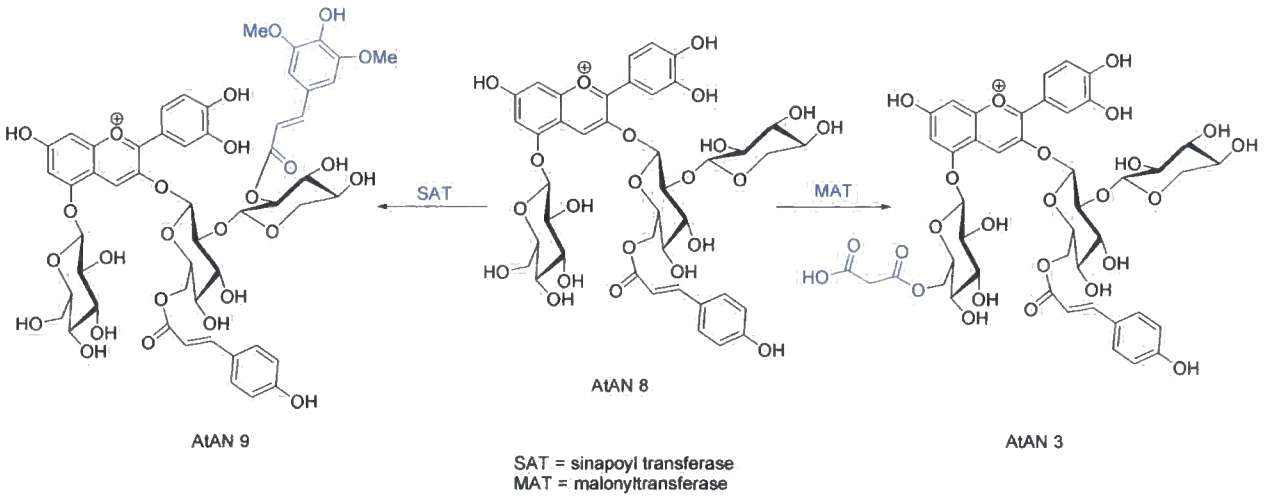


Figure 7.12 The anthocyanin AtAN 8 is substrate for both a sinapoyl transferase (Chapter 4) and a malonyl transferase[125]

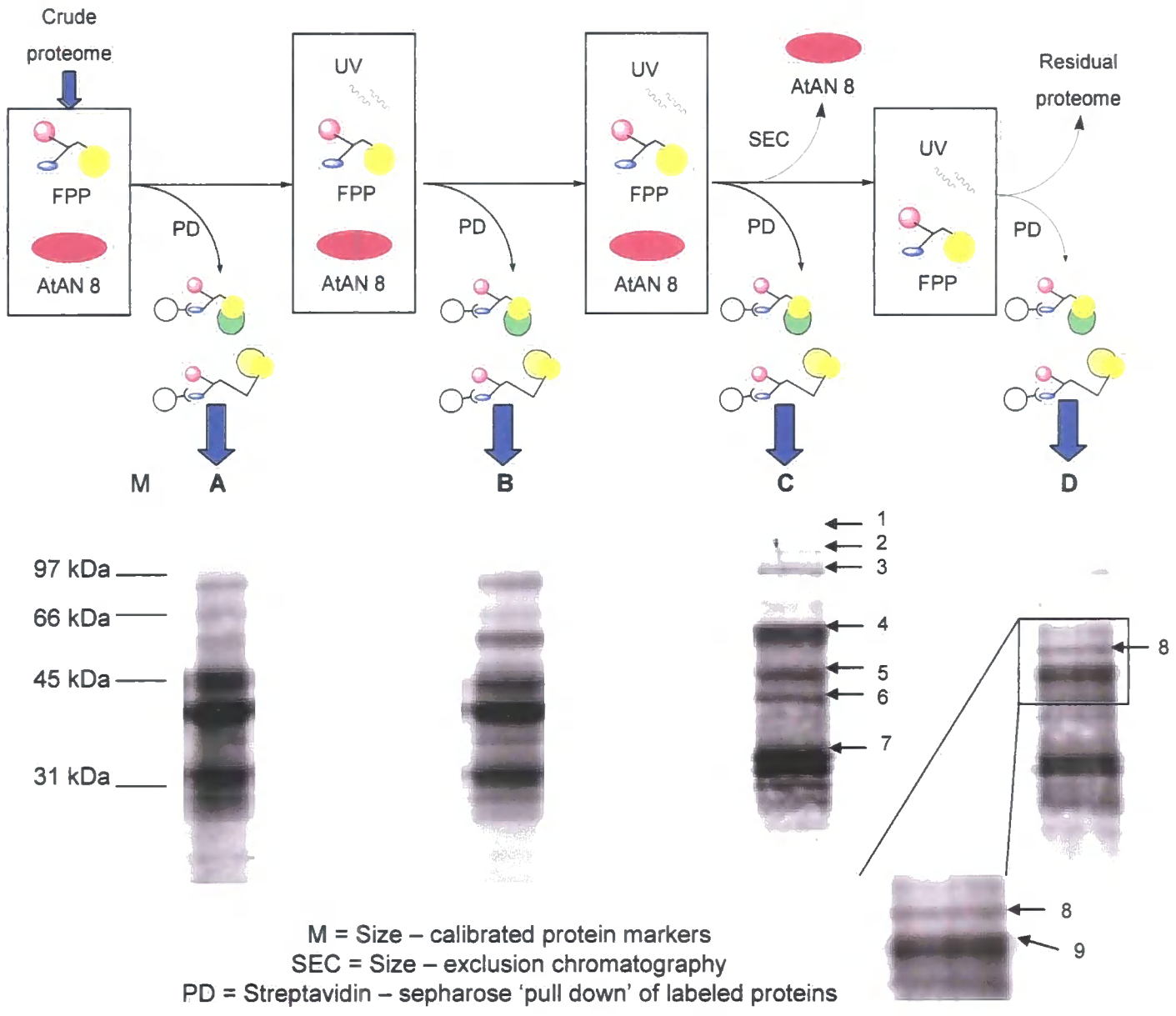


Figure 7.13 Trifunctional fluorophosphonate probe (FPP) proteomic profiling of the catalytic serine – based proteome in high light treated *Arabidopsis thaliana*. A combination of UV irradiation and specific inhibition by the acyltransferase substrate AtAN 8, allowed for activation (lane D) and prevention (lanes A – C) of the labeling of acyltransferases which lie in the molecular weight range 48- 60 kDa (bands 8 (56.2 kDa) and 9 (50.1 kDa) proposed)

Band	Protein identification	TAIR accession	Comments
1 + 2	-	-	Low abundance proteins only detected when the abundance of other protein is low.
3	Prolyl oligopeptidase	At1g76140	Serine peptidase
4	Ribulose bis phosphate carboxylase (RuBisCo)	-	Identified as RuBisCo, likely to be contaminant
5	Glyceraldehyde 3-phosphate dehydrogenase	At1g13440	-
6	AtCXE12	At3g48690	Carboxyl esterase known to react with TFFP probes
7	-	-	-
8	Ribulose bis phosphate carboxylase (RuBisCo)	-	Was only labelled in the absence of AtAN 8. RuBisCo likely to be contaminant
9	Ribulose bis phosphate carboxylase (RuBisCo)	-	Was only labelled in absence of AtAN 8. RuBisCo likely to be contaminant

Figure 7.14 Identification of the polypeptides indicated in figure 7.13 (1 to 9) based upon MALDI-TOF proteomic analysis. Several polypeptides were identified, however, those in the molecular weight region 48 – 60 kDa were due to contamination with RuBisCo

The two polypeptides corresponding to band 8 and band 9 were found to be highly sensitive to the presence of AtAN 8 and as such were only observed to be labelled after removal of the acyl acceptor (Lane D – Figure 4.13). However, MALDI analysis of all bands corresponding to polypeptides of molecular weight 49 – 60 kDa (bands 4, 8 and 9) were found to be contaminated with ribulose 1, 5, - bisphosphate carboxylase / oxygenase large subunit (RuBisCo) – a common contaminant of green plant

proteomic studies[52]. RuBisCo is not labelled by FPP, but it is susceptible to non - specific interaction with streptavidin sepharose beads due to its abundance. Extensive washing of immobilised proteins with detergent and water was incorporated into the methodology. However, although this was beneficial, it was not sufficient to remove RuBisCo from the relatively low abundance proteins. RuBisCo contamination was concluded to be an inherent problem with the method of isolation of specifically labelled proteins via biotin - streptavidin sepharose immobilisation.

The two polypeptides of interest did, however, correspond to the approximate molecular weight expected for acyl transferases. Band 8 was found to have an estimated molecular weight of 54 kDa and band 9 an estimated molecular weight of 50 kDa, by calibration of the protein size markers. These are in the correct range for the native anthocyanin malonyl transferase present in *Arabidopsis thaliana*, which has a native mass of 49.55 kDa[44]. However, further studies are required to enable the efficient identification of acyltransferases by this method.

7.3 Discussion

Two contrasting proteomic strategies were explored in this chapter, with the aim being to isolate and identify BAHD acyltransferases efficiently. A UV-activated substrate affinity probe was employed as a highly specific inhibitor and was found to have potential as a probe toward phenylpropanoyl coenzyme A ligases. However, relatively weak inhibition of acyltransferase activity and a lack of bio-compatible reactive sites for synthetic attachment of a probe scaffold, meant that an alternative strategy was explored. The mechanism-based inhibitor, ethyl fluorophosphonate, was

found to effect selective inhibition of a high percentage of acyltransferase activity, particularly in the presence of UV light, and was therefore incorporated into a trifunctional probe scaffold. This proteomic strategy was successful in isolation of serine-based enzyme activities and from within this highly-enriched protein fraction, both known and unknown acyltransferases were identified by the selective out-competing of the probe from acyltransferase active sites by their respective acyl acceptor. This apparently enzyme-specific proteomic approach represents a novel method for the future profiling of CADATs.

7.3.1 Mechanism of UV 'activation' of fluorophosphonate inhibitor

The mechanism by which UV radiation enhances the reactivity of fluorophosphonate toward Gent5AT was not obvious. To the best of our knowledge, there is no previous literature on the topic of enzyme photoinhibition which specifically relates to fluorophosphonate inhibitors or enzymes with serine catalytic dyads/triads. In general terms, inhibitors that require 'photoactivation' possess chromophores that absorb in the UV-A/B/C range, for example azides[117] or *ortho*-nitrobenzyl compounds[156], and following absorption subsequently decompose to form highly reactive species. However, none of the residues that are directly involved in FP inhibition – FP, serine, histidine or aspartate – are known to absorb UV irradiation at 254 nm.

The residues in enzymes that absorb radiation at this wavelength are phenylalanine, tyrosine and tryptophan[159]. Although none of these residues are directly involved in catalysis it is possible that, if proximate to the active site, absorbance by these residues could cause localised heating and enhance the rate of reaction through this

medium. However, widespread absorption of irradiation by these residues throughout the protein would ultimately lead to photo-oxidation of the protein through the formation of reactive excited states and radical species[159]. This eventuality causes irreversible damage to a protein's infrastructure and subsequently loss of activity. Although significant loss of enzyme activity was not observed under the control conditions utilised in this study it would be difficult to ascertain the level of absorption achieved by amino acid residues proximate to the active site.

Protein absorbtion of UV-light can also be facilitated by bound prosthetic groups[159]. Gent5AT activity is known to be dependent upon manganese (II) ions, although it is not known how this ion interacts with the mechanism of acyltransfer[81]. However, several manganese-dependent proteins have been characterised with respect to their photosensitivity[157] and manganese (II) complexes were shown to be particularly sensitive to light in the UV-C range (100-280 nm)[157]. One such divalent manganese is present in photosystem II as a histidine-manganese-glutamate complex[158], which could be comparable to a possible histidine-manganese-aspartate complex in CADATs (Figure 7.15). Alternatively manganese is also known to co-ordinate to carboxylic acid-functionalised amino acids, such as aspartate (Figure 7.15).

Figure 7.15 Co-ordination of Manganese (II) to histidine and aspartate amino acid residues

Excess absorption of light in the UV-C range by manganese-dependent enzymes is generally thought to cause inhibition by breaking the relatively weak co-ordination bonds that bind the metal to the enzyme[157]. In terms of FP inhibition of manganese (II)-dependent CADATs, it is possible that removal of the metal ion increases the reactivity (nucleophilicity) of the enzyme toward FP. For instance, photo-dissociation of Mn co-ordination to aspartate could reinstate aspartate hydrogen bonding to histidine, in turn increasing the reactivity of serine toward FP. Alternatively, light-induced Mn (II) dissociation may enhance both FP inhibition and acyltransfer activity. Therefore it may not be coincidence that acylated flavonoid photoprotectants are favourably produced under high-light conditions *in planta* (Chapter 4) and that the activity of the acyltransferase mechanism is potentially photosensitive.(Chapter 7).

'Activation' of the phosphonate moiety to form a reactive species, such as a radical, was thought to be a further possibility. This could be initiated directly or through reaction with photoinduced oxygen radicals. Electron spin resonance studies could be performed upon fluorophosphonates to assess their ability to form radicals under UV-irradiating conditions. However, if formation of a highly-reactivephosphonate species was the reason for the observed UV-dependent rate enhancement of FP inhibition, then a large increase in non-specific labelling of proteins would be expected. This would occur through rapid reaction with non-specific amino acid residues, the chances of which would be increased relative to affinity-based probes owing to the low affinity of FP-probes for target enzymes[144].

7.3.2 Improvement upon proteomic strategy

Contamination of the isolated polypeptides corresponding to bands 8 and 9 with RuBisCo, was the major flaw in the novel proteomic approach. All of the acyltransferases characterised to date have been of the molecular weight which are likely to be affected by this contamination and therefore future identification of these enzymes is likely to be similarly hindered. Further preparation of labelled proteins (denatured) by HPLC would allow the separation of labelled proteins from RuBisCo and is the most obvious solution. In addition, the bulky fluorescent tag would facilitate purification by affording facile identification. However, a potentially more elaborate approach would be to employ a non-denaturing specific release mechanism of immobilised labelled proteins, thus leaving non-labelled proteins and intact streptavidin behind. To this end, a considerable amount of research has already been undertaken in order to develop mild release mechanisms of biotin-streptavidin interactions[145]. For example, a directed mutagenesis study upon streptavidin was explored in an attempt to identify mutant proteins able to bind biotin reversibly[146]. A mutant of streptavidin was developed with an inherently weaker interaction with biotin, which could be displaced with successive additions of biotin, effectively reversing the original interaction[145]. The development of cleavable linkers has also been intensely researched and this could be applied to proteomic probes. The most commonly utilised is a disulphide bridge which is susceptible to cleavage under reducing conditions[147]. The reversal of the original inhibition is also an option, although this is a less studied approach. It is thought that there is scope for the reversal of the fluorophosphonate inhibition in this instance (Figure 7.16).

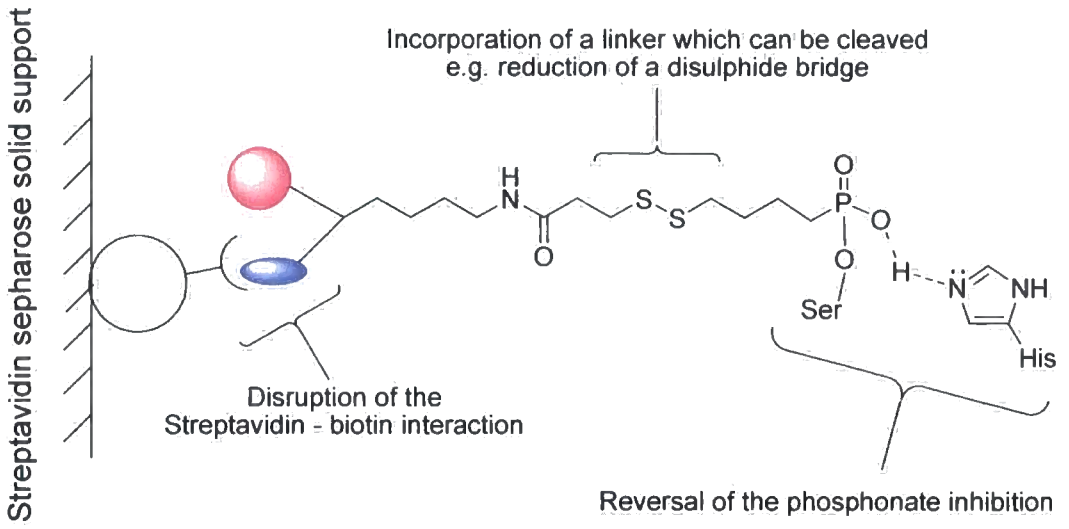
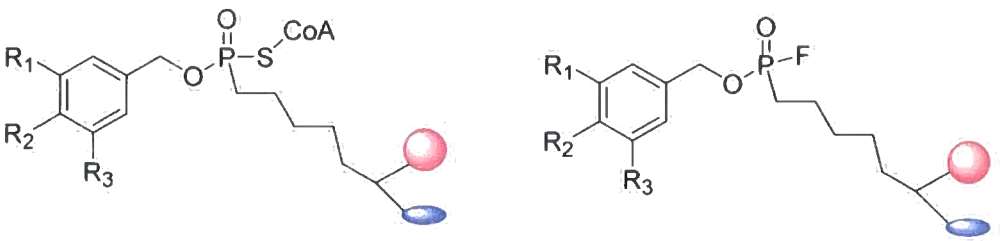


Figure 7.16 Mechanisms of *specific* release of labelled protein post-immobilisation upon sepharose-bound streptavidin supports.

Of the three possibilities, the ability to selectively reverse the covalent inhibition of the catalytic dyad would have application aside from achieving molecular identification of enzyme activities. In particular, the released protein would be active and could therefore be characterised biochemically. As the protein would not be denatured, analysis of proteins by 2-D electrophoresis would be possible. This would enhance separation from non-labelled proteins and increase detection of low abundance enzymes.

Phosphonate inhibition of serine catalytic dyads is due to the mechanisms inability to *catalyse* a second nucleophilic addition of water / alcohol onto the phosphonate. Therefore the introduction of a reactive nucleophile to FP-immobilised enzyme in sufficient concentration could theoretically reverse the inhibition. In support of this theory, oximes have been found to be able to reverse the fluorophosphonate inhibition of acetyl choline esterase in humans[148]. However, it may also be possible stimulate hydrolysis of the phosphonate ester simply by increasing the pH of the surrounding environment, depending upon protein stability.

Synthesis of a more specific inhibitor would potentially facilitate the proteomic probe-based isolation of acyltransferases, by reducing the amount of non-targeted labelling, as well as other applications such as *in vivo* chemical genetics[149]. Such probes would preferentially be based upon structures of acyl donors, as the least variable substrate in enzymatic acylation and as they posses the carbonyl chemistry successfully mimicked by the phosphonyl chemotype probe. The probe scaffold is necessarily conjugated through a carbon-phosphorous bond to prevent its hydrolysis upon inhibition. This dictates that the aromatic moiety be an alkoxy substituent to allow its elimination in preference to the serine phosphanyl ester. Incorporation of acyl donor motifs and phosphonate inhibitors are shown in Figure 7.17.



R = H, OH, OMe, or F

Figure 7.17 Design of bioisosteric chemotype probes toward BAHD acyltransferases based upon phenylpropanoyl coenzyme A acyl donor and fluorine leaving group analogue

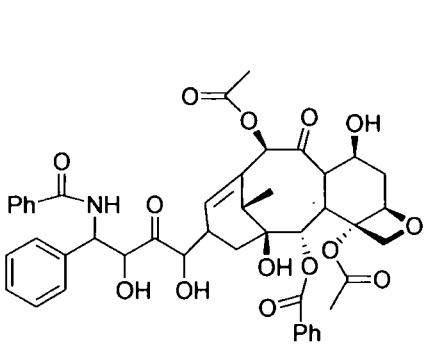
The use of chemical probes for the discovery of BAHD acyltransferases would benefit greatly from achievement of the discussed improvements to methodology and specificity. Use of these methods would facilitate isolation of target enzymes for specific application, however the major advantage of this approach is the production of a diverse library of acyltransferase enzyme activity, which could only be achieved with a method for their rapid and selective isolation. Such a resource would enable

comparative structural and biochemical characterisations to be made, yielding much needed understanding of aspects of acyltransferase activity and substrate selectivity. This in turn would open up the potential for application of BAHD acyltransferases on a new scale, not only in biosynthesis of natural products, but also in biosynthesis of second generation natural products and modification of synthetic molecules.

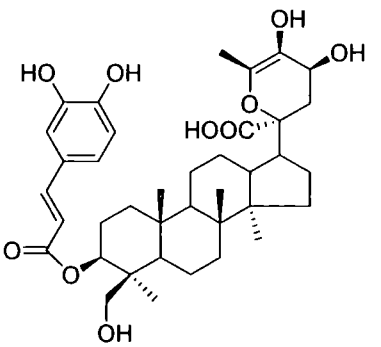
Chapter 8

Discussion and future work

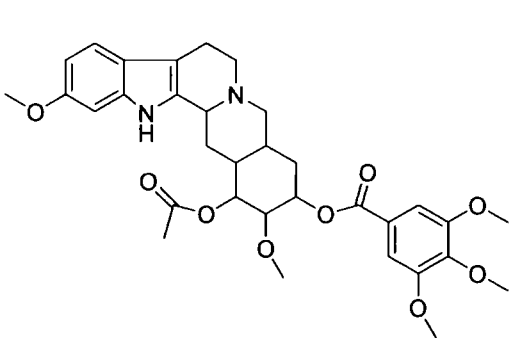
An increasing number of acylated plant natural products are being identified as having pharmaceutical properties where the acylating moiety is identified as being crucial for the bioactivity of these natural products (Figure 8.1). This study has explored the potential for the biological incorporation of acylating moieties onto these natural products, which has proved to be the only viable means of production of complex natural products such as Taxol (section 1.6).



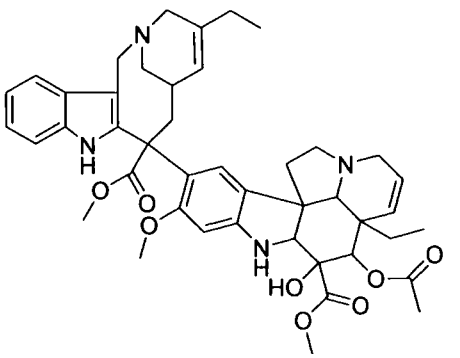
Taxol
Commercial anti - cancer pharmacon



Trihydroxydammarene caffeate
Strong anti tumor activity toward cervical cancer



Reserpine
Rauvolfia serpentina
Decreases blood pressure



Navelbine
Rauvolfia serpentina
Anti - cancer pharmacon

Figure 8.1 Natural products bearing acyl moieties which have been shown to be crucial for bioactivity[78,79,150,151]

A unique insight was gained into the application of coenzyme A ligase and BAHD acyltransferase activity *in vitro,* to achieve biosynthesis of natural products bearing an array of acylating moieties. Through a one - pot biosynthetic approach, inherent problems such as the enzymatic hydrolysis of the acyl donor and competitive inhibition by coenzyme A were overcome. The importance of coenzyme A ligase activity was evident in biosynthesis of these natural products. In particular, their biased activity toward the acylation precursors was thought to have resulted in the lack of, or low level, accumulation of novel acylated flavonoids in *Arabidopsis thaliana* and *Petunia hybrida*, as described in Chapter 4. Thus coenzyme A ligase activity is likely to hinder any future attempt to preferentially biosynthesise natural products bearing novel acylations within the respective plant. Therefore incorporation of non – natural acylating moieties must be performed either *in vitro* or in a secondary host organism, which incorporates a tolerant coenzyme A ligase (Chapter 5). To this effect, Loncaric *et. al.* demonstrated the bioconversion of exogenously supplied 10-deacetyl baccatin III to baccatin III (Taxol biosynthetic pathway) in cultures of *E. coli*, which were over - expressing the respective acetyltransferase. They relied on the endogenous biosynthesis of acetyl coenzyme A by the bacterium to provide the acyl donor. However, adopting a combinatorial 4CL - AT in cultures of *E. coli* would allow both acyl acceptor and a series of novel acyl precursors to be exogenously provided. The modification of the coenzyme A ligase to incorporate the novel acylation precursors, as discussed in chapter 5, would enhance exclusive acylation by novel acyl moieties. Additionally, the *in vivo* bioavailability of the expensive coenzyme A and ATP would significantly reduce the cost of production. Thus incorporation of both acyltransferase and the desired coenzyme A ligase into the same expression plasmid in sequence would be of interest (Figure 8.2).

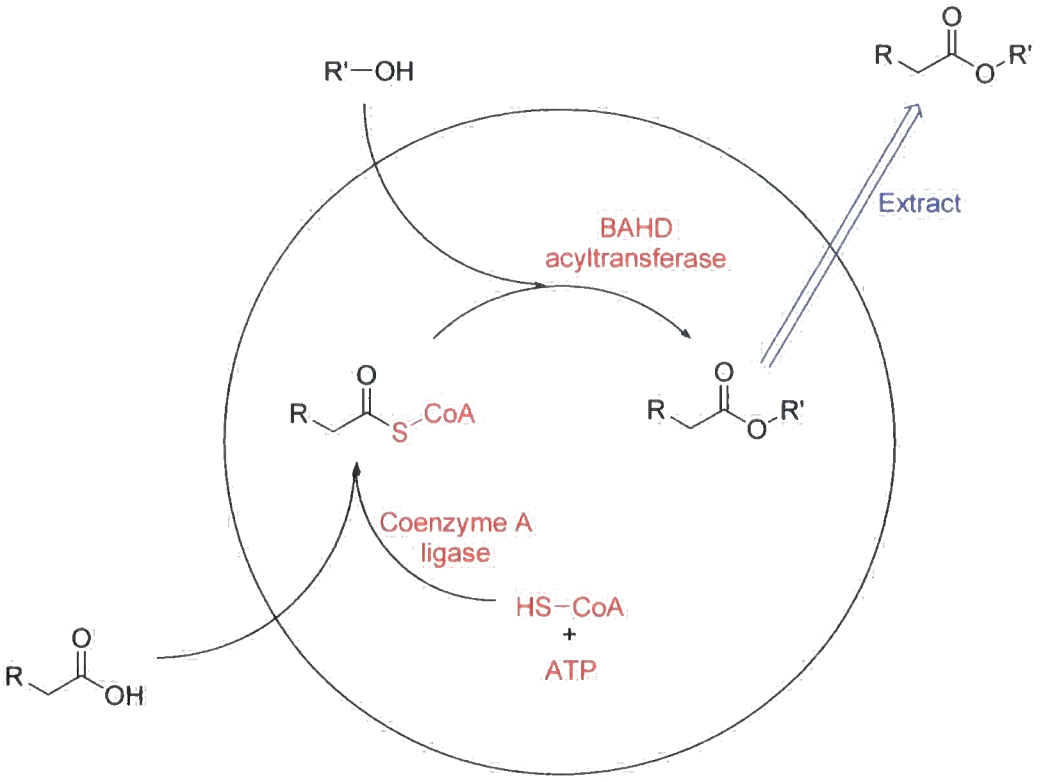


Figure 8.2 Co – expression of both acyltransferase and coenzyme A ligase in cells of *E. coli* would allow biosynthesis of acylated natural products *in vivo*. Components in red are found inside the bacterium, whilst the acyl acceptor (R' – OH) and acylation precursor must be supplied exogenously

In order to biosynthesise natural products bearing non – natural acylations by this route, the acyl acceptor must first be extracted from the respective plant, due to the complexity of many natural product scaffolds. In order to prevent plants from imparting the endogenous acyl group onto the natural product, the respective acyltransferase activity must be nullified. There are several methods which could achieve this, including RNA interference[152] or tDNA insertion mutagenesis[153]. However, it would be of interest to use the acyltransferase inhibitors developed in Chapter 7, in particular the fluorophosphonate ester, in a reverse chemical genetic approach[154]. Exogenous supply of the fluorophosphonate warhead to cell cultures expressing a given natural product pathway which incorporates acylation, would

allow the selective inhibition of the acyltransferase activity and cause accumulation of the acyl acceptor (Figure 8.3).

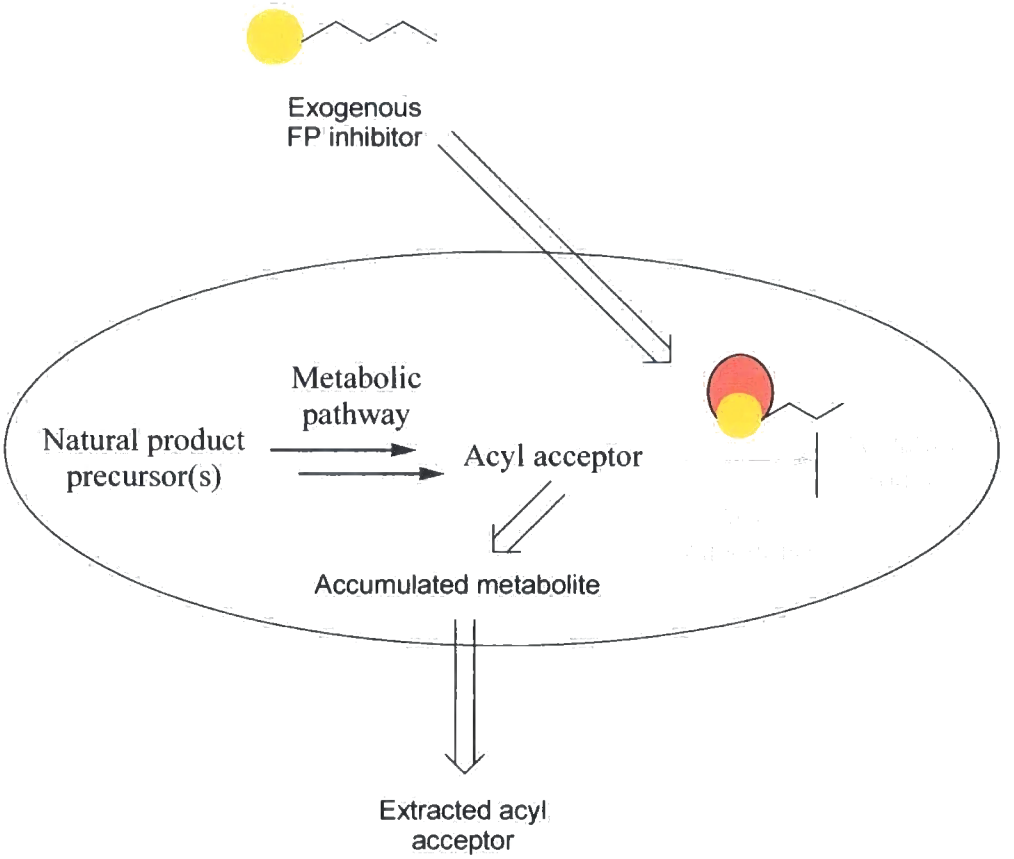


Figure 8.3 Use of a fluorophosphonate inhibitor *in vivo* to effect accumulation of an acyl acceptor

A more selective array of inhibitors based upon the acyl donors were envisaged for a chemical genetics approach. The acyl donors were thought to be the best candidates to form the structural basis of a chemical probe, as they are relatively conserved in comparison to the acyl acceptors in natural product acylation. It has been shown that the fluorophosphonate moieties are turned over by acyltransferases as mimics of carboxyl groups. Therefore the replacement of the carboxyl group present in acylating moieties with a fluorophosphonate group, would potentially introduce increased specificity of the mechanism – based probe toward acyltransferases, as opposed to other serine – based enzyme activities. It may also be possible to target specific

acyltransferases by incorporating aromatic R substituents onto phenylpropanoid and benzoyl probes (Figure 8.4).

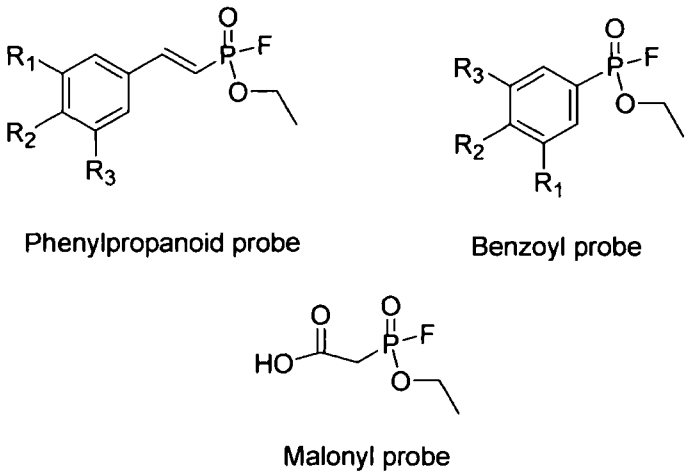


Figure 8.4 Fluorophosphonate probes based upon acyl donor structures for use in enzyme – selective chemical genetics studies

Currently, one of the major drawbacks of *in vivo* production of natural products is their subsequent separation from the resulting crude extract, which often contains other compounds of a highly similar chemical nature e.g. precursors or diversified natural products. To this effect, an unnatural acylating group bearing a recognition or affinity tag could be selectively incorporated onto the desired natural product scaffold utilising BAHD acyltransferase activity. Research in the field of targeted drug delivery has identified a cascade - cleavable linker group[85], which is conjugated to its pharmacon by esterification. If a coenzyme A ligase – acyltransferase combination could be modified to enable selctive incorporation of this molecule into the structure of a bioactive natural product it would afford facile purification (Figure 8.5).

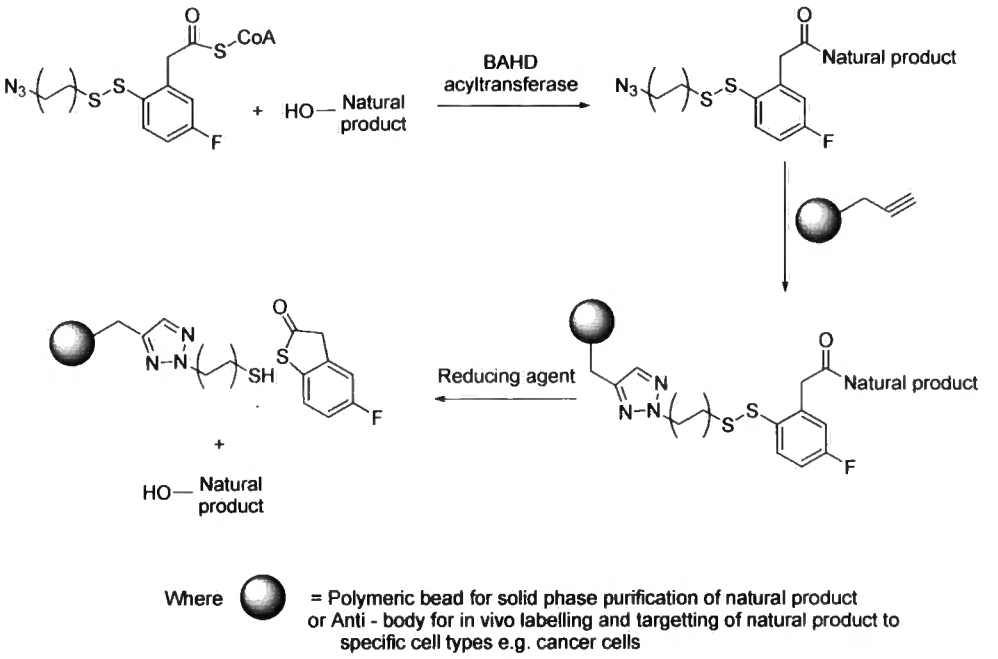


Figure 8.5 Potential use of BAHD acyltransferases in the derivitisation of natural products with selectable affinity molecules

The mechanism by which enzymes are inhibited by fluorophosphonate esters occurs by the mechanism outlined in figure 8.6 (Chapter 7). Thus serine is concluded to play a catalytic role in acyltransferase activity as part of a serine – histidine catalytic dyad, as opposed to the serine – histidine – aspartate catalytic triad, which is employed in carboxyl esterases.

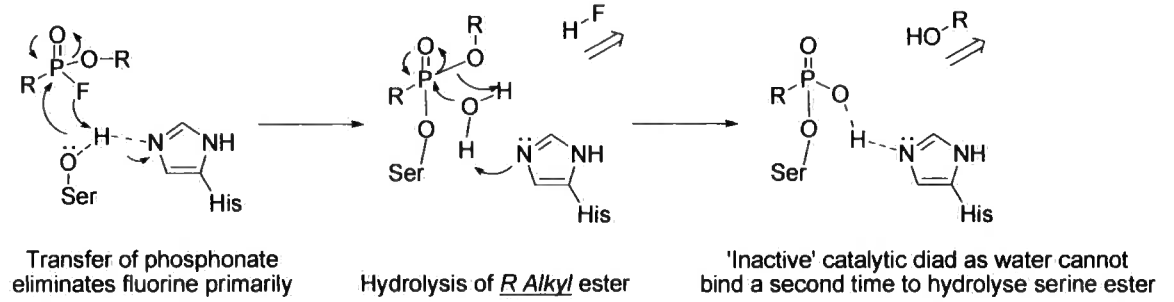


Figure 8.6 Inhibition of acyltransferase activity by ethyl fluorophosphonate

It was observed that inhibition of acyltransferases by FPP occurred at a slower rate in comparison to esterases. This was concluded to be an indicator of the nucleophilic strength of serine within its catalytic environment. The positive effect of UV irradiation upon the reactivity of FPP toward serine in Gent5AT is also indicative of a relatively soft nucleophile in acyltransferases. UV irradiation is thought to cause excitation of the phosphonyl moiety (reducing the bond order) and in turn increase the electrophillicity of phosphorous, thus promoting nucleophillic attack. A less nucleophilic serine is well suited to the transacylation of labile coenzyme A thioesters and adds weight to the probability of coenzyme A tagging of acyl moieties to ensure specific acyl transfer (Chapter 6). Therefore other serine - based enzyme activities such as esterases, amidases, dehydratases, hydrolases, *N*- acyltransferases, may have a serine of a particular nucleophilic strength to suit their respective role. To this end, the synthesis of a series of inhibitors of varying electrophillic strength would therefore be able to probe for all manner of serine – based enzyme activities and be designed with the chemical nature of the substrate in mind. Such inhibitors may include carbamates, boronic acids, sulphonyl fluorides amongst others.

In order to realise many of the described elaborations on the potential application of acyltransferases, a greater understanding of these enzymes and their substrate selection must be achieved. The use of chemical probes to increase the rate at which acyltransferases are able to be isolated and characterised would enhance progress in this field and enable comparative structural characterisation of acyltransferases with activity toward varying acyl donors and / or acyl acceptors. Although the difficulties encountered in Chapter 7 must first be overcome.

References

1. Villorbina, G. Solid-phase synthesis of a combinatorial library of dihydroceramide analogues and its activity in human alveolar epithelial cells. *Bioorg. & Med. Chem.* **15**, 50-62 (2007).
2. Deka, K., Laskar, M., Baruah, J. Carbon-nitrogen bond cleavage by copper(II) complexes. *Polyhedron* **25**, 2525-2529 (2006).
3. Somu, R., Boshoff, H., Qiao, C., Bennett, E. M., Aldrich, C. C. Rationally designed nucleoside antibiotics that inhibit siderophore biosynthesis of mycobacterium tuberculosis. *J. Med. Chem.* **49**, 31-34 (2006).
4. Knights, K. M. Role of hepatic fatty acid: coenzyme A ligases in the metabolism of xenobiotic carboxylic acids. *Clin.Exp. Pharm. Phys.* **25**, 776-782 (1998).
5. Diedrich, G. Ribosomal protein L2 is involved in the association of the ribosomal subunits, tRNA binding to A and P sites and peptidyl transfer. *Embo J.* **19**, 5241-5250 (2000).
6. D'Auria, J. C. Acyltransferases in plants: a good time to be BAHD. *Curr. Opin. Plant Biol.* **9**, 331-340 (2006).
7. Li, A. X., Steffens, J. C. An acyltransferase catalyzing the formation of diacylglucose is a serine carboxypeptidase-like protein. *Proc. Nat. Acad. Sci.* **97**, 6902-6907 (2000).
8. Cammenberg, M., Hult, K., Park, S. Molecular basis for the enhanced lipase-catalyzed N-acylation of L-phenylethanamine with methoxyacetate. *Chembiochem* **7**, 1745-1749 (2006).
9. Santaniello, E., Ferraboschi, P., Grisenti, P. Lipase-catalyzed transesterification in organic solvents - Applications to the preparation of enantiomerically pure compounds. *Enz. Micr. Tech.* **15**, 367-382 (1993).
10. Schellenberger, V., Jakubke, H. D. Protease-catalyzed kinetically controlled peptide-synthesis. *Ang. Chem.* **30**, 1437-1449 (1991).
11. Faber, K. Biotransformations of non-natural compounds: State of the art and future development. *Pure Appl. Chem.* **69**, 1613-1632 (1997).
12. Dodson, G., Wlodawer, A. Catalytic triads and their relatives. *Trend Bioch. Sci.* **23**, 347-352 (1998).
13. Szerszen, J. B., Szczyglowski, K. & Bandurski, R. S. Ia_{glu}, a gene from zeamays involved in conjugation of growth hormone indole-3-acetic. *Science* **265**, 1699-1701 (1994).
14. Barak, R., Abouhamad, W. N., Eisenbach, M. Both acetate kinase and acetyl coenzyme A synthetase are involved in acetate-stimulated change in the direction of flagellar rotation in *Escherichia coli. J. Bact.* **180**, 985-988 (1998).
15. Ehlting, J. Three 4-coumarate : coenzyme A ligases in *Arabidopsis thaliana* represent two evolutionarily divergent classes in angiosperms. *Plant J.* **19**, 9-20 (1999).
16. Watkins, P. A. Fatty acid activation. *Prog. Lip. Res.* **36**, 55-83 (1997).
17. Ma, X. Y. Vinorine synthase from Rauvolfia: the first example of crystallization and preliminary X-ray diffraction analysis of an enzyme of the BAHD superfamily. *Bioch. Biophys.* **1701**, 129-132 (2004).

18. Suzuki, H. cDNA cloning, heterologous expressions, and functional characterization of malonyl-coenzyme A : anthocyanidin 3-*O*-glucoside-6 "-*O*-malonyltransferase from dahlia flowers. *Plant Phys.* **130**, 2142-2151 (2002).
19. Jaeger, K. E., Dijkstra, B. W., Reetz, M. T. Bacterial biocatalysts: Molecular biology, three-dimensional structures, and biotechnological applications of lipases. *Ann. Rev. Microb.* **53**, 315-319 (1999).
20. Zocher, R. Biosynthesis of taxol: Enzymatic acetylation of 10-deacetylbaccatin-III to baccatin-III in crude extracts from roots of *Taxus baccata. Biochem. Biophys. Res. Comm.* **229**, 16-20 (1996).
21. Nakayama, T., Suzuki, H., Nishino, T. Anthocyanin acyltransferases: specificities, mechanism, phylogenetics, and applications. *J. Mol. Cat. B-Enzymatic* **23**, 117-132 (2003).
22. Hertweck, C., Jarvis, A. P., Xiang, L. K., Moore, B. S., Oldham, N. J. A mechanism of benzoic acid biosynthesis in plants and bacteria that mirrors fatty acid β-oxidation. *Chembiochem* **2**, 784-791 (2001).
23. Winkel-Shirley, B. Evidence for enzyme complexes in the phenylpropanoid and flavonoid pathways. *Physiologia Plantarum* **107**, 142-149 (1999).
24. Harborne, J. B., Williams, C. A. Anthocyanins and other flavonoids. *Nat. Prod. Rep.* **15**, 631-652 (1998).
25. Negrel, J., Martin, C. The biosynthesis of feruloyltyramine in *Nicotiana tabacum. Phytochem.* **23**, 2797-2801 (1984).
26. Vessey, D., Kelley, M. Characterization of the reaction mechanism for the XL-I form of bovine liver xenobiotic/medium-chain fatty acid:CoA ligase. *Biochem. J.* **357**, 283-288 (2001).
27. Beuerle, T., Pichersky, E. Purification and characterization of benzoate : coenzyme A ligase from *Clarkia breweri. Arc. Biochem. Biophys.* **400**, 258-264 (2002).
28. An, J. H., Lee, G. Y., Jung, J. W., Lee, W., Kim, Y. S. Identification of residues essential for a two-step reaction by malonyl-CoA synthetase from *Rhizobium trifolii. Biochem. J.* **344**, 159-166 (1999).
29. Stuible, H. P., Buttner, D., Ehlting, J., Hahlbrock, K., Kombrink, E. 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 (2000).
30. Hamberger, B., Hahlbrock, K. The 4-coumarate : CoA ligase gene family in *Arabidopsis thaliana* comprises one rare, sinapate-activating and three commonly occurring isoenzymes. *Proc. Nat. Acad. Sci.* **101**, 2209-2214 (2004).
31. Stuible, H. P., Kombrink, E. Identification of the substrate specificity-conferring amino acid residues of 4-coumarate : coenzyme A ligase allows the rational design of mutant enzymes with new catalytic properties. *J. Biol. Chem.* **276**, 26893-26897 (2001).
32. Harborne, J. B., Williams, C. A. Anthocyanins and other flavonoids. *Nat. Prod. Rep.* **18**, 310-333 (2001).
33. Winkel-Shirley, B. Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. *Plant Phys.* **126**, 485-493 (2001).
34. Bloor, S. J. Deep blue anthocyanins from blue Dianella berries. *Phytochem.* **58**, 923-927 (2001).

35. Bloor, S. J., Abrahams, S. The structure of the major anthocyanin in *Arabidopsis thaliana*. *Phytochem.* **59**, 343-346 (2002).
36. Abe, F., Iwase, Y., Yamauchi, T., Yahara, S., Nohara, T. Flavonol sinapoyl glycosides from leaves of *Thevetia peruviana*. *Phytochem.* **40**, 577-581 (1995).
37. Bloor, S. J., Bradley, J. M., Lewis, D. H., Davies, K. M. Identification of flavonol and anthocyanin metabolites in leaves of *Petunia mitchell* and its LC transgenic. *Phytochem.* **49**, 1427-1430 (1998).
38. Nielsen, J. K., Olsen, C. E., Petersen, M. K. Acylated flavonol glycosides from cabbage leaves. *Phytochem.* **34**, 539-544 (1993).
39. Nielsen, J. K., Norbaek, R., Olsen, C. E. Kaempferol tetraglucosides from cabbage leaves. *Phytochem.* **49**, 2171-2176 (1998).
40. Schnitzler, J. P. Tissue localization of UV-B screening pigments and of chalcone synthase mRNA in needles of Scots pine seedlings. *New Phytol.* **132**, 247-258 (1996).
41. Goto, T., Kondo, T. Structure and molecular stacking of anthocyanins - flower color variation. *Ang. Chem.* **30**, 17-33 (1991).
42. Jenkins, J. N., Hedin, P. A., Parrott, W. L., McCarty, J. C., White, W. H. Cotton allelochemics and growth of tobacco budworm larvae. *Crop Science* **23**, 1195-1198 (1983).
43. Griesbach, R. J., Asen, S., Leonnarat, B. A. *Petunia hybrida* anthocyanins acylated with caffeic acid. *Phytochem.* **30**, 1729-1731 (1991).
44. D'Auria, J. C., Reichelt, M., Luck, K., Svatos, A. Gershenzon, J. Identification and characterization of the BAHD acyltransferase malonyl CoA: anthocyanidin 5-*O*-glucoside-6 "-*O*-malonyltransferase (At5MAT) in *Arabidopsis thaliana*. *Febs Letters* **581**, 872-878 (2007).
45. Friedman, M. Overview of antibacterial, antitoxin, antiviral, and antifungal activities of tea flavonoids and teas. *Mol. Nutr. Food Res.* **51**, 116-134 (2007).
46. Repka, V. Early defence responses induced by two distinct elicitors derived from a *Botrytis cinerea* in grapevine leaves and cell suspensions. *Biol. Plantar.* **50**, 94-106 (2006).
47. Glassgen, W. E. Regulation of enzymes involved in anthocyanin biosynthesis in carrot cell cultures in response to treatment with ultraviolet light and fungal elicitors. *Planta* **204**, 490-498 (1998).
48. Hirner, A. A., Veit, S., Seitz, H. U. Regulation of anthocyanin biosynthesis in UV-A-irradiated cell cultures of carrot and in organs of intact carrot plants. *Plant Sci.* **161**, 315-322 (2001).
49. Li, J. Y., Oulee, T. M., Raba, R., Amundson, R. G., Last, R. L. Arabidopsis flavonoid mutants are hypersensitive to UV-B irradiation. *Plant Cell* **5**, 171-179 (1993).
50. Harborne, J. B., Corner, J. J. Plant polyphenols 4- hydroxycinnamic acid sugar derivatives. *Biochem. J.* **81**, 242-246 (1961).
51. Dougall, D. K., Baker, D. C., Gakh, E. G., Redus, M. A., Whittemore, N. A. Studies on the stability and conformation of monoacylated anthocyanins part 2 - Anthocyanins from wild carrot suspension cultures acylated with supplied carboxylic acids. *Carbohydrate Res.* **310**, 177-189 (1998).
52. Hashimoto, M., Komatsu, S. Proteomic analysis of rice seedlings during cold stress. *Proteomics* **7**, 1293-1302 (2007).
53. Suzuki, H., Nakayama, T., Nishino, T. Proposed mechanism and functional amino acid residues of malonyl-CoA : anthocyanin 5-*O*-glucoside-6 "'-*O*-

malonyltransferase from flowers of *Salvia splendens*, a member of the versatile plant acyltransferase family. *Biochem.* **42**, 1764-1771 (2003).
54. Fujiwara, H. Anthocyanin 5-aromatic acyltransferase from *Gentiana triflora* - Purification, characterization and its role in anthocyanin biosynthesis. *Eur. J. Biochem.* **249**, 45-51 (1997).
55. Saito, N., Tatsuzawa, F., Kasahara, K., Iida, S., Honda, T. Acylated cyanidin 3-sophorosides in the brownish-red flowers of *Ipomoea purpurea*. *Phytochem.* **49**, 875-880 (1998).
56. Fujiwara, H. Purification and characterization of anthocyanin 3-aromatic acyltransferase from *Perilla frutescens*. *Plant Sci.* **137**, 87-94 (1998).
57. Gerasimenko, I. Purification and partial amino acid sequences of the enzyme vinorine synthase involved in a crucial step of ajmaline biosynthesis. *Bioorg. Med. Chem.* **12**, 2781-2786 (2004).
58. St-Pierre, B., Laflamme, P., Alarco, A. M., De Luca, V. The terminal *O*-acetyltransferase involved in vindoline biosynthesis defines a new class of proteins responsible for coenzyme A-dependent acyl transfer. *Plant J.* **14**, 703-713 (1998).
59. Grothe, T., Lenz, R., Kutchan, T. M. Molecular characterization of the salutaridinol 7-*O*-acetyltransferase involved in morphine biosynthesis in opium poppy *Papaver somniferum*. *J. Biol. Chem.* **276**, 30717-30723 (2001).
60. Bayer, A., Ma, X. Y., Stockigt, J. Acetyltransfer in natural product biosynthesis - functional cloning and molecular analysis of vinorine synthase. *Bioorg. Med. Chem.* **12**, 2787-2795 (2004).
61. Ma, X. Y. Crystallization and preliminary X-ray analysis of native and selenomethionyl vinorine synthase from *Rauvolfia serpentina*. *Acta Cryst. Biol. Cryst.* **61**, 694-696 (2005).
62. Ma, X. Y., Koepke, J., Panjikar, S., Fritzsch, G., Stockigt, J. Crystal structure of vinorine synthase, the first representative of the BAHD superfamily. *J. Biol. Chem.* **280**, 13576-13583 (2005).
63. Rao, K. V. Taxol and related taxanes: Taxanes of *Taxus brevifolia* bark. *Pharm. Res.* **10**, 521-524 (1993).
64. Walker, K., Croteau, R. Taxol biosynthesis: Molecular cloning of a benzoyl-CoA : taxane 2 α-*O*-benzoyltransferase cDNA from Taxus and functional expression in *Escherichia coli*. *Proc. Nat. Acad. Sci.* **97**, 13591-13596 (2000).
65. Ojima, I., Inoue, T., Chakravarty, S. Enantiopure fluorine-containing taxoids: potent anticancer agents and versatile probes for biomedical problems. *J. Fluor. Chem.* **97**, 3-10 (1999).
66. Jennewein, S.. Croteau, R. Taxol: biosynthesis, molecular genetics, and biotechnological applications. *Appl. Microb. Biotech.* **57**, 13-19 (2001).
67. Pezzuto, J. Taxol production in plant cell culture comes of age. *Nature Biotech.* **14**, 1083-1083 (1996).
68. Luo, J., He, G. Y. Optimization of elicitors and precursors for paclitaxel production in cell suspension culture of *Taxus chinensis* in the presence of nutrient feeding. *Process Biochem.* **39**, 1073-1079 (2004).
69. Walker, K., Fujisaki, S., Long, R., Croteau, R. Molecular cloning and heterologous expression of the C-13 phenylpropanoid side chain-CoA acyltransferase that functions in Taxol biosynthesis. *Proc. Nat. Acad. Sci.* **99**, 12715-12720 (2002).

70. Walker, K., Long, R., Croteau, R. The final acylation step in Taxol biosynthesis: Cloning of the taxoid C-13 side-chain *N*-benzoyltransferase from *Taxus*. *Proc. Nat. Acad. Sci.* **99**, 9166-9171 (2002).
71. Chau, M. D., Walker, K., Long, R., Croteau, R. Regioselectivity of taxoid-*O*-acetyltransferases: heterologous expression and characterization of a new taxadien-5-ol-*O*-acetyltransferase. *Arc. Biochem. Biophys.* **430**, 237-246 (2004).
72. Walker, K., Croteau, R. Molecular cloning of a 10-deacetylbaccatin III-10-*O*-acetyl transferase cDNA from Taxus and functional expression in *Escherichia coli*. *Proc. Nat. Acad. Sci.* **97**, 583-587 (2000).
73. Fleming, P. E., Mocek, U., Floss, H. G. Biosynthesis of taxoids - Mode of formation of the taxol side-chain. *J. Am. Chem. Soc.* **115**, 805-807 (1993).
74. Uchida, K., Yokoshima, S., Kan, T., Fukuyama, T. Total synthesis of morphine. *Org. Lett.* **8**, 5311-5313 (2006).
75. Lenz, R., Zenk, M. H. Acetyl coenzyme A:salutaridinol-7-*O*-acetyltransferase from *Papaver somniferum* plant cell cultures - The enzyme catalyzing the formation of thebaine in morphine biosynthesis. *J. Biol. Chem.* **270**, 31091-31096 (1995).
76. Choi, Y. G. Total synthesis of vindoline. *Org. Lett.* **7**, 4539-4542 (2005).
77. Zarate, R., Dirks, C., van der Heijden, R., Verpoorte, R. Terpenoid indole alkaloid profile changes in *Catharanthus pusillus* during development. *Plant Sci.* **160**, 971-977 (2001).
78. Wang, K. W., Sun, C. R., Wu, X. D., Pan, Y. J. Novel bioactive dammarane caffeoyl esters from *Celastrus rosthornianus*. *Planta Med.* **72**, 370-372 (2006).
79. Powers, J. F., Brachold, J. M., Ehsani, S. A., Tischler, A. S. Up-regulation of ret by reserpine in the adult rat adrenal medulla. *Neuroscience* **132**, 605-612 (2005).
80. Yonekura-Sakakibara, K. Molecular and biochemical characterization of a novel hydroxycinnamoyl-CoA: Anthocyanin 3-*O*-glucoside-6 "-*O*-acyltransferase from *Perilla frutescens*. *Plant Cell Phys.* **41**, 495-502 (2000).
81. Fujiwara, H. cDNA cloning, gene expression and subcellular localization of anthocyanin 5-aromatic acyltransferase from *Gentiana triflora*. *Plant J.* **16**, 421-431 (1998).
82. Suzuki, H. Malonyl-CoA : anthocyanin 5-*O*-glucoside-6 "'-*O*-malonyltransferase from scarlet sage (*Salvia splendens*) flowers - Enzyme purification, gene cloning, expression, and characterization. *J. Biol Chem.* **276**, 49013-49019 (2001).
83. Suzuki, H., Nakayama, T., Yamaguchi, M. A., Nishino, T. cDNA cloning and characterization of two *Dendranthema x morifolium* anthocyanin malonyltransferases with different functional activities. *Plant Sci.* **166**, 89-96 (2004).
84. Loncaric, C., Merriweather, E., Walker, K. D. Profiling a taxol pathway 10 β-acetyltransferase: Assessment of the specificity and the production of baccatin III by *in vivo* acetylation in *E-coli*. *Chem. Biol.* **13**, 309-317 (2006).
85. Ojima, I. Use of fluorine in the medicinal chemistry and chemical biology of bioactive compounds - A case study on fluorinated taxane anticancer agents. *Chembiochem* **5**, 628-635 (2004).

86. Ruppert, M., Ma, X. Y., Stockigt, J. Alkaloid biosynthesis in Rauvolfia - cDNA cloning of major enzymes of the ajmaline pathway. *Curr. Org. Chem.* **9**, 1431-1444 (2005).
87. Wipf, P. Natural product based inhibitors of the thioredoxin-thioredoxin reductase system. *Org. Biomol. Chem.* **2**, 1651-1658 (2004).
88. Wessjohann, L. A. Synthesis of natural-product-based compound libraries. *Curr. Op. Chem. Biol.* **4**, 303-309 (2000).
89. Altmann, K. H., Florsheimer, A., Bold, G., Caravatti, G., Wartmann, M. Natural product-based drug discovery - Epothilones as lead structures for the discovery of new anticancer agents. *Chimia* **58**, 686-690 (2004).
90. Schnarr, N. A., Khosla, C. Biological chemistry: Just add chlorine. *Nature* **436**, 1094-1095 (2005).
91. Yeh, E., Garneau, S., Walsh, C. T. Robust in vitro activity of RebF and RebH, a two-component reductase/halogenase, generating 7-chlorotryptophan during rebeccamycin biosynthesis. *Proc. Nat. Acad. Sci.* **102**, 3960-3965 (2005).
92. Ismail, F. M. D. Important fluorinated drugs in experimental and clinical use. *J. Fluor. Chem.* **118**, 27-33 (2002).
93. O'Hagan, D., Harper, D. B. Natural products containing fluorine and recent progress in elucidating the pathway of fluorometabolite biosynthesis in *Streptomyces cattleya*. *Asym. Fluoro. Chem.* **746**, 210-224 (2000).
94. Smart, B. E. Fluorine substituent effects (on bioactivity). *J. Fluor. Chem.* **109**, 3-11 (2001).
95. Moroni, M. First synthesis of totally orthogonal protected α-(trifluoromethyl)- and α-(difluoromethyl)arginines. *J. Org. Chem.* **66**, 130-133 (2001).
96. Wetterau, J. R. An MTP inhibitor that normalizes atherogenic lipoprotein levels in WHHL rabbits. *Science* **282**, 751-754 (1998).
97. Kosoglou, T. Pharmacodynamic interaction between cerivastatin and the selective cholesterol absorption inhibitor ezetimibe. *Eur. Heart J.* **22**, 252-252 (2001).
98. Young, H. Measurement of clinical and subclinical tumour response using [F-18]-fluorodeoxyglucose and positron emission tomography. *Eur. J. Cancer* **35**, 1773-1782 (1999).
99. Katzenellenbogen, J. A. Steroids labeled with F-18 for imaging tumors by positron emission tomography. *J. Fluor. Chem.* **109**, 49-54 (2001).
100. Middleton, E., Kandaswami, C. The effect of plant flavonoids on immune and inflammatory cell function. *Abs. Am. Chem. Soc.* **212**, 125-127 (1996).
101. Rice-Evans, C. Flavonoid antioxidants. *Curr. Med. Chem.* **8**, 797-807 (2001).
102. Tian, Q. G., Aziz, R. M., Stoner, G. D., Schwartz, S. J. Anthocyanin determination in black raspberry (*Rubus occidentalis*) and biological specimens using liquid chromatography-electrospray ionization tandem mass spectrometry. *J. Food Sci.* **70**, C43-C47 (2005).
103. Lu, C. M., Yang, J. J., Wang, P. Y., Lin, C. C. A new acylated flavonol glycoside and antioxidant effects of *Hedyotis diffusa*. *Planta Med.* **66**, 374-376 (2000).
104. Mellou, F. Biocatalytic preparation of acylated derivatives of flavonoid glycosides enhances their antioxidant and antimicrobial activity. *J. Biotech.* **116**, 295-304 (2005).
105. van der Meer, I. M., Bovy, A. G., Bosch, D. Plant-based raw material: improved food quality for better nutrition via plant genomics. *Curr. Op. Biotech.* **12**, 488-492 (2001).

106. Yu, O. Production of the isoflavones genistein and daidzein in non-legume dicot and monocot tissues. *Plant Phys.* **124**, 781-793 (2000).
107. Muir, S. R. Overexpression of petunia chalcone isomerase in tomato results in fruit containing increased levels of flavonols. *Nature Biotech.* **19**, 470-474 (2001).
108. Oufedjikh, H., Mahrouz, M., Lacroix, M., Amiot, M. J., Taccini, M. The influence of gamma irradiation on flavonoid content during storage of irradiated clementina. *Rad. Phys. Chem.* **52**, 107-112 (1998).
109. Tian, Q. G., Giusti, M. M., Stoner, G. D., Schwartz, S. J. Characterization of a new anthocyanin in black raspberries (*Rubus occidentalis*) by liquid chromatography electrospray ionization tandem mass spectrometry. *Food Chem.* **94**, 465-468 (2006).
110. Tian, Q. G., Konczak, I., Schwartz, S. J. Probing anthocyanin profiles in purple sweet potato cell line (*Ipomoea batatas*) by high-performance liquid chromatography and electrospray ionization tandem mass spectrometry. *J. Ag. Food Chem.* **53**, 6503-6509 (2005).
111. Yoshimoto, M., Okuno, S., Yamaguchi, M., Yamakawa, O. Antimutagenicity of deacylated anthocyanins in purple-fleshed sweetpotato. *Biosci. Biotech. Biochem.* **65**, 1652-1655 (2001).
112. Harada, K., Kano, M., Takayanagi, T., Yamakawa, O., Ishikawa, F. Absorption of acylated anthocyanins in rats and humans after ingesting an extract of *Ipomoea batatas* purple sweet potato tuber. *Biosci. Biotech. Biochem.* **68**, 1500-1507 (2004).
113. Nakajima, N., Ishihara, K., Hamada, H., Kawabe, S., Furuya, T. Regioselective acylation of flavonoid glucoside with aromatic acid by an enzymatic reaction system from cultured cells of *Ipomoea batatas*. *J. Biosci. Bioeng.* **90**, 347-349 (2000).
114. Adam, G. C., Burbaum, J., Kozarich, J. W., Patricelli, M. P., Cravatt, B. F. Mapping enzyme active sites in complex proteomes. *J. Am. Chem. Soc.* **126**, 1363-1368 (2004).
115. Adam, G. C., Sorensen, E. J., Cravatt, B. F. Trifunctional chemical probes for the consolidated detection and identification of enzyme activities from complex proteomes. *Mol. Cell. Proteom.* **1**, 828-835 (2002).
116. Kidd, D., Liu, Y. S., Cravatt, B. F. Profiling serine hydrolase activities in complex proteomes. *Biochemistry* **40**, 4005-4015 (2001).
117. Battaglia, E. Photoaffinity labeling studies of the human recombinant UDP-glucuronosyltransferase, UGT16, with 5-azido-UDP-glucuronic acid. *Drug Metab. Disp.* **25**, 406-411 (1997).
118. Chivasa, S. Proteomic analysis of the *Arabidopsis thaliana* cell wall. *Electrophoresis* **23**, 1754-1765 (2002).
119. Devreese, B. Van Beeumen, J. Nanoflow LC Q-TOF MS for *De novo* peptide sequencing in microbial proteomics. *Lc Gc Europe* **15**, 658-662 (2002).
120. Edwards, R., Kessmann, H. Isoflavonoid phytoalexins and their biosynthetic enzymes. *Mol. Plant Path.: A practical approach* **2**, 50 - 61.
121. Loutre, C. Isolation of a glucosyltransferase from *Arabidopsis thaliana* active in the metabolism of the persistent pollutant 3,4-dichloroaniline. *Plant J.* **34**, 485-493 (2003).
122. Thompson, J. D., Higgins, D. G., Gibson, T. J. Clustal improving the sensitivity of progressive multiple sequence alignment through sequence

weighting, position specific gap penalties and weight matrix choice. *Nucleic Acids Research* **22**, 4673-4680 (1994).

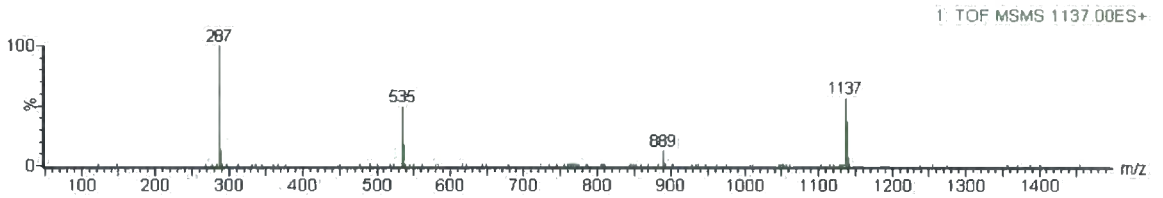
123. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., Lipman, D. J. Basic Local Alignment Search Tool. *J. Mol. Biol.* **215**, 403-410 (1990).
124. Beuerle, T., Pichersky, E. Enzymatic synthesis and purification of aromatic coenzyme A esters. *Anal. Biochem.* **302**, 305-312 (2002).
125. Stockigt, J., Zenk, M. H. Chemical syntheses and properties of hydroxycinnamoyl coenzyme A derivatives. *Zeit. Natur. J. Biosci.* **30**, 352-358 (1975).
126. Kaffarnik, F., Heller, W., Hertkorn, N., Sandermann, H. Flavonol 3-*O*-glycoside hydroxycinnamoyltransferases from Scots pine (*Pinus sylvestris* L.). *Febs J.* **272**, 1415-1424 (2005).
127. Raacke, I. C., von Rad, U., Mueller, M. J., Berger, S. Yeast increases resistance in Arabidopsis against *Pseudomonas syringae* and *Botrytis cinerea* by salicylic acid-dependent as well as independent mechanisms. *Mol. Plant Micr. Int.* **19**, 1138-1146 (2006).
128. Suzuki, H. cDNA cloning and functional characterization of flavonol 3-*O*-glucoside-6 "-*O*-malonyltransferases from flowers of *Verbena hybrida* and *Lamium purpureum*. *J. Mol. Cat. B-Enz.* **28**, 87-93 (2004).
129. Gershater, M. C., Cummins, I., Edwards, R. Role of a carboxylesterase in herbicide bioactivation in *Arabidopsis thaliana*. *J. Biol. Chem.* **282**, 21460-21466 (2007).
130. Gershater, M. C., Edwards, R. Regulating biological activity in plants with carboxylesterases. *Plant Sci.* **173**, 579-588 (2007).
131. Cuyckens, F. The application of liquid chromatography-electrospray ionization mass spectrometry and collision-induced dissociation in the structural characterization of acylated flavonol *O*-glycosides from the seeds of *Carrichtera annua*. *Eur. J. Mass Spec.* **9**, 409-420 (2003).
132. Merali, Z. Metabolic diversion of the phenylpropanoid pathway causes cell wall and morphological changes in transgenic tobacco stems. *Planta* **225**, 1165-1178 (2007).
133. Sawada, S. UDP-glucuronic acid : anthocyanin glucuronosyltransferase from red daisy (*Bellis perennis*) flowers - Enzymology and phylogenetics of a novel glucuronosyltransferase involved in flower pigment biosynthesis. *J. Biol. Chem.* **280**, 899-906 (2005).
134. Panza, J. L., Russell, A. J., Beckman, E. J. Synthesis of fluorinated NAD as a soluble coenzyme for enzymatic chemistry in fluorous solvents and carbon dioxide. *Tetrahedron* **58**, 4091-4104 (2002).
135. Suzuki, H. Identification and characterization of a novel anthocyanin malonyltransferase from scarlet sage (*Salvia splendens*) flowers: an enzyme that is phylogenetically separated from other anthocyanin acyltransferases. *Plant J.* **38**, 994-1003 (2004).
136. Schein, C. H. Production Of Soluble Recombinant Proteins In Bacteria. *Biotechnology* **7**, 1141-1147 (1989).
137. Simmons, T. L., Andrianasolo, E., McPhail, K., Flatt, P., Gerwick, W. H. Marine natural products as anticancer drugs. *Mol. Cancer Ther.* **4**, 333-342 (2005).
138. Roobol, A., Carden, M. J. ATP hydrolysis causes extensive disassembly of eukaryotic CCT chaperonin at intracellular potassium levels. *Mol. Biol. Cell* **9**, 453A-453A (1998).

139. Bennett, J., Scott, K. J. Quantitative staining of fraction 1 protein in polyacrylamide gels using coomassie brilliant blue. *Anal. Biochem.* **43**, 173-& (1971).
140. Nandi, D. L., Lucas, S. V., Webster, L. T. Benzoyl-coenzyme-A glycine *N*-acyltransferase and phenylacetyl-coenzyme-A glycine *N*-acyltransferase from bovine liver mitochondria - Purification and characterization. *J. Biol. Chem.* **254**, 7230-7237 (1979).
141. Ghangas, G. S. A survey of the nature of glucose acylation reactions in plant extracts. *Phytochem.* **52**, 785-792 (1999).
142. Arndt, J. W. Crystal structure of an α/β serine hydrolase from *Saccharomyces cerevisiae* at 1.85 angstrom resolution. *Proteins - Struct. Funct. Bioinf.* **58**, 755-758 (2005).
143. Gershater, M., Cummins, I., Edwards, R. Role of a carboxylesterase in herbicide bioactivation in *Arabidopsis thaliana*. *J. Biol. Chem.* **282**, 21460-21466 (2007).
144. Liu, Y. S., Patricelli, M. P., Cravatt, B. F. Activity-based protein profiling: The serine hydrolases. *Proc. Nat. Acad. Sci.* **96**, 14694-14699 (1999).
145. Doerr, A. Protein biochemistry - Streptavidin surprises. *Nature Methods* **2**, 573-573 (2005).
146. Wu, S. C., Wong, S. L. Engineering soluble monomeric streptavidin with reversible biotin binding capability. *J. Biol. Chem.* **280**, 23225-23231 (2005).
147. Grimsley, P. G., Amos, T. A. S., Gordon, M. Y., Greaves, M. F. Rapid positive selection of Cd cells using magnetic microspheres coated with monoclonal-antibody linked via a cleavable disulfide bond. *Leukemia* **7**, 898-908 (1993).
148. Ashani, Y., Bhattacharjee, A. K., Leader, H., Saxena, A., Doctor, B. P. Inhibition of cholinesterases with cationic phosphonyl oximes highlights distinctive properties of the charged pyridine groups of quaternary oxime reactivators. *Biochem. Pharm.* **66**, 191-202 (2003).
149. Kawasumi, M., Nghiem, P. Chemical genetics: Elucidating biological systems with small-molecule compounds. *J. Inves. Derm.* **127**, 1577-1584 (2007).
150. Depierre, A. A phase-II study of Navelbine (Vinorelbine) in the treatment of non small-cell lung-cancer. *Am. J. Clin. Onco. Cancer Clin. Trials* **14**, 115-119 (1991).
151. Tu, J., Zhu, P., Cheng, K., Meng, C. Cloning and sequencing of hydroxylase genes involved in taxol biosynthesis. *Zeit. Nat. C-J. Biosci.* **59**, 561-564 (2004).
152. Azorsa, D. O., Mousses, S., Caplen, N. J. Gene silencing through RNA interference: Potential for therapeutics and functional genomics. *Letters In Peptide Science* **10**, 361-372 (2003).
153. AzpirozLeehan, R., Feldmann, K. A. t-DNA insertion mutagenesis in Arabidopsis: Going back and forth. *Trends In Genetics* **13**, 152-156 (1997).
154. Elphick, L. M., Lee, S. E., Gouverneur, V., Mann, D. J. Using chemical genetics and ATP analogues to dissect protein kinase function. *ACS. Chem. Biol.* **2**, 299-314 (2007).
155. Hammett, L. P. The effect of structure upon the reactions of organic compounds. *J. Am. Chem. Soc.* **59**, 96-103 (1937).
156. Loudwig, S., Nicolet, Y., Masson, P., Bon, S., Nachon, F., Goeldner, M. Photoreversible inhibition of cholineesterases. *Chembiochem* **4**, 762-767 (2003).

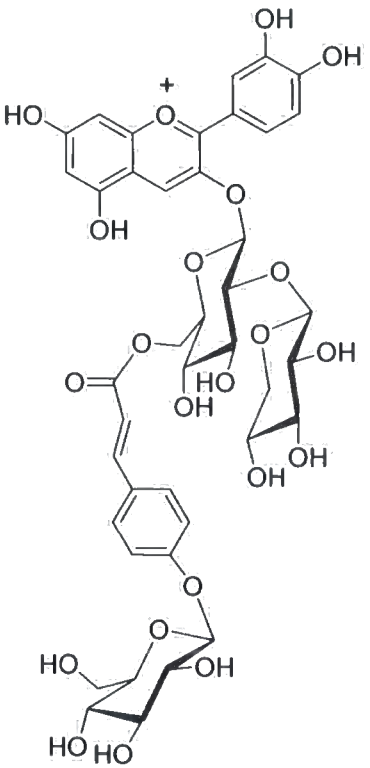
157. Hakala, M., Rantamaki, S., Puputti, E., Tyystjarvi, T., Tyystjarvi, E. Photoinhibition of manganese enzymes: insights into the mechanism of photosystem II inhibition. *J. Exp. Botany* **57**, 1809-1816 (2006).
158. Grasbolle, M., Haumann, M., Muller, C., Liebisch, P., Holger, D. Rapid loss of structural motifs in the manganese complex of oxygenic photosynthesis by x-ray irradiation at 10-300 K. *J. Biol. Chem.* **281**, 4580-4588 (2006).
159. Davies, M. J., Truscott, R. J. W. Photooxidation of proteins and its role in cataractogenesis. *J. Photochem. Photobiol. B. Biol.* **63**, 114-125 (2001).

Appendix A

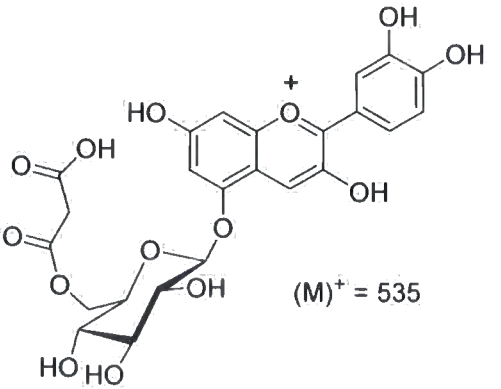
AtAN 1 – MS/MS (ES^+) Fragmentation (1137 Da)



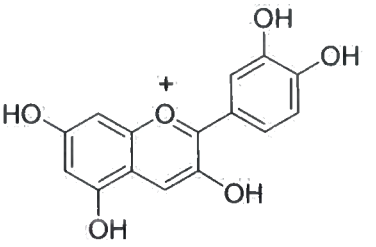
Proposed Fragments



$(M)^+ = 889$

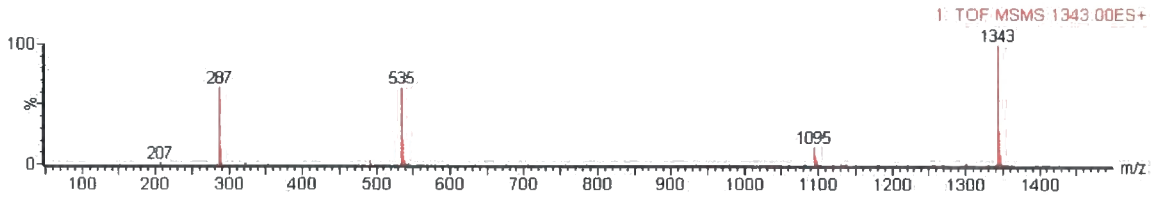


$(M)^+ = 535$

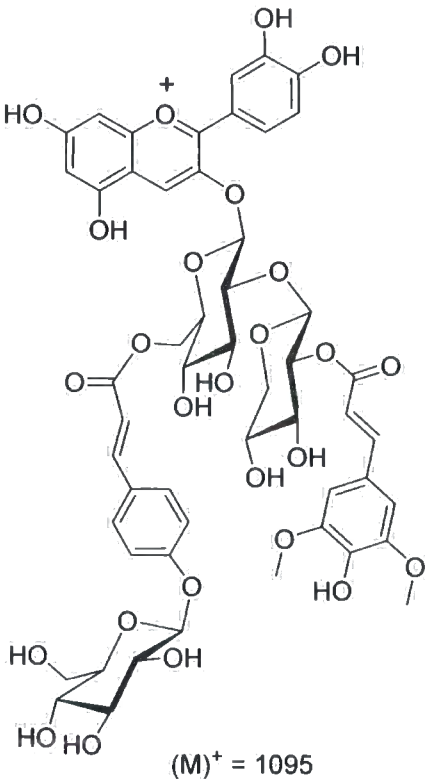


$(M)^+ = 287$

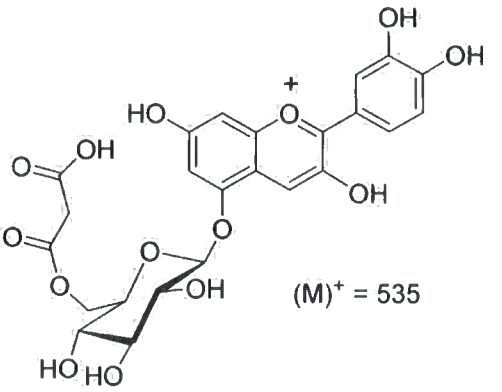
AtAN 2 – MS/MS (ES^+) Fragmentation (1343 Da)



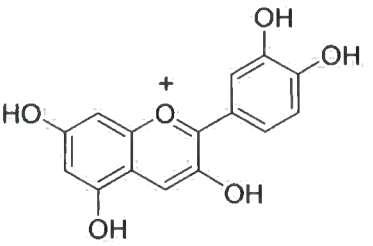
Proposed Fragments



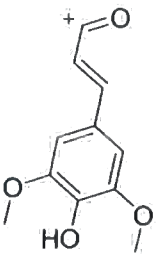
$(M)^+ = 1095$



$(M)^+ = 535$

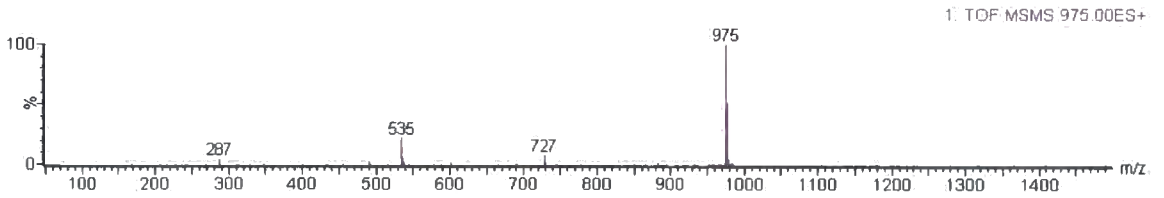


$(M)^+ = 287$

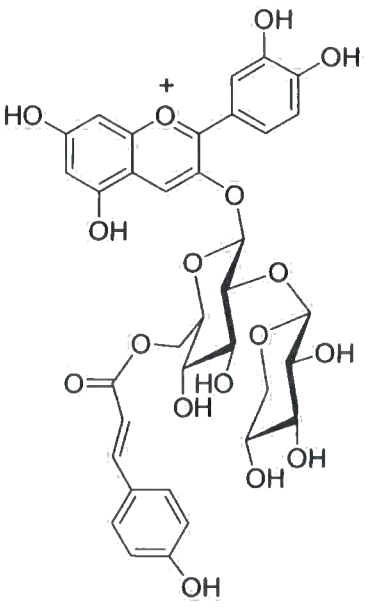


$(M)^+ = 207$

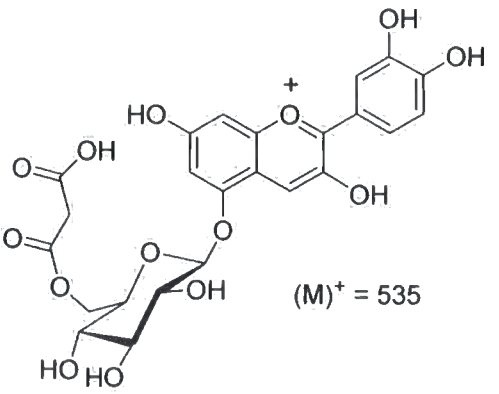
AtAN 3 – MS/MS (ES^+) Fragmentation (975 Da)



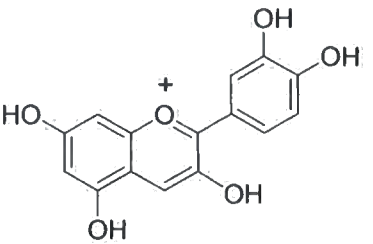
Proposed Fragments



$(M)^+ = 727$

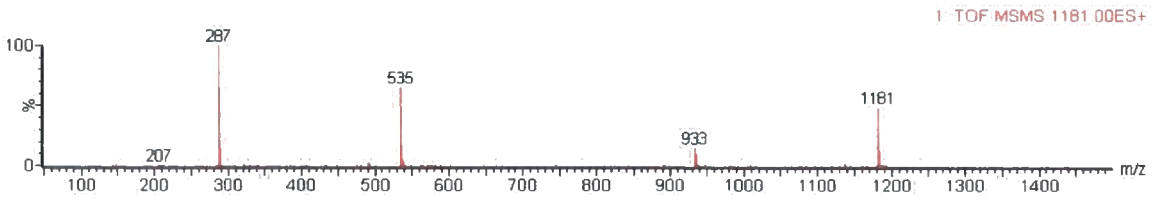


$(M)^+ = 535$

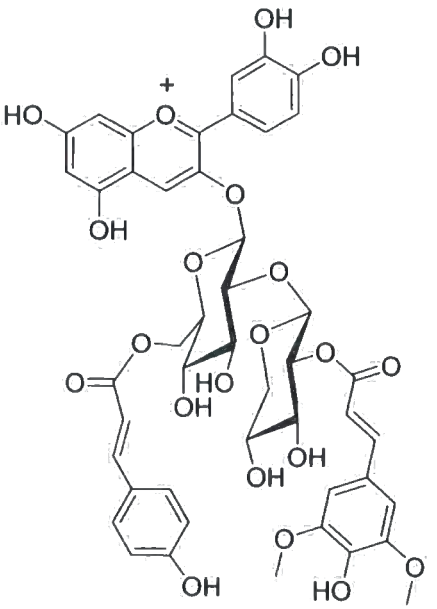


$(M)^+ = 287$

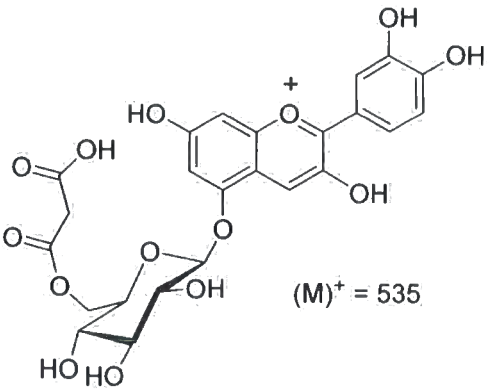
AtAN 4 – MS/MS (ES^+) Fragmentation (1181 Da)



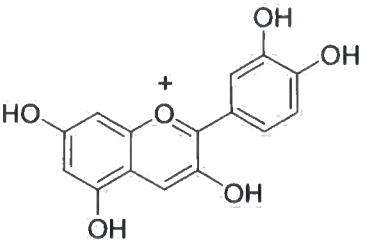
Proposed Fragments



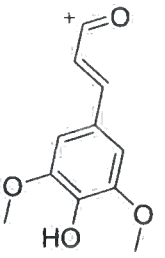
$(M)^+ = 933$



$(M)^+ = 535$

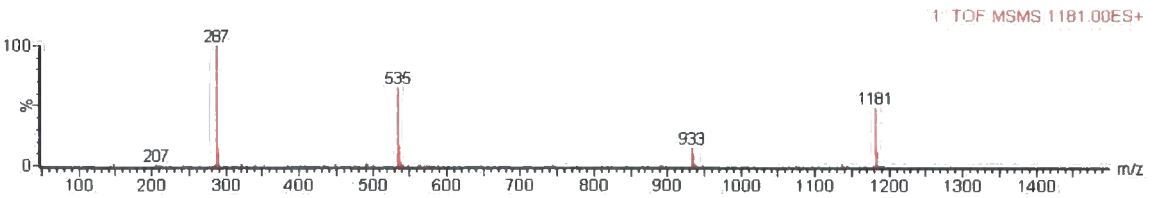


$(M)^+ = 287$

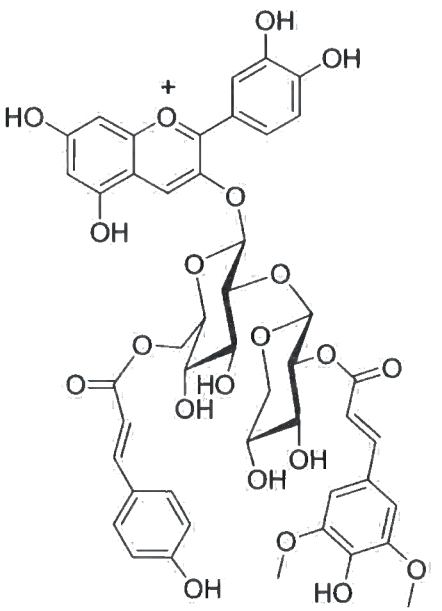


$(M)^+ = 207$

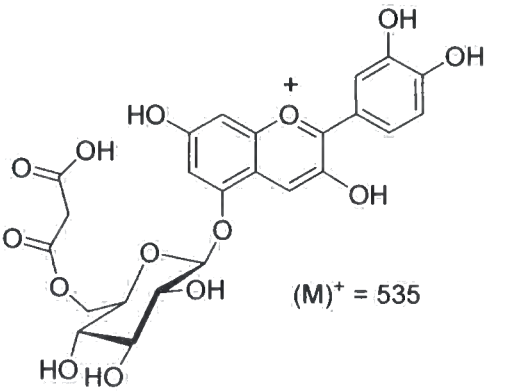
AtAN 5 – MS/MS (ES^+) Fragmentation (1181 Da)



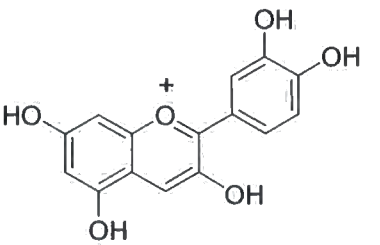
Proposed Fragments



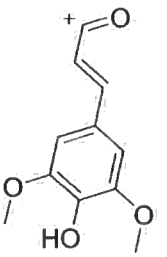
$(M)^+ = 933$



$(M)^+ = 535$

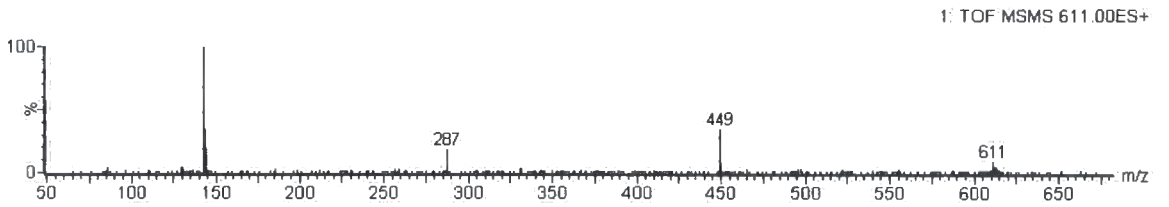


$(M)^+ = 287$

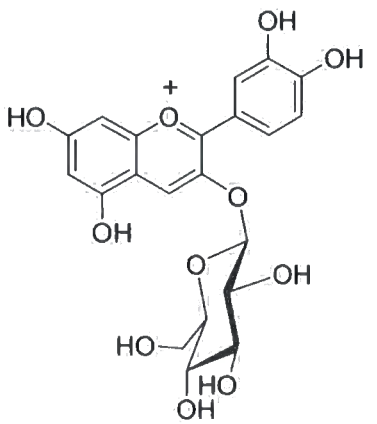


$(M)^+ = 207$

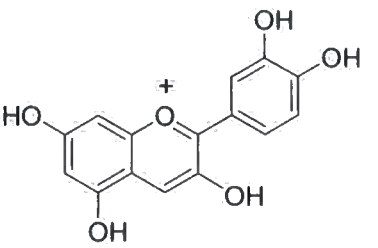
AtAN 6 – MS/MS (ES^+) Fragmentation (611 Da)



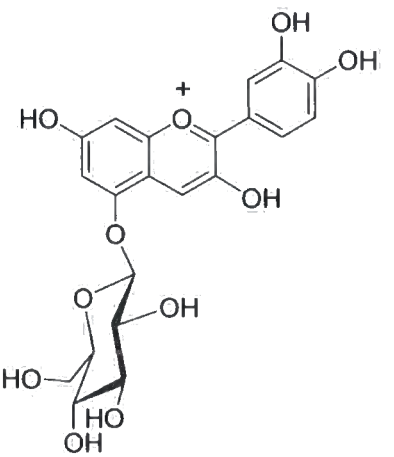
Proposed Fragments



$(M)^+ = 449$

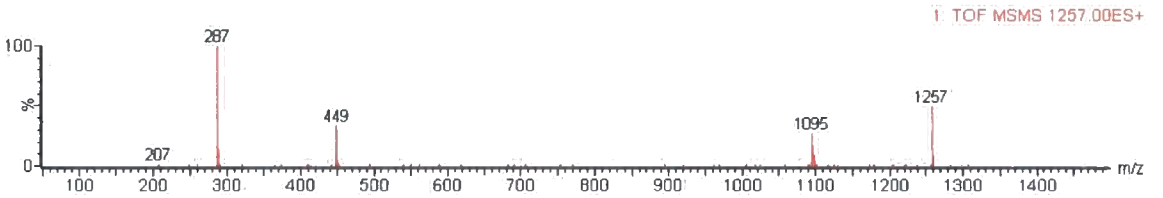


$(M)^+ = 287$

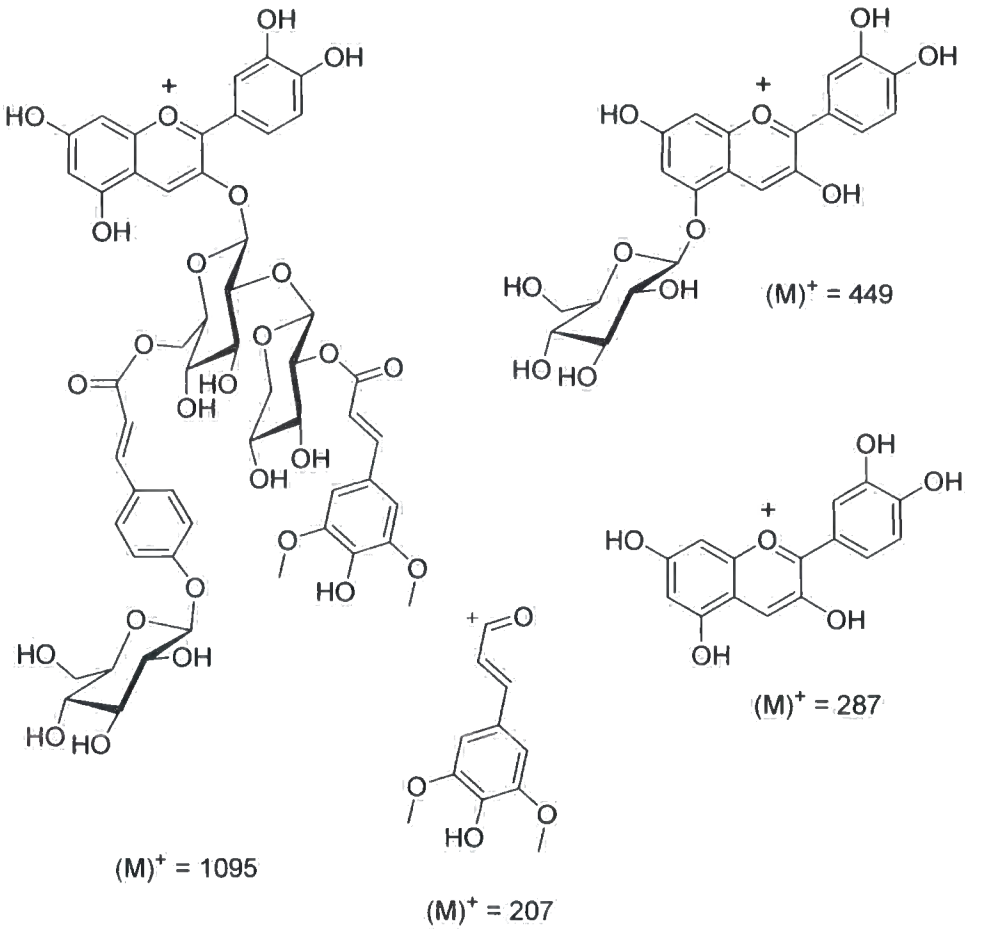


$(M)^+ = 449$

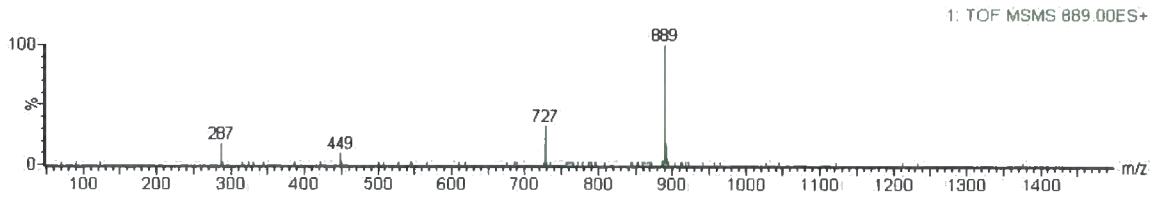
AtAN 7 – MS/MS (ES^+) Fragmentation (1257 Da)



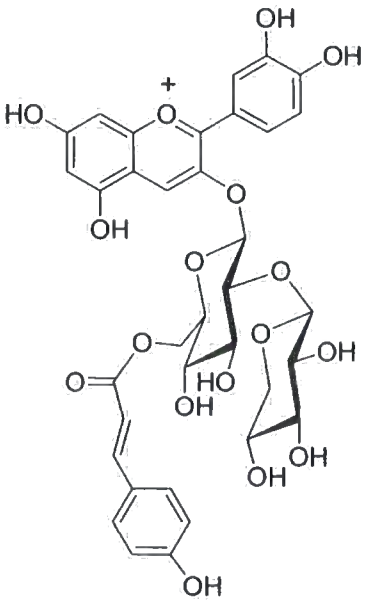
Proposed Fragments



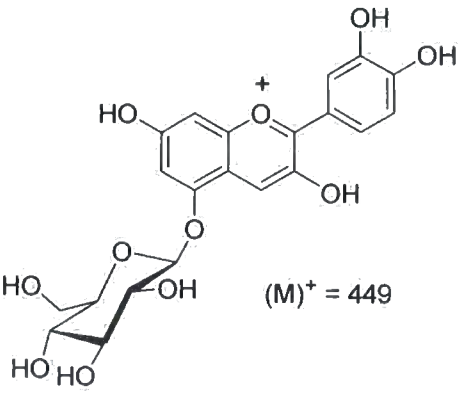
AtAN 8 – MS/MS (ES^+) Fragmentation (889 Da)



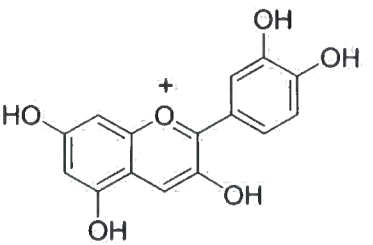
Proposed Fragments



$(M)^+ = 727$

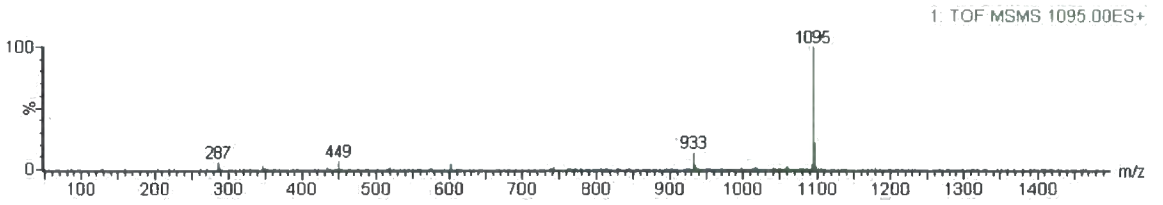


$(M)^+ = 449$

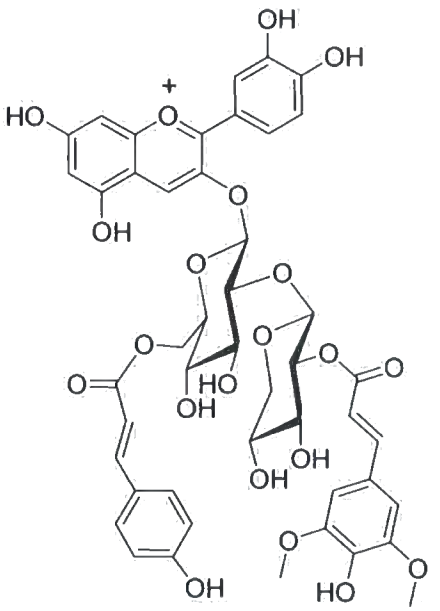


$(M)^+ = 287$

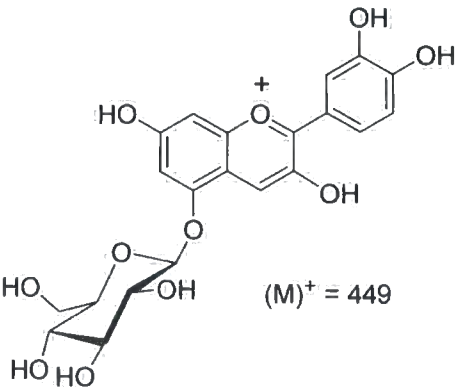
AtAN 9 – MS/MS (ES^+) Fragmentation (1095 Da)



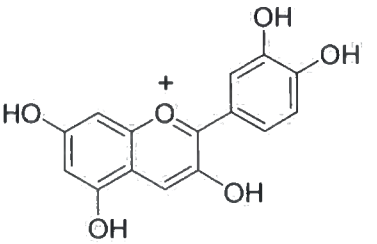
Proposed Fragments



$(M)^+ = 933$



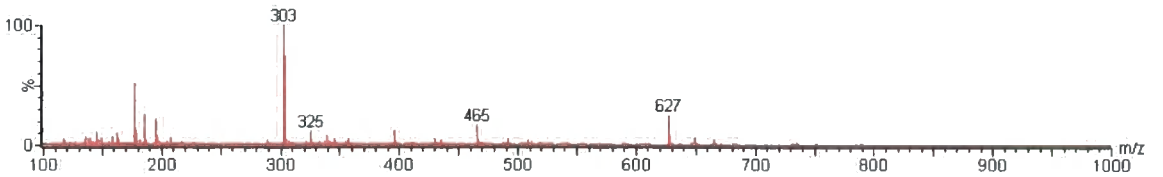
$(M)^+ = 449$



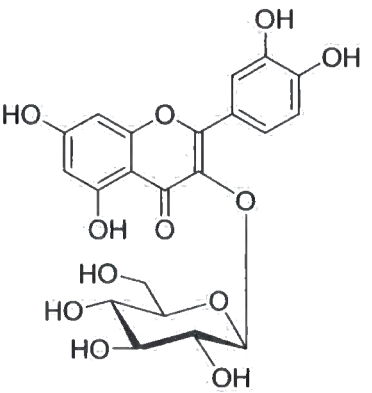
$(M)^+ = 287$

Appendix B

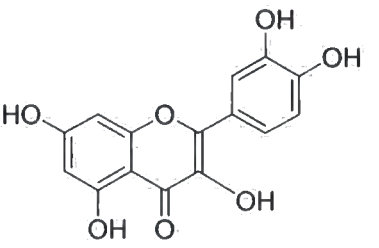
QDG (1) - MS (ES^+) Fragmentation (627 Da)



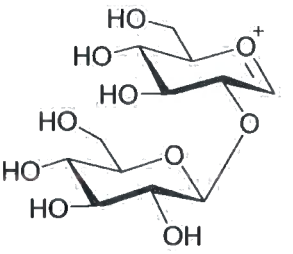
Proposed Fragments



$(M + H)^+ = 465$

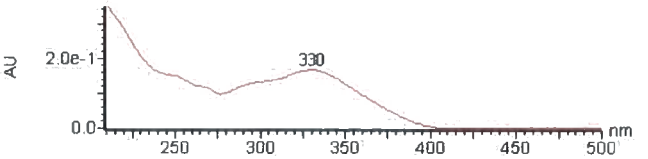


$(M + H)^+ = 303$

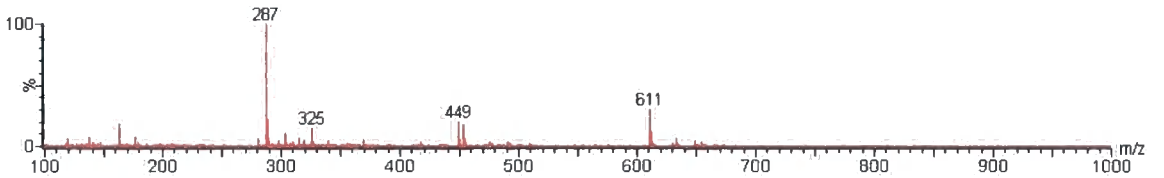


$(M)^+ = 325$

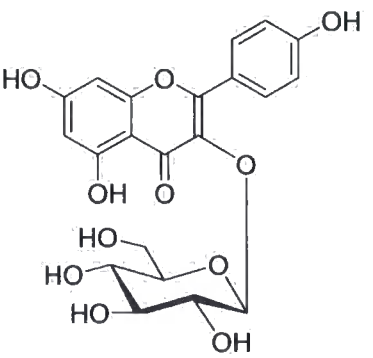
UV-Vis absorption spectrum



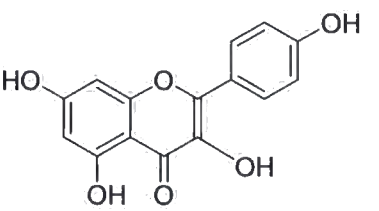
KDG (2) - MS (ES^+) Fragmentation (611 Da)



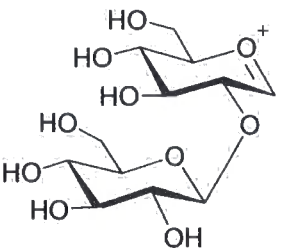
Proposed Fragments



$(M + H)^+ = 449$

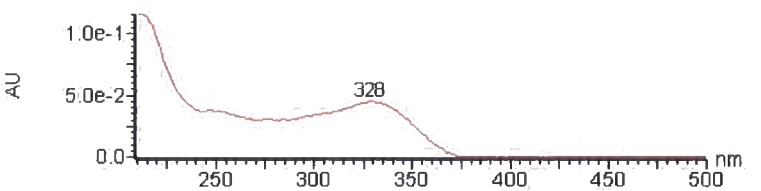


$(M + H)^+ = 287$

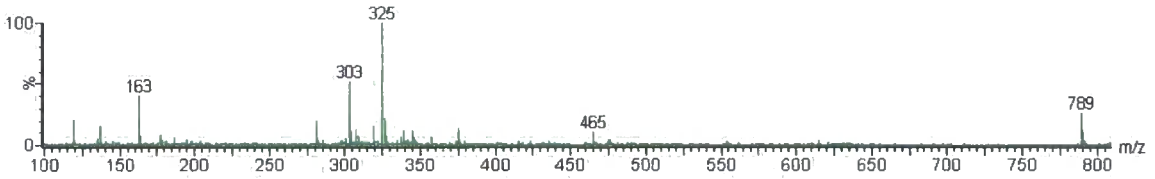


$(M)^+ = 325$

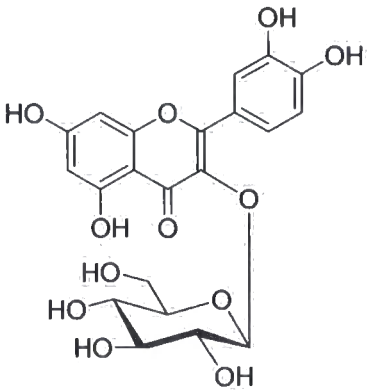
UV-Vis absorption spectrum



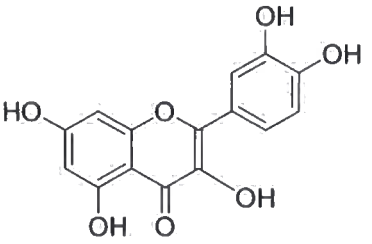
QDG-Caff (3) – MS (ES^+) Fragmentation (789 Da)



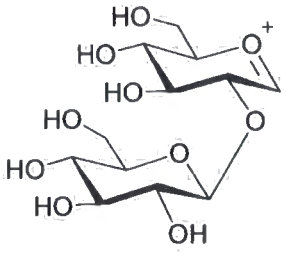
Proposed Fragments



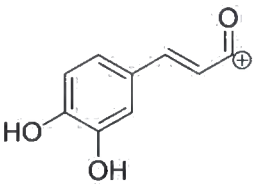
$(M + H)^+ = 465$



$(M + H)^+ = 303$

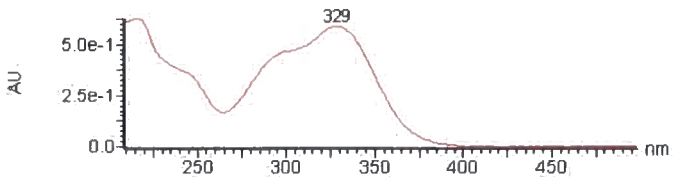


$(M)^+ = 325$

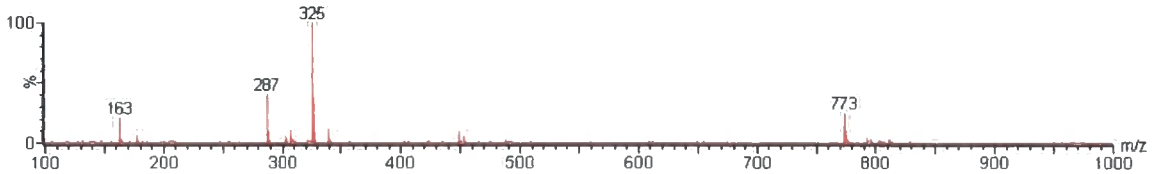


$(M)^+ = 163$

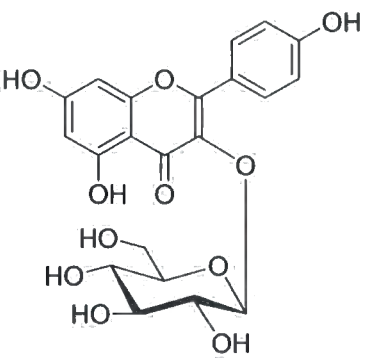
UV-Vis absorption spectrum



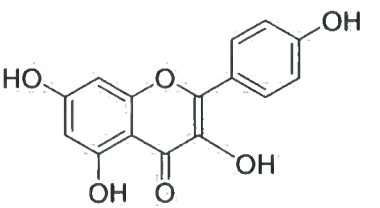
KDG-Caff (4) – MS (ES^+) Fragmentation (773 Da)



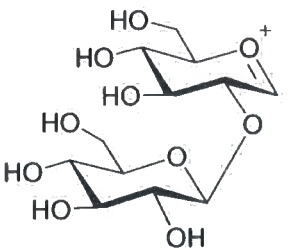
Proposed Fragments



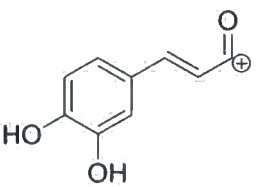
$(M + H)^+ = 449$



$(M + H)^+ = 287$

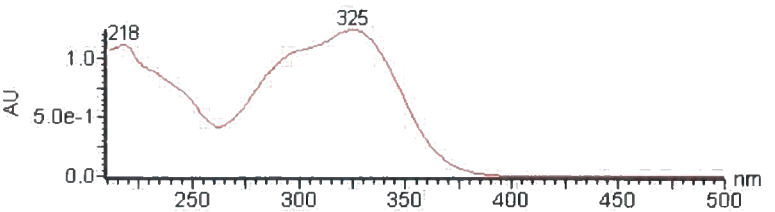


$(M)^+ = 325$

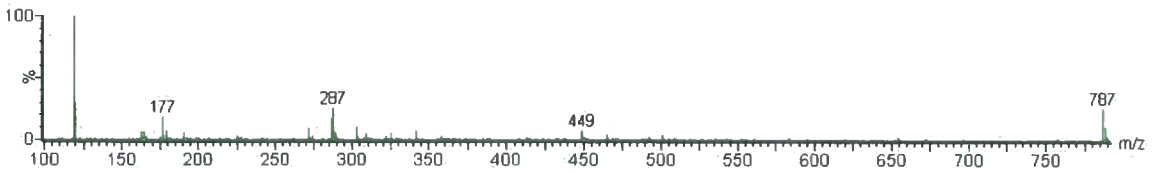


$(M)^+ = 163$

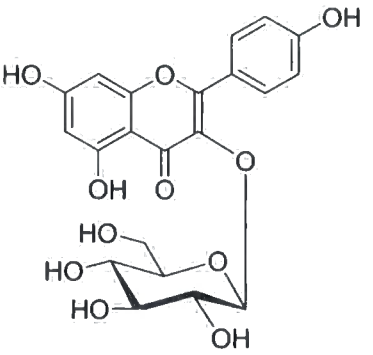
UV-Vis absorption spectrum



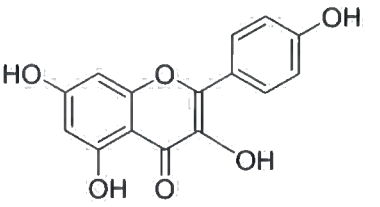
KDG-Fer (5) – MS (ES^+) Fragmentation (787 Da)



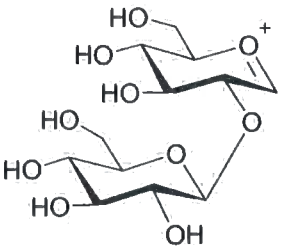
Proposed Fragments



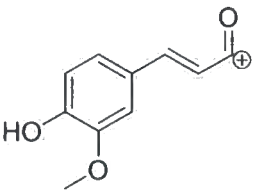
$(M + H)^+ = 449$



$(M + H)^+ = 287$

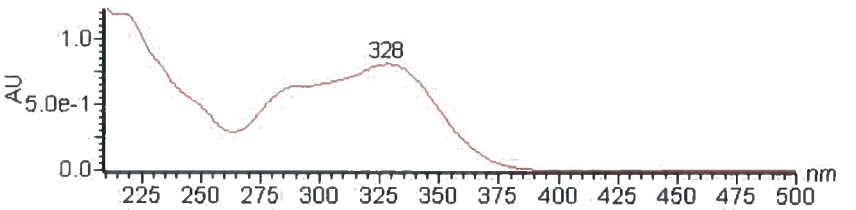


$(M)^+ = 325$



$(M)^+ = 177$

UV-Vis absorption spectrum



Appendix C

Raw data for calculation of the extinction coefficient of quercetin 3-*O*-diglucoside at 330 nm (10 µl injection volume)

Conc. (mM)	*Absorbance 1*	*Absorbance 2*
2.5	2052	2100
1.25	1016	1029
0.625	477	537
0.3125	236	273
0.15625	113	128

Coefficient 1 = 815.54 mM^{-1} cm^{-1}
Coefficient 2 = 838.08 mM^{-1} cm^{-1}

Average +/- error = 826.80 +/- 11.27 mM^{-1} cm^{-1}

Raw data for calculation of the extinction coefficient of cyanidin 3-5- *O*-diglucoside at 520 nm (10 µl injection volume)

Conc. (mM)	*Absorbance 1*	*Absorbance 2*
1	448.6	465.4
0.1	43.9	47.8
0.05	21.1	24.09
0.01	4.09	5.11

Coefficient 1 = 448.44 mM^{-1} cm^{-1}
Coefficient 2 = 465.57 mM^{-1} cm^{-1}

Average +/- error = 457.0 +/- 8.6 mM^{-1} cm^{-1}

Raw data for calculation of the extinction coefficient of cyanidin 3-*O*-glucoside 5-*O*-(caffeoyl)glucoside at 520 nm (10 µl injection volume)

Conc. (mM)	*Absorbance 1*	*Absorbance 2*
1	501	523
0.5	255	260
0.1	46.2	42.09
0.01	5.61	6.39

Coefficient 1 = 502.33 mM^{-1} cm^{-1}
Coefficient 2 = 522.94 mM^{-1} cm^{-1}

Average +/- error = 510.3 +/- 10.3 mM^{-1} cm^{-1}

Raw data for calculation of the extinction coefficient of cyanidin 3-*O*-glucoside 5-*O*-(feruloyl)glucoside at 520 nm (10 µl injection volume)

Conc. (mM)	*Absorbance 1*	*Absorbance 2*
1	488	489
0.5	246	259
0.1	43.7	52.1
0.01	5.10	5.41

Coefficient 1 = 488.87 mM^{-1} cm^{-1}
Coefficient 2 = 496.73 mM^{-1} cm^{-1}

Average +/- error = 492.8 +/- 3.93 mM^{-1} cm^{-1}

Raw data for calculation of the extinction coefficient of cyanidin 3-*O*-glucoside 5-*O*-(coumaroyl)glucoside at 520 nm (10 µl injection volume)

Conc. (mM)	*Absorbance 1*	*Absorbance 2*
1	480	512
0.5	236	249
0.1	42	53.6
0.01	3.79	4.45

Coefficient 1 = 477.93 mM^{-1} cm^{-1}
Coefficient 2 = 509.42 mM^{-1} cm^{-1}

Average +/- error = 493.7 +/- 15.7 mM^{-1} cm^{-1}

Raw data for calculation of the extinction coefficient of cyanidin 3-*O*-glucoside 5-*O*-(4-Fluorocinnamoyl)glucoside at 520 nm (10 µl injection volume)

Conc. (mM)	*Absorbance 1*	*Absorbance 2*
1	441	453
0.5	216	234
0.1	33.2	47.2
0.01	3.21	4.09

Coefficient 1 = 456.08 mM^{-1} cm^{-1}
Coefficient 2 = 438.38 mM^{-1} cm^{-1}

Average +/- error = 447.23 +/- 8.8 mM^{-1} cm^{-1}

Raw data for calculation of the extinction coefficient of cyanidin 3-*O*-glucoside 5-*O*-(3-4-Difluorocinnamoyl)glucoside at 520 nm (10 µl injection volume)

Conc. (mM)	*Absorbance 1*	*Absorbance 2*
1	447	469
0.5	227	242
0.1	38.6	52.4
0.01	2.82	4.78

Coefficient 1 = 472.45 mM^{-1} cm^{-1}
Coefficient 2 = 447.85 mM^{-1} cm^{-1}
Average +/- error = 460.15 +/- 12.3 mM^{-1} cm^{-1}

Raw data for calculation of the extinction coefficient of coumaroyl coenzyme A at 333 nm

Conc. (mM)	*Absorbance 1*	*Absorbance 2*
0.001	0.018	0.020
0.01	0.198	0.210
0.05	0.998	1.005
0.1	2.033	2.047

Coefficient 1 = 20.40 mM^{-1} cm^{-1}
Coefficient 2 = 20.25 mM^{-1} cm^{-1}
Average +/- error = 20.33 +/- 0.075 mM^{-1} cm^{-1}

Raw data for calculation of the extinction coefficient of caffeoyl coenzyme A at 347 nm

Conc. (mM)	*Absorbance 1*	*Absorbance 2*
0.001	0.017	0.018
0.01	0.176	0.179
0.05	0.872	0.888
0.1	1.815	1.835

Coefficient 1 = 18.005 mM^{-1} cm^{-1}
Coefficient 2 = 18.229 mM^{-1} cm^{-1}
Average +/- error = 18.117 +/- 0.11 mM^{-1} cm^{-1}

Raw data for calculation of the extinction coefficient of feruloyl coenzyme A at 350 nm

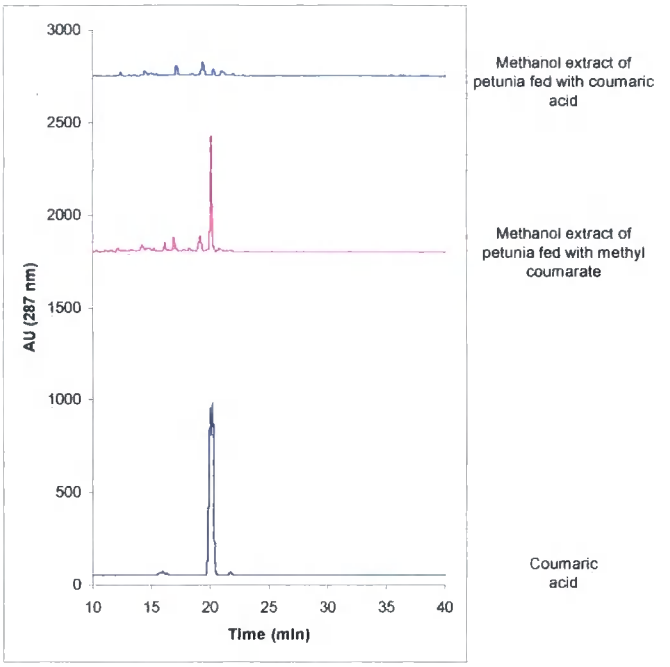
Conc. (mM)	*Absorbance 1*	*Absorbance 2*
0.001	0.019	0.019
0.01	0.188	0.194
0.05	0.947	0.955
0.1	1.897	1.911

Coefficient 1 = 18.963 mM^{-1} cm^{-1}
Coefficient 2 = 19.110 mM^{-1} cm^{-1}
Average +/- error = 19.037 +/- 0.07 mM^{-1} cm^{-1}

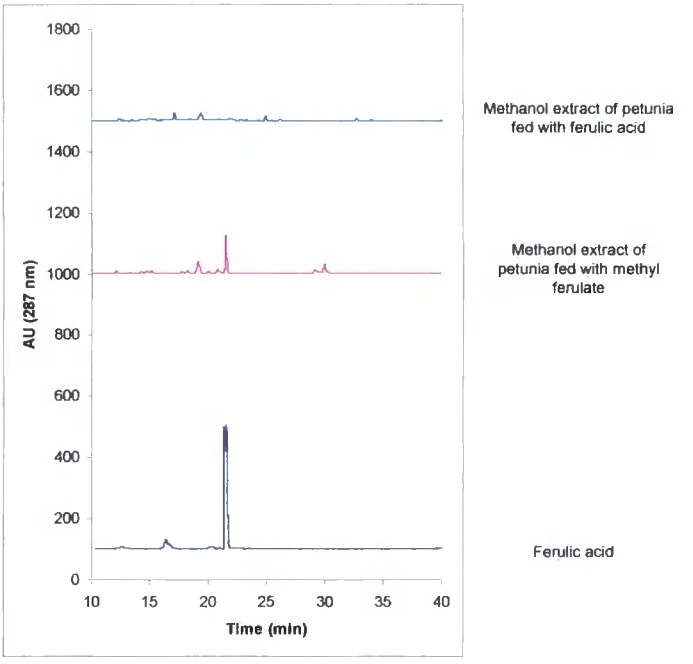
Appendix D

Hydrolysis of exogenously-fed methyl phenylpropanoid esters in petunia

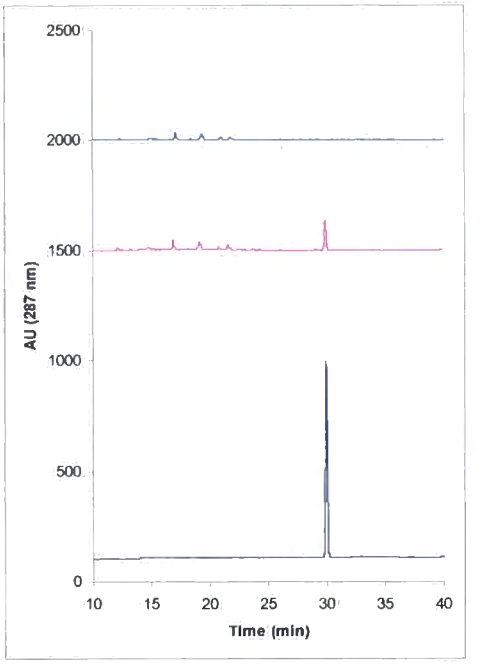
Coumaric acid



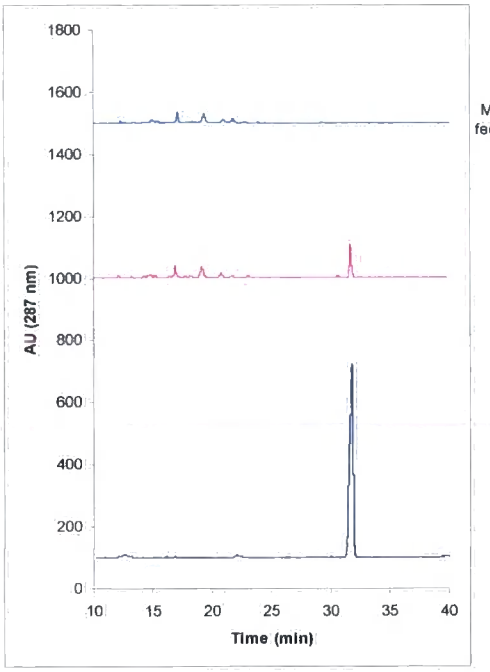
Ferulic acid



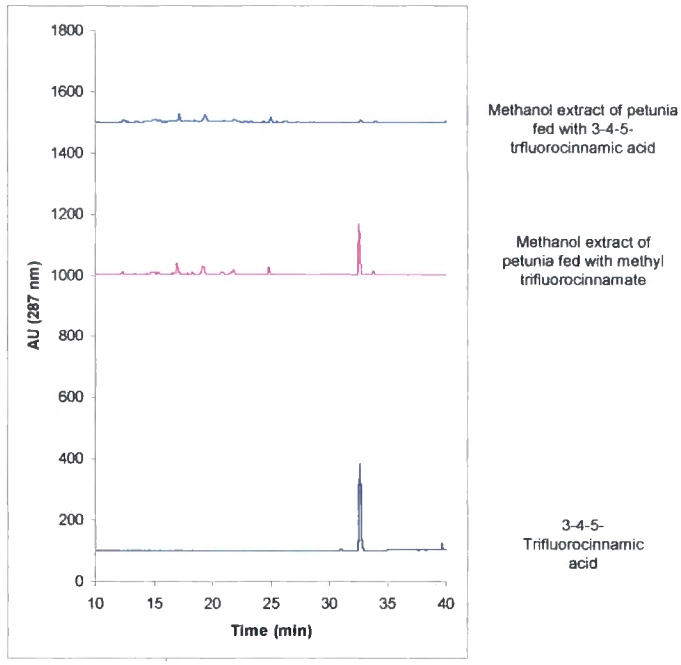
4-Fluorocinnamic acid



3-4-Difluorocinnamic acid



3-4-5-Trifluorocinnamic acid

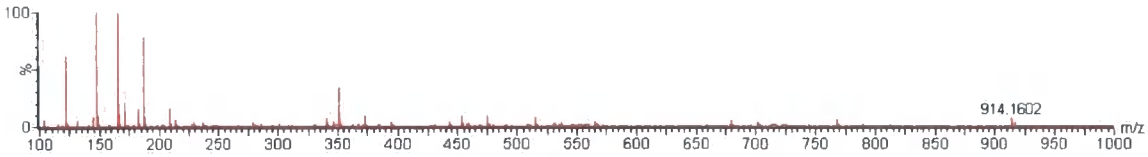


Appendix E

High-resolution mass spectrometry (ES^+) of biosynthesised coenzyme A esters

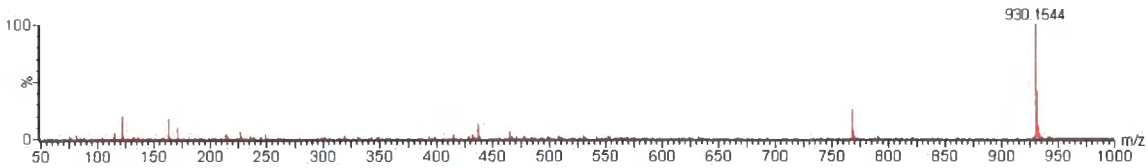
Coumaroyl CoA

$(M + H)^+ = 914.1598$



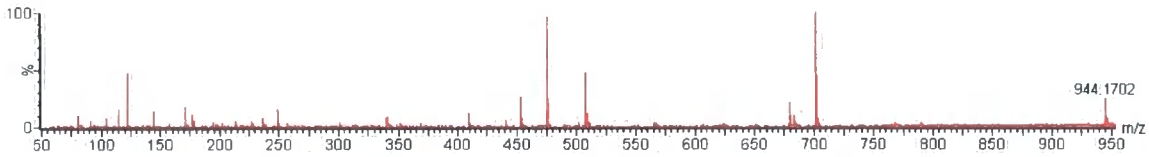
Caffeoyl CoA

$(M + H)^+ = 930.1547$



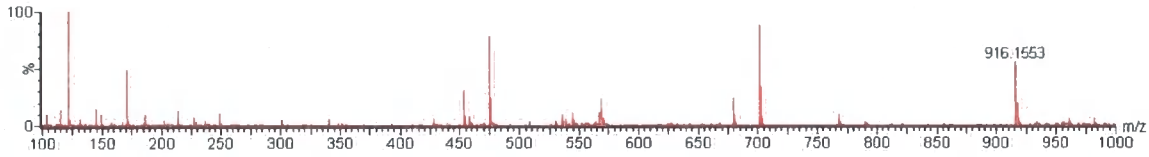
Feruloyl CoA

$(M + H)^+ = 944.1704$



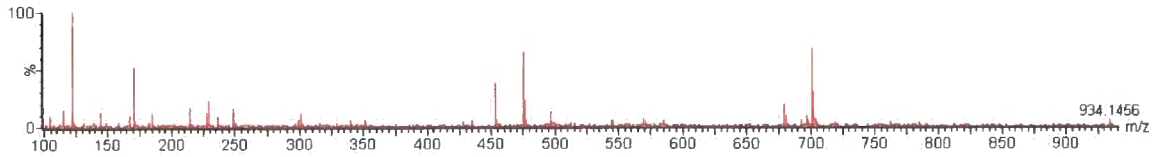
4- Fluorocinnamoyl CoA

$(M + H)^+ = 916.1555$



3-4- Difluorocinnamoyl CoA

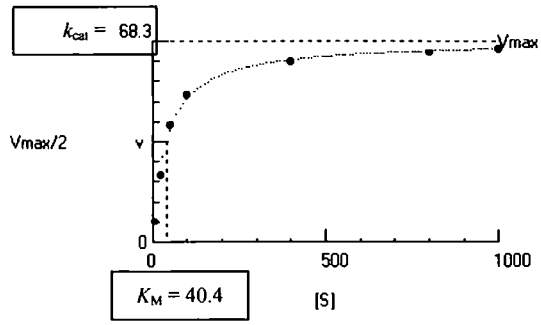
$(M + H)^+ = 934.1461$



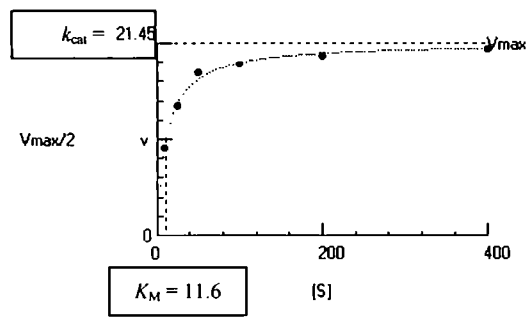
Appendix F

Kinetic data for activity of At4CL1 toward phenylpropanoid substrates

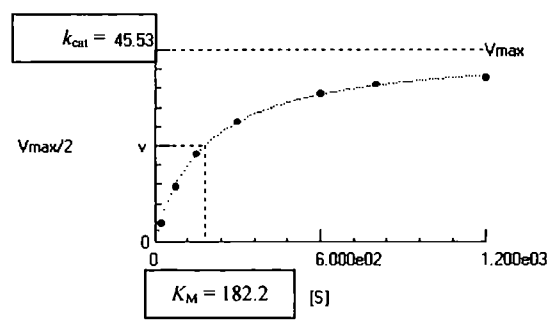
Coumaric acid



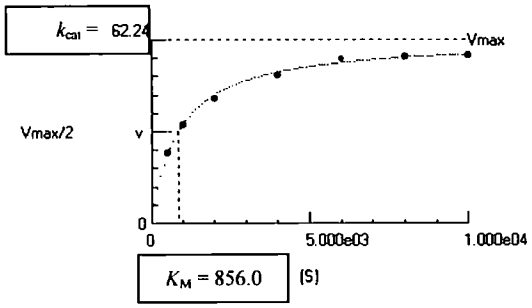
Caffeic acid



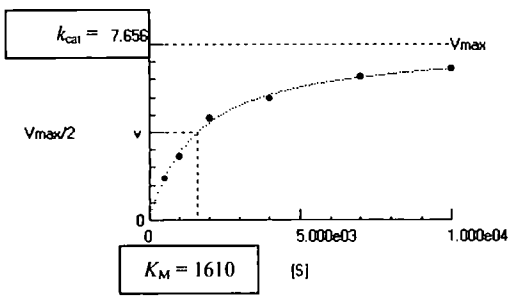
Ferulic acid



4- Fluorocinnamic acid



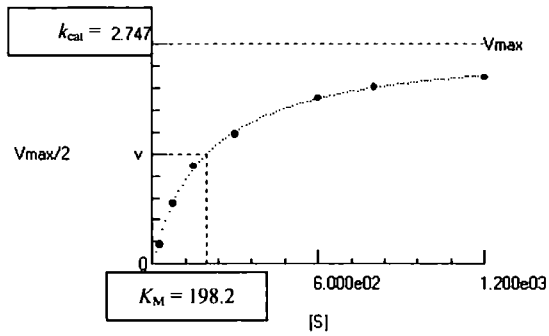
3-4- Difluorocinnamic acid



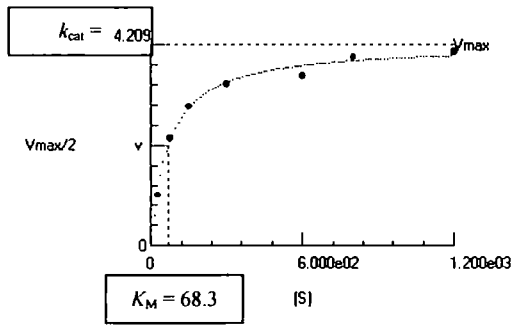
Appendix G

Kinetic data for activity of Gent5AT toward phenylpropanoyl CoA substrates

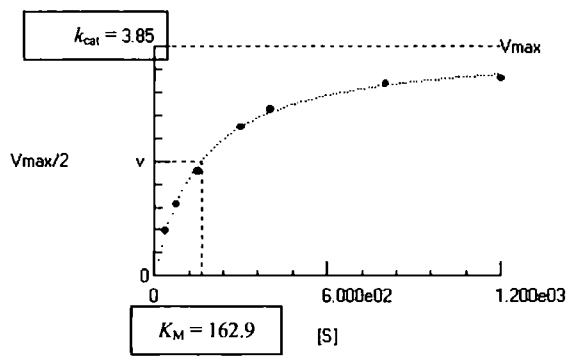
Coumaric acid



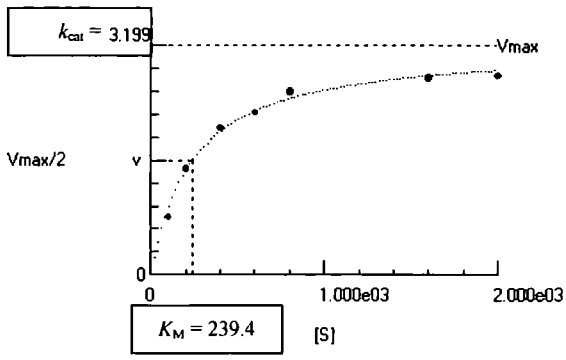
Caffeic acid



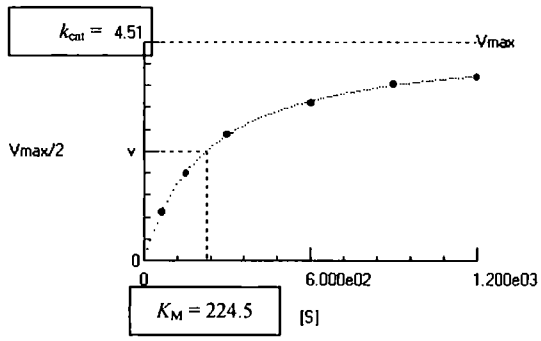
Ferulic acid



4-Fluorocinnamic acid



3-4-Difluorocinnamic acid

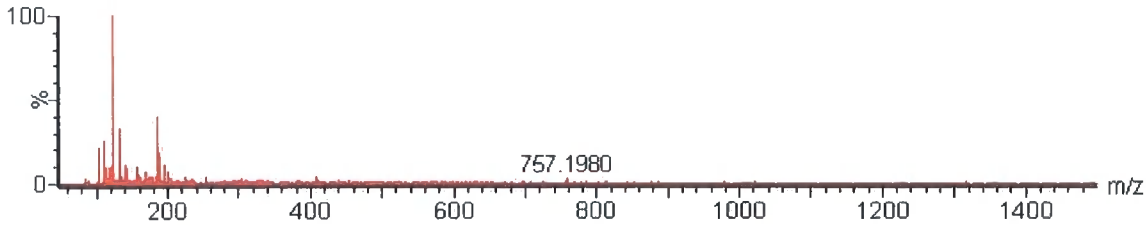


Appendix H

High Resolution Mass Spectrometry of Biosynthesised Acylated Anthocyanins

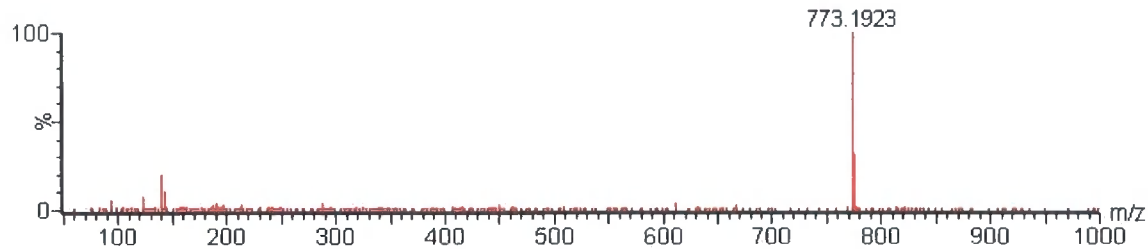
Cyanidin 3-*O*-glucoside 5-*O*-(Coumaroyl)glucoside

$(M)^+$ requires 757.1974



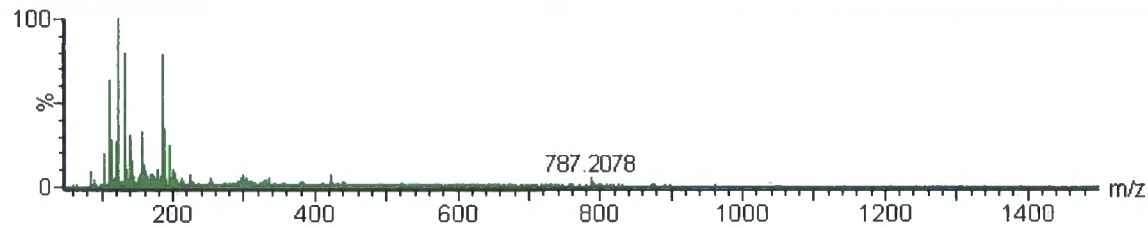
Cyanidin 3-*O*-glucoside 5-*O*-(Caffeoyl)glucoside

$(M)^+$ requires 773.1924



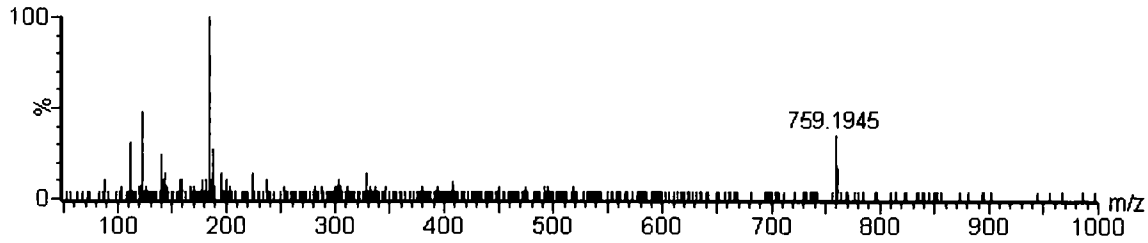
Cyanidin 3-*O*-glucoside 5-*O*-(Feruloyl)glucoside

$(M)^+$ requires 787.2080



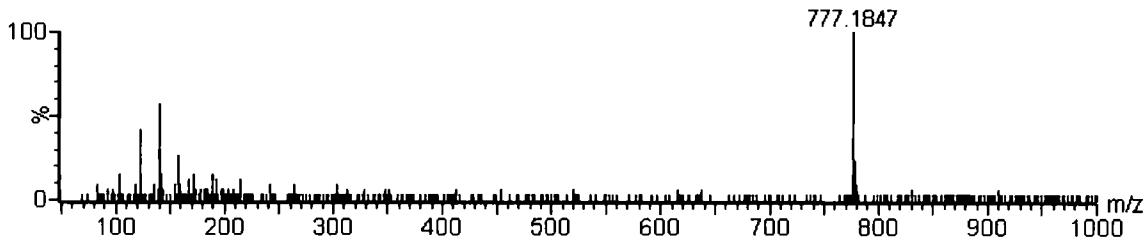
Cyanidin 3-*O*-glucoside 5-*O*-(4-Fluorocinnamoyl)glucoside

$(M)^+$ requires 759.1935



Cyanidin 3-*O*-glucoside 5-*O*-(3-4-Difluorocinnamoyl)glucoside

$(M)^+$ requires 777.1837

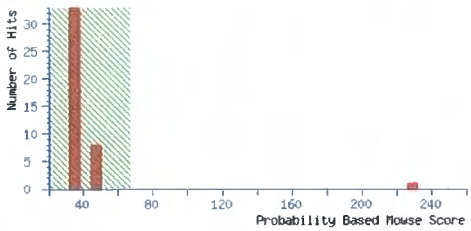


Appendix I
MALDI-TOF identification of peptides following tryptic digestion

Gent5AT

Probability Based Mowse Score

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 67 are significant (p<0.05).



1. gi|4185599 **Mass:** 53162 **Score:** 229 **Expect:** 3.4e-18 **Queries matched:** 30
Anthocyanin 5-aromatic acyltransferase [Gentiana triflora]

Protein View

Nominal mass (M_r): **53235**; Calculated pI value: **6.15**
NCBI BLAST search of gi|51971299 against nr
Unformatted sequence string for pasting into other applications

Taxonomy: Gentiana scabra var. buergeri

Fixed modifications: Carbamidomethyl (C)
Variable modifications: N-Acetyl (Protein),Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Number of mass values searched: **90**
Number of mass values matched: **22**
Sequence Coverage: **51%**

Matched peptides shown in **Bold Red**

```
    1 MEQIQMVKVL EKCQVTPPFD TTDVELSLPV TFFDIPWLHL NKMQSLLFYD
   51 FPVPKTHFLD TVVPNLKASL SLTLKHYVPL SGMLLMPIKS GKMPKFRYSR
  101 DEGDSITLIF AESDQDFDNL KGHQLVDSND LHALFYVMPR VIRTMQDYKV
  151 IPLVAVQVTV FPNRGIAVAL TAMHSIADAK SFVMFINAWA YINKFGKDAD
  201 LLSANLLPSF DRSIIKDPYG LEETFWNEMQ HVLEMFSRFG SKPPRFNKVR
  251 ATYVLSLAEI QKLKNKVLNL RGSEPTIRVT TFTVTCGYVW TCMVKSKDDV
  301 VSEESSNDEN ELEYFSFTAD CRGLLTPPCP PNYFGNCLAS CVAKATHKEL
  351 VGNKGLLVAV AAIVEAIDKR VHNERGVLAD AKTWLSESNG IPSKRFLGIT
  401 GSPKFDSYGV DFGWGKPAKF DITSVDYAEL IYVIQSRDFE KGVEIGVSLP
  451 KIHMDAFAKI FEEGFCSLS
```

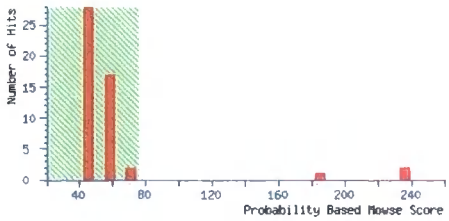
Start - End	Observed	Mr(expt)	Mr(calc)	ppm	Miss	Sequence
1 - 1			149.0510		0	M
1 - 8			1005.4987		0	MEQIQMVK
1 - 12			1474.7887		1	MEQIQMVKVLEK
1 - 12	1533.8590	1532.8517	1532.7942	38	1	-.MEQIQMVKVLEK.C N-Acetyl (Protein); Oxidation (M)
2 - 8			874.4582		0	EQIQMVK
2 - 12			1343.7482		1	EQIQMVKVLEK
9 - 12			487.3006		0	VLEK
9 - 42			3998.0539		1	VLEKCQVTPPFDTTDVELSLPVTFFDIPWLHLNK
13 - 42			3528.7639		0	CQVTPPFDTTDVELSLPVTFFDIPWLHLNK
13 - 55			5158.5540		1	CQVTPPFDTTDVELSLPVTFFDIPWLHLNKMQSLLFYDFPYPK
43 - 55			1647.8007		0	MQSLLFYDFPYPK
43 - 67			3012.5459		1	MQSLLFYDFPYPKTHFLDTVVPNLK
56 - 67			1382.7558		0	THFLDTVVPNLK
56 - 75			2196.2518		1	THFLDTVVPNLKASLSLTLK
68 - 75			831.5066		0	ASLSLTLK
68 - 89			2394.3708		1	ASLSLTLKHYVPLSGNLLMPIK
76 - 89	1581.8470	1580.8397	1580.8748	-22	0	K.HYVPLSGNLLMPIK.S
76 - 89	1597.8250	1596.8177	1596.8697	-33	0	K.HYVPLSGNLLMPIK.S Oxidation (M)
76 - 92			1853.0233		1	HYVPLSGNLLMPIKSGK
90 - 92			290.1590		0	SGK
90 - 95			646.3472		1	SGKMPK
93 - 95			374.1988		0	MPK
93 - 97			677.3683		1	MPKFR
96 - 97			321.1801		0	FR
96 - 100			727.3766		1	FRYSR
98 - 100			424.2070		0	YSR
98 - 121			2777.2668		1	YSRDEGDSITLIFAESDQDFDNLK
101 - 121			2371.0703		0	DEGDSITLIFAESDQDFDNLK
101 - 140			4564.1492		1	DEGDSITLIFAESDQDFDNLKGHQLVDSNDLHALFYVHPR
122 - 140			2211.0895		0	GHQLVDSNDLHALFYVHPR
122 - 143			2579.3430		1	GHQLVDSNDLHALFYVHPRVIR
141 - 143			386.2641		0	VIR
141 - 149			1152.5961		1	VIRTHQDYK
144 - 149			784.3425		0	THQDYK
144 - 164			2417.3140		1	THQDYKVIPLVAVQVTVFPNR
150 - 164	1651.9620	1650.9547	1650.9020	-17	0	K.VIPLVAVQVTVFPNR.G
150 - 180			3206.8291		1	VIPLVAVQVTVFPNRGIAVALTAHHSIADAK
165 - 180	1574.8020	1573.7947	1573.8576	-40	0	R.GIAVALTAHHSIADAK.S
165 - 194			3258.7011		1	GIAVALTAHHSIADAKSFVHFINAWAYINK
181 - 194	1703.8430	1702.8357	1702.8541	-11	0	K.SFVHFINAWAYINK.F
181 - 194	1719.8230	1718.8157	1718.8490	-19	0	R.SFVHFINAWAYINK.F Oxidation (M)
181 - 197			2035.0389		1	SFVHFINAWAYINKFGK
195 - 197			350.1954		0	FGK
195 - 212			1978.0159		1	FGKDADLLSANLLPSFDR
198 - 212	1646.7930	1645.7857	1645.8311	-28	0	K.DADLLSANLLPSFDR.S
198 - 216			2087.1262		1	DADLLSANLLPSFDRSIIK
213 - 216			459.3057		0	SIIK
213 - 238			3198.5154		1	SIIKDPYGLEETFWNEMQHVLEMFSR
217 - 238			2757.2203		0	DPYGLEETFWNEMQHVLEMFSR
217 - 245			3526.6438		1	DPYGLEETFWNEMQHVLEMFSRFGSKPPR
239 - 245	788.4390	787.4317	787.4340	-3	0	R.FGSKPPR.F
239 - 248			1176.6404		1	FGSKPPRFNK
246 - 248			407.2169		0	FNK
246 - 250			662.3864		1	FNKVR
249 - 250			273.1801		0	VR
249 - 262			1589.9141		1	VRATYVLSLAEIQK
251 - 262	1335.7390	1334.7317	1334.7445	-10	0	R.ATYVLSLAEIQK.L
251 - 264			1575.9236		1	ATYVLSLAEIQKLK
263 - 264			259.1896		0	LK
263 - 266			501.3275		1	LKNK
265 - 266			260.1485		0	NK
265 - 271	856.5220	855.5147	855.5290	-17	1	K.NKVLNLR.G
267 - 271	614.3950	613.3877	613.3911	-6	0	K.VLNLR.G
267 - 278			1353.7728		1	VLNLRGSEPTIR
272 - 278	759.4010	758.3937	758.3922	2	0	R.GSEPTIR.V
272 - 295			2792.3335		1	GSEPTIRVTTFTVTCGYVWTCNVK
279 - 295			2051.9519		0	VTTFTVTCGYVWTCNVK
279 - 297			2267.0788		1	VTTFTVTCGYVWTCNVKSK
296 - 297			233.1376		0	SK
296 - 322	3172.2500	3171.2427	3171.3098	-21	1	K.SKDDVVSEESSNDENELEYFSFTADCR.G
298 - 322			2956.1828		0	DDVVSEESSNDENELEYFSFTADCR
298 - 344			5373.3158		1	DDVVSEESSNDENELEYFSFTADCRGLLTPPCPPNYFGNCLASCVAK
323 - 344	2436.1420	2435.1347	2435.1435	-4	0	R.GLLTPPCPPNYFGNCLASCVAK.A
323 - 348			2872.3822		1	GLLTPPCPPNYFGNCLASCVAKATHK
345 - 348			455.2492		0	ATHK
345 - 354			1095.6036		1	ATHKELVGNK
349 - 354			658.3650		0	ELVGNK
349 - 369			2121.2408		1	ELVGNKGLLVAVAAIVEAIDK
355 - 369			1480.8864		0	GLLVAVAAIVEAIDK
355 - 370			1636.9875		1	GLLVAVAAIVEAIDKR
370 - 370			174.1117		0	R
370 - 375			809.4256		1	RVHNER
371 - 375			653.3245		0	VHNER
371 - 382			1307.6946		1	VHNERGVLADAK
376 - 382			672.3806		0	GVLADAK
376 - 394			1972.0265		1	GVLADAKTWLSESNGIPSK
383 - 394	1310.6630	1317.6557	1317.6564	-1	0	K.TWLSESNGIPSK.R
383 - 395	1474.7640	1473.7567	1473.7576	-1	1	K.TWLSESNGIPSKR.F
395 - 395			174.1117		0	R
395 - 404			1074.6186		1	RFLGITGSPK

396 - 404	919.5190	918.5117	918.5174	-6	0	R.FLGITGSPK.F
396 - 419			2573.2954		1	FLGITGSPKFDSYGVDFGWGKPAK
405 - 419	1673.7030	1672.7757	1672.7885	-8	0	R.FDSYGVDFGWGKPAK.F
405 - 437			3785.8617		1	FDSYGVDFGWGKPAKFDITSVDYAELIYVIQSR
420 - 437	2132.0790	2131.0717	2131.0837	-6	0	K.FDITSVDYAELIYVIQSR.D
420 - 441			2650.3166		1	FDITSVDYAELIYVIQSRDFEK
438 - 441			537.2435		0	DFEK
438 - 451			1516.8137		1	DFEKGVEIGVSLPK
442 - 451	998.5710	997.5637	997.5807	-17	0	K.GVEIGVSLPK.I
442 - 459			1911.0287		1	GVEIGVSLPKIHMDAFAK
452 - 459	932.4640	931.4567	931.4585	-2	0	K.IHMDAFAK.I
452 - 469			2100.9648		1	IHMDAFAKIFEEGFCSLS
460 - 469			1187.5169		0	IFEEGFCSLS

At1g76140

Probability Based Mowse Score

Ions score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 76 are significant ($p<0.05$).
Protein scores are derived from ions scores as a non-probabilistic basis for ranking protein hits.



	Accession	Mass	Score	Description
1.	gi|30699145	82187	236	prolyl oligopeptidase, putative / prolyl endopeptidase, putative / post-proline cleaving enzyme

Protein View

Match to: **gi|30699145** Score: **236** Expect: **5.9e-018**
prolyl oligopeptidase, putative / prolyl endopeptidase, putative / post-proline cleaving enzyme

Nominal mass (M_r): **82187**; Calculated pI value: **5.18**
NCBI BLAST search of gi|30699145 against nr
Unformatted sequence string for pasting into other applications

Taxonomy: Arabidopsis thaliana

Variable modifications: Carboxymethyl (C),Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Sequence Coverage: **11%**

Matched peptides shown in **Bold Red**

```
    1 MGSSSVFGEQ LQYPATRRDD SVVDDYHGVK IGDPYRWLED PDAEEVKEFV
   51 QSQVKLTDSV LEKCETKEKL RQNITKLIDH PRYDSPFRQG DKYFYFHNTG
  101 LQAQSVLYMQ DNLDAEPEVL LDPNTLSDDG TVALNTFSVS EDAKYLAYGL
  151 SSSGSDWVTI KLMKIEDKKV EPDTLSWVKF TGITWTHDSK GFFYGRYPAP
  201 KEGEDIDAGT ETNSNLYHEL YYHFIGTDQS QDILCWRDNE NPKYMFGAEV
  251 TDDGKYLIMS IGESCDPVNK LYYCDMTSLS GGLESFRGSS SFLPFIKLVD
  301 TFDAQYSVIS NDETLFTFLT NKDAPKYKLV RVDLKEPNSW TDVVEEHEKD
  351 VLASACAVNG NHLVACYMSD VKHILQIRDL KSGSLLHQLP LDIGSVSDVS
  401 ARRKDNTFFF SFTSFLTPGV IYKCDLANES PEVKVFREVT VPGFDREAFQ
  451 AIQVFYPSKD GTKIPMFIVA KKDIKLDGSH PCLLYAYGGF NISITPSFSA
  501 SRIVLSKHLG VVFCFANIRG GGEYGEEWHK AGSLAKKQNC FDDFISGAEY
  551 LVSAGYTQPS KLCIEGGSNG GLLVGACINQ RPDLYGCALA HVGVMDMLRF
  601 HKFTIGHAWT SDYGCSENEE EFHWLIKYSP LHNVKRPWEQ QTDHLVQYPS
  651 TMLLTADHDD RVVPLHSLKL LATLQHVLCT SLDNSPQMNP IIGRIEVKAG
  701 HGAGRPTQKM IDEAADRYSF MAKMVNASWT E
```

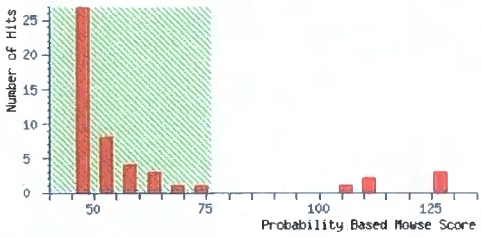
Start - End	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Sequence
1 - 1			149.0510		0	M
1 - 17			1856.8727		0	MGSSSVFGEQLQYPATR
1 - 18			2012.9736		1	MGSSSVFGEQLQYPATRR
2 - 17			1725.8322		0	GSSSVFGEQLQYPATR
2 - 18			1881.9333		1	GSSSVFGEQLQYPATRR
18 - 18			174.1117		0	R
18 - 30			1503.6954		1	RDDSVVDDYHGVK
19 - 30			1347.5942		0	DDSVVDDYHGVK
19 - 36			2048.9439		1	DDSVVDDYHGVKIGDPYR
31 - 36			719.3602		0	IGDPYR
31 - 47			2030.9585		1	IGDPYRWLEDPDAEEVK
37 - 47			1329.6088		0	WLEDPDAEEVK
37 - 55			2275.1008		1	WLEDPDAEEVKEFVQSQVK
48 - 55			963.5025		0	EFVQSQVK
48 - 63			1848.9833		1	EFVQSQVKLTDSVLEK
56 - 63			903.4913		0	LTDSVLEK
56 - 67			1364.6857		1	LTDSVLEKCETK
64 - 67			479.2050		0	CETK
64 - 69			736.3425		1	CETKER
68 - 69			275.1481		0	ER
68 - 71			544.3333		1	ERLR
70 - 71			287.1957		0	LR
70 - 76			871.5239		1	LRQNITK
72 - 76			602.3388		0	QNITK
72 - 82			1333.7466		1	QNITKLIDHPR
77 - 82			749.4184		0	LIDHPR
77 - 88			1514.7630		1	LIDHPRYDSPFR
83 - 88			783.3551		0	YDSPFR
83 - 92			1211.5571		1	YDSPFRQGDK
89 - 92			446.2125		0	QGDK
89 - 144			6222.9167		1	QGDKYFYFHNTGLQAQSVLYMQDNLDAEPEVLLDPNTLSDDGTVALNTFSVSEDAK
93 - 144			5794.7148		0	YFYFHNTGLQAQSVLYMQDNLDAEPEVLLDPNTLSDDGTVALNTFSVSEDAK
93 - 161			7622.6191		1	YFYFHNTGLQAQSVLYMQDNLDAEPEVLLDPNTLSDDGTVALNTFSVSEDAKYLAYGLSSSGSDWVTIK
145 - 161			1845.9149		0	YLAYGLSSSGSDWVTIK
145 - 164			2218.1344		1	YLAYGLSSSGSDWVTIKLMK
162 - 164			390.2301		0	LMK
162 - 168			875.4786		1	LMKIEDK
165 - 168			503.2591		0	IEDK
165 - 169			631.3541		1	IEDKK
169 - 169			146.1055		0	K
169 - 179			1300.7027		1	KVEPDTLSWVK
170 - 179			1172.6077		0	VEPDTLSWVK
170 - 190			2446.2168		1	VEPDTLSWVKFTGITWTHDSK
180 - 190	**1292.6232**	**1291.6159**	**1291.6197**	**-0.0038**	**0**	**K.FTGITWTHDSK.G** (Ions score 33)
180 - 196			2018.9639		1	FTGITWTHDSKGFFYGR
191 - 196			745.3547		0	GFFYGR
191 - 201			1301.6557		1	GFFYGRYPAPK
191 - 196			745.3547		0	GFFYGR
191 - 201			1301.6557		1	GFFYGRYPAPK
197 - 201			574.3115		0	YPAPK
197 - 237			4788.1714		1	YPAPKEGEDIDAGTETNSNLYHELYYHFIGTDQSQDILCWR
202 - 237			4231.8705		0	EGEDIDAGTETNSNLYHELYYHFIGTDQSQDILCWR
202 - 243			4929.1736		1	EGEDIDAGTETNSNLYHELYYHFIGTDQSQDILCWRDNENPK
238 - 243			715.3137		0	DNENPK
238 - 255			2028.8734		1	DNENPKYMFGAEVTDDGK
244 - 255			1331.5703		0	YMFGAEVTDDGK
244 - 270			2981.3496		1	YMFGAEVTDDGKYLIMSIGESCDPVNK
256 - 270			1667.7898		0	YLIMSIGESCDPVNK
256 - 287			3590.6441		1	YLIMSIGESCDPVNKLYYCDMTSLSGGLESFR
271 - 287			1940.8648		0	LYYCDMTSLSGGLESFR
271 - 297			3004.4350		1	LYYCDMTSLSGGLESFRGSSSFLPFIK
288 - 297			1081.5808		0	GSSSFLPFIK
288 - 322			3943.9771		1	GSSSFLPFIKLVDTFDAQYSVISNDETLFTFLTNK
298 - 322			2880.4069		0	LVDTFDAQYSVISNDETLFTFLTNK
298 - 326			3291.6187		1	LVDTFDAQYSVISNDETLFTFLTNKDAPK
323 - 326			429.2223		0	DAPK
323 - 328			720.3806		1	DAPKYK
327 - 328			309.1689		0	YK
327 - 331			677.4224		1	YKLVR
329 - 331			386.2641		0	LVR
329 - 335			841.5385		1	LVRVDLK
332 - 335			473.2849		0	VDLK
332 - 349			2153.0276		1	VDLKEPNSWTDVVEEHEK
336 - 349			1697.7533		0	EPNSWTDVVEEHEK
336 - 372			4058.8448		1	EPNSWTDVVEEHEKDVLASACAVNGNHLVACYMSDVR
350 - 372			2379.1021		0	DVLASACAVNGNHLVACYMSDVK
350 - 378			3139.5728		1	DVLASACAVNGNHLVACYMSDVKHILQIR
373 - 378			778.4813		0	HILQIR
373 - 381			1134.6873		1	HILQIRDLK
379 - 381			374.2165		0	DLK
379 - 402			2506.3391		1	DLKSGSLLHQLPLDIGSVSDVSAR
382 - 402	**2151.1240**	**2150.1167**	**2150.1331**	**-0.0164**	**0**	**K.SGSLLHQLPLDIGSVSDVSAR.R** (Ions score 13)
382 - 403			2306.2342		1	SGSLLHQLPLDIGSVSDVSARR
403 - 403			174.1117		0	R
403 - 404			302.2066		1	RK
404 - 404			146.1055		0	K
404 - 423			2358.1936		1	KDNTFFFSFTSFLTPGVIYK
405 - 423			2230.0986		0	DNTFFFSFTSFLTPGVIYK
405 - 434			3415.6322		1	DNTFFFSFTSFLTPGVIYKCDLANESPEVK
424 - 434			1203.5441		0	CDLANESPEVK
424 - 437			1605.7621		1	CDLANESPEVKVFR
435 - 437			420.2485		0	VFR
435 - 446			1420.7463		1	VFREVTVPGFDR
438 - 446	**1019.5148**	**1018.5075**	**1018.5083**	**-0.0008**	**0**	**R.EVTVPGFDR.E** (Ions score 16)

438 - 459			2527.2747		1	EVTVPGFDREAFQAIQVFYPSK
447 - 459	1527.7773	1526.7700	1526.7769	-0.0068	0	R.EAFQAIQVFYPSK.D (Ions score 66)
447 - 463			1927.9679		1	EAFQAIQVFYPSKDGTK
460 - 463			419.2016		0	DGTK
460 - 471			1318.7319		1	DGTKIPHFIVAK
464 - 471			917.5408		0	IPHFIVAK
464 - 472			1045.6358		1	IPHFIVAKK
472 - 472			146.1055		0	K
472 - 475			502.3115		1	KDIK
473 - 475			374.2165		0	DIK
473 - 502			3228.5913		1	DIKLDGSHPCLLYAYGGFNISITPSFSASR
476 - 502			2872.3854		0	LDGSHPCLLYAYGGFNISITPSFSASR
476 - 507			3412.7489		1	LDGSHPCLLYAYGGFNISITPSFSASRIVLSK
503 - 507			558.3741		0	IVLSK
503 - 519			1915.0865		1	IVLSKHLGVVFCFANIR
508 - 519			1374.7230		0	HLGVVFCFANIR
508 - 530			2604.2331		1	HLGVVFCFANIRGGGEYGEEWHK
520 - 530	1248.5321	1247.5248	1247.5206	0.0042	0	R.GGGEYGEEWHK.A (No match)
520 - 536			1774.8274		1	GGGEYGEEWHKAGSLAK
531 - 536			545.3173		0	AGSLAK
531 - 537			673.4123		1	AGSLAKK
537 - 537			146.1055		0	K
537 - 561			2767.2799		1	KQNCFDDFISGAEYLVSAGYTQPSK
538 - 561			2639.1849		0	QNCFDDFISGAEYLVSAGYTQPSK
538 - 599			6536.0829		1	QNCFDDFISGAEYLVSAGYTQPSKLCIEGGSNGGLLVGACINQRPDLYGCALAHVGVHDHLR
562 - 599			3914.9085		0	LCIEGGSNGGLLVGACINQRPDLYGCALAHVGVHDHLR
562 - 602			4327.1308		1	LCIEGGSNGGLLVGACINQRPDLYGCALAHVGVHDHLRFHK
600 - 602			430.2328		0	FHK
600 - 627			3410.5454		1	FHKFTIGHAWTSDYGCSENEEEFHWLIK
603 - 627			2998.3232		0	FTIGHAWTSDYGCSENEEEFHWLIK
603 - 635			3936.8205		1	FTIGHAWTSDYGCSENEEEFHWLIKYSPLHNVK
628 - 635	957.5123	956.5050	956.5079	-0.0029	0	K.YSPLHNVK.R (Ions score 28)
628 - 661			4089.9755		1	YSPLHNVKRPWEQQTDHLVQYPSTHLLTADHDDR
636 - 661			3151.4781		0	RPWEQQTDHLVQYPSTHLLTADHDDR
636 - 669			4025.0217		1	RPWEQQTDHLVQYPSTHLLTADHDDRVVPLHSLK
662 - 669			891.5542		0	VVPLHSLK
662 - 694			3606.9741		1	VVPLHSLKLLATLQHVLCTSLDNSPQHNPIIGR
670 - 694			2733.4305		0	LLATLQHVLCTSLDNSPQHNPIIGR
670 - 698			3202.7206		1	LLATLQHVLCTSLDNSPQHNPIIGRIEVK
695 - 698			487.3006		0	IEVK
695 - 709			1547.8532		1	IEVKAGHGAGRPTQR
699 - 709	1079.5671	1078.5598	1078.5631	-0.0033	0	K.AGHGAGRPTQK.H (Ions score 1)
699 - 717			1979.9595		1	AGHGAGRPTQKHIDEAADR
710 - 717			919.4069		0	HIDEAADR
710 - 723			1646.7432		1	HIDEAADRYSFHAK
718 - 723			745.3469		0	YSFHAK
718 - 731			1663.7374		1	YSFHAKHVNASWTE
724 - 731			936.4011		0	HVNASWTE

At1g13440

Probability Based Mowse Score

Ions score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 76 are significant ($p<0.05$).
Protein scores are derived from ions scores as a non-probabilistic basis for ranking protein hits.



	Accession	Mass	Score	Description
1.	gi|19699140	36890	127	At1g13440/F13B4_8 [Arabidopsis thaliana]

Protein View

Match to: **gi|19699140** Score: **87** Expect: **0.0043**
At1g13440/F13B4_8 [Arabidopsis thaliana]

Nominal mass (M_r): **36890**; Calculated pI value: **6.67**
NCBI BLAST search of gi|19699140 against nr
Unformatted sequence string for pasting into other applications

Taxonomy: Arabidopsis thaliana
Links to retrieve other entries containing this sequence from NCBI Entrez:
gi|15222848 from Arabidopsis thaliana
gi|15294160 from Arabidopsis thaliana
gi|15146236 from Arabidopsis thaliana
gi|9958054 from Arabidopsis thaliana

Variable modifications: Carboxymethyl (C),Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Sequence Coverage: **28%**

Matched peptides shown in **Bold Red**

```
    1 MADKKIRIGI NGFGRIGRLV ARVVLQRDDV ELVAVNDPFI TTEYMTYMFK
   51 YDSVHGQWKH HELKVKDDKT LLFGEKPVTV FGIRNPEDIP WGEAGADFVV
  101 ESTGVFTDKD KAAAHLKGGA KKVVISAPSK DAPMFVVGVN EHEYKSDLDI
  151 VSNASCTTNC LAPLAKVIND RFGIVEGLMT TVHSITATQK TVDGPSMKDW
  201 RGGRAASFNI IPSSTGAAKA VGKVLPSLNG KLTGMSFRVP TVDVSVVDLT
  251 VRLEKAATYD EIKKAIKEES EGKMKGILGY TEDDVVSTDF VGDHRSSIFD
  301 AKAGIALSDK FVKLVSWYDN EWGYSSRVVD LIVHMSKA
```

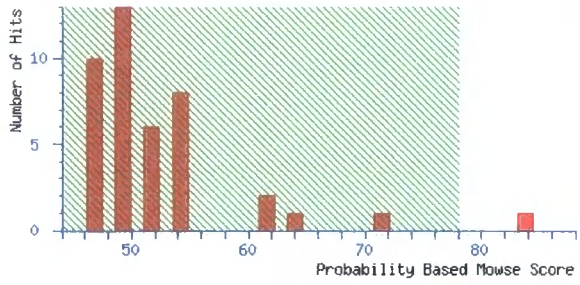
Start - End	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Sequence
1 - 1			149.0510		0	M
1 - 4			463.2101		0	MADK
1 - 5			591.3050		1	MADKK
2 - 4			332.1696		0	ADK
2 - 5			460.2645		1	ADKK
5 - 5			146.1055		0	K
5 - 7			415.2907		1	KIR
6 - 7			287.1957		0	IR
6 - 15			1101.6407		1	IRIGINGFGR
8 - 15			832.4555		0	IGINGFGR
8 - 18			1158.6621		1	IGINGFGRIGR
16 - 18			344.2172		0	IGR
16 - 22			783.5079		1	IGRLVAR
19 - 22			457.3013		0	LVAR
19 - 27			1052.6818		1	LVARVVLQR
23 - 27			613.3911		0	VVLQR
23 - 50			3335.6457		1	VVLQRDDVELVAVNDPFITTEYMTYMFK
28 - 50			2740.2652		0	DDVELVAVNDPFITTEYMTYMFK
28 - 59			3840.7691		1	DDVELVAVNDPFITTEYMTYMFKYDSVHGQWK
51 - 59	1119.5154	1118.5081	1118.5145	-0.0064	0	K.YDSVHGQWK.H (No match)
51 - 64			1762.8539		1	YDSVHGQWKHHELK
60 - 64			662.3500		0	HHELK
60 - 66			889.5134		1	HHELKVK
65 - 66			245.1739		0	VK
65 - 69			603.3228		1	VKDDK
67 - 69			376.1594		0	DDK
67 - 84			2034.1149		1	DDKTLLFGEKPVTVFGIR
70 - 84	1676.9668	1675.9595	1675.9660	-0.0065	0	K.TLLFGEKPVTVFGIR.H (No match)
70 - 84	1676.9668	1675.9595	1675.9660	-0.0065	0	K.TLLFGEKPVTVFGIR.H (No match)
70 - 109			4337.1895		1	TLLFGEKPVTVFGIRNPEDIPWGEAGADFVVESTGVFTDK
85 - 109			2679.2340		0	NPEDIPWGEAGADFVVESTGVFTDK
85 - 111			2922.3559		1	NPEDIPWGEAGADFVVESTGVFTDKDK
110 - 111			261.1325		0	DK
110 - 117			852.4817		1	DKAAAHLK
112 - 117			609.3598		0	AAAHLK
112 - 121			922.5348		1	AAAHLKGGAK
118 - 121			331.1856		0	GGAK
118 - 122			459.2805		1	GGAKK
122 - 122			146.1055		0	K
122 - 130			927.5753		1	KVVISAPSK
123 - 130			799.4803		0	VVISAPSK
123 - 145	2516.1755	2515.1682	2515.2780	-0.1098	1	K.VVISAPSKDAPMFVVGVNEHEYK.S (No match)
131 - 145			1733.8083		0	DAPMFVVGVNEHEYK
131 - 145	1750.8771	1749.8698	1749.8031	0.0667	0	K.DAPMFVVGVNEHEYK.S Oxidation (M) (Ions score 5)
131 - 145	1750.8771	1749.8698	1749.8031	0.0667	0	K.DAPMFVVGVNEHEYK.S Oxidation (M) (No match)
131 - 166			3850.8215		1	DAPMFVVGVNEHEYKSDLDIVSNASCTTNCLAPLAK
131 - 166			3850.8215		1	DAPMFVVGVNEHEYKSDLDIVSNASCTTNCLAPLAK
146 - 166			2135.0238		0	SDLDIVSNASCTTNCLAPLAK
146 - 171			2732.3473		1	SDLDIVSNASCTTNCLAPLAKVINDR
167 - 171			615.3340		0	VINDR
167 - 190			2629.3897		1	VINDRFGIVEGLMTTVHSITATQK
172 - 190			2032.0663		0	FGIVEGLMTTVHSITATQK
172 - 198			2847.4510		1	FGIVEGLMTTVHSITATQKTVDGPSMK
191 - 198			833.3953		0	TVDGPSMK
191 - 201			1290.6026		1	TVDGPSMKDWR
199 - 201			475.2179		0	DWR
199 - 204			745.3620		1	DWRGGR
202 - 204			288.1546		0	GGR
202 - 219			1703.8954		1	GGRAASFNIIPSSTGAAK
205 - 219			1433.7514		0	AASFNIIPSSTGAAK
205 - 223			1788.9734		1	AASFNIIPSSTGAAKAVGK
220 - 223			373.2325		0	AVGK
220 - 231			1181.7132		1	AVGKVLPSLNGK
224 - 231			826.4912		0	VLPSLNGK
224 - 238			1618.8865		1	VLPSLNGKLTGMSFR
232 - 238			810.4058		0	LTGMSFR
232 - 252			2290.2355		1	LTGMSFRVPTVDVSVVDLTVR
239 - 252	1498.8373	1497.8300	1497.8402	-0.0102	0	R.VPTVDVSVVDLTVR.L (No match)
239 - 255			1868.0618		1	VPTVDVSVVDLTVRLEK
253 - 255			388.2322		0	LEK
253 - 263			1279.6660		1	LEKAATYDEIK
256 - 263			909.4443		0	AATYDEIK
256 - 264			1037.5393		1	AATYDEIKK
264 - 264			146.1055		0	K
264 - 267			458.3217		1	KAIK
265 - 267			330.2267		0	AIK
265 - 273			989.5029		1	AIKEESEGK
268 - 273			677.2868		0	EESEGK
268 - 275			936.4222		1	EESEGKMK
274 - 275			277.1460		0	MK
274 - 295			2430.1373		1	MKGILGYTEDDVVSTDFVGDNR
276 - 295	2172.0093	2171.0020	2171.0018	0.0002	0	K.GILGYTEDDVVSTDFVGDNR.S (Ions score 35)
276 - 295	2172.0093	2171.0020	2171.0018	0.0002	0	K.GILGYTEDDVVSTDFVGDNR.S (No match)
276 - 302			2919.3774		1	GILGYTEDDVVSTDFVGDNRSSIFDAK
296 - 302			766.3861		0	SSIFDAK
296 - 310			1521.8038		1	SSIFDAKAGIALSDK
303 - 310			773.4283		0	AGIALSDK
303 - 313			1147.6601		1	AGIALSDKFVK

311 - 313			392.2423		0	FVK	
311 - 327			2135.0112		1	FVKLVSWYDNEWGYSSR	
314 - 327	**1761.7831**	**1760.7758**	**1760.7794**	**-0.0036**	**0**	**K.LVSWYDNEWGYSSR.V**	(Ions score 8)
314 - 327	**1761.7831**	**1760.7758**	**1760.7794**	**-0.0036**	**0**	**K.LVSWYDNEWGYSSR.V**	(No match)
314 - 337			2882.4061		1	LVSWYDNEWGYSSRVVDLIVHMSK	
328 - 337			1139.6372		0	VVDLIVHMSK	
328 - 338			1210.6743		1	VVDLIVHMSKA	
338 - 338			89.0477		0	A	

At3g48690

Probability Based Mowse Score

Protein score is -10*Log(P), where P is the probability that the observed match is a random event. Protein scores greater than 78 are significant ($p<0.05$).



1. gi|50198972 **Mass:** 35773 **Score:** 84 **Expect:** 0.016 **Queries matched:** 16
   At3g48690 [Arabidopsis thaliana]

Protein View

Match to: **gi|50198972** Score: **84** Expect: **0.016**
At3g48690 [Arabidopsis thaliana]

Nominal mass (M_r): **35773**; Calculated pI value: **5.23**
NCBI BLAST search of gi|50198972 against nr
Unformatted sequence string for pasting into other applications

Taxonomy: Arabidopsis thaliana
Links to retrieve other entries containing this sequence from NCBI Entrez:
gi|6523100 from Arabidopsis thaliana
gi|50198972 from Arabidopsis thaliana

Fixed modifications: Carbamidomethyl (C)
Variable modifications: N-Acetyl (Protein),Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Number of mass values searched: **138**
Number of mass values matched: **16**
Sequence Coverage: **46%**

Matched peptides shown in **Bold Red**

```
    1 MDSEIAVDCS PLLKIYKSGR IERLMGEATV PPSSEPQNGV VSKDVVYSAD
   51 MNLSVRIYLP EKAAAETDSK LPLLVYFHGG GFIIETAFSP TYHTFLTTSV
  101 SASNCVAVSV DYRRAPEHPI SVPFDDSWTA LKWVFTHITG SGQEDWLMKH
  151 ADFSRVFLSG DSAGANIVHH MAMRAAKEKL SPGLNDTGIS GIILLHPYFW
  201 SKTPIDEKDT KDETLRMKIE AFWRMASPNS KDGTDDPLLN VVQSESVDLS
  251 GLGCGKVLVM VAEKDALVRQ GWGYAAKLEK SGWKGEVEVV ESEGEDHVFH
  301 LLKPECDNAI EVMHKFSGFI KGGN
```

Start - End	Observed	Mr(expt)	Mr(calc)	ppm	Miss	Sequence
1 - 1			149.0510		0	M
1 - 14			1576.7476		0	MDSEIAVDCSPLLK
1 - 17			1980.9900		1	MDSEIAVDCSPLLKIYK
1 - 17	2040.0740	2039.0667	2038.9955	35	1	-.MDSEIAVDCSPLLKIYK.S N-Acetyl (Protein); Oxidation (M)
2 - 14			1445.7072		0	DSEIAVDCSPLLK
2 - 17	1050.9720	1049.9647	1049.9495	0	1	M.DSEIAVDCSPLLKIYK.S
15 - 17			422.2529		0	IYK
15 - 20			722.4075		1	IYKSGR
18 - 20			318.1652		0	SGR
18 - 23	717.4290	716.4217	716.3929	40	1	K.SGRIER.L
21 - 23			416.2383		0	IER
21 - 43			2424.2318		1	IERLMGEATVPPSSEPQNGVVSK
24 - 43			2026.0041		0	LMGEATVPPSSEPQNGVVSK
24 - 56			3458.6987		1	LMGEATVPPSSEPQNGVVSKDVVYSADNNLSVR
44 - 56	1451.7750	1450.7677	1450.7052	43	0	K.DVVYSADNNLSVR.I
44 - 62			2194.1270		1	DVVYSADNNLSVRIYLPEK
57 - 62	762.4660	761.4587	761.4323	35	0	R.IYLPEK.A
57 - 70	1535.8290	1534.8217	1534.7870	22	1	R.IYLPEKAAAETDSK.L
63 - 70			791.3661		0	AAAETDSK
63 - 113			5509.7180		1	AAAETDSKLPLLVYFHGGGFIIETAFSPTYHTFLTTSVSASNCVAVSVDYR
71 - 113			4736.3624		0	LPLLVYFHGGGFIIETAFSPTYHTFLTTSVSASNCVAVSVDYR
71 - 114			4892.4635		1	LPLLVYFHGGGFIIETAFSPTYHTFLTTSVSASNCVAVSVDYRR
114 - 114			174.1117		0	R
114 - 132	2166.1710	2165.1637	2165.0903	34	1	R.RAPEHPISVPFDDSWTALK.W
115 - 132	2010.0540	2009.0467	2008.9894	29	0	R.APEHPISVPFDDSWTALK.W
115 - 149			4007.9482		1	APEHPISVPFDDSWTALKWVFTHITGSGQEDWLNK
133 - 149	2018.0350	2017.0277	2016.9693	29	0	K.WVFTHITGSGQEDWLNK.H
133 - 155			2730.2939		1	WVFTHITGSGQEDWLNKHADFSR
150 - 155	732.3780	731.3707	731.3351	49	0	K.HADFSR.V
150 - 174			2725.2965		1	HADFSRVFLSGDSAGANIVHHMAMR
156 - 174			2011.9720		0	VFLSGDSAGANIVHHMAMR
156 - 177			2282.1412		1	VFLSGDSAGANIVHHMAMRAAK
175 - 177			288.1797		0	AAK
175 - 179			545.3173		1	AAREK
178 - 179			275.1481		0	EK
178 - 202			2784.4850		1	EKLSPGLNDTGISGIILLHPYFWSK
180 - 202	2528.4330	2527.4257	2527.3474	31	0	K.LSPGLNDTGISGIILLHPYFWSK.T
180 - 208			3210.6964		1	LSPGLNDTGISGIILLHPYFWSKTPIDEK
203 - 208			701.3596		0	TPIDEK
203 - 211			1045.5291		1	TPIDEKDTK
209 - 211			362.1801		0	DTK
209 - 216			976.4825		1	DTKDETLR
212 - 216			632.3129		0	DETLR
212 - 218			891.4484		1	DETLRMK
217 - 218			277.1460		0	MK
217 - 231	1770.8960	1769.8887	1769.8302	33	1	R.MKIEAFWMMASPNSK.D
217 - 231	1706.9000	1705.8927	1705.8252	38	1	R.MKIEAFWMMASPNSK.D Oxidation (M)
219 - 231			1510.6948		0	IEAFWMMASPNSK
219 - 256			4066.8961		1	IEAFWMMASPNSKDGTDDPLLNVVQSESVDLSGLGCGK
232 - 256			2574.2119		0	DGTDDPLLNVVQSESVDLSGLGCGK
232 - 264			3443.7163		1	DGTDDPLLNVVQSESVDLSGLGCGKVLVMVAEK
257 - 264			887.5150		0	VLVMVAEK
257 - 269	1442.8000	1441.0727	1441.8326	20	1	K.VLVMVAEKDALVR.Q
257 - 269	1458.8810	1457.8737	1457.8275	32	1	K.VLVMVAEKDALVR.Q Oxidation (M)
265 - 269			572.3282		0	DALVR
265 - 277			1433.7415		1	DALVRQGWGYAAK
270 - 277	880.4670	879.4597	879.4239	41	0	R.QGWGYAAK.L
270 - 280			1249.6455		1	QGWGYAAKLEK
278 - 280			388.2322		0	LEK
278 - 284			846.4599		1	LEKSGWR
281 - 284			476.2383		0	SGWK
281 - 315			4032.8985		1	SGWKGEVEVVESEGEDHVFHLLKPECDNAIEVMHK
285 - 315			3574.6707		0	GEVEVVESEGEDHVFHLLKPECDNAIEVMHK
285 - 321			4254.0401		1	GEVEVVESEGEDHVFHLLKPECDNAIEVMHKFSGFIK
316 - 321			697.3799		0	FSGFIK
316 - 324			925.4657		1	FSGFIKGGN
322 - 324			246.0964		0	GGN