Thioesterases in fruit and the generation of high impact aroma chemicals

Tapp, Edward

Thioesterases in Fruit and the Generation of High Impact Aroma Chemicals

Edward Tapp

PhD Thesis 2008

School of Biological and Biomedical Sciences

Durham University

Abstract:

Volatile organosulphur compounds (VOSCs) are key ingredients in the aroma of tropical fruit where they are active as both free thiols and the respective thioesters. Present as trace high-impact flavourings, VOSCs are however problematic to extract and have therefore become targets for bioproduction.

This work focuses on the endogenous thioesterases in tropical fruit believed to catalyze the liberation of thiol VOSCs from thioester precursors. Using a simple and sensitive colourimetric assay, a thioesterase activity toward VOSCs was identified in purple passion fruit (*Passiflora edulis* Sims). The enzyme was identified as a cell wall-bound protein in the mesocarp of the fruit. Following extraction with salt solutions, the thioesterase was purified 150-fold and shown to be associated with a 43 kDa polypeptide. Affinity labelling with a biotinylated fluorophosphonate suicide probe showed the enzyme to be a serine hydrolase, with MS-MS sequencing of tryptic digests identifying it as a pectin acetylesterase (PAE). Putative thioesterase PAEs were subsequently cloned from passion fruit and *Arabidopsis thaliana*. The observation that an esterase involved in cell wall modification had a secondary role in hydrolysing esterified VOSCs led to the consideration of further fruit species as a source of the enzyme.

Orange (*Citrus sinensis*) was particularly abundant in thioesterase activity. The enzyme was purified 85-fold and identified as a homologous 43 kDa basic (pI: 9) PAE. The enzyme was stable ($t_{1/2}$: 7 days 22 hours) and demonstrated a high turnover toward VOSCs (k_{cat}: 7.85 sec^{-1}). Freeze-dried orange peel was found to retain activity (>90% activity, 3 months 4°C) and demonstrated comparable productivities to those of immobilized microbial enzymes. Here we have initiated a programme for developing processes for the bioproduction of VOSCs, in which the potential of plant glycohydrolases has been demonstrated.

.

"Man cannot discover new oceans,

unless he has the courage to lose sight of the shore"

André Gide

Table of Contents

List of Figures

List of Tables

Abbreviations

AAT	Alcohol acyltransferase
ADH	Alcohol dehydrogenase
At	*Arabidopsis thaliana*
ATP	Adenoside triphosphate
cDNA	Complementary DNA
CoA	Co-enzyme A
CRL	*Candida rugosa* lipase
Cs	*Citrus sinensis*
*Cs*WH	*Citrus sinensis* wall-bound hydrolase
Da	Dalton
DTT	1,4-dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EST	Expressed sequence tag
FA	Fatty acid
F&F	Flavour and Fragrance
FT	Furfuryl thiol
FTA	Furfuryl thioacetate
GC	Gas chromatography
GCMS	Gas chromatography mass spectrometry
GSH	Glutathione
GST	Glutathione-S-transferase
HPLC	High-performance liquid chromatography
IEF	Isoelectric focusing
IPTG	Isopropylthiogalactoside
LOX	Lipoxygenase
MALDI	Matrix assisted laser desorption ionisation
mBrB	Monobromobimane
MFT	2-methyl furan-3-thiol
MFTA	2-methyl furan-3-thioacetate
mRNA	Messenger RNA
MS	Mass spectrometry
MS-MS	Tandem mass spectrometry

MT	Methyl transferase
NADPH	Reduced nicotinamide adenine dinucleotide phosphate
Pa	Pascal
PAE	Pectin acetylesterase
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
Pe	*Passiflora edulis*
*Pe*WH	*Passiflora edulis* wall-bound hydrolase
PLE	Pig liver esterase
PME	Pectin methylesterase
PVPP	Polyvinylpolypyrrolidone
RT-PCR	Reverse transcription PCR
SDS	Sodium dodecyl sulphate
SPME	Solid phase micro extraction
RACE	Rapid amplification of cDNA ends
ToF	Time of flight
TriFP	Trifunctional probe
VOSCs	Volatile organosulphur compounds

Declaration

No material presented in this thesis has previously been submitted as part of any other degree. Except where acknowledged through referencing, all material is the work of the author.

Statement of copyright

Publications arising from work described in this thesis

Tapp, E.J., Brassington, D. 2007. Detecting volatiles with Ellman's reagent. Speciality Chemicals 27(9): 38-39.

Tapp, E.J., Cummins, I., Brassington, D., Edwards, R. 2008. Determination and isolation of a thioesterase from passion fruit (*Passiflora edulis* Sims) that hydrolyzes volatile thioesters. Journal of Agricultural and Food Chemistry 56: 6623-6630.

Acknowledgements

Firstly my thanks go to my supervisor, Rob Edwards, for his endless support, guidance and encouragement; his tutoring and enthusiasm made my research both enjoyable and exciting.

I also owe a great deal to the team at Oxford Chemicals Ltd. for making me feel a part of their company, in particular, Richard Smith, (Managing Director) whose great energy initiated this project. His drive to make things happen will always be an inspiration to me. Thankyou to my industrial supervisors, David Brassington and Simon Jameson, for their time and expertise in flavour chemistry. I was fortunate to have such a supportive team provided through the Knowledge Transfer Partnership scheme, including Val Wooff and Richard Parker Smith.

This work would not have been possible without the help of my many colleagues. Thankyou to my labmates (Ian, David and Mellisa) for guiding me through the past three years in both biochemistry and molecular biology. The hard work of all those in the analytical services both in Biology and Chemistry made this work possible. In particular, I would like to thank Joanne and Jackie for their help and hard work with proteomic analysis, I kept you both very busy.

On a financial note I would like to acknowledge the BBSRC and Oxford Chemicals for their funding.

Thankyou to my friends, Ollie, Markus, Andrjez, Louise, Al, Blondie and Fran, with whom I enjoyed walking in the mountains, snowboarding, the occasional beer and even a marathon. Others too, Aggie, Vicky, Phil, Caroline, Heather and Milena, who made my time in Durham great fun. A special thank you to my good friend Ian, who taught me what it means to be a biochemist and with whom I shared many exciting, memorable and often terrifying adventures down holes in Yorkshire or up cliffs in Weardale.

A final thankyou goes to my family, Mum, Dad, Will and Izzy, who always support me in everything I do.

To Mum and Dad

1. Introduction

1.1 Flavour and Fragrance

Our world is rich in aroma, a consequence of small volatile chemicals found in abundance throughout nature that evoke sensations we recognise as flavour and fragrance. Humans, animals and insects have all evolved developed senses of taste and smell as a means of detecting the presence of such compounds, some being perceived as pleasant, others repulsive. Hence, odourant chemicals essentially function as attractants or deterrents, signals of communication between organisms that have remarkable influences on behaviour.

Certain tastes and smells warn of danger. Putrescine and cadaverine (Figure 1.1) are generated through the microbial degradation of proteins and give the pungent foul smell to rotten meat, chemicals that we associate with disease-ridden food. A further example are the bitter-tasting alkaloids produced by plants, such as the lethal poison atropine (Figure 1.1) from deadly nightshade, which acts as an indicator of the plant's toxicity and makes it unpalatable. This association between chemical perception and danger found originally in nature is being utilized in our modern synthetic world. For example, small sulphur volatiles such as *tert*-butyl thiol (Figure 1.1) are purposely infused with odourless natural gas as humans can detect this chemical at less than one part per billion levels and instantly associate it with the dangers of a gas leak.

Flavours and fragrances (F&Fs) can however also be pleasant and act as attractants. This is elegantly illustrated in insects, organisms that utilize aroma chemicals as a means of communication. Ipsdienol and ipsenol (Figure 1.1) are sex pheromones of the male bark beetle and trace levels, too low for human detection, have been shown to attract female beetles from afar[1]. The male ingests the terpene myrcene from the phloem of the host tree *Pinus ponderosa* and through a number of biotransformations is able to convert it into aromatic pheromones[2]. Plants also utilize aroma chemicals to interact with their environment.

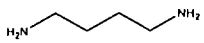
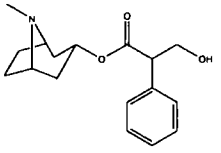
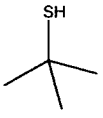
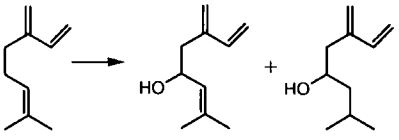
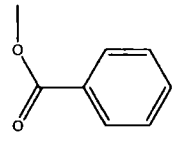
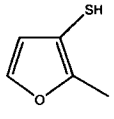
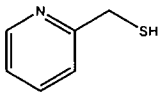
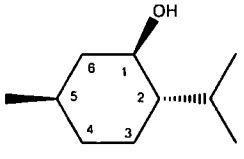
Aroma Chemical	Occurrence	Description
 Putrescine  Cadaverine	Rotten meat, urine and sweat	These toxic diamines are foul-smelling indicators of rotten diseased foods[8].
 Atropine	*Atropa belladonna* (deadly nightshade)	*Atropa* is a member of the Solanaceae family including tomato, potato and tobacco. All wild species are known to contain similar bitter tasting phenolic toxins[6].
 *tert*-butyl thiol	Synthetic chemical added to natural gas	Infused with natural gas to give it odour. Sulphur volatiles are extremely potent and detectable at low concentrations. Hence why they are referred to as "high-impact" aromas[7].
 Myrcene Ipsdienol Ipsenol	*Ips paraconfusus* (male bark beetle)	Sex pheromones generated by the biotransformation of the terpene myrcene ingested from the phloem of the host tree[2].
 Methyl benzoate	*Antirrhinum australe (snapdragon)*	The biosynthesis of this phenylpropanoid plant secondary metabolite is developmentally regulated so as to only attract insect pollinators to mature flowers[3].
 2-methyl-3-furanthiol  2-pyridine methanethiol	Maillard reaction products: generated upon roasting beef and lamb respectively	Cyclic structures containing sulphur, nitrogen and oxygen produced from sugars and amino acids upon heating foods. Sulphur volatiles are important roasted notes[4].
 Menthol	*Mentha piperita* (peppermint)	A chemical that evokes a cooling sensation, a bitter taste and a minty smell. Natural menthol exists as a single enantiomer (1R,2S,5R), highlighting the importance of chirality for the sensory properties of chemicals[12].

Figure 1.1 Structures of a number of representative aroma chemicals referred to in the text. The chemical diversity of flavouring compounds is great and is apparent here even through this small number of examples.

Phenolic volatiles emitted from flowers serve to attract insect pollinators. For example the volatile ester methyl benzoate (Figure 1.1) from snapdragon flowers is biosynthesised in specific regions of the petal where pollinators (bumblebees) come into contact with the mature flower[3]. The plant also regulates the biosynthesis of attractants in a rhythmic manner to maximise emission during the day, which coincides with pollinator activity.

Aromatic chemicals are also generated by microbes, such as yeasts which are used in fermentation processes in the cheese and brewing industries. In addition, flavourings are generated through thermal reactions, whereby the high temperatures and pressures associated with cooking generate a unique set of flavours[4] (Figure 1.1). Today many food, perfume and cosmetic products are artificially flavoured with synthetic chemicals, a result of the F&F industry that utilizes chemical processes, catalysts and cheap feedstock to recreate natural flavours.

In nature, the major producers of natural F&Fs are plants. This results from their diverse secondary metabolic pathways that generate a vast array of unusual chemical metabolites, a considerable subset of which have aroma properties[5]. Both floral and vegetative tissues are well known to sequester or emit aroma constituents which include alcohols, aldehydes, esters, terpenes, phenolics as well as sulphur- and nitrogen-containing compounds[6]. Originally, such chemicals were believed to be by-products of metabolism, or a result of overflow from central pathways. However, it is now believed they serve specific physiological roles in plants; attracting insect pollinators to flowers, promoting seed dispersal in animals attracted to fruit or detering insect and animal herbivores.

Humans (and animals) detect the presence of aromatic chemicals through receptor proteins embedded in specialised epithelial cells of the mouth and nose which pass electrical signals through the trigeminal nerve to the brain[7]. The human sense of smell is far more developed than that of taste, being able to recognise and distinguish between thousands of chemicals present at low concentrations (0.001 parts per billion). Conversely, only five major taste

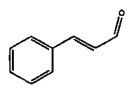
sensations can be perceived; bitter, sweet, salty, sour and savoury (umami). Bitter tastes are detectable at the lowest threshold of around 1 part per million, a likely result of the evolutionary importance of being able to detect toxins. Aromatic chemicals interact with different receptors, causing overlap between the senses. For example menthol (Figure 1.1) found in leaves of peppermint (*Mentha spicata*) evokes a cooling sensation, a bitter taste and a minty smell.

As described, F&F principles have unique functions in nature and the amazing ability to evoke memory and change behaviour. The importance of flavour to human life is apparent through the establishment of a multi billion pound global industry that produces F&F fine chemicals[8].

1.1.1 The Flavour and Fragrance Industry

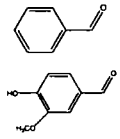
Plant natural extracts have long been used as flavouring ingredients, with the earliest written records referring to the use of herbs and spices in food[8]. However, it was not until the rise of the Roman Empire in the 1st century BC when methods for extraction such as distillation were developed and trade routes formed that the industrial production of F&Fs was established. Plant extracts continue to be important flavouring materials but something of a revolution occurred in the mid 19th century with the development of the chemicals industry and the introduction of synthetic materials.

Chemists were able to purify aroma constituents from essential oils and identify active ingredients, such as benzaldehyde from almond oil, cinnamaldehyde from cinnamon bark and vanillin from vanilla beans (Figure 1.2). Due to rapid developments in organic chemistry during this time it meant such natural products could be synthetically recreated in the lab. The first perfume containing a synthetic aroma chemical was Houbigant's Fougere Royale which went on sale in 1881 and contained synthetic coumarin[9]. This process of identification, lab synthesis and scale up is essentially the same today, only due to technological advances such as the introduction of gas chromatography, mass spectrometry and sensitive methods for trapping volatiles, such as Solid Phase Micro Extraction (SPME), it has been speeded up considerably (Figure 1.2).



Identification of cinnamaldehyde - the first flavour compound to be identified **1834**

Identification of benzaldehyde **1837**

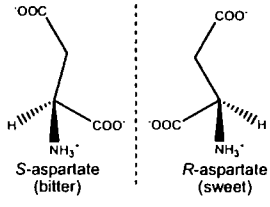


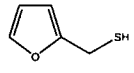
Identification of vanillin **1871**

Synthesis of vanillin - marking the beginning of the modern F&F industry **1872**

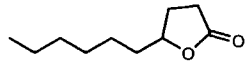
Synthesis of cinnamaldehyde **1884**

Enantiomers of aspartate found to have different tastes **1887**



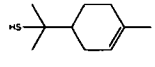


furfuryl mercaptan identified in coffee - highlighting the importance of trace sulphur constituents in food flavour **1926**



Bioproduction of γ-decalactone (peach) **1980**

Grapefruit mercaptan identified - 1-p-menthene-8-thiol **1982**



1820 Foundation of the first flavourings company: Roure Bertrand, France

1881 1st perfume to contain a synthetic: Houbigant's Royale

1888 1st determination of an odour threshold. Ethyl mercaptan.

1896 Foundation of Givuadan, Switzerland

1912 Maillard reaction described

1921 Chanel No. 5 by Coco Chanel

1939 Lavoslav Ruzicka wins nobel prize for work on sex pheromones

1952 Gas Chromatography

1959 International Flavors and Fragrances founded, USA

1964 GC-Olfactometry introduced

1989 Solid Phase Micro Extraction (SPME)

1991 Merger of Givuadan and Roure

Figure 1.2 Important milestones in the development of the flavourings industry (right) and the identification of important aroma chemicals (left). Vanillin is the worlds most important flavouring in terms of production (12,000 tonnes/ year), second is benzaldehyde. Many sulphur volatiles are important flavourings but most have only been identified over the past 30 years due to sensitive sulphur-GC detectors. Figure adapted from Rowe, D.J.[8], 2005.

Much of the research into F&F has been carried out by the industry and driven by commercial interest, as a result the aroma constituents from different natural sources have been isolated and characterised, with a total of more than 7,000 documented to date (Flavourings industry database; www.vcf-online.nl). The industry is worth £9 billion UK (F&F industry consultants: www.leffingwell.com) and is dominated by five major companies: Givuadan and Firmenich (Switzerland), IFF (USA), Symrise (Germany) and Takasago (Japan). In addition, a collection of smaller companies manufacture specialised flavourings, such as Oxford Chemicals Limited (UK) that produce high impact sulphur-containing ingredients.

The industry was established and has grown to the size it is today as a result of consumer demand for exciting flavour-rich products. With eighty percent of sales in developed countries, there is a clear correlation between the affluence of society and the demand for flavourings[8]. Although flavourings have improved the quality of life for many people in the developed world, their inappropriate use can be harmful. The dissociation of flavours from their natural nutritional source can lead to undesirable health consequences. For example, snack companies often artificially flavour food products, attracting consumers to fatty unnutritional foods they would otherwise find unpalatable, a likely contributor to the high levels of obesity in developed countries[10,11].

An increasingly health conscious society is concerned over what is in its food. The media have highlighted the dangers of certain constituents, such as salt, monosodium glutamate, acrylamide, toxic metals, saturated fats and genetically modified products. Reports are however often contradictory, and with the labelling of food products often hard to interpret and misleading, consumers are unsurprisingly confused and worried. Natural unprocessed foods are now the preferred choice of many.

These demands have fed down from the food industry to the flavour houses and ultimately to the F&F producers and as a result the global demand for natural products is on the rise[12]. Other factors such as legislation on food additives and the pressure to develop environmentally-friendly chemical

processes will also have contributed toward this trend. As a result the industry has turned back to nature to source F&F chemicals from plants. However, this dependence on agriculture means yields and quality are unpredictable and very much dependent on climate, consequently the value of such products are somewhat unpredictable. In addition, the availability of land for growing aromatic plants is heavily contested, using the land for growing food for animals and humans or for biofuel feedstock is likely to give greater returns[13]. Contamination of extracts with pesticide residues is another major problem and renders the product worthless[14]. In addition the isolation of single flavouring chemicals from a mixture requires costly methods such as fractional distillation and is often not cost effective when an active ingredient is found at trace levels. Hence, the range of flavours derived from plant extracts is limited.

The high demand for natural flavours in combination with costs associated with their isolation gives them high price tags as premium chemicals. Hence, the difference in price between synthetic and natural products is substantial. For example, synthetic vanillin is worth U.S. \$12 kg^{-1} as opposed to natural vanillin extracted from vanilla pods which has a value of U.S. \$4,000 kg^{-1} [15]. The use of plant and microbial enzymes for bioproduction is now seen by the industry as an attractive means of producing such valuable naturals and overcoming the difficulties associated with their extraction from plants.

1.2 Flavour and Fragrance Principles in Exotic Fruit

One group of aroma chemicals of particular interest to the industry are the tropical flavours of exotic fruits. Tropical fruits are native to equatorial regions where they develop in warm tropical climates and are characteristically spectacular in colour and form and also rich in unusual flavours and aromas. They form an important resource for trade in developing countries, accounting for 98% of global production, with developed countries consuming 80% in the form of imports[16]. In terms of production, mango forms the major tropical fruit crop followed by pineapple, papaya and avocado, with the trade in lychees, rambutan, guavas and passion fruit smaller but increasing[16].

The food and drinks industry is witnessing a substantial growth in demand for tropical fruits and flavours[17]. This could be a result of the cultural diversity in our cities or the more adventurous modern consumer who enjoys the exciting flavours found in foreign foods. In addition, concerns over diet and health may have also contributed towards this trend. Although more and more unusual foreign fruits are appearing on supermarket shelves these are only a small proportion of those grown in the tropics; most cannot survive the journey to higher latitudes as they over-ripen and spoil making their export impractical.

The flavourings industry has acted on the demand for tropical flavours and has artificially synthesized a large number of "nature identical" tropical compounds. However, with the industry well aware of the consumer demand for natural products, aromatic oils and juice extracts are also becoming big business. As has been described this can be problematic, climate in the tropics is particularly variable, floods and hurricanes can destroy crops making fruit supply unreliable. Other difficulties include trade restrictions and political instabilities. In addition, the chemicals responsible for tropical notes are present at trace levels within the fruit and as a result they are nearly impossible to isolate in pure form from natural materials, hence giving them high price tags in the flavourings industry. As a result the biotechnological production of such single aroma metabolites is a far cheaper and more attractive alternative to extraction from natural sources.

1.2.2 The Role of Aroma Chemicals in Fruit

Exotic fruit plants would have originally been found growing in dense tropical jungles competing for light, nutrients and space. The single purpose of the fruit is to aid in seed dispersal, something that gives the new plant the greatest chance of survival. The elegant colours, forms and flavours of the fruit are all adaptations to serve this purpose; they are attractants for specific species of seed-dispersing jungle animals and birds.

Flavours in fruit are derived from an array of nutrients, including amino acids, sugars and fats. Hence, it has been proposed that F&Fs in fruit act as indicators of their nutritional value[18]. The fruit offers a rich source of

carbohydrate, fat and water and in return animals will transport the plants seed long distances within their gut. Seed dispersal is a crucial stage in the life cycle of these tropical plants, and when the parent plant passes its genes on to the next generation it invests heavily to ensure the survival of its offspring. Although avoiding competition for resources may be important, the most likely reason for dispersing seed is to move the new plant away from diseases that can often plague a group of related plants growing in close proximity[19]. Passing the seed through the digestive system of an animal has the added advantage of killing any insect-borne diseases, allowing the seed to germinate and start a new life disease free in a rich mound of fertiliser.

The competition for animals (vectors) is the driving force behind the evolution of variability in fruit flavours, something that has led to differences in the fruit flavour compositions between species. A plant that accumulates mutations and evolves a new flavour profile runs the risk of not being as attractive as its neighbours, its genes will not be selected and eventually it will die out. However through mutations, such as gene duplication and divergence or silencing, new flavour profiles more attractive to animals could evolve. Hence, this battle to be the most "attractive" is carried out at the level of biochemical pathways with flavour production and diversification important in the evolution of tropical fruit.

In addition, many plant secondary metabolites play an important role in defence against insect pests[20], with these signalling compounds often found to be volatile and aromatic. There is a battle between plants to produce insecticidal toxins and insects to develop resistance to them, and as a consequence this drives the evolution of different secondary metabolites and ultimately the evolution of different flavour volatiles.

1.2.3 Flavour and Fragrance Principles in Tropical Fruit

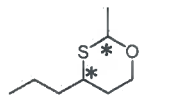
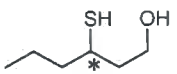
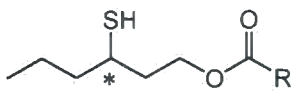
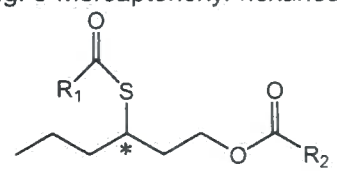
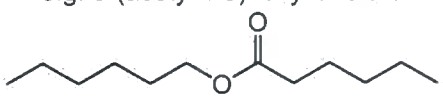
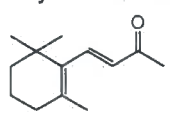
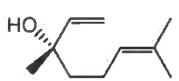
The aroma contributing volatiles in tropical fruits are of particular interest to the flavourings industry. Due to the increase in demand for tropical notes the F&F industry set out to discover what were the odour-active chemicals and what was so unique about these fruits that gave them such unusual exotic notes. In

answering these questions, the volatile constituents of exotic fruits were extensively studied and found like other fruits to contain predominantly esters and significant levels of alcohols, aldehydes, terpenes, lactones and amino acid derivatives[21]. However, the exotic notes remained elusive until sensitive GC detectors were developed and an array of low abundance (<10 ppb levels) sulphur constituents were discovered. These were later found to be responsible for the true "tropical" flavour in these fruits. The character impact volatiles (chemicals responsible for characteristic fruit flavour) from a number of important tropical fruit varieties are illustrated in figure 1.3.

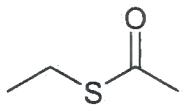
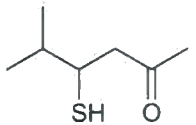
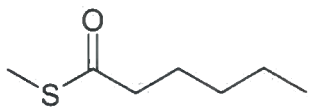
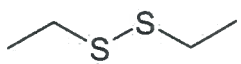
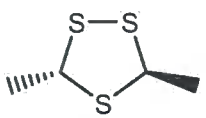
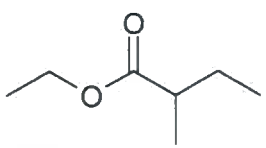
Due to its delicate tropical flavour and its value in the fruit processing industry the yellow passion fruit (*Passiflora edulis* f. *flavicarpa*) has been extensively studied. The fruit was found to contain some 270 volatile compounds (www.vcf-online.nl). However, not all of these contribute to flavour, in fact less than a quarter (24%) have aroma character at levels present in the fruit[22]. The initial identification of volatile chemicals found that esters comprised 95% of the volatile oil, with hexylhexanoate (Figure 1.3) the principal constituent[23]. The chemicals responsible for the attractive tropical notes were however not identified until the development of flame photometric GC detectors, which identified a number of character impact sulphur volatiles[24,25], the most important being 2-methyl-4-propyl-1,3-oxathaine and 3-mercaptohexanol, with ester and thioester derivatives of the latter also forming important odour volatiles (Figure 1.3). As with many natural aromas, chirality was found to be important for aroma character[26,27](Figure 1.3); only the correct optical isomers will interact with the active sites on receptor proteins.

The Durian fruit is highly prized by the people of South East Asia but due to its powerful odour it is repulsive to most Europeans. Volatile flavouring constituents from this fruit were first identified in 1971[30], and again sulphur compounds were found to have high odour impact. In this case the sulphur constituents were small and emitted from the fruit, with thioesters, disulsphides and sulphides all found to contribute to odour[31], 3,5-dimethyl-1,2,4,-trithiolane (Figure 1.3) was found to have the highest odour impact [32].

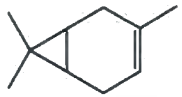
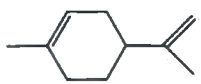
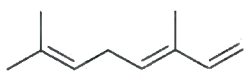
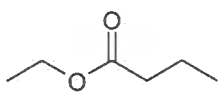
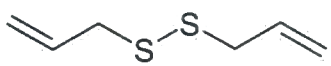
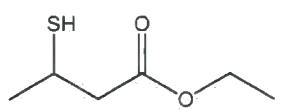
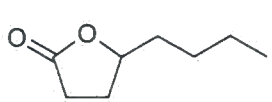
Figure 1.3 Character impact volatiles in tropical fruit

Fruit Plant	Character Impact Compound/ Chemical structure	Description
 *Passiflora edulis* F. *flavicarpa* (yellow passion fruit) Purple fruit are preferred fresh whereas the yellow is used for juice pressing and preserves. A characteristic and appealing tropical flavour. The purple passion fruit is native to Southern Brazil. The yellow is more disease resistant and thought to be a hybrid variety of unknown origin.	 2-methyl-4-propyl-1,3-oxathiane	Said to be the key character impact compound in passion fruit, commonly called tropathiane. 2*R*,4*S* is found to be the natural configuration. Sulphurous and green with a slight fruity, roasty note.
	 3-mercapto-1-hexanol	A juicy, tropical fruit aroma reminiscent of grapefruit, blackcurrant and mango. Commonly called passion fruit mercaptan. The *S* configuration is most abundant in nature (58-81%).
	 e.g. 3-mercaptohexyl-hexanoate	R= CH_3, C_3H_7, C_5H_{11}. Tropical grapefruit, blackcurrant and passion fruit like note. The *S* configuration is most abundant in nature (96%).
	 e.g. 3-(acetylthio)hexyl acetate	Thioesters are important aroma constituents in a number of tropical fruit. R_1/R_2= CH_3, C_3H_7, C_5H_{11}. Fatty, fruity, sweet mango, passion fruit and durian like. The *S* configuration is most abundant in nature (90-98%)
	 hexylhexanoate	4 esters were found to make up 95% of volatiles from passion fruit oil; ethyl butyrate, ethyl hexanoate, hexyl butyrate and hexyl hexanoate, the latter being the major constituent. A Fruity note.
	 β-ionone	Found in very low concentrations but with a low odour threshold it will contribute to passion fruit flavour. Responsible for the rose-like notes.
* Chiral centre	 linalool	The major terpene component of passion fruit aroma. Found in many fruit types contributing a fruity citrus note.

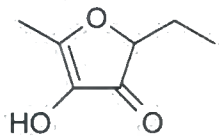
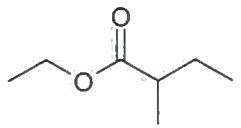
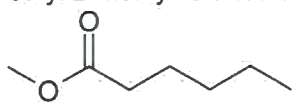
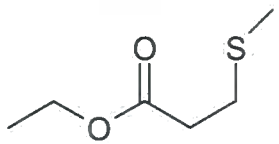
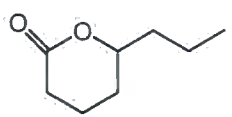
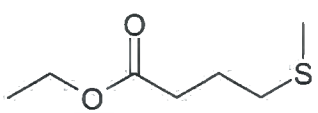
Refs: (Jordan *et al.*, 2002)[22], (Parliment, T.H., 1972)[23], (Engel *et al.*, 1991)[24], (Werkhoff *et al.*, 1998)[25], (Singer *et al.*, 1986)[26], (weber *et al.*, 1995)[27], (Brat *et al.*, 2000)[28], (Chen *et al.*, 1982)[29].

Fruit Plant	Character Impact Compound/ Chemical structure	Description
 *Durio zibethinus murr.* (durian fruit) The flavour is unpleasant to most that did not grow up in South East Asia. It has 2 characteristic notes; strong sulphury onion-like and delicately fruity. Native to Borneo and Sumatra. Cultivated in South Eastern India, Thailand and Burma where they are prized in the royal palace.	 S-ethyl thioacetate	This sulphur compound is in highest concentration in durian. The thioesters are important sulphur constituents in durian with fruity and sulphury notes.
	 5-methyl-4-mercapto-2-hexanone	A range of sulphur volatiles are present in durian; thioesters, disulphides, sulphides and trithiolanes but only one mercaptan (5-methyl-4-mercapto-2-hexanone) has been identified. It is believed they are oxidized during extraction.
	 methyl thiohexanoate	This compound has an intensely sweet, ripe odour with tropical and soft-fruit character
	 ethyl disulphide	A number of ethyl sulphides such as ethyl disulphide contribute to the sulphurous gassy note given off by ripe durian fruit.
	 *trans*-3,5-dimethyl-1,2,4-trithiolane	The trithiolanes are major sulphur components in durian. Interestingly *cis* and *trans* isomers of this compound were found to be present in equal amounts, suggesting a non-enzymatic origin.
	 ethyl 2-methylbutanoate	In contrast to the sulphurous gassy notes durian is also said to have a delicate fruity taste. This is predominantly due to the branched-chain ester, ethyl 2-methylbutanoate, found in many fruit species.

Refs: (Baldry *et al.*, 1972)[30], (Moser *et al.*, 1980)[31], (Weenen *et al.*, 1996)[32].

Fruit Plant	Character Impact Compound/ Chemical structure	Description
 *Mangifera indica* L. (mango) A member of the plant family Anacardiaceae, containing many poisonous varieties. Under the waxy red/ green/ yellow peel is a fibrous, yellow sweet flesh with a pleasant aroma. Native to Southern Asia.	 δ-3-carene	Terpene volatiles are by far the predominant group of volatiles in mango. In particular δ-3-carene is found to be the major compound in the majority of mango varieties. Sweet flavour similar to limonene.
	 limonene	50 terepenes have been identified in mango; 19 monoterpenes and 31 sesquiterpenes. Limonene is an important monoterpene contributing to this aroma with a lemon/fruity note
	 (Z)-β-ocimene	Responsible for the green aroma in raw mangos.
	 ethyl-butanoate	The esters are another important class of volatiles contributing to mango flavour. Ethyl-butanoate was found to be dominant in most varieties. This gives the fruity flavour to mangos.
	 diallyl disulphide	Sulphur volatiles are also found to be important contributors to aroma in mango. Diallyl disulphide, trisulphide and benzyl isothiocyanate are of particular note.
	 ethyl 3-mercaptobutyrate	Only recently identified due to sensitive sulphur-GC detectors, present at only trace levels but nonetheless an important contributor to aroma character in mango.
	 γ-octalactone	The lactones are also important fruit flavours derived from fatty acid metabolism. In mango γ-octalactone is the major form and found well above its odour threshold of 7 ppb. It has a sweet, herbaceous, coconut like aroma and is thought to be important in the ripe odour.

Refs: (Pino *et al.*, 2005)[33], (Dewis and Kendrick, 2002)[34].

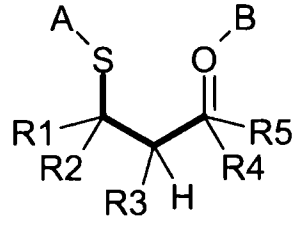
Fruit Plant	Character Impact Compound/ Chemical structure	Description
 *Ananas comosus* (pineapple) A multiple fruit made up from mature ovaries of many flowers. Juicy tropical flesh. Native to southern Brazil and Paraguay. Domesticated by the Indians and carried up to central America and the West Indies.	 2,5-dimethyl-4-hydroxy-3-[2H]-furanone	The key character impact compound in pineapple, commonly called pineapple furanone. A strong pineapple note.
	 ethyl 2-methylbutanoate	The second largest odour contributor to pineapple after furaneol. A fruity estery note, found in many fruits often as a key character impact compound.
	 methyl hexanoate	Another major ester of the pineapple, with distinctive pineapple aroma and a sweet taste. Both the methyl and ethyl esters are found in pineapple. Note the abbundance of C6 straight-chain volatiles in fruit.
	 ethyl 3-(methylthio)propanoate	Methyl and ethyl thiopropanoates are major sulphur constituents found in pineapple, especially in under ripe fruit. They are both known to have characteristic pineapple notes.
	 δ-octalactone	A number of reports of glycosidically bound volatiles have been made. One of the major groups of glycosidically bound volatiles in pineapple are the lactones, in particular δ-octalactone . A sweet, tropical coconut, dairy note.
	 ethyl 4-(methylthio)butanoate	In its synthetic form this chemical is known to enhance pineapple character in flavourings and perfumes. However, it has not been identified in pineapple to date and hence cannot be sold as a nature-identical ingredient, it must be labelled as an artificial flavouring.

Refs: (Umano *et al.*, 1992)[35], (Rodin *et al.*, 1965)[36], (Wu *et al.*, 1991)[37], (Takeoka *et al.*, 1991)[38], (Berger *et al.*, 1985)[39].

Mango is the major tropical fruit crop in terms of production and trade. For this reason the volatile constituents have been thoroughly characterised, with more than 300 identified to date (www.vcf-online.nl). Terpene hydrocarbons were found to be the predominant volatiles, with some 50 types identified[33], δ-3-carene (Figure 1.3) forms the major character impact terpene in most varieties. Mango is also found to contain a number of sulphur aroma constituents, namely diallyl disulphide and ethyl 3-mercaptobutyrate, the latter only recently discovered but found to be an important contributor to overall exotic flavour[34].

Pineapple is another major tropical fruit variety and has therefore attracted attention, with over 200 volatiles identified[35]. Of particular mention are pineapple furanone[36], esters (ethyl 2-methylbutanoate and methyl hexanoate), δ-octalactone[37] and sulphur volatiles (methyl and ethyl thiopropanoates)[38].

The British flavourings company Oxford Chemicals, who specialise in the manufacture of high impact flavourings, realised that it was the position of sulphur with respect to oxygen within volatiles that produced the tropical note and proposed the structure of the tropical olfactophore[40] (Figure 1.4). This is a chemical structure that must have both the correct conformation and vibrational bond energies to trigger a nerve impulse from receptor proteins in the nose and mouth which we perceive as "exotic" or "tropical". Tropical fruit have therefore utilized sulphur and its remarkable chemistry for the generation of unique aroma profiles. They only need invest in the synthesis of small quantities of such compounds for them to have a huge impact on fruit aroma, hence saving valuable nutrients and metabolic energy. It is the ability of exotic fruits to assimilate sulphur and incorporate it into volatiles that makes them unique.



A = H, CH_3, acyl, ring
B = H, CH_3, acyl, absent if carbonyl, ring
R1, R2 = H, alkyl
R3 = H, alkyl, ring
R4 = H, CH_3, ring
R5 = H, absent if carbonyl

Figure 1.4 The Tropical Olfactophore[40]

1.3 Aroma Generation in Tropical Fruit

1.3.1 The Biosynthesis of Aroma in Fruit

The early research into fruit flavour concentrated on identifying the F&F principles and assessing their contribution to overall fruit aroma. However, with the technological developments in molecular biology and genetics over the past thirty years research is now also focussing on identifying the biochemical pathways of flavour generation. An understanding of fruit flavour biosynthesis will enable directed plant breeding programmes and the genetic engineering of fruit plants to improve flavour quality, something that has deteriorated through intensive fruit production. The aroma chemicals in fruit are predominantly generated through three major biochemical pathways; fatty acid degradation, terpene metabolism and amino acid metabolism, which are reviewed by Morton and MacLeod[41], 1990, and Parliment[42], 1986. Much of this work has focussed on a number of model fruits, predominantly apple (*Malus pumila*), tomato (*Lycopersicon esculentum*), melon (*Cucumbis melo*), strawberry (*Fragaria ananassa*), banana (*Musa sapientum*) and to a lesser extent mango (*Mangifera indica*). Aroma generation in tropical fruit will undoubtedly share many of these conserved pathways, whereby sulphur volatiles are generated as derivatives of these core aroma metabolites. Hence, their biosynthesis must be considered first.

The methodology for elucidating aroma biogenesis initially focussed on identifying precursors and intermediates and studying their accumulation and depletion during the ripening process[43,44], with feeding studies and radiolabelling confirming many of these proposed pathways[45,46]. A number of genes predicted to be involved in aroma generation have been isolated and cloned through methods such as Reverse Transcription Polymerase Chain Reaction (RT-PCR), with *in vitro* characterisation of such heterologously expressed enzymes determining substrate and product specificity[47,48]. Genomics approaches have also proved successful, with both random and targeted screens of fruit cDNA libraries identifying further genes and enzymes of interest[49], such as those identified in papaya[50]. Microarray analysis in apple[51] and strawberry[52] has identified genes and enzymes that are

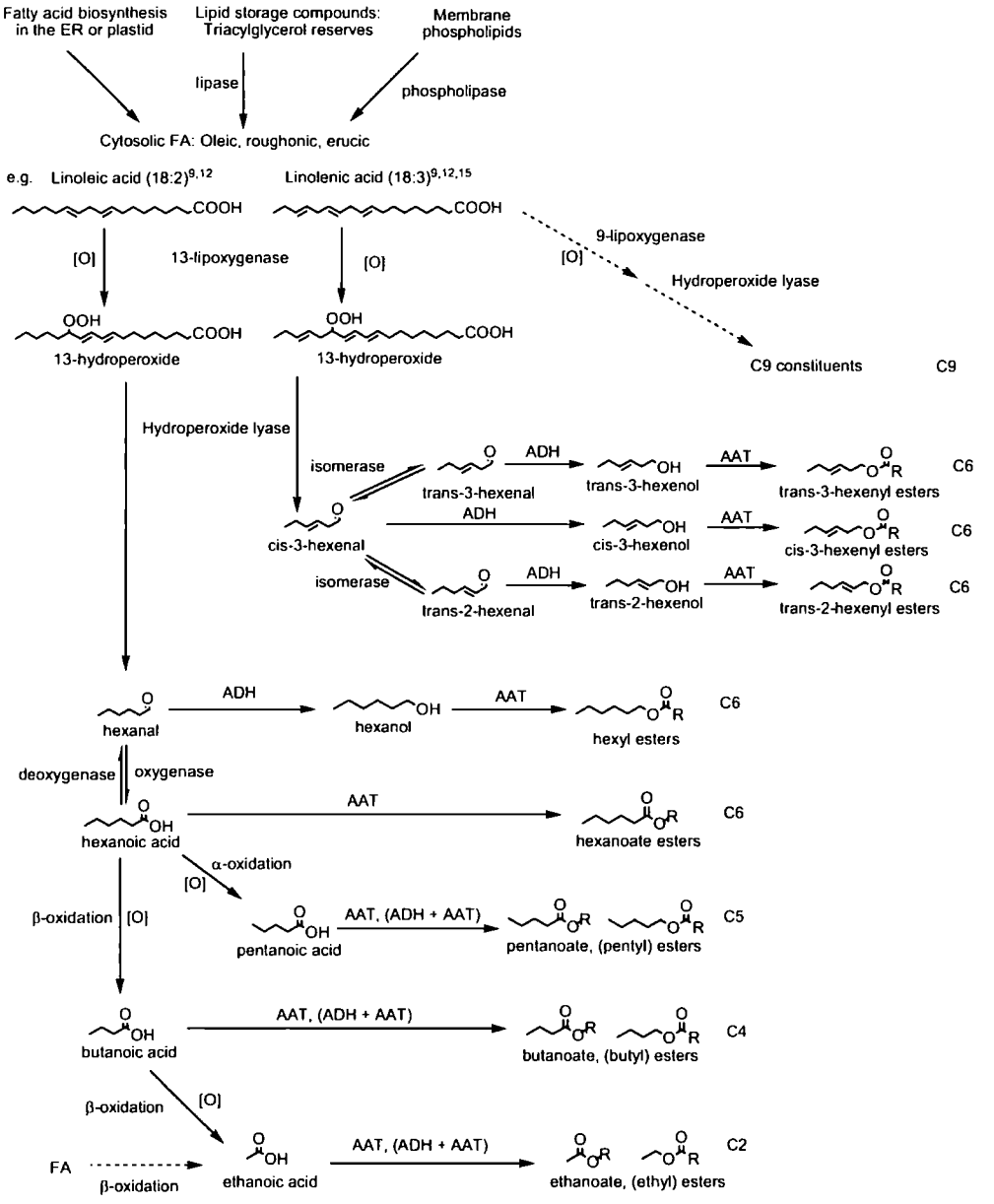
upregulated during ripening, a considerable subset of which have been shown to be involved in aroma biogenesis. Collectively such studies have provided a sound understanding of the core pathways of aroma biosynthesis in fruit.

1.3.1.1 Fatty Acids as Precursors to Straight Chain Alcohols, Aldehydes and Esters

Fatty acids (FAs) and lipids are important structural and metabolic constituents of plant cells, forming cell membranes (phospholipids) and also serving as energy storage compounds (triacylglycerols)[44]. However, with the observation that many fruit volatiles were composed of even numbered straight-chain organic structures, and with the high proportion of C6 aroma metabolites in fruits, FAs (being even numbered carbon compounds themselves) were chosen as primary candidates as precursors to these aroma constituents[41].

Climacteric fruits, including mango and apple, continue to ripen when picked from the tree: they show a characteristic increase in respiration, ethylene release, protein synthesis and metabolic activity during maturation (the climacteric respiratory rise) which results in the softening of the fruit, colour changes and aroma generation. They can be picked from the tree and studied under controlled conditions, and as the stage of ripening can be calculated from levels of ethylene production, they serve as useful models for studying aroma biosynthesis. Such work on climacteric species revealed an accumulation of lipids and free FAs in developing fruit, especially during the respiratory rise. In mango, levels of linoleic, linolenic and stearic acids accumulated during the rise and then decreased following the respiratory peak, due to their metabolism into further products[43], with ester biosynthesis correlating with this rise in free FAs. A similar trend was observed in apple[44]. In addition it was demonstrated that feeding of FA to homogenates of apple[53] and tomato[54] could increase the production of straight chain aroma volatiles. Radiolabelling studies confirmed deuterated FAs were metabolised to straight chain aldehydes, alcohols, acids and esters in apple[46]. In addition, this study demonstrated that 18:3 FAs, such as linolenic acid, were precursors to unsaturated C6 volatiles (hexenal and hexenol) whereas 18:2 FAs, such as linoleic acid, were precursors to saturated constituents (hexanal and hexanol).

A.



B.

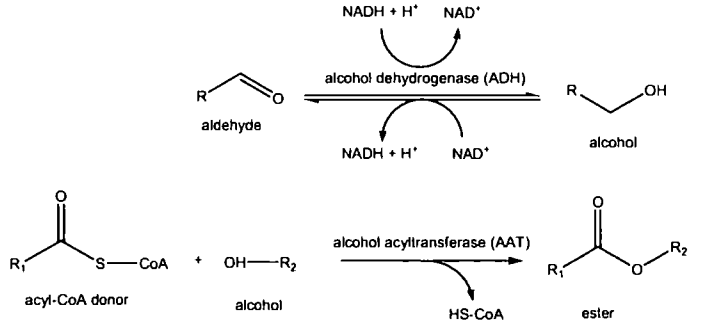


Figure 1.5 A. Biosynthetic pathways leading to straight-chain aldehydes, alcohols and esters in fruit. **B.** The reactions catalysed by alcoholdehydrogenases (ADH) and alcohol acyltransferases (AAT).

FAs are metabolised into short chain alcohols, aldehydes and acids (prior to ester formation) through oxidative degradation; β-oxidation and/or lipoxygenase cleavage[55] (Figure 1.5A). FA degradation through β-oxidation occurs in peroxisomes and involves the repeated cleavage of C2 units (acetyl co-A) from the carboxyl end of fatty acyl-CoAs. This is believed to be responsible for the generation of acetic, butanoic and hexanoic acids, which may then be reduced to the corresponding aldehydes and alcohols before esterification[46]. Odd numbered carbon structures are a result of α-oxidation generating the C5 constituents found in a number of fruit[46]. The high proportion of C6 volatiles, such as hexanal and hexenal are believed to be derived from the lipoxygenase (LOX) pathway[55].

Cytosolic lipoxygenases catalyse the oxidation of unsaturated FA at specific points along the carbon chain, generating hydroperoxides. Linoleic or linolenic acids are the major unsaturated FA in plants, and they are oxidised by LOX enzymes at the 12 or 9 positions and subsequently cleaved by hydroperoxide lyases to generate C6 or C9 aldehydes respectively, which may be oxidised to acids or reduced to alcohols prior to esterification[46] (Figure 1.5A).

The regulation of these pathways is partly through transcriptional control over enzymes that metabolise FAs and those that generate esters, however, the major controlling factor is believed to be the availability of free FA precursors[44]. It was therefore proposed that the accumulation of free FAs could result from the biosynthesis (anabolism) of *de novo* FAs within the endoplasmic reticulum or through their release from phospholipid pools through the action of lipases or lipoxygenases.

It was observed that apples stored under anaerobic conditions only generated very low levels of free FA and subsequently did not develop a natural ripe aroma[53], this dependence upon oxygen and respiration indicates that the anabolic pathway is responsible for FA aroma generation. Furthermore, the dependence of free FA accumulation on adenoside triphosphate (ATP)

strongly reinforces the anabolism theory[56]. Hence, it is important to consider that fruit storage conditions will affect fruit maturation and aroma development.

The enzymes responsible for the interconversion of aldehydes and alcohols (Figure 1.5B) are alcohol dehydrogenases (ADH, EC 1.1.1.1) which have been identified in a number of fruits. Considerable work on their effects on aroma biogenesis has been carried out in tomato[57,58], where it was observed that ADH accumulated in ripe fruit coinciding with flavour volatile generation. In addition, a study using transgenic tomato lines with reduced or increased ADH activities were shown to affect fruit flavour[57], whereby reduced ADH activity resulted in an imbalance of aldehydes and alcohols (low levels of hexanol or hexenol) and as a result less diversity and abundance of esters. Hence, ADH activity is important in the reduction of short chain aldehydes to alcohols, a critical step in the generation of hydroxyl functionality required for subsequent ester formation.

Esters, the major aroma constituents in fruit, are generated through the linkage of acyl moieties from acyl co-enzyme A (CoA) to alcohols, a reaction catalysed by alcohol acyltransferases (AAT, EC 2.6.1.1) (Figure 1.5B). AATs have been studied in apple[59,60], banana[61], melon[62,63] and strawberry[52,64]. Such studies found that AATs generally show substrate specificities that correlate with the volatile esters present[64], although it is argued that substrate availability rather than enzyme specificity controls the types of esters formed[59].

Hence, FA metabolism generates important core aroma volatiles in fruit. The oxidative degradation of FAs initially results in an accumulation of short chain aldehydes which give the green notes in underipe fruit, with their subsequent metabolism into alcohols, acids and esters providing the fruity notes later in maturation. The types of FAs in fruit, and the specific enzymes of FA oxidation (LOX enzymes and hydroperoxide lyases) with different regio-specificities, account for the accumulation of different degraded FA metabolites in different fruits. In addition, the substrate specificities of ADHs and AATs will channel these degraded FA metabolites down further alternate pathways resulting in the accumulation of different end products, such as hexylhexanoate in passion

fruit. As a result of repeated rounds of duplication and divergence ADH and AAT gene families in fruits are large and diverse, which allows for the continual evolution of fruit flavour.

1.3.1.2 Terpene Metabolism

The terpenes (also called isoprenoids) are the most functionally and structurally diverse group of plant secondary metabolites[65] and form another important collection of aroma volatiles. They are classified according to the number of C5 isoprenoid units they contain and named hemiterpenes (C_5), monoterpenes (C_{10}), sesquiterpenes (C_{15}) and the diterpenes (C_{20}). They are highly aromatic and function as defence agents against insect pests and pathogens or as attractants towards pollinators and seed-dispersing animals. Their importance in tropical fruit flavour has already been demonstrated through their abundance in mango.

The pathways leading to the isoprenoids are well characterized, with the basic C5 units isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP) (Figure 1.6A) generated via two alternative pathways; the mevalonate pathway (cytosol) and the methylerythritol phosphate pathway (plastid), reviewed by Rodriguez-Concepcion and Boronat[65], 2002. However, the diversity of terpene metabolites is achieved through the action of terpene synthases which catalyse transformations and cyclisation reactions[66]. This is demonstrated through monoterpene biosynthesis where GPP is converted to various skeletal structures through the action of large families of monoterpene synthases. Limonene present in mango and linalool from passion fruit are common monoterpenes found throughout the plant kingdom and serve as useful examples (Figure 1.6A). One remarkable property of terpene synthases is their ability to take a single substrate (e.g. GPP) and convert it into numerous products[66]. Although the larger terpene (C15, C20) metabolites are non-volatile, their cleavage can generate important aroma constituents, such as β-ionone in passion fruit, formed through the cleavage of carotenoid compounds[67] (Figure 1.6A).

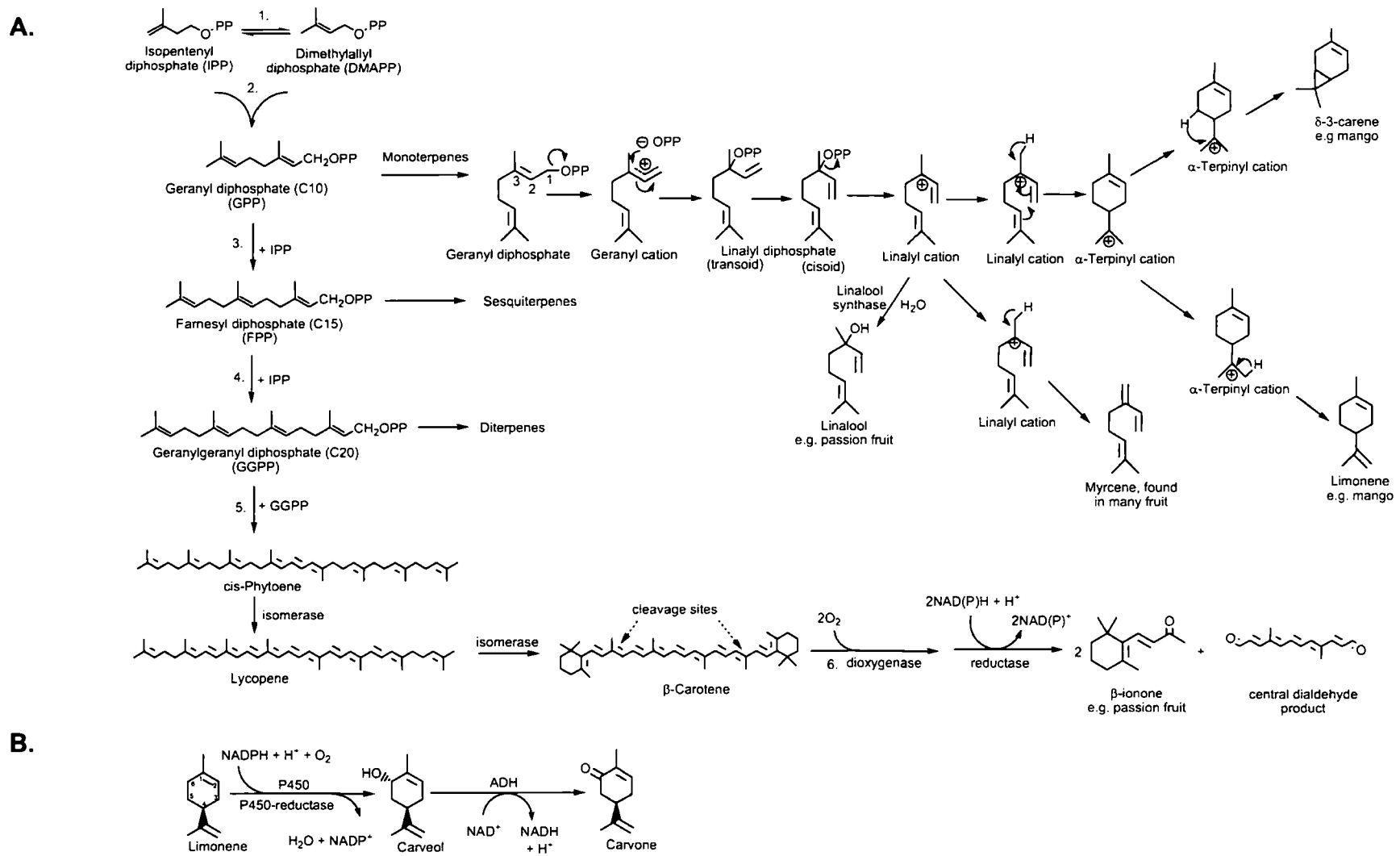


Figure 1.6 A. Terpene metabolic pathways leading to aroma chemicals. 1. IPP isomerase, 2. GPP synthase, 3. FPP synthase, 4. GGPP synthase, 5. tetraterpene synthase , 6. caratenoid dioxygenase. **B.** The biotransformation of limonene demonstrating the action of cytochrome P450 oxidases (P450) and NADP/NAD dependent oxidoreductase, alcohol dehydrogenase (ADH).

Further diversity still is created through secondary transformations. The large family of cytochrome P450 oxidases catalyse hydroxylation reactions[68] and are involved in numerous metabolic pathways in plants, where they exist as multiple isoforms able to hydroxylate diverse chemistries. They have attracted particular interest due to their involvement in the initial stages of detoxification of xenobiotics (foreign compounds) in plants and animals[68]. However, in plants they are also recruited to catalyse the many reactions of volatile aroma generation. Hydroxylation is substrate, regio-, and stereo-specific and generates a multitude of products from a single terpene structure. For example, in caraway (*Carum carvi*) fruit, limonene is hydroxylated at the 6 position to produce carveol[69] (Figure 1.6B). This is metabolised further through the action of NADP/ NAD-dependent oxidoreductases, such as ADH, which oxidises the hydroxyl functionality into a carbonyl group. Hence the aroma characteristic is changed through a single biotransformation to produce carvone, the major aroma in this fruit. Collectively, secondary transforming enzymes such as AATs, ADHs, P450s and terpene synthases create the great chemical diversity found in plant aroma constituents and hence maximise the odour impact of nutrients channelled away from central metabolism.

There is considerable interest in both terpene synthase genes and the secondary transforming genes for the genetic engineering of plants to improve flavour. Due to the importance of terpenoid constituents in essential oils, a number of terpene synthases have been isolated and characterised from these aromatic plants, including peppermint and spearmint[49]. Such studies hope to be the first steps in genetically engineering commercial essential oil plants to obtain improved essential oil yield and composition.

1.3.1.3 Amino Acid Derivatives- Branched Chain Aromas

Many fruit ester volatiles contain branched-chain structures, such as the 2-methyl butanoate esters found in durian and pineapple (Figure 1.3). The occurrence of such compounds can not be explained through their metabolism from straight-chain FA derivatives. The accepted route for their generation is through metabolism of aliphatic amino acids.

Much of the work in elucidating branched-chain ester biosynthesis has been carried out in apple[45], banana[70,71] and strawberry[72,73]. In strawberry levels of amino acids were shown to increase during the ripening period and then fall away as branched-chain ester concentrations increased late in ripening[72]. In addition, *in vitro* feeding of L-isoleucine to strawberry fruit led to a seven-fold increase in levels of 2-methylbutanoate esters and a two-fold increse in 2-methyl butyl esters compared to control fruit[73]. Similar observations were made in banana[70].

More recently, radiolabelling studies have confirmed such pathways. Incubation of ripe banana discs with radiolabelled L-leucine resulted in its biotranformation into branched-aroma volatiles[71], predominantly 3-methyl butanol (81%) and 3-methyl butyl esters (10%). Similar studies in apple using deuterium-labelled L-isoleucine confirmed its bioconversion into 2-methylbutyl and 2-methylbutanoate esters[45]. Collectively these studies have shown that the biosynthesis of 3-methylbutyl, 2-methylbutyl and 2-methylpropyl esters proceeds from the amino acids leucine, isoleucine and valine respectively (Figure 1.7). The first step is the deamination of precursors catalysed by aminotransferases which generate branched-chain oxo acids; their subsequent reduction yields branched-chain alcohols or acyl-CoA metabolites. These substrates contribute to the pool of alcohols and acyl-CoA compounds available for utilization by AATs in ester biosynthesis (Figure 1.7).

It is evident that the availability of nutrients will have a significant effect on fruit flavour: plants growing in poor soils will not be able to invest as much resourse in aroma production and consequently will not develop full aroma character. An evolutionary driver for plants is to generate maximum aroma impact to attract animals, with minimal loss of nutrients to themselves. Interestingly, tropical plants have achieved this through incorporating sulphur into metabolites. Perhaps the greater competition for resources in dense jungle environments puts greater pressure on these species to maximise the use of their valuable nutrients in F&F generation. Alternatively, the availability of sulphur may be greater in warmer equatorial regions.

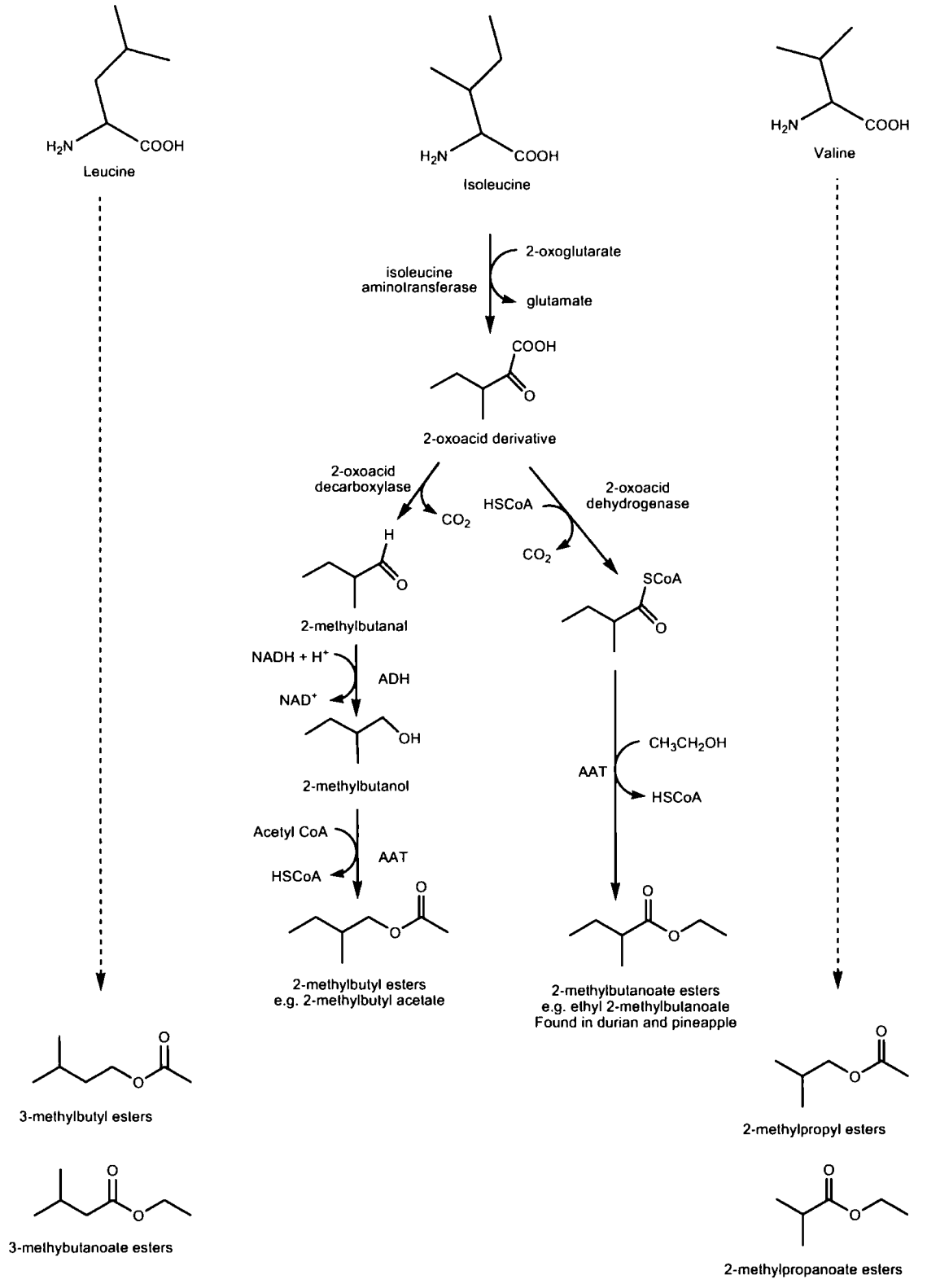


Figure 1.7 The biosynthesis of branched-chain aroma volatiles from amino acid precursors. Tropical plants evidently share these conserved pathways found in model fruit. Figure produced with information from Rowan *et al.*,[45] 1996 and Myers *et al.*,[71] 1970.

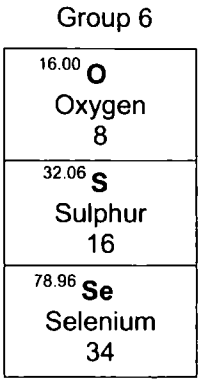
1.3.2 The Biosynthesis of High-Impact Flavourings in Tropical Fruit

In addition to the core aroma volatiles described above, tropical fruits also contain trace levels of sulphur constituents which impart the characteristic "tropical" notes[24]. The biosynthetic pathways leading to their generation have not been studied in such detail in tropical fruit. However, the identification of sulphur volatile intermediates, such as cysteine-conjugated aroma precursors found in passion fruit[74], have allowed possible pathways to be proposed.

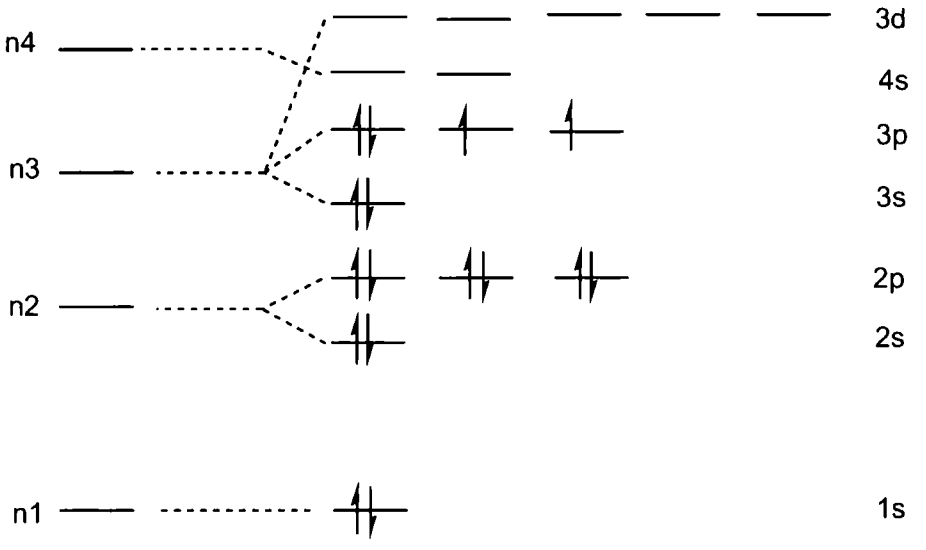
Sulphur is essential for life. The element is electronically very active and therefore generally has catalytic or electrochemical function in biomolecules. Situated in group six of the periodic table (Figure 1.8A), below oxygen and above selenium, its chemistry is often compared to that of oxygen-containing compounds. However, a number of important differences must be noted. Sulphur has an electronic configuration (Figure 1.8B) with four high energy 3p orbital electrons making it a strong nucleophile, a characteristic of the thiol group in cysteine. However, unlike oxygen, its highest occupied molecular orbitals (HOMO) are further from the nucleus and more diffuse, making it a soft nucleophile which favours attack at saturated carbon atoms (sp3 hybridized carbon centres) in S_N2 reactions, as opposed to oxygen that holds its HOMO electrons close to the nucleus, making it hard and favouring attack at carbonyl carbons (sp2 hybridized) in S_N1 reactions. It has an electronegativity comparable to that of carbon (Figure 1.8C) so it is not possible to describe its chemical reactions in terms of polarisation of carbon sulphur bonds. Sulphur also has empty p, s and d outer orbitals which allow it to have coordination numbers from 1 to 7 and variable oxidation states; 2,4,6. Hence, it is a versatile element found in a range of functional groups (Figure 1.8D).

The reason sulphur-containing volatiles are such potent aromas is something of a mystery. Humans and animals have evidently evolved to recognise such chemicals specifically, perhaps as a means of avoiding unpleasant or toxic foods, or conversely as attractants (as is the case in exotic fruits). Sulphur is an essential nutrient to animals and important in cellular protection and foreign compound detoxification, it is possible we are so sensitive to it because we need it in our diet and a deficiency could be fatal.

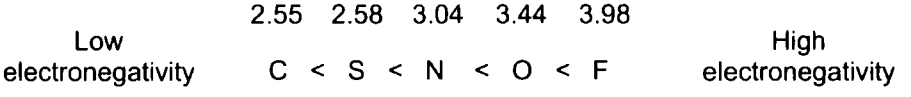
A.



B.



C.



D.

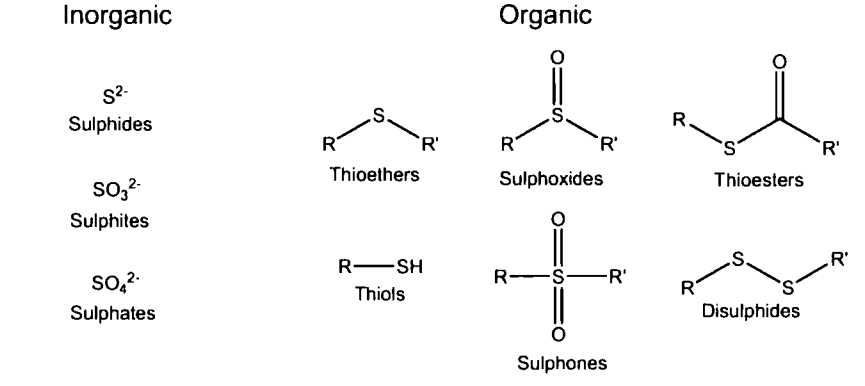


Figure 1.8 Sulphur; **A.** Sulphur is a group six element situated below oxygen, **B.** The electronic configuration of sulphur, note the empty 3p, 4s and 3d orbitals, **C.** The Pauling electronegativity scale, **D.** Sulphur compounds.

1.3.2.1 Sulphur Metabolism in Plants

Plants carry out out the important role of fixing inorganic sulphate (SO_4^{2-}) and reducing it to sulphide (S^{2-}), prior to incorporation into organic compounds. With anionic sulphate in abundance in soil, it forms the primary source of sulphur for plants. Sulphate is actively transported into roots through the action of plasma membrane H^+/SO_4^{2-} co-transporters driven by an electrochemical gradient generated by ATPase proton pumps[75]. It is subsequently transported throughout the plant and actively taken up by cells of the leaves. Once inside the plant cells sulphate can be stored in vacuoles or metabolised in plastids through reduction and assimilation (Figure 1.9). The pathways of sulphur uptake and assimilation have been well characterised in the model plant *Arabidopsis thaliana* where regulation of the pathways and adaptation to low sulphur environments have been studied, reviewed by Leustek[75] *et al.*, 2000.

Adenylation of sulphate is catalysed by ATP sulphurylase which generates 5'adenylsulphate (APS). APS is subsequently reduced in two enzymatic steps: firstly APS reductase transfers 2 electrons (believed to be from glutathione) to generate sulphite (SO_3^{2-}) and secondly sulphite reductase transfers six electrons from ferredoxin to produce sulphide (S^{2-})[75]. Sulphide or thiosulphide is then exported from the plastid, as enzymes for cysteine biosynthesis are located in the cytosol. The assimilation of sulphur into cysteine requires the coordination of both serine metabolism and the sulphate reduction pathway.

Cysteine forms a central intermediate at a crossroads in primary metabolism from which further sulphur biomolecules are synthesised. It can be channelled into protein synthesis or metabolised further to form the amino acid methionine or secondary metabolites (Figure 1.10). Hence, free cysteine levels are low (<10 μM) but its flux is high due to its incorporation into these further compounds[75]. In addition, cysteine forms the active constituent of the tripeptide glutathione (Figure 1.11), which functions to protect cells from biotic and abiotic stress[76]. It is an important redox buffer in cells and through the glutathione-S-transferase (GST)-detoxification system it functions as a scavenger of electrophilic oxidising agents[77].

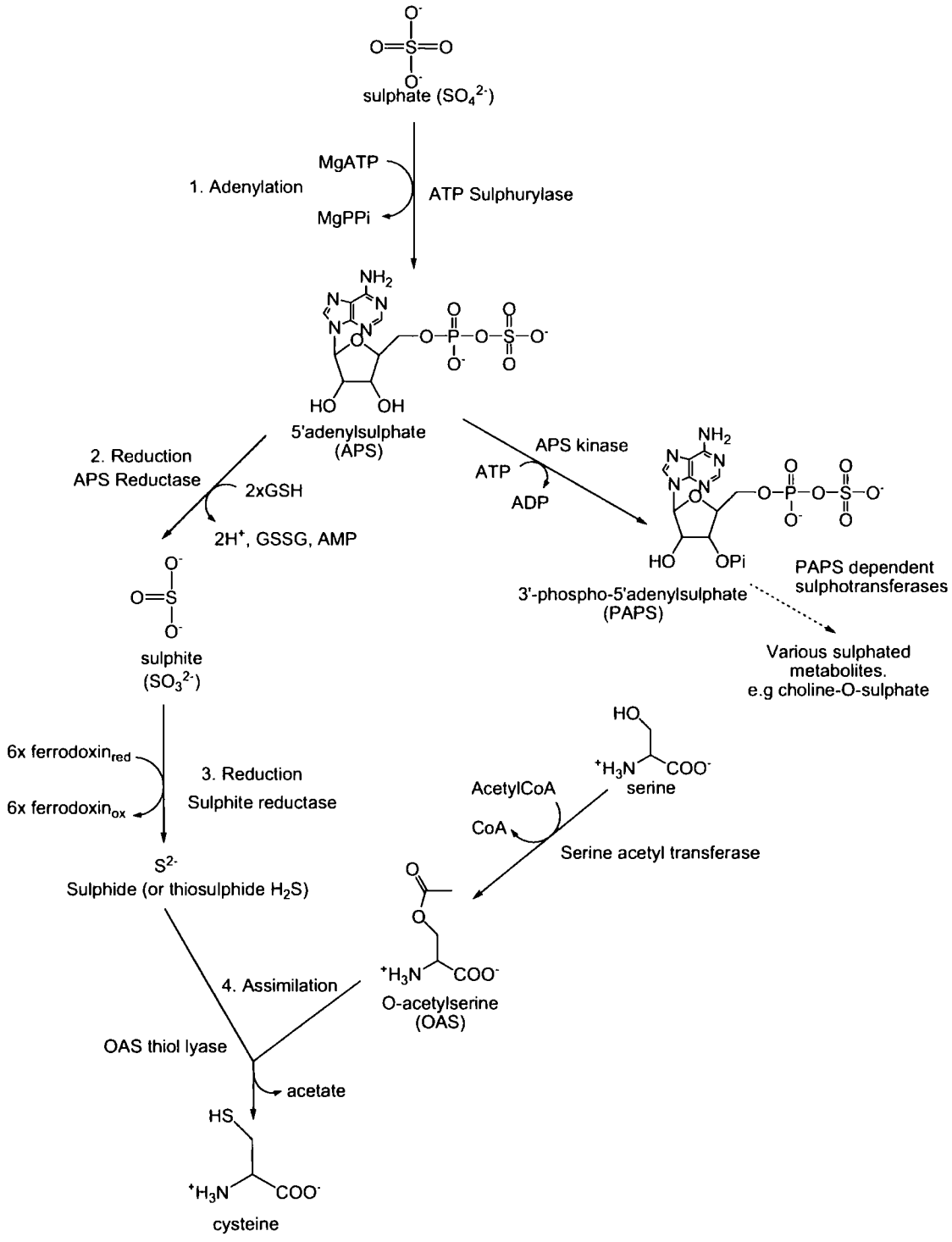


Figure 1.9 The biochemical pathways of sulphur reduction and assimilation by plants. APS can be reduced and incorporated into OAS to form cysteine or conversely it can be phosphorylated to form 3'-phospho-5'-adenylsulphate (PAPS), for the sulphation of various metabolites catalysed by PAPS-dependent sulphotransferases. Figure adapted from Leustek[75], 2000.

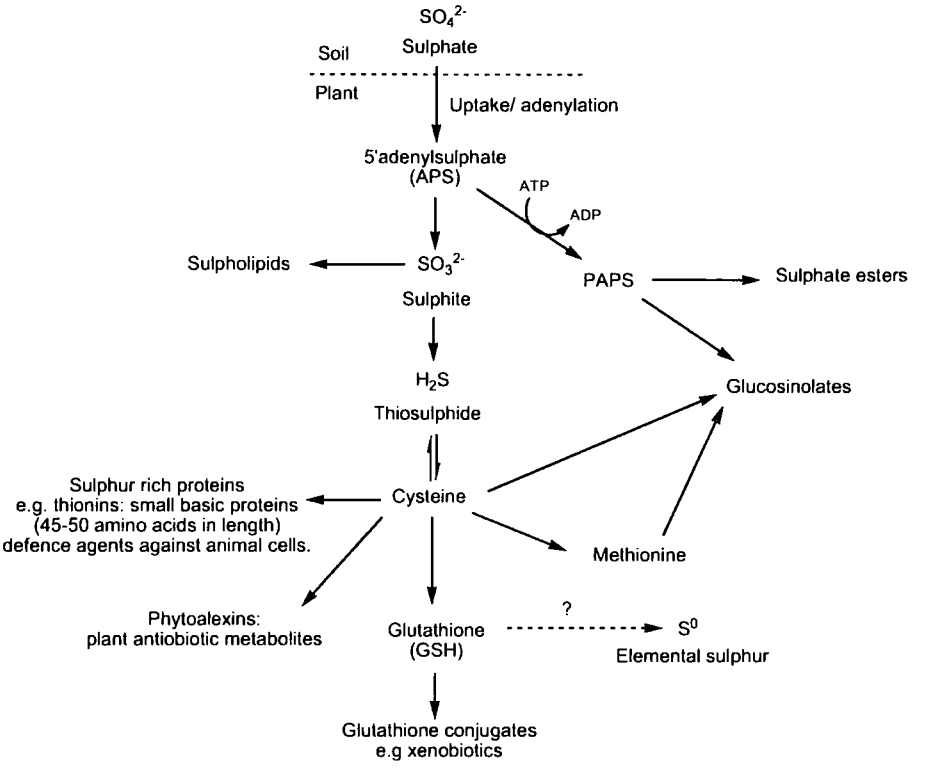


Figure 1.10 An overview of sulphur metabolism in plants. Figure adapted from Rausch and Wachter[78], 2005.

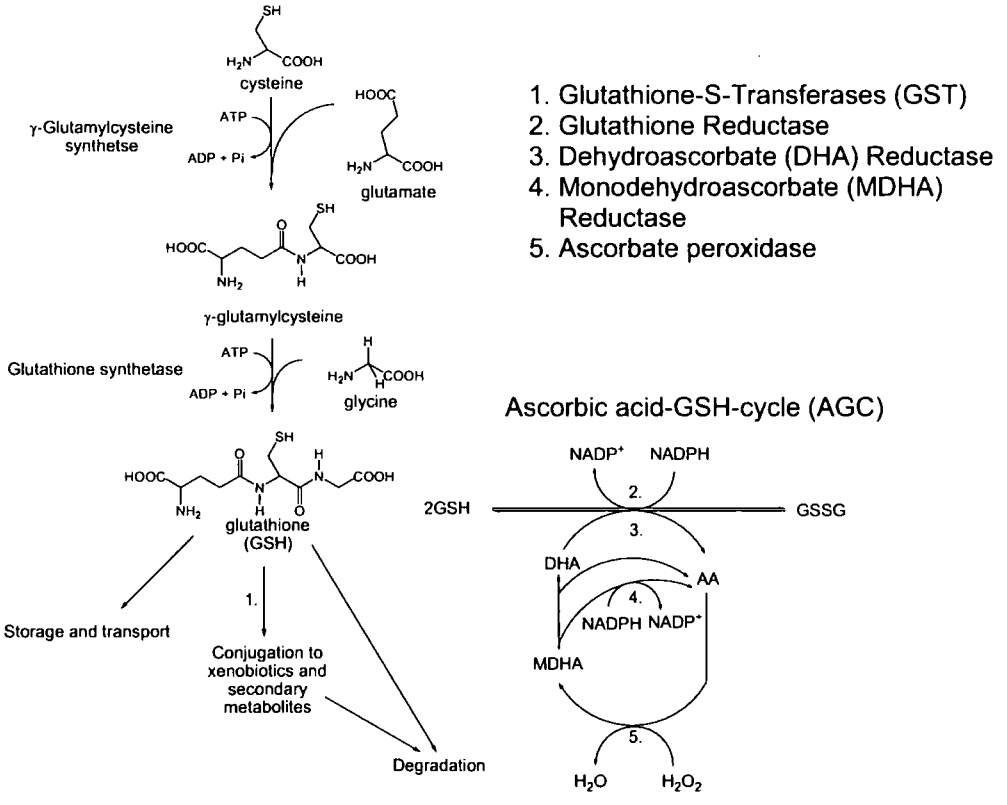


Figure 1.11 An overview of glutathione (GSH) biosynthesis and metabolism in plants. The ascorbic acid (AA) GSH-cycle is present in numerous cellular compartments and is responsible for detoxifying reactive oxygen species; O_2^-, H_2O_2. Figure adapted from Rausch and Wachter[78], 2005.

Many of the sulphur metabolites in plants are believed to function as defence agents[78]. For example the glucosinolates in *Brassicaceae* act as precursors to various volatile sulphur compounds, with pathogen attack inducing their hydrolysis through the action of thioglucosidase which generates the "defence-active" compounds[79] (e.g. isothiocyanates and thiocyanates, Figure 1.12A). A similar system is found in *Allium* species (onion and garlic) that sequester high levels of stable, odourless *S*-alk(en)yl cysteine sulphoxide aroma precursors, which can account for up to 1% of the bulb's dry weight[80]. Again, upon herbivore attack, compartmentalised enzymes, allinases, are released which cleave precursors generating defensive sulphur volatiles[81] (Figure 1.12B).

The accumulation of sulphur aromatic defence agents appears to be a characteristic of vegetables (e.g. onion, garlic, broccoli, cabbage). Fruits however, are produced to be eaten and their aromas function as attractants, not deterrents. Tropical plants appear to have incorporated the sulphur defence-pathways into their own fruit aroma biosynthetic machinery for the generation of pleasant exotic notes.

1.3.2.2 The Generation of High-Impact Aroma Chemicals in Tropical Fruit

The amino acid methionine forms a likely precursor to sulphur volatiles in a number of tropical fruits, such as the generation of methyl and ethyl thiopropanaotes in pineapple (Figure 1.3). The acid-contributing parts of such esters have an identical structure to the sulphur-containing functional side chain of methionine (Figure 1.13). As has been described, amino acid metabolism is important in the generation of branched-chain aroma esters found in many fruits[45,71]. However, pineapple has evolved to utilize methionine specifically for the generation of higher impact volatiles. Perhaps through repetitive mutations, these pathway enzymes have evolved to accept methionine; conversely, genes that encode enzymes responsible for metabolism of other amino acids have been silenced, hence channelling methioine into aroma biosynthetic pathways. A likely pathway is depicted in figure 1.13. This does however remain to be confirmed through feeding and radiolabelling studies similar to those used for the elucidation of similar pathways in model fruits.

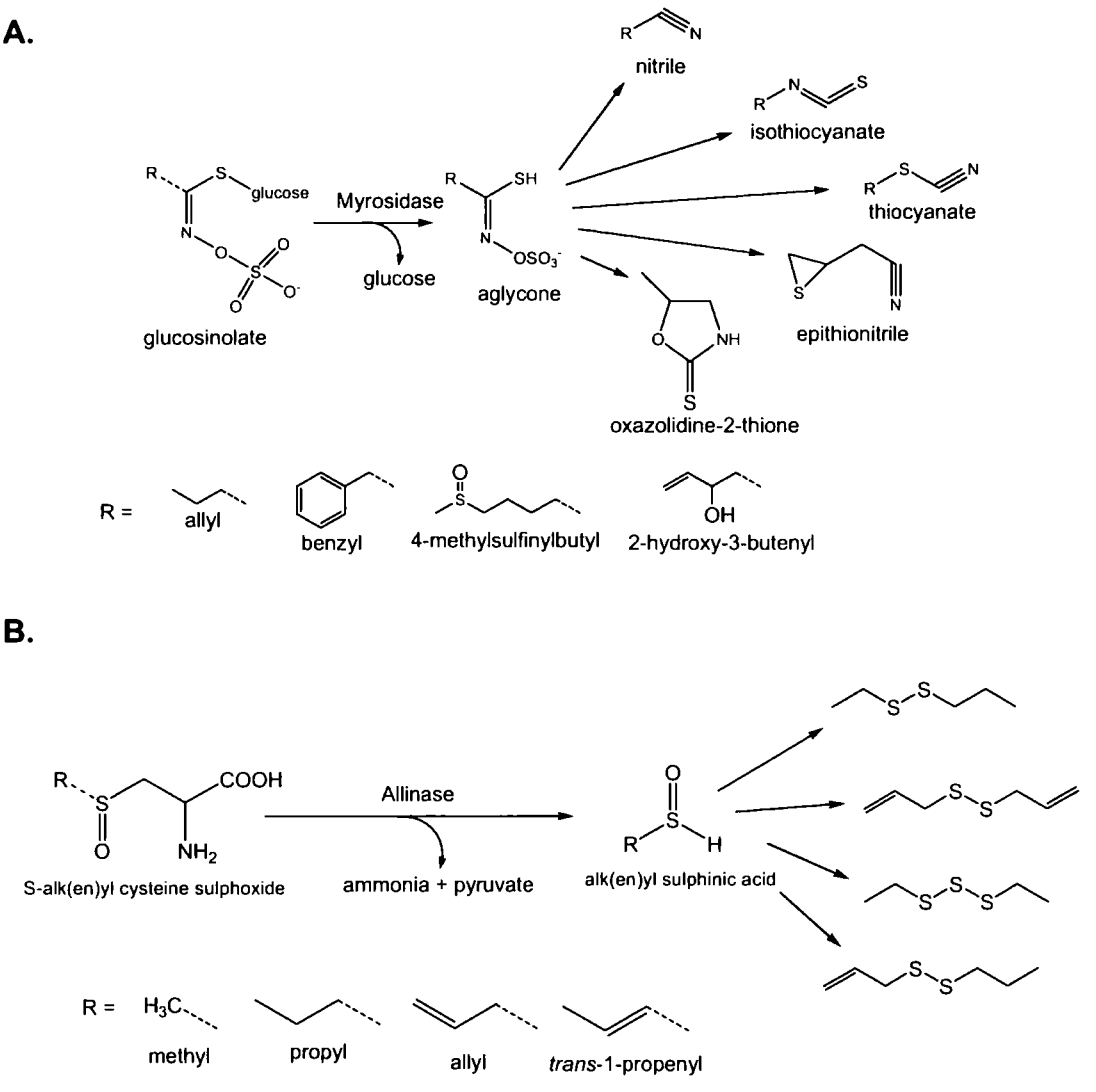


Figure 1.12 A. Glucosinolate defence agents in *Brassicaceae*. The core structure of the glucosinolates is shown, with variation due to different side chain structures (R) which are derived from amino acid metabolism and subsequent incorporation[79]. Glucosinolates are hydrolysed upon plant damage by the thioglucosidase, myrosinase. Hydrolysis generates glucose and reactive aglycone which rearranges to form different products (dependent on the side chain R). The sulphur volatiles produced have been proposed to function as both defence agents and attractants.

B. *S*-alk(en)yl cysteine sulphoxide defence agents in *Allium*. The core structure of *S*-alk(en)yl cysteine sulphoxide precursors is shown, variation is due to different side chain structures (R). The origin of such alky or alkenyl substituents remains to be confirmed however studies have indicated they are generated through amino acid metabolism, e.g. valine[81]. Cysteine is derived from glutathione incorporation and degradation[82]. Upon plant damage precursors are cleaved by the CS-β-lyase, allinase, which generates reactive alk(en)yl sulphinic acids. Condensation of such acids generates both symmetrical and asymmetrical thiosulphinites, which undergo further reactions and decompositions to produce disulphide and trisulphide volatiles.

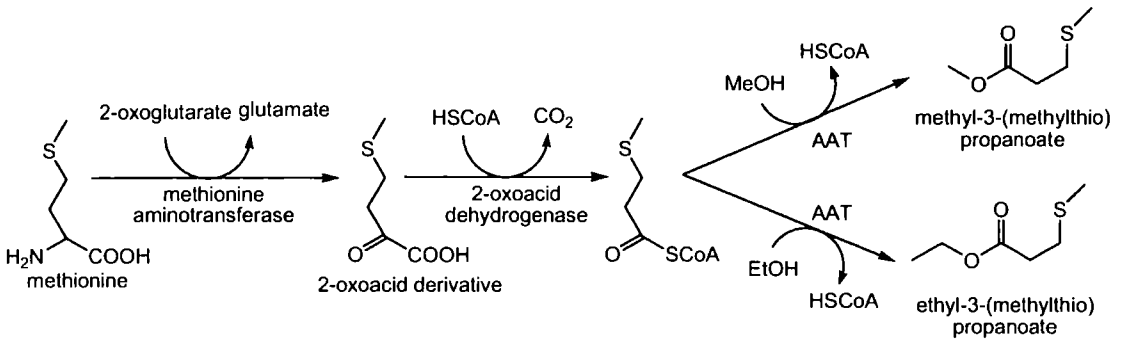


Figure 1.13 A hypothetical biosynthetic pathway for the generation of thiopropanoates in pineapple.

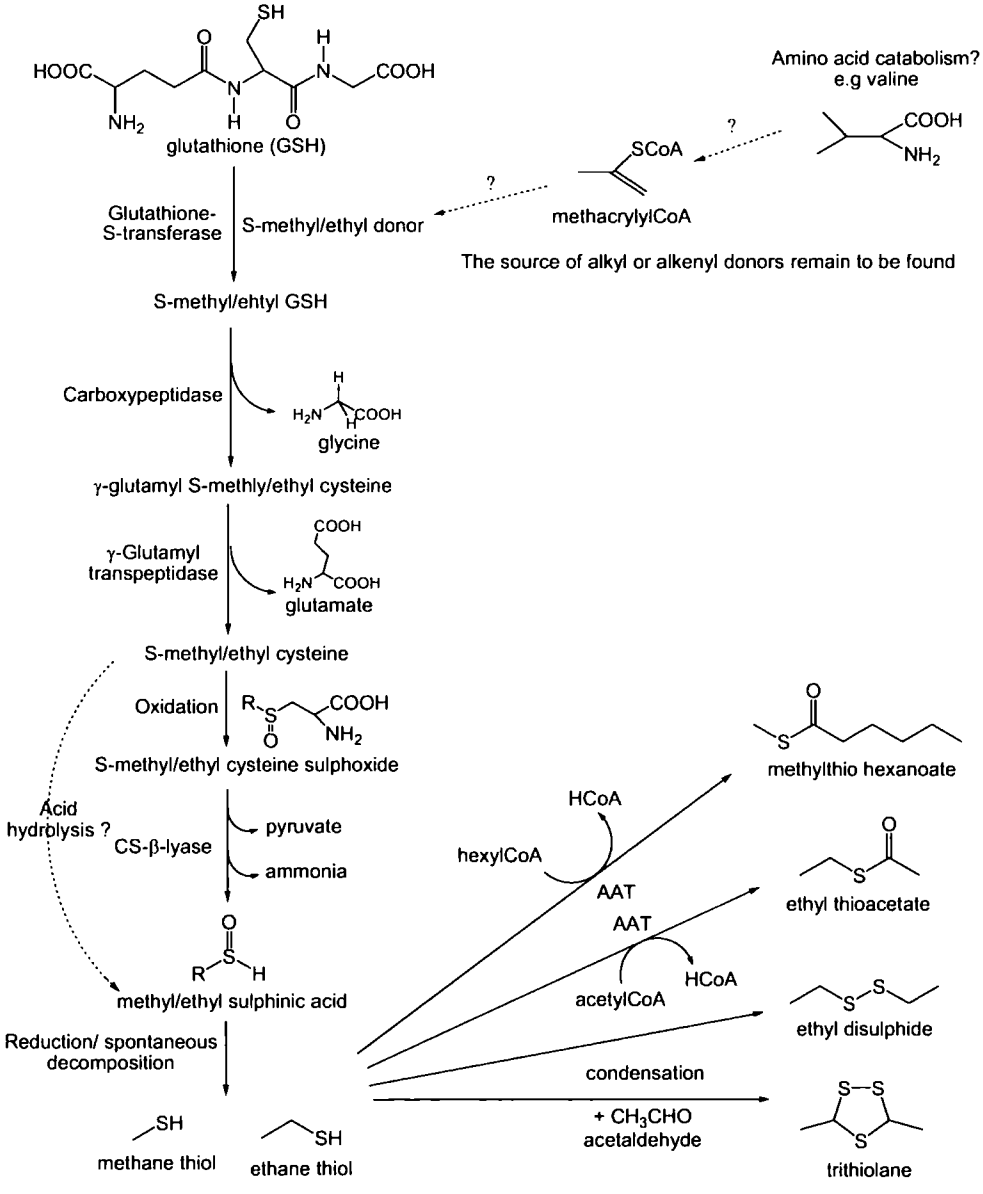


Figure 1.14 A hypothetical biosynthetic pathway for the generation of sulphur high-impact aroma chemicals in durian fruit, based upon similar metabolic pathways described in *Alliums* for the generation of *S*-alk(en)yl cysteine sulphoxides[82].

Glutathione is another sulphur-containing biomolecule believed to be important in incorporating sulphur into aroma chemicals, as illustrated by the generation of the *S*-alk(en)yl cysteine sulphoxide aroma precursors in *Allium* species[82]. Similar sulphur flavourings are found to be emitted from tropicals such as bread fruit and durian fruit, where the biosynthetic pathways are likely to be conserved between these species of plants (Figure 1.14). The thiol group of cysteine is a strong nucleophile and favours attack at electron rich regions of chemicals, such as alkene double bonds. Although the source of the alkyl and alkenyl substituents is somewhat unknown, radiolabelling studies in *Allium* using ^{14}C-valine showed this amino acid was metabolised into methacrylic acid, which is known to react with glutathione[81]. The *S*-alk(en)yl sulphoxides produced are sequestered in *Allium*, whereas in durian or breadfruit they are emitted, presumably through the action of CS-lyase enzymes under developmental regulation.

The incorporation of sulphur into terpenoid constituents is not common in tropical plants and hence it is not discussed here. However, it is important to note that in grapefruit the key character impact compound, grapefruit mercaptan (Figure 1.2), appears to be derived from sulphur addition to terpenes. This is the most potent of all high-impact flavourings and the isolated natural product would justify a high price tag in the industry. Hence, a metabolic understanding of sulphur metabolite biosynthesis is important for the development of systems for the bioproduction of such natural products.

Another biosynthetic pathway that utilizes glutathione as a means of incorporating sulphur into flavourings is found in passion fruit. The evidence for this first stemmed from the observation that certain white wine grape varieties, Sauvignon blanc grapes in particular, contained unusual tropical notes contributing to their overall aroma. The odour active chemicals were isolated and found to be 3-mercaptohexan-1-ol (3MH) and its acetate ester, 3-mercaptohexyl acetate[83], the same odour active chemicals found in passion fruit. In addition a cysteinylated precursor of 3MH was also identified[84] and led researchers to look for similar compounds in passion fruit itself where the cysteine precursor was identified as *S*-3-(hexan-1-ol)-L-cysteine[74].

The identification of this intermediate provided the first strong evidence toward how such tropical sulphur chemicals are produced in passion fruit. 3MH is a C6 volatile, a likely product of FA oxidative metabolism. Furthermore, the electron rich regions of the double bonds in unsaturated hexenals would provide the functionality for the incorporation of sulphur. Hence, the early biosynthetic steps in the generation of 3MH appear to be a consequence of the FA oxidation of linolenic acid through the LOX pathway which generates *cis*-3-hexenal (Figure 1.15A). Sulphur is subsequently incorporated through the glutathione detoxification pathway, whereby the thiol group of cysteine (as part of GSH) undergoes nucleophilic attack across the double bond in a Michael addition reaction catalysed by glutathione-S-transferase. The decomposition of GSH through the release of glycine and glutamyl generates *S*-3-(hexan-1-ol)-L-cysteine. The cysteine conjugate is not however found at considerably high levels in passion fruit (most as free 3MH), as opposed to grapes where 3MH is predominantly in the conjugated form[74]. This could be a consequence of constitutive CS-β-lyase activity in passion fruit which hydrolyses conjugates upon their formation or through the chemical acid hydrolysis of conjugates as a consequence of the low pH of the fruit juice. CS-β-lyases in passion fruit remain to identified.

Both the alcohol and thiol functionalities of 3MH are reactive and as a consequence the compound is metabolised further through the action of secondary transforming enzymes, such as AATs, which generate a collection of stable exotic sulphur volatiles (Figure 1.15B). Thioesters are likely to be formed through the action of ester synthases, such as AATs, as the thiol group will undergo esterification just like alcohols when presented with an activated acylCo-A donor in the active site of the enzyme.

Tropathiane is another key character impact aroma in passion fruit and it appears to be generated through the formation of a stable cyclic intermediate of 3-mercatohexyl acetate (Figure 1.15B). However, radiolabelling studies and the cloning and characterisation of enzymes proposed to be involved in such pathways are required for the confirmation of such metabolism.

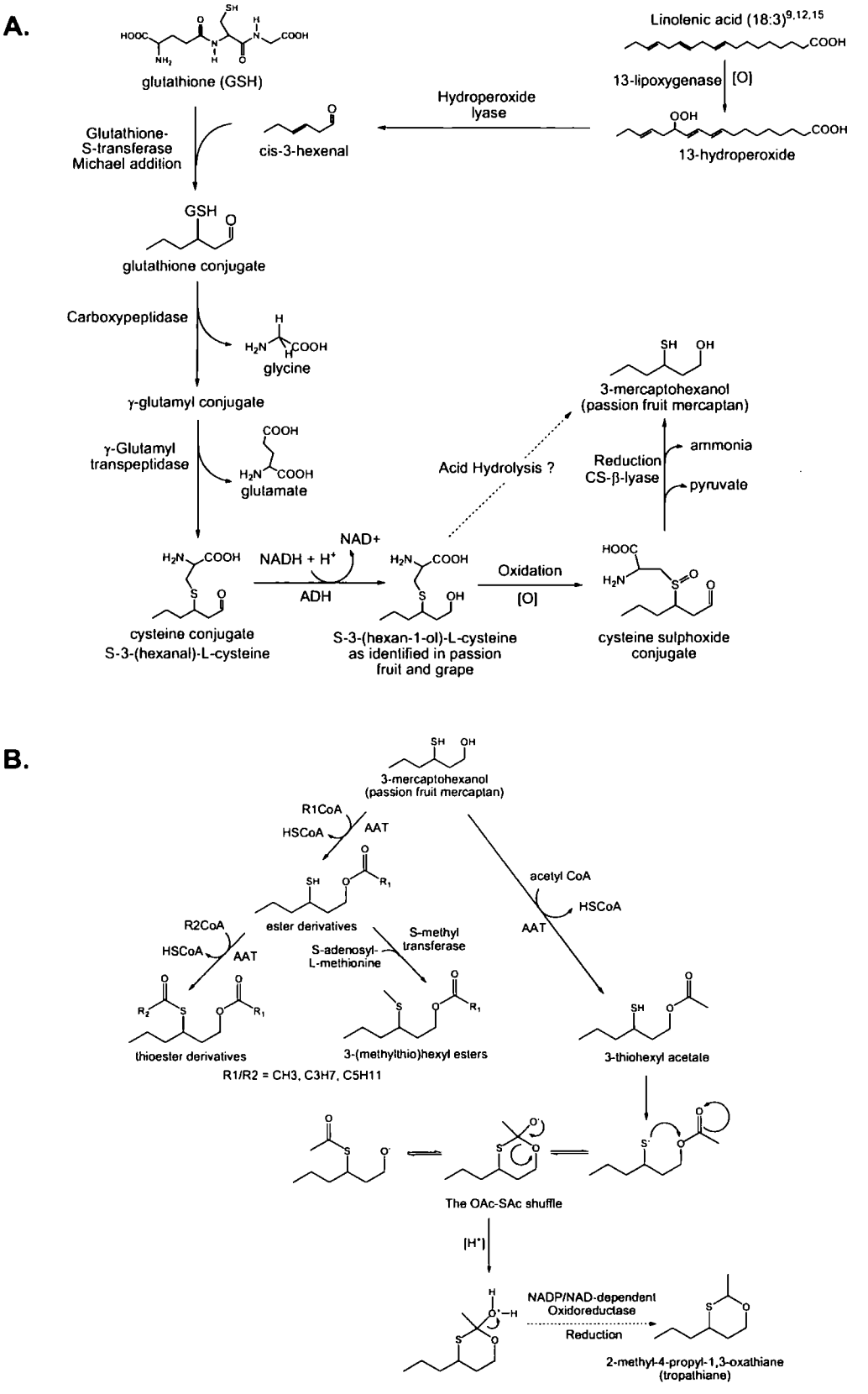


Figure 1.15 A. A hypothetical biosynthetic pathway for the generation of 3-mercaptohexanol in passion fruit. **B.** The generation of sulphur high-impact aroma chemical derivatives of 3-mercaptohexanol in passion fruit. Initially proposed by Rowe, D.,[40] 1999.

It is possible that these sulphur aroma biosynthetic pathways found in tropical plants (and *Allium* and *Bracaceae*) may be the ancient origins of the sulphur detoxification pathways recognised today for the metabolism of synthetic chemicals (e.g. agrochemicals) by plants. Such pathways would provide a number of possible endogenous roles for enzymes of secondary metabolism, such as the GSTs and cytochrome P450 oxidases, and may explain why they are found as such divergent gene families. How such sulphur metabolic pathways are incorporated into aroma biogenesis at the cellular level remains to be determined. The compartmentalisation of enzymes, protein-protein interactions and transport of intermediates will evidently be important in controlling such metabolism.

1.4 Storage and Release of Aroma Volatiles in Fruit

In contrast to aroma biogenesis, less is understood about how fruits store and emit aroma chemicals. However, the majority of studies have indicated that odourant compounds are synthesized exclusively in epidermal cells of the fruit cortex where they are either sequestered or emitted[6]. The aroma constituents containing reactive side groups, such as alcohols or thiols, could potentially be damaging to plant cells. For example, free thiol groups, which are over 500 times more reactive than the corresponding oxygen analogue[85], are likely to inactivate proteins through forming covalent disulphide bonds, interfering with catalysis and cell signalling. Hence, they are metabolised into the less reactive ester and thioester derivatives. However, esters and thioesters still remain lipophilic and are likely to interfere with membrane lipid components of plant cells, and a system for their storage or excretion is likely to be present. Recent studies have found that a considerable proportion of aroma constituents (40% in mango[86]) are locked up as sugar conjugates and exported to the cell wall. Similar non-volatile aroma precursors have also been characterised in passion fruit[87,88] and melon[89]. In addition, further chemicals may be transported to the vacuole for storage or excreted into the juice of the fruit. Isolating aroma chemicals is vital for the protection of cellular metabolism, however, very little

is known about such excretory processes or the metabolic trafficking between sub-cellular compartments.

In addition to sequestering aroma chemicals, fruit are also known to emit volatiles, a characteristic that first became apparent through the discovery of the gaseous hormone ethylene emitted by plants[90]. This brought about the realisation that some of the volatiles released by plants may be of physiological importance, even at concentrations too low for human detection. The release of alcohols and esters observed in numerous fruits may be due to their passive diffusion out of epidermal cells of the peel. However, it has also been proposed that enzymatic release is important in controlling volatile emission from fruit[93,94]. One family of enzymes which are upregulated in the later stages of fruit maturation and believed to serve this function are the esterases.

1.4.1 Esterases Identified in Fruits Involved in Flavour Release

Esterases are members of the hydrolase superfamily of enzymes (EC 3) that cleave ester bonds through the addition of water (Figure 1.16).

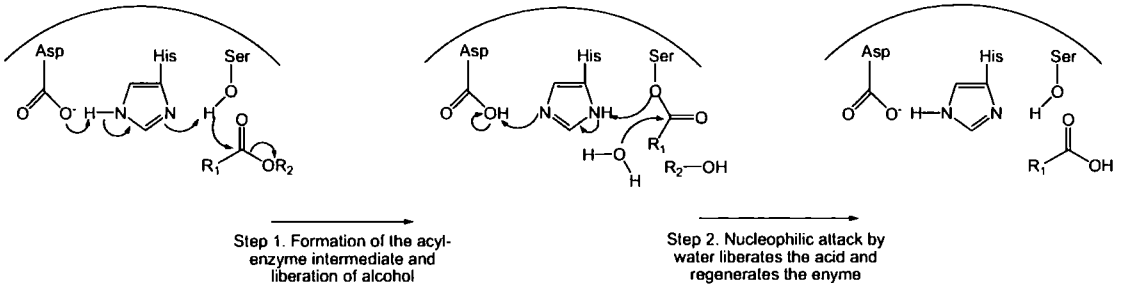


Figure 1.16 The catalytic mechanism of a serine hydrolase.
Serine hydrolase enzymes are a large sub-family of esterases which catalyse the hydrolysis of ester bonds, including amide esters, thioesters and carboxylesters, the latter illustrated here. Three amino acid residues in the active site; aspartate (Asp), histidine (His) and Serine (Ser) make up the "catalytic triad" which catalyzes hydrolysis. The triad reduces the pKa of serine thus making it a more effective nucleophile for attack at the carbonyl of the substrate which liberates alcohol and forms the acyl-enzyme intermediate (step 1). This is subsequently attacked by a nucleophile, water (under cellular conditions), to release the acid and regenerate the enzyme (step 2).

They have been found in abundance throughout nature, with plants known to express multiple isoforms in different tissues. Their involvement in xenobiotic metabolism has attracted particular attention, such as their role in agrochemical activation[91] and detoxification[92]. They serve to hydrolyse chemical groups on compounds introducing new functionality (e.g. hydroxyl) and increasing their solubility, the first stage in plant detoxification pathways. Their specific endogenous roles (before the advent of synthetic chemicals) is however harder to determine, they show broad specificities and are involved in the metabolism of various secondary metabolites and as such they may have numerous functions within the plant.

However, one specific endogenous role for esterases that has been proposed is that of generating various volatile aromas which are emitted from the peel of ripe fruit. An increase in esterase activity in late ripening was first observed in apple, with a total of five distinct carboxylic ester hydrolases being identified in the peel[93]. This study, and many other assays for carboxylic ester hydrolases used synthetic substrates (*p*-nitrophenyl acetate and α-napthyl acetate), and as a result no direct link to an endogenous role could be claimed. However, further studies in apple demonstrated that these fruit esterases were also active toward natural substrates, including butyl and hexyl acetates[94]. Alcohols, and in particular butanol, are found at far higher concentrations outside the fruit than within its tissues[93]. This is believed to be a consequence of esterases being found predominantly in the cortex and peel such that they hydrolyse esters as they diffuse out from the fruit. Such small alcohol volatiles may function as attractants indicating that the fruit has reached maturity. It appears a further level of chemical communication important for plant-plant, plant-insect and plant-animal interactions is found below that of human perception. Hence, the difficulties associated with understanding such interactions may explain why much of the metabolism involved in aroma release is not so well studied in fruit.

However, such undefined fruit esterases will be of particular interest to the flavourings industry as a novel source of biocatalysts due to their potential

substrate, regio- and stereo-specificities. With the tools of modern biochemistry and genetics, scientists have moved beyond just understanding plant aroma metabolism and our now looking at how to utilize such enzymes to reproduce the diverse chemistries of plant volatiles. The earliest work in organic chemistry focussed on identifying F&Fs and designing processes to produce them synthetically, now this is being taken a stage further through looking to plants (and microbes) to source the catalytic machinery required to generate aroma chemicals on a large scale through natural processes. This will undoubtedly revolutionise the F&F industry over the coming decade.

1.5 The Boioproduction of Flavours and Fragrances - Lipases, Esterases and Thioesterases

The use of Nature's catalysts enzymes for biotransformations in organic chemistry is a rapidly expanding area of research. Originally chemists would have been apprehensive to consider the use of enzymes in synthetic reactions due to their concerns over handling biological materials. However, this is changing with the development of efficient enzymatic processes that in many cases have become superior to classical methods of organic synthesis. There are numerous advantages of using enzymes (Figure 1.17), and they have become particularly attractive for use in industrial processes due to the pressures to develop environmentally acceptable systems for chemical manufacture. Hence, enzymatic processes are being developed through the use of whole cell systems (plant cell or microbes), or as isolated enzymes (crude preparations or in pure immobilized form) for the generation of specific chemicals. The latter are becoming increasingly available commercially and are attractive to the classic organic chemist as they can be handled just like any synthetic catalyst.

The major advantage of using biocatalysts is however their ability to carry out stereo and regioselective chemical manipulations, a useful tool for producing fine chemicals, such as pharmaceuticals, agrochemicals, neutraceuticals (health-benefiting compounds) and flavours and fragrances.

Advantages	Disadvantages
Selectivity: Enzymes are chemoselective; acting on particular functional groups, regioselective; distinguishing between functional groups at different sites on a molecule and stereoselective; able to distinguish between chiral substrates. The latter is termed the "specificity" of an enzyme and is a major advantage.	*Greatest activity in water:* The high boiling point and heat of vaporization of water makes it an unsuitable solvent for recovering product. In addition many organic compounds are insoluble in water. Hence immobilising enzymes to tolerate organic solvents is an active area of research.
Efficiency: Enzymatic rates of catalysis are high with accelerated rates of reaction as great as a factor of 10^{12}. This far exceeds those of synthetic catalysts, and hence less enzyme is required in comparison to chemical catalysts.	*Some require cofactors:* Some enzymes require cofactors such as ATP or NAD(P)H and they will not accept man-made substitutes. This is a major difficulty as recycling cofactors is problematic. However, systems are being developed.
Compatible catalysts: As enzymes function under the same reaction conditions a number can be used in a cascade. Such a system removes the requirement to isolate materials between reactions and hence prevents the loss of intermediates.	*Provided in only one form:* Enzymes are made up of L-amino acids and hence are only able to catalyse one form of stereoselective synthesis. If the product is the undesired enantiomer a search for the same function with the reverse selectivity must be sought in nature, a difficult task.
Environmentally acceptable: Enzymes are totally degradable and non-toxic, unlike chemical catalysts that are harmful if residues are released in effluent.	*Narrow parameters for operation:* If yields are low under the mild conditions of temperatures and pH there is little room for modification. However, remarkable enzymes have been found to work at extremes of temperature (0°C or 100°C).
Act under mild conditions: Optimum conditions at ph 5-8, temperature 20-40°C and in water, which reduces decomposition of product and prevent undesirable side-reactions. In addition less energy is required for such processes and hence environmental impact and energy costs are reduced.	*Prone to inhibition:* High substrate and product concentrations inhibit many enzymes and hence reduce productivity. Substrate levels can be kept low by drip feeding, but removal of product is more difficult to achieve.
Catalyse a broad spectrum of reactions: There is an enzymatic-process for almost every organic reaction known.	*Expensive:* Some are, but so are many chemical catalysts. Enzymes are however increasingly being produced on a large scale through microbial expression.
Not bound to natural role: It is often a common misconception that enzymes only function on their natural substrate, most are also able to act on non-endogenous chemicals.	*Allergies:* Enzymes can cause allergic reactions. However, this is irrelevant if handled with care and most synthetic catalysts will be far more harmful to humans and the environment.

Figure 1.17 Advantages and disadvantages of biocatalysts. Figure produced with information from Faber, K.[95], 2004.

Such bioactive compounds often only function as single enantiomers, with the other possessing little or even undesirable effects. Hence, the different enantiomers should be regarded as distinct species. Concerns over the side effects of the "wrong" enantiomers in humans and their impact upon the environment has put pressure on these industries to produce enantiomerically pure products. In 1990 only 11% of drugs sold globally were single chiral compounds whereas today this has increased to around 36%, a consequence of stereoselective synthesis often through enzymatic approaches[95].

Enzymes do however have their disadvantages (Figure 1.17), in particular they only perform optimally under natural "mild" conditions of pH, temperature, pressure and in an aqueous environment. Such reactions conditions are often not favourable for conducting certain chemical transformations, overcoming these obstacles are the greatest challenges facing biotechnologists. However, this discipline is still (relatively) in its infancy and through screening nature for novel selective enzymes and through directed evolution, a greater range of biocatalysts with further favourable properties will become available.

As described F&Fs are bioactive compounds with many found in nature as single chiral constituents. However, the non-selective synthetic processes used industrially generate racemic mixtures (containing equal quantities of each enantiomer), with the different forms showing distinct odour properties. The subsequent isolation of the desirable enantiomer is costly and difficult to achieve. As a result enzymes are now seen by the industry as a novel source of chiral catalysts for selective aroma production. Such enzymatic approaches should however not be regarded as competing with classical synthetic methods or replacing plant aromatic extraction, they are simply an additional tool available for manufacture. Ultimately economics will decide which processes are adopted. However, the generation of certain aroma constituents in their natural form appear only achievable through enzymatic approaches. For example, important flavouring constituents, such as raspberry ketone, are found at only trace levels in plants (4mg/kg of berries[96]) making their extraction costly and unsuitable, and hence natural biocatalytic systems are very attractive in such situations.

With the advantages of enzymatic flavour production and the industries' desire to introduce such technologies it may appear somewhat surprising that it is not widely in use. This is however a likely consequence of the great chemical diversity found in aromatic chemicals, the "simple" enzymatic procedures (like those available commercially) will not currently suffice to generate this level of complexity. In addition, volatile flavouring ingredients are not produced on an extremely large scale, only 400 are produced at over one tonne a year, and as will be discussed, bioprocesses have therefore only been developed for those products in greatest demand (e.g. vanillin, benzaldehyde and menthol)[15]. This is in contrast to the bioproduction of non-volatile food ingredients, such as sweeteners, savoury constituents and flavour enhancers which are produced on a scale of over 1,000 tonnes per year through enzymatic means[15]. Their production only requires "simple" biotransformations, with commercial lipases and microbiological processes found to be very effective. However, there is a considerable amount of academic literature on the biosynthesis of volatile flavours and many processes have now been developed for successful lab scale synthesis. The processes that have been scaled up for industrial production are predominantly based on the use of plant and microbial whole cell systems for the generation of "mass produced" flavours.

The observation that microbes (yeast, fungi and bacteria) were able to biosynthesise aroma volatile constituents identical to those found in many plants led researchers to explore (or exploit) such organisms for flavour production, reviewed by Vandamme, E.J.[97] (2003). For example, the plant terpenoid pathway appears to be mimicked in the yeast *Kluyveromyces lactis* which generates fruity floral terpenes such as linalool, citronellol and geraniol[97]. Hence, microbes can be used to duplicate plant secondary metabolism either through *de novo* synthesis or by adding a substrate/ precursor to enhance levels of desirable volatiles. Of all microorganisms, fungi, and in particular basidiomycetes, are found to have the closest volatile spectrum to that of plants and hence many have been utilized for industrial scale production[15]. Such an example is the generation of natural benzaldehyde by *Ischnoderma benzoinum*. Microbes fed on media rich in L-phenylalanine (cheap and plentiful) were found to degrade it almost completely to the

flavouring constituents benzaldehyde and 3-phenylpropanol[98] (Figure 1.18). In addition microbes have also been utilised for their ability to mimic plant FA metabolism for the generation of flavour lactones[97] and phenylpropanoid metabolism for the generation of vanillin[99]. As described, plant secondary transforming enzymes are important for generating diversity in aroma chemicals, enzyme activities such as P450, ADH and AAT are also evident in such microbial processes. The latter being illustrated through feeding branched-chain alcohols from fusel oil (a cheap by-product of alcohol rectification) to yeast, *Williopsis saturnus*, generating fruity acetate esters[100].

However, there are certain cases where the desired plant metabolism simply can not be found in microbes or it only generates trace quantities of products. In such cases plant whole-cell systems have been favoured. The use of plant tissue or cell cultures for the production of vanillin serves as a useful example. Currently the world demand for natural vanillin can not be met through extraction from the plant[99] and hence many alternative processes for natural production have been developed (Figure 1.19). Systems using *Vanilla planifolia* tissue cultures are of particular interest. Feeding of phenylpropanoid precursors, such as cinnamic acid or ferulic acid to tissue or cell cultures enhanced the generation of vanillin and vanillic acid. For example a 1.7 fold increase in vanillin was observed in cell cultures fed with ferulic acid as compared to untreated cells[101]. Yields however remained low (15 mg/kg tissue callus/day). A novel method is taking aerial roots from the plant and culturing them in media rich in ferulic acid combined with charcoal[102]. The charcoal acts as an adsorbent of vanillin and increased yields to 400 mg/kg tissue/day, this is 5-10 times the amount that can be extracted from an equivalent weight of vanilla-beans[102]. A further area of research in this field is the use of phytohormones, such as naphthalene acetic acid or cytokinins, to increase cell division and upregulate the desired metabolism[99].

Hence, whole cell plant or microbial processes are effective for aroma production when conversion rates are high, cheap feedstock is readily available and subsequent extraction of product is not too difficult or costly.

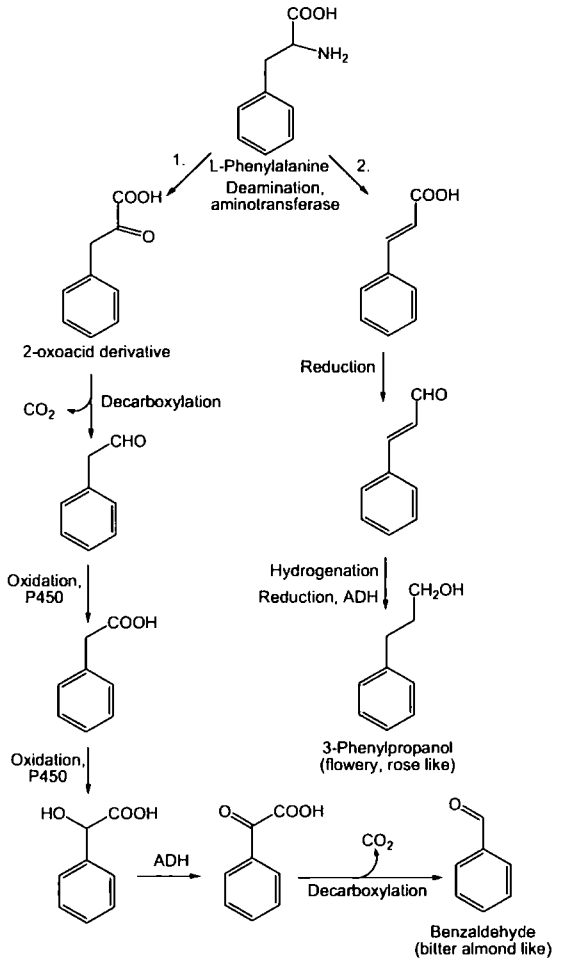


Figure 1.18 The bioproduction of benzaldehyde (1.) and 3-phenylpropanol (2.) by the fungi *Ischnoderma benzoinum.* Radiolabelled phenylalanine was added to submerged cultures of the yeast and degradation products and intermediates identified[98]. This natural process is however only viable due to the plentiful and cheap supply of L-phenylalanine which is an intermediate in the production of the sweetener, aspartame.

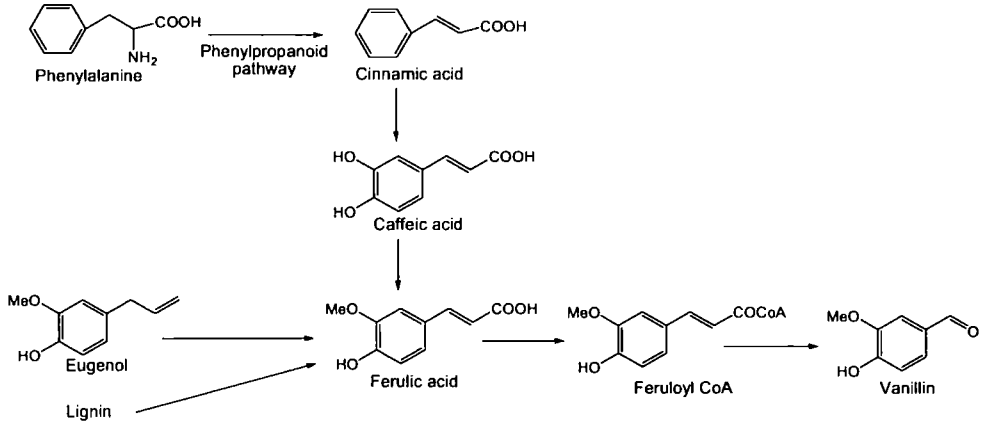


Figure 1.19 The bioproduction of Vanillin. Biotechnological processes for vanillin production have utilised plant and microbial whole-cell systems. Feeding of precursors such as phenylalanine, phenylpropanoids (cinnamic acid, caffeic acid, ferulic acid, eugenol) or lignin (a complex polymer of C9 phenylpropanoids) have enhanced the generation of vanillin.

There are however limitations to this approach. High levels of substrate or product can be toxic to cells and hence must be drip fed into the reaction or products continuously removed. Yields are generally low, rarely greater than 100 mg/Litre reaction, due to by-product formation and inhibition[97]. However, screening for further microbes and optimising novel forms of plant tissue culture will be likely to increase the repertoire of products they can generate on an industrial scale.

Now that the metabolic understanding of aroma biogenesis is becoming clear, it is also possible to look to single enzyme systems for "cleaner" production. Specific biocatalysts with desirable properties and specificities can be screened for in plants and microbes and subsequently genes encoding such proteins can be isolated and transformed in to bacteria or yeast for overexpression. The enzyme catalyst can then be used as a crude cell lysate or can be purified and immobilised to enhance stability and aid in enzyme recovery. Such technologies are particularly attractive for the generation of volatiles not achievable through whole cell transformations. One group of enzymes that have attracted particular interest due to their exceptional properties as biocatalysts and that have found there way into numerous industrial processes are the lipases.

1.5.1 Lipases for the Bioproduction of Flavours and Fragrances

The lipases (EC 1.1.1.3), are believed to account for 40% of industrial bioreactions[103]. They are found in abundance throughout nature with the biological role of hydrolysing triacylglycerols to release FAs. They function through a similar catalytic manner to that described for carboxyl ester hydrolases. However, they are unique in their ability to work on insoluble substrates at the aqueous/ organic interface, a property that distinguishes them from esterases. It is their ability to catalyse different types of chemical reactions, such as hydrolysis, esterification and trans-esterification that makes them so versatile (Figure 1.20).

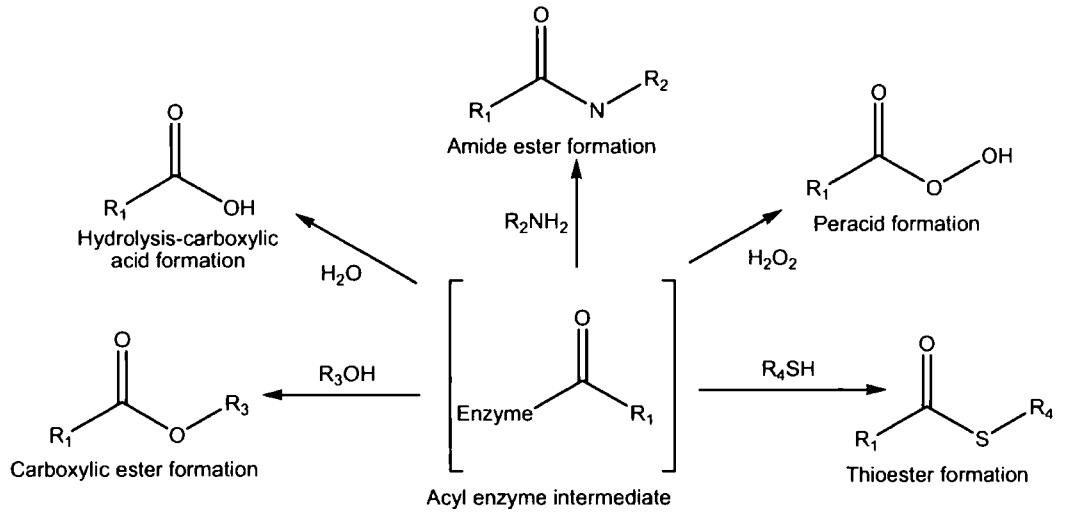


Figure 1.20 The versatility of lipase catalysed reactions. Different products are formed dependent upon the nucleophile that attacks the acyl-enzyme intermediate. Such nucleophiles must be in excess of water, hence why such reactions are conducted using immobilised lipases in non-aqueous solvents. Figure adapted from Faber K.[103], 2004.

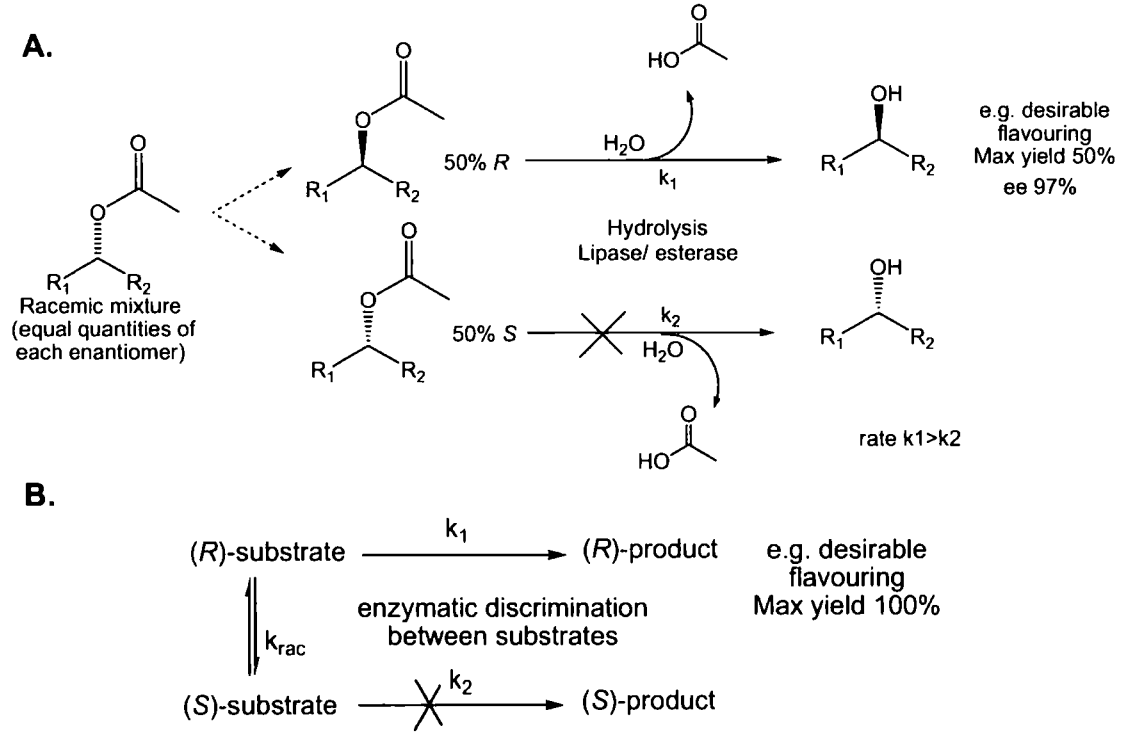


Figure 1.21 A. Kinetic resolution. Hydrolases such as lipases and esterases are able to discriminate between chiral chemicals, as is illustrated above through the stereoselctive cleavage of a carboxyl ester. In theory only 50% yield of a single enantiomer can be achieved (the other remains unreacted). ee: enantiomeric excess (% of defined enantiomer). In practice the ratios of each enantiomeric product is a consequence of rates k1 and k2. **B.** Dynamic kinetic resolution. In dynamic kinetic resolution the unreactive enantiomer is converted to the reactive form through racemisation. The active form is then converted to product (*R*-product in this case), with theoretical yields of 100%.

Lipases are readily available, as they are secreted by various microbes and can be produced on a large scale in fermentations. Immobilised lipases have been demonstrated to function in organic solvents[104]. Those sourced from thermophilic microbes can tolerate high temperatures[105], and importantly they do not require expensive cofactors. Hence, they are a versatile set of selective enzymes that have unsurprisingly gained so much interest as biocatalysts.

Lipases have therefore become targets for the bioproduction of aroma chemicals, with many efficient lab-scale processes now developed. For example, various immobilised microbial lipases have been utilized for the synthesis of short chain esters in organic media, a means of reversing natural hydrolysis. Fruity esters such as ethyl pentanoate (green apple flavour) and hexylacetate (pear flavour) have been produced through the use of immobilised crude *Staphylococcus simulans* lipase (fungi) in a biphasic reaction[106] (20% water by weight). Through the careful control of reaction conditions; water content, pH, solvent type, temperature and acid/alcohol molar ratios, yields of 41% hexylacetate were achieved[106]. Immobilised lipase was recovered from the completed reaction and reused in up to five cycles with no loss of activity. Similar studies using lipases from *Mucor miehei*[107] and *Candida cylindracea*[108] immobilised on nylon supports were also found to be effective for the production of short-chain esters. Such studies demonstrate the importance of careful choice of biocatalyst, immobilisation support and defining the optimum reaction conditions for obtaining maximum yields.

Lipases have also been utilised for their ability to produce optically active products through kinetic resolution and dynamic kinetic resolution (Figure 1.21). Theoretically lipases will only act on one enantiomer at significant rates and leave the other unreacted, a tool for the kinetic resolution of racemic mixtures. Such an example is the kinetic resolution of (*R*)-menthol, an important flavouring produced at over 5,000 tonnes a year. The *R*-enantiomer is found in nature and said to be cooling, fresh, sweet and minty, whereas the *S*-enantiomer is less cooling and bitter to taste[96]. Pure (*R*)-menthol has been shown to be generated through the transesterification of racemic (*R*/*S*)-menthol using *Burkholderia cepacia* lipase[109] (Figure 1.22).

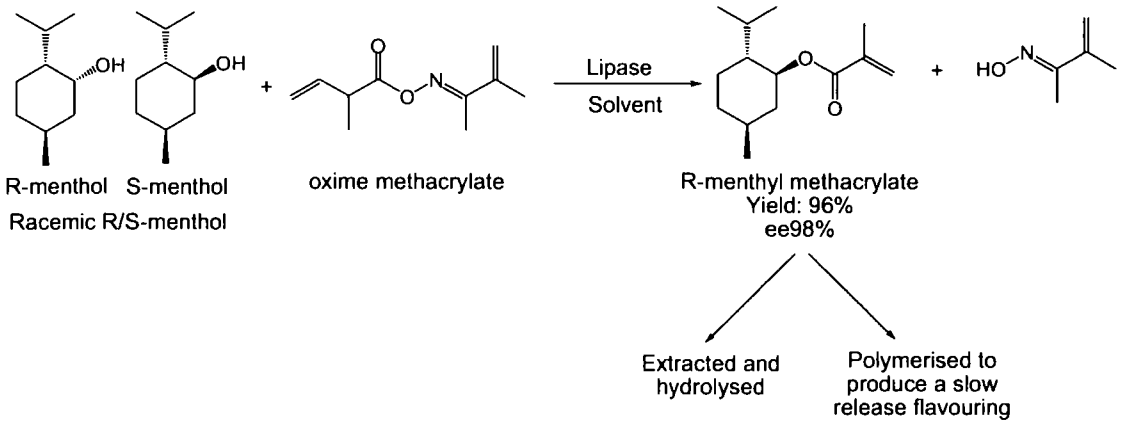


Figure 1.22 The kinetic resolution of (+/-)-menthol using *Burkholderia cepacia* lipase. The optimal conditions for the stereoselective generation of *R*-menthol (desirable form) were achieved using oxime methacrlate as the acyl donor and diisopropylether as the solvent. *R*-menthyl methacrylate can be extracted from the reaction mixture due to its different properties and hydrolysed to generate *R*-menthol or polymerised to produce a slow release flavouring. Figure adapted from Athawale *et al.*,[109] 2001.

A drawback of kinetic resolution is that one enantiomer remains unreacted, hence theoretically 50% of the feedstock is wasted. This is however overcome through dynamic kinetic resolution, whereby the unreacted enantiomer is racemised using synthetic catalysts or racemase enzmes (EC 5) to provide further substrate for the selective enzyme. This process should theoretically convert 100% of starting materials to product. An example is the synthesis of chiral δ-lactones, a process which utilises the selectivity of lipases in combination with a ruthenium-catalysed alcohol racemisation[110] (Figure 1.23).

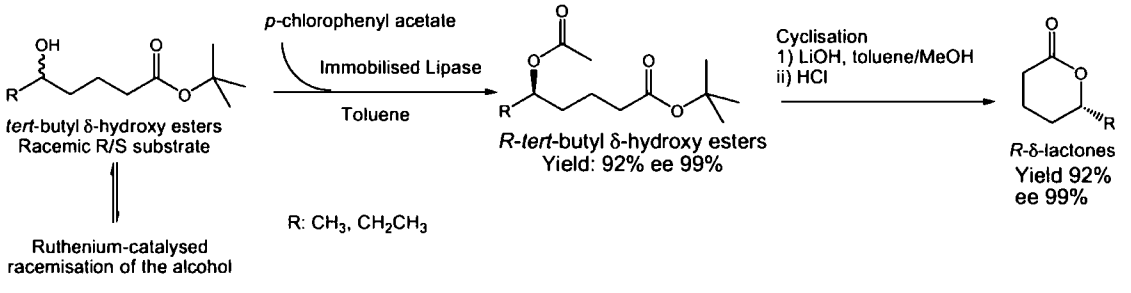


Figure 1.23 The dynamic kinetic resolution of δ-lactones through lipase-catalysed transesterification of δ-hydroxyesters in combination with ruthenium-catalysed alcohol racemisation. Immobilised *Candida antarctica* lipase (Novozyme-435) selectively transesterified the ester precursors to generate *R*-δ-hydroxyesters. In combination with a ruthenium catalyst for the racemisation of substrate high yields of desired product were achieved. Figure adapted from Pamies and Backvall[110] 2002.

Hence, the stereoselectivity of lipases in combination with their relaxed substrate specificities makes them exceptional biocatalysts. They clearly have application for the bioproduction of chiral flavourings and through optimisation of such processes they are likely to be used on an industrial scale.

1.5.2 Esterases for the Bioproduction of Flavours and Fragrances

One drawback of lipases for aroma bioproduction is that they work optimally on larger (non-volatile) substrates. The esterases however show much potential for use in systems using short-chain soluble substrates. In contrast to the large number of readily available microbial lipases, only a small number of esterases are in use commercially, predominantly pig liver esterase (PLE) and horse liver esterase (HLE)[103]. As a consequence the range of esters that can be acted upon is limited to the specificity shown by these enzymes. More recently efforts have been made to screen microbes for novel esterases, in particular those from thermophilic organisms that can withstand higher temperatures[111].

Of all the esterases available, PLE remains the best in terms of versatility and enantioselectivity and has therefore been used in a number of biotransformations. It is particularly useful for the regiospecific hydrolysis of protection groups on delicate compounds, as is used for the generation of the anti-inflammatory drug prostoglandin E1, that would otherwise be destroyed through classical hydrolytic conditions[103]. The ability of the enzyme to generate enantiopure products from racemic mixtures has also been utilised for the generation of optically pure starting materials for drug synthesis, such as the synthesis of carboxylic nucleoside analogs used as antiviral drugs[103]. However, the use of PLE in flavour bioproduction is somewhat prohibited as the use of pig enzymes will contravene Halal and Kosher requirements.

A further potential source of esterases is however fruit. As described, they are abundant in esterases. These enzymes have been proposed as useful biocatalysts but as they have not been cloned into microbes for overexpression, they are not readily available. However, the substrate specificities of such enzymes makes them very attractive as a novel source of biocatalyst for aroma production.

1.5.3 Thioesterases in Plants, a New Source of Biocatalyst for the Flavourings Industry

As described, high-impact sulphur volatiles form the major character impact compounds in numerous tropical fruits and are important aroma constituents in cooked foods. As a consequence of their high-odour impact, they are however only generated in trace quantities in fruit (e.g. 3-mercaptohexanol 5 μg/kg passion fruit)[112] and hence are difficult and costly to extract in isolated form; most are therefore produced as synthetic nature identical chemicals. Many of the natural tropical aromas are however present as chiral compounds, for example 3-mercaptohexanol and its derivatives are predominantly found in the *S*-form[27] (90-98%), with the other enantiomers posing significantly different sensory properties[113]. Hence their synthesis through classical synthetic methods without resolution of enantiomers can result in reduced flavour properties. As a result sulphur volatiles are targets for bioproduction.

However, surprisingly few studies have been conducted on the biosynthesis of such valuable aroma materials. Those of particular mention have predominantly focussed on the use microbial enzymes, such as lipases, esterases and CS-β-lyases. These studies have investigated the potential of microbial enzymes for catalysing thioesterification, transesterification and the hydrolysis of thioesters and cysteine conjugates.

A number of microbial lipases have been shown to possess thioesterase activity, such as the yeast *Candida rugosa* lipase (EC 3.1.1.4, type VII) which is one of the most widely used industrial biocatalysts[103]. It was shown to catalyse the formation of important thiol-containing food constituents, such as 2-methyl-3-furanthiol and 2-furfurylthiol through the hydrolysis of the respective thioacetate precursors[114] (Figure 1.24). Through optimising reaction conditions (temperature, solvent, pH, substrate and enzyme concentrations) yields at equilibrium of 88% and 80% were achieved respectively. A common problem encountered with thiol-containing aroma chemicals is their reactivity, causing them to readily oxidise to the corresponding disulphides upon storage. Hence, it was proposed that the thiol could be protected through acetylation and

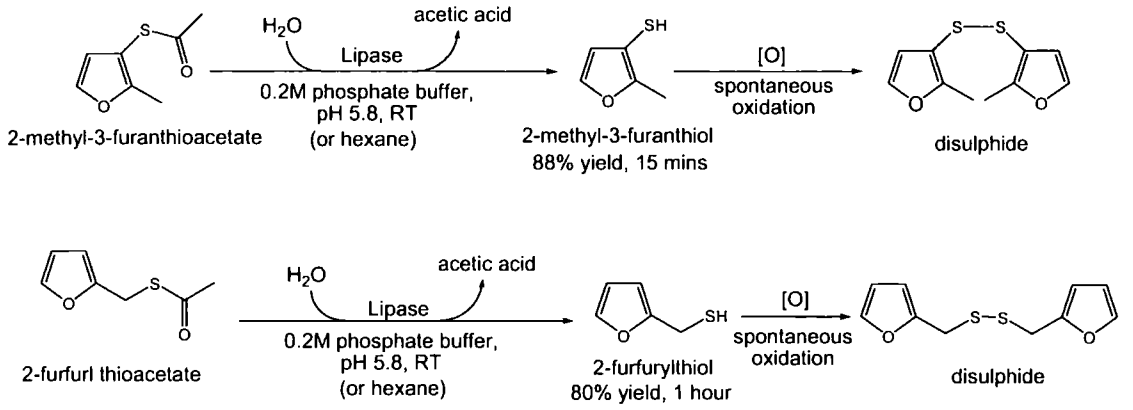


Figure 1.24 The generation of 2-methyl-3-furanthiol (roast meat) and 2-furfurylthiol (roast coffee) through lipase catalysed hydrolysis of thioacetates. The yeast lipase from *Candida rugosa* was used in free and immobilised form. Both thiols were generated in good yields, however, oxidation of product to the disulphide is a problem in water with all of 2-methyl-3-furanthiol converted to the disulphide over 2 hours. Hence the reactions were repeated in organic solvents (hexane, pentane) with immobilised lipase, rates were slower but yields were the same and the thiol product was more stable. Figure produced with information from Rhylid *et al.*,[114] 2002.

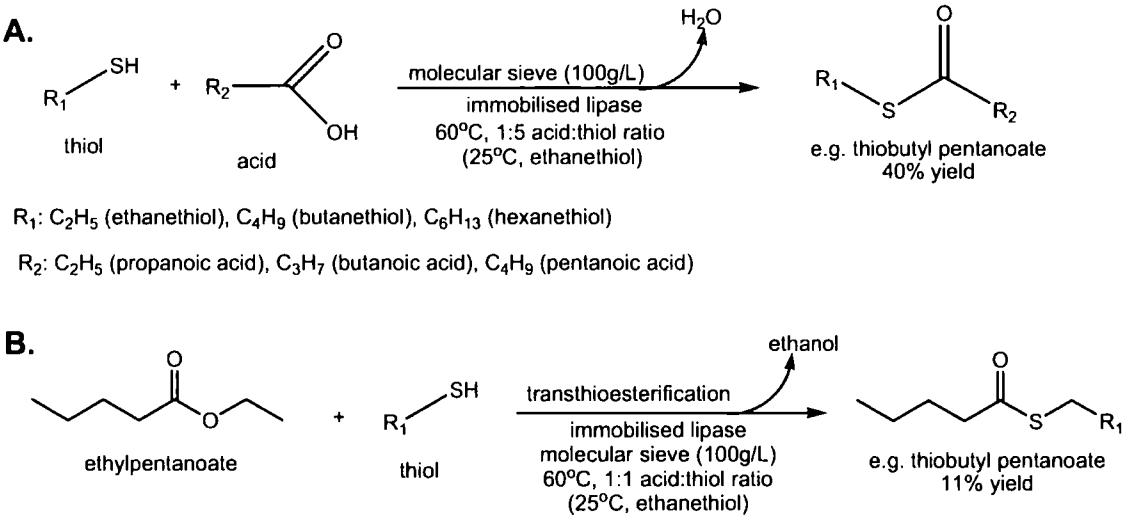


Figure 1.25 Lipase synthesis of short-chain thioesters in solvent free medium. Immobilised yeast lipases from *Candida antarctica* and *Mucor miehei* were used to reverse the natural hydrolysis and generate thioesters in solvent-free medium. Reaction conditions were optimised for the generation of thiobutyl pentanoate as illustrated above. Thioesterification (A.) had a slower reaction rate but yields at equilibrium were greater than the reaction through transthioesterification (B.). Figure produced with information from Cavaille-Lefebvre and Combes[115], 1997.

subsequently cleaved enzymatically through such a process to generate the free mercaptan prior to use, in a sense mirroring what plants do in Nature.

Immobilised lipases in non-aqueous solvents were studied for their ability to reverse the natural hydrolytic reaction. For example, immobilised lipases from *Mucor miehei* and *Candida antarctica* were shown to produce short-chain thioesters in solvent-free medium[115] (i.e. no water, only thiol and acid) (Figure 1.25). For example a thioester yield of 40% was achieved for the condensation of pentanoic acid and butanethiol at 60°C, 1:5 (acid: thiol) ratio with addition of molecular sieves to remove the water generated upon condensation.

An important emerging area of flavour research is that of CS-β-lyase cleavage of cysteinylated aroma precursors for the selective liberation of odour active thiols. For example, the stereoselective generation of 3-mercaptohexanol (passion fruit) and 8-mercato-*p*-menthan-3-one (thiomenthone, buchu leaf) by microbial CS-β-lyase cleavage of the respective cysteine precursors has been investigated[116,117]. The process for generation of 3-mercaptohexanol was investigated using bacterial lipases, tryptophanase from *Escherilia coli* and a crude enzyme extract from *Eubacterium limosum*, and yeast cells, all known to posses CS-β-lyase activity (Figure 1.26A). Cleavage of the 3-L-cysteinylhexanal derivative only produced low yields of product (20%) with a slight preference for the *S*-configuration (ee 57%). However, the cleavage of 3-L-cysteinylhexanol through the two processes demonstrated slightly higher selectivities (ee 65%), with tryptophanase generating the *S*-enantiomer and *E.limosum* selective for the *R*-constituent. The same biocatalysts were used to study the stereoselective generation of pulegone from cysteine precursors[117] (Figure 1.26B). Pulegone can exist in four possible enantiomeric forms (2 enantiomers and 2 diastereroisomers), each with significantly different odour properties[117]. Both enzymes showed little stereoselectivity over the different forms of thioether bond. However, discrimination between the two diastereoisomers was more significant with a clear preference for the *cis*-configuration. Yields however remained low in both cases, possibly due to inhibitory effects.

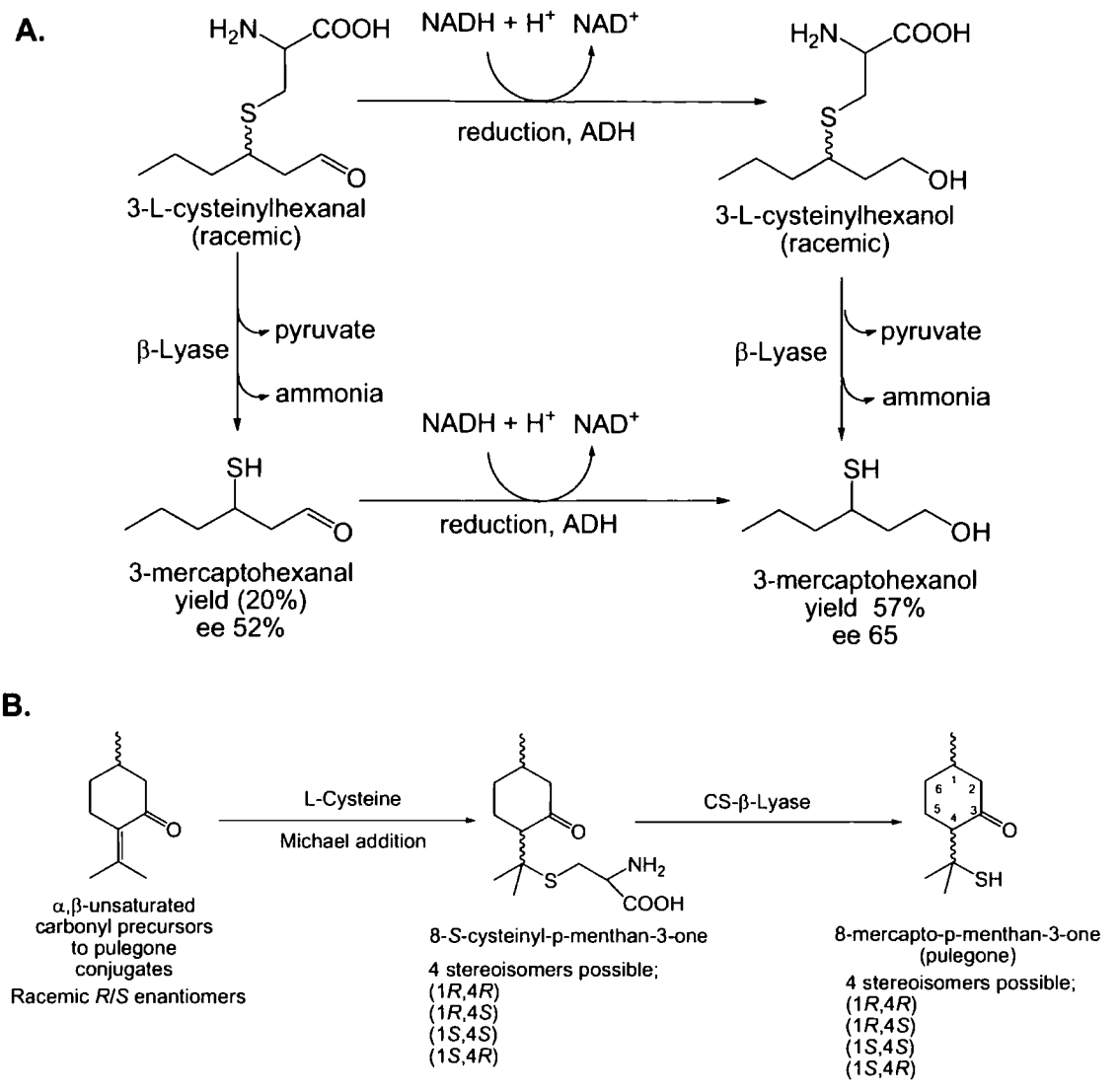


Figure 1.26 A. The stereoselective generation of 3-mercaptohexanol. Synthetically produced L-cysteine precursors to 3-mercaptohexanal and 3-mercaptohexanol were used to determine the stereoselective cleavage using various CS-β-lyase enzymes. Tryptophanase from *E.coli* and a crude enzyme extract from *E.limosum* showed little selectivity toward the hexanal derivative and slightly improved yield and selectivity toward the hexenol derivative (found in nature). Yeast cells were found to metabolise the hexenal derivative to 3-mercaptohexanol, due to endogenous ADH activity, yields and ee remained low.

B. The stereoselective generation of pulegone through CS-β-lyase cleavage of cysteine precursors. Pulegone cysteine conjugates were synthetically produced and were a racemic mixture of four stereo forms. The same enzymes as above were used to determine selectivity for substrates and product. Both tryptophanase and *E.limosum* showed a preference for the cis form of the substrate (1*R*,4*S*) ee 64% and (1*S*,4*R*) ee 88%.

On an industrial scale, thioester functionality can be introduced to chemicals through the addition of sulphinic acid across double bonds, with subsequent cleavage through acid/base hydrolysis generating the free thiol[118] (Figure 1.27). However, such synthetic approaches do not discriminate between enantiomeric forms and a racemic mixture is consequently produced. An enzymatic means of kinetic resolution would be desirable. A possible approach is illustrated through the selective generation of 3-mercaptohexanal by lipase-catalysed hydrolysis of the thioacetate[113] (Figure 1.28). A total of 16 commercially available enzymes (15 microbial lipases and one esterase, PLE) were screened for their ability to selectively generate enantiomericly pure product at high yields. The lipase from *Candida antarctica* was found to be optimal generating a yield of 36% and with high enantioselectivity for the *S*-product (ee 91%). The immobilised form of this enzyme provided even greater selectivity for the *S*-product (ee 97%), believed to be a result of the increased rigididity of the enzyme in this form. This approach demonstrates the potential of thioesterases for the selective generation of sulphur aroma chemicals with enhanced odour properties. Yields however still remain relatively low and more specific biocatalysts are desirable.

Thioesterase enzymes with their ability to transform sulphur aroma chemicals and generate them in a sterospecific manner are hence of interest to the flavouring industry. However, the studies described account for almost all of the published academic research in this area todate and the potential for further development is great. Such processes have predominantly used microbial enzymes or the mammalian lipase (PLE) for biotransformations and plant systems have not been considered. As described, the substrate and product specificities of lipases are limited, a property that may account for many of the low yields described. Hence lipases are not the optimal catalysts for transforming small volatile aroma constituents.

With the identification of VOSC thioester derivatives in tropical fruit, including 3-(thioacetyl) hexylacetate identified in passion fruit[119] and 8-acetylthio-*p*-menthan-3-one reported in buchu plant extracts[120], it is now a logical step to look to plant sources, and in particular tropical fruit plants, as a novel source of

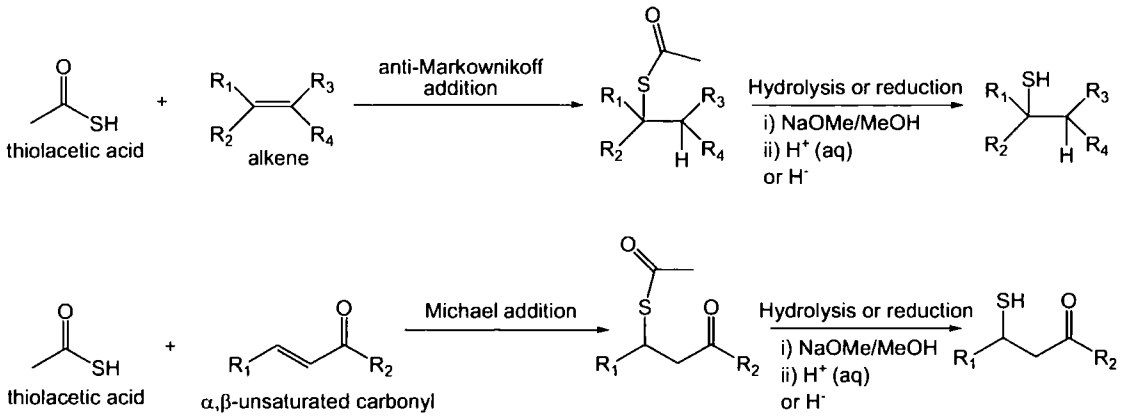


Figure 1.27 The synthetic synthesis of thioesters and thiols. Thiolacetic acid forms the principle reagent for introduction of sulphur into organic molecules. It can be achieved through addition across a double bond via a radical anti-Markovnikoff addition or by Michael addition to α/β-unsaturated carbonyls. The thioester can subsequently be hydrolysed to yield the free thiol functionality. Note hydrolysis is not stereoselective through this procedure. Figure adapted from Jameson, S.B.[118], 2001.

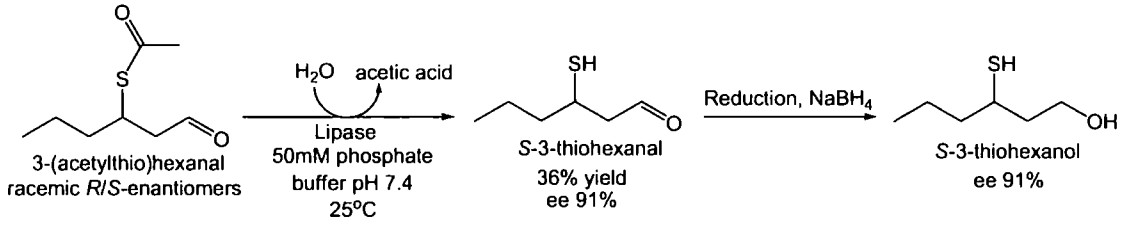


Figure 1.28 The stereoselective generation of 3-mercaptohexanol through lipase catalysed hydrolysis of thioacetate precursors. The greatest yield and stereoselectivity was produced by *Candida antarctica* lipase (as above). The immobilised form of this enzyme increased the stereoselectivity of the reaction (ee product 97%), however, much of the product was lost (up to 75%) through adsorption onto the immobilisation resin[113].

thioesterases. They will have evolved over generations to produce the selective and specific enzymes required for thioester and thiol production. Thioesterases from tropical fruits would potentially show greater "specificity" for short-chain substrates important in flavour biosynthesis. In addition they are likely to be "selective" for enantiomeric forms with the desired odour characteristics.

The identity of thioesterases in plants involved in F&F release are however unknown, and it would be of interest to elucidate how they function. The only other reports on plant thioesterases are of those involved in fatty acid synthesis[121]. In plants FA biosynthesis occurs in plastids where acyl chains are covalently bound to an acyl carrier protein (ACP) through thioester bonds. The extension cycle can be terminated by selective thioesterases which generate free FAs and ACP. These thioesterases work through a catalytic cysteine in the active site, perhaps this functionality is optimal for hydrolysing thioester links as opposed to the serine hydrolase functionality found in many carboxylesterases and lipases.

Fruit thioesterases would form potentially "versatile" biocatalysts, being able to catalyse many of the important transformations described; thioesterification, trans-esterification and would complement synthetic processes through stereoselective thiol formation. As these biocatalysts are found in living organisms, this system is regarded as "natural" under current European[122] and US legislation[123], as long as the feedstocks are also from a natural source and reaction conditions are mild (temperatures and pressures found in traditional home cooking). As a result the products generated would be considered premium flavouring agents. The desire to isolate and utilise such unknown enzymes from tropical fruit forms the focus of this work.

1.6 Aims and Objectives

The aim of this work was to isolate a thioesterase from tropical fruit for use in the bioproduction of natural volatile organosulphur chemicals. Hence, the potential of using plant enzymes as improved biocatalysts was initiated. Specific attention was paid to the specificity, selectivity and versatility of thioesterases to determine if they could be used as industrial biocatalysts in the flavourings industry. The work was broken down into the following objectives:

- Develop an assay for measuring thioesterase activity and use this to test for the proposed activity in tropical fruit. Screen tropical fruit for the optimal activity.
- Characterise the thioesterase activities in the optimal source through the use of natural and synthetic substrates. Conduct initial studies on the performance of the enzyme as a biocatalyst; stability, temperature tolerance and yields with the crude enzyme.
- Purify and clone the thioesterase from the optimal tropical fruit. Test the performance of the purified thioesterase.
- Optimise the production of the biocatalyst.
- Investigate the potential of using tropical fruit thioesterases for the bioproduction of volatile organosulphur compounds and optimise the processes through immobilisation, defining reaction conditions and looking at yields of product and enantioselectivities. Investigate the potential for scale-up to an industrial process.
- It was hoped this work would initiate a programme for developing systems for the bioproduction of sulphur aroma chemicals and through random screening of a tropical fruit cDNA library further biocatalysts were identified for the continuation of this work.

2. Materials and Methods

2.1 Fruit and Plant Materials

Fruit was supplied by H2H wholesalers (Newcastle Upon Tyne, UK). The following fruit varieties were used; purple passion fruit (*Passiflora edulis* Sims, Kenya), yellow passion fruit (*Passiflora edulis* F. *flavicarpa*, Brazil), mango (*Mangifera indica*, South Africa), Valencia orange (*Citrus sinensis* L., California, US), lychee (*Litchi chinensis*), banana (*Musa sapientum*), grapefruit (*Citrus paridisi*), blackcurrant (*Ribes nigrum*), raspberry (*Rubus ideaus*), lime (*Citrus limettioides*) and lemon (*Citrus limon*). Fruit was stored at the wholesalers at 5-8°C under high humidity (85%) to delay ripening. Subsequently, fruit was ripened under controlled conditions in growth rooms at the Centre for Bioactive Chemistry (Durham, UK), at 25°C with a 16 hour photoperiod for a duration stated in the text (5-10 days), dependent upon fruit type. Fruit was splashed with water each day and stored under clear plastic to increase ambient humidity. All fruit material used for protein extraction was fresh and not stored at -80°C prior to use.

The model plant *Arabidopsis thaliana* used for gene cloning was grown under controlled conditions for 3-4 weeks in plant growth rooms at the Biological Sciences Department (Durham, UK). Seeds were germinated in moist potting compost and grown under a 16 hour photoperiod (24°C) and an 8 hour dark period (22°C). The light intensity was 80 μEinstein/m^2 /sec. Compost was kept moist with regular watering. At harvest, foliage tissue was removed and flash frozen in liquid nitrogen before storage at -80°C.

2.2 Protein Extraction

Fruit tissue was homogenised in a chilled Waring blender, with all subsequent steps carried out on ice. Protein was extracted in 4 v/w 0.1 M potassium phosphate buffer, pH 7.2 containing 0.25 M NaCl, 2 mM EDTA, 4 mM DTT and 5% w/w PVPP. The homogenate was stirred at 4°C for 2 hours to ensure the complete extraction of wall bound protein before filtering through two layers of

Miracloth (Calbiochem, Nottingham, UK) and centrifuging at 10,000 g for 15 min (4°C) to remove insoluble material. Protein precipitating between 40-80% ammonium sulphate $(NH_4)_2SO_4$ saturation was then recovered using centrifugation (10,000 g, 15 min, 4°C) with the resulting pellets stored at -80°C until required. Prior to use, protein pellets were resuspended in 10 v/w 20 mM potassium phosphate buffer, pH 7.2 and dialysed overnight against the same buffer in a semi-permeable cellulose membrane (Sigma-Aldrich) to desalt the extract and remove any additives.

2.2.1 Protein Determination

Protein content was determined using the Bradford protein dye-binding reagent (BIO-RAD) as previously described (Bradford, 1976)[124], with γ-globulin used as a reference protein.

2.3 Assay Methods

The high-impact thioester and thiol chemical flavourings were donated by Oxford Chemicals Limited (Hartlepool, UK). Enzyme assays were run in triplicate with mean specific activities (nkat mg^{-1} protein) determined. The level of variability (error) is displayed as +/- one standard deviation from the mean. Heat treated protein samples, denatured at 95°C for 5 min, were run in parallel to determine the chemical rate which was subtracted from the total rate (+ enzyme) to determine the true enzymatic rate.

2.3.1 Ellman's Thioesterase Assay

Protein samples were dissolved in 490 µl of 0.1 M potassium phosphate buffer, pH 7.2. Thioester substrates (0.1 M in acetone) were added to a final concentration of 1 mM and incubated at 30°C. An equal volume (500 µl) of Ellman's reagent (dithio-*bis*-(2-nitrobenzoic acid)), 4 mg/ml in 0.1 M phosphate buffer, pH 7.2 was added to the assay, mixed and analysed on a Beckman Coulter DU 530 UV/Vis spectrophotometer. Product formation was determined by measuring the change in absorbance at 412 nm over 1 min (extinction coefficient: 14150 $M^{-1}cm^{-1}$). Authentic standards were used to produce a

standard curve for quantifying product formation, refer to chapter 3 for the validation of the Ellman's thioesterase assay.

2.3.2 Gas Chromatography Thioesterase Assay

Thioesterase activity was also determined using gas chromatography (GC). Protein samples were dissolved in 490 μl of 0.1 M potassium phosphate buffer, pH 7.2 in air-tight 1 ml glass vials. Thioester substrates (0.1 M in acetone) were injected using a Hamilton syringe through a septum to final concentration of 1 mM, sealed with Parafilm and incubated at 30°C for 5 min (unless otherwise stated). Products were partitioned into dichloromethane (0.5 ml) and the organic phase dried over sodium sulphate prior to injection (5 μl) onto a Varian capillary GC column, CP SIL coated (Theromo). A Thermo Focus GC system was used with a splitless injection under the following parameters; injector temperature: 200°C, Initial oven temperature: 80°C, ramp rate 5°C/min to a maximum temperature: 200°C, carrier gas: Nitrogen, column pressure 30 kPa. Products were identified with a flame ionisation detector, with the identity of sulphur volatiles confirmed and quantified using authentic standards.

2.3.3 Fluorimetric Thioesterase Assay

Protein samples in 0.1 M potassium phosphate buffer, pH 7.2 (2 ml) were incubated with 1 mM thioester substrates at 30°C in a 2 ml fluorimetric cuvette. The reaction was started by the addition of 10 μM bromobimane (3-bromomethyl-2,5,6-trimethyl-1H). The increase in fluorescence emission at 475 nm was measured over 2 min using an FP-6200 spectrofluorimeter (Jasco), with excitation at 400 nm. A standard curve was established using authentic thiol standards to quantify product formation.

2.3.4 Spectrophotometric Thioesterase Assay

The emission spectra for authentic thioester and thiol standards (50 μM) in methanol were scanned between 200-600 nm on a spectrophotometer (as above) as a means of determining the maximum absorbance wavelength for the thiol groups. Protein samples were dissolved in 990 μl of 0.1 M phosphate buffer, pH 7.2 and incubated with 1 mM thioester substrates at 30°C. The

increase in absorbance at the chosen wavelength (e.g. 377 nm for 4-methylumbelliferyl thioacetate) was measured over 2 minutes.

2.3.5 HPLC Thioesterase Assay

Protein samples in 0.1 M potassium phosphate buffer, pH 7.2 (500 µl) were incubated with 1 mM thioester substrates at 30°C for 15 min. The reaction was stopped by mixing an equal volume of methanol and centrifuging at 12,000 g for 5 min. Samples (200 µl) were transferred to HPLC vials and 10 µl injected onto a Dionex UltiMate 3000 HPLC with a C18 column (Waters, Symmetry 4.6 x 30 mm, 3.5 micron beads). Solvent A: 0.5% formic acid in water, Solvent B: 0.5% formic acid in acetonitrile at a flow rate of 0.8 ml/min starting at 10%B to 100% B over 12 ml. The identity of the sulphur volatiles was confirmed and quantified using authentic standards.

2.3.6 Carboxylesterase Assays

Esterase assays were either run on a spectrophotometer or a fluorimeter. For the spectrophotometric assay protein samples were incubated with 1 mM substrate in 0.1 M phosphate buffer, pH 7.2 at 30°C to a total volume of 1 ml. Either *p*-nitrophenol acetate (*p*-nitrophenol extinction coefficient: 17000 $M^{-1}cm^{-1}$) or α-naphthol acetate (α-naphthol extinction coefficient: 2300 $M^{-1}cm^{-1}$) were used as substrates and the rate of reaction was determined through measuring the change in absorbance at 400, or 310 nm respectively over 1 min.

When using the 4-methyl umbelliferly esters a fluorescence assay was used. Protein samples in 0.1 M potassium phosphate buffer, pH 7.2 (2 ml) were incubated with 1 mM methyl umbelliferyl esters at 30°C in a 2 ml fluorimetric cuvette. The increase in fluorescence emission at 450 nm was measured over 2 min using a fluorimeter (as above), with excitation at 370 nm. A standard curve was established using methyl umbelliferol to quantify product formation.

2.3.7 Amide-esterase Assays

The measurement of amide esterase activity was determined using a fluorimetric assay (as above), except the substrates used were amides of 4-

methyl umbelliferol. Excitation and emission wavelengths were as stated above.

2.4 Protein Purification Methods

Chromatography materials and columns were supplied by GE-Healthcare (formerly Amersham Biosciences). Initial chromatography (DEAE and Octyl sepharose) was carried out at 4°C using the Akta Basic liquid chromatography system (GE-Healthcare). Subsequent chromatography using pre-packed high-performance columns (Phenyl superose, Mono S, Superdex 200) was carried out at room temperature on an Akta FPLC system (GE-Healthcare). The elution of protein was monitored through UV absorbance at 280 nm.

2.4.1 Large Scale Protein Extraction

For the purification of thioesterases from purple passion fruit and orange, large scale protein extracts were made. For the extract from passion fruit the peel of 25 ripe fruits (30% dry weight) was used (245 g tissue), whereas in the extract from orange the peel of 4 ripe fruits was used (200 g tissue). The extraction protocol was then followed as described in section 2.2.

2.4.2 Anion Exchange Chromatography Using DEAE Sepharose

Proteins preparations were resuspended in 10 v/w 20 mM potassium phosphate buffer, pH 7.2 and dialysed overnight against this buffer before being centrifuged (10,000 g, 15 min, 4°C). The supernatant was loaded onto a pre-equilibrated DEAE sepharose anion exchange column (43 ml) in the above buffer at 4 ml/min. The column was washed with 100 ml of phosphate buffer before proteins were eluted with a linearly increasing concentration of salt (0-0.5 M NaCl) over 200 ml. Following an 80 ml wash with phosphate buffer + 0.5 M NaCl a pulse of high salt phosphate buffer (2 M, 5 ml) was passed through the column to elute any remaining bound protein. Fractions (8 ml) were collected and 100 µl aliquots assayed for thioesterase activity using the

Ellman's method (Section 2.3.1). Active fractions were pooled and applied to the following column or precipitated with $(NH_4)_2SO_4$ and stored at -20°C.

2.4.3 Hydrophobic Interaction Chromatography Using Octyl Sepharose

Active fractions eluting from the DEAE sepharose column were pooled and $(NH_4)_2SO_4$ added to a final concentration of 1 M prior to loading onto an octyl sepharose column (47 ml) pre-equilibrated in 50 mM phosphate buffer, pH 7.2 plus 1 M $(NH_4)_2SO_4$ at 4 ml/min. The column was washed with 100 ml of this buffer before proteins were eluted with a linearly decreasing concentration of $(NH_4)_2SO_4$ (1-0 M) over 200 ml. Following an 80 ml wash with phosphate buffer a pulse of ethylene glycol (5 ml) was passed through the column to elute any remaining bound protein. Fractions (8 ml) were collected and 100 µl aliquots were assayed for thioesterase activity using the Ellman's method. Active fractions were pooled and applied to the following column or precipitated with $(NH_4)_2SO_4$ and stored at -20°C.

2.4.4 Hydrophobic Interaction Chromatography Using Phenyl Superose

Pools of active proteins eluting from the octyl sepharose column were concentrated down to a 10 ml volume through ultrafiltration using centrifugal Vivaspin columns (Sartorius; at 3900 g for 30 min). The concentrated protein in 1 M $(NH_4)_2SO_4$ was then applied to a high-performance phenyl superose column (0.69 ml) in 50 mM phosphate buffer + 1 M $(NH_4)_2SO_4$ at 0.5 ml/min. The column was washed with 5 ml of the above buffer before proteins were eluted with a linearly decreasing concentration of $(NH_4)_2SO_4$ (1-0 M) over 20 ml. Fractions (1 ml) were collected and 20 µl aliquots were assayed for thioesterase activity using the Ellman's method. This column provided sufficient resolution of proteins for the purification of thioesterases from passion fruit and pooled fractions were kept on ice during protein characterisation or precipitated with $(NH_4)_2SO_4$ before storage at -20°C. When purifying thioesterases from orange a further Mono S column purification step was required.

2.4.5 Cation Exchange Chromatography Using Mono S

Active fractions eluting from the phenyl superose column were pooled and desalted through overnight dialysis against 20 mM Bis-Tris buffer, pH 7.0. Due to the small sample volume (2 ml) dialysis was conducted using a Slide-A-Mini dialysis unit (Perbio, formerly Pierce). The sample was then applied to a high-performance cation exchange Mono S column in the above buffer at 1 ml/min. Following a wash with 5 ml of Bis-Tris buffer, proteins were eluted with a linearly increasing concentration of salt (0-0.5 M NaCl) over 20 ml. Fractions (1 ml) were collected and 20 μl aliquots were assayed for thioesterase activity using the Ellman's method. Pooled fractions were kept on ice during protein characterisation or precipitated with $(NH_4)_2SO_4$ and stored at -20°C.

2.4.6 Gel Filtration Chromatography Using Superdex 200

In characterizing the thioesterase from passion fruit, a gel filtration column was used to determine the molecular mass of the native protein. A semi-purified fraction of activity retained following octyl sepharose chromatography (10 ml) was applied to a superdex gel filtration column (7.8 ml) in 50 mM phosphate buffer, pH 7.2, + 0.15 M NaCl at a flow rate of 0.5 ml/min. 0.5 ml fractions were collected and 20 μl aliquots were assayed for thioesterase activity. In addition, 200 μl samples of protein standards (1 mg/ml); cytochrome C, ovalbumin and bovine serum albumin made up in the same buffer were applied to the column in subsequent runs as above.

2.4.7 Glycoprotein Interaction Chromatography Using Concanavalin A Sepharose

In determining the post translational modification of the thioesterase from passion fruit, concanavalin A sepharose chromatography was used. A concanavalin A sepharose 4B column was packed (10 ml) and following a pre-wash with 5 volumes of 1 M NaCl, 5 mM $MgCl_2$, 5 mM $MnCl_2$, 5 mM $CaCl_2$, the column was equilibrated with 20 mM Tris-HCl, pH 7.4 containing 0.5 M NaCl. A 1 ml aliquot of active protein eluting from the phenyl superose column was dialysed overnight in 20 mM Tris-HCl, pH 7.4 using the mini dialysis unit. Following the addition of 0.5 M NaCl to the protein sample it was applied to the

column in the same buffer at 1 ml/min. Following a 10 ml wash protein was eluted using a linearly increasing concentration of methyl-α-D-mannopyranoside (0-0.1 M). A pulse of 0.5 M methyl-α-D-mannopyranoside (5 ml) was used to release any remaining bound protein. 1 ml fractions were collected and assayed for thioesterase activity.

2.5 Protein Analysis

2.5.1 SDS-PAGE

SDS-PAGE (sodium dodecyl sulphate-polyacrylamide gel electrophoresis) slab gels were prepared using the Mini-PROTEAN 3 kit (BIO-RAD) according to the accompanying cell assembly guide. The resolving gel was polymerised from 10% or 12% acrylamide *bis*-acrylamide in 375 mM Tris-HCl, pH 9.0 containing 0.1% v/v TEMED (Promega), 0.1% w/v SDS and 0.1% w/v $(NH_4)_2SO_4$. 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 $(NH_4)_2SO_4$. Gel cassettes were assembled in a Mini-PROTEAN 3 electophoresis module according to the manufacturer's guidelines. Protein samples were diluted with an equal volume of 2 X SDS loading buffer (100 mM Tris-HCl, pH 6.8, 20% v/v 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 onto gels. The samples were 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.

2.5.2 Two-Dimensional SDS-PAGE

The purified thioesterase from passion fruit was analysed by two-dimendional (2D) SDS-PAGE, with all reagents and equipment obtained from GE-Healthcare. Protein samples (1 ml) were concentrated 10-fold using a centrifugal evaporator (Jouan RC 1022) to ensure sufficient protein was present in the 100 μl sample used for analysis (~100 μg). The sample was passed through a mini desalting spin column (Sephadex G-25, Sigma) to remove any salt and other impurities which may have interfered with

separation. Subsequently the protein sample was pre-treated with an equal volume (100 µl) of rehydration solution (0.5% IPG buffer, 9 M urea, 2 M thiourea, 4% w/v CHAPS, 1% w/v DTT and 2% v/v pH 3-10 ampholytes) and pipetted evenly along the well of an Immobiline Drystrip rehydration tray. An Immobiline DryStrip pH 3-10 (11 cm) was placed over the solution gel side down, covered with mineral oil and rehydrated overnight. The rehydrated strip was washed with water, blotted dry and placed in the running tray of a MultiPhor II electrophoresis unit (Pharmacia). The cooling plate was set to 25°C and moistened IEF electrode strips were placed over the DryStrip at the anode and cathode. Electrodes were placed over the IEF electrode strips and mineral oil was poured into the tray to completely cover the strip. The gel was run at 500 V for 1 hr and then at 3500 V for a further 3 hours. The DryStrip was then equilibrated in 2 stages, firstly for 10 min in 1 M Tris-HCl, pH 6.8, 10 mg/ml DTT and secondly for 10 min in 1 M Tris-HCl pH 6.8, 6 M Urea, 30% v/v glycerol, 10% w/v SDS, 45 mg/ml iodoacetamide plus a trace of bromophenol blue dye. Proteins were then run in the second dimension using standard SDS-PAGE. A 1 mm slab gel (12%) was used to accommodate for the thickness of the DryStrip which was blotted dry and inserted across the top of the resolving gel. The gel was sealed with agarose sealer (0.1% low melting point agarose in SDS-PAGE running buffer containing a trace of bromophenol blue) and run as described in section 2.5.1.

2.5.3 Western Blotting and Immunodetection

Western blotting was used to detect recombinant streptactin fusion proteins in bacterial extracts. The soluble and insoluble protein fractions resulting from the expression of recombinant protein in bacteria (Section 2.8) were separated using standard SDS-PAGE and then electroblotted onto a PVDF membrane (Hybond-P, GE-Healthcare) using a mini Trans-Blot cell (BIO-RAD) according to manufacturer's instructions. After blotting the membrane was rinsed in Tris-buffered saline (TBS) 0.12% w/v Tris, 0.88% w/v NaCl and then blocked for 1 hr in TBS 3% skimmed milk powder. Anti Streptactin antibody, Strep-Tactin alkaline phosphatase conjugate (IBA BioTAGnology, Germany) was added at a 1: 5000 dilution and incubated for 1-2 hr at room temperature or overnight at

4°C. The membrane was subsequently washed twice for 20 min in TBST (TBS plus 0.1% Triton X-100) and then for a further 20 min in TBS. No secondary antibody was required. The membrane was then rinsed in 100 mM Tris, pH 9.5 and developed in the dark in 3% v/v 5-bromo-4-chloro-3-indolyl phosphate (BCIP; 50 mg/ml) dissolved in *N,N,N',N'*-dimethylformamide (DMF) and 0.3% v/v nitro blue tetrazolium (NBT; 100 mg/ml) dissolved in 70% DMF. The reaction was stopped by rinsing the membrane in a large volume of water.

2.5.4 Protein Staining

Different protein stains were used dependent upon the level of sensitivity required. SYPRO Ruby protein gel stain was the preferred choice due to its high degree of sensitivity (minimum detection limit of 1-2 ng protein per gel band) and its ease of use. Coomasie staining with its minimum detection limit of 10 ng of protein per band was used to ensure sufficient protein was in gel bands used for protein identification through mass-spectrometry. In all cases a sufficient volume of solution was used to completely cover the gel (25-30 ml) with incubation steps carried out under mild agitation on a rocker. Steps carried out in the dark required covering the tray with aluminium foil.

2.5.4.1 SYPRO Staining

Following electrophoresis SDS gels were fixed in 40% methanol, 10% acetic acid for 1 hour (not required for 1D gels). Gels were then submerged in SYPRO Ruby protein gel stain (Invitrogen) and incubated in the dark for 4 hours-overnight. Following incubation gels were washed twice in 10% methanol, 7% acetic acid for 1 hour before imaging under a Gel Doc UV transilluminator (BIO-RAD).

2.5.4.2 Coomassie Staining

SDS gels were washed thoroughly with water (2 x 15 min) under gentle agitation to remove the SDS and then submerged in Coomassie blue reagent (0.006% w/v coomassie G250, 5% ethanol, 10% v/v phosphoric acid) until protein bands were visible.

2.5.4.3 Krypton Glycoprotein Staining

Glycoprotein staining of SDS gels was carried out using the Krypton glycoprotein staining kit (Perbio). Following the separation of proteins by standard SDS-PAGE, gels were subsequently rinsed thoroughly in ultrapure water (5 x 15 min) and transferred to fixing solution (40% ethanol, 30% glacial acetic acid) for 2 x 30 min incubation. Gels were soaked in wash solution (15% glacial acetic acid) for 10 min. Prior to staining, it was necessary to oxidise the carbohydrate residues on proteins by soaking in oxidizing solution (supplied with kit) in the dark for 20 min. Gels were subsequently rinsed twice in wash solution for 10 min before adding the Krypton stain solution and incubated in the dark for 1 hr (2 hrs for 2D gels). Gels were rinsed in ultrapure water and before the destaining solution (supplied with kit) was added and incubated for 20 min in the dark. The gels were then imaged using an Advanced Image Data Analyzer (AIDA) FLA-3000 UV transilluminator (Fujifilm) exciting at 654 nm and reading the emission at 673 nm. Following staining and imaging gels were destained in 3% acetic acid for 5 min and then stained with SYPRO Ruby (Section 2.5.4.1) for the comparison of total protein.

2.5.5 Protein Inhibition Studies

Protein samples dissolved in 0.1 M potassium phosphate buffer pH 7.2 were incubated with 1 mM inhibitors (EDTA, paraoxon, iodoacetamide, mercuric chloride, fp-biotin (10 μM)) for 2 hours at 30°C. Samples were subsequently desalted using Sephadex G-25 spin columns to remove unbound inhibitor and assayed using the Ellman's method. Control reactions with acetone (-inhibitor) were run in parallel.

2.5.6 Protein Labelling and Visualisation Using the Fluorophosphonate Trifunctional Probe (TriFP)

TriFP was synthesized as described previously (Gershater *et al.*, 2007)[125] and made available for use in our laboratory. 100 μl aliquots of the crude protein extract from passion fruit (dialysed fraction) and the purified thioesterase (phenyl superose fraction) were incubated with 10 μM FPP for one hour at room temperature. The solution was then passed through a PD-10 desalting

column (GE-Healthcare) to remove any excess probe before resolving peptides by SDS-PAGE. The fluorescently labelled proteins in the gel were visualised using the Fujifilm FLA-3000, exciting at 532 nm and reading the emission at 580 nm. Gels were subsequently destained in 3% acetic acid for 5 min and then restained with SYPRO Ruby for the comparison of total protein.

2.5.7 Protein Identification Using MALDI-ToF Mass Spectrometry

Bands of protein excised from SDS gels were digested with trypsin and analysed by MALDI-ToF-based proteomics on a Voyager-DE STR BioSpectrometry Workstation (Applied Biosystems, Warrington, UK), using the Durham University proteomic service. The resulting calibrated ion spectra generated a peak list which was used to screen non-redundant protein and EST databases using Mascot software (www.matrixscience.com).

2.5.8 Protein Sequencing Using Tandem Mass Spectrometry

Peptides were extracted from SDS-PAGE gel slices prior to trypsin digestion. Gel plugs were placed in sterile eppendorf tubes and equilibrated with 100 µl of 25 mM ammonium bicarbonate. Samples were reduced with 30 µl of 10 mM DTT for 30 min followed by alkylation with 30 µl of 100 mM iodoacetamide for 15 min. The gel was subsequently destained and dessicated with 2 x washes in 50 µl of acetonitrile. Gel slices were subsequently soaked for 1 hr 30 min in 15 µl of 25 mM ammonium bicarbonate containing 6.6 w/v trypsin (Promega). 10 µl of water was added to rehydrate the gel slice and left to incubate overnight. The gel slice was rinsed in 10 µl of 25 mM ammonium bicarbonate for 10 min before peptides were extracted in 20 µl of 10% formic acid for 10 min. The solution was retained and 2 further extractions carried out, firstly with 20 µl of acetonitrile for 15 min and secondly with 30 µl of acetonitrile for 15 min. Fractions were pooled and freeze dried before resuspending the digested peptides in 10 µl of 0.1% formic acid. Samples were submitted for MS-MS *de novo* proteomic sequencing on an ion trap mass spectrometer (Thermo LTQ FT) using the chemistry department mass spectrometry sevice (Durham, UK). The resulting peptide mass fingerprints were used to determine partial amino

acid sequence through the automated *de novo* sequencing software DeNovoX (Thermo) and checked by manual calculation for confirmation of sequence.

2.5.9 Protein Bioinformatics

Partial amino acid sequence identified through tandem mass spectrometry was used to interrogate the plant genome and EST databases using short match BLAST (NCBI; www.ncbi.nlm.nih.gov/blast/Blast.cgi). The resulting protein sequence matches were aligned using CLUSTALW (Thompson *et al.*, 1994)[126]. Bioinformatic analysis of proteins was carried out using BRENDA software (www.brenda-enzymes.info/) to obtain data on the isoelectric points and molecular masses of homologous enzymes identified through BLAST searches. In addition, further bioinformatic analysis on proteins was carried out using the ExPASy proteomic tools (www.expasy.ch/tools/) such as SignalP 3.0 for the determination of predicted signal peptide cleavage sites and NetNGly 1.0 for the prediction of glycosylation sites in proteins.

2.6 Gene Cloning

Custom oligonucleotides were obtained from MWG-biotech AG (Ebersberg, Germany). The sequences of the specific primers used in each PCR reaction are detailed in Chapter 5. All reagents and enzymes were supplied by Promega unless stated otherwise.

2.6.1 mRNA Extraction

Buffers and water used in the extraction of RNA were autoclaved at 120°C for 15 min to ensure they were RNAase free. For the same reason sterile plastic ware was used and the pestle and mortar baked in an oven at 200°C overnight. Passion fruit total RNA was extracted from mesocarp tissue, foliage and secretary glands using TRI Reagent (Sigma) as described in the accompanying protocol. Samples of passion fruit foliage were obtained from plants maintained in the glass-houses of the University botanical gardens (Durham, UK). 100 mg of each tissue was processed with 2 ml of TRI Reagent and the resulting RNA pellet was washed with 75% ethanol before being

stored in ethanol at -20°C or dried and redissolved in 10 µl of ultrapure sterile water for further processing. The concentration and quality of the RNA was determined by measuring the $OD_{260/280}$ ratio ($OD_{260/280}$ ratio > 1.7: pure RNA without DNA contamination, OD_{260} 1.0 = 40 µg/ml RNA). RNA samples were mixed with RNA sample loading buffer (50% glycerol, 1 mM EDTA, 0.25% bromophenol blue, 0.25% xylene cyanol) in a 5:1 ratio, heated to 65°C for 10 min and analysed by agarose gel electrophoresis (Section 2.6.5).

2.6.2 Synthesis of cDNA

Total RNA (5 µg) was mixed with 20 pmols of either an oligo dT primer (Og2) or a specific primer and heated to 65°C for 10 min before immediately cooling on ice. The following reagents were added; 4 µl 5x AMV (Avian myeloblastoma leukaemia virus) reverse transcriptase buffer, 8 µl dNTP mix (1 mM each), 0.5 µl RNasin (20 U), 1 µl AMV reverse transcriptase and ultrapure water to a total volume of 20 µl. Reverse transcription was then carried out at 42°C for 1 hr.

When amplifying the 5' end of truncated cDNA through 5'-RACE (Section 2.6.4) it was necessary to increase the reverse transcription temperature to 45-50°C in order to promote the denaturation of RNA. The more heat stable reverse transcriptase RevertAid M-MuLV (Fermentas) was therefore used. 5 µg of total RNA was mixed with 10 pmols of primer plus water to 12.5 µl before heating to 70°C for 5 min. The following reagents were added; 4 µl 5x reverse transcriptase buffer (supplied with enzyme), 2 µl dNTP mix (1 mM each), 0.5 µl RNasin (20 U), and ultrapure water to a total volume of 19 µl. The sample was heated at 37°C for 5 min before 1 µl of M-MuLV reverse transcriptase was added and incubated at 45°C/50°C for 1 hr. The reaction was stopped by heating to 70°C for 10 min before cooling on ice.

2.6.3 Polymerase Chain Reaction (PCR)

Standard 50 µl PCR reactions were performed in 0.2 ml PCR tubes. The following reagents were added to the reaction mix;

4.5 µl 11x PCR buffer (0.5 M Tris-HCl pH 8.8, 125 mM $(NH_4)_2SO_4$, 25 mM $MgCl_2$, 3 µM mercaptoethanol, 50 µM EDTA, each dNTP at 2.25 mM, 1.25 mg/ml BSA)

5 µl 10 mM forward primer (50 pmols)

5 µl 10 mM reverse primer (50 pmols)

1 µl *Taq* DNA polymerase

1-5 µl template DNA

The reaction mixture was made up to 50 µl with water and subjected to a standard PCR programme as detailed in table 2.1, unless otherwise stated, using a Mastercycler Gradient PCR machine (Eppendorf, Hamburg, Germany).

Step	Temperature (°C)	Duration (min:sec)
1. Denature	94	2.00
2. Denature	94	0:15
3. Anneal	55	0:45
4. Extend	72	2:00
	Steps 2-4 repeated over 25 cycles	
5. Extend	72	5:00
	Hold 4°C, End	

Table 2.1 The standard PCR Programme used for the amplification of DNA

For PCR requiring the use of a proof reading DNA polymerase, KOD HiFi DNA polymerase was used (Novagen). A 50 µl reaction mix was made up with 1 Unit (0.4 µl) of enzyme (1 unit will catalyze the incorporation of 10 nmol of dNTP into acid insoluble form in 30 min at 75°C), 20 pmols primers, a suitable quantity of template, 20 µM dNTPs, 2 mM $MgCl_2$ and 5 µl 10 x KOD HiFi DNA polymerase buffer. PCR amplification conditions were used as above.

2.6.4 5' RACE (Rapid Amplification of cDNA Ends)

5' RACE was used to extend the 5' end of truncated cDNAs, the exact sequence of primers used are illustrated in chapter 5. Two antisense oligonucleotide primers were designed from the 5' end of truncated cDNAs; primer 1 (41 bp) and nested primer 2 (24 bp). cDNA was synthesized from 5 μg of total RNA using primer 1 and RevertAid M-MuLV reverse transcriptase as detailed in section 2.6.2. The cDNA template was then tailed at the 5' end with polyA in the following reaction. Total synthesized cDNA (17 μl) was mixed with 2.5 μl dATP (2 mM), 2.5 μl terminal transferase buffer (New England Biolabs, USA) and water to a total volume of 24.5 μl before heating to 94°C for 3 min. 0.5 μl (10U) of terminal transferase (New England Biolabs) was added and incubated for 15 min at 37°C before heat inactivation (70°C, 10 min). The tailed cDNA (5 μl) was then amplified by PCR using primer 1 and an oligo dT adaptor primer (5'- CTT ATA CGG ATA TCC TGG CAA TTC GGA CTT TTT TTT TTT TTT TTT AGC –3') with 35 cycles at an annealing temperature of 55°C. A 1:20 dilution of the PCR product was prepared and then 5 μl subjected to a second round of PCR using an adaptor primer (5'- CTT ATA CGG ATA TCC TGG CAA TTC GGA CTT -3') and the nested primer 2. PCR conditions as above. PCR reaction products were resolved by agarose gel electrophoresis (Section 2.6.5) and following purification of DNA from the gel slice (2.6.6) products were directly ligated into pGEM-T Easy vector (Promega) (2.6.8), cloned in Top 10 competent cells (2.6.10) before plasmid preparations were made (2.6.11) and sent for DNA sequence analysis (2.9).

2.6.5 DNA Agarose Gel Electrophoresis

Agarose gels were prepared by microwaving 50 ml of TAE (4.84% w/v Tris, 1.14 v/v glacial acetic acid and 1% molecular biology grade agarose (Helena BioSciences) until the agarose had melted. The gel mix was cooled to approximately 60°C and 0.5 μl of ethidium bromide was added before pouring into a UV-transparent plastic casting tray (BIO-RAD). 6X loading buffer (0.25% w/v bromophenol blue, 0.25% v/v xylene cyanol, 15% w/v Ficoll) was added to DNA samples in a 1:5 ratio, mixed and loaded onto the gel (5-20 μl). DNA markers (1 Kb DNA ladder, Invitrogen) were added to an adjacent lane for

reference prior to electrophoresis in TAE buffer at 120 V for 20 min using the Sub-Cell GT agarose gel electrophoresis system (BIO-RAD).

2.6.6 Purification of DNA

PCR products and digested DNA samples were resolved on agarose gels and the desired products excised. Gel slices were placed in sterile tubes and dissolved in 500 µl of binding buffer (6 M Sodium Perchlorate, 50 mM Tris-HCl, 10 mM EDTA, pH 8.0) at 55°C for 10 min. 10 µl of resuspended Silica fines (166 mg/ml in water, World Mineral Ltd, Hessle, UK) were added, the sample vortexed and incubated for 5 min at room temperature. The solution was microfuged for 1 min at maximum speed and the resulting pellet washed in 0.4 M NaCl, 2 mM EDTA, 2 mM Tris-HCl pH 7.5. Following a second wash in 70% ethanol the fines were dried, resuspended in 15 µl of ultrapure water and incubated at 50°C for 5 min to solubilise the DNA. The sample was centrifuged for 1 min to remove the fines and the supernatant containing DNA was transferred to a clean tube.

2.6.7 Restriction Digests

PCR products and plasmid DNA were digested with various restriction enzymes obtained from Promega. 2-10 Units of enzyme(s) were added to the sample (0.5-2 µg DNA) with the appropriate buffer to 20 µl volume and incubated at 37°C for 1-3 hours. The digests were analysed by agarose gel electrophoresis and the desired products excised.

2.6.8 Ligation

PCR products produced using *Taq* polymerase were directly cloned into the pGEM-T Easy vector (Promega). The following were mixed in a sterile tube; 1 µl T4 DNA ligase, 1 µl (50 ng) pGEM-T Easy vector, 5 µl 2 x ligation buffer, 1-3 µl purified PCR product to a total volume of 10 µl with water. The reaction was incubated for 1 hour at room temperature or 4°C overnight.

For cloning into other vectors, such as pET Strep, digested vector and insert DNA were mixed in a 3:1 molar ratio with 1 µl of 10 x ligation buffer (as

supplied with the enzyme) and 1 µl (3 U) of T4 DNA ligase (Promega) in a reaction volume of 10 µl and incubated at 17°C overnight.

2.6.9 Bacterial culture

Luria Bertani (LB) medium (1% w/v NaCl, 1% w/v tryptone, 0.5% w/v yeast extract) was used for liquid culturing of bacteria. Starter cultures (10 ml) were inoculated with single colonies picked from LB agar plates and grown overnight at 37°C on an orbital shaker (Kuhner) at 200 rpm with antibiotic if required. Bacterial cultures were grown until an adequaute cell density was reached (OD_{600}: 0.6-0.7). Bacteria grown on LB agar plates (LB medium plus 2% agar) were incubated overnight at 37°C with the appropriate antibiotic if required; ampicillin; 100 µg/ml, chloramphenicol; 34 µg/ml, tetracycline; 12.5 µg/ml, kanamycin; 50µg/ml, carbenicillin; 50µg/ml, rifampicin; 50µg/ml, gentomicin; 26µg/ml. In the construction of cDNA libraries SM buffer (5.8% NaCl, 0.2% $MgSO_4$, 5% (v/v) 1M Tris-HCl pH 7.5, 0.5% gelatin) was used for the dilution and storage of phage solutions.

2.6.10 Transformation

Transformation of chemically competent *E. coli* cells (One shot Top 10 cells, Invitrogen) was used for cloning PCR products in the pGEM-T Easy vector. 25 µl of competent cells were incubated on ice with 1 µl of plasmid (ligation mix) for 5 minutes. The cells were then heat-shocked at 42 °C for exactly 30 seconds before being returned to ice for a further 2 min. Cells were mixed with 150 µl LB medium and incubated at 37°C for 1 hr before plating 20-200 µl on LB agar containing ampicillin (100 µg/ml).

For transformation prior to expression Tuner DE3 pRARE (Tunette) E.coli cells were used. Transformation was carried out as for Top 10 cells except kanamycin (50 µg/ml) and chloramphenicol (34 µg/ml) selection was used.

2.6.11 Plasmid Preps

10 ml cultures of transformed bacteria were grown overnight (OD_{600}: 0.6-0.7) and the cells pelleted through centrifugation at 3900 g for 10 min. Plasmid

DNA was subsequently purified from the cells using the Wizard Plus SV miniprep DNA purification system (Promega) according to manufacturer's instructions.

2.7 cDNA Libraries

Passion fruit cDNA libraries were prepared from mesocarp and secretary gland tissue (100 mg) taken from ripe fruit (10 days ripening period). Total RNA (0.25 μg) was used to construct a cDNA library in bacteriophage λ TriplEx2 vector using the SMART cDNA library construction kit (BD Biosciences). Reagents and *Escherichia coli* bacterial strains (XL1-Blue and BM25.8) were supplied with the kit.

2.7.1 Preparation of Phage Libraries

First-strand cDNA was synthesised using LD-PCR (long distance-PCR)[127] to ensure full-length cDNA templates were produced. 3 μl of total RNA (0.25 μg) was mixed with 1 μl of the SMART IV oligonucleotide primer (10 μM) and 1 μl of the CDS III/3' PCR primer (10 μM), both containing Sfi I restriction sites. The sample was incubated for 2 min at 72°C before immediately cooling on ice. The following was added; 2 μl first strand buffer, 1 μl DTT (20 mM), 1 μl dNTP mix (10 mM), 1 μl PowerScript reverse transcriptase and incubated for 1 hr at 42°C. First-strand cDNA was subsequently amplified by LD-PCR in a 100 μl volume reaction. 2 μl of first strand cDNA was mixed with 2 μl of 5' PCR primer (10 μM), 2 μl of CDS III/3' primer in a standard PCR mix with 2 μl of 50 x Advantage 2 polymerase mix (containing *Taq* polymerase). cDNA was amplified using 24 PCR cycles with an annealing temperature of 68°C. A 5 μl sample of amplified cDNA was analysed by agarose gel electrophoresis.

50 μl of amplified cDNA (3 μg) was digested with 2 μl of proteinase K (20 μg/μl) for 20 min at 45°C. cDNA was extracted using a phenol, chloroform, isoamylalcohol mix as described in the user manual. The cDNA pellet was

resuspended in 79 μl of water and digested with SfiI (10 μl 10 x Sfi I buffer, 10 μl Sfi I enzyme, 1 μl 100 x BSA) for 2 hrs at 50°C.

DNA was dyed by adding 2 μl of 1 % cyanol dye before being fractionated on CHROMA SPIN-400 columns (supplied with kit). The column matrix was resuspended and conditioned with 700 μl of column buffer before applying sample (100 μl). Following adsorption of the sample cDNA was eluted with column buffer (600 μl), and single drip fractions (~40 μl) were collected and analysed by agarose gel electrophoresis, removing small cDNA fragments (<0.4 kb) and unincorporated primers. Fractions 4-7 were pooled and DNA precipitated. The digested cDNA pellet was resuspended in 7 μl of water. A 1.5 μl aliquot was mixed with 1 μl λ TriplEx2 vector, 0.5 μl 10 x ligation buffer, 0.5 μl ATP (10 mM), 0.5 μl T4 DNA ligase and water. The ligation mix was incubated at 16°C overnight. Vector was packaged into E.coli XL1-blue using the Gigapack III gold packaging extract (Stratagene). 25 μl of packaging extract was mixed with vector (2 μl) and incubated at room temperature for 2 hrs. 500 μl of SM buffer, 20 μl chloroform was added and mixed before microfuging at full speed for 30 sec. The supernatant containing phage was titered.

The phage packing extract was diluted (5 x, 10 x, 20 x) in SM buffer and 1 μl added to 200 μl of XL1-Blue bacteria (grown overnight, OD_{600}: 2) and left to adsorb for 15 min at 37°C. 2 ml of melted LB top agar (45°C) was added and poured onto 90 mm LB/$MgSO_4$ agar plates at 37°C. After setting the top agar at room temperature for 10 min the plates were inverted and incubated at 37°C overnight. Plaques were counted and the following titers determined:

endocarp library: 640000 pfu (plaque forming units)/ml

secretory gland library: 540000 pfu/ml

2.7.1.1 Amplification of Phage Libraries

300μl of unamplified library was mixed with 2 ml of overnight XL1-Blue bacterial culture, incubated at 37°C for 15 min before 22 ml of melted top agar (45°C) was added, mixed and poured onto large (590 cm^2) plates. Growth as

above. 33 ml of SM buffer was poured onto the plates and incubated at 4°C with gentle rocking overnight. The solution containing bacteriophage was poured into a sterile tube plus 5% chloroform and mixed before centrifuging (5000 g, 10 min). The supernatant containing the amplified libraries was aliquoted (1 ml) into sterile tubes plus 7% DMSO and stored at -80°C. The amplified library was again titered:
endocarp library: 2.6×10^8 pfu/ml
secretary gland library: 1.6×10^8 pfu/ml

2.7.2 Preparation of Plasmid Libraries

A plasmid library was constructed by converting the λ TriplEx2 clones into pTriplEx2 plasmids through *in vivo* excision and circularisation of a complete plasmid in BM25.8 bacteria expressing Cre recombinase activity. 150 μl of BM25.8 bacterial cell culture (OD_{600}: 1.4) was mixed with 15 μl of unamplified libary (or 200 μl cells + 20 μl gland library) and incubated at 31°C for 30 min. The culture was diluted in 20 ml of LB medium and incubated for a further hour at 31°C with shaking (Cre recombinase activity is expressed at 31°C). Bacteria were pelleted through centrifugation (5000 g, 10 min) before the pellet was resuspended in 20 ml of SM buffer and aliquoted (1 ml) into sterile tubes plus 15% glycerol and stored at -80°C.

The plasmid libraries were titered by diluting (10 x, 100 x) in SM buffer and growing 10 μl on agar plates plus carbenicillin (50 μg/ml) overnight at 37°C.
endocarp library: 1.6×10^7 cfu (colony forming units)/ml
secretary gland library: 1.12×10^7 cfu/ml

2.7.3 Library Screening

Plasmid libraries (1 μl) were diluted 100 fold in SM buffer and plated out on large agar plates (490 cm^2) plus carbenicillin (50 μg/ml) and grown overnight at 37°C. 50 colonies were picked at random from each library and grown in 10 ml cultures (plus carbenicillin) overnight at 37°C. Plasmid preps were made (Section 2.6.11) and inserts sequenced using forward and reverse pTriplEx2 sequencing primers (supplied with the kit).

2.8 Recombinant Expression - Bacteria

Single colonies of transformed Tunette bacteria (Section 2.6.10) were used to inoculate 10 ml mini starter cultures containing kanamycin (kan, 50 μg/ml) and chloramphenicol (cam, 34 μg/ml). Following overnight growth at 37°C with shaking a 1 ml aliquot was used to inoculate a larger 100 ml LB kan/cam culture. Cells were grown as before until dense (OD_{600}: 0.7), when 0.1-1 mM IPTG was added to induce expression over 3 hours. Different growth temperatures (15-37°C) and +/- IPTG induction was used to optimise expression. Following induction and expression cells were centrifuged at 3900 g for 5 min and the pellet resuspended in 2 ml of Tris-HCl, pH 7.5, achieving a 50 x concentration. Samples were sonicated 3 times at 30% intensity for 20 seconds using an HD 2070 sonicator (Bandelin), microfuged at full power for 5 min and the supernatant (soluble protein) collected. The pellet was resuspended in 1 ml of Tris-HCl, pH 7.5, vortexed and stored on ice (insoluble protein). Both soluble and insoluble fractions were analysed for recombinant expression by western blotting (Section 2.5.3) and assaying for thioesterase activity using the Ellman's method.

2.9 DNA Sequencing and Analysis

Double stranded cDNAs were sequenced using an Applied Biosystems 3730 DNA Analyser by the University of Durham sequencing service. DNA sequences were edited, translated 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, Centre for Bioactive Chemistry, University of Durham, UK). DNA and protein sequences were aligned using CLUSTALW and sequence similarities were determined using BLAST. RNA secondary structure predictions were determined using Mfold version 3 (www.mfold.bioinfo.rpi.edu/) as developed by M.Zuker, (2003)[128].

2.10 Bioreactions

Commercially available immobilized enzymes were supplied by Sigma; Lipozyme (lipase from *Mucor miehei* immobilized on a macroporous ion-exchange resin) and Novozyme (lipase from *Candida antarctica* immobilized on acrylic resin).

2.10.1 Preparation of Fruit Enzymes

Orange or passion fruit peel was homogenized in a waring blender with 4 v/w 0.1 M potassium phosphate buffer, pH 7.2 and subsequently dried using an LP3 freeze drier (Jouan) overnight. Dried material was crushed to a fine powder with a pestle and mortar and stored at -20°C.

2.10.2 Reaction Conditions

Bioreactions were conducted on a 50 ml scale in round bottomed flasks at 37°C with stirring. 100 mg of fruit enzyme preparation or immobilized enzymes were used in each reaction, unless specified otherwise. Further specific reaction conditions are described in chapter 6.

2.10.3 Gas Chromatography Mass Spectrometry (GCMS) Analysis

500 µl aliquots of reaction products were partitioned into an equal volume of dichloromethane and dried over sodium sulphate. 5 µl samples were analysed using a Thermo-Finnigan Trace GCMS (Thermo) with a Phenonemex Zebron ZB-5 capillary GC column (30 m x 0.25 mm), by the department of chemistry analytical services (Durham, UK). The following parameters were used; splitless injection, injector temperature: 200°C, Initial oven temperature: 80°C, ramp rate 2°C/min to a maximum temperature: 200°C, carrier gas: Nitrogen, column pressure 30 kPa.

3. Identification and Characterisation of Thioesterases in Tropical Fruit

3.1 Introduction

Volatile organosulphur chemicals (VOSCs) make a major contribution to flavour and fragrance in tropical fruits, where they are found as free thiols or the respective thioester derivatives[24]. The proportion of each is determined by the action of thioesterases which are believed to be important in the generation of ripe fruit flavour. As has been described, these thioesterase enzymes from exotic fruits form a novel source of potential biocatalyst for use in the manufacture of high-impact flavourings. The search for such unknown enzymes in fruit must however begin with the development of a sensitive and simple assay for determining thioesterase activity in fruit tissues.

VOSCs being high impact aromas, are subsequently found at only trace levels in fruit tissue, for example the major character impact volatile in passion fruit, 3-mercaptohexanol (passion fruit mercaptan), is present at less than 10μg/kg passion fruit[24] but is still able to make a major contribution to overall fruit aroma. When handling fruit VOSCs in their pure form they have an attractive fruity note at low concentrations (10-100 ppb), any higher and they smell sulphury, unpleasant and are hard to distinguish. Fruit therefore generate trace levels of these metabolites with flux through their pathways of biosynthesis low and hence difficult to investigate. The enzymes involved in VOSC biosynthesis may subsequently be found to have low total activities (e.g. high specific activities but with low protein abundance) and a careful choice of assay to determine activity was required. Simply determining thioesterase activity in tropical fruits was likely to be a challenge and for the subsequent purification of the associated protein(s) the assay of choice had to be sensitive and rapid.

The properties of thioesterases in tropical fruit have not been studied before and it was therefore important to determine the location of activity within fruit (e.g. peel, juice, pith) and its regulation during maturation. Developing fruit undergo many changes both at the physiological level (colour, texture, F&F) and at the genetic and biochemical level in terms of gene regulation, protein

synthesis, respiration, cell wall degradation, secondary metabolite accumulation and nutrient depletion[43]. Essentially, many developmental changes occur over a short period (1-2 weeks) and it was therefore important to monitor thioesterase activity over the full course of ripening so as to determine when activity was greatest.

In order to monitor thioesterase activity, an assay for determining the hydrolysis of thioester VOSCs was required (Figure 3.1). This could be achieved through quantifying the loss of starting materials (thioester) or the generation of products (thiol or acid).

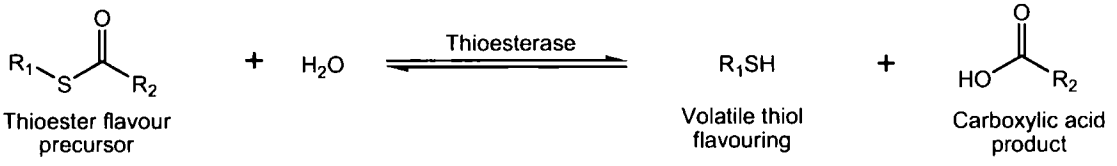


Figure 3.1 The enzymatic hydrolysis of thioesters. It was necessary to develop an assay for determining the extent of this reaction. Although this is a reversible reaction, *in vitro* assays with buffered aqueous solutions favour thioester hydrolysis (right).

Quantifying the acid formed is achievable through titrating in base to keep the pH static, with the amount of base required correlating with the extent of the reaction. However, with the requirement of a pH stat and titration syringe this is a somewhat laborious method for carrying out routine assays and a simpler alternative was to measure the generation of thiol (SH) groups. Due to the biological importance of SH groups in proteins, glutathione, cystiene and secondary metabolites, a great variety of methods for their determination have been published, reviewed by Jocelyn[129] (1972). The difficulty was therefore in deciding which assay was best suited to our needs. The classical approach for determining the generation of VOSCs in enzyme assays is to partition the volatile products into a suitable organic solvent (e.g. diethyl ether, dichloromethane) prior to quantification by gas chromatography[113,114]. Such assays are sensitive and can be linked to a mass spectrometer for the analysis

of products, however, they are prone to the loss of VOSCs during processing of the organic phase, are time consuming and require access to GC equipment. Hence, simpler alternative assays were desirable.

A further assay for the SH group is based upon measuring the reducing potential of the product solution. Thiol containing compounds are well known to be important reducing agents in cells (e.g. glutathione as involved in redox homeostasis[76]), a property that can be assayed. For example, titration with iodine has been used for the free thiol determination of biological samples[130].

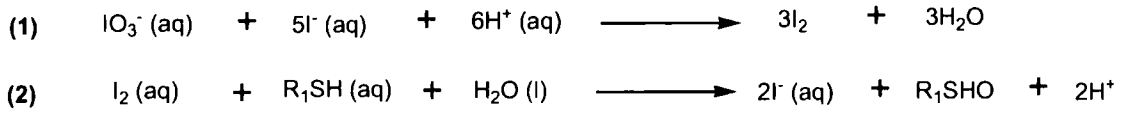


Figure 3.2 The iodometric determination of thiols in solution. **(1)** When iodate (IO_3^-) is added to an iodide solution (I^-) the iodate oxidises the iodide to form iodine (yellow). However, when free thiol is added **(2)** this acts as a reducing agent and converts iodine back to iodide. Only at the endpoint of the reaction when the reducing capacity of the solution is saturated, does free iodine form. The volume of iodate required to reach the end point correlates with the reducing capacity and hence the thiol content of the solution.

This method is however somewhat time consuming for our needs and it does not account for the presence of further reducing agents within a sample, such as endogenous ascorbic acid or glutathione.

Assays specific to thiol groups are based on the reactivity of the thiolate anion (RS^-) which is believed to be one of the most reactive groups found in cells, over 500 times more reactive than the corresponding oxygen analogue[85]. Hence, thiolate anions can be reacted with further chemical compounds as a means of quantifying their presence (Figure 3.3).

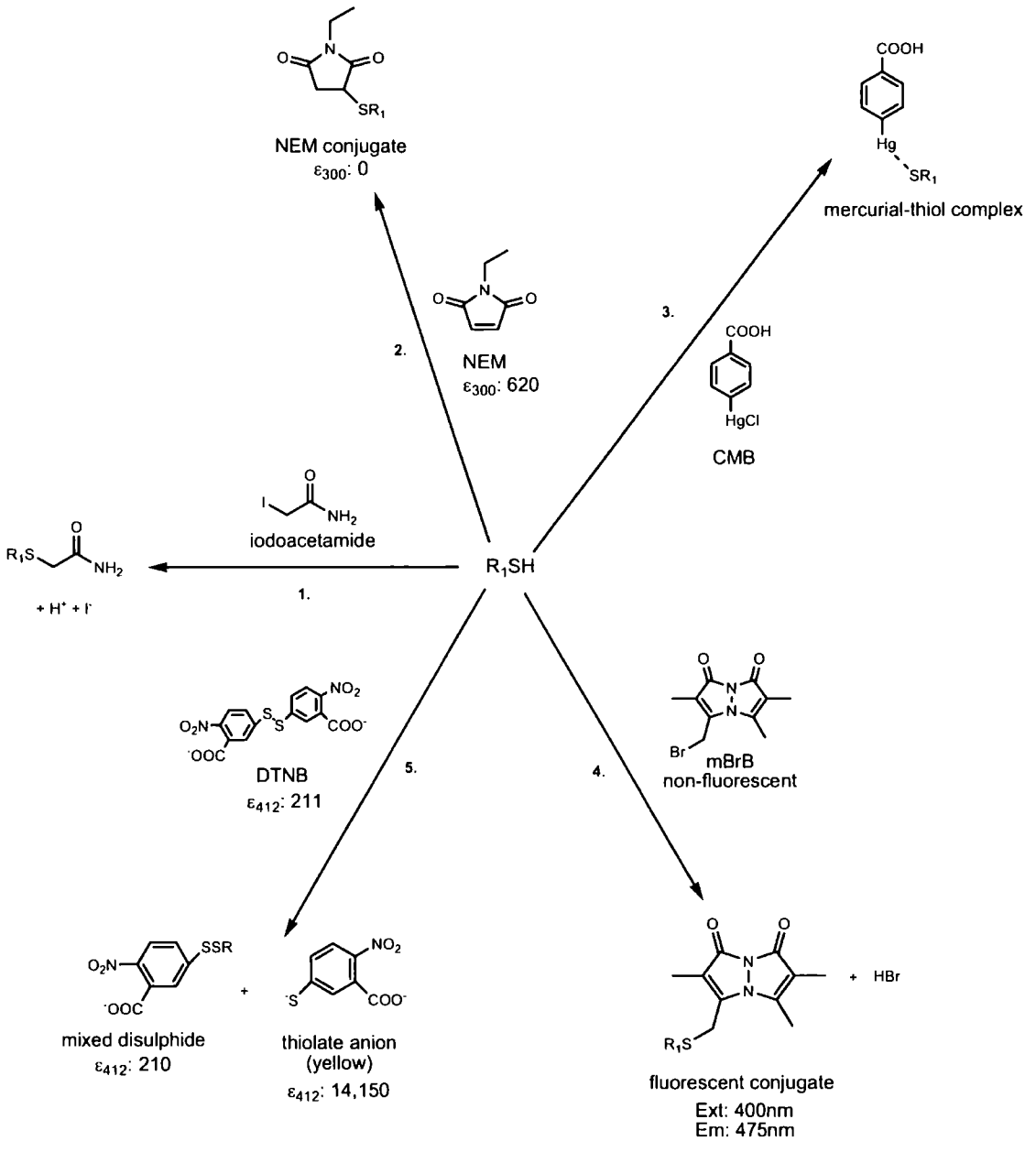


Figure 3.3 Assaying for thiol groups through chemical derivatisation. **1.** Alkylation; iodoacetamide can be used in a titrimetric assay as an equivalent of hydrogen ions are displaced during the reaction[131], **2.** Spectrophotometric (spec) assay; the addition of thiols across the double bond in maleimides (e.g. *N*-ethylmaleimide, NEM) produces an adduct with different absorption properties[132,133], **3.** Spec assay; mercurials, such as *p*-chloromercuribenzoic acid (CMB) complex with thiols to produce a product with greatly increased absorption at 250 nm[134], **4.** Fluorescence assay; monobromobimane (mBrB) reacts with thiols to generate a fluorescent conjugate (excitation 400 nm, emission 475 nm)[135], **5.** Thiol-disulphide exchange spec assay; 5,5'-dithio-bis(2-nitrobenzoic acid) (DTNB, Ellman's reagent) reacts with free thiol to generate a mixed disulphide and an aromatic thiolate anion which has a strong absorption at 412 nm[136].

Assays for thiols can be specific, such as those that are selective for thiol groups on glutathione or cysteine residues in proteins. However, here we required a general assay for thiol groups not a specific one. Such derivatisation techniques include titrimetric, spectrophotometric and fluorescent assays, each with their own limitations which must be considered. Long or complex sample analysis does not lend itself to repetitive assays (titrimetric assays), certain reactions can only be achieved under acidic or basic pH and hence can not be used for following continuous enzymatic rates (they must be stopped first). The spontaneous hydrolysis of the thioester bond is slow at neutral or acidic pH but is rapid above pH 8, hence such high pH must be avoided to reduce chemical hydrolysis[137]. In addition, the solubility and decomposition of adducts can be problematic (e.g. NEM[129]) and some reagents, such as the mercurials are believed to promote thioester hydrolysis[138]. Hence, it was of importance to test several assays so as to determine which would be most effective in quantifying the enzymatic generation of thiols from thioester precursors. Assay development and subsequent identification and partial characterisation of thioesterase activity in purple passion fruit therefore forms the focus of this initial work.

3.2 The Development of a Simple Assay for Measuring Thioesterase Activity in Fruit Tissue Extracts

Before testing and developing various assays it was first important to ensure thioesterase activity was present in tropical fruit using a method known to work. It was therefore decided to start by using the standard gas chromatography system. In addition, two commercial enzymes, pig liver esterase (PLE) and *Candida rugosa* lipase (CRL), were selected to assay alongside the fruit extracts as positive controls, both enzymes had previously been demonstrated to posses thioesterase activity[113]. CRL is a widely used industrial lipase[103] and PLE is a remarkably versatile esterase which has attracted great academic interest for potential use in bioprocessing[103]. It was therefore useful to gauge the efficiency of the fruit enzymes against these tried and tested biocatalysts.

3.2.1 Gas Chromatography Assay

The high degree of sensitivity achieved using gas chromatography (GC) has made it the method of choice for analyzing volatile aroma chemicals. This is a likely reason why it was initially chosen for determining the enzymatic generation of aromatic volatiles. Here the enzyme reactions were incubated with the thioester substrate 2-methyl furan-3-thioacetate (MFTA) as PLE and CRL had previously been demonstrated to act on this substrate[114]. MFTA is a high-impact sulphur volatile generated upon roasting beef, although not a natural fruit metabolite it will be a useful tool for determining activity in fruit at this stage. The delocalised furan ring stabilises the thiolate anion making it a good leaving group which favours enzymatic hydrolysis and will aid in determining what is predicted to be a relatively low abundance activity. The general reaction scheme is depicted below.

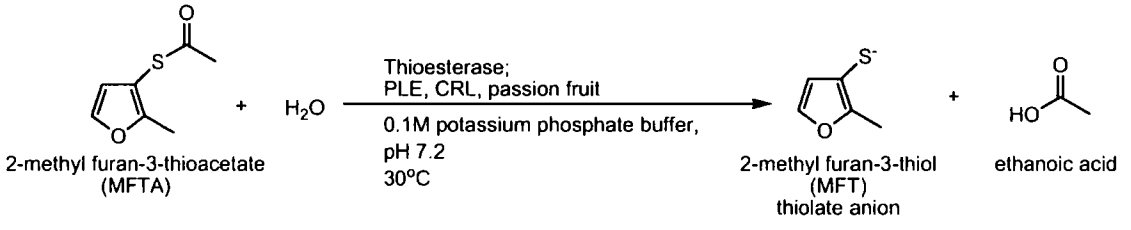


Figure 3.4 A reaction scheme illustrating the enzymatic hydrolysis of MFTA. Thioesterases including a microbial lipase (CRL), a mammalian esterase (PLE) and a fruit crude protein extract (passion fruit) were assayed for activity against this substrate.

The identification of substrate and product was confirmed through running the thioester (MFTA) and thiol (MFT) standards through the GC and determining their retention times (Figure 3.5A). In addition, MFT was run through the GC at varying concentrations (3-50 nmols) for calibration, illustrating the direct relationship between GC peak area and the amount of product formed (Figure 3.5B+C). It was also important to take into account the incomplete partitioning of volatiles into the solvent phase during processing, the uptake of MFT into DCM was determined at 60% +/- 1% and was corrected for when calculating specific activities.

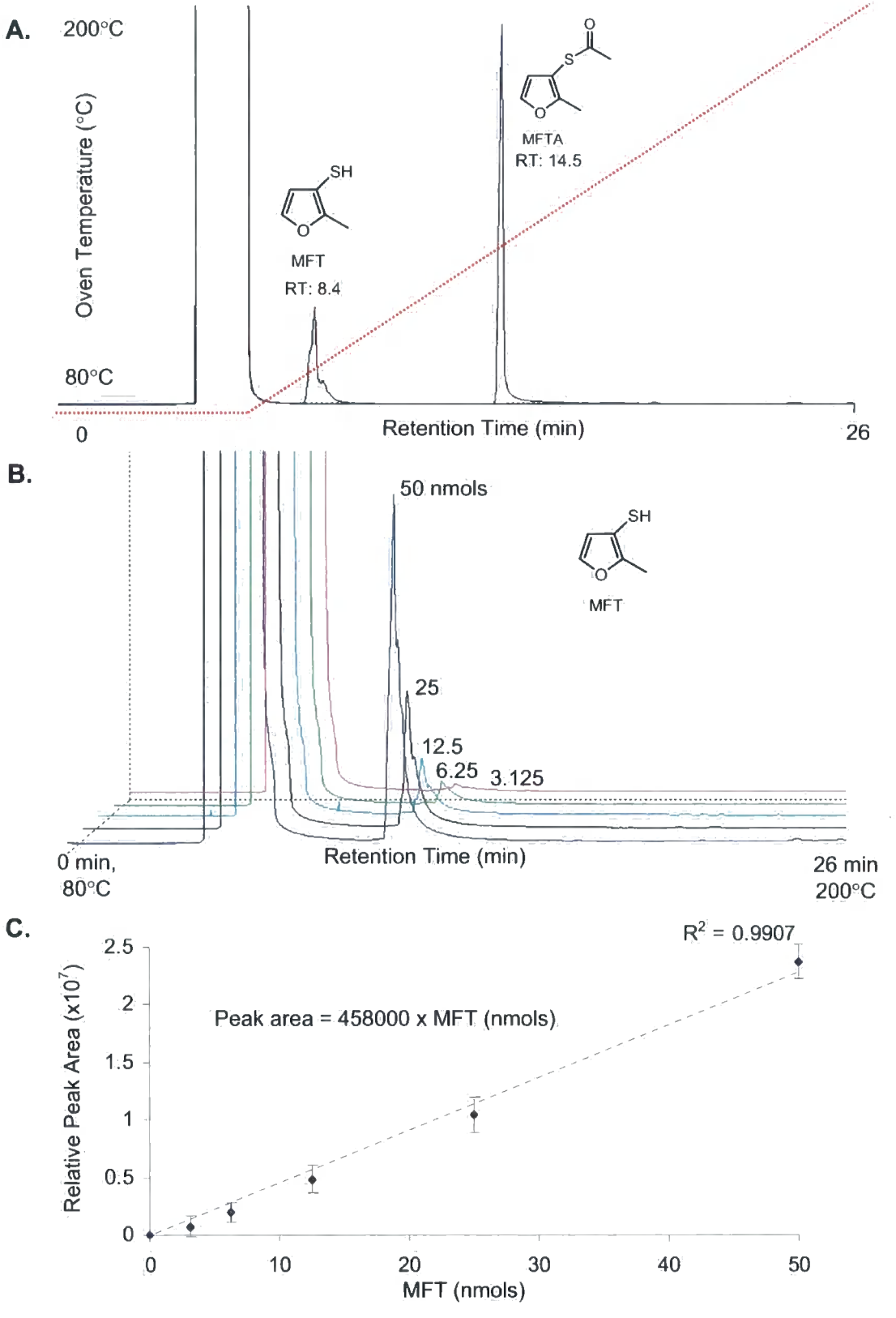


Figure 3.5 A. GC separation and identification of the standards 2-methyl furan-3-thiol (MFT) and 2-methyl furan-3-thioacetate (MFTA). **B.** A GC trace illustrating the relationship between GC peak area and the amount of thiol product MFT. **C.** The direct linear relationship between peak area and the amount of MFT over the range 3-50 nmols.

At this stage a simple total protein extract was made from 5 ripened (10 days incubation at 25°C) purple passion fruit (*Passiflora edulis* Sims) through homogenising whole fruit in a 0.1M potassium phosphate buffer prior to $(NH_4)_2SO_4$ precipitation (40-80% cut) and overnight dialysis against the same buffer. Protein preparations were dissolved in 490 µl of 0.1 M potassium phosphate buffer, pH 7.2 in air tight 1 ml glass vials before thioester substrate (10 µl) was injected using a Hamilton syringe through a septum to final concentration of 20 mM. The vial was sealed with parafilm and incubated for 12 minutes at 30°C before the reaction was stopped by the addition of an equal volume of DCM (500 µl), mixed and the solvent removed and dried with anhydrous sodium sulphate before analysis on the GC (Figure 3.6). The amount of thiol product formed could be determined from the size of the peak area using the calibration curve (Figure 3.5C) and specific activities (nmol product/sec/mg) for each enzyme preparation calculated (Table 3.1).

Thioesterase	CRL	PLE	P.e. Crude Extract
Specific Activity (nkat/mg)	0.31 +/- 0.08	510.99 +/- 84.28	0.38 +/- 0.08

Table 3.1 A comparison of thioesterase activities toward MFTA as determined using the GC assay. Assays were run in triplicate along with a heat treated (95°C, 5min) protein sample used for determining the non-enzymatic chemical rate. Errors +/- 1 standard deviation from the mean.

Thioesterase activity toward MFTA was determined in crude passion fruit extracts (0.38 nkat/ mg) and forms the first evidence that a thioesterases are present in tropical fruit. The thioesterase activity of PLE and CRL was confirmed, further supporting the identification of the activity in the fruit extract. Thioesterase activity in the fruit was comparable to that of CRL, and although considerably lower than that of PLE this is a crude protein preparation and the commercial enzymes are in partially purified form. The assay was sensitive and reproducible, however, processing of the organic phase was likely to result in a loss of volatiles (sulphur volatiles were smelled in the air during extraction) and an investigation of further thiol assays was required.

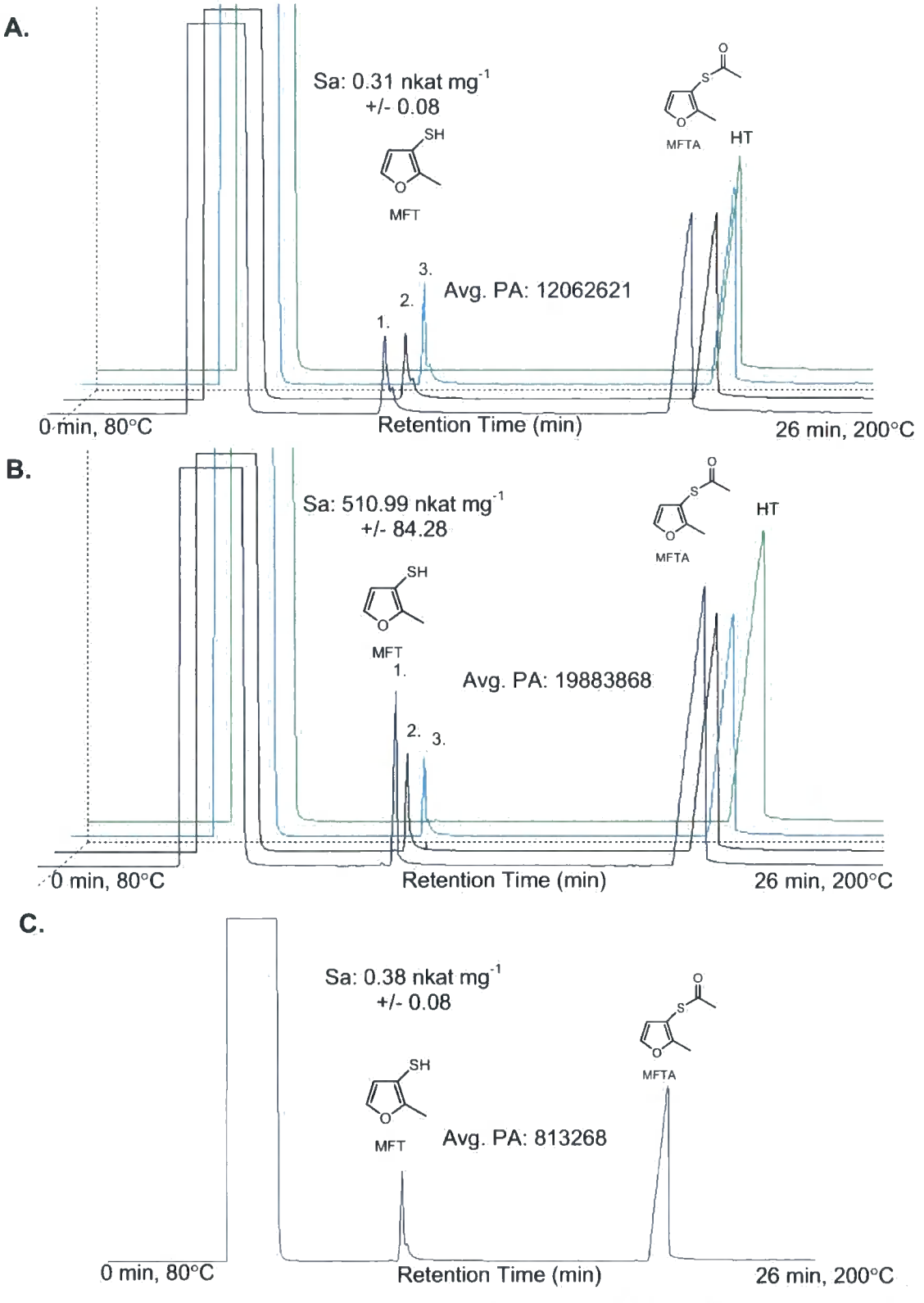


Figure 3.6 The enzymatic hydrolysis of MFTA (20 mM) by **A.** CRL (2 mg/ml), **B.** PLE (0.002 mg/ml) and **C.** Passion fruit crude protein extract (0.11 mg/ml). Each trace shows the substrate peak (right) and product peak (left) following a 12 min incubation with enzyme. Traces 1-3 are triplicates of the assay as compared to a heat treated protein sample (HT, 95°C, 5 min). From the average peak area (Avg. PA) the corresponding amount of MFT can be determined and a specific activity (Sa) calculated.

3.2.2 Fluorimetric Assay

It was desirable to carryout the thioesterase assay in an air tight tube, preferably in a single solution so as to prevent the loss of product or starting materials noted using the GC assay. A method for the analysis of biological thiols had been previously developed using monobromobimane (mBrB, (3-bromomethyl-2,5,6-trimethyl-1H))[139], a non-fluorescent compound that reacts with thiols to generate a fluorescent product. This was developed to tag pico-mol levels of thiols in biological extracts (e.g. red blood cells) and through separation using reverse-phase HPLC single metabolites could be identified[135]. This assay had the potential to be adapted for determining thioester hydrolysis in enzymatic reactions, the basis of which is depicted in below.

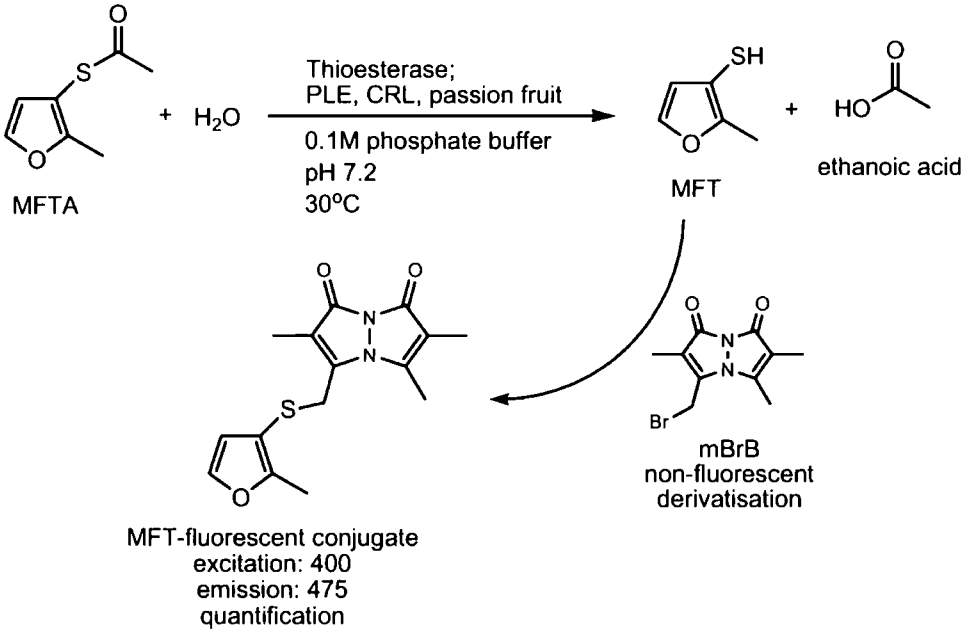


Figure 3.7 The determination of thioesterase activity using the fluorometric monobromobimane (mBrB) assay and the model substrate MFTA.

This theory was tested by incubating the thiol standard MFT at various concentrations (1-10 μM) with mBrB (50 μM) for 1 min at 30°C and measuring fluorescence (excitation: 400, emission: 475). An MFT product standard curve was established (Figure 3.8), with the relationship between fluorescence and thiol content linear over the range 1.5-10 μM product.

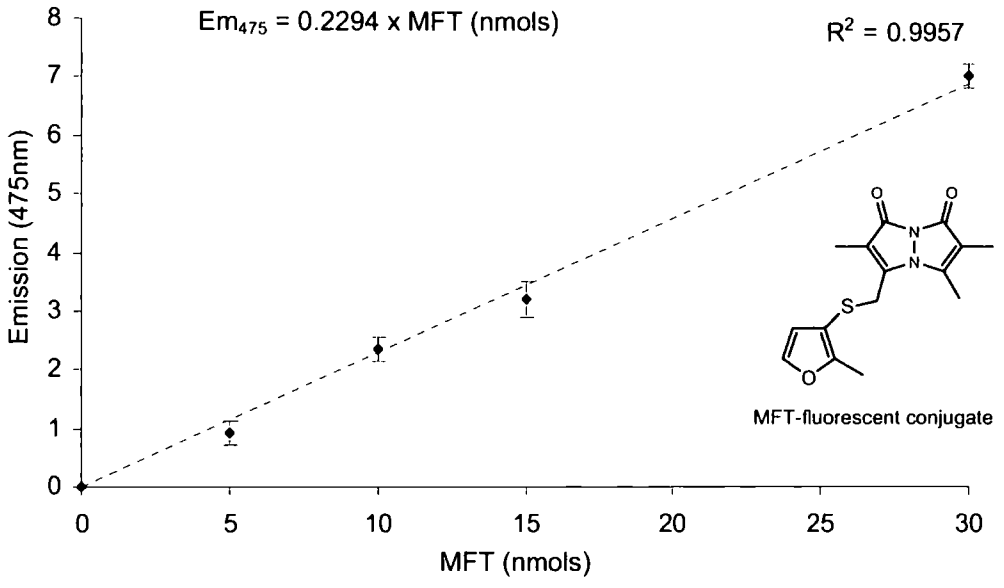


Figure 3.8 An MFT standard curve illustrating the direct linear relationship between MFT concentration and fluorescence emission following derivatisation with mBRB.

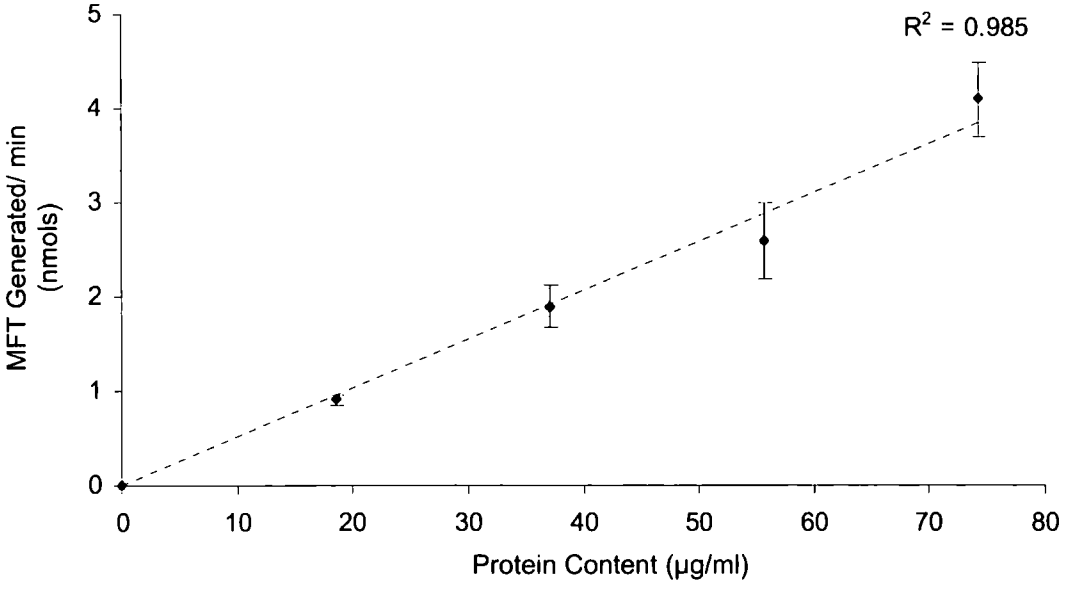


Figure 3.9 An illustration of the protein dependence of the mBrB assay. Varying amounts of passion fruit crude extract (0.06-0.2 mg) were incubated with 1 mM MFTA and the amount of product determined using derivatisation with mBrB.

Having established the assay could theoretically work it was tested on the protein extract from passion fruit, which we had shown to posses thioesterase activity through the GC assay. Protein samples in 0.1M potassium phosphate buffer pH 7.2 were incubated with 1 mM MFTA at 30°C in a 2 ml fluorimetric cuvette. The reaction was started by the addition of 50 μM mBrB, injected through a Hamilton syringe, and the increase in fluorescence emission at 475 nm measured over 2 minutes on a fluorimeter. The reaction could be monitored continuously and it was not required to stop and extract product. The rate was linear over 2 minutes and through adding increasing amounts of protein fruit extract (0.03-0.1 mg/ml), the assay was demonstrated to be protein dependent (Figure 3.9). The specific activity of the fruit preparation was determined at 0.31 nkat/mg (+/- 0.03), comparable to that determined through the GC assay (0.38 nkat/mg +/- 0.08).

3.2.3 Spectrophotometric Assay

The hydrolysis of thioesters can also be followed spectrophotometrically as thioester bonds have an absorption peak at approximately 230 nm, somewhat lower than that of the corresponding thiol[140]. It was therefore postulated that through determining the absorption maximum of thiol product and then following the increase in absorption at this wavelength the thiol content of the solution could be measured. However, if protein in solution also absorbed over the same maximum then thiol quantification may be masked. In addition, different chemical types may shift the absorption maxima or aromatic constituents may mask it completely. Although this assay was theoretically the simplest to perform as no derivatisation or solvent partitioning was required it was however likely to have limitations in practice.

In this case, the artificial substrate 4-methylumbelliferyl thiopropanoate (MUTP) (Figure 3.10) was chosen as it contained a heavily delocalised double ring structure which would in theory stabilise the thiolate anion and cause a strong absorption of visible light. The chemical structure was larger (Mr: 249, MFTA: Mr: 156) and was unlikely to be lost through vaporisation into the surrounding air.

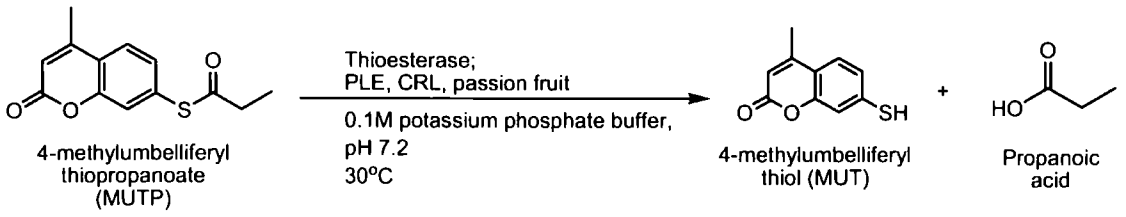


Figure 3.10 A reaction scheme for the enzymatic hydrolysis of 4-methylumbelliferyl thiopropanoate (MUTP). The extent of the reaction was monitored through determining the absorption maximum for the thiol product and following absorption at this wavelength using a spectrophotometer which theoretically would directly correlate with product concentration.

The absorbance spectra for both MUTP and MUT (45 μM) in phosphate buffer, pH 7.2 were independently scanned between 200-600 nm (Figure 3.11). A distinct peak in the product spectra (MUT) at 377 nm was not seen in the spectra for the substrate (MUTP) and was likely to be due to the thiol functionality. Hence, following the reaction at 377 nm enabled the determination of thiol product concentration.

An enzymatic assay was set up by dissolving passion fruit protein preparation (50 μl, 2.5 mg/ml) in 1 ml of potassium phosphate buffer, pH 7.2 with 100 μM MUTP. The cuvette was mixed and the increase in absorbance at 377 nm determined over 5 minutes at 30°C (Figure 3.12). The rate observed was non-linear over the first 2 minutes and then levelled over the following three minutes. This may have been caused by the incomplete solubilisation of MUTP and therefore its concentration was reduced to 50 μM and the assay repeated (Figure 3.12), again the rate was non-linear. Hence, following the enzymatic reaction at this wavelength does not directly correlate with product formation, a linear relationship would be expected for a typical enzymatic reaction. It is also possible absorption by compounds (e.g. protein) in the fruit extracts may be interfering with the assay. In conclusion it is not a reliable method and as further difficulties were likely to be encountered when different chemical structures were assayed (which may alter λ_{max}) it was decided to focus on alternative methods.

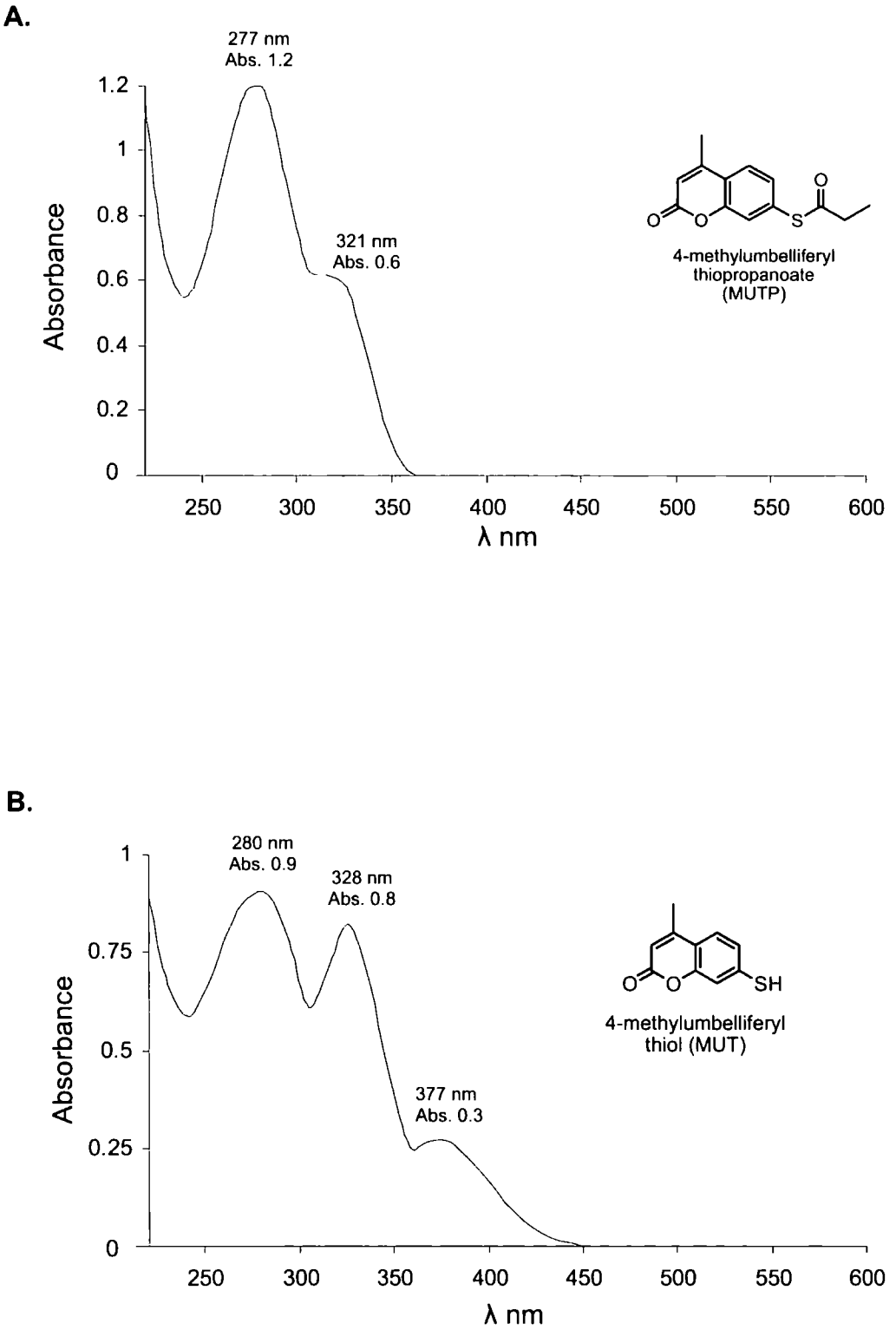


Figure 3.11 Spectral scans (pH: 7.2) of **A.** MUTP and **B.** MUT over the visible range (220-600nm). An absorption peak in the MUT spectra at 377 nm is not present in the MUTP trace and may be accounted for by the thiol functionality.

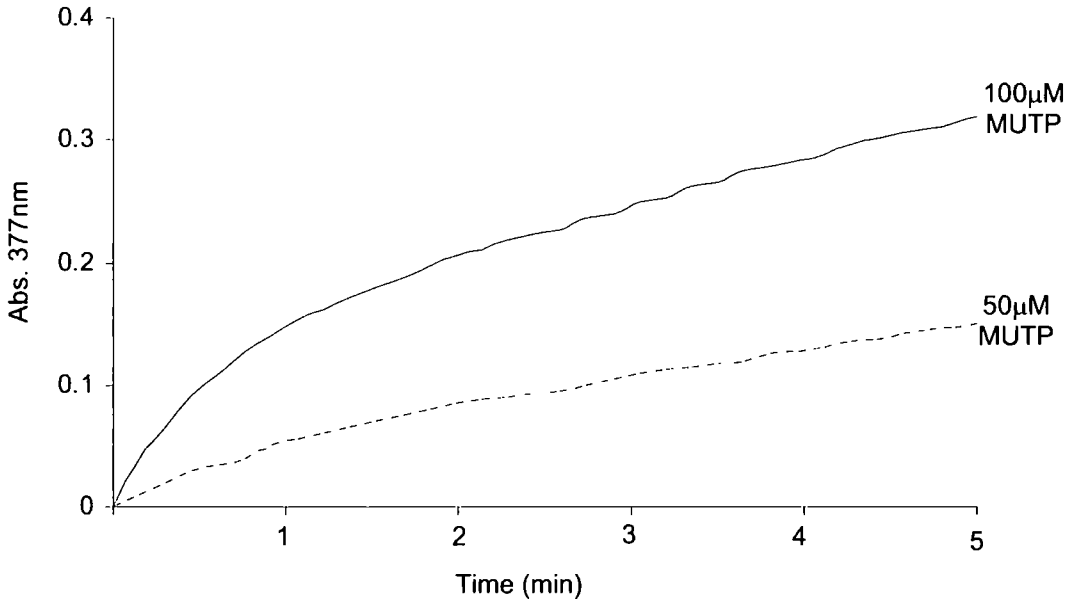


Figure 3.12 Spectrophotometric traces for the enzymatic cleavage of 4-methylumbelliferyl thiopropanoate (MUTP) by the passion fruit extract (0.13 mg/ml). Thiol product formation was determined by following the change in absorbance at 377 nm (MUT generation). Reactions were tested with both 100 μM and 50 μM substrate to see if the incomplete solubilisation of MUTP may be causing the non-linear rate.

3.2.4 Ellman's Assay

One of the most widely used methods for assaying thiols is the thiol-disulphide exchange method, whereby aromatic disulphides react quantitatively with free thiols to generate a mixed disulphide and an intensely coloured thiolate anion. Such an example is 5,5'-dithio-*bis*(2-nitrobenzoic acid) (DTNB, Ellman's reagent)[136]. Ellman's reagent was introduced in 1959 as a biochemical tool for quantifying sulfhydryl groups on proteins and primary metabolites such as glutathione provided by the amino acid cysteine[136]. As illustrated in figure 3.13, the assay is based upon the organosulphur compound reacting in the form of the thiolate anion (RS^-) with DTNB to generate a mixed disulphide (RS-TNB) and a by-product TNB which absorbs visible light intensely at 412 nm.

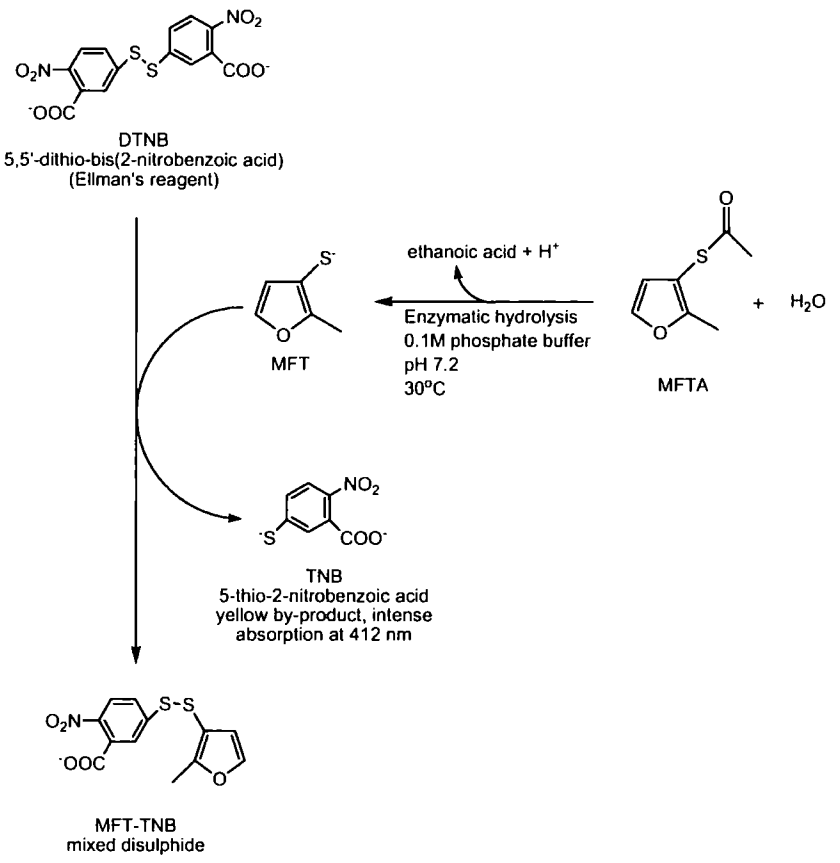


Figure 3.13 The use of Ellman's reagent for measuring thioesterase activity based on the hydrolysis of 2-methylfuran-3-thioacetate (MFTA).

Passion fruit protein samples were dissolved in 490 µl of 0.1 M potassium phosphate buffer, pH 7.2. with 10 µl of MFTA (final concentration of 1 mM) and incubated at 30°C in a sealed cuvette. An equal volume (500 µl) of Ellman's reagent (4 mg/ml in same buffer) was added, mixed and product formation determined by measuring the change in absorbance at 412 nm over 5 minutes (Figure 3.14A).

The rate was linear over this period and due to the simplicity and speed of the assay it looked very attractive, however further validation was first required. The enzymatic reaction was re-run using a different thioester substrate, furfuryl thioacetate (FTA, Figure 3.14A). Authentic standards of MFT and furfuryl thiol (FT, furfuryl mercaptan) were used to produce standard curves for quantifying product formation (Figure 3.14B), theoretically both curves should be identical if the derivatisation with Ellman's reagent was not compound specific.

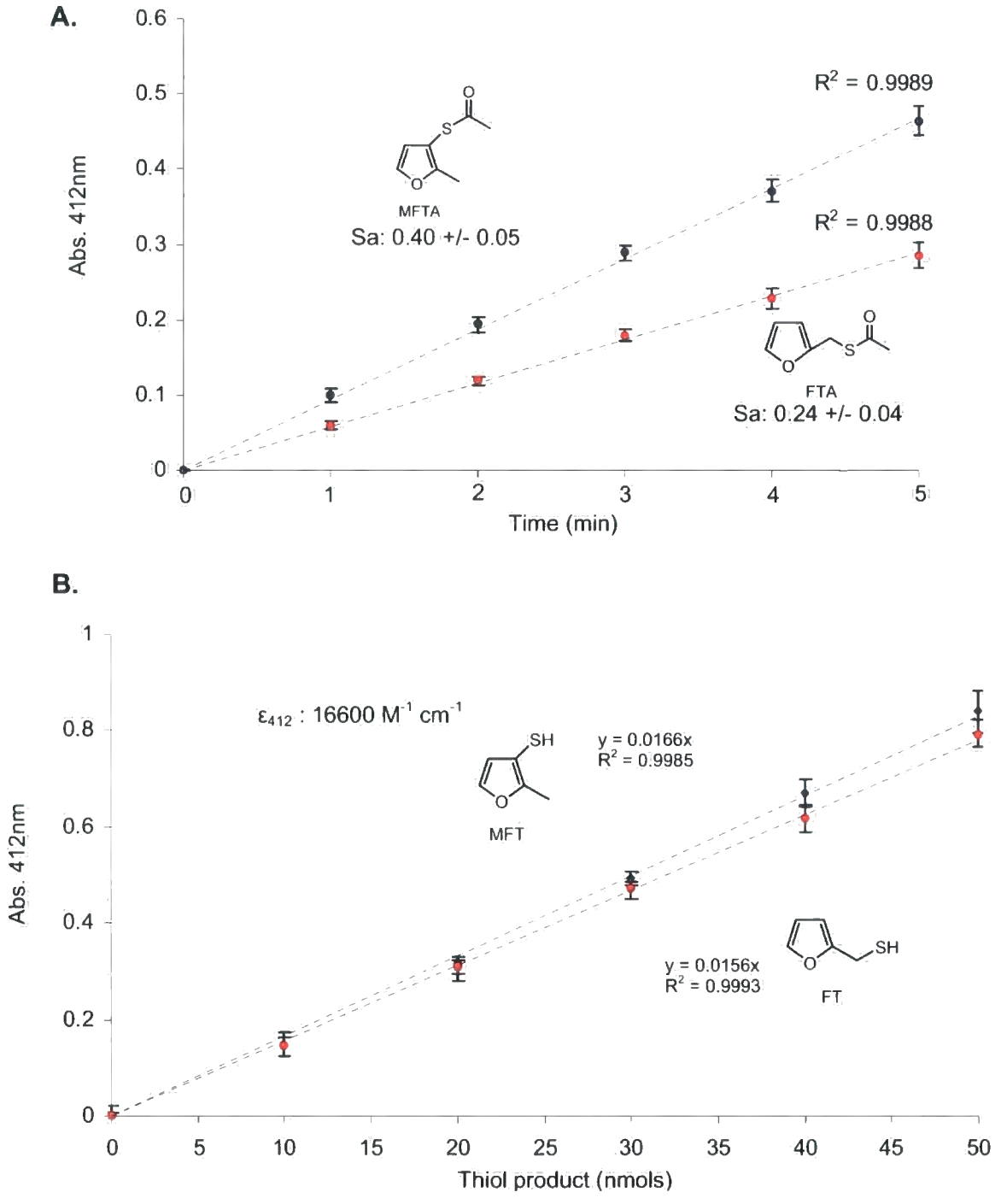


Figure 3.14

A. The enzymatic hydrolysis of MFTA and FTA by passion fruit crude protein extract (0.25 mg/ ml). The increase in absorbance (thiol generation) is linear over 5 minutes. The specific activities (Sa) of the extract toward both thioesters are shown.

B. Calibration curves generated using furfuryl thiol (FT, red spots) and 2-methyl furan-3-thiol (MFT, blue spots). When errors are taken into account the curves are superimposable. The extinction coefficient for MFT was determined at 16600 $M^{-1}cm^{-1}$.

Both curves were linear indicating that absorbance at 412 nm was directly proportional to the concentration of thiol in solution. When considering the different chemical types (MFT and FT) both standard curves were superimposable when the slight error between triplicates was considered. Essentially, the Ellman's assay is able to determine the concentration of thiol in solution regardless of the compound conformation on which the group is attached. This is supported by the same observation in the literature[141].

For further validation of the assay the thioesterase activity of PLE and CRL toward MFTA was also determined (Table 3.2) and specific activities were found to be comparable to those determined through the GC assay.

Thioesterase	CRL	PLE	P.e. Crude Extract
GC Assay Specific Activity (nkat/ mg)	0.31 +/- 0.08	510.99 +/- 84.28	0.38 +/- 0.08
Ellman's Assay Specific Activity (nkat/ mg)	0.35 +/- 0.02	537.88 +/- 24.65	0.40 +/- 0.05

Table 3.2 Thioesterase activities toward MFTA determined using the GC and Ellman's assay.

In fact, as no processing of the product solution was required, the error resulting from multiple pipetting steps and the loss of volatiles was reduced. As a means of validating the assay, increasing concentrations of PLE (0.1-0.8 μg/ml) were incubated with substrate to ensure product formation was protein-dependent (Figure 3.15A). Furthermore, the time dependence of the assay was investigated by incubating 1 mM MFTA with PLE (0.1 μg) and following the absorbance at 412 nm over 10 minutes (Figure 3.15B). The rate was initially linear for the first 3 minutes before a slight drop in gradient (rate) was observed, presumably as the temperature of the reaction equilibrated to room temperature (RT). However, it was important to carry out assays slightly above RT (30°C) so as to avoid fluctuations caused by varying air temperature. Over the initial minute the rate was linear and suitable for our requirements.

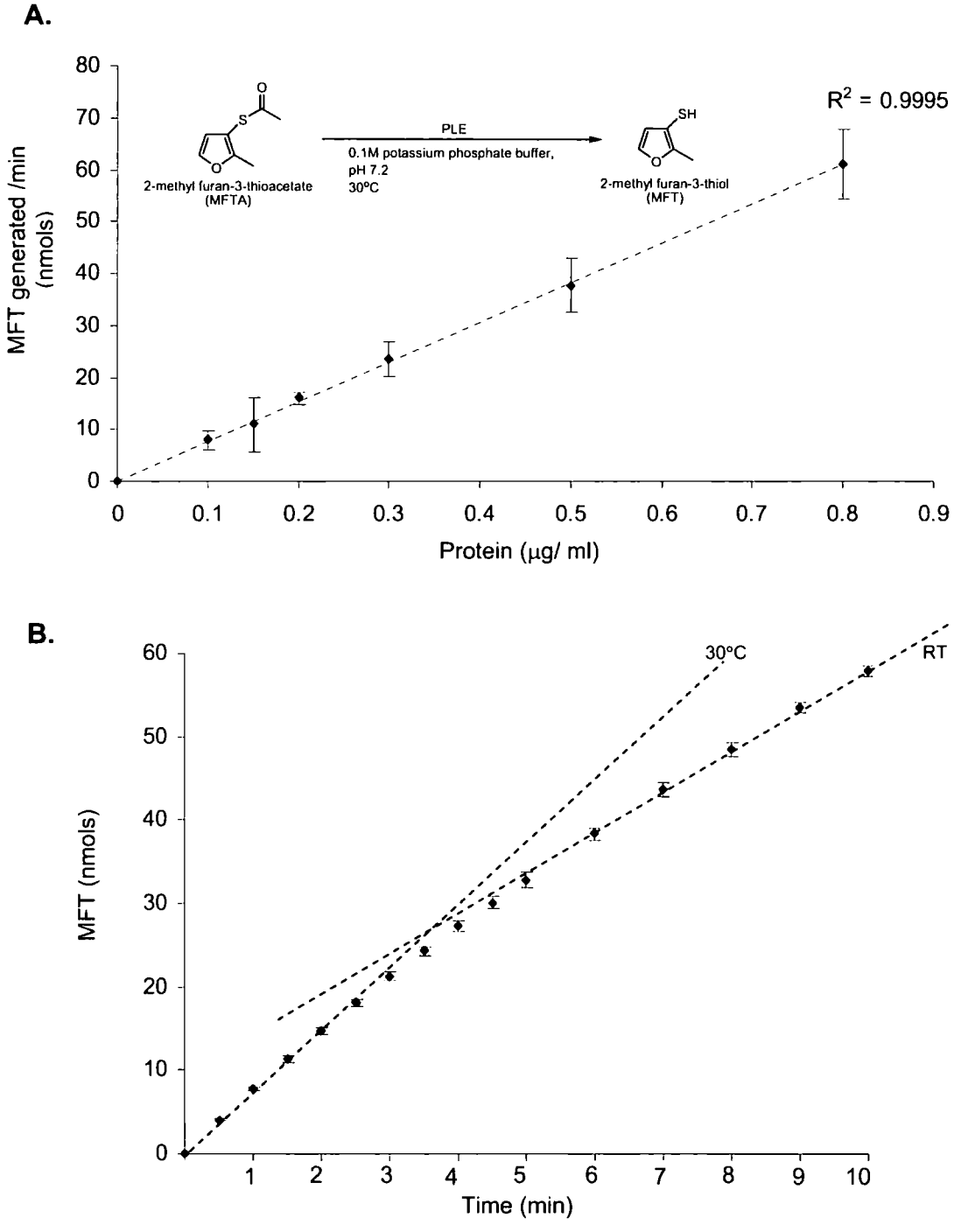


Figure 3.15 Ellman's assay validation

A. An illustration of the protein dependence of the Ellman's assay using PLE. Varying amounts of PLE (0.1-0.8 µg/ml) were incubated with 1mM MFTA (total: 1000 nmols).

B. An illustration of the time dependence of the Ellman's assay using PLE. 0.1 µg/ml PLE was incubated with 1 mM MFTA for 10 min. The enzymatic rate was linear for the first 3 minutes before the gradient fell slightly over the remaining time, presumably as the temperature of the reaction dropped to RT, and consequently slightly reducing the reaction rate.

The Ellman's assay was simple, fast to perform and sensitive enough for our requirements with a published minimum detection limit of 0.6 nmol thiol/ml solution[141]. The assay gave a linear response over the range 10-50 μM thiol with excellent reproducibility as compared with GC protocols. The only limitation to the assay was that reactions had to be performed at close to neutral pH with the maximum stability of DTNB at pH 7.2 (most enzymes are however active around this pH). Under alkaline conditions (pH > 8.6) decomposition of DTNB occurs[141]. Importantly, while GC analysis takes 50 minutes per sample and includes solvent partitioning, the one step Ellman's assay reaction was complete in 5 minutes with the developed colour being stable over 1 hour. The fluorescence assay using mBrB was also simple and highly sensitive and formed an equally efficient means for determining thioesterase activity in enzyme assays. However, the Ellman's assay could be carried out on microtitre plates (96 wells) and run on a flashscan so as multiple assays could be conducted over a short time (5 min), a useful approach for the large number of assays required when conducting protein purification. Although the fluorimetric mBrB assay could also be conducted on such a scale using fluorescent imaging, attempts using a FLASHscan® spectrometer (Analytik Jena AG, Germany) for conducting multiple assays were found to be variable, lacking the sensitivity and reproducibility of the Ellman's method. For this reason the Ellman's protocol was chosen as the optimal assay for determining thioesterase activity in further work.

3.3 Sourcing Thioesterases from Tropical Fruit

Having developed an effective assay for quantifying free thiol generation and establishing that thioesterase activity was present in purple passion fruit it was important to look to other exotic fruits as possible sources of the enzyme. The fruit type was not a concern in this case as the only requirement was finding the optimal source of an efficient biocatalyst. A number of fruit were screened, including mango, lychee, banana and yellow passion fruit. In each case total protein was extracted from whole fruit (Methods 2.2) and assayed for activity toward MFTA using the Ellman's protocol.

It soon became apparent that purple passion fruit was a good first choice of material as it was very problematic extracting protein from other fruit types. The high levels of fat and pectin in each variety caused difficulties with the ammonium sulphate precipitation of protein and in each case no pellet formed and the supernatant formed a thick jelly. Attempts were made to clarify the solution prior to salting out the protein, including the addition of 10% (w/v) protamine sulphate to the extraction buffer for removal of nucleic acids. However, protein recovery was low and no activities could be determined from this material. It was therefore decided that at this stage it was best to concentrate on the purple variety of passion fruit as activity had been determined in this fruit. In addition, it is known to contain numerous characteristic tropical notes, which have also been identified as thioester precursors[119]. Work concentrated on following this activity as it was always possible to return to other fruit types as a source of the enzyme at a later stage through a genetic approach (e.g. PCR cloning and expression).

3.4 Characterizing Thioesterase Activity in purple passion fruit (*Passiflora edulis* Sims)

Before embarking on the purification of thioesterase activity from purple passion fruit it was important to characterize the observed activity further. The properties of the enzyme (e.g. stability and temperature tolerance) would be important factors to consider in evaluating its potential application in bioprocessing. In addition, it was important to determine where in the fruit activity was localised and how it was regulated during maturation as a means of optimising extraction and for understanding more about the physiological role of the protein.

3.4.1 Localisation of Thioesterases to the Fruit Mesocarp (Peel)

Initial assays used homogenates made from whole fruit and as such they did not tell us anything about the localisation of activity. In order to examine this fruit were separated into their component tissues; seed, endocarp, mesocarp, exocarp and juice (Figure 3.15).

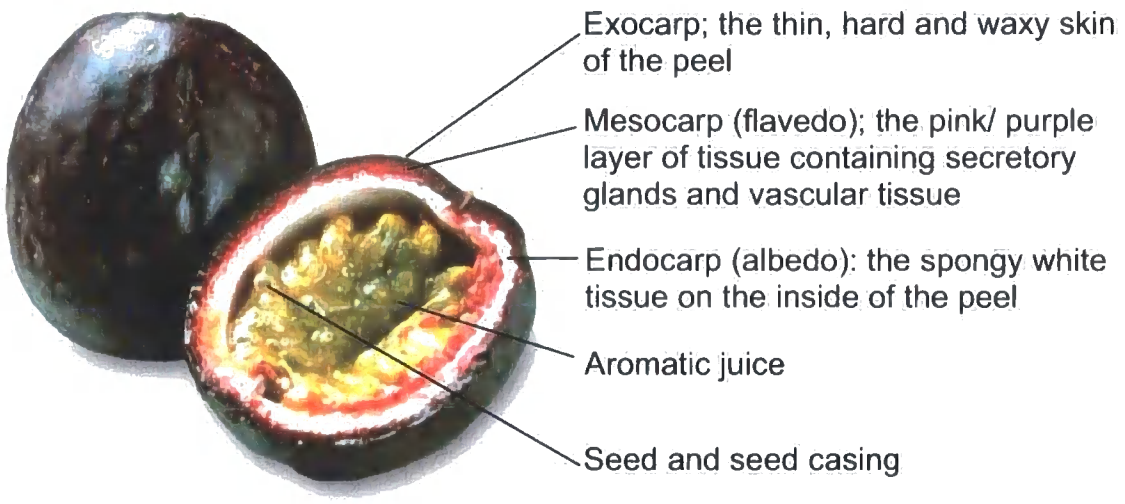


Figure 3.15 Purple passion fruit - physiology and tissue type

Five ripened fruit (10 days, 25°C) were separated into component tissue types and processed (Methods 2.2) before being assayed for activity toward MFTA using the Ellman's method (Table 3.3).

Fruit tissue	Specific activity (nkat mg^{-1})
seed	ND
juice	ND
endocarp	0.06 +/- 0.03
mesocarp	0.36 +/- 0.04
exocarp	ND

Table 3.3 The localisation of thioesterase activity in passion fruit

It was apparent that thioesterase activity was predominantly confined to the mesocarp tissue of the peel. The mesocarp is made of cellular material and has a higher water content than the inner spongy endocarp which is made up of structural material and contained only trace thioesterase activity (0.06 nkat/mg). The bulk of the peel is composed of the mesocarp and for this reason it was not necessary to remove the waxy peel or endocarp tissues. Interestingly the mesocarp also contains numerous secretory glands which protrude into the interior of the fruit. Such structures could potentially contain enzymes of flavour biosynthesis.

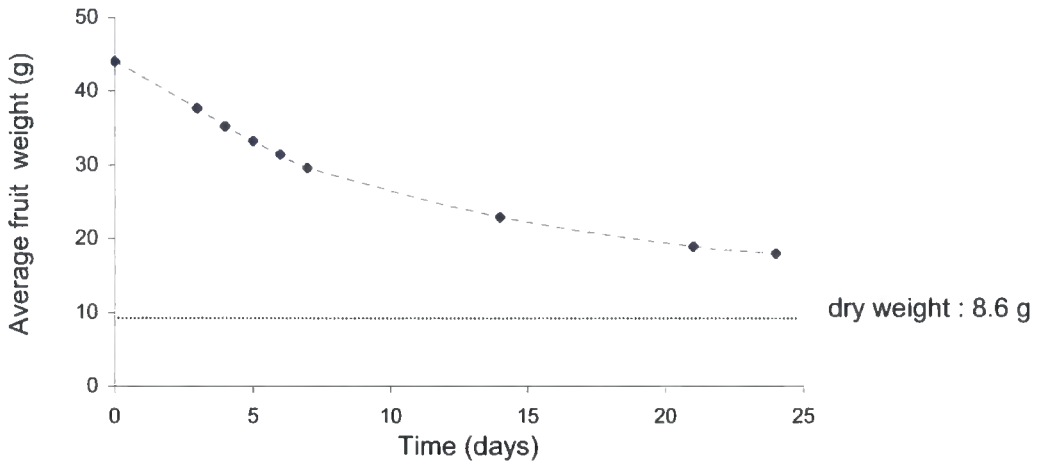
The peel tissue of fruit appears a very dry environment in which to find enzyme activity, especially later in ripening when the fruit loses much of its water content. It was therefore speculated that the observed activity may be a consequence of a wall-bound enzyme. For this reason, further protein extracts were made from the fruit with the addition of salt (0.25-1 M NaCl). Fruit peel homogenates were left to soak in the salt extraction buffer with stirring on ice for 2 hours to overnight prior to filtering and centrifugation. In each case the supernatant was assayed for thioesterase activity. The addition of salt (0.25 M) caused a 2.5-fold increase in extractable thioesterase activity (0.97 +/- 0.03 nkat/mg), with 0.25 M NaCl adequate for the optimal extraction of activity over 2 hours. Longer incubation times or higher salt concentrations yielded comparable activities. Hence, the activity was found to be associated with a partially wall bound protein in the outer cortex (mesocarp) of the fruit.

3.4.2 The Developmental regulation of Thioesterase Activity

As fruit develop, the numerous biochemical and genetic changes that occur cause them to become attractive, flavoursome and nutritious. The sharp increase in the production of the gaseous plant hormone ethylene in climacteric fruit (including passion fruit) is believed to initiate a cascade of genetic and biochemical changes which bring about ripening[142]. It is a genetically programmed event involving the regulation of specific genes, some being expressed constitutively and others activated only when ethylene levels reach a specific concentration. An example of the latter are the alcohol acyl transferase (AAT) genes, which in apple were found to regulated by ethylene which controls this final stage of ester biosynthesis[60]. Hence, the regulation of thioesterase activity in fruit may tell us more about the function of these enzymes. However, the developmental stage of the fruit must be determined.

Fruit from the local markets were stored so as to prevent the onset of ripening (5°C, 85% humidity), with development subsequently initiated by transferring fruit to a growth room under controlled conditions (25°C, 16 hour photoperiod). As the fruit matured they lost weight, shrivelled, their water content dropped and the juicy interior became more fragrant (Figure 3.16A+B).

A.



B.

Figure 3.16 Developmental changes in ripening purple passion fruit. **A.** 10 fruit were ripened under controlled conditions (25°C), over the 24 day study they lost water and their weight fell. **B.** This was associated with the skin shrivelling and drying out. Day 10 was considered to be when fruit was ripe, they degraded and lost quality after this.

As a means of determining the developmental stage of each batch of fruit, their average dry weight was calculated, and the % dry weight determined as a simple means of measuring the degree of ripening. Twelve fruit were ripened over 24 days, with three fruit processed (peel + 0.25 M NaCl extraction) at regular intervals and thioesterase activity determined (Figure 3.17).

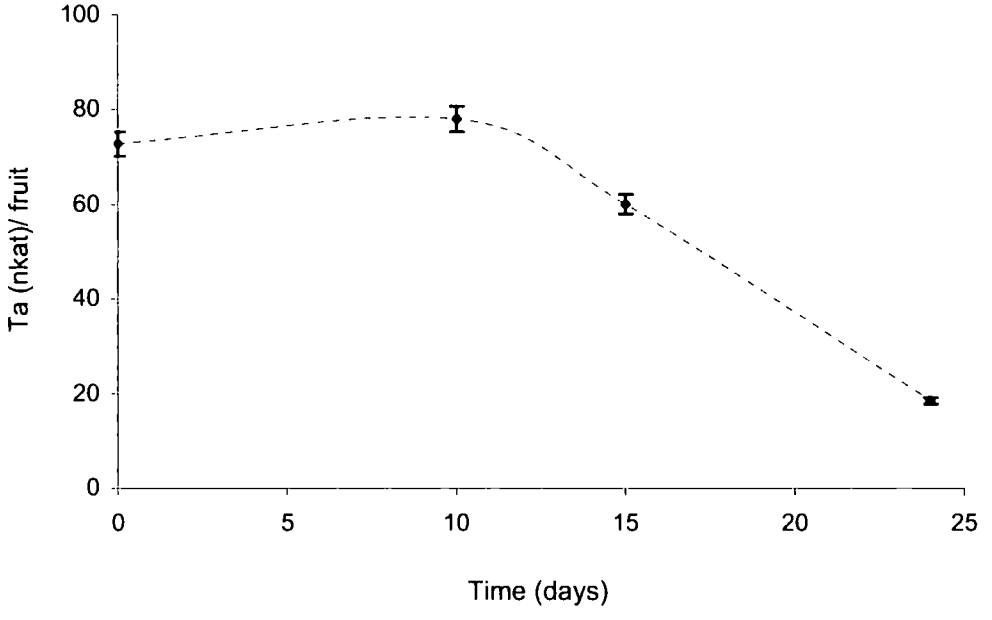


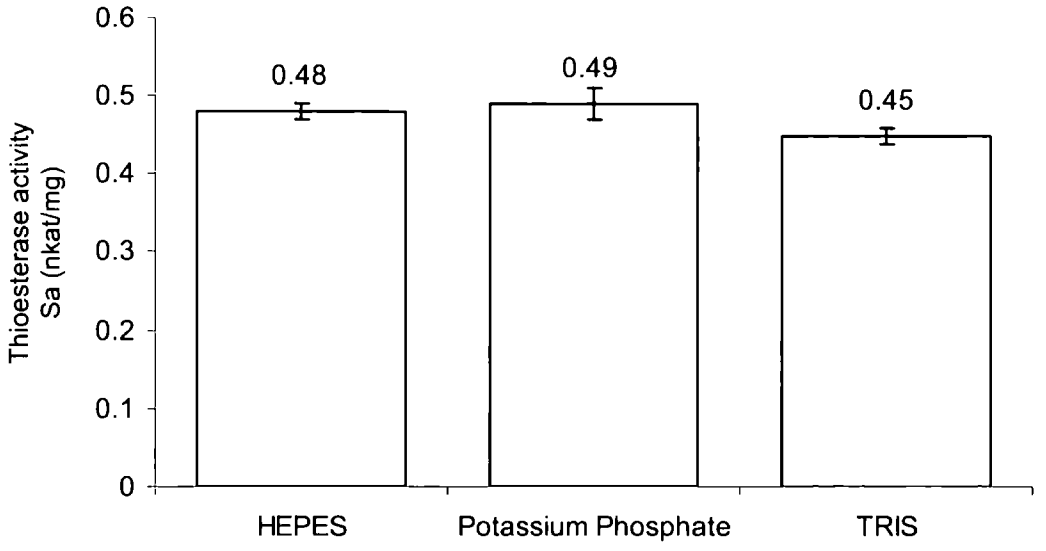
Figure 3.17 The developmental regulation of thioesterase activity.

Thioesterase activity was present in under-ripe fruit and its expression remained static for the first 10 days of ripening. The subsequent rapid loss of activity was likely to be a consequence of the low water content of the fruit. It appeared that thioesterase activity was expressed constitutively throughout ripening, with a loss of activity accounted for by the degradation of the fruit when over-ripe.

3.4.3 Optimizing the Ellman's Assay: Buffers and Co-factors

In order to ensure optimal thioesterase activity was being quantified in the Ellman's assay, a range of different buffers were tested and the enzymes dependence on co-factors was determined. Figure 3.18 illustrates the effect of different buffers and co-factors on enzyme activity.

A.



B.

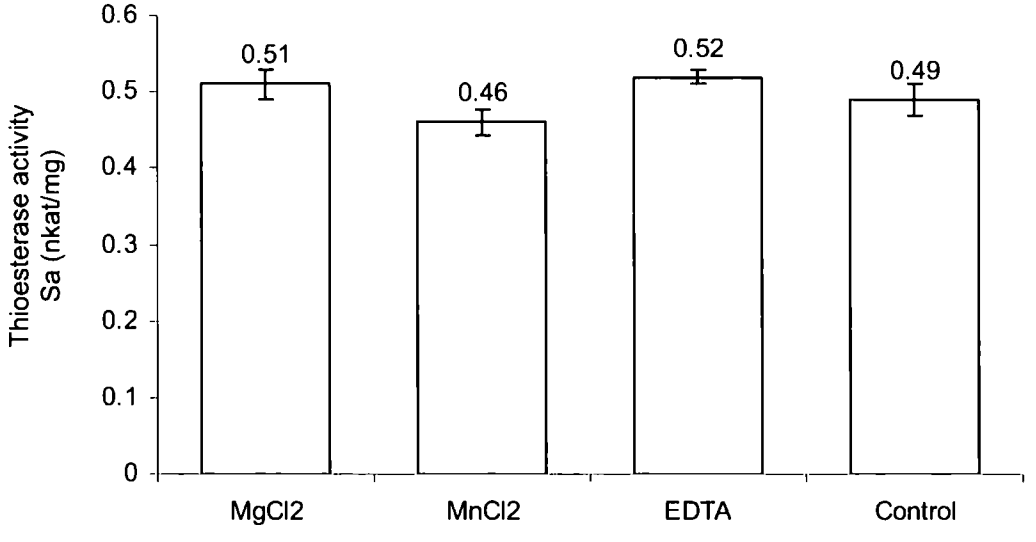


Figure 3.18 The effect of buffers (**A.**) and co-factors (**B.**) on thioesterase activity. Precipitated protein was resuspended (0.12 mg/ml) in different buffers at 0.1 M concentration and assayed for thioesterase activity toward MFTA. Co-factor salts were added to protein samples (0.12 mg/ml) in 0.1M phosphate buffer to a final concentration of 10 mM and assayed for activity.

Changing the buffer or adding co-factor had no effect on enzyme activity and for this reason future assays were conducted in 0.1 M potassium phosphate buffer, pH 7.2 without co-factor.

3.4.4 Substrate Specificities

A useful biocatalyst should accept a variety of substrates, allowing it to work on various starting materials, and ideally work with high regio- or stereo-selectivities. The hydrolytic properties of the passion fruit crude protein extract was therefore investigated by assaying toward various thioester substrates, including straight chain natural products, cyclic compounds and various artificial esters and amides (Table 3.4). It was of interest to determine the range of chemical reactions achievable with crude extracts. Would the extract function as a carboxylic ester hydrolase or amidase as well as a thioesterase? Furthermore, it was important to determine any preference for specific chemical types (e.g. straight-chain, bulky or cyclic). Crude protein solutions were therefore incubated with various substrates and specific activities determined (Table 3.4), the activities of PLE and CRL were also determined for comparison.

The passion fruit extract showed greatest activity toward the thioester 2-methyl furan-3-thioacetate (MFTA, **2**, 0.97 nkat/mg), the model substrate chosen for determining thioesterase activity. When the delocalisation of the furan ring is removed as in **3** activity was dramatically reduced (7-fold), presumably as the thiolate anion is no longer a stable leaving group. The second greatest activity was toward the ester substrate **8**, *p*-nitrophenylacetate (*p*-NA, 0.67 nkat/mg), again a relatively bulky cyclic compound with a short C2 acetate moiety as the acyl component. Only trace activities toward 4-methylumbelliferyl esters were detected through a sensitive fluorimetric assay (Methods 2.3.6), it appears the protein will not accommodate these larger compounds. This is unlike PLE, which accommodates these bulky substrates and was most active toward **10** (1589.26 nkat/mg). Interestingly, the fruit extract appeared to show a preference for short-chain acyl components in thioesters/ esters, as demonstrated by the reduction in activities between the ester series **9-11** (C2: 0.04 nkat/mg, C3: 0.03 nkat/mg, C6: 0.0005 nkat/mg).

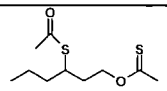
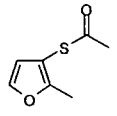
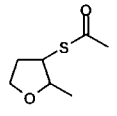
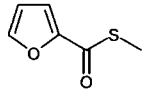
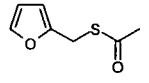
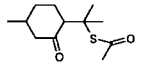
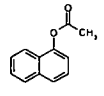
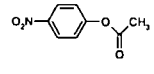
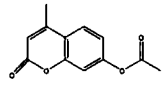
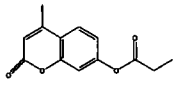
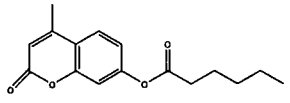
Compound number	Thioester/ ester substrates	Crude passion fruit extract (S.a. nkat/mg)	Porcine liver esterase (S.a. nkat/mg)	*Candida rugosa* lipase (S.a. nkat/mg)
1	 3-(thioacetyl) hexylacetate	0.04 +/- 0.002	14.02 +/- 0.72	ND
2	 2-methyl furan-3-thioacetate	0.97 +/- 0.03	537.88 +/- 24.65	0.35 +/- 0.02
3	 Tetra hydro 2-methyl furan-3-thioacetate	0.14 +/- 0.01	58.71 +/- 5.68	ND
4	 Methyl thiofuroate	ND	100.0 +/- 9.75	ND
5	 Furfuryl thioacetate	0.24 +/- 0.04	35.04 +/- 1.98	0.05 +/- 0.01
6	 8-(acetylthio) menthone	0.13 +/- 0.02	4.17 +/- 0.02	0.02 +/- 0.01
7	 α-Naphthyl acetate	0.26 +/- 0.9	159.4 +/- 16.14	0.58 +/- 0.01
8	 *p*-Nitrophenyl acetate	0.67 +/- 0.10	2117.6 +/- 87.5	5.14 +/- 0.23
9	 4-Methylumbelliferyl acetate	0.04 +/- 0.003	426.15 +/- 72.55	4.26 +/- 0.18
10	 4-Methylumbelliferyl propanoate	0.03 +/- 0.008	1589.26 +/- 102	118.36 +/- 21.39
11	 4-Methylumbelliferyl hexanoate	0.0005 +/- 0.0001	1302.97 +/- 97	36.59 +/- 4.00

Table 3.4 Enzymatic hydrolysis of thioester and ester substrates by a passion fruit crude protein extract, a mammalian esterase (PLE) and a microbial lipase (CRL). Activities are means of triplicates +/- standard deviation from the mean. ND; no activity detected.

The passion fruit extract showed less carboxylester hydrolase activity than PLE but was active toward various thioesters including 3-(thioacetyl)hexyl acetate, a precursor to the major character impact volatile in this fruit, 3-mercatohexanol. Activities were however low at this stage, a likely consequence of the low protein content of fruit peel. A large scale protein extraction and purification of the desired activity was therefore required when comparing activities with the commercial enzymes. However, at this stage the fruit preparation demonstrated different hydrolytic properties to the commercial enzymes with a relatively relaxed specificity for chemical types, a useful property for an efficient biocatalyst. It was unclear whether activities were due to specific thioesterases or multiple enzyme activities derived from different families of enzymes, this was investigated further through the chemical inhibition of specific enzyme types.

3.4.5 Enzyme Inhibition

Having demonstrated the crude protein preparation from passion fruit expressed both thioesterase and carboxylesterase activity it was of interest to investigate the catalytic mode of action of the enzyme pools which accounted for the observed activities. This was achieved using enzyme specific inhibitors, such as the insecticide paraoxon, which covalently deactivates serine catalytic residues in serine hydrolases[143]. Protein preparations were incubated with different inhibitors (1 mM) for 2 hours at 30°C before desalting to remove the unbound inhibitor and assaying toward MFTA and *p*-NA (Table 3.5).

To test for the presence of a catalytic cysteine, mercuric chloride and iodoacetamide were tested as inhibitors of thioesterase activity, resulting in 25% and 0% inhibition respectively toward MFTA, suggesting a catalytic cysteine was not present in the majority of thioesterase activity. Conversely, the organophosphate insecticide paraoxon reduced hydrolysis by 87%, indicating a catalytic serine was present in the enzyme(s). The serine hydrolase inhibitor fp-biotin only knocked out 38% of activity, a possible consequence of the long chain structure of the inhibitor which may not be accommodated so readily in the active site as paraoxon.

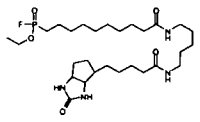
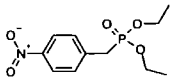
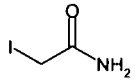
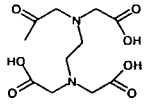
Inhibitor (hydrolase class inhibited)	Percentage Inhibition (%) thioesterase (MFTA)	carboxylesterase (*p*-NA)
 Fp-biotin (serine)	38 +/- 1	83 +/- 1
 Paraoxon (serine)	87 +/- 3	76 +/-2
 Iodoacetamide (cysteine)	ND	ND
$HgCl_2$ (cysteine) Mercuric chloride (cysteine)	25 +/- 1	17 +/-1
 EDTA (metallo hydrolase)	ND	ND

Table 3.5 The inhibition of thioesterase and carboxylesterase activity in crude passion fruit preparations. Structures of inhibitors and the hydrolase class they are active toward. Percentage inhibition was calculated from the loss of activity and is presented as the means of triplicates +/- standard deviation from the mean. ND; no inhibition detected.

It does however appear likely a serine residue is important in the catalytic mechanism of these fruit thioesterases, although further studies on a purified fraction were required to confirm this.

Carboxylesterase activity in the peel extracts showed a similar pattern of inhibition, being predominantly effected by serine hydrolase inhibitors such as paraoxon (77% inhibition) and fp-biotin (83% inhibition), with cysteine and metallo hydrolase inhibitors having little or no effect on activity. The greater inhibition of esterases with fp-biotin (83%) compared to the thioesterase pool (38%) indicated these were a separate family of enzymes with different active site pockets. The thioesterase(s) appeared to be more accommodating of small, cyclic, short chain compounds, whereas the carboxylesterases appeared less selective and accommodated the longer fatty straight chain compounds. Both however appear to use a catalytic serine, as demonstrated by the near complete inhibition of both activities by paraoxon.

3.4.6 Thioesterase Stability

One of the major drawbacks of biocatalysts is their reduced shelf-life compared to synthetic catalysts. This has been overcome in many cases by immobilising enzymes on resin supports[144]. However, only so much improvement to enzyme stability can be achieved and essentially a robust protein is desirable from the offset. For this reason it was important to determine the stability of the thioesterase activity in passion fruit. Protein preparations were incubated at 30°C and assayed every 24 hours for both thioesterase (MFTA) and esterase (*p*-NA) activity (Figure 3.19). Thioesterase activity toward MFTA was remarkably stable with a half-life of 4.5 days (30°C), whereas the esterase activity fell sharply over the first day and then decayed with a half life of 14 hours. Interestingly thioesterase activity showed a different stability decay curve to that of the carboxylesterases, a further indication separate pools of hydrolase activity were present in the fruit. Although each pool may demonstrate both carboxylester and thioester hydrolase activities there appeared to be significant differences between enzymes that acted on carboxylic esters and those that acted on thioesters.

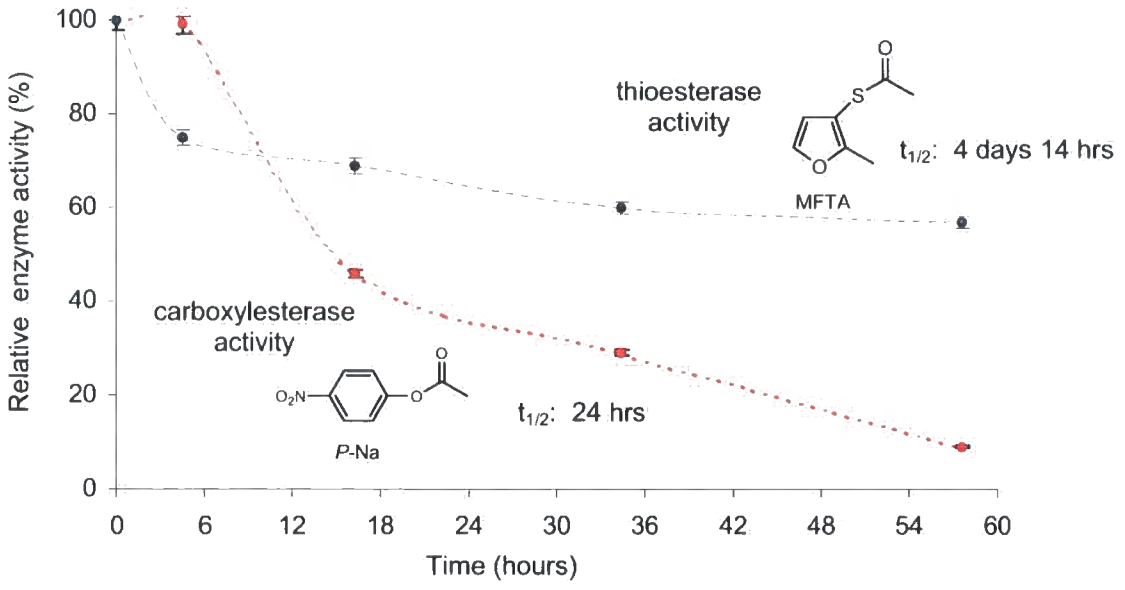


Figure 3.19 The stability of thioesterases and esterases in passion fruit protein preparations. Protein samples (0.15 mg) in 0.1 M phosphate buffer were incubated at 30°C and assayed for thioesterase activity toward MFTA (blue spots) and esterase activity toward *p*-NA (red spots). The rate of loss of activity is illustrated by half life values ($t_{1/2}$). Assays were run in triplicate and corrected for chemical rate by subtracting the activity of a heat treated sample (95°C, 5min).

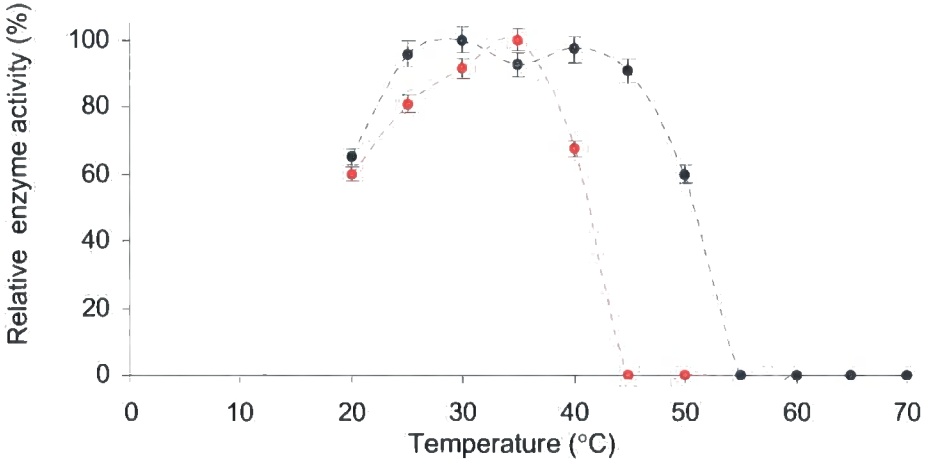


Figure 3.20 The temperature tolerance of carboxylesterases (red) and thioesterases (blue) in passion fruit. Protein samples (0.15 mg) in 0.1 M phosphate buffer were incubated at 20-70°C and assayed for thioesterase activity toward MFTA and esterase activity toward *p*-NA. Samples were incubated for 5 min at each 5°C increment. Heat treated sample were run in parallel as a means of determining relative chemical rates.

The initial sharp drop in thioesterase activity may be a consequence of a separate pool of thioesterase enzymes or a result of the loss of the carboxylesterase pool which also had this activity.

The temperature tolerance of biocatalysts is also an important property, those enzymes that can tolerate higher temperatures can be used in reactions at increased temperature which promotes faster reactions and hence greater productivity. This is why numerous biocatalysts are sourced from thermophilic microbes which contain proteins able to tolerate remarkably high temperatures (> 70°C)[111]. In the case of manufacturing flavouring volatiles, the extraction of products from the reaction mixture can be facilitated by increasing the temperature to favour volatilisation and extraction of products from the reaction medium. An enzyme that can tolerate an increased temperature would favour such processes. Protein samples were therefore incubated in a water bath at increasing temperatures (20-70°C) and following equilibration at each temperature carboxylesterase and thioesterase activity was determined (Figure 3.20). Both pools of enzymes show optimal activity around 37°C, with the thioesterase(s) able to tolerate slightly higher temperatures with over 50% activity remaining at 50°C.

3.5 Discussion

Identifying and quantifying a potentially trace activity in an unusual (and problematic) plant source was a challenge and required the development of a sensitive assay. However, through adapting the Ellman's disulphide exchange method it was possible to accurately and reproducibly determine thioesterase activity in fruit extracts. Such a simple quantification technique forms a novel means of analysing the enzymatic hydrolysis of VOSCs and is also likely to have further application in the flavourings and food industries. It serves as a simple means of determining the thiol content of food or flavouring products, raw materials, and effluent streams[145]. The decomposition of thioester flavourings is a common problem encountered by the industry[114] and could be

tested for using this simple visual test with nothing more than a colorimeter or spectrophotometer.

Partial characterisation of the thioesterases in purple passion fruit indicated activity was partially cell wall bound, requiring the addition of NaCl for its complete dissociation from the insoluble matrix. Furthermore the activity was found to be predominantly localised to the mesocarp tissue of the peel. A network of vascular tissue runs throughout the mesoocarp which presumably feeds the fruit with water and nutrients for development. It is possible the nutrients fed into the cortex are metabolized further here into flavours and fragrances before secretion into the aromatic juice of the interior. It is interesting activity is found predominantly in this tissue where secretory glands are also present.

Thioesterase activity appears to be distinct from further pools of hydrolases (carboxylesterases) within the fruit peel as indicated by the differences in their stabilities and inhibition patterns. However, further characterisation of the purified activity is required in order to determine how thioesterase activity differs from that of other fruit hydrolases. Activity was demonstrated to act on a wide spectrum of chemical thioesters including the plant natural products 3-(thioacetyl)hexylacetate (compound **1**, passion fruit[119]) and 8-(acetylthio) menthone (**6**, buchu plant[120]). Such diversity of accommodation in combination with the enzymes stability (4.5 days, 30°C) provides promising characteristics required of any successful biocatalyst.

Work will now concentrate on the purification of thioesterase activity from purple passion fruit as a means of characterising the enzyme(s) further.

4. Purification of a Thioesterase from Purple Passion Fruit (*Passiflora edulis* Sims)

4.1 Introduction

The presence of thioesterase activity in the mesocarp tissue of ripe purple passion fruit provided the first evidence that hydrolases could be involved in the release of VOSCs from thioester precursors in tropical fruit. Furthermore, this activity was demonstrated to hydrolyse natural thioester substrates including 3-thioacetyl hexylacetate, a precursor of the distinctive aroma volatile in passion fruit, 3-mercaptohexanol. However, further characterization of the enzyme(s) responsible for this activity was required in order to confirm a physiological role in the fruit. Isolating the enzyme(s) associated with thioesterase activity through protein purification would therefore aid in the characterisation and identification of the novel activity.

Localisation of thioesterase activity to the cortex of the fruit is a characteristic of several families of fruit hydrolases. For example, in apple (*Malus pumila*) butyl esters are hydrolysed to release butanol into the surrounding air by undefined esterases in the cortex and peel[93]. In addition, the outer cortex of fruit is also known to contain cell wall hydrolases which control fruit softening during ripening through the degradation of structural polysaccharides. Hydrolases have previously attracted particular attention as targets for controlling fruit softening and increasing shelf life through genetic engineering[146]. Tomato has become the model system for studying fruit glycosidases (EC 3.2.1), but these enzymes have also been purified and characterized from further fruits including polygalacturonases from peach (*Prunus persica*)[147], β-glucosidases and β-galactosidases from sweet cherry (*Prunus avium*)[148], β-1,3-glucanases from banana (Musa acuminate)[149] and pectin esterases from Valencia orange (*Citrus sinensis*)[150]. The fruit exocarp is also found to contain chitinases, as characterized in grapefruit (*Citrus paradisi*)[151], where they serve as a first line of defence against fungal and insect pathogens. In addition to being localised to the periphery of fruit, such carbohydrate esterases are also found to be cell wall bound enzymes which

require the addition of salt (0.25-1 M NaCl) to liberate them from the complex polysaccharide matrix[148,150].

Hence, the thioesterase activity determined in passion fruit shares many common characteristics with other fruit hydrolases. The identity of the enzyme will therefore help in determining whether this is a novel esterase or an adaptation to an existing form recruited here for a different function.

Determining the catalytic mode of action of thioesterases in passion fruit will also help in the characterization of the enzyme. Through studying the inhibition of crude protein activity it is already apparent a catalytic serine is likely to be involved in catalysis, as observed through inhibition of thioesterase activity (87%) following treatment with the organophosphate insecticide paraoxon. The use of chemical probes provides a powerful tool for isolating specific enzymes and they will be recruited here to characterize the active site of the purified protein. A further property of the enzyme which is somewhat unusual is its preference for short chain acyl components in thioesters, such as acetyl groups, which is an indication of a constrained acyl-binding pocket. It is hoped the isolation of activity will help to resolve the mode of action of the enzyme and the cause of such tight chemoselectivity.

Having developed a reliable and simple assay for monitoring thioesterase activity in fruit homogenates, the next step was to purify the enzyme(s) using classical biochemical techniques. The purification of fruit cell wall hydrolases has previously utilised techniques such as ion-exchange chromatography (anion and cation exchange), gel filtration and affinity to concanavalin A-sepharose[150,152]. However, the specific properties of hydrolases vary greatly between families and even different isoforms of the same enzymes are known to posses considerably different characteristics, as observed with the acidic and basic pools of pectin esterases[150]. Hence, the properties of thioesterase enzymes can not be presumed to be similar to those of other hydrolases in fruit and a purification protocol must be developed empirically.

This chapter describes the methodology used to maximise the extraction of thioesterase activity from passion fruit peel, followed by the biochemical techniques used to purify it to homogeneity. Through further characterisation of purified fractions in combination with protein sequencing using tandem mass spectrometry it was hoped the identity of thioesterase activites could be determined. The generation of partial protein sequence would also enable the design of degenerate oligonucleotide primers for amplifying the corresponding gene sequence from passion fruit, thus enabling the properties of the recombinant protein to be examined in a bacterial or yeast expression system.

4.2 Extraction and Purification of Thioesterases from *Passiflora edulis* Sims

4.2.1 Optimising the Extraction of Thioesterase Activity

Prior to embarking on the purification of thioesterase enzymes it was important to maximise the extraction of the associated activity from the fruit peel. As has been described, this activity may be accounted for by a relatively small amount of protein in the mesocarp (low abundance, high activity), due to the high-impact nature of the products formed. Protein purification protocols inevitably result in a loss of recovery of the desired enzyme at each stage and in order to ensure sufficient protein is present in the final preparation required for enzyme identification through mass spectrometry, it was important to start with maximum protein activity. In addition, for the reliable calculation of enzyme activity per gram of fruit tissue it was important to ensure complete extraction of active protein. The extraction and purification of thioesterase activity toward the model substrate MFTA was chosen due to the sensitivity of the assay, which would aid in isolating low abundance activities.

As has been described, initial studies determined that activity was greatest in fruit at a maturity corresponding to 30% dry weight, being predominantly confined to the mesocarp. Hence, this fruit tissue was harvested and the juice and seed removed prior to homogenisation using a waring blender on ice in 0.1 M phosphate extraction buffer (4:1, volume:weight). As noted previously,

the addition of NaCl (0.25 M) increased the liberation of thioesterase enzymes (2.5-fold) from the peel, and hence it was added to the extraction buffer.

The binding of thioesterase activity to the insoluble cell wall matrix in homogenates was similar to the way in which proteins adsorb to ion-exchange columns and it was therefore postulated that the release of bound protein from this material may form a useful purification step in itself. Peel was therefore homogenized in a waring blender without salt and then filtered through Miracloth (Calbiochem) in order to wash off any unbound protein. The insoluble peel tissue was retained and resuspended in extraction buffer containing 0.25 M NaCl and stirred for 2 hours on ice to solubilise the adsorbed activity before filtering through Miracloth to yield a supernatant containing solubilised cell-wall bound protein (Figure 4.1). The supernatant fraction was subsequently analysed for activity and although this technique did remove much of the unwanted protein resulting in a 3 fold purification of activity, only 32% of activity was recovered. Thioesterase activity appeared to be present in two pools, an insoluble cell-wall bound fraction and a soluble fraction. It was therefore decided that at this stage it was important to retain both activities as a single pool and to separate them more efficiently during column chromatography. Hence, a single wash containing 0.25 M salt was used to extract both soluble and wall-bound proteins together.

Passion fruit, like most fruit, contains a lot of pectin which forms a major structural component of the cell wall. This material is however problematic for protein purification as it blocks up columns interfering with the even loading of samples and enzyme affinities for the solid matrix. A number of protocols were therefore tested to remove the pectin, including adding protamine sulphate, a highly cationic peptide which often enhances sedimentation of poorly soluble macromolecules in solution[153]. Following extraction with 10% (w/v) protamine sulphate and centrifugation (10,000g, 15 min) the supernatant was partially clarified with the pellet containing a jelly like layer of pectin. However, no detectable thioesterase activity remained in solution and the activity that precipitated in the pellet was low (~5% total crude activity). This protocol appeared to deactivate or denature the enzyme.

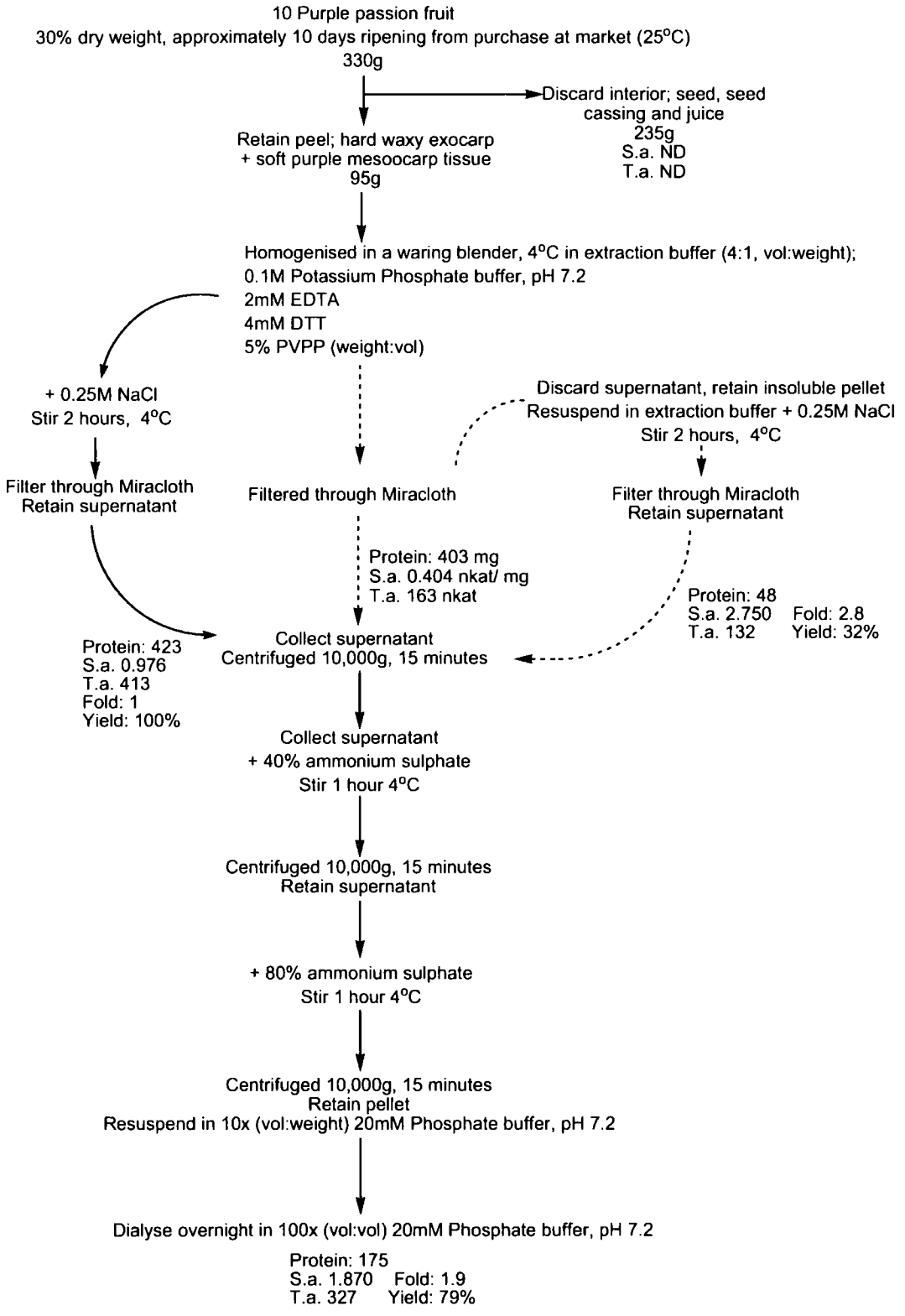


Figure 4.1 Optimising the extraction of thioesterase activity from passion fruit. Solid arrows indicate the extraction protocol used prior to column purification, dotted arrows indicate protocols tested but not used due to greater losses of activity or reduced fold purification in comparison to the method of choice. The specific thioesterase activity toward MFTA (S.a., nkat mg^{-1} protein) and the total activity (T.a., nkat) following protein extraction from 10 fruit is shown. The fold purification (fold) and recovered yields (yield) of activity were determined.

The most efficient solution was found to be the simple "salting out" (precipitation) of protein using ammonium sulphate. The recovery of proteins precipitating between 40-80% salt saturation was found to remove a proportion of unwanted protein, which was recovered in the 0-40% fraction, and cleared the supernatant of the majority of pectin. This was an efficient purification step, resulting in a 2-fold enrichment of activity with a 79% recovery. The salting out of proteins using ammonium sulphate is a widely used purification technique and here it was employed to increase fold purification and remove unwanted polysaccharide material. Some pectin did however remain in solution and it could often be seen as a thin layer on top of the sepharose following chromatography. This small amount was not a problem on large columns, however, it was important to clarify the extract prior to loading onto the expensive high-performance columns later in the purification.

A common problem encountered when making protein extracts from plant material is the deactivation and precipitation of proteins through oxidation, a destructive process that is apparent through the "browning" of an extract. These undesirable effects can however be avoided through the addition of additives to the extraction buffer which reduce oxidative processes[154]. The metal ion chelator ethylenediaminetetraacetic acid (EDTA) was added (2 mM) to sequester heavy metal ions (often present as contaminants in buffer salts and ammonium sulphate) which catalyse the non-enzymatic oxidation of thiol groups on proteins leading to their inhibition or inactivation. In addition, phenolic compounds in plant extracts readily form covalent bonds with proteins and lead to their aggregation and precipitation, in order to prevent this the phenolic metabolite adsorbent polyvinylpolypyrrolidone (PVPP) was added (5% w/v). The final measure was to add 1,4-dithiothreitol (DTT, 4 mM) which functions as an inhibitor of polyphenol oxidases which account for the enzymatic oxidation and inactivation of enzymes. In combination these additives prevented the browning of fruit extracts and the associated deactivation of enzymes, this was of particular importance as the crude extract was left to stir in a salt solution for two hours. Following precipitation of thioesterase activity in the 40-80% ammonium sulphate fraction, the protein pellet was resuspended in a low molarity buffer (20 mM phosphate, pH 7.2)

and dialysed overnight to remove any traces of salt and other additives (EDTA, PVPP, DTT) which were no longer required. It was of particular importance to remove both salt and DTT which were observed to interfere with the Ellman's assay, the latter a consequence of thiol functionality in the compound.

An overview of the optimal extraction protocol used for the purification of thioesterase activity from passion fruit is depicted in figure 4.1. In this case the procedure enabled 175 mg of protein to be extracted from 10 passion fruit (330 g whole fruit/ 95 g peel) with a final thioesterase activity recorded as 1.870 nkat mg^{-1}, a result of a 2-fold purification and 79% recovery of activity as compared with the original crude extract. A 2-fold purification step giving 79% recovery was not as efficient as could be hoped but it should be noted that it was required for the removal of pectin and clarification of the extract. In addition, simply removing the interior of the fruit (seed and juice) removed 70% of unwanted tissue (by weight) and resulted in a 3.5-fold purification. The level of soluble protein extracted from the peel was low. However, through starting with a large scale extract, it was hoped that thioesterase activity could be purified to a level that could be analysed by tandem mass spectrometry (MS/MS sequencing). It was calculated that 25 fruit would provide approximately 1 mg of pure protein at a purification efficiency of 0.1%, more than enough to ensure characterisation and identification of the protein.

4.2.2 Column Chromatography

The wide range of adsorbent materials suitable for trapping proteins has made column chromatography the most efficient means of separating enzymes[154]. A protein's affinity to adsorb to a variety of solid phases is dependent upon its surface properties (charge, hydrophobicity) and these can be used to separate single constituents from the protein mixture. In addition, the elution of bound protein can be controlled through changing the buffer conditions (ionic strength or pH) over a continuous gradient as a means of increasing the resolution of proteins further.

A purification procedure for isolating thioesterase enzymes was essentially developed through trial and error. However, it was known that a proportion of

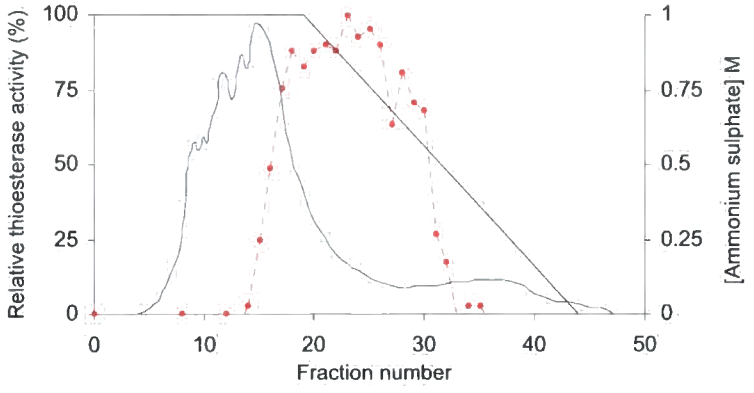
activity was cell-wall bound suggesting it may be highly charged (ionic) in nature. A number of different column adsorbents were therefore tested prior to establishing an efficient sequential purification protocol.

Firstly the dialysed crude protein extract was applied to hydrophobic columns which separate proteins based on differences in their surface hydrophobicities (Figure 4.2A). Different ligands immobilised on sepharose beads were tested for their ability to resolve thioesterase activity, including butyl, octyl and phenyl substituents. In each case the dialysed crude protein extract (8 ml) was passed through a relatively small column (10 ml) at a flow rate of 1 ml/minute with bound protein eluted with a linearly decreasing concentration of ammonium sulphate (1-0 M) over 50 ml. The elution of MFTA-thioesterase activity in 1ml fractions was monitored using the Ellman's assay. A broad peak of activity (20 ml) was partially bound to the butyl sepharose column with some eluting in the flowthrough (void) and the remainder eluting between 1-0.5 M ammonium sulphate. Hence, much of the unwanted protein was not removed and only a 3.5-fold purification was achieved with recovery of 42% of the activity. The longer aliphatic chain of octyl sepharose provided a more hydrohobic matrix and was more efficient at adsorbing activity, a single tight peak eluted in 0.05 M ammonium sulphate, after the majority of soluble protein. Hence, this was an efficient column purification step resulting in a 9-fold enrichment of activity with 77% recovery. On the phenyl sepharose column, activity eluted midway through the gradient of mobile phase in 0.36 M ammonium sulphate, achieving a 4.2-fold purification with 56% of the activity recovered.

The behaviour of thioesterase activity on ion-exchange chromatography was also investigated, whereby proteins were separated on the basis of differences in their charge characteristics (cationic/anionic), in this case on a diethylaminoethyl (DEAE) weak anion-exchange column (Figure 4.2B). A dialysed crude sample (140 ml) was applied to a larger (43 ml) column at a flow rate of 4 ml/minute as a means of determining the capacity of the column. The bound protein was eluted using a linearly increasing gradient of sodium chloride (0-0.5 M) over 300 ml.

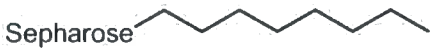
A. Hydrophobic column chromatography

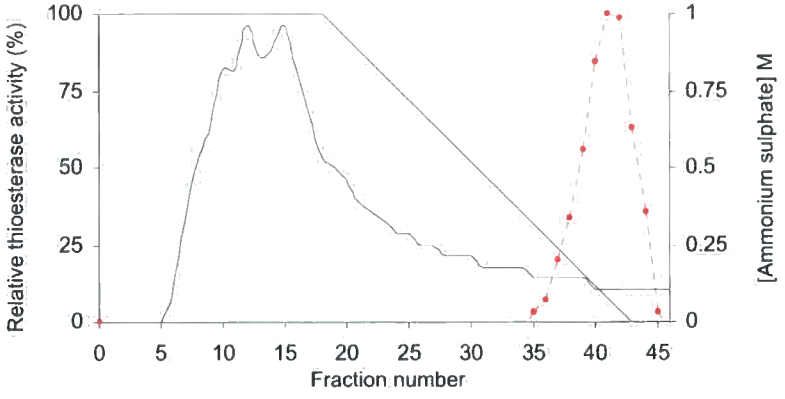
Butyl sepharose:



Fraction ID	Vol (ml)	Total protein (mg)	S.a. (nkat mg^{-1})	T.a. (nkat)	Fold purification	Recovery
Crude	8	12.05	0.72	8.31	-	100%
Column fractions 15-35	42	1.46	2.5	3.65	3.6	44%

Octyl sepharose:

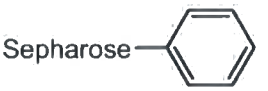


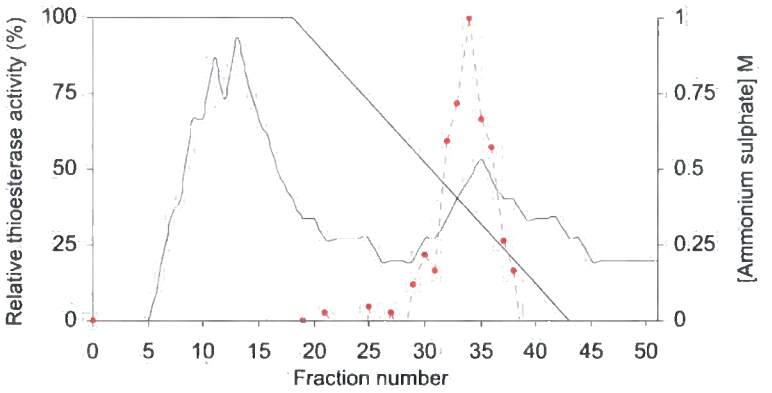


Fraction ID	Vol (ml)	Total protein (mg)	S.a. (nkat mg^{-1})	T.a. (nkat)	Fold purification	Recovery
Crude	8	12.05	0.72	8.67	-	100%
Column fractions 35-45	22	1.02	6.67	6.67	9.1	77%

Figure 4.2 Protein purification - column chromatography. The behaviour of thioesterase activity on hydrophobic interaction columns (**A.**) and on a DEAE anion exchange column (**B.**). Blue lines show the elution of UV absorbing protein (280nm), red dotted lines indicate thioesterase activity and black lines represent the gradient of mobile phase.

Phenyl sepharose:

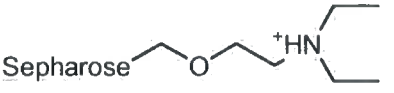


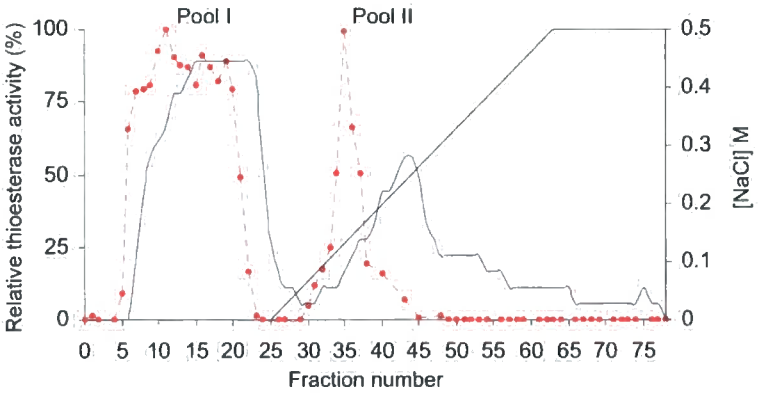


Fraction ID	Vol (ml)	Total protein (mg)	S.a. (nkat mg^{-1})	T.a. (nkat)	Fold purification	Recovery
Crude	8	12.05	0.72	8.67	-	100%
Column fractions 29-38	20	1.61	3.02	4.87	4.2	56%

B. Anion exchange chromatography

DEAE sepharose:





Fraction ID	Vol (ml)	Total protein (mg)	S.a. (nkat mg^{-1})	T.a. (nkat)	Fold purification	Recovery
Crude	140	145	0.37	53.65	-	100%
Pool I: column fractions 5-22	144	43	0.67	28.8	1.8	54%
Pool II: column fractions 30-43	112	14	1.63	22.78	4.4	42%

Interestingly this chromatography split the activity into two peaks, a basic fraction eluting in the void volume (0 M salt) and a more acidic fraction eluting with 0.14 M NaCl (Figure 4.2B). To test for the possibility that the lack of retention was due to over-loading of the column, this fraction was reapplied to freshly conditioned DEAE sepharose. Again the activity was unretained suggesting that the respective enzyme was basic in nature. The first pool (54% activity) eluted with the majority of unbound protein and hence a fold purification of only 1.8 was achieved as opposed to the second fraction which eluted later with a greater fold purification of 4.4.

After examining the characteristics of thioesterase activity on the different columns, a sequential purification protocol was chosen. Twenty-five ripened purple passion fruit were used in a large scale purification (Methods 2.4.1). Following the removal of juice and seed material, this left 245 g of peel tissue from which the enzyme would be extracted. A crude protein extract was prepared (Figure 4.1) and following precipitation of the enzyme with ammonium sulphate and overnight dialysis the thioesterase was sequentially purified using DEAE anion exchange chromatography followed by hydrophobic interaction chromatography; firstly using octyl sepharose and then a high performance phenyl suparose column (GE Healthcare).

The crude protein preparation was applied to the DEAE sepharose column (43 ml) at a flow rate of 4 ml/min with 8 ml fractions collected and assayed for activity. As before thioesterase activity eluted in an unbound and bound pool, referred to as pools 1 and 2 respectively (Figure 4.3). The unretained pool 1 activity accounted for slightly more of the recovered total activity (59%) and eluted over multiple fractions in the void volume. The bound pool eluted with 0.196 M salt and was found to have a considerably higher specific activity (8.07 nkat mg^{-1}) than that of pool 1 (3.47 nkat mg^{-1}). However, at this stage the purification of both pools was pursued so as to distinguish and identify both forms of activity.

Ammonium sulphate (1 M) was added to both pools of activity before they were separately applied onto an octyl sepharose column (47 ml) at a flow rate

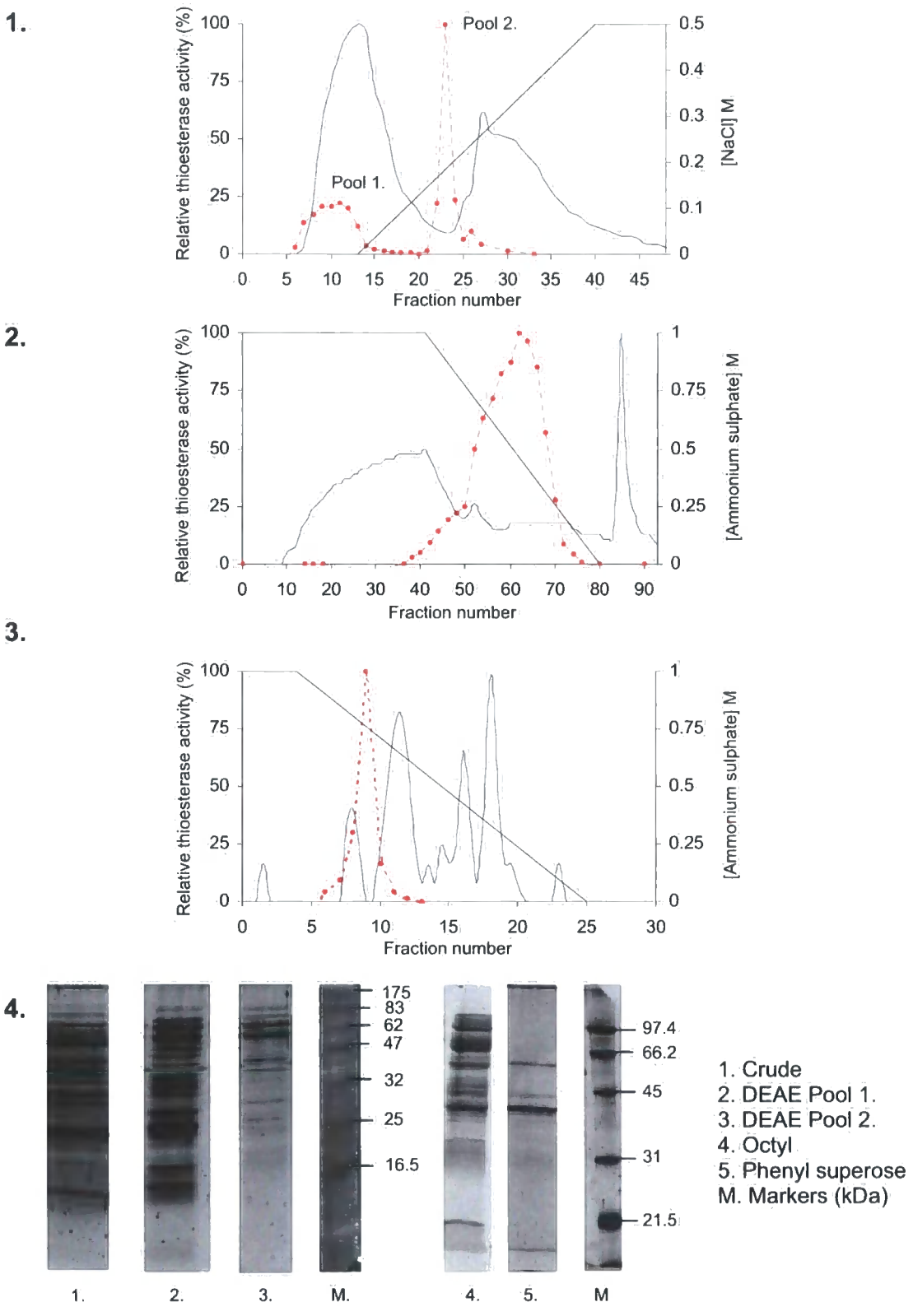


Figure 4.3 A. Sequential purification of pool 1 thioesterase activity by chromatography on (**1.**) DEAE Sepharose, (**2.**) Octyl Sepharose and (**3.**) Phenyl Superose. Blue lines show the elution of UV absorbing protein (280 nm), red dotted lines indicate thioesterase activity and black lines represent the gradient of mobile phase. (**4.**) Analysis of polypeptides at each stage of purification via SDS-PAGE, lanes 1-3 12% gels, lanes 4+5 10% gels.

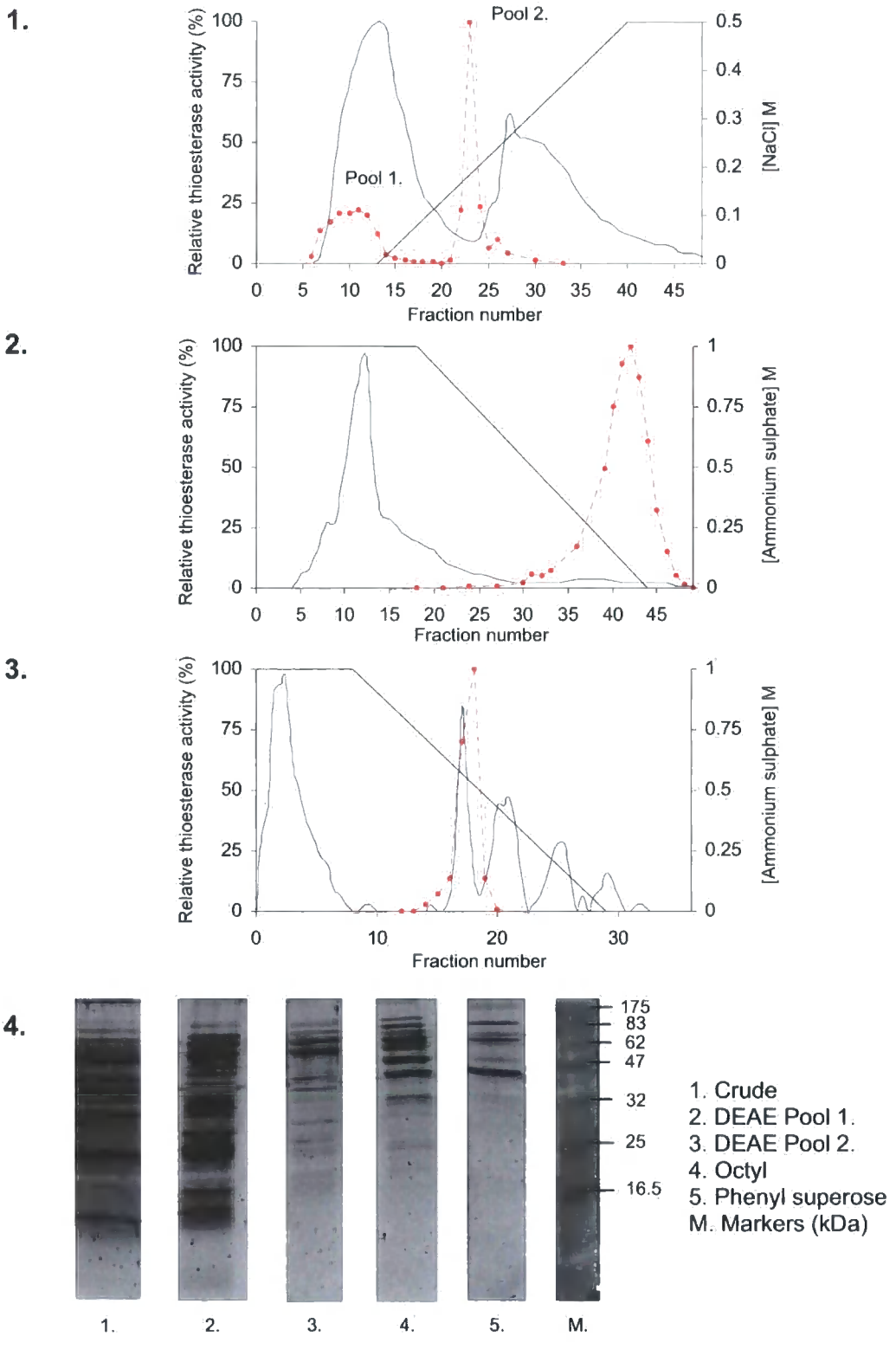


Figure 4.3 B. Sequential purification of pool 2 thioesterase activity by chromatography on (**1.**) DEAE Sepharose, (**2.**) Octyl Sepharose and (**3.**) Phenyl Superose. Blue lines show the elution of UV absorbing protein (280 nm), red dotted lines indicate thioesterase activity and black lines represent the gradient of mobile phase. (**4.**) The analysis of polypeptides at each stage of purification via SDS-PAGE (12% gel).

Preparation step	Total protein (mg)	Specific activity (nkat mg^{-1})	Total activity (nkat)	Purification (fold)	Yield (%)
Crude	1018	0.976	994	1	100
40-80% $(NH_4)_2SO_4$	421	1.870	787	1.9	79
DEAE Sepharose					
Pool 1	135	3.47	468	3.6	47
Pool 2	41	8.07	331	8.3	33
Octyl Sepharose					
Pool 1	14	16.1	225	16.5	22
Pool 2	6	38.8	233	40.0	23
Phenyl superose					
pool 1	0.8	55.5	44.4	57	1
Pool 2	0.2	149.38	30.9	154	3

Table 4.1 Summary of the purification of MFTA-thioesterase activity from purple passion fruit. 25 passion fruit (95 g peel) were processed as described in figure 4.1 to provide a crude protein extract. This was subsequently passed though DEAE, octyl and phenyl sepharose columns. Two separate pools of activity were resolved with 154 and 57 fold purification respectively.

of 4 ml/min with 8 ml fractions collected. Both pools of activity demonstrated similar characteristics, eluting with 0.49 M and 0.09 M ammonium sulphate respectively, an indication they were both relatively hydrophobic proteins, pool 2 slightly more so than 1. Octyl sepharose chromatography was efficient at removing much of the unwanted protein, however, it did result in a large elution volume (168 ml and 56 ml respectively). It was therefore necessary to concentrate the protein solution to 10 ml by ultrafiltration on centrifugal Vivaspin columns (Sartorius) prior to loading on the final column.

The final purification step using high-pressure liquid chromatography (HPLC) with a phenyl-superose column (0.69 ml) provided exceptional resolution of the remaining proteins. Pool 1 thioesterase activity bound to phenyl superose and was recovered with 0.82 M ammonium sulphate (figure 4.3A). Based on the final specific activity of the thioesterase (55.5 nkat mg^{-1}), the enzyme had been purified 57-fold in 1% yield. Pool 2 activity also bound to phenyl superose and eluted slightly later in 0.55 M ammonium sulphate (figure 4.3B). Based on the final specific activity of the thioesterase (149.38 nkat mg^{-1}), the enzyme had been purified 154-fold in 3% yield. A summary of the purification procedure is depicted in table 4.1 where the purification fold and recovery of total activity are shown for both pools of activity at each subsequent purification step.

4.2.3 SDS-PAGE Analysis of Purified Activity

The analysis of polypeptides present at each stage of the purification was achieved using SDS-PAGE (sodium dodecyl sulphate polyacrylamide gel electrophoresis). Protein samples denatured in SDS were resolved on either 10% or 12% acrylamide gels, the latter used when the resolution of smaller polypeptides was required. Polypeptides were subsequently stained with SYPRO Ruby protein gel stain (Invitrogen) and visualised using a UV transilluminator. This stain was chosen due to its ease of use and its high degree of sensitivity (minimum detection limit of 1-2 ng protein per gel band).

Gel images showing the total protein extracted from fruit peel and polypeptides in subsequent purified fractions highlight the efficiency of each purification step (Figure 4.3A+B). A number of polypeptides can be seen in the final preparation

of each pool indicating that both thioesterase enzymes had been purified to near homogeneity. However, to determine which polypeptides were associated with thioesterase activity individual fractions from the terminal purification steps were analysed by SDS-PAGE (Figure 4.4A+B). In both cases a 43 kDa polypeptide was identified whose relative abundance mirrored eluting thioesterase activity.

4.3 Characterisation of Purified Thioesterase

The purified thioesterases were termed *Passiflora edulis* Sims wall-bound hydrolase A and B (*Pe*WHa/ *Pe*WHb). Due to the greater specific activity demonstrated by *Pe*WHb this activity was selected to be characterised further. However, the identity of both pools was of interest, especially in determining whether *Pe*WHa was a separate isoform of the protein. Both purified polypeptides were therefore excised from the SDS-gels and processed prior to protein sequencing (Section 4.4).

Numerous biochemical techniques were used to characterise purified *Pe*WHb further. Firstly, the size of the active protein was investigated using gel filtration (Figure 4.5). This technique also referred to as size exclusion chromatography, separates proteins on the basis of mass, but unlike SDS-PAGE, it does not require proteins to be denatured. They are separated in solution in their physiological conformation. This was important in determining the oligomeric state of the protein, resolving whether activity was due to a single polypeptide or multiple sub-units. Therefore a semi-purified fraction of activity retained following octyl sepharose chromatography (10 ml) was applied to a superdex gel filtration column (GE Healthcare, 7.8 ml) in 50 mM phosphate buffer, pH 7.2 with 0.15 M NaCl at a flow rate of 1 ml/min. The elution of activity relative to standards of known size indicated that the native enzyme was of a size of approximately 20-32 kDa (Figure 4.5). Individual proteins of the same size are known to behave differently on size exclusion columns dependent upon their surface properties and conformation. This provided evidence that the active thioesterase enzyme was a monomer.

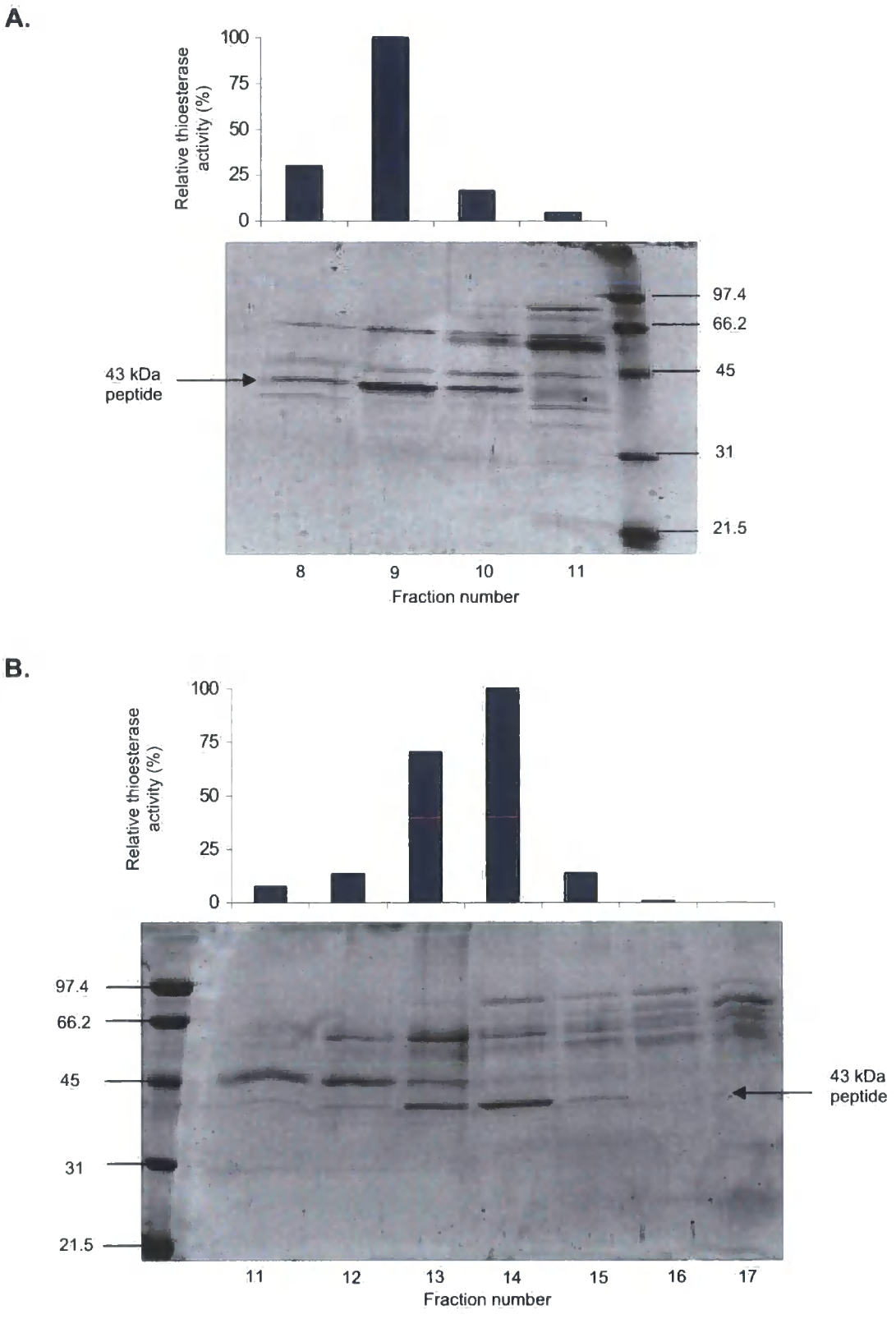


Figure 4.4 SDS-PAGE identification of active polypeptides in pool 1 (**A.**) and pool 2 (**B.**). The active fractions eluting from the final purification column (phenyl superose) were analysed via SDS-PAGE and when aligned with relative thioesterase activity a polypeptide of 43 kDa mirrored eluting activity, as highlighted by the arrows above.

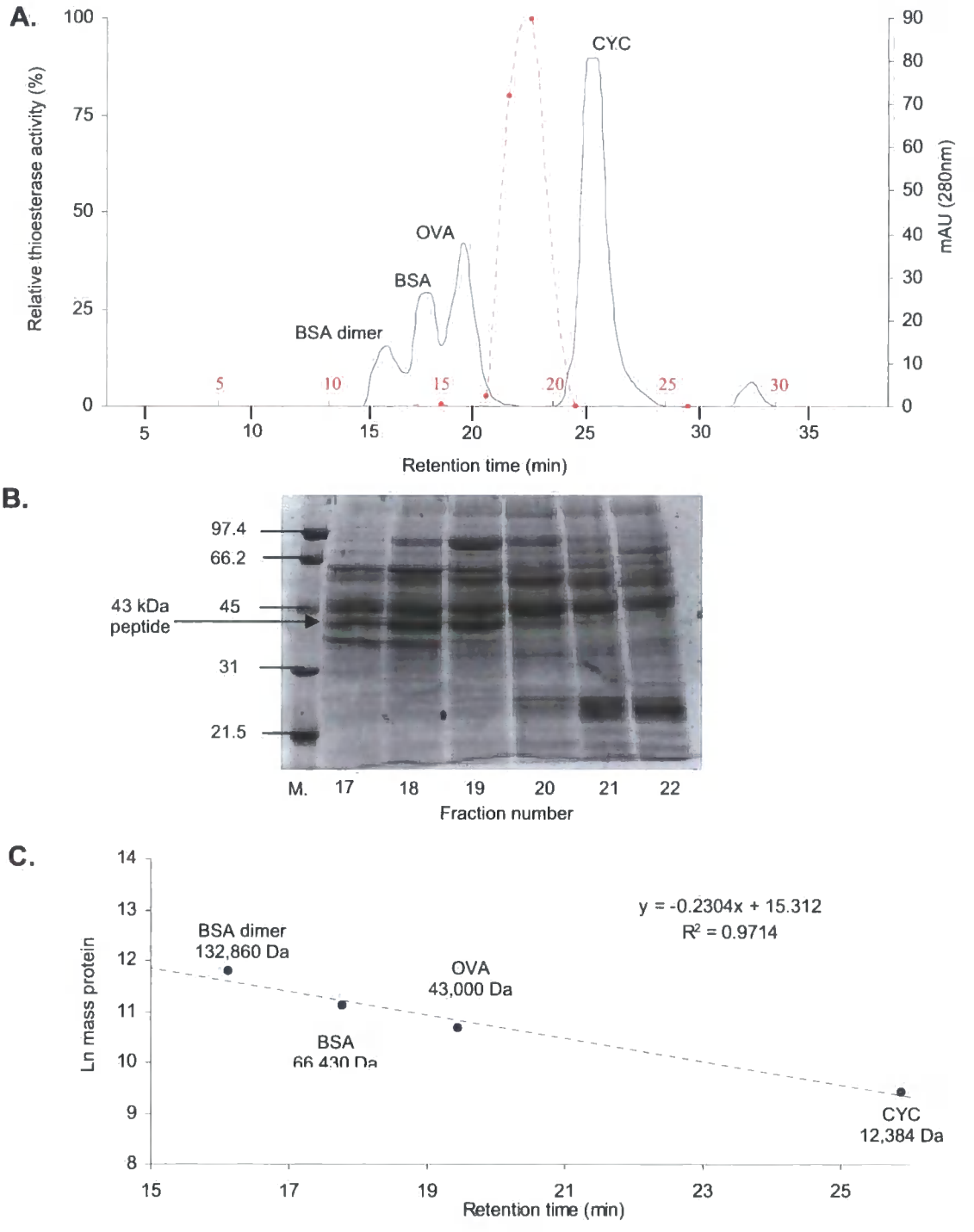


Figure 4.5 Characterising the size of *Pe*WHb by gel filtration. **A.** Gel filtration chromatography trace showing the separation of standards (200 μg); bovine serum albumin (BSA), ovalbumin (OVA) and cytochrome C (CYC) and the elution of thioesterase activity. The elution of activity was in fractions 17-21 (retention time: 20.5-24.5), with maximum activity in fraction 19. **B.** A 12% SDS-gel showing the polypeptides across the active peak, note the peptide at 43 kDa (arrow), the abundance of which mirrors elution of activity. **C.** A standard curve showing the linear relationship between Ln protein size (Da) and retention time on the size exclusion column. Thioesterase activity corresponds to a protein approximately 20-32 kDa in size, a monomer.

However, the exact molecular mass was difficult to ascertain as the protein eluted later than expected giving it a lower mass than that observed on SDS-gels. This anomaly was likely to be a result of partial adsorption of the protein to the solid matrix through ionic or hydrophobic interactions. Hence, this study indicated the active thioesterase was as a monomer with SDS-PAGE analysis of the active peak eluting from the gel filtration column confirming its mass as a 43 kDa protein (Figure 4.5B).

The isoelectric point of the enzyme was determined using two-dimensional gel electrophoresis whereby proteins were resolved by isoelectric focussing over a pH gradient (pH: 3-10) in the first dimension and subsequently resolved by size in the second dimension. Purified *Pe*WHb eluting from the final phenyl superose column was concentrated 10-fold (SpeedVac) to ensure sufficient loading of protein in the 100 μl (100 μg) sample applied onto an Immobiline dry strip (Amersham) for electrophoresis. The gel strip was inserted onto an agarose plate gel (12%) containing SDS and run in the second dimension (Figure 4.6).

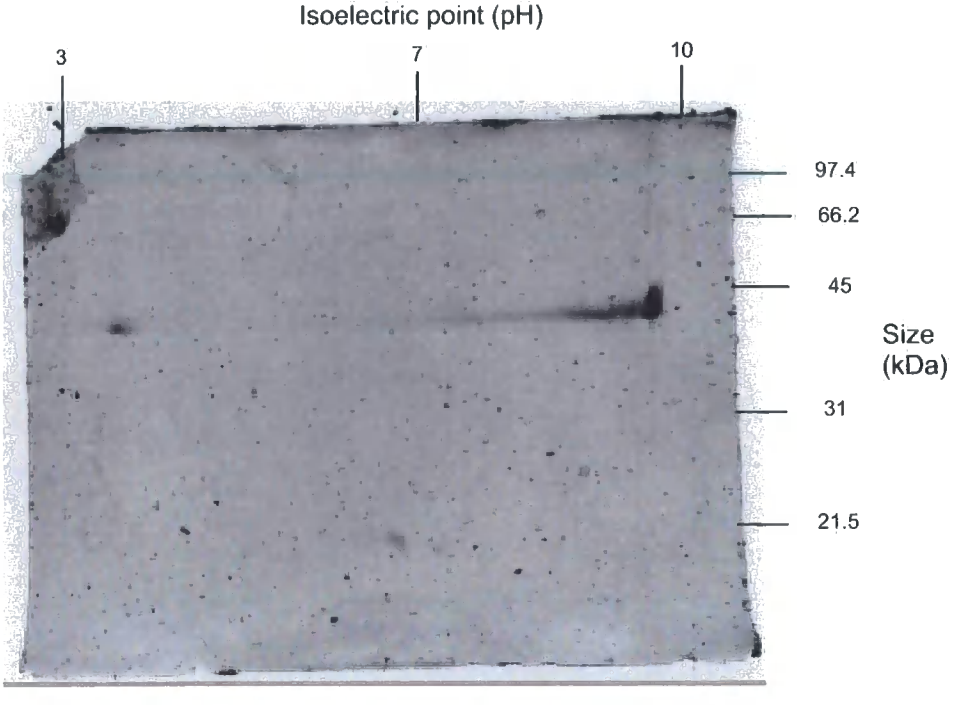
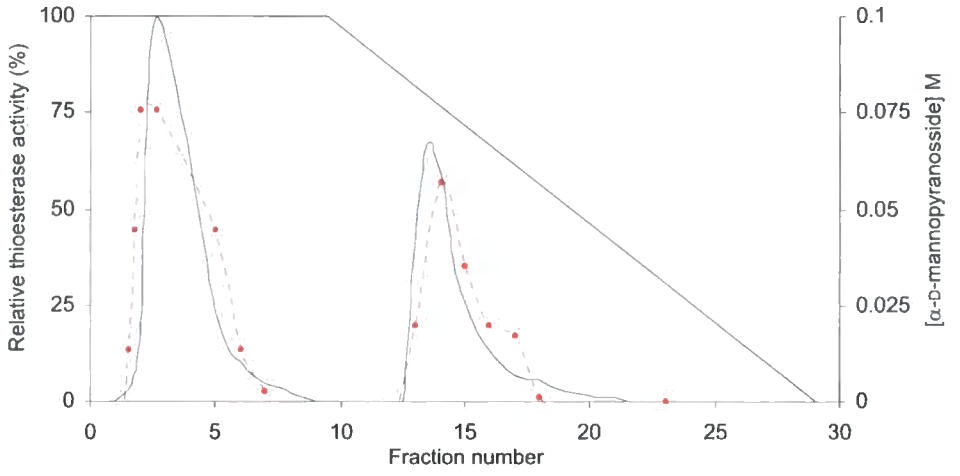


Figure 4.6 Characterising the isoelectric point of *Pe*WHb by resolution on a 2D gel. Polypeptides eluting from the final phenyl superose column (pool 2) were resolved by isoelectric point in the first dimension (horizontal) and then by size in the second dimension (vertical).

Although partially purified *Pe*WHb was the slightly more acidic of the two pools of activity, it was still an unusually basic protein with a pI (isoelectric point): 9.5-10. *Pe*WHb was resolved as multiple spots over this pH and in order to determine whether this was due to different forms of the protein or unequal resolution on the gel each spot was excised and analysed using MALDI-ToF mass spectrometry (data not shown). Each peptide fragmented in a similar pattern and it was therefore confirmed each spot was the same protein, *Pe*WHb. This is an occasional problem noted previously in isoelectric focussing[154], whereby proteins associated with ampholytes in the gel acquire isoelectric points which differ from that of the free protein.

*Pe*WHb was subsequently analysed for post-translational modifications, namely glycosylation using a combination of gel staining and concanavalin A chromatography. Purified *Pe*WHb appeared to be partially glycosylated based on its incomplete adsorption to concanavalin A sepharose (Con-A). Only 43% of activity was selectively retained by Con-A and eluted with 0.075 M methyl-α-D-mannopyronoside (Figure 4.7A). In order to determine whether the bound fraction was in fact glycosylated and to rule out that adsorption may be due to ionic or hydrophobic interactions with the matrix, the fractions eluting from the column were analysed with Krypton glycostain (PIERCE) (Figure 4.7B). The stain reacts with periodate oxidised carbohydrate groups on glycoproteins, creating a fluorescent dye. Reactivity to the stain suggests the partially bound fraction was glycosylated. Hence, a fraction of this activity has undergone glycosylation, perhaps as a means of regulating the properties and localisation of the protein between soluble and cell-wall fractions.

A.



Fraction ID	Vol (ml)	Total protein (mg)	S.a. (nkat mg^{-1})	T.a. (nkat)	Fold purification	Recovery
Phenyl superose	1	0.025	149.38	3.73	-	100%
Con-A 8-10	3	0.011	145.45	1.60	0.97	43%

B. Sypro stain:

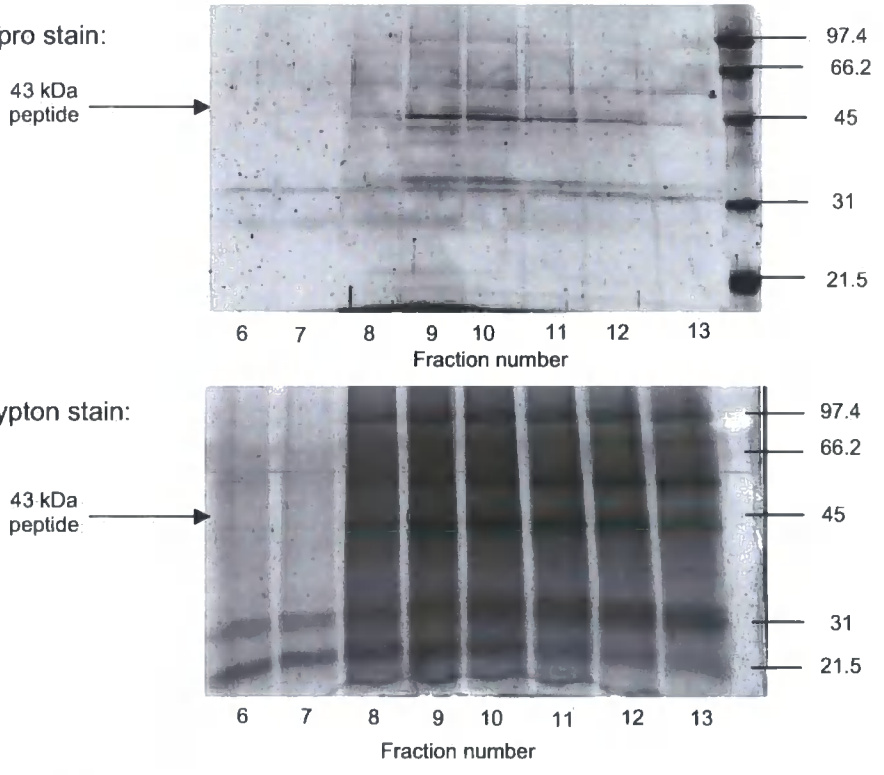


Figure 4.7 Examining glycosylation. **A.** The binding of thioesterase activity to Conconavalin A sepharose, only partial binding was observed (43%). **B.** The fluorescent detection of glycoproteins in fractions eluting from the Con-A column using Krypton stain (PIERCE), minimal detection limit 15 ng of protein per band. A comparison of total protein stained with Sypro ruby (top) and with Krypton stain (bottom). Glycoproteins appear as darker bands, the 43 kDa peptide of interest is brighter in respect to other polypeptides following Krypton staining.

The sensitivity of the enriched thioesterase (phenyl superose fraction) to inhibition by the different classes of hydrolase inhibitors was then re-examined (Table 4.2). As determined with the crude enzyme, the partially purified preparation was selectively inhibited by the organophosphate paraoxon (100% inhibition). In addition, the purified enzyme was more susceptible to inhibition with the serine hydrolase inhibitor fp-biotin (86% inhibition) as compared to the crude protein preparation. The suggestion that the protein was a serine hydrolase was confirmed by labelling *Pe*WHb with a trifunctional fluorophosphonate activity probe (TriFPP). The custom-prepared TriFPP, bearing a reactive fluorophosphonate to covalently modify reactive active site serines, and biotin and rhodamine components for recovery and recognition respectively (figure 4.8A), had been previously shown to label and inactivate a serine hydrolases in the model plant *Arabidopsis thaliana*[125]. Following a 60 minute incubation with 10 μM probe, 86% of the *Pe*WH's thioesterase activity was lost, with the concomitant fluorescent labelling of the 43 kDa polypeptide (Figure 4.8B).

The enzyme was then screened for activity against the panel of thioester and carboxylester substrates in comparison with the commercial hydrolases (Table 4.3). The greatest activity of purified *Pe*WHb was toward MFTA (**2**); 149.38 nkat mg protein, followed by the carboxyl ester *p*-nitrophenylacetate (**8**) (11.39 nkat mg^{-1}). This is somewhat different to PLE which had greatest activity toward carboxyl ester substrates. Interestingly, the position of the thioester functionality on the cyclic ring appears important: activity toward **2** where the thioester moiety is at position 3 is some 15 times greater than the activity toward **5** where the thioester is at position 2. In addition, the hydrolysis of the thioester bond in **1** occurs at a greater rate (9x) than that of hydrolysis of the carboxylester bond in the same compound (as determined using the GC assay). This could be a consequence of the higher energy C-S bond compared to the C-O bond, however it is equally as likely to be a consequence of the enzymes preference for thioester substituents, an indication that the enzyme shows chemoselectivity.

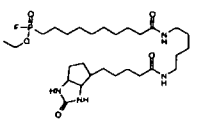
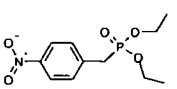
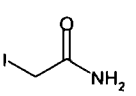
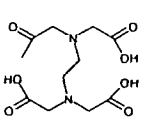
Inhibitor (hydrolase class inhibited)	Percentage Inhibition (%)	
	Crude protein	Partially purified protein (*Pe*WHb)
 10μM Fp-biotin (serine)	38 +/- 1	86 +/- 3
 1mM Paraoxon (serine)	87 +/- 3	100 +/- 1
 1mM Iodoacetamide (cysteine)	ND	ND
$HgCl_2$ (cysteine) 1mM Mercuric chloride (cysteine)	25 +/- 1	ND
 1mM EDTA (metallo hydrolase)	ND	ND

Table 4.2 The inhibition of thioesterase activity. Structures of inhibitors and the hydrolase class towards which they are active. Inhibition of *Pe*WHb is shown in comparison to the crude extract from whole fruit. Percentage inhibition values were calculated from the loss of activity after treatment with the inhibitor for 2 hours at 30°C and stated as the average value from triplicates +/- the standard deviation from the mean. ND; no inhibition detected. Control studies using acetone (- inhibitor) were found to have no effect on enzyme activity.

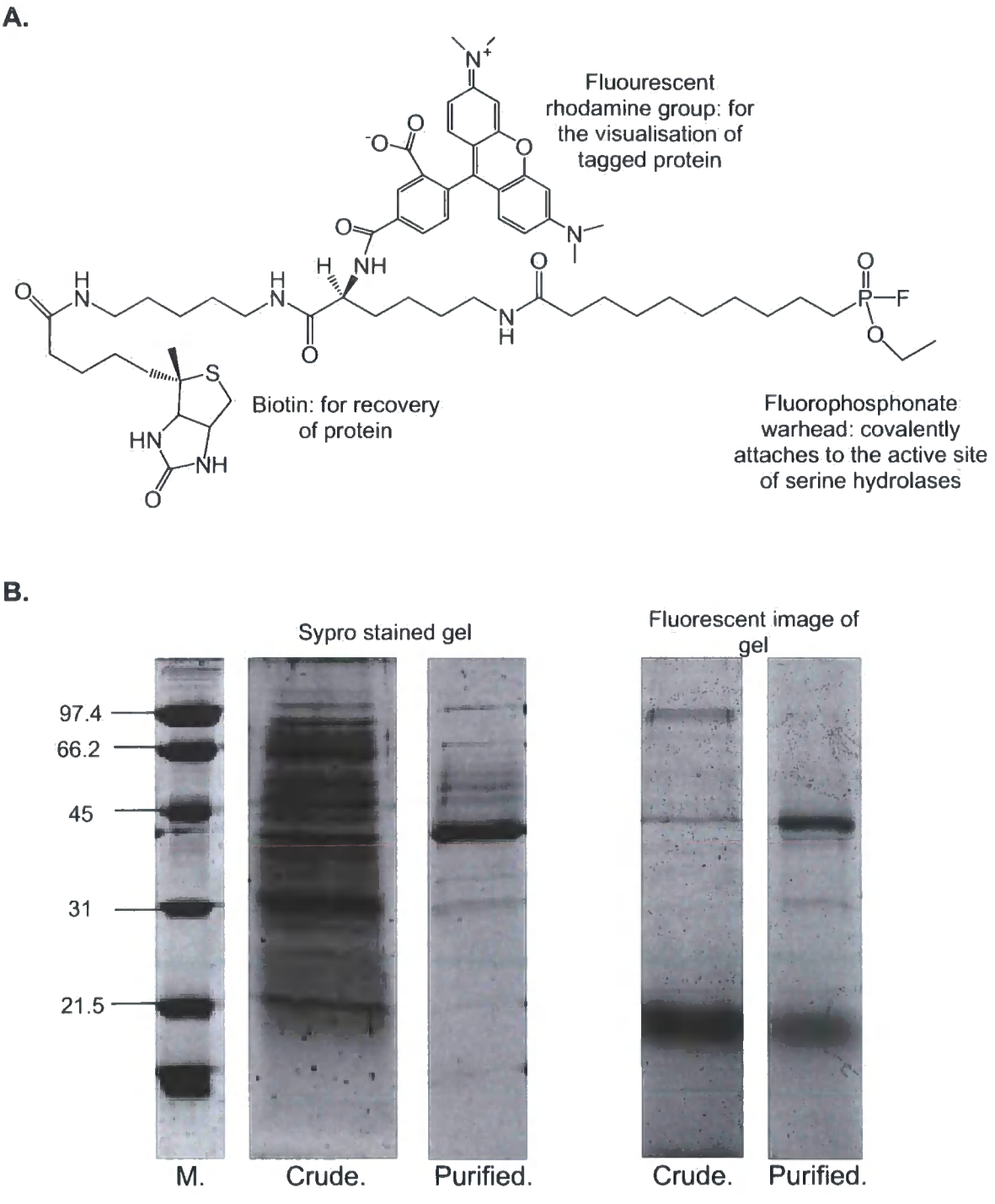


Figure 4.8 A. The structure and function of the biotinylated fluorophosphonate suicide probe. **B.** Labelling with a fluorophosphonate suicide probe. Gel images of crude passion fruit extract and purified PeWHb (Phenyl superose preparation) showing total protein (Sypro stain, left) compared to labelled protein as seen under fluorescent light (right). The 43 kDa polypeptide is labelled with the probe confirming it is a serine hydrolase. M: markers kDa. The diffuse bands in the fluorescent images at 21 kDa are likely to be due to unbound probe. Note the larger peptide bands (97 kDa) highlighted in the labelled crude sample indicating the presence of a further set of passion fruit hydrolases.

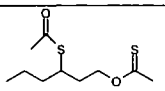
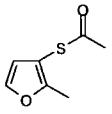
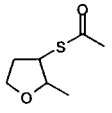
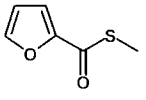
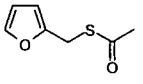
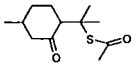
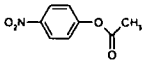
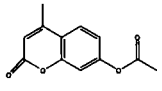
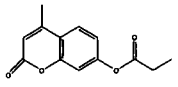
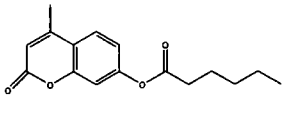
Compound number	Thioester/ ester substrates	Crude passion fruit extract (S.a. nkat/mg)	Pure passion fruit thioesterase (S.a. nkat/mg)	Porcine liver esterase (S.a. nkat/mg)	*Candida rugosa* lipase (S.a. nkat/mg)
1	 3-(thioacetyl) hexylacetate	0.04 +/- 0.002	1.12 +/- 0.02	14.02 +/- 0.72	ND
2	 2-methyl furan-3-thioacetate	0.97 +/- 0.003	149.38 +/-0.0021	537.88 +/- 24.65	0.35 +/- 0.02
3	 Tetra hydro 2-methyl furan-3-thioacetate	0.14 +/- 0.01	0.54 +/- 0.05	58.71 +/- 5.68	ND
4	 Methyl thiofuroate	ND	ND	100.0 +/- 9.75	ND
5	 Furfuryl thioacetate	0.24 +/- 0.04	10.06 +/-	35.04 +/- 1.98	0.05 +/- 0.01
6	 8-(acetylthio) menthone	0.13 +/- 0.02	8.84 +/- 0.08	4.17 +/- 0.02	0.02 +/- 0.01
7	 α-Naphthyl acetate	0.26 +/- 0.9	4.16 +/- 0.72	159.4 +/- 16.14	0.58 +/- 0.01
8	 *p*-Nitrophenyl acetate	0.67 +/- 0.10	11.39 +/- 1.2	2117.6 +/- 87.5	5.14 +/- 0.23
9	 4-Methylumbelliferyl acetate	0.04 +/- 0.003	0.66 +/- 0.08	426.15 +/- 72.55	4.26 +/- 0.18
10	 4-Methylumbelliferyl propanoate	0.03 +/- 0.008	0.57 +/- 0.09	1589.26 +/- 102	118.36 +/- 21.39
11	 4-Methylumbelliferyl hexanoate	0.0005 +/- 0.0001	0.006 +/- 0.0004	1302.97 +/- 97	36.59 +/- 4.00

Table 4.3 Enzymatic hydrolysis of thioester and ester substrates by passion fruit (crude protein and purified enzyme), mammalian esterase (PLE) and a microbial lipase. Activities are means of triplicates +/- standard deviation from the mean. ND; no activity detected.

PLE demonstrated broader substrate specificities than the *C.rugosa* lipase which was more selective for longer chain, bulky substituents, an observation noted previously[103]. It is likely the lipase would show greater specific activities if substrate concentrations were increased above their solubility limits. This may account for the relatively low activities observed. In contrast, the thioesterase from passion fruit, *Pe*WHb predominantly favoured short chain acyl components on the substrate, optimally acetate, as demonstrated by activities toward the increasing chain length of methylumbelliferyl esters (**9**, **10**, **11**) (Table 4.3). Interestingly many natural odorant volatiles are found to be acetylated. It will however accept bulky substituents as the alcohol component of the compound. In addition *Pe*WHb appears to show a preference for cyclic constituents especially those containing oxygen. *Pe*WHb is likely to show improved selectivities toward these short-chain volatiles in a regio- and chemo-specific manner.

4.4 Identification of Purified Thioesterase

It was important to establish the identity of *Pe*WHa and *Pe*WHb. However, this was likely to be challenging as the enzymes were purified from a non-model plant source which had attracted little scientific attention to date. The protein data for passion fruit on genomic and EST (expressed sequence tag) databases was therefore likely to be limited, meaning the identification of *Pe*WHa+b would be based upon homology to proteins from other plant species. However, with sufficient protein sequence data generated it was hoped the identity of each enzyme could be determined. In addition, being able to associate this activity with homologous proteins may shed further light onto the mechanism of action of the enzyme, and the associated scientific literature may indicate further endogenous roles.

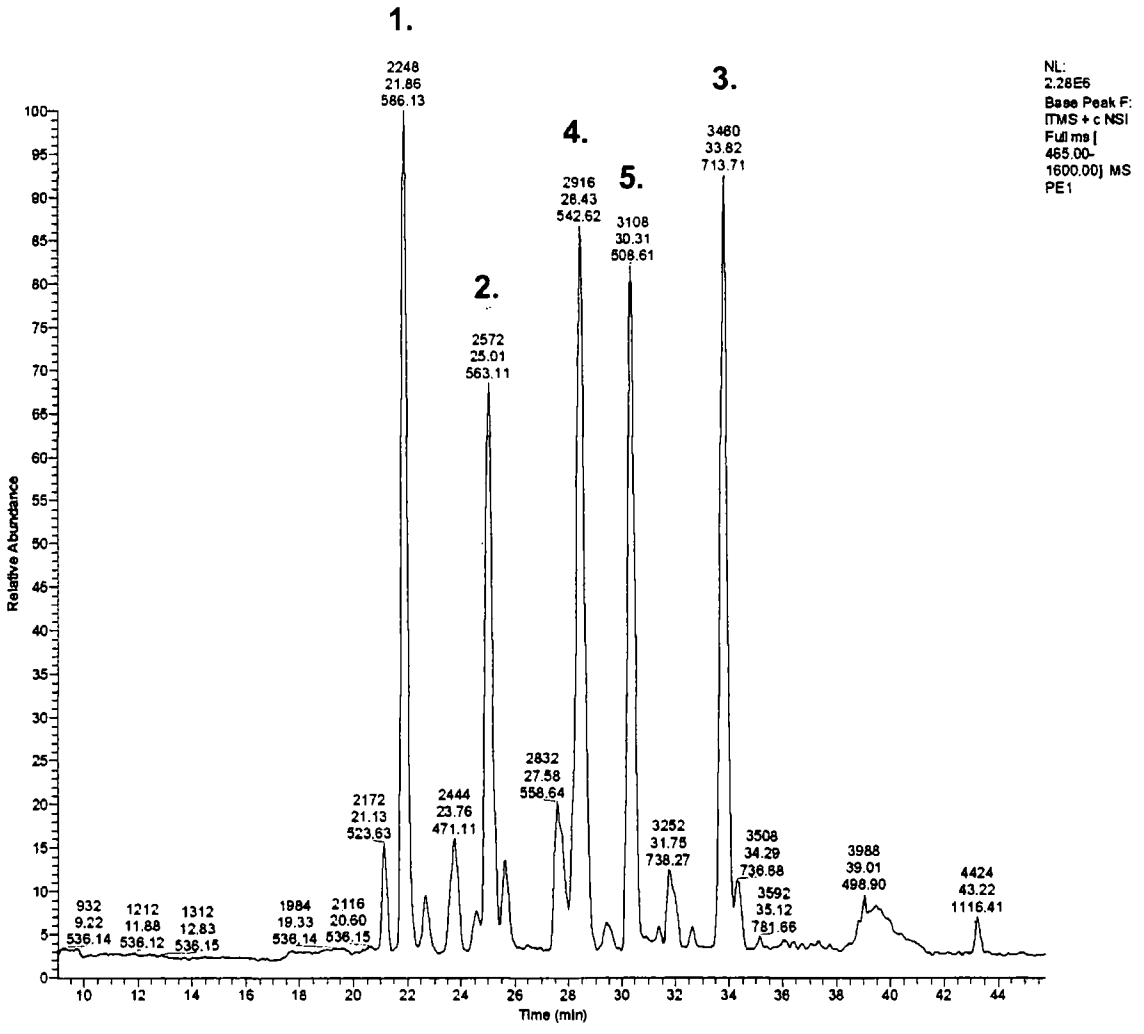
Mass spectrometry is an outstanding analytic tool which is used to identify compounds with remarkable levels of sensitivity[155]. Initially such instruments were limited to the analysis of small compounds due to the requirement of generating gas-phase ions from a sample. However, over the past twenty

years instruments have also been optimised for analysis of larger biomolecules such as proteins[155]. Through improved ionisation techniques such as MALDI (matrix assisted laser desorption ionisation) and novel mass analysers, used for separating ions, such as time of flight (ToF) spectrometers the problems associated with analysing high-mass multiply charged ions has been overcome. Such an instrument offers the only truly reliable determination of the molecular weight of a protein. In addition, through tandem mass spectrometry whereby parent ions are fragmented further and analysed on a second mass spectrometer the structure and sequence of a protein can be determined. Such techniques were therefore used to determine the identity of *Pe*WHa+b.

As described the sample concentration is vital for successful analysis. Micro litres of sample are required for MALDI and only part of this is consumed in analysis. Through visualising *Pe*WHa and *Pe*WHb on Coomassie stained gels it was apparent enough protein (10 ng+) was present. Peptides were extracted from the gel slices and digested with trypsin (30 μg/ml) prior to separation using liquid chromatography and analysis through mass spectrometry (LC-MS). LC-MS using a MALDI-ToF mass spectrometer (Durham proteomics) was used to analyse *Pe*WHa+b. The MALDI ion traces for *Pe*WHa+b (Annex I: *Pe*WH Protein Analysis) illustrate similar fragmentation of the original peptides, indicating they are in fact the same protein. Their different properties may therefore be a result of subtle post-translational modifications (e.g. glycosylations).

*Pe*WHb was also analysed further through tandem mass spectrometry (MS-MS) sequencing on a ToF-ToF mass spectrometer (Thermo). Such an instrument has been previously used for successful tandem MS based sequencing of unknown proteins from non-model plants[143]. Initial ionisation through MALDI produced a spectrum with five major ion peaks (Figure 4.9). Hence, due to the nature of the protein only these five distinct peptides could be further fragmented to provide amino acid sequence, somewhat limiting the amount of amino acid coverage generated.

Figure 4.9 MALDI spectra for *Pe*WHb on the Thermo ToF-ToF tandem mass spectrometer. The five major ion peaks were subjected to further fragmentation and used to generate partial protein sequence.



However, it was hoped these five ions would be sufficient and they were therefore fragmented through collision induced dissociation (CID) and analysed on a further ToF mass spectrometer (Annex I B). Determining the mass difference between consecutive daughter ions allowed the determination of amino acid sequence for all five major ions (Table 4.4).

These sequences were then used to interrogate the plant genome and EST databases using short match BLAST (basic local alignment search tool, www.ncbi.nlm.nih.gov/blast/Blast.cgi). From the five major hits it was apparent the enzyme was homologous to a carbohydrate binding hydrolase. Three families of enzymes were matched as top hits; polygalacturonase (EC 3.2.1.15), β-galactosidase (EC 3.2.1.23) and pectin acetylesterase (EC 3.1.1.6). Of these proteins identified, pectin acetylesterase was the primary candidate, with three of the five fragments generating significant hits toward this family of enzymes. The hits were considerably more significant than those toward the other enzymes, with higher score values and lower E (Expect). In addition, the properties of *Pe*WH did not match well with those of the other enzyme families. For example, through comparing β-galactosidase enzymes isolated from plants using the enzyme database BRENDA (http://www.brenda-enzymes.info/index.php4) it was apparent their molecular mass was generally larger (>70 kDa), often due to multiple domains, and their isoelectric points were lower as determined in mung bean (*Vigna radiata*) (pI: 7-8)[156]. A comparable search on BRENDA for polygalacturonases showed that those isolated and characterized from plants ranged considerably in size (30-330 kDa) but generally had isoelectric points between pH 7-8, with the greatest from Lily (*Lilium longiflorum*), pI: 8.1[157]. Hence, the high isoelectric point of *Pe*WH (9.5-10) could not be accounted for by these enzymes and our focus concentrated on the more significant match of pectin acetylesterases (PAEs). *Pe*WH (43 kDa) had a comparable molecular mass to other PAEs (42-45 kDa) and a similar isoelectric point (pH 8-10)[158]. In addition, the occurrence of *pe*WH as a cell wall bound protein, with a preference for acetylated substrates, being basic in nature and with supporting protein MS-MS data all indicated the purified thioesterase, *Pe*WH, was a PAE.

Peptide fragment	Database hit (E value / Score)	Identification	Predicted Mr/ Isoelectric point (kDa)/ pI	Accession
1. VG(I/L)GNSPSR (9 amino acids)	GNSPSR GNSPSR GNSPSR *Lactuca sativa* (21.0/ 692)	Pectin acetylesterase, putative (member of pectin acetylesterase family)	43.7 8.76	AAP72959
2. N(I/L)(I/L)APSAVD (11 amino acids)	NILAPSAVD N+LAP+AVD NVLAPTAVD *Arabidopsis thaliana* (25.2/ 37)	Pectin acetylesterase, putative (member of pectin acetylesterase family)	42.1 9.31	AAK96722
3. VP(I/L)T(I/L)VQSAVA (11 amino acids)	VPLTLVQSA VPLTL+Q+A VPLTLIQAA *Arabidopsis thaliana* (25.2/ 41)	Pectin acetylesterase, putative (member of pectin acetylesterase family)	45.9 6.9	NP182216
4. TFGFAW(I/L)GGK (10 amino acids)	GFAWLGG GFAWL G GFAWLNG *Trifolium pratense* (23.1/ 159)	β-galactosidase, putative	47.3 8.04	BAE71266
5. AV(I/L)DD(I/L)(I/L)FK (9 amino acids)	VIDDLLFK VIDD LFK VIDDNLFK *Solanum lycopersicum* (23.1/ 159)	Polygalacturonase (pectinase)	50.1 6.4	P05117

Table 4.4 Sequencing of the purified thioesterase. Database hits obtained from MS-MS sequencing of polypeptides. For each peptide the top BLAST hits are presented for known or putative proteins. Both score and expect values are shown to highlight the significance of the match. Score: Higher the better, measure of how close the match is, taking into account miss matches (space) or similar amino acid types (+). Expect: Lower the better, the chance such an alignment could occur by chance in the database being searched.

4.5 Discussion

Through the purification and characterisation of thioesterase activity from the peel of ripened purple passion fruit it has been demonstrated that a protein characteristic of a pectin acetylesterase (PAE) is also able to selectively hydrolyse acetylated thioesters to release VOSCs.

The observed activity was found to be a consequence of two pools of acidic and basic enzyme fractions which were independently purified and shown to be the same protein through MALDI fragmentation. The differences in their properties and activities was therefore a likely consequence of post-translational modifications, possibly through glycosylation which may alter the surface properties of the enzyme thus increasing or decreasing its affinity for the cell-wall matrix. Interestingly this observation has been noted before in Valencia orange (*Citrus* sinensis), whereby hydrolase activity was separated into two pools of activity on an anion exchange column (Mono S), one pool which bound to the matrix and the other which was unbound and eluted in the void volume[150]. This is the same observation noted in our studies and often the relationship between the size of the bound and unbound pools varied dependent upon the extract, as described in *Citrus*[150]. The regulation of polygalacturonase (PG) activity in tomato has previously been linked to post translational modification. It was proposed that the binding of the enzyme to a further heavily glycosylated peptide, the PG β-subunit, altered the enzyme's characteristics through localising it to particular regions of the cell wall[146]. *Pe*WH was shown to be a monomeric enzyme, and it therefore appears likely that glycosylation influences localisation of the protein in the cell wall. The active enzyme may only become glycosylated and recruited to the cell wall late in ripening, when it is required for softening the cell-wall matrix, or in the release of volatiles from precursors as we postulate here.

Pectin acetylesterases (PAEs), despite their importance in plant cell wall metabolism, have not been well studied at the biochemical level. However, what is known about these enzymes does support the identification of *Pe*WH as a PAE. Firstly, as determined in *Citrus* species including grapefruit, orange

and lemon PAE activity like *Pe*WH was found predominantly in the peel, but not the flesh of the fruit[159]. PAEs have only been purified from mung bean and orange, with the extraction conditions mirroring those required to isolate *Pe*WH. For example, high salt concentrations in the extraction buffer (0.2-1 M NaCl) were required to solubilise the ionically bound cell wall proteins in mung bean[158] and orange[150]. Like *Pe*WH, PAEs are unusually basic proteins and are of similar molecular mass; the PAE from mung bean was a 43 kDa protein with an isoelectric point pH > 9 [158].

A 43 kDa cell wall bound PAE was first purified from mung bean (*Vigna radiata*) hypocotyls in 1994[158] and then sequenced and cloned[160]. Due to the enzyme's preference for hydrolysing acetylated pectin it was termed a "pectin acetylesterase". Similar enzyme activities had been previously identified in fruit peels, such as *Citrus* (grapefruit, orange and lemon)[150,159], however, the only enzyme cloned remains that from mung bean. Based on sequence similarity PAEs were grouped into a sub-family of carbohydrate esterases referred to as family 13 pectin acetylesterases[161,162]. The protein entries are catalogued on the carbohydrate-active enzyme (CAZy) database which is regularly updated and available at www.cazy.org/. This family contains 12 putative PAEs in *Arabidopsis thaliana* and 11 in rice (*Oryza sativa*). The relatively large number of different isoforms of PAEs in plants suggests the spectrum of functions they may serve. Those in tropical fruit may well have evolved through divergence and diversification to become selective for the varied chemical forms of flavour and fragrance metabolites. However, this hypothesis needs to be examined further.

The inhibition studies on the purified enzyme demonstrated that it is a classic serine hydrolase, being selectively inhibited by organophosphates which covalently modify this catalytic residue. *Citrus* PAE was also demonstrated to be inhibited by organophosphates[163] which is consistent with the evidence that *Pe*WH is a PAE. The presence of an active site serine was clearly illustrated using the fluorophosphonate trifunctional probe, which has been also shown to act as a suicide inhibitor of carboxylesterases involved in xenobiotic metabolism in black-grass[143] and Arabidopsis[125]. Interestingly, the trifunctional

probe also labelled and inactivated a bifunctional thioesterase/ carboxylesterase *S*-formylglutathione hydrolase[164]. *Pe*WH was also a bifunctional carboxylesterase and thioesterase, though it was unable to hydrolyse amide bonds. The trifunctional probe has been demonstrated to be an important tool for characterising hydrolases in these plants and it has great potential here to be utilised further in fruit for the specific pull-down of hydrolases. Such studies could be used to characterise the different forms of cell-wall hydrolases in peel at different developmental stages. This may form a useful directed screen for further hydrolases in fruit with potential use as biocatalysts in F&F synthesis.

An examination of plant PAE protein sequences, including the mung bean sequence and those with significant similarity with *Pe*WHb, highlight the amino acid residues making up the His-Asp-Ser catalytic triad of a serine hydrolase (Figure 4.10). In addition, as determined using the SignalP 3.0 server (ExPASy proteomics tools, www.cbs.dtu.dk/services/SignalP/), the PAE sequences were predicted to contain a signal peptide leader sequence which targets the protein to the cell wall (Figure 4.10). To date none of the plant PAEs have been crystallized and the presence of an oxyanion hole in the active site remains to be confirmed. However, a microbial PAE homologue was crystallized and found to be a classic α/β serine hydrolase enzyme with the oxyanion hole[165]. The most unusual feature of *Pe*WH was its selectivity for acetyl groups as the acyl components. This was suggestive of a constrained acyl-binding pocket, which is in complete contrast to the accommodation of large acyl chains seen with PLE and *Candida rugosa* lipase[103]. Relatively little is known about the structure of PAEs and class 13 hydrolases which can account for this selectivity for acetyl-esters. However, the specificity of PAEs toward short-chain acetyl groups as the acyl components has also been observed in *Citrus* species (grapefruit, orange and lemon), where increasing the acyl chain length to C4 butyrate esters reduced activity by around 50%[159]. Plant PAEs are regiospecific, hydrolysing ester bonds at C2 and C3 positions on galacturonic acid residues[166].

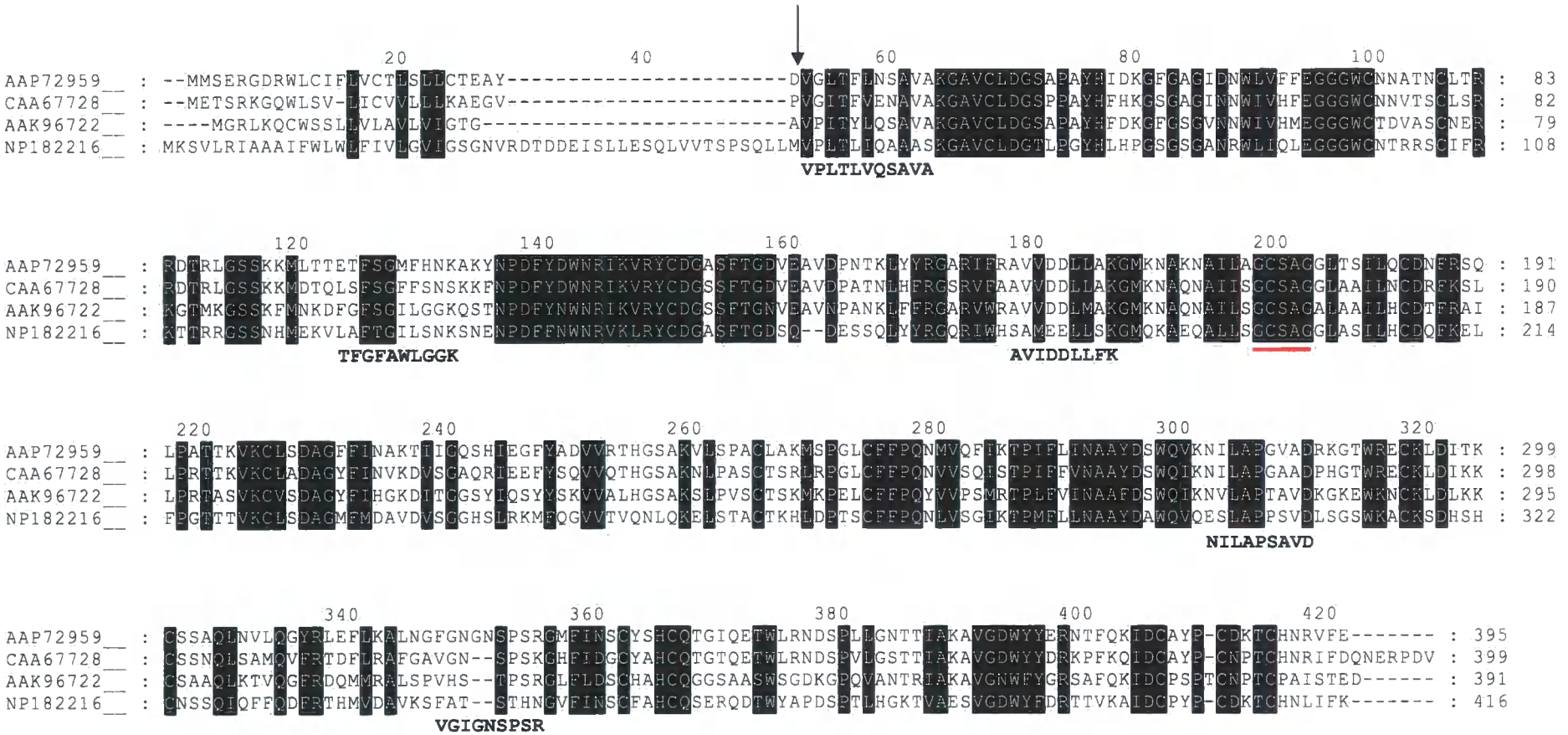


Figure 4.10 The alignment of selected plant PAEs. Sequences include the 3 major hits from BLAST searches; AAP72959 (*Lactuca sativa*), NP182216 and AAK96722 (*Arabidopsis thaliana*). For comparison the protein sequence of PAE from mung bean (*Vigna radiata*) is also aligned (accession: CAA67728). Black shaded sequence shows conserved amino acids (including amino acids of similar chemical type). The alignment highlights the serine hydrolase catalytic motif GXSXG (red), a predicted signal peptide cleavage site (arrow) and the alignment of PeWHb peptide fragments (underlying text).

It is possible the oxygen in the cyclic structure of substrates (and sugars) is important for the orientation of the compound in the active site of the enzyme. This may account for the activity observed toward VOSCs where the conformation of the tropical olfactophore (Figure 1.4) has oxygen in a specific orientation to the hydrolysed thioester bond.

The endogenous role of PAEs is somewhat unresolved, however, a number of suggestions have been put forward. The enzyme is important in regulating the degree of acetylation in pectin and hence has been proposed to be involved in cell wall modification, for example in softening fruit tissue[167], or in controlling cell growth and elongation in developing hypocotyls[160]. However, in *Citrus* PAEs are not only localised to the cell-wall material but are also found at high concentrations in juice vesicles, suggesting a possible further role[167].

We have therefore highlighted a possible novel function for PAEs in VOSC generation in tropical fruit. Work now needs to focus on cloning and expressing the enzyme both as a means of confirming identity and for characterizing the enzyme further. This will form the focus of the following chapter.

5. Cloning of Putative thioesterases from *Passiflora edulis* Sims and *Arabidopsis thaliana*

5.1 Introduction

Following from confirming the presence of thioesterase activity in passion fruit and the identification of the purified enzyme, *Pe*WH, as a pectin acetylesterase (PAE), work now focussed on the cloning and expression of the respective protein. Determination of the complementary DNA (cDNA) coding sequence for *Pe*WH and its subsequent expression would enable further characterization of the enzyme and confirm the protein's identity, which so far had been described through proteomics. Furthermore, the cloning and expression of the protein would form the basis of a system for obtaining the purified biocatalyst in abundance. An overview of the technical strategy for sourcing the enzyme through genetic engineering is depicted in figure 5.1. This process is dependent upon an expression construct containing the full length cDNA open reading frame for *Pe*WH and its transformation into a suitable expression system.

The demand for enzymes for use in bioprocessing has never been greater, as is reflected by the fact that in 1991 8% of all papers published in synthetic organic chemistry described the use of enzymatic synthesis[168]. Hence, the commercial availability of enzymes both in purified or crude form has increased dramatically with over thirty five major global suppliers competing today[169]. The majority of commercial biocatalysts are supplied in crude form (1-30% actual protein) where they are shipped by the kilo and often derived from microbial sources through fermentation[95]. Conversely, pure or partially purified enzyme preparations are produced on a smaller scale and sold by the unit, however, this is increasing as systems for recombinant expression of proteins and their subsequent isolation (e.g. affinity purification) are becoming easier to perform[95]. Obtaining purified enzymes through recombinant DNA technology is not novel and has been developed over the past 30 years due to the wealth of technological advances made in molecular biology and genetics.

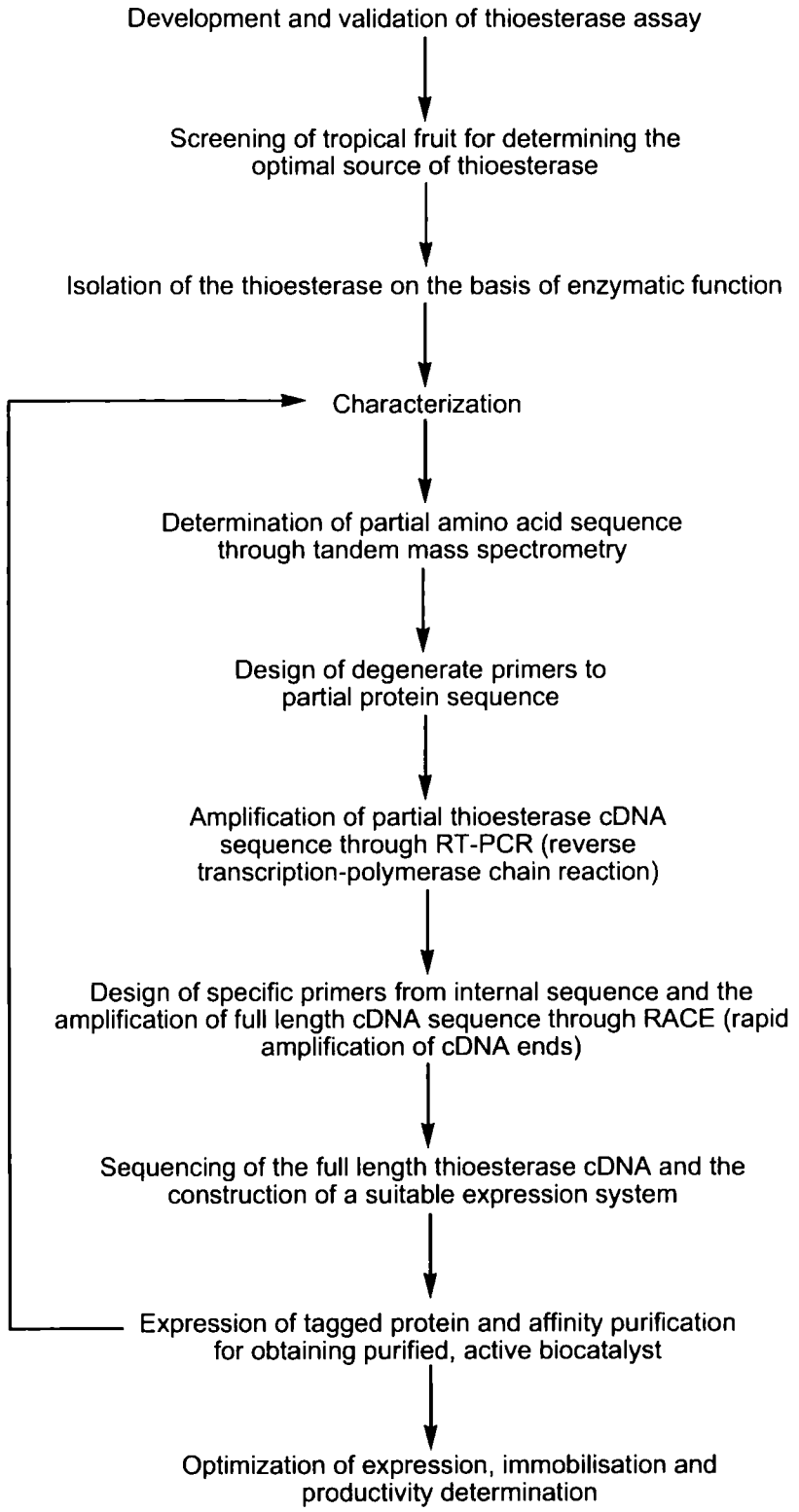


Figure 5.1 Sourcing thioesterases through recombinant DNA technology. An overview of the technical approach employed. A variation on the above would be to design oligonucleotide probes for screening a cDNA library prepared from the optimal fruit source.

Such practices are commonplace in today's biological laboratories where gene cloning and expression are powerful experimental tools, however, the manufacture of enzymes (biocatalysts) industrially through such an approach is relatively in its infancy. The purified enzyme should show improved selectivities and reduce unwanted side reactions as compared with chemical transformations. A system for immobilisation would be preferable, so as to improve the stability of the protein and aid in its recovery. The enzyme could be tagged and pulled out of the bacterial lysate mix prior to immobilisation, or simply used in its crude form. As fine chemical companies see the clear advantages of enzymatic manufacture, they are likely to change their systems of production and the classic chemist is now collaborating with biochemists, molecular biologists and geneticists to bring such technologies to industry.

There is a demand for novel hydrolases (carboxylesterases/ thioesterases) for use in flavour and fragrance manufacture[113,114], and the possibility of obtaining such a purified biocatalyst from fruit through recombinant expression is the focus of this chapter. To date only a single plant pectin acetylesterase gene has been cloned (accession number: X99348). This sequence was obtained from three day old mung bean (*Vigna radiata*) seedlings through an RT-PCR approach[160] using degenerate oligonucletide primers designed to N-terminal and internal amino acid sequence derived from the purified protein[158]. A 317 bp internal cDNA fragment was amplified and used to design specific primers for obtaining the missing 5' and 3' ends through RACE[160]. There is however no description of the expression of this clone. A similar approach will be employed here for the cloning of putative thioesterases from purple passion fruit. The challenge in working with a non-model plant is that degenerate primers must be designed from partial amino acid sequence from the purified protein or from regions of conservation identified in putative PAEs in other plants. In many cases it may not be possible to find the corresponding nucleotide cDNA sequence in the databases, as would have been available if working with model plants, such as *Arabidopsis* or rice. Searching the passion fruit genome for the *Pe*WH coding sequence will set a path for further researchers looking to clone and express putative genes of flavour biosynthesis from unusual plant sources in this collaborative project with Oxford Chemicals Ltd.

5.2 Partial PAE Thioesterase Sequence From Passion Fruit Using Degenerate Primers

The design of degenerate oligonucleotide primers to the amino acid sequences determined from *Pe*WH was considered (Table 5.1). Oligonucleotides of 20 residues in length designed to the regions of least degeneracy still contained too many possible combinations (768+) to be considered specific. This is ultimately a consequence of the high leucine/ isoleucine content of each peptide which are encoded by multiple codons, and greatly increase the degeneracy of the sequence. For comparison the degenerate primers (20 mers) used for the successful cloning of the PAE from mung bean both contained 512 combinations[160], and hence were specific.

Peptide fragment	Sequence Number of possible codons	Lowest number of combinations over 20 mer
1	VG(I/L)GNSPSR 44(9)426466	24576
2	N(I/L)(I/L)APSAVD 2(9)(9)446442	32768
3	VP(I/L)T(I/L)VQSAVA 44(9)4(9)446444	36864
4	TFGFAW(I/L)GGK 424241(9)442	768
5	AV(I/L)DD(I/L)(I/L)FK 44(9)22(9)(9)21	4608

Table 5.1 Designing degenerate primers toward partial amino acid sequences from *Pe*WH. Peptides **1-5** all possessed too high a level of degeneracy, shown by the number of different possible codons below the sequence. The lowest number of possible nucleotide combinations in a 20 mer primer is illustrated (right).

In an attempt to obtain greater amino acid sequence from *Pe*WH for subsequent primer design, a different trypsin digestion protocol, optimised toward low amounts of protein, was employed[170]. It was postulated that the hydrophobic nature of *Pe*WH may hinder its digestion by trypsin and subsequently reduce the number of ions of sufficient intensity required for MSMS sequencing. However, the nature of *Pe*WH meant no further protein sequence could be determined under the conditions tested. A further approach

would have been to determine the N-terminal amino acid sequence of the protein through Edman degradation, as was employed in mung bean[158]. However, the carbohydrate active enzyme database (www.cazy.org) contained 28 full length PAE protein sequences from which degenerate primers could be designed toward homologous non-redundant sections of coding sequence for the amplification of partial PAE sequence.

5.2.1 The Alignment of Plant PAEs, Conserved Regions and Degenerate Primer Design

An alignment of all known non-redundant plant PAE protein sequences (n =28) from the carbohydrate active enzyme database highlighted conserved regions of homology (Annex II: Protein and DNA Sequence Analysis). Such conserved amino acids are likely to be involved in maintaining the structural conformation of the protein (α-helices/ β-sheets) or are important for catalytic activity, such as the GXSXG catalytic motif and the acidic aspartic acid and the basic histidine residues which make up the catalytic triad. Designing primers to the corresponding coding sequence of these conserved regions would greatly increase the chance of specifically amplifying PAE thioeterases from the fruit. Sequence homology is however fragmented and for this reason it was postulated that sub-families of PAEs may exist, a phylogenetic tree was therefore constructed from all of the 28 full length proteins identified (Figure 5.2). Sequences clustered into three distinct clades of similarity, which when aligned were found to have far greater conservation of amino acid sequence (Annex II B). It therefore appears plants contain several isoforms of PAEs, at least three distinct sub-families, perhaps differing in their surface properties or catalytic preference for different metabolites, further evidence this family of wall bound hydrolases have diverged to serve different functions within plants.

To determine the sub-family (clade) to which the partially purified thioesterase *Pe*WH belonged, the five peptide sequences from the purified protein determined through tandem mass spectrometry (Table 5.1) were used to search for matches in each clade. Four of the five peptides (**1**,**2**,**3**,**5**) matched with significant homology to sequences from clade 1 (Figure 5.3), with only three of the five peptides matching in the two other sub families, and often with

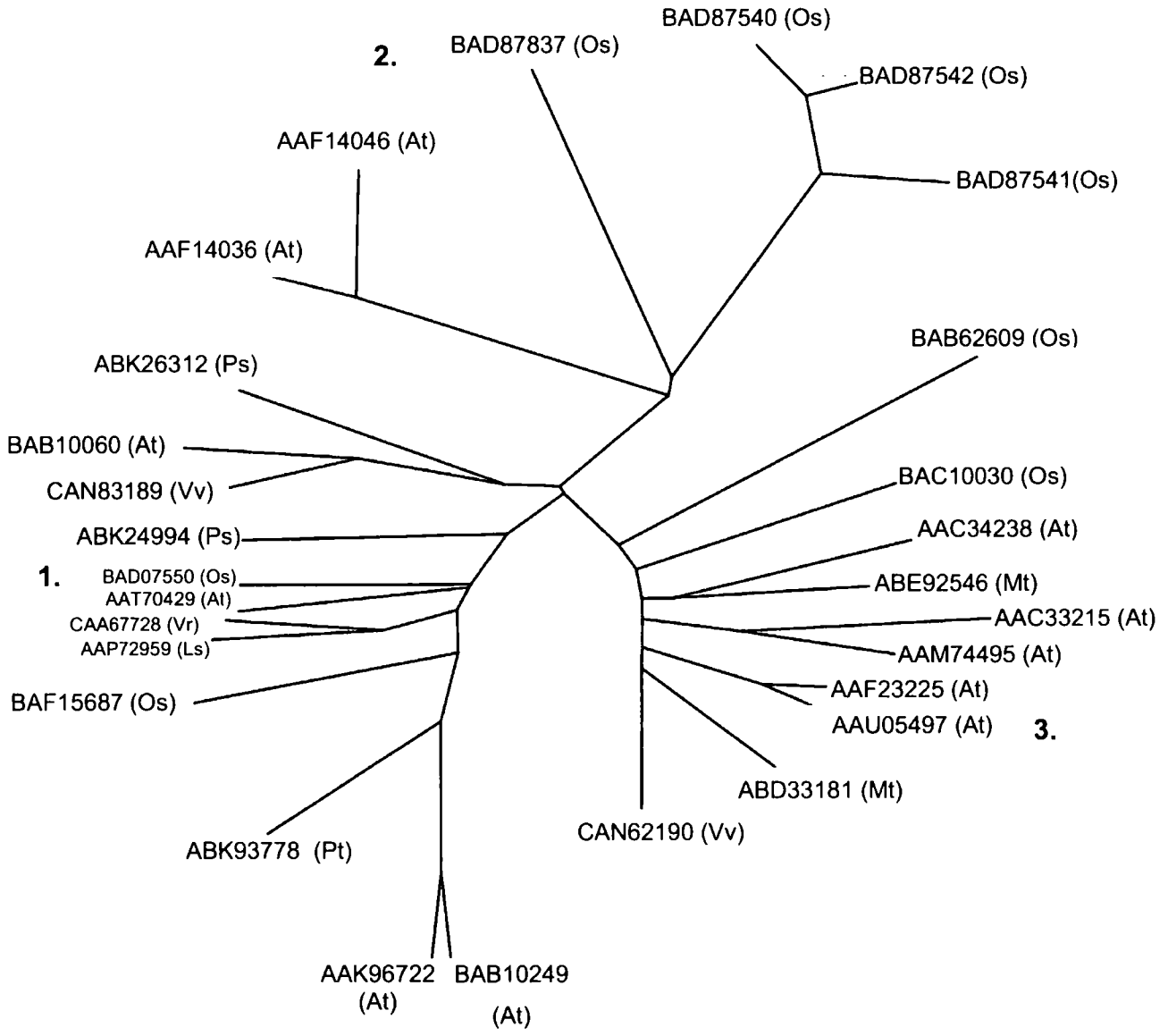


Figure 5.2 A phylogenetic tree of all known full length plant PAEs as catalogued on the carbohydrate active enzyme database (family 13). The tree was generated using the phylogeny package PHYLIP (Felsentein, 1989)[171]. Sequence homology matches cluster the 28 PAE protein sequences into three distinct sub-families (clades **1.-3.**). Entries are illustrated by accession number with the plant species type below; rice (*Oryza sativa*, *Os*), Arabidopsis (*Arabidopsis thaliana*, *At*), mung bean (*Vigna radiata*, *Vr*), Barrel medic (*Medicago trunculata*, *Mt*), Lettuce (*Lactuca sativa*, *Ls*), grape (*Vitis vinefera*, *Vv*), polplar (*Populus trichocarpa*, *Pt*) and spruce (*Picea sitchensis*, *Ps*).

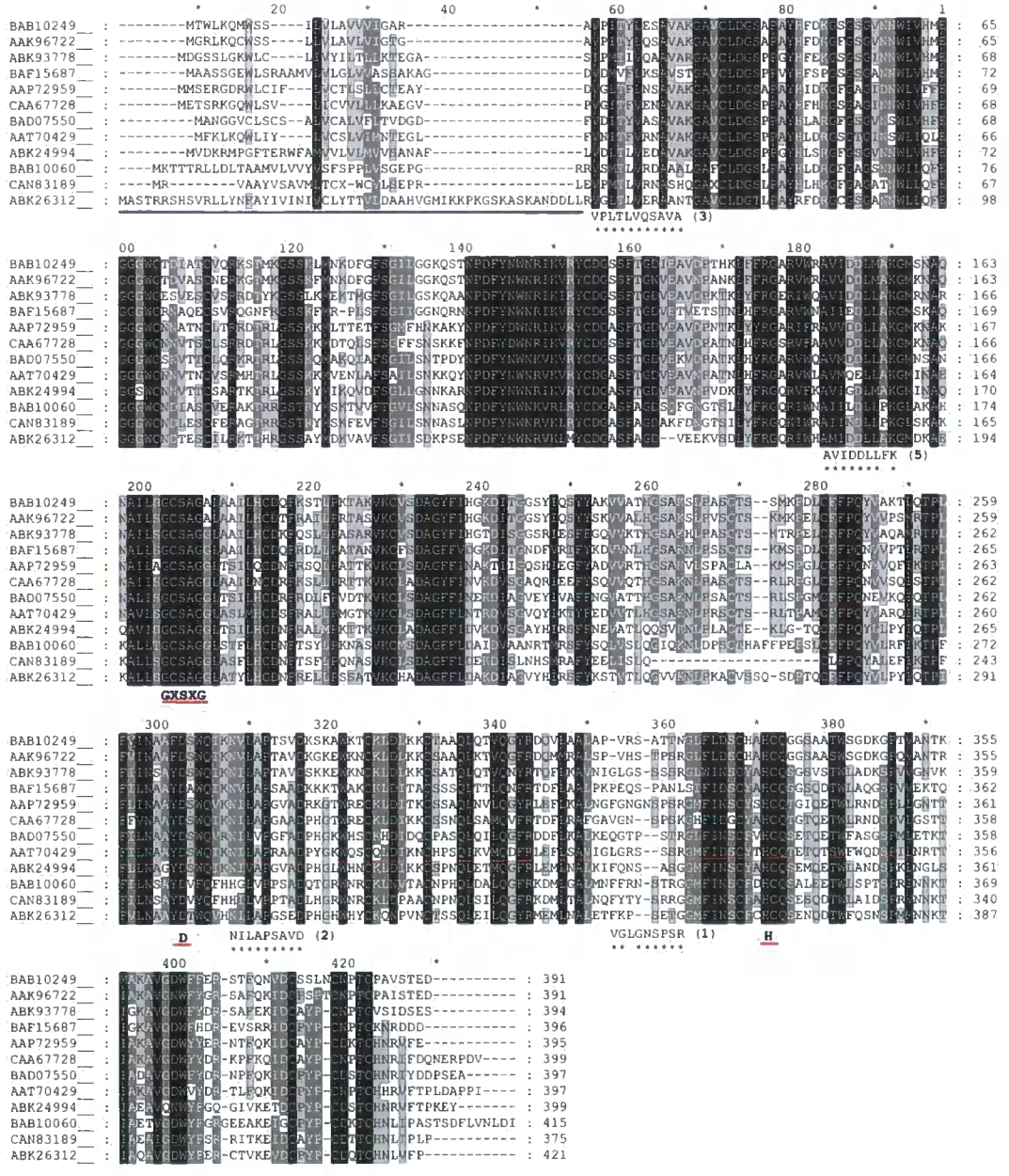


Figure 5.3 The alignment of clade 1 plant PAEs, demonstrating the greater sequence homology within this sub-family. Matches toward *Pe*WH peptide fragments (**1**,**2**,**3**,**5**) are shown. The catalytic motif GXSXG and the conserved aspartate and histidine residues which make up the catalytic triad are underlined (red). The predicted siganl peptide is underlined in blue. Each protein sequence is identified by its Genbank accession number. Black shading highlights complete sequence conservation, dark grey shading shows 80% or greater conservation and light grey shows 60% or greater conservation. All peptide matches are flanked by K or R amino acid trypsin cleavage sites, except **3**, which is the N terminus of the mature protein (- signal sequence).

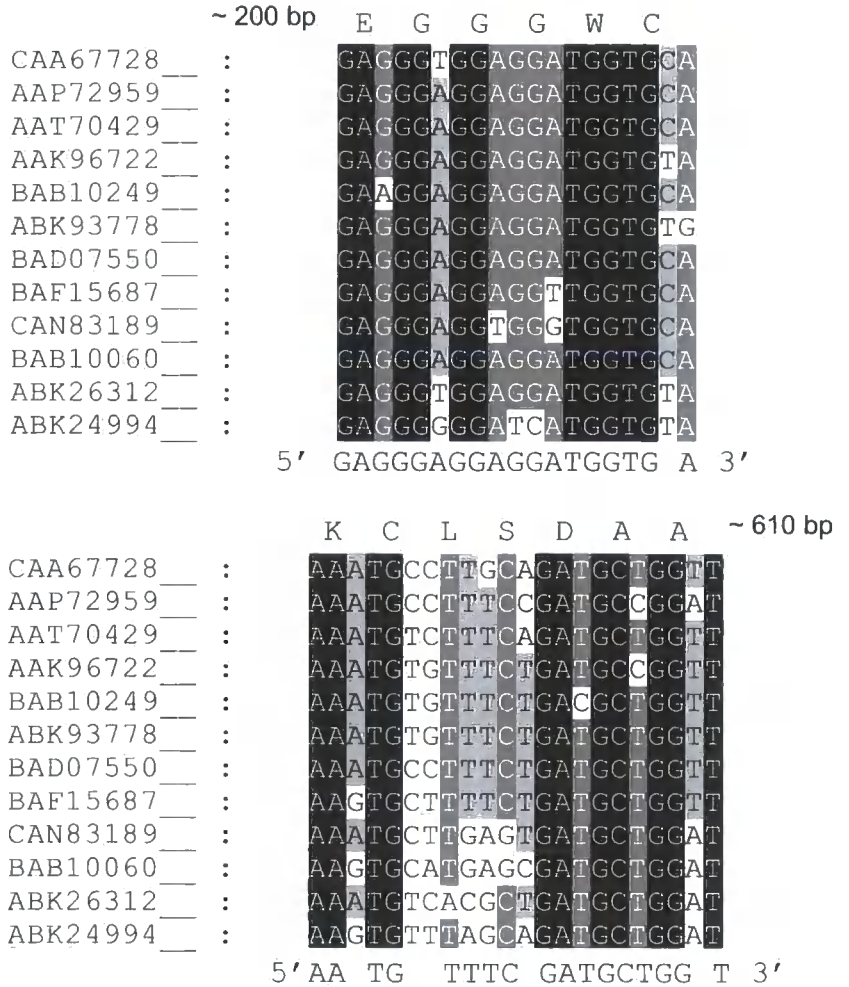
numerous mismatches over conserved regions. The unidentified peptide (**4**) may be from a non-conserved region of the protein and hence why it did not align with sequences in sub-family one. The greater sequence conservation within this family suggested that it would be possible to design primers to the corresponding coding sequences. All clade 1 full length cDNA sequences were therefore aligned (Annex II C) and the regions of least degeneracy (greatest homology) considered for the design of primers.

It was important to ensure the PCR primers were specific for PAE cDNAs, so that they would not amplify non-target sequences. For primers to anneal with sufficient specificity they had to meet a number of requirements:

- 18-24 nucleotides in length (ensuring specificity of binding to the correct sequence)
- 40-60% GC content (ensuring a high annealing temperature could be used for PCR to prevent binding to homologous non-target sequences)
- high GC content at the 5' end (ensuring strong binding where the DNA polymerase initiates amplification)
- minimal mismatches especially at the 3' end
- ensure primer pairs are non complementary

Several regions within the aligned sequences showed significant homology, however, only two regions matched the requirements above for the design of effective primers. Hence, a forward (*Pe*For) and reverse (*Pe*Rev, anti-sense) primer were designed at 200 and 610 base pairs from the 5' end respectively from these two optimal stretches of DNA (Figure 5.4A+B). PCR amplification resulting from the combination of *Pe*For and *Pe*Rev would theoretically amplify a 410 bp internal fragment, accounting for approximately 35 % coverage of the full sequence (Figure 5.4C).

A.



B.

*Pe*For (sense): 5' GAGGGAGGAGGATGGTGYA 3'
Length: 19 mer, GC content: 61%, Annealing temperature: 60-62°C

*Pe*Rev (antisense): 5'AWCCAGCATCWGAAAVRCAYTT3'
Length: 22 mer, GC content: 45%, Annealing temperature: 58-64°C

C.

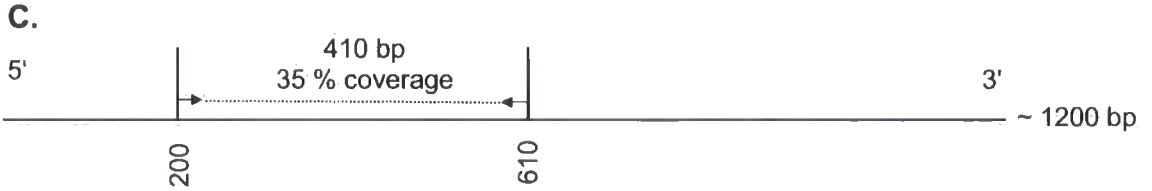


Figure 5.4 A. The alignment of cDNA PAE sequences from clade 1 over conserved amino acid regions selected for primer design. **B.** Forward and reverse primers, *Pe*For and *Pe*Rev used for the amplification of internal sequence. Y = C/T, W = A/T, R = A/G, V = A/C/G. **C.** A map of the predicted fragment amplified using *Pe*For and *Pe*Rev.

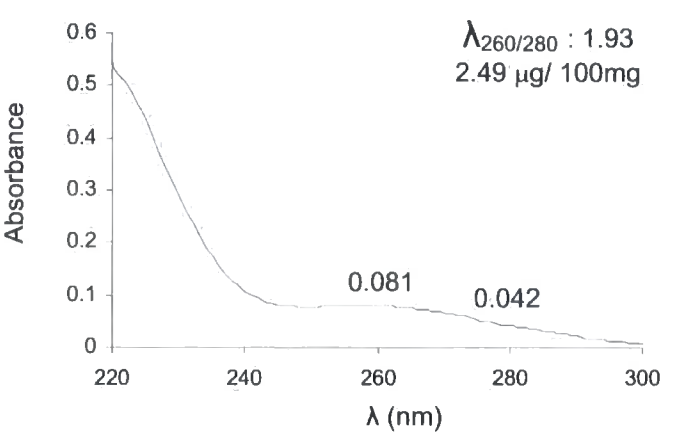
5.2.2 RNA Extraction From *Passiflora edulis* Sims and the Amplification of Partial PAE Sequence

RNA was extracted from both the mesoocarp (flavedo) and secretary gland tissues of purple passion fruit. Due to the high carbohydrate content of the fruit tissue, RNA extraction was however problematic and several different experimental approaches were considered, such as the chloroform extraction protocol developed for RNA extraction from mango[172]. It was far less problematic extracting RNA of good purity and abundance from leaf tissue so this was also run in parallel. In addition, it was of interest to determine whether transcripts encoding thioesterase PAEs were expressed in the foliage. The optimal method for extraction of RNA was using TRI-reagent[173] (Sigma, Methods 2.6.1), where RNA was purified from 100 mg of each tissue type using acid guanidinium thiocyanate-phenol-chloroform extraction.

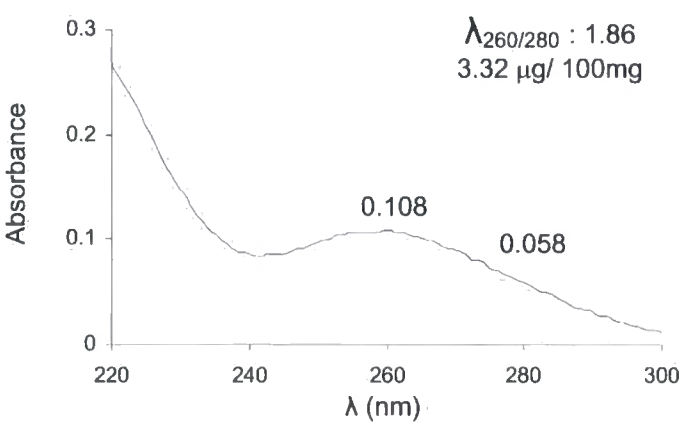
The purity of isolated RNA was determined through spectrophotometric analysis in order to ensure there was no genomic DNA contamination (Figure 5.5). RNA samples run on 1% agarose gels showed the ribosomal RNA doublet (18S and 28S rRNAs) and a background smear of varying sized mRNAs (Figure 5.5). Extracted RNA from leaf tissue (24 μg/100 mg tissue) was ten fold greater than from fruit samples (2.5-3.5 μg/100 mg tissue), however all were sufficiently pure for reverse transcription ($\lambda_{260/280} > 1.7$).

Reverse transcription of total RNA (1 μg) was performed using an oligo dT (OG2) 3' primer and AMV (Avian myeloblastoma leukaemia virus) reverse transcriptase. PCR on template cDNA (leaf, gland, endocarp) using *Taq* DNA polymerase with *Pe*For and *Pe*Rev (10 μM) primers was run using the standard PCR programme (Methods 2.6.3) with 30 cycles of amplification. The analysis of PCR products on agarose gels illustrated that several fragments had been amplified from each tissue type (Figure 5.6A). A 400 bp product was amplified in each case, and was absent in the PCR controls in which no reverse transcriptase was added when preparing template cDNA. The larger product (750 bp) appears to be due to genomic DNA contamination as it was also present in the controls (- reverse transcriptase).

Endocarp tissue:



Gland tissue:



Leaf tissue:

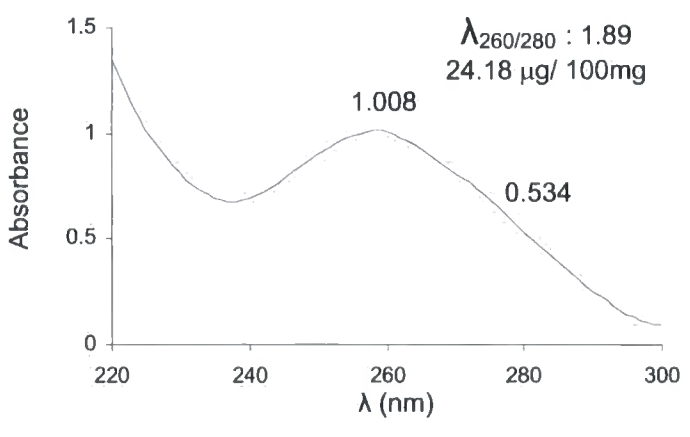
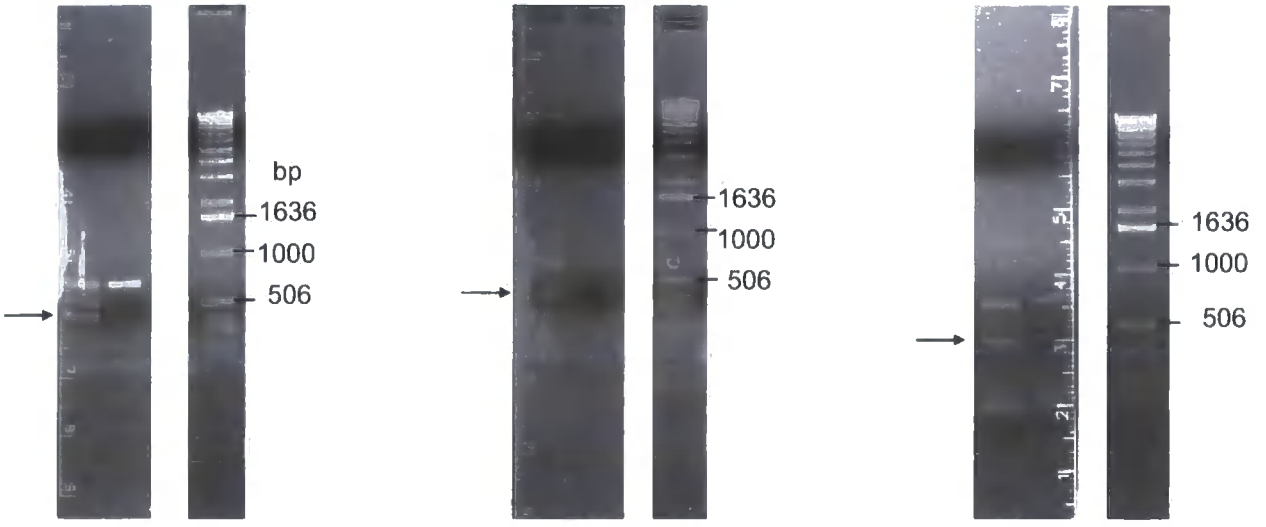


Figure 5.5 RNA extraction. Spectral scans used to determine the purity and abundance of RNA extracted from passion fruit. Agarose gel images of extracted total RNA. Note the doublet resulting from abundant 18 and 28S ribosomal RNA and the background smear caused by different sized mRNAs. $1A_{260}$unit: 40 µg/ml RNA, $\lambda_{260/280}$ ratio >1.7: good quality RNA with no genomic DNA contamination. Gels were loaded equally with 1 µl of each RNA sample.



1. Leaf +RT, 2. Leaf –RT, 3. Gland +RT, 4. Gland –RT, 5. Mesocarp +RT, 6. Mesocarp –RT.

B.

```
Leaf  5'  GAGGGAGGAGGATGGTGCAACAATGTCACGACCTGCCTTTATCGTAAGAA 50
Meso      GAGGGAGGAGGATGGTGCAATAATGTAACAAATTGCGTTAGTCGGATGCA
          ********************* ***** ** *  *** **  *** * * *
Leaf      TACTCATTTGGGTACATCTAAGCTAATGGGACAACCACTTGCTTTTTCTG 100
Meso      TACTCGATTAGGTTCATCGAAGAAAATGGTGGAGAACCTTGCTTTCTCAG
          *****  ** *** **** ***  *****   *    ******** ** *
Leaf      GGATTCTGAGCAACAGAGGGACATTTAATCCTGATTTCTATAATTGGAAT 150
Meso      CTATTCTTAGCAATAAGAAACAATATAATCCTGATTTTTACAATTGGAAT
            ***** ***** *     1. ** ************ ** *********
Leaf      AGAGTCAAGATTAGATACTGTGATGGTTCCTCTTTTACCGGTGATGTGCA 200
Meso      AGAGTGAAAGTTAGATACTGTGACGGGGCATCATTCACAGGAGATGTAGA
          ***** **  ************* **  * ** ** ** ** *****  *
Leaf      GAAAGTAAATCCTGCTACTAACCTTCACTTCAGAGGAGCAAGGATTTGGC 250
Meso      AGCAGTGAACCCTGCTACTAATCTTCACTTCAGAGGTGCTCGAGTTTGGT
            *** ** *********** ************** **  *  *****
Leaf      TAGCAGTTGTTCAGGAATTGTTGGCTAAGGGCATGAAAAATGCTGAAAAT 300
Meso      TAGCCGTTATGCAAGAGCTGCTAGCTAAAGGCCTGATAAACGCCGAGAAT
          **** *** * ** **3. ** * ***** *** *** *** ** ** ***
Leaf      GCACTTCTCTCTGGCTGTTCGGCCGGTGGATTGGCTTCGATTCTGCATTG 350
Meso      GCTGTTTTGTCTGGCTGTTCTGCTGGCGGGTTAACTTCGCTGATGCATTG
          **  ** * *********** ** ** ** **  ***** *  *******
Leaf      TGATAGCTTCCGAGCTCTGCTACCTATGGGTGCTAAAGTGAAATGCCTTT 400
Meso      TGATAGTTTCCGTGCTCTATTACCGATGGGAACCAAAGTAAAATGCCTTT
          ****** ***** *****  **** *****  * ***** **********
Leaf      CAGACGCTGGTT 3'                  2.                412
Meso      CTGACGCTGGTT
          * **********
```

Figure 5.6 A. PCR products following amplification with degenerate primers *Pe*For and *Pe*Rev as visualised on 1% agarose gels. The predicted PAE internal sequence (410 bp) is highlighted (arrow). Control reactions were run minus reverse transcription (-RT). **B.** The alignment of partial PAE cDNAs from leaf and mesocarp passion fruit tissues, highlighting consensus matches (*). The sequence used for designing further specific primers is shown (boxes), including sense primers (**1.**) and antisense primers (**2.**, **3.**).

The 400 bp products were excised from the gels and subcloned using the pGEM T-Easy vector system (Promega) with Top 10 *E.coli* competent cells (Invitrogen) prior to DNA sequencing (T7/ SP6 sequencing primers). The cDNA sequences from each tissue type (Figure 5.6B) were used to screen the databases using nucleotide BLAST and translated before screening using protein BLAST. The leaf and mesocarp sequences matched with putative pectin acetylesterase family proteins, whereas the gland sequence was a 3' product matching with an unknown cDNA clone from *Populus trichocarpa* (poplar), accession number: EK146327. Hence, the degenerate primers had successfully amplified a PAE internal sequence from leaf and mesocarp tissue. However, it was not possible to amplify PAEs from gland cDNA even under different PCR conditions. The leaf and mesocarp cDNA sequences shared 75 % sequence conservation over this internal region with the predicted translation products (Annex II D) sharing 79 % homology. It is therefore apparent purple passion fruit, like the model plants, contains multiple (at least 2) isoforms of PAE thioesterases.

Both protein sequences were subsequently screened for matches toward the five peptides from *Pe*WH, although it was predicted that only one of the five fragments (**5**) would fall into this stretch of sequence. The amino acid matches were similar but not identical (Annex II D). It appears the two cDNAs isolated from leaf and mesoocarp tissue do not encode *Pe*WH, but rather correspond to closely-related gene products. Further PCR reactions at increasing annealing temperatures with varying amplification cycles using tissues at different stages of ripening produced the same products. A "shotgun" approach was also used to sequence multiple clones (10 +) from each PCR 410 bp product, however no further PAE sequence could be found under such conditions. It was therefore decided to obtain the full length nucleotide sequences for the PAE products described above.

5.3 Full Length Passion Fruit PAE Sequence Through RACE

The missing 5' and 3' ends of the coding sequences were determined through rapid amplification of cDNA ends (RACE).

5.3.1 3' RACE

Nested gene specific primers were designed to the internal sequences from both leaf and mesocarp tissues (Table 5.2). For 3' RACE, sense primers termed *Pe*L1 (*Passiflora edulis* leaf 1) and *Pe*M1 (mesocarp 1) were designed to anneal to the 5' end of the internal fragments (Figure 5.6). A single round of PCR using the sense primer plus the complementary 3' primer (Og9) on the corresponding tissue cDNAs was sufficient to obtain 1 kp products, predicted to be the missing 3' ends (Figure 5.7). In this case a second round of nested PCR was not required as a single product was produced. Although smeary, cutting this band and employing a shotgun PCR approach using primers *Pe*L1+3 or *Pe*M1+3 to select for positive PAE products provided multiple clones for sequencing. The leaf cDNA was extended 752 bp (575 bp of open reading frame) whilst the mesocarp sequence was extended 765 bp (575 bp of open reading frame).

5.3.2 5' RACE

5' RACE was employed to determine the full length passion fruit PAE thioesterase sequences. In each case two nested antisense primers, *Pe*L2+3 (leaf) and *Pe*M2+3 (mesocarp) (Table 5.2) were designed at 600 and 530 bp from the 5' ends respectively. Hence two rounds of PCR using 5' primers was predicted to provide the missing 200 bp of sequence.

Initial attempts to amplify the 5' ends of PAEs from passion fruit resulted in truncated sequence, some 200 bp short of the full open reading frame. Interestingly, products were always truncated at this same point in the sequence and optimising PCR conditions (annealing temperatures, cycles) did not overcome this. It was therefore postulated this could be a consequence of tight RNA secondary structure which prevented the reverse transcriptase from reading the full sequence, hence generating truncated cDNA templates.

Name of oligonucletide	Experiment	Direction	Sequence
*Pe*For	internal sequence (leaf/ meso)	sense (degenerate)	5'- GAG GGA GGA GGA TGG TGY A -3′
*Pe*Rev	internal sequence (leaf/ meso)	antisense (degenerate)	5'- AWC CAG CAT CWG AAA VRC AYT T -3'
*Pe*L1	3' RACE (leaf)	sense	5′- TTA GAT ACT GTG ATG GTT CCT CTT TTA -3′
*Pe*M1	3' RACE (meso)	sense	5′- TTA GAT ACT GTG ACG GGG CAT CAT TCA -3′
Og2	oligo dT	antisense	5′- GAG AGA GGA TCC TCG AGT TTT TTT TTT TTT TTT T -3′
Og9	complementary to Og2	antisense	5′- CGC ACT GAG AGA GGA TCC TCG AG 3′
*Pe*L2	5' RACE (leaf)	antisense	5′- GAA AGG CAT TTC ACT TTA GCA CCC ATA -3′
*Pe*L3	5' RACE nested (leaf)	antisense	5′- ATC CAC CGG CCG AAC AGC CAG AGA GAA -3′
*Pe*M2	5' RACE (meso)	antisense	5′- GAA AGG CAT TTT ACT TTG GTT CCC ATC -3′
*Pe*M3	5' RACE nested (meso)	antisense	5′- ACC CGC CAG CAG AAC AGC CAG ACA AAA -3′
Adaptor dT	5' RACE (leaf/ meso)	sense	5′- CTT ATA CGG ATA TCC TGG CAA TTC GGA CTT TTT TTT TTT TTT TTT V -3′
Adaptor	5' RACE nested (leaf/ meso)	sense	5′- CTT ATA CGG ATA TCC TGG CAA TTC GGA CTT -3′
*Pe*L4	full sequence leaf	sense	5′- GCG CGC TTA ATT AAC CAT ATG GTA GAT TCA AGA CTT TCA TG -3′
*Pe*L5	full sequence leaf	antisense	5′- CGC GCG GTC GAC CTA TAC TTG GAA TGG TTG AGG AAC TTT GC -3′
*Pe*L6	full sequence leaf (- signal sequence)	sense	5′- GCG CGC TTA ATT AAC CAT ATG CTC TAT GTG AAC ATT ACA TAT G -3′
T7	sequencing		5′- TAA TAC GAC TCA CTA TAG GG -3′
SP6	sequencing		5′- CGT ATT TAG GTG ACA GTA TAG -3′

Table 5.2 Sequences of oligonucleotides used in the cloning of putative PAEs from *Passiflora edulis* Sims.
Y = C or T, W = A or T, R = A or G, V = A or C or G

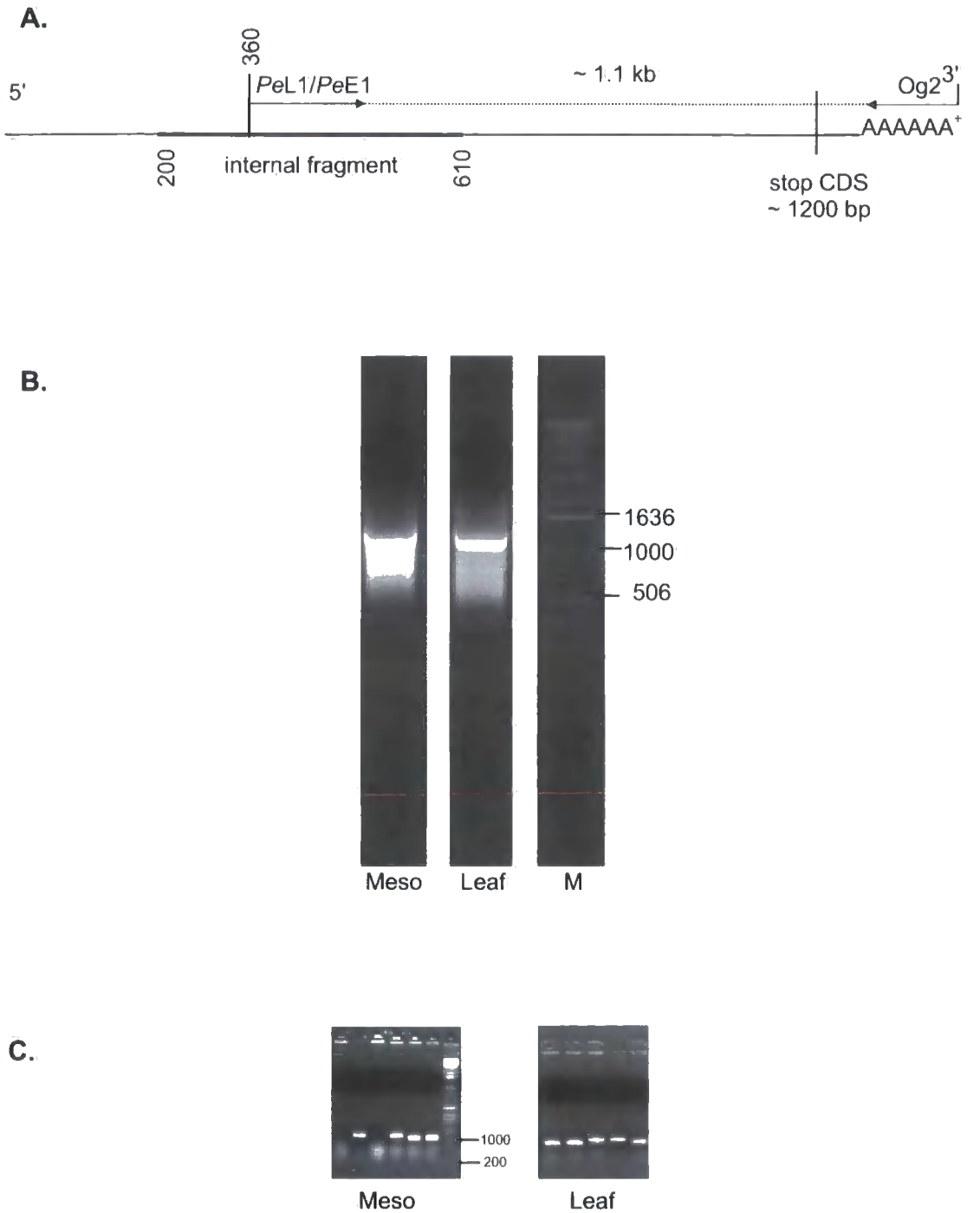


Figure 5.7 A. 3' RACE map of predicted PAE products. Extension is predicted to generate 840 bp of open reading frame, plus a section of non-coding 3' sequence before the poly A tail. **B.** PCR products following amplification with gene specific primers (*Pe*L1 or *Pe*M1) and Og9 as visualised on an agarose gel. Product at 1.1 kb is predicted PAE 3' sequence. **C.** PCR test for positive clones using *Pe*L1+3 or *Pe*M1+3, expected product 170 bp.

This was further supported by an Mfold[128,174] RNA secondary structure prediction (Figure 5.8) based on the known PAE coding sequence from mung bean (Accession: X99348.1). A major branch point can be seen in the folded sequence at approximately 200 bp, the cut off point for the truncated sequences. A new approach for the synthesis of first-strand cDNA was therefore required. A genetically modified reverse transcriptase, RevertAid M-MuLV (Fermentas) which could tolerate higher extension temperatures was used and the reactions run at 50°C. This overcame the difficulties in template production from leaf tissue but not from mesocarp material, where it is possible an even tighter secondary folding was present. The 5' RACE used to complete the leaf sequence is described below.

Reverse transcription was carried out on total RNA using the antisense primer *Pe*L2 before a homopolymer (polyA) tail was subsequently added to the 5' end of the reaction products using a terminal deoxynucleotidyl transferase. The first round of PCR using *Pe*L2 and an oligo dT (adaptor dT) primer, which annealed to the extended 5' end, generated a single (smeary) 600 bp product (Figure 5.9B). The nested primer *Pe*L3 in combination with a 5' adaptor primer were used for the second round of PCR in which a single sharp band at 530 bp could be seen on an agarose gel (Figure 5.9B). Products were cloned and PCR used to select for positives (Figure 5.9C). The full 5' leaf cDNA was obtained, where the sequence was extended 219 bp, including the ATG start site (methionine) and 200 bp of extended open reading frame (Figure 5.10). The partial PAE sequence from mesocarp tissue is depicted in Annex II E.

To obtain the full-length sequence from *Pe* leaf, primers *Pe*L4 and *Pe*L5 (Table 5.2) were designed to the 5' and 3' termini plus restriction sites (5': *Pac*I, 3': *Sal*I) to allow sub-cloning into a modified pET-41a vector (Novagen) such that the sequence would contain an N-terminal Strep tag. The vector termed, pET-STRP3, was constructed by Dixon *et al.*,[175] (2008) to encode an N-terminal Strep fusion protein and further restriction sites (*Pac*I and *Bst*XI) to enable subcloning between bacterial and plant vector systems. A further primer was designed (*Pe*L 6) for the amplification of a truncated 5' product, minus the predicted signal peptide sequence (Figure 5.10, Table 5.2).

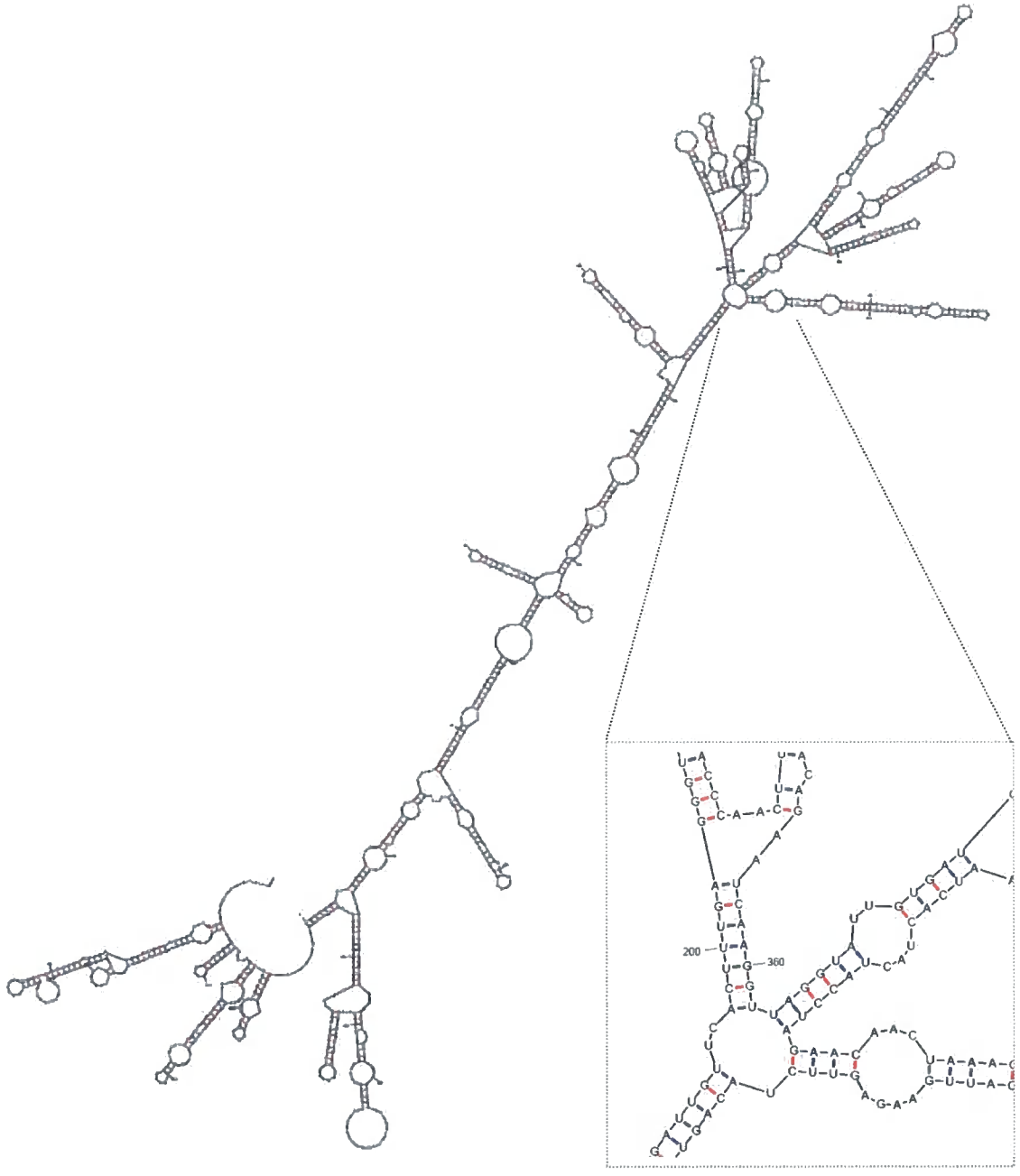


Figure 5.8 RNA secondary structure prediction for mung bean (*Vigna radiata*) PAE mRNA sequence (Accession: X99348). RNA folding was determined using mfold version 3.2 (http://mfold.bioinfo.rpi.edu/), which predicts RNA folding upon thermodynamic properties[128,174]. The major branch point at approximately 200 bp is likely to require a higher melting temperature and hence a heat stable reverse transcriptase was required (>45°C).

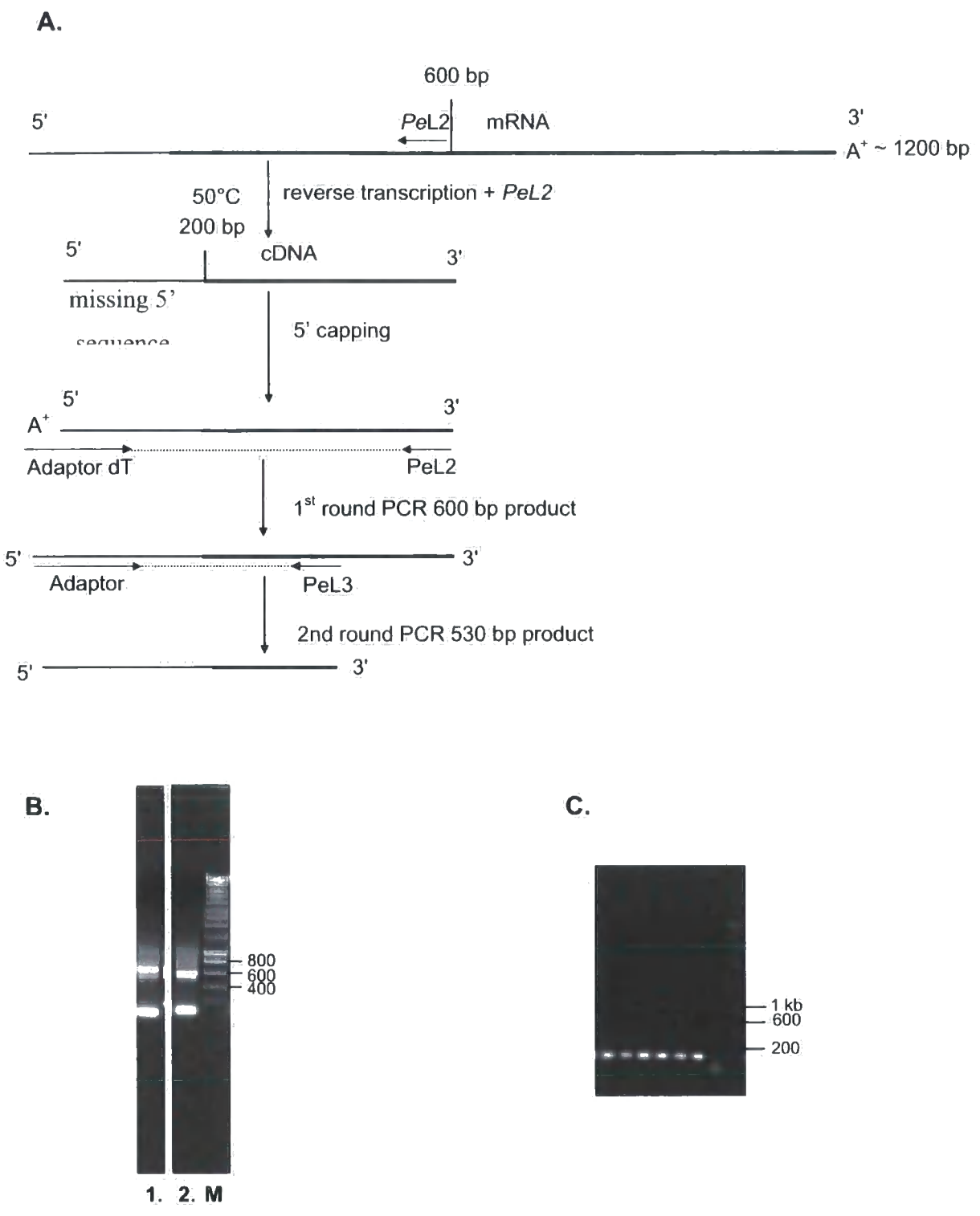


Figure 5.9 A. 5' RACE map of predicted PAE products. The initial PCR reaction is predicted to produce a 600 bp product. A second round of nested PCR on this is predicted to give a 530 bp product. Extension is predicted to generate 200 bp of missing open reading frame. **B.** PCR products; **1.** 1st round PCR following amplification with the gene specific primer *Pe*L2 and adaptor dT, **2.** 2nd round PCR using *Pe*L2 + adaptor primer. **C.** PCR test for positive clones using *Pe*L1 + *Pe*L3 , expected product 170 bp.

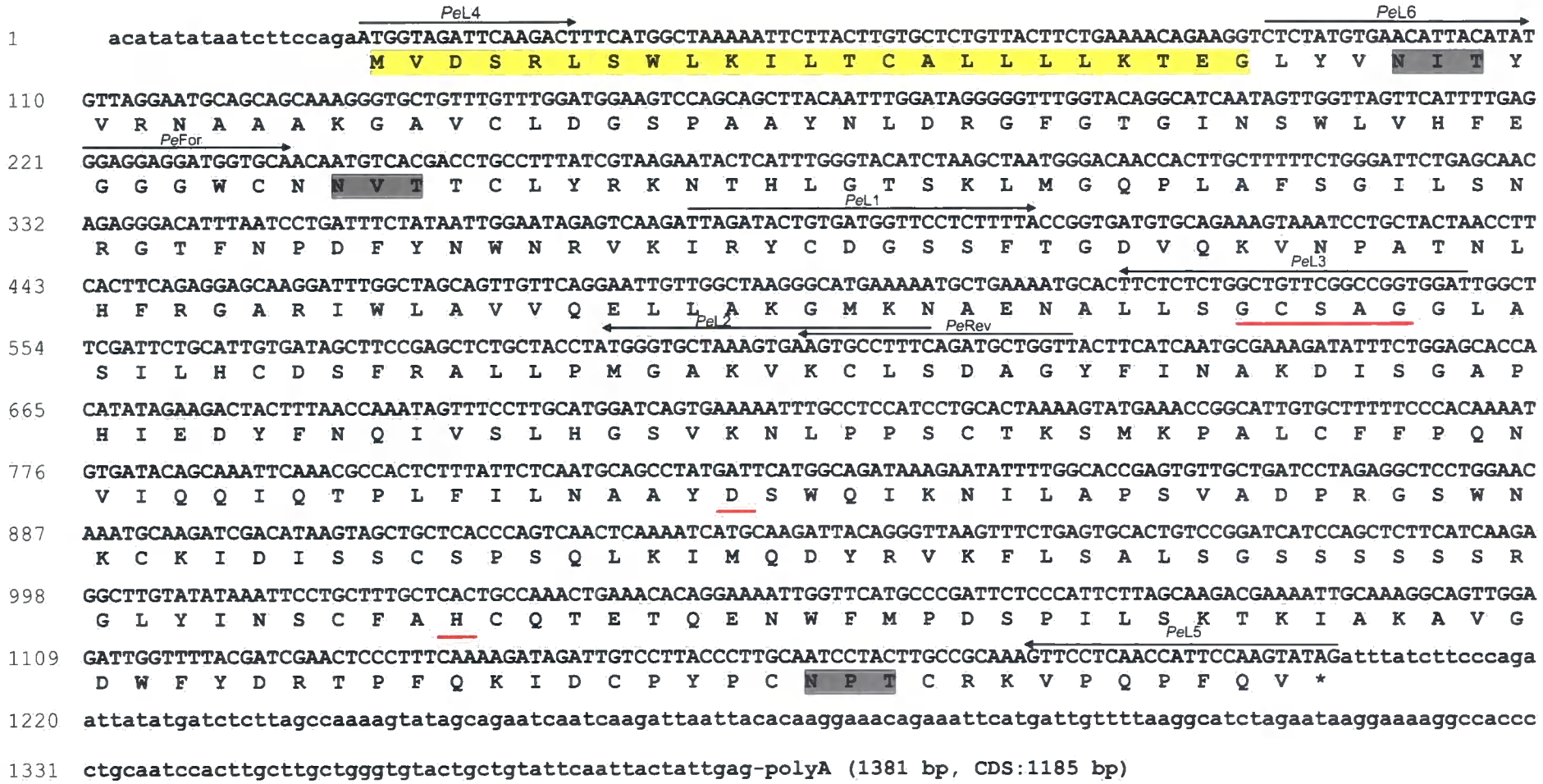


Figure 5.10 Full length nucleotide and deduced amino acid sequence of the *Pe*PAE cDNA from passion fruit leaf tissue. The GXSXG serine hydrolase catalytic motif and the histidine and aspartate residues which make up the catalytic triad are underlined (red). Signal peptide is highlighted (yellow). The three potential *N*-glycosylation sites are highlighted (grey). The regions used for primer design are underlined (black arrows). Predicted pI: 9.1, Mr: 42876 Da (mature protein). Predicted disulphide bonds between cysteine residues (49+55, 215+224, 354+358, 268+275, 149+160, 314+362). Predictions made using ExPASy proteomics tools.

As bacterial expression in E.coli would not recognise plant signal sequences and cleave them to produce a mature protein it was important to clone and express both forms (+/- signal sequence) to determine which would be stable and active using such an expression system. Both full length clones were amplified using the proof reading DNA polymerase, KOD, and sequenced in both directions to ensure correct sequence. Overlapping sequences were aligned to ensure the correct ORF had been produced. The clones were named *Pe*PAE+ and *Pe*PAE- (-signal sequence).

5.4 Expression of Putative Thioesterases

5.4.1 *Arabidopsis thaliana* PAE clone construction

Cloning of putative PAEs from *Passiflora edulis* was challenging and for this reason it was only possible to produce a single full length clone from leaf tissue for subsequent expression studies. Although this clone, *Pe*PAE, did not match with the purified thioesterase *Pe*WH, it encoded a homologous enzyme likely to have similar catalytic activity. In addition, it was decided to clone a further putative PAE from the model plant *Arabidopsis thaliana*.

The mung bean protein (Accession: CAA67728) previously purified and partially characterized was shown to possess acetylesterase activity toward pectin and triacetin, and shared many of the characteristics of the thioesterase *Pe*WH (size 43 kDa, pI: 9)[158]. The closest *Arabidopsis* homologue to this was therefore chosen, accession number: AAT70429 (Figure 5.11). Primers were designed to the 5' (*At*PAE1) and 3' (*At*PAE2) ends of the coding sequence (accession number: BT014978), and a further nested 5' primer (*At*PAE3) used for amplification minus signal peptide (Figure 5.12A+B). 5' Primers contained *Nde*I sites and the 3' primer a *Sal*I site for sub-cloning into the pET-STRP3 vector such that the proteins would contain N-terminal Strep tags.

Total RNA extracted from leaf tissue was reverse transcribed using the heat stable reverse transcriptase, RevertAid M-MuLV, at 50°C plus *At*PAE2 to generate first strand cDNA template.

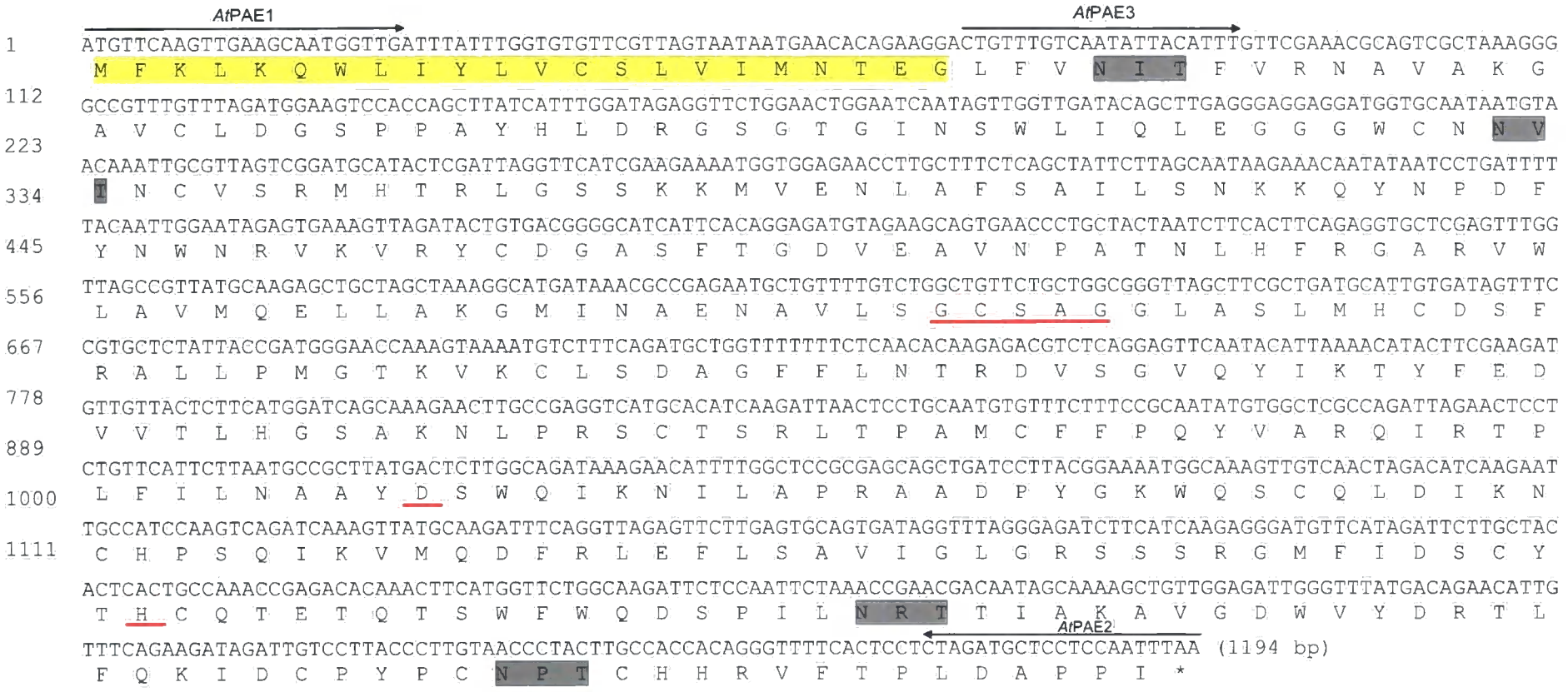
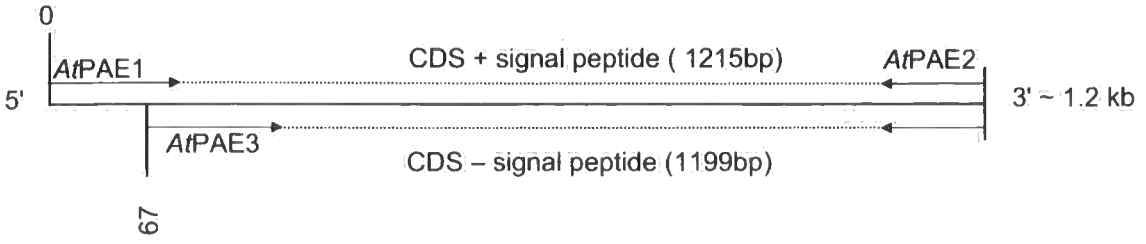


Figure 5.11 Full length nucleotide and deduced amino acid sequence of the *At*PAE cDNA from *Arabidopsis thaliana* (Accession number: BT014978). The GXSXG serine hydrolase catalytic motif and the histidine and aspartate residues which make up the catalytic triad are underlined (red). Signal peptide is highlighted (yellow). The regions used for primer design are underlined (black arrows). Predicted pI: 9.0, Mr: 41827 Da (mature protein). Predictions made using ExPASy proteomics tools[15]. The nucleotide sequence shares 69% homology with *Pe*PAE and 66% with the purified mung bean PAE.

A.



B.

*At*PAE1	sense	5′- GCG CGC CAT ATG TTC AAG TTG AAG CAA TGG TTG -3′
*At*PAE2	antisense	5′- GCG CGC GTC GAC TTA AAT TGG AGG AGC ATC TAG -3′
*At*PAE3	sense	5′- GCG CGC CAT ATG CTG TTT GTC AAT ATT ACA TTT G -3′

C.

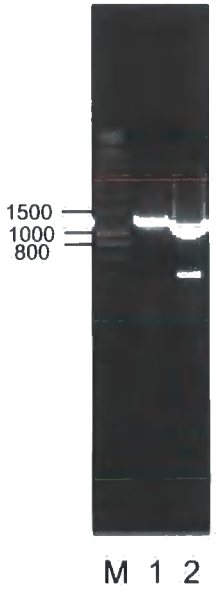


Figure 5.12 **A.** A map of the predicted fragments amplified using *At*PAE1 and *At*PAE2 (or *At*PAE3 and *At*PAE2). **B.** Forward and reverse primers, *At*PAE1, *At*PAE2 and *At*PAE3 used for the amplification of *At*PAE sequences. **C.** PCR products; **1.** full length sequence amplification using *At*PAE1 and *At*PAE2 (1215 bp product), **2.** truncated sequence amplification with *At*PAE3 and *At*PAE2 used to generate cDNA –signal sequence (1199 bp product).

PCR reactions using the proof reading DNA polymerase KOD under the standard conditions amplified products of the expected size, 1190 and 1130 bp (+/- signal respectively) (Figure 5.12C). Products were cloned into the pET-STRP3 vector and sequenced in both directions to ensure the correct coding sequence had been amplified. The full length clones were error free and named *At*PAE+ and *At*PAE- (-signal sequence).

5.4.2 Bacterial Expression

Four clones were used in expression studies. The putative thioesterase from purple passion fruit, with and without signal peptide (*Pe*PAE+, *Pe*PAE-) and the homologue from *Arabidopsis*, again with and without signal sequence (*At*PAE+, *At*PAE-). Each clone was overexpressed in *E.coli* with the aim of characterising the activities of the recombinant enzymes toward various natural and synthetic thioesters. It was also hoped this would support the identification of purified fruit thioesterases as PAEs. All four expression constructs were used to transform *E. coli* competent cells, strain Tuner (DE3) (Novagen) containing the pRARE plasmid from strain Rosetta (Novagen), and the resulting colonies grown in 10 ml LB medium starter cultures overnight at 37°C before inoculating larger 100 ml cultures. Initially, standard expression conditions were used whereby cultures were grown at 37°C to OD_{600} ~ 0.7 when 1 mM IPTG was added to initiate recombinant protein expression for 3 hours. Bacteria were then harvested, sonicated and the lysates analysed by SDS-PAGE and western blotting. A comparison of the crude soluble and insoluble fractions suggested that the recombinant proteins were being expressed, however, they were present only as inclusion bodies (Figure 5.13A). No thioesterase activity could be determined in any fraction.

Efforts were subsequently made to optimise expression including slower growth (15-30°C overnight) and reducing the IPTG concentration (0-0.1 mM). In addition, terrific broth, a nutrient rich medium was tested[176]. A slower induction may enable the correct folding of the protein and prevent it accumulating to high levels were it may be aggregating into inclusion bodies.

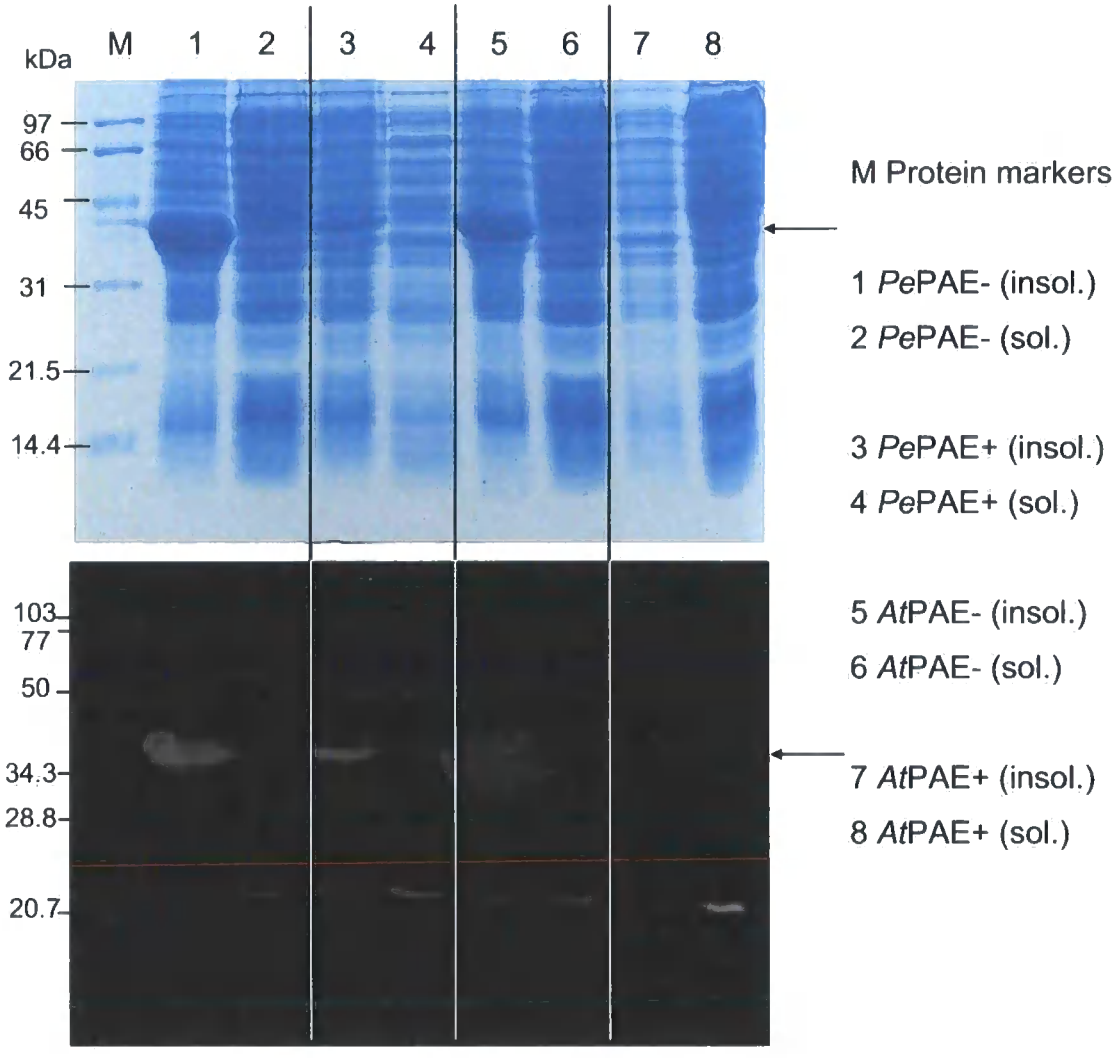


Figure 5.13 Recombinant protein expression in *E.coli.* **A.** Expression under standard conditions; 100 ml cultures grown at 37°C until dense (OD_{600} ~ 0.7) + 1 mM IPTG for 3 hours. Total crude protein (100-fold concentrate) is illustrated in the coomasie stained gel (top). The detection of recombinant protein was achieved through western blotting the same gel (below) using an anti Strepdtactin antibody, Strep-Tactin alkaline phosphatase conjugate (IBA BioTAGnology), 1:5000 dilution and incubation for 1-2 hr at room temperature. Expected protein size marked (arrow). The accumulation of insoluble recombinant protein can be seen in each case, except *At*PAE+ where the clone appears to have failed to express. Minimal soluble expression can be detected in lanes 2 (*Pe*PAE-) and 4 (*Pe*PAE+), however this may be due to protein diffusion during blotting.

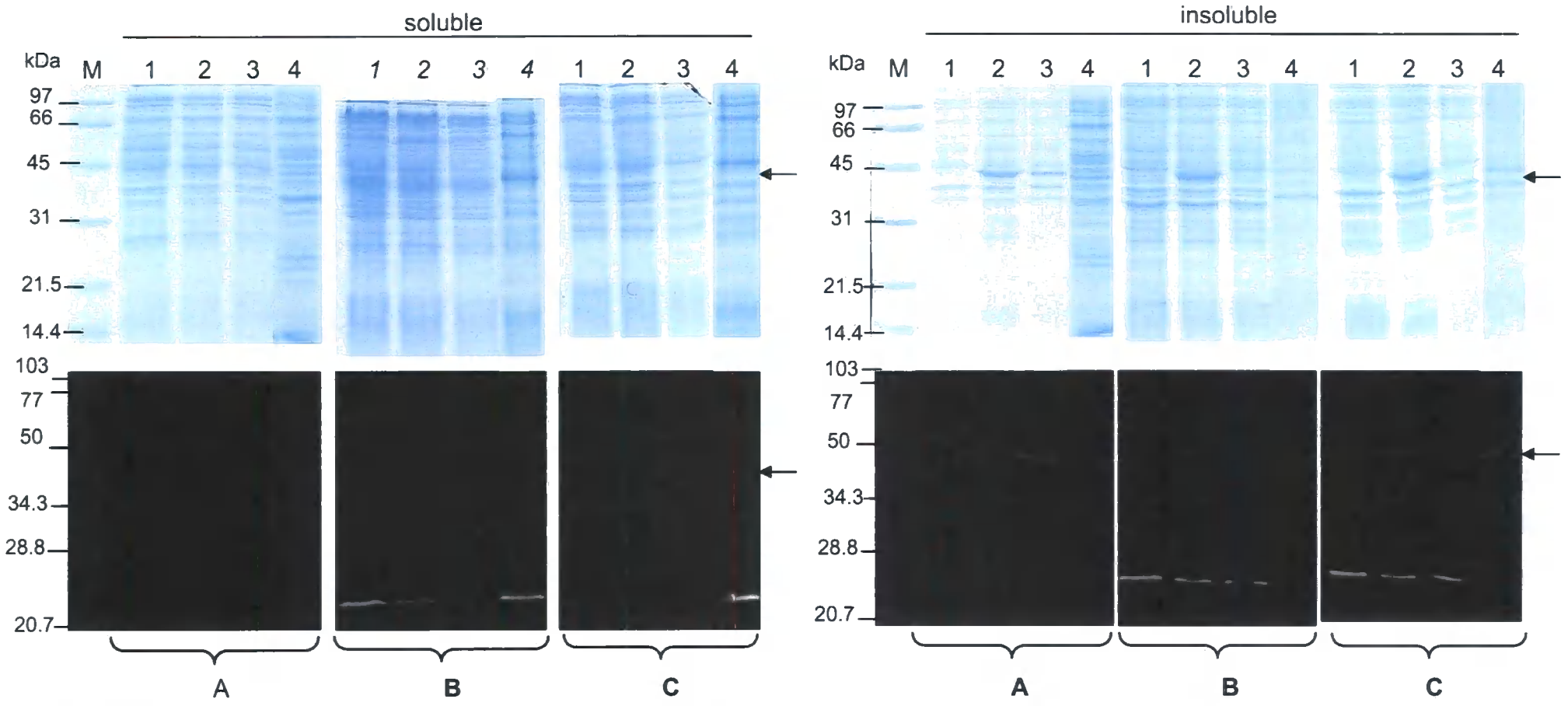
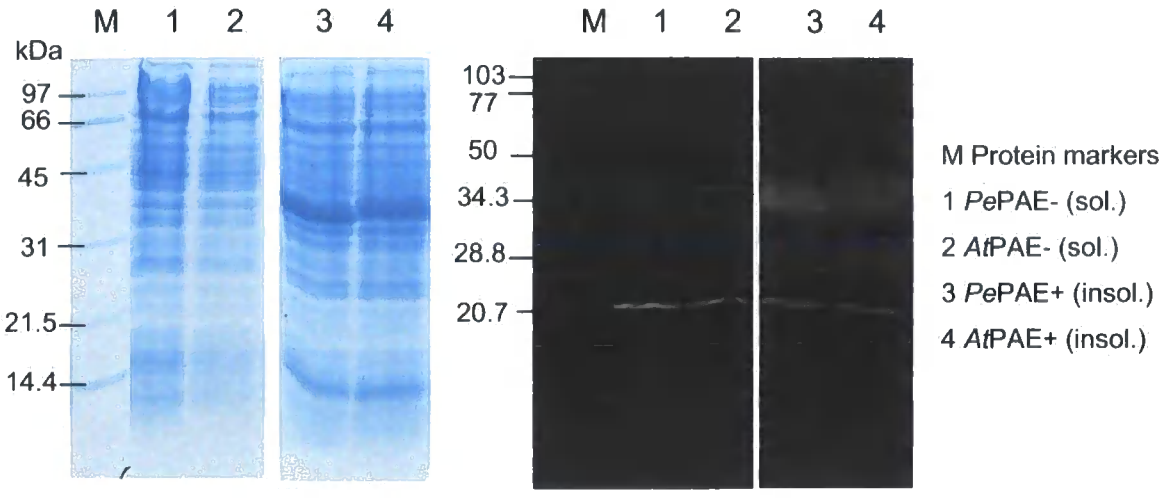


Figure 5.13 Recombinant protein expression in *E.coli.* **B.** Optimising expression. Three different expression conditions were chosen; 30°C overnight growth + 1 mM IPTG (**A**), 30°C overnight growth + 0.1 mM IPTG in terrific broth (**B**), 15°C overnight growth - IPTG (**C**). Total crude protein (100-fold concentrate) is illustrated in the coomasie stained gels (top). The detection of recombinant protein through western blotting (below). The expression of recombinant protein using each clone is shown; *At*PAE+(lane **1**), *At*PAE- (**2**), *Pe*PAE+ (**3**), *Pe*PAE- (**4**). Soluble protein left, insoluble right. Expected protein size marked (arrow). The accumulation of insoluble recombinant protein can be seen in each case, except *At*PAE+ where the clone appears to have failed to express. Minimal soluble expression can be detected in lane 3 (*Pe*PAE+).

SDS-PAGE analysis and western blotting indicated that this did reduce the accumulation of the recombinant protein in the insoluble fractions. However, only minimal protein could be determined in the soluble fractions (*Pe*PAE+) (Figure 5.13B). No fractions were found to posses thioesterase activity.

Optimising the expression conditions above did produce low levels of soluble protein from the passion fruit PAE clone + signal peptide (*Pe*PAE+). However, this protein did not demonstrate the associated thioesterase activity and the signal sequence may prevent correct protein folding. As a final approach, large cultures (500 ml) were used for the expression of both the *Arabidopsis* and passion fruit clones minus signal sequence (*At*PAE- and *Pe*PAE-). Expression under slow growth conditions (15°C + 0.1 mM IPTG) did generate detectable levels of soluble recombinant protein from the larger cultures in each case (Figure 5.14A). Most was however present as inclusion bodies. The soluble recombinant protein was loaded onto a Strep-tactin column (IBA technologies) and following washing was eluted as a pure protein in a single purification step (Figure 5.14B). Such a system for the purification of recombinant strep tagged plant proteins had been demonstrated previously for the isolation of glutathione transferases (GST) from *Arabidopsis thaliana* (Dixon et al., 2008)[177]. However, in this case only minimal soluble recombinant protein was recovered (< 1.5 mg) and when the thioesterase activity of the corresponding fractions was determined using the Ellman's method no significant enzymatic rate could be found. A different expression system was therefore required.

A.



B.

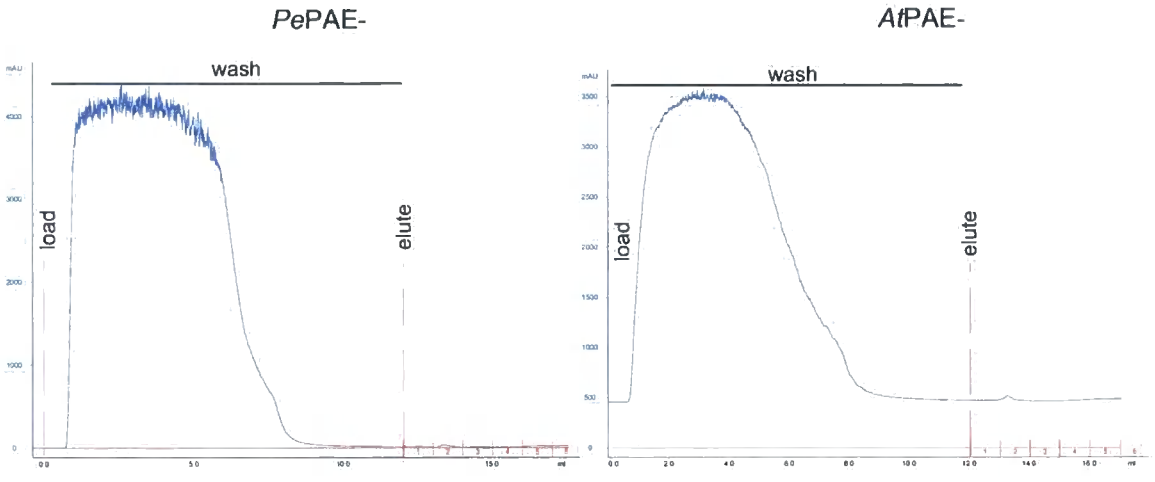


Figure 5.14 Large scale recombinant protein expression in *E.coli*. **A.** 500 ml cultures were expressed at 30°C overnight + 0.1 mM IPTG until dense (OD_{600} ~ 0.7). Total crude protein is illustrated in the Coomassie stained gel (left) and the detection of recombinant protein through western blotting (right). The clones form passion fruit and *Arabidopsis* – signal were used (*Pe*PAE- and *At*PAE-) and their expression in soluble (lanes 1 + 2) and insoluble (lanes 3 + 4) fractions is shown. Soluble expression can be detected in both cases (blots lanes 1 and 2). **B.** The purification of Strep-tagged soluble protein through affinity purification. Minimal recovery of recombinant protein was achieved.

5.5 Discussion

Cloning of putative thioesterases from the tropical fruit, purple passion fruit, led to a study of known and putative pectin acetylesterases in plants. The phylogenetic analysis indicated that from the collection of full length plant sequences (n=28), 3 major sub-families existed, each with significant differences in protein primary sequence. It is possible that such enzymes of primary metabolism predicted to be involved in cell wall modification[160,167], have evolved through divergence to form multiple isoforms involved in further plant metabolism. This supports our studies which suggest plant PAEs may serve a further role in secondary metabolism through the liberation of thiol and alcohol volatile flavouring agents from thioester and ester precursors in the cortex of fruit. The isoforms may have evolved subtly different active site structures which allow for the accommodation of different chemical substrates (i.e. flavour esters and thioesters). Changes in the surface properties of the enzyme may also be found which influence trafficking of the protein and localisation within the cell wall where they could potentially become involved in novel metabolism.

The cloning of a full length PAE from passion fruit leaf tissue, *Pe*PAE, adds a further sequence to the collection of plant acetylesterases. Aligning with similar protein sequences to members of sub-family one, *Pe*PAE, shares 68% nucleotide sequence and 75% amino acid conservation with the acetylesterase purified and cloned from mung bean[158,160]. Both sequences share many similarities, including 5' signal sequences which direct the protein to the cell wall and the GXSXG serine hydrolase catalytic motif. A predicted secondary structure analysis on translated *Pe*PAE indicates the mature protein contains 8 major α-helices and 6 β-strands; such folding is indicative of α/β hydrolase family enzymes[178] (Figure 5.15). In addition *Pe*PAE is predicted to contain multiple glycosylation sites (Figure 5.10), which fits with the observation that the purified thioesterase *Pe*WH bound to conconavalin A Sepharose. Glycosylation of the protein may localise it to specific regions of the cell and act as a means of controlling the regulation of enzyme activity between the cytosol and cell wall, however, this remains to be studied.

A.

```
AA      LYVNITYVRNAAAKGAVCLDGSPAAYNLDRGFGTGINSWLVHFEGGGWCNNVTTCLYRKN
Pred    CEEEEEEECCCCCCCCEECCCCCCCEEEECCCCCCCCCEEEEECCCCEECCCCHHHHHCC
                1                       2               3          A

AA      THLGTSKLMGQPLAFSGILSNRGTFNPDFYNWNRVKIRYCDGSSFTGDVQKVNPATNLHF
Pred    CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCEEEEEEECCCCCCCCCCCCCEECCCCEEEE
                                                4

AA      RGARIWLAVVQELLAKGMKNAENALLSGCSAGGLASILHCDSFRALLPMGAKVKCLSDAG
Pred    ECHHHHHHHHHHHHHHCCCCCCEEEEECCCHHHHHHHHHHHHHHHCCCCCEEEEEECCE
                    B               5               C          6

AA      YFINAKDISGAPHIEDYFNQIVSLHGSVKNLPPSCTKSMKPALCFFPQNVIQQIQTPLFI
Pred    ECCCCCCCCCCCCHHHHHHHHHHHCCCCCCCHHHHHCCCCHHHCCHHHHHHHCCCCEEE
                                D

AA      LNAAYDSWQIKNILAPSVADPRGSWNKCKIDISSCSPSQLKIMQDYRVKFLSALSGSSSS
Pred    EECHHHHHHHCCEECCCCCCCCCCCEEEEECCCCCCCHHHHHHHHHHHHHHHHHHHHHCC
            7                                                   E

AA      SSRGLYINSCFAHCQTETQENWFMPDSPILSKTKIAKAVGDWFYDRTPFQKIDCPYPCNP
Pred    CCCEEECHHHHHHHHHHCCCCCCCCCCCCCCCCEEHHHHHHHHHHCCCCEEEECCCCCCCC
            8                                       F

AA      TCRKVPQPFQV
Pred    CCCCHHCCCCC
```

B.

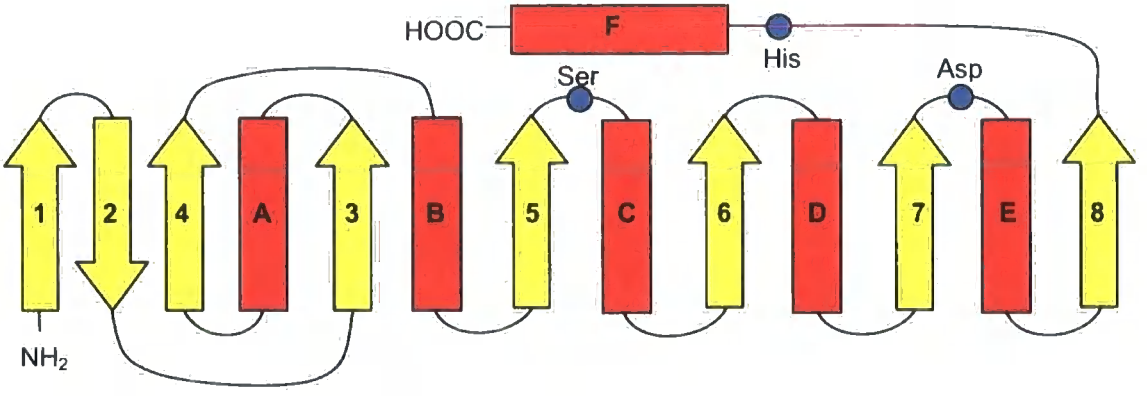


Figure 5.15 A. Protein secondary structure prediction for translated *Pe*PAE. H: predicted α-helix, E: predicted β-strand, C: predicted coil. Active site residues are highlighted in blue (serine, histidine and aspartate). The structure prediction was made using the protein structure prediction server PSIPRED (http://bioinf.cs.ucl.ac.uk/psipred/), developed by Jones,[179] 1999. **B.** The arrangement of secondary structure in α/β hydrolase fold enzymes, adapted from Ollis *et al.*,[178] 1992. The assigned β-strands (1-8) and α-helices (A-F) can be closely matched on the *Pe*PAE predicted structure (top).

Although *Pe*PAE does not encode the purified thioesterase *Pe*WH it is a member of the same family of proteins and is likely to demonstrate similar catalytic properties. The Edman degradation of purified *Pe*WH may provide the further amino acid sequence required for the design of degenerate primers for cloning the corresponding cDNA. However, the PAE degenerate primers used to clone *Pe*PAE form important molecular tools for the amplification of further putative PAE sequences from other plant sources. Cloning of further PAEs from plants would increase the collection of acetyl esterases available for researchers interested in plant cell wall metabolism and in determining the role of hydrolases in fruit ripening.

Efforts toward the expression of PAEs from both passion fruit and *Arabidopsis* were unsuccessful, a possible consequence of the hydrophobic nature of this wall bound enzyme. Attempts were made to express each clone in *Nicotiana benthamiana*, however, no recombinant protein was detected in each case. Expression of the clones in further systems is required (e.g. yeast; *Pichia pastoris*) and if successful would be very informative. The purified enzyme could be fully characterized and if possible the first crystal structure for a plant PAE could be produced so as to elucidate the cause of such tight chemoselectivity. It is likely a constrained binding pocket may account for the accommodation of short chain (predominantly acetyl) components as the acyl moieties. Interestingly, microbes are also found to contain acetylesterase enzymes, believed to function in the degradation of plant cell walls as a means of obtaining nutrients and breaking this tough physical barrier[180]. Of the 59 entries on the pectin acetylesterase database two include sequences from the thermophilic bacteria, *Thermus thermophilus*, (Accessions: AAS81248.1, BAD71091.1). Cloning and expression of these proteins may yield heat-stable, selective enzymes with great potential for application as biocatalysts.

It remains unlikely, however, that the recombinant expression of plant PAEs could provide sufficient yield of the enzyme for it to be considered a means of obtaining this potential biocatalyst and hence a further source is required.

6. Thioesterases from Orange (*Citrus sinensis*) and the Bioproduction of High-Impact Flavourings

6.1 Introduction

Attempts to express the thioesterases *Pe*WH and *At*WH in bacteria and plants were unsuccessful, a possible consequence of the unusually hydrophobic nature of the proteins which may explain their aggregation into inclusion bodies. Recombinant expression of the biocatalysts was therefore non-viable and an alternative source of the enzyme had to be sought. Initially it was proposed that tropical fruit would be unique in containing thioesterases which were believed to be present in only trace amounts, however, it appears this is not the case. The identification of *Pe*WH as a homologue to pectin acetylesterase (PAE) enzymes, a sub-family of α,β-hydrolases, which are found throughout the plant kingdom[150,158] indicated that these proteins have an important function in addition to their proposed role in primary metabolism. Furthermore, the identification of putative PAEs in *Arabidopsis* and rice supports the generalist role of such enzymes in plants and it is therefore likely thioesterase PAEs could be sourced from several sources; plant or fruit material. In particular, PAEs have been noted in *Citrus* where they are found at significantly high levels in peel mesocarp[167]. Hence, a screen of further fruit species, including *Citrus*, was conducted as a means of identifying the optimal source of this biocatalyst.

It was important to consider the source of the plant material, if it was ultimately going to be used in bioprocessing it had to be cheap, readily available and of reliable quality. The screen therefore concentrated on fruit widely used in the juicing industries (e.g. *Citrus*) and those grown in the UK. Fruit processing for juicing is a big industry with a global market value of $56.7 billion (US)[181], and following flavour extraction peel is regarded as a by-product, often discarded or used as animal feed[182]. However, with the realisation that such material is rich in hydrolases with potential application as biocatalysts a novel end use may emerge.

This alternative approach using waste fruit material has a number of merits: it does not require the use of microbes or genetic modification and is favourable from a consumer perspective. In addition, costly fermentation equipment and changes to manufacturing using microbes are not required; the fruit peel preparation could be regarded just like any synthetic catalyst. Once the life-span of the catalyst is reached the material could be easily discarded as it is non-toxic and fully biodegradable.

There are however a number of challenges to this approach: it is likely the enzyme will have to be processed and immobilised to increase stability and aid in recovery. Furthermore, the catalytic content must be abundant enough in the initial material for it to be a cost-effective source of the enzyme. Finally, crude enzymatic preparations are prone to catalyzing undesirable side reactions which must be avoided. This chapter will look at screening fruit for an optimal source of thioesterase enzymes before investigating how to achieve maximum productivity from the chosen material.

6.2 Screening Fruit for Thioesterase Activity

Crude protein extracts were prepared from nine fruits including *Citrus* (lemon, lime, orange and grapefruit), tropicals (mango, yellow and purple passion fruit) and berries (raspberries and blackcurrant). In each case 100 g of peel tissue was homogenised, with the exception of berries where whole fruit was used. To ensure the complete extraction of wall-bound proteins, 0.25 M salt was added to the extraction buffer (Methods 2.2). Dialysed protein samples in 0.1 M phosphate buffer, pH 7.2 were subsequently assayed for both thioesterase and esterase activity toward the panel of substrates using the Ellman's method (Figure 6.1). The use of microtitre plates (96 wells) in combination with a FLASHscan® spectrometer (Analytik Jena AG, Germany) enabled multiple assays to be run over short time periods (10 min). Such a system could be used in further high-throughput screens for hydrolases or lyases involved in organosulphur volatile transformations.

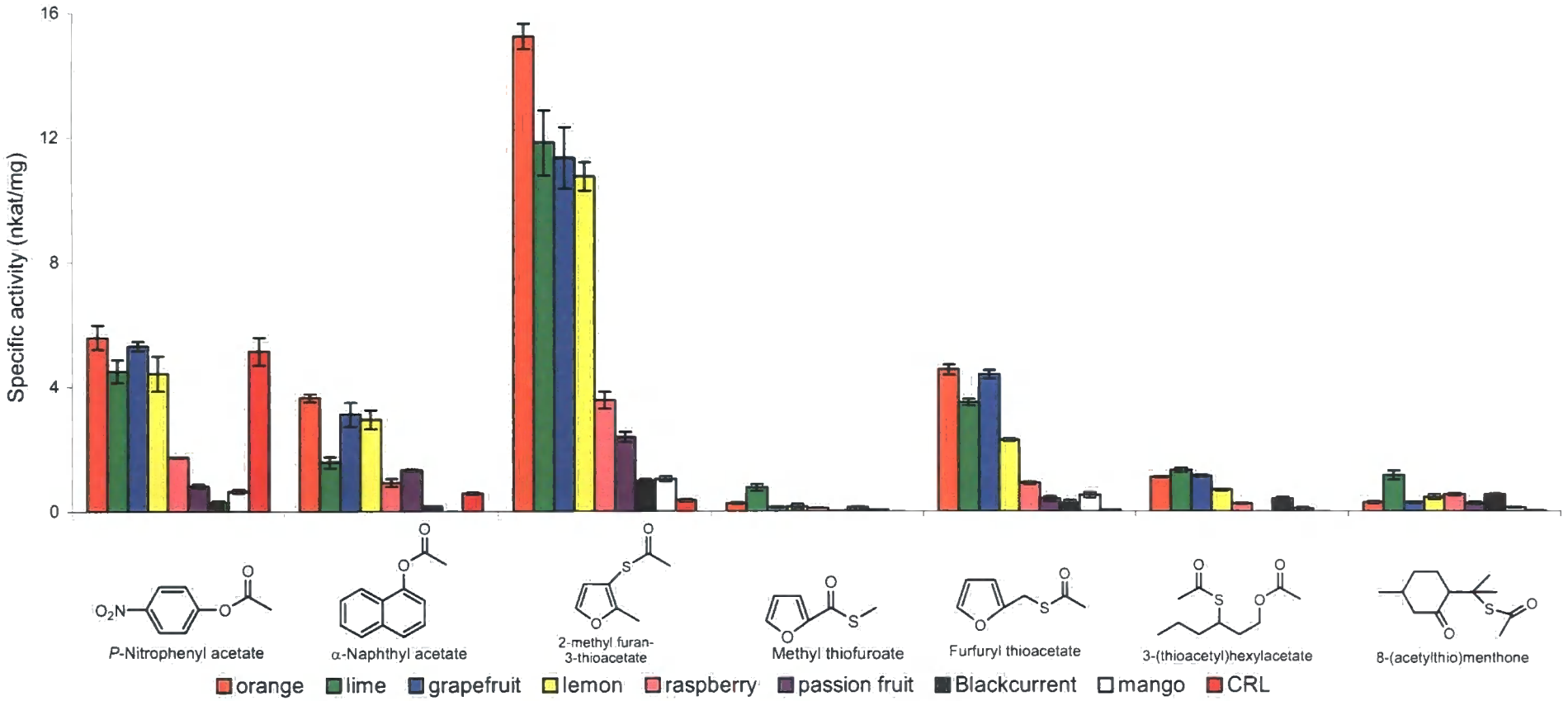


Figure 6.1 A Comparison of fruit hydrolase activities toward thioester and carboxylester substrates. Protein preparations (0.5-3 mg/ml) from each fruit type were incubated with 1 mM substrate at 30°C. For comparison the specific activities of the microbial lipase (CRL) are also shown. Results are averages of triplicates with errors +/- 1 standard deviation from the mean. No active protein was extracted from the pectin rich yellow passion fruit and hence no activities could be determined.

Orange had the greatest specific activity toward MFTA (15.24 nkat/mg), with its activities toward all other substrates comparable to, or greater than, the other fruit preparations. Grapefruit, lemon and lime all demonstrated relatively high levels of activity with the different substrates, illustrating that *Citrus* species are particularly rich in both thioesterase and carboxylesterase activities. Interestingly, the pattern of activities determined for the fruit enzymes were comparable. Whereas CRL showed a greater preference for carboxylester substrates such as *p*-NA and α-NA. It therefore appeared that although the fruit extracts were particularly active toward cyclic thioacetyls, such as MFTA, the energy of the thioester/ carboxylester bonds and the stability of the respective anions were also important factors in determining the pattern of activities observed.

It was also important to consider total activities per 100 g of fruit peel, rather than just specific activities which do not reflect the catalytic capacity of each tissue type by weight. Figure 6.2 illustrates the amount of protein extracted from each fruit (100g tissue); this was found to vary considerably with 200 mg of protein extracted from 100 g of orange peel compared to only 20 mg extracted from an equivalent of blackcurrant fruit (10-fold lower). The amount of protein and biocatalyst available per 100 g of fruit tissue was an important consideration when sourcing the enzyme. A comparison of total thioesterase and carboxylesterase activity toward MFTA and *p*-NA is illustrated in figure 6.3. Orange peel demonstrated the highest total thioesterase and carboxylesterase activities, 3109 nkat and 1140 nkat respectively, making it the optimal source of the biocatalyst. The orange juicing industry is large ($2.2 billion US), and generates 2.3 million metric tonnes of juice per year[183]. Therefore peel by-product is likely to be readily available as a cheap feedstock. Oranges are predominantly grown in Spain (Valencia), US (Florida), Brazil and Spain so a year round supply is available. However, prior to optimizing the use of this peel biocatalyst and testing its productivity it was of interest to determine the identity of the enzyme(s) responsible for the observed activities.

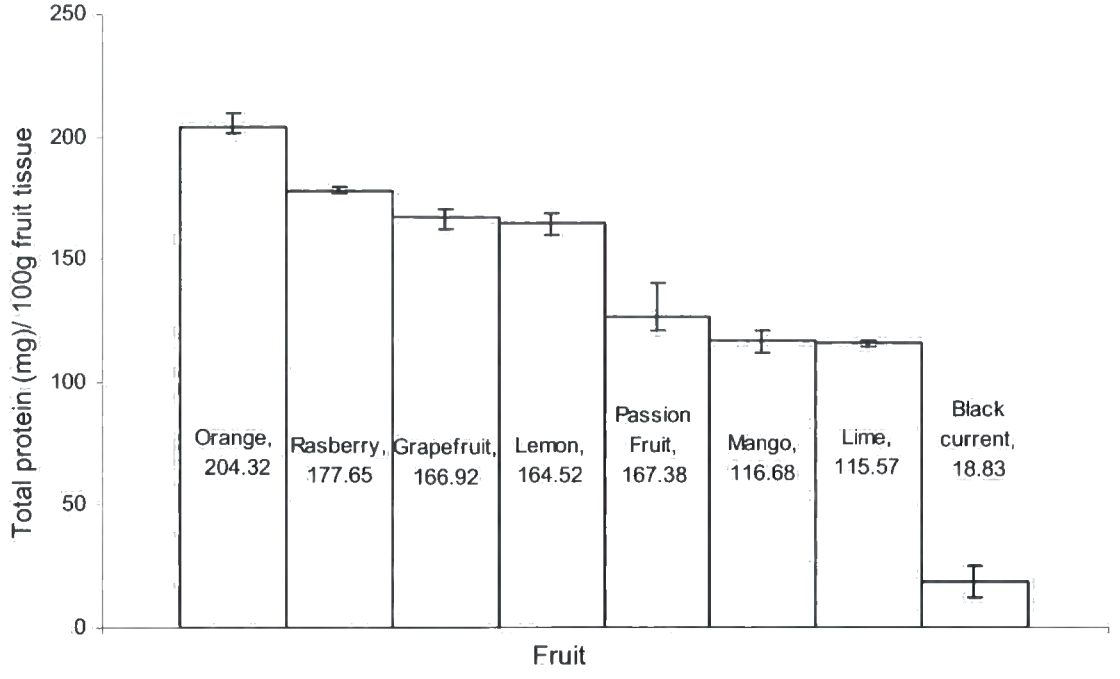


Figure 6.2 Protein extracted from 100 g of fruit tissue following 0.25 M salt extraction, filtration through Mira Cloth and 40-80% ammonium sulphate precipitation. Results are average protein determinations (triplicates) from extracts of multiple fruits.

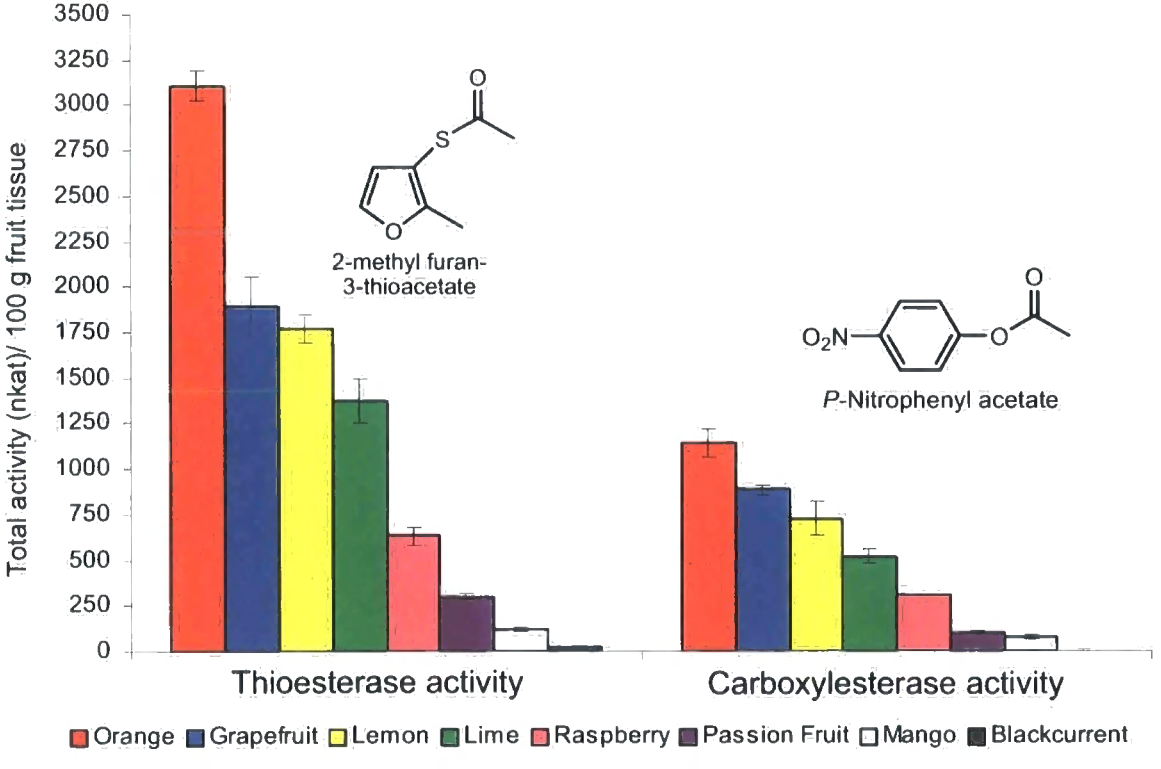


Figure 6.3 A comparison of total thioesterase and carboxylesterase activity in 100 g of fruit tissue. Through combining both specific activities and protein extracted the total activities could be determined. This serves as a more informative measure of biocatalytic productivity per weight of fruit material.

6.3 Purification and Identification of Thioesterases in Orange (*Citrus sinensis*)

The pattern of chemical preference determined for the orange extract mirrored that of the purple passion fruit sample (Figure 6.1), but whether this was due to the action of related PAE enzymes was unclear. Hence, a sequential protein purification protocol similar to that used for isolating the passion fruit enzyme, *Pe*WH, was employed as a means of determining the identity of the enzyme(s).

6.3.1 The Sequential Purification of Thioesterases from Orange Peel

The peel of four ripe Valencia oranges (200 g) was homogenized and protein extracted in an identical manner to that used for the large scale extraction of protein from passion fruit (Methods 2.4.1). The dialysed extract was then passed through three sepharose columns, DEAE, octyl and phenyl sepharose with the elution of MFTA-thioesterase activity determined using the Ellman's assay (Figure 6.4). The purification protocol to this stage was identical to that used for the isolation of *Pe*WH and hence the properties of the orange thioesterases on each column could be compared to that of *Pe*WH (Chapter 4, Figure 4.3). Columns were run under the same parameters of flow rate, fraction size and buffers.

The crude protein preparation was firstly applied to DEAE sepharose where thioesterase activity eluted in two distinct pools, an unbound and a bound pool referred to as pools 1 and 2 respectively (Figure 6.4 A1). The unbound pool which eluted over multiple fractions in the void volume accounted for approximately 85% of recovered activity. The smaller bound pool of activity was more acidic and eluted with 0.08 M NaCl. This separation of activity into two pools based upon differences in isoelectric properties was also observed during purification of *Pe*WH from passion fruit, where activity was separated equally between two pools, with pool 2 showing the greater specific activity. However, in *Citrus* the opposite was noted, pool 1 had a greater specific activity (22.69 nkat/mg) than that of 2 (7.64 nkat/mg). For this reason work primarily concentrated on characterizing pool 1 activity.

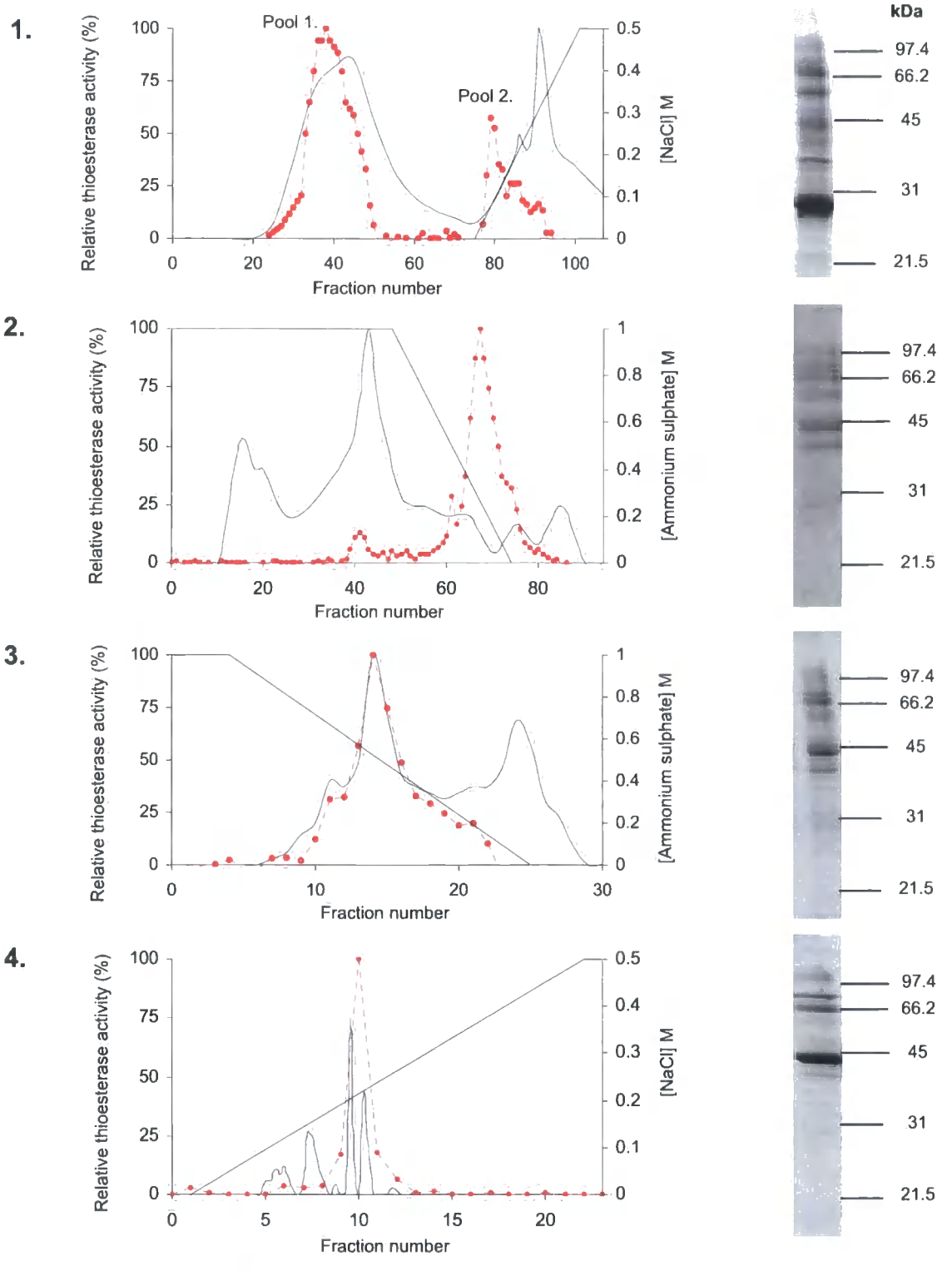


Figure 6.4 A. The sequential purification of Pool 1 thioesterase activity from Valencia orange peel by chromatography on (**1.**) DEAE Sepharose, (**2.**) Octyl Sepharose, (**3.**) Phenyl Superose and (**4.**) Mono S Sepharose. Blue lines show the elution of UV absorbing protein (280 nm), red dotted lines indicate thioesterase activity and black lines represent the gradient of mobile phase. Analysis of polypeptides at each stage of purification via SDS-PAGE (right). The abundant protein of size 28 kDa in the DEAE pool 1 fraction (gel top right) was excised and identified through MALDI-Tof based proteomics as a germin like protein, an abundant allergen found in orange (data not shown).

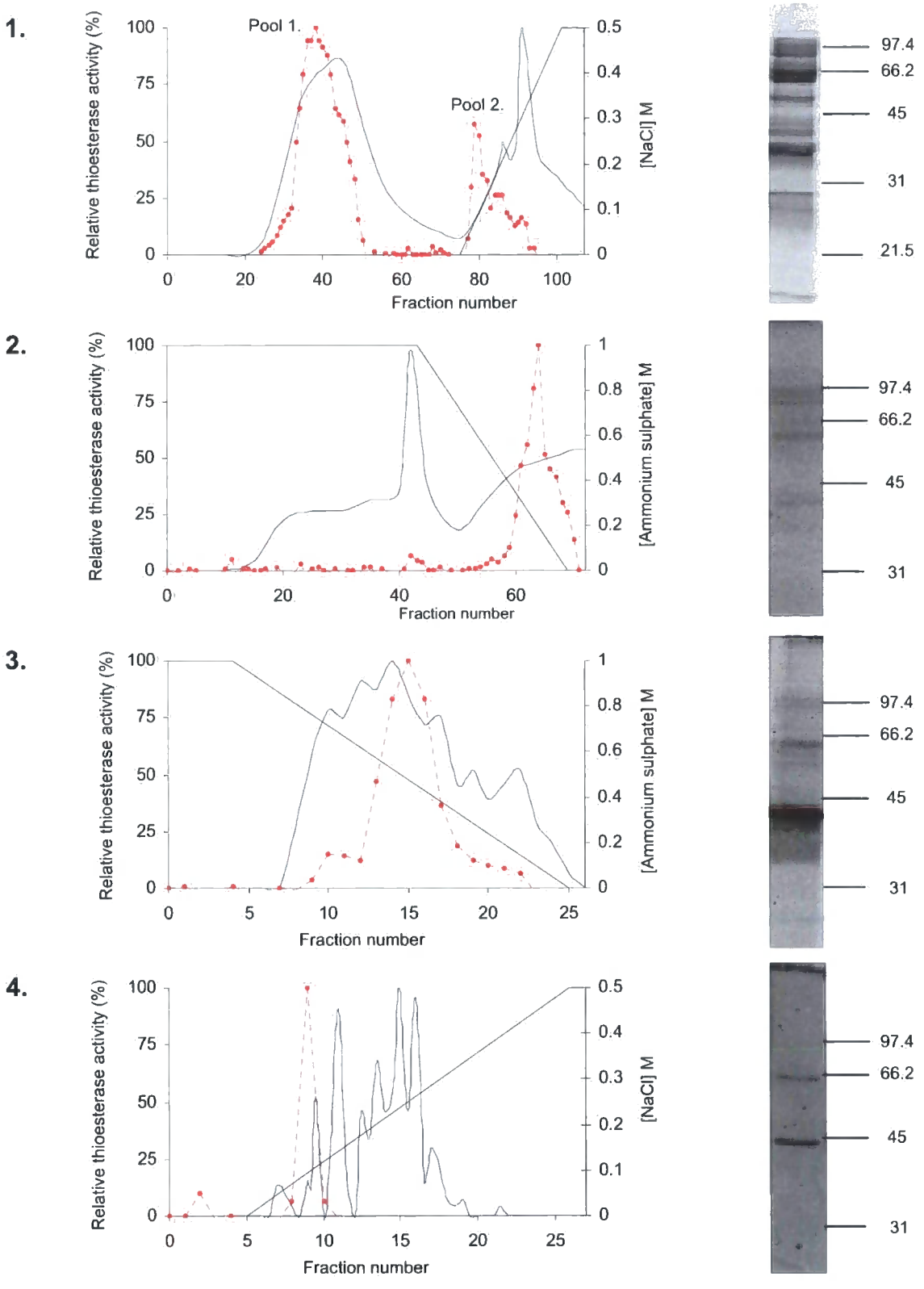


Figure 6.4 B. The sequential purification of Pool 2 thioesterase activity from Valencia orange peel by chromatography on (**1.**) DEAE Sepharose, (**2.**) Octyl Sepharose, (**3.**) Phenyl Superose and (**4.**) Mono S Sepharose. Blue lines show the elution of UV absorbing protein (280 nm), red dotted lines indicate thioesterase activity and black lines represent the gradient of mobile phase. Analysis of polypeptides at each stage of purification via SDS-PAGE (right).

Ammonium sulphate (1 M) was added to both pools of activity before they were separately applied onto the octyl sepharose column. Both activities demonstrated similar elution characteristics being recovered in 0.27 M and 0.19 M ammonium sulphate respectively (Figure 6.4 A2 + B2); an indication they were both relatively hydrophobic proteins. It was necessary to concentrate the protein solution to 10 ml by ultrafiltration prior to loading onto the phenyl superose column. Both pools of activity again demonstrated similar properties on this high-performance column with activity eluting in 0.52 M and 0.48 M ammonium sulphate respectively (Figure 6.4 A3 + B3).

The analysis of polypeptides eluting from the phenyl superose column (Figure 6.4, gel images) indicated the protein preparations were not adequately pure as multiple polypeptide bands were present. Hence, a further purification step was required. Due to the basic nature of the homologous thioesterase purified from passion fruit (*Pe*WH, pI: 9-10), it was decided that a cation exchange column composed of Mono S sepharose may retain the enzyme(s). Protein samples were therefore dialysed overnight against 20 mM Bis-Tris, pH 7.0 before loading onto the high-performance Mono S column (1 ml) at a flow rate of 1 ml/min with 1 ml fractions collected. Both pools of activity bound to the column with the slightly less basic pool 2 activity eluting with 0.10 M NaCl and pool 1 eluting with 0.19 M NaCl (Figure 6.4 A4 + B4). Based on the final specific activity (403 nkat mg^{-1}), the pool 1 enzyme had been purified 85-fold in 1% yield, whereas, pool 2 with a specific activity of 127 nkat mg^{-1}, was purified 27-fold in 0.3 % yield. A summary of the purification procedure is depicted in table 6.1 where the purification fold and recovery of activity is shown for both pools of enzymes at each subsequent purification step.

Preparation step		Vol (ml)	Protein concentration (mg/ml)	Total protein (mg)	Specific activity (nkat mg^{-1})	Total activity (nkat)	Purification (-fold)	Yield (%)
Crude		1300	0.54	702	4.74	3327	-	-
Crude + NaCl		1365	0.61	833	12.77	10637	2.69	100
40-80%		250	1.66	415	11.33	4702	2.39	44
DEAE	1.	212	0.83	176	22.69	3993	4.79	38
	2.	128	0.69	89	7.64	682	1.61	6.4
Octyl	1.	112	0.34	38	70.97	2697	14.97	25
	2.	72	0.27	19.5	22.94	448	4.83	4.2
P.Sup	1.	14	0.23	3.22	281.17	905	59.32	9.0
	2.	3	1.02	3.04	83.64	255	17.65	2.4
Mono S	1.	1	0.32	0.32	403.08	129	85.04	1.2
	2.	1	0.25	0.25	127.36	31.9	26.87	0.3

Table 6.1 Summary of the purification of MFTA thioesterase activity from Valencia orange. The peel of 4 fruit (200 g) was processed to provide a crude extract. Following four rounds of purification two separate pools of activity were resolved with 85- and 27-fold purification respectively.

6.3.2 The Identification of Thioesterases in Oranges as Pectin Acetylesterases (PAEs)

SDS-PAGE was used for the analysis of purified peptides eluting from the final Mono S columns (Figure 6.5). In both cases a 43 kDa polypeptide was identified whose relative abundance mirrored eluting thioesterase activity. The two pools of purified thioesterases from orange were named *Cs*WHa (*Citrus sinensis* wall-bound hydrolase a) and *Cs*WHb. The associated polypeptides were excised from the SDS gels, digested with trypsin and analysed by MALDI-ToF based proteomics and tandem mass spectrometry (MS-MS sequencing).

The MALDI analysis of *Cs*WHa generated a clear spectrum (Figure 6.6) with multiple ions identified above the background level. The ion fragmentation pattern was similar to that of *Cs*WHb and through comparing the peak ion lists it was apparent *Cs*WHa and b were the same protein (Annex III: *Cs*WH Protein Analysis), with the differences in their overall charge properties likely to be accounted for by subtle post-translational modifications. The calibrated ion peak list for *Cs*WHa was then used to screen the non-redundant protein and EST databases using Mascot software (www.matrixscience.com). Mascot takes the mass spectrometry data for a digested protein (peak ion list) and searches the protein databases for sequences with similar (or exact) peptide mass fingerprints. Such methods are however limited to the identification of proteins of known sequence stored within the databases. A search of the plant non-redundant protein and EST (expressed sequence tag) databases was conducted using experimental mass values between 1000-3500 (i.e. removing predominantly trypsin ions) with a low peptide tolerance of 20 ppm (providing a high level of sensitivity). No significant matches were found in the non-redundant protein database, however significant hits were found in the plant EST database (Figure 6.7A). The best match was toward a cDNA EST clone form sweet orange (accession: BQ624005), Ridge pineapple cultivar (*Citrus sinesis* L. Osbeck), with a significant Mowse (**Mo**lecular **w**eight **se**arch) score of 125. The translated protein sequence gave a predicted tryptic fragmentation of 14 peptides, 12 of which had masses identical to the experimental mass values (Figure 6.7B), accounting for 77% coverage of the sequence.

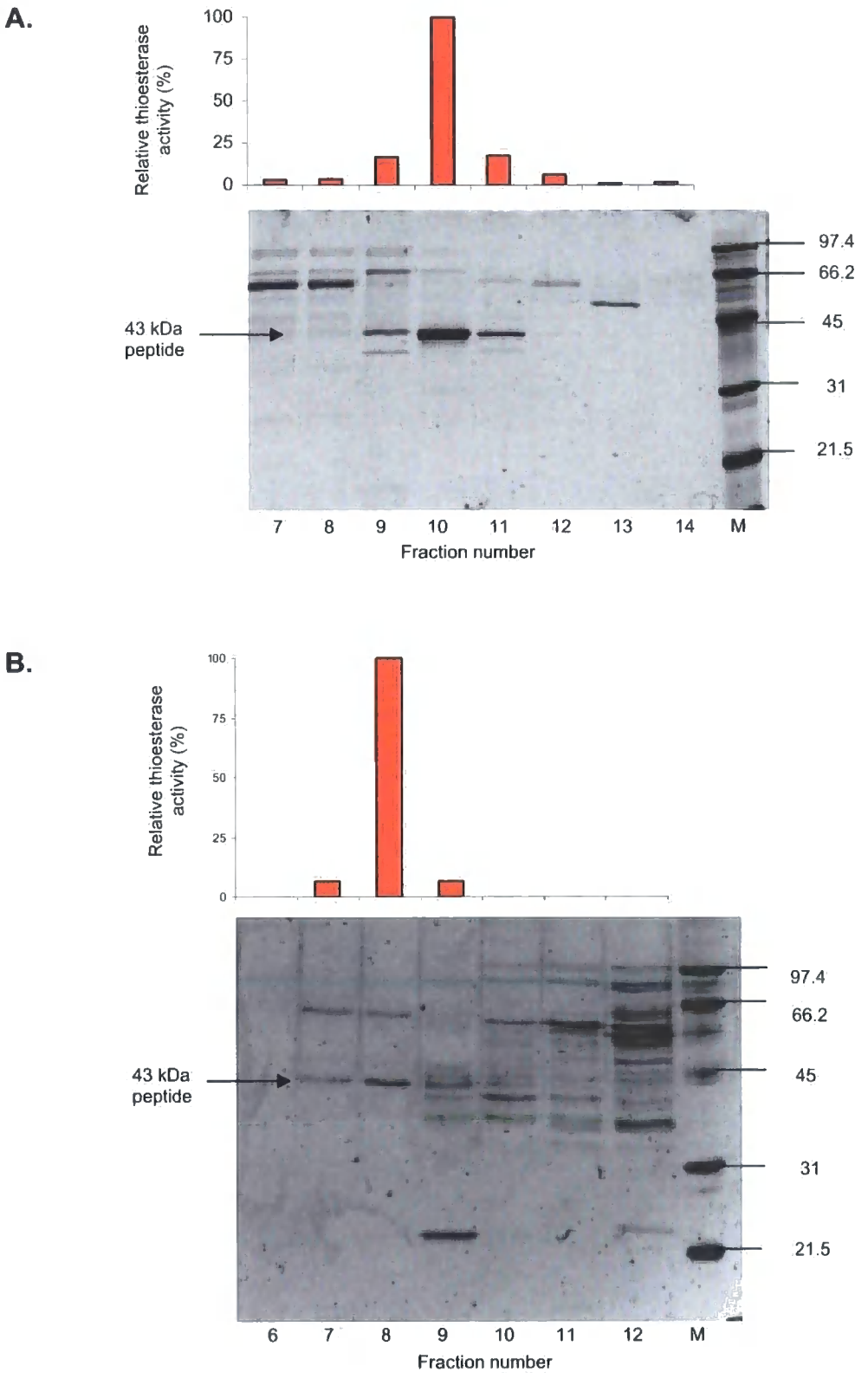


Figure 6.5 SDS-PAGE identification of active polypeptides in pool 1 (**A.**) and pool 2 (**B.**). The active fractions eluting from the final Mono S columns were analysed via SDS-PAGE and when aligned with relative thioesterase activity a polypeptide of 43 kDa mirrored eluting activity in each case, as highlighted by the arrows above.

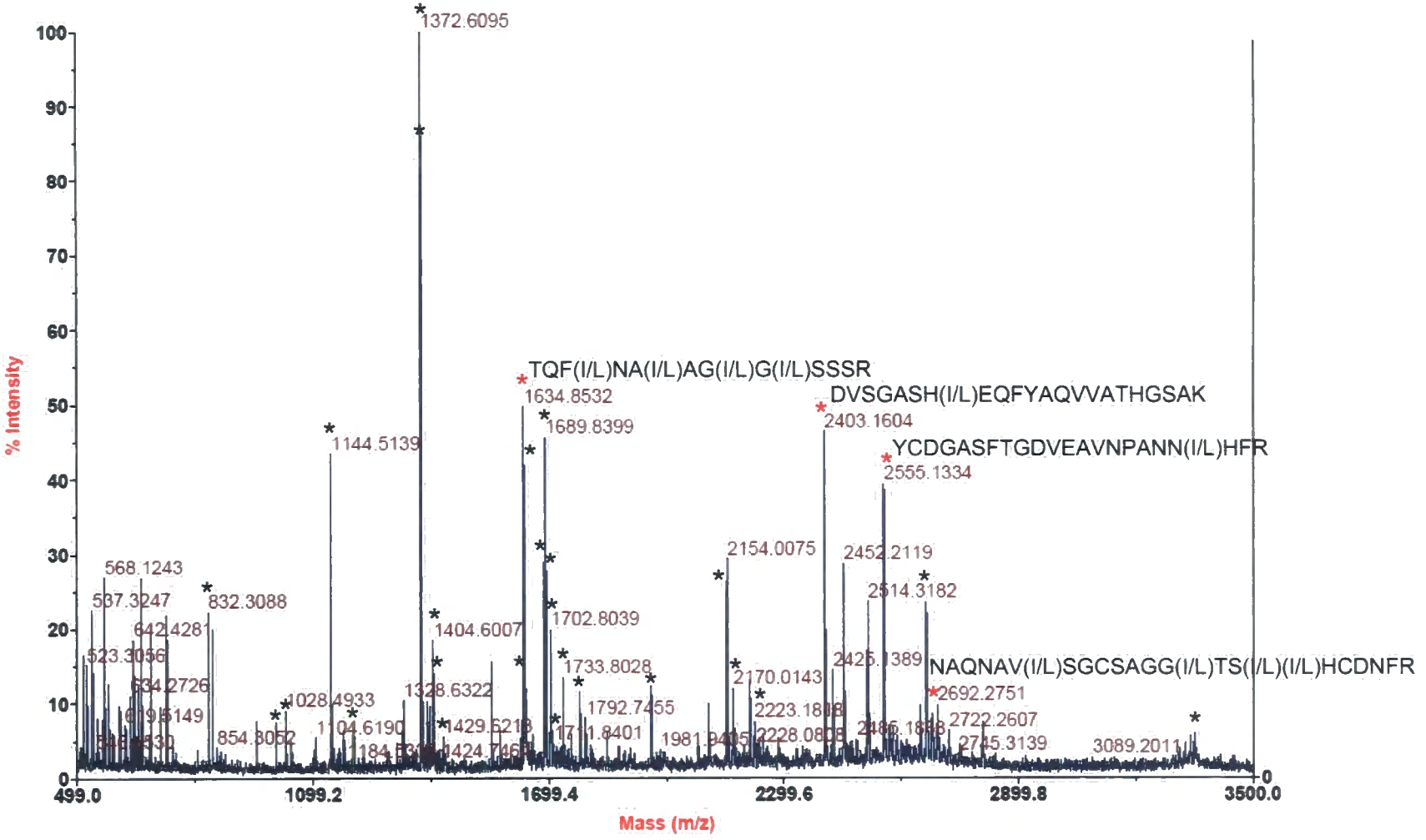
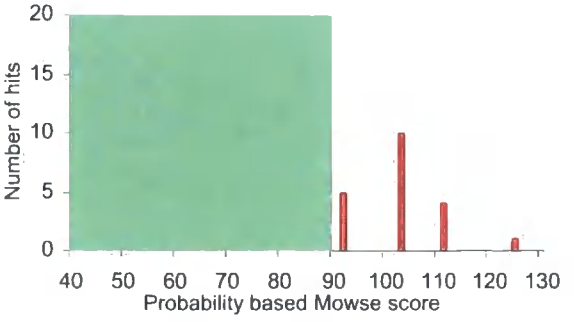


Figure 6.6 MALDI ion spectra for *Cs*WHa. The ions matched through Mascot searches (*) and those sequenced through tandem mass spectrometry (*) are highlighted. Where determined the amino acid sequence is shown.

A.



1. gi21651174 Mass: 22565 Score: 125 Expect: 1.8e-5 Queries matched: 12
USDA-FP_01096 Ridge pineapple sweet orange, *Citrus sinensis* cDNA clone, mRNA
2. gi63074511 Mass: 19907 Score: 113 Expect: 2.8e-4 Queries matched: 11
C18008E01Rv Drought2 *Citrus reshni* cDNA clone C18008E01, mRNA
3. gi42623376 Mass: 20918 Score: 112 Expect: 3.6e-4 Queries matched: 11
USDA-FP_5969 Ridge pineapple sweet orange, *Citrus sinensis* cDNA clone, mRNA

B. Top hit BQ624005: Translated EST clone

1 TKVKCFADAG YFINAKDVSG ASHIEQFYAQ VVATHGSAKH LPASCTSRLS
51 PGLCFFPQYM ARQITTPLFI INAAYDSWQI KNILAPGVAD PHGTWHSCKL
101 DINNCSPTQL QTMQSFRTQF LNALAGLGIS SSRGMFIDAC YAHCQTEMQE
151 TWLRTDSPVL GKMSIAKAVG DWYYDRSPFQ KIDCAYPCNP TCHNRGF

Peptide residue No.	Experimental mass (Mr)	Expected mass (Mr)	Error (ppm)	Sequence
5-16	1376.6320	1375.6247	1	CFADAGYFINAK
17-39	2402.1590	2401.1517	-6	DVSGASHIEQFYAQVVATHGSAK
40-48	1028.4900	1027.4827	-4	HLPASCTSR
49-62	1686.8100	1685.8027	-2	LSPGLCFFPQYMAR
49-62	1702.8120	1701.8047	2	LSPGLCFFPQYMAR oxidation M
82-99	1959.9480	1958.9407	-1	NILAPGVADPHGTWHSCK
100-117	2153.0080	2152.0007	-2	LDINNCSPTQLQTMQSFR
100-117	2169.0110	2168.0037	2	LDINNCSPTQLQTMQSFR oxidation M
118-133	1634.8560	1633.8487	-18	TQFLNALAGLGISSSR
134-154	2647.1510	2646.1437	12	GMFIDACYAHCQTEMQETWLR
168-176	1144.5180	1143.5107	11	AVGDWYYDR
182-195	1777.7230	1776.7157	2	IDCAYPCNPTCHNR

C.

Accession	Identification	Species	Predicted Mr/pI	Score/ E value
CAA67728	Pectin acetylesterase	*Vigna radiata*	43.3 / 9.5	327 / 3e-89
NP193677	Pectin acetylesterase	*Arabidopsis thaliana*	41.8 / 9.0	305 / 1e-82
AAP72959	Pectin acetylesterase	*Lactuca sativa*	43.7 / 8.76	295 / 2e-79

Figure 6.7 A. Mascot search results for *Cs*WHa. The graph illustrates the number of hits on the plant EST database and their significance based upon Mowse scores (-10*log(p), p: probability the observed match is a random event). Mowse score > 100 is commonly accepted as significant. Top hits are listed. **B.** The translated EST sequence BQ624005, illustrating predicted trypsin cleavage sites (underlined) and peptide mass matches accounting for 77% sequence coverage (Red). **C.** The top 3 BLAST hits when searching the plant protein databases with translated BQ624005.

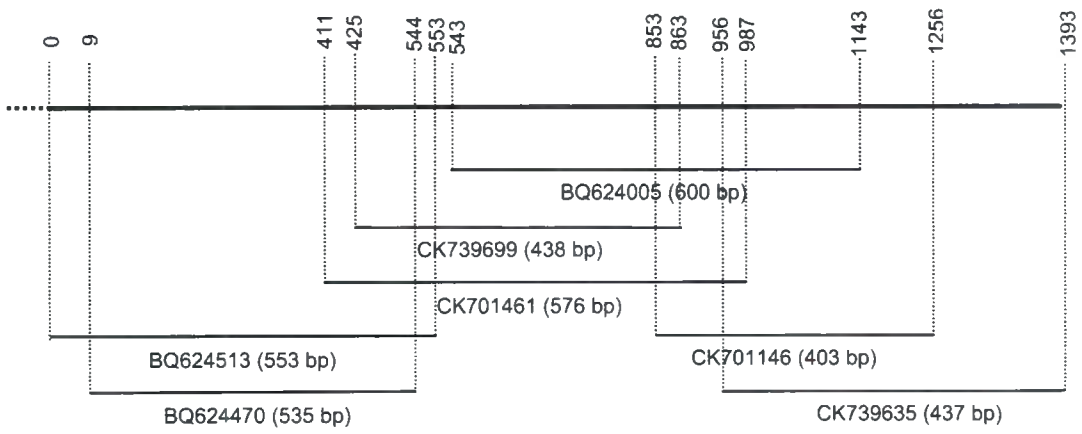
Two further matches in the plant EST database were also found to be significant, however one was from the closely related mandarin fruit (*Citrus reshni*) and the other was also from ridge pineapple sweet orange. Interestingly, Valencia orange is a sweet variety and the combination of fruit match and high Mowse score indicated that the top hit, BQ624005, was significant.

The EST clone for BQ624005 was translated into all six possible reading frames before screening against the plant protein databases using BLAST (Figure 6.7C). The top hit with a score of 349 was toward the pectin acetylesterase identified in mung bean (*Vigna radiata*, accession: CAA67728). Hence, the purified protein *Cs*WH gave a similar ion fragmentation pattern following trypsin digestion and MALDI to that predicted for an EST clone encoding a pectin acetylesterase from orange. It appears the purified thioesterases from orange are of the same family as those isolated from passion fruit, members of the pectin acetylesterase (PAE) protein family.

BQ624005 is a partial cDNA sequence and does not encode the full length PAE open reading frame. Therefore, the short sequence was used to search the EST nucleotide database for further overlapping clones from the same fruit as a means of building up a near complete translated protein sequence (Figure 6.8A). This larger and more representative sequence was subjected to a predicted MS-tryptic digest using the prospector software (www.prospector.ucsf.edu). The predicted ion peak list was compared to the experimental ion peak list for *Cs*WHa (Annex III B), where 20 matches were found accounting for 76% coverage of this larger sequence (Figure 6.8B). The unidentified sequence may be due to a number of factors, namely the incomplete digestion of the protein caused by its unusually hydrophobic nature, or the fact that this is a homologous protein from a related species where the true protein sequence although similar, may differ in exact amino acid content which will result in differences in peptide fragment sizes.

The identification of the protein as a PAE was supported by tandem mass spectrometry analysis. The parent ions 1635, 2402, 2554 and 2662 were all of

A.



B.

```
1    ESFXYVHSXX LLKADGFNVG ITYVENAVVK GAVCLDGSPP AYHFDKGFGA GINNWLVHIE GGGWCNNVTT CLERKKTRLG SSKQMVKVVA FSGMLSNKQK
101  FNPDFYNWNR IKVRYCDGAS FTGDVEAVNP ANNLHFRGAR VFQAVMEDLM AKGMKNAQNA VLSGCSAGGL TSILHCDNFR ALFPVGTKVK CFADAGYFIN
201  AKDVSGASHI EQFYAQVVAT HGSAKHLPAS CTSRLSPGLC FFPQYMARQI TTPLFIINAA YDSWQIKNIL APGVADPHGT WHSCKLDINN CSPTQLQTMQ
301  SFRTQFLNAL AGLGISSSRG MFIDACYAHC QTEMQETWLR TDSPVLGKMS IAKAVGDWYY DRSPFQKIDC AYPCNPTCHN RVFDSNVHSE V      391
```

Figure 6.8 A. EST sequence overlap used to build a translated open reading frame for *Citrus sinesis* L. Osbeck PAE. This larger translated sequence was predicted to account for 99 % coverage of the full protein. **B.** Translated *Citrus sinesis* L. Osbeck PAE sequence. Illustrating predicted trypsin cleavage sites and highlighting sequence coverage (bold) corresponding to identified ion fragments (76%). Predicted pI: 8.5, Mr: 43 kDa.

sufficient intensity to provide clear secondary fragmentation and could be used to determine partial protein sequence (Annex III C, Table 6.2). Each sequence was used to search the plant protein and EST databases using protein BLAST (short nearly exact match), with each peptide matching pectin acetylesterase proteins (Table 6.2).

Peptide fragment	Database hit (Score / E value)	Identification	Species Accession
1. Mr: 1635 TQF(I/L)NA(I/L)AG (I/L)G(I/L)SSSR (16 amino acids)	QFLNALAGLGLSSS QFL+AL+GLG S+S QFLDALSGLGNSTS (29.9/ 1.6)	Pectin acetylesterase	*Litchi chinensis* ACF05806
2. Mr: 2402 ASH(I/L)EQFYAQVV ATHGS (16 amino acids)	ASHIEQFYAQVVATHGS A HIE FY +VVATHGS AQHIEAFYNEVVATHGS (39.7/ 0.002)	Pectin acetylesterase	*Litchi chinensis* ACF05806
3. Mr: 2555 DGASFTGDVEAVNP ANN(I/L)H (19 amino acids)	DGASFTGDVEAVNPANNLH DGASFTGDVEAVNPA NLH DGASFTGDVEAVNPATNLH (57.5/ $8e^{-09}$)	Pectin acetylesterase,	*Litchi chinensis* ACF05806
4. Mr: 2663 QNAV(I/L)SGCSAG G(I/L)TS(I/L)(I/L)HCD NFR (23 amino acids)	QNAVLSGCSAGGLTSLLHCDNFR +NAVLSGCSAGGL SL+HCD+FR ENAVLSGCSAGGLASLMHCDSFR (62.1/ $3e^{-10}$)	Pectin acetylesterase, putative	*Arabidopsis thaliana* NP974575

Table 6.2 Sequencing of the purified thioesterase *Cs*WH. Database hits obtained from MS-MS sequencing of polypeptides. For each peptide the top BLAST hits are presented for known or putative proteins. Both score and expect values are shown to highlight the significance of the match. Score: Higher the better, measure of how close the match is, taking into account miss matches (space) or similar amino acid types (+). Expect: Lower the better, the chance such an alignment could occur by chance in the database being searched.

These two independent methods of protein identification, MALDI analysis and MS-MS sequencing, both identified the purified thioesterase as a PAE like hydrolase. Furthermore, the two analytical techniques can be combined to ensure both the peptide mass finger print and the sequence of selected ions match (Figure 6.6), as a thorough means of confirming protein identity.

6.3.3 The Characterisation of *Cs*WH Thioesterase

When characterizing the purified enzyme particular attention focused on determining the properties of the protein with respect to its use in bioprocessing, such as stability, chemical activity and kinetic rates. The enzyme was screened against the commercial hydrolases (CRL and PLE) and the purified thioesterase from passion fruit (*Pe*WH) for activity against the panel of thioester and ester substrates (Table 6.3). The purified orange enzyme had a high thioesterase activity toward **2** (MFTA; 403.08 nkat mg^{-1}), was 2.7-fold more active than *Pe*WH and had similar activity to PLE (537.88 nkat mg^{-1}). The enzymes pattern of chemical preference was similar to that of *Pe*WH, but generally 5-10 fold greater in activity. Partially purified *Cs*WH showed preference for short chain acyl components in esters/ thioesters as demonstrated through the reduction in activity caused by the increasing acyl chain length (C2, C3, C6) in the umbelliferyl ester series **9-11**.

The stability of the enzyme at 30°C in 0.1 M phosphate buffer, pH 7.2 was investigated and found to decay steadily over 14 days, with a half life of 7 days 22 hours. Hence, *Cs*WH was found to be a particularly stable protein, even more so than *Pe*WH which had a half life of 4 days 14 hours under comparable conditions.

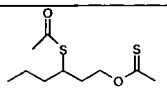
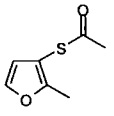
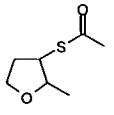
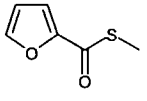
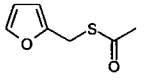
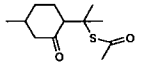
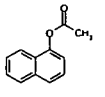
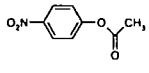
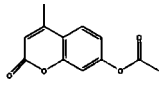
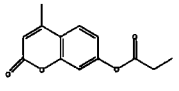
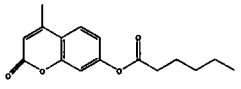
Compound number	Thioester/ ester substrates	Purified *Pe*WH thioesterase (S.a. nkat/mg)	Purified *Cs*WH thioesterase (S.a. nkat/mg)	Porcine liver esterase (S.a. nkat/mg)	*Candida rugosa* lipase (S.a. nkat/mg)
1.	 3-(thioacetyl) hexylacetate	1.12 +/- 0.02	3.01 +/- 0.12	14.02 +/- 0.72	ND
2.	 2-methyl furan-3-thioacetate	149.38 +/-0.0021	403.08 +/- 18.37	537.88 +/- 24.65	0.35 +/- 0.02
3.	 Tetra hydro 2-methyl furan-3-thioacetate	0.29 +/- 0.03	1.08 +/- 0.09	31.12 +/- 3.01	ND
4.	 Methyl thiofuroate	ND	ND	100.0 +/- 9.75	ND
5.	 Furfuryl thioacetate	10.06 +/-	43.62 +/- 2.13	35.04 +/- 1.98	0.05 +/- 0.01
6.	 8-(acetylthio) menthone	8.84 +/- 0.08	6.29 +/- 0.17	4.17 +/- 0.02	0.02 +/- 0.01
7.	 α-Naphthyl acetate	4.16 +/- 0.72	42.11 +/- 1.98	159.4 +/- 16.14	0.58 +/- 0.01
8.	 P-Nitrophenyl acetate	11.39 +/- 1.2	135.36 +/- 11.03	2117.6 +/- 87.5	5.14 +/- 0.23
9.	 4-Methylumbelliferyl acetate	0.66 +/- 0.08	1.48 +/- 0.19	426.15 +/- 72.55	4.26 +/- 0.18
10.	 4-Methylumbelliferyl propanoate	0.57 +/- 0.09	0.93 +/- 0.04	1589.26 +/- 102	118.36 +/- 21.39
11.	 4-Methylumbelliferyl hexanoate	0.006 +/- 0.0004	0.01 +/- 0.002	1302.97 +/- 97	36.59 +/- 4.00

Table 6.3 Enzymatic hydrolysis of thioester and carboxylester substrates by *Cs*WH as compared to the thioesterase from passion fruit, *Pe*WH, the mammalian esterase (PLE) and a microbial lipase (CRL). Activities are means of triplicates +/- standard deviation from the mean. ND; no activity detected.

The enzyme kinetics of *Cs*WH were also of interest, both as a means of investigating the enzymatic mechanism of the protein and for determining its productivity (e.g. turnover number). The effect of substrate concentration on reaction velocity was therefore investigated using the model substrate MFTA. The variation of reaction velocity with substrate concentration followed classic Michaelis-Menten kinetics (Figure 6.9) as confirmed through the linear double reciprocal plot (Lineweaver-Burk plot, Figure 6.9B). The kinetic constants for *Cs*WH were determined as;

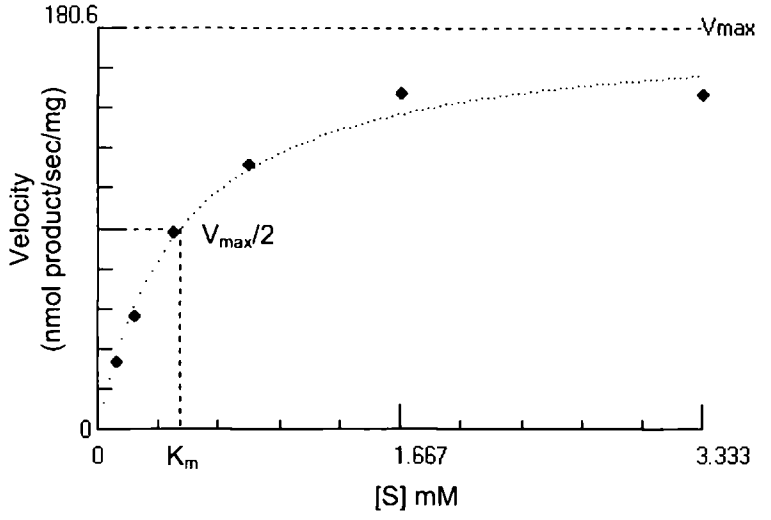
V_{Max}: 181 +/- 26 nmol/ sec/mg protein
K_m: 0.5 mM MFTA
K_{cat}: 7.85 nmol product/sec/nmol protein

The turnover number (K_{cat}: 7.85 Sec^{-1}) was comparable to other plant enzymes involved in the metabolism of xenobiotic compounds, such as an esterase from the weed, black-grass (*Alopecurus myosuroides*), which bioactivates aryloxyphenoxypropionate herbicides[143]. Although turnover number is not high enough for VOSCs to be considered the natural substrate of *Cs*WH, it is nonetheless significantly productive for consideration in bioprocessing.

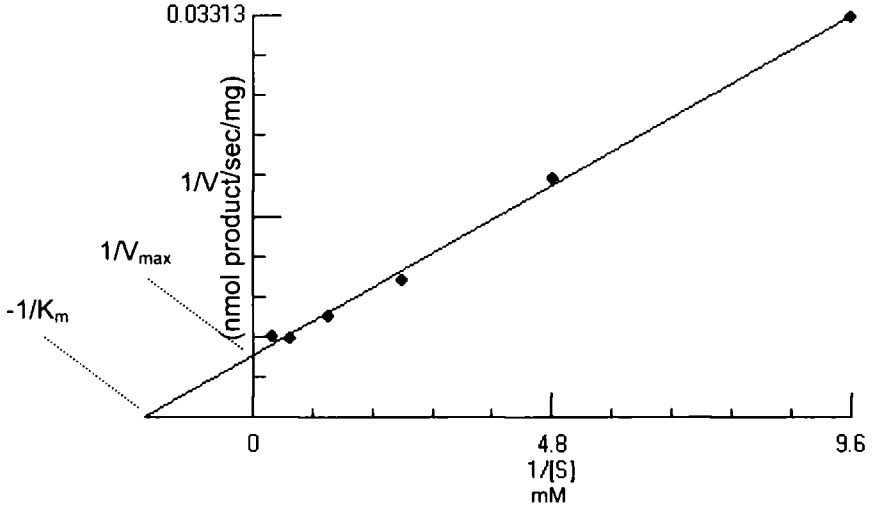
Using the turnover number for purified *Cs*WH, it is possible to predict the amount of thiol product (MFT) that can be generated per kg of peel, per day, per 1 litre reaction (space/ time/ yield value).
400g peel: 6680 nmol product/ sec
400g peel: 48 g/ day/ L
1kg peel: 120 g product/ day/ L

Due to the high-impact nature of the product only small volumes need be produced, the predicted productivity of this biocatalyst is theoretically attractive. However, through optimising the extraction and processing of the enzyme yields could be improved.

A.



B.



V_{max}: 180.6 +/- 26.07 nmol product/sec/mg protein

K_m: 0.4548 +/- 0.2047 mM MFTA

K_{cat}: 7.85 +/- 1.14 nmol product/sec/nmol protein

Figure 6.9 Enzyme kinetics of *Cs*WH thioesterase. The dependence of reaction velocity (V) on substrate concentration (MFTA) follows a classic Michaelis-Menton curve (**A.**). Through plotting a double reciprocal plot the kinetic constants V_{max} and K_m could be determined and the turnover number K_{cat} defined (**B.**). 1.5 µg purified protein (final column) was incubated with varying substrate concentrations 0.007-3.3 mM MFTA at 30°C and product formation determined over 1 min using the Ellman's assay. Values shown are averages of triplicates with correction for chemical rates.

6.4 The Enzymatic Production of High-Impact Aroma Chemicals Using Orange Peel

Due to the high acetylesterase content of *Citrus* peel there has previously been interest in optimising the material as a novel source of hydrolases[182,184]. Partially purified acetylesterase from *Citrus sinensis* demonstrated regioselectivity in hydrolysing the acetyl group from the C6 position of penta-*O*-acetyl-α-D-glucose[184], a precursor in agrochemical synthesis (Figure 6.10A). Complete hydrolysis of the ester at position 6 occurred within one hour, where longer incubation times did not give rise to further deacetylation products. A recent study looked to optimize the acetyl esterases from *Citrus sinensis* through partial purification (190-fold) and immobilisation of the enzyme on Sepa-beads-EP (Resindion, Milano, Italy)[182]. The immobilized enzyme removed the acetyl group at the 3 position of β-lactam antibiotics, such as cephalosporin C, with > 98% conversion and subsequent product yields of 91-93% (Figure 6.10B).

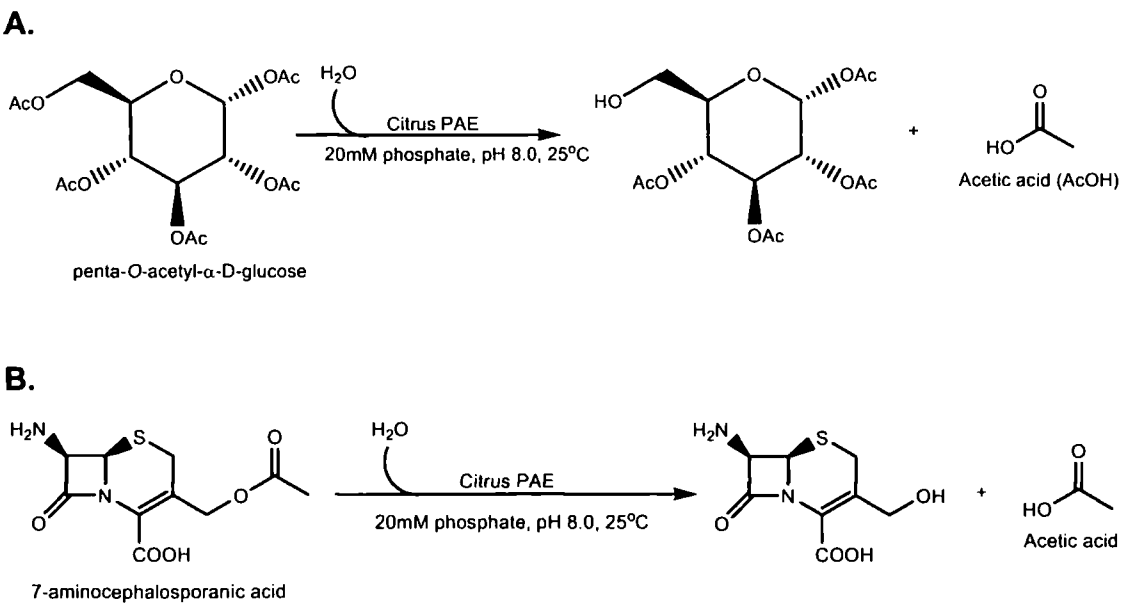


Figure 6.10 Citrus acetylesterase catalysed hydrolysis of penta-*O*-acetyl-α-D-glucose (**A.**) and 7-aminocephalosporanic acid, a precursor in cephalosporin C synthesis (**B.**). Figure produced with information from Pasta *et al.*,[182] 2004 and Waldmann *et al.*,[184] 1994.

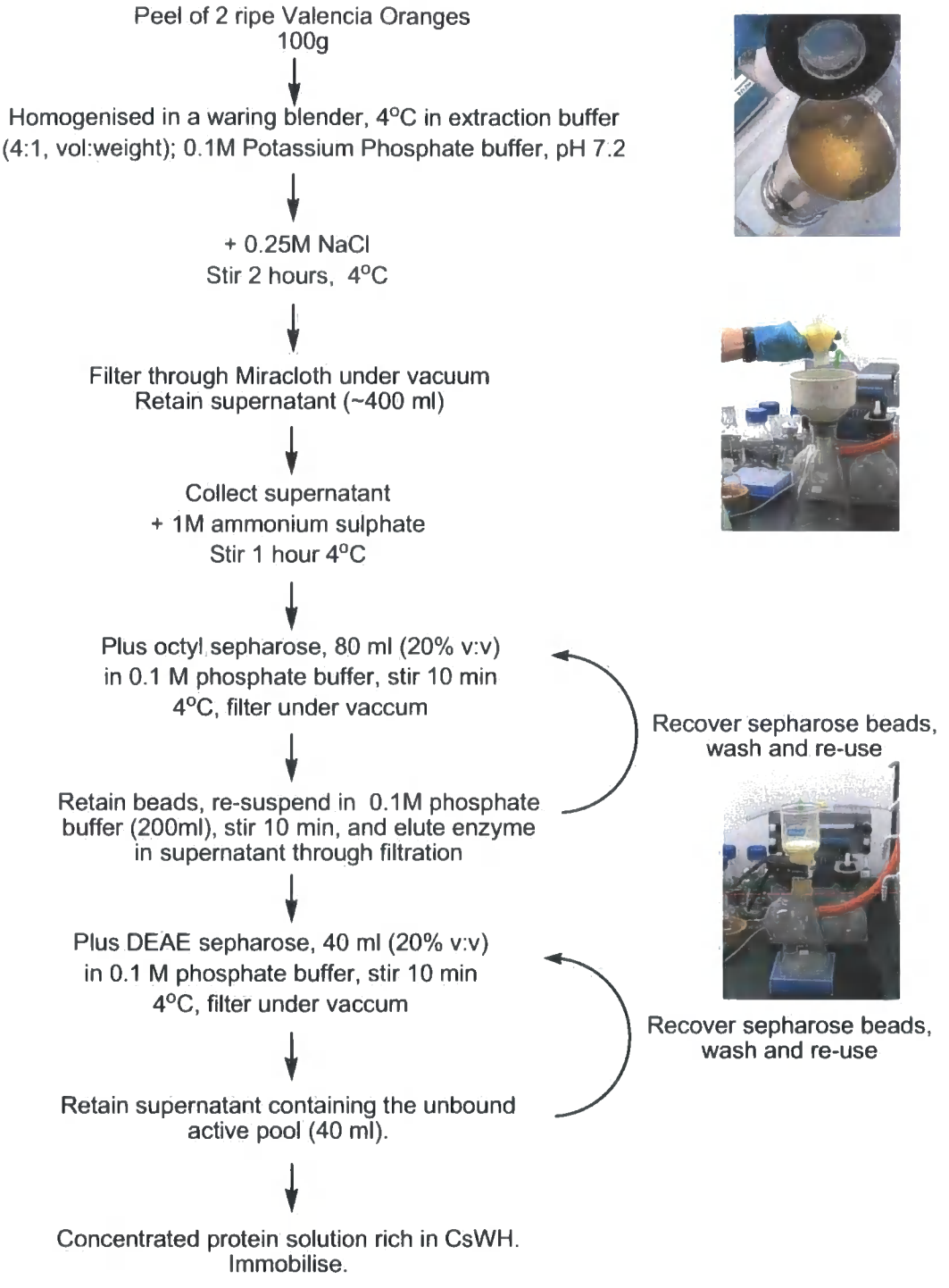
These studies are consistent with our observations with *Cs*WH and *Pe*WH, which demonstrated the enzymes' preference for acetylated substrates. There is a clear application for PAEs and PMEs in the regioselective deprotection of blocking groups on intermediates used in drug or agrochemical synthesis. Furthermore, enzymatic hydrolysis is conducted under mild conditions and prevents the decomposition of delicate intermediates which may otherwise occur through acid or base catalysis[103]. However, the application of *Cs* peel has to date not been considered for the biotransformation of VOSCs.

6.4.1 Optimisation of *Cs* Biocatalyst

When developing a process for optimizing a biocatalyst one must weigh up the pros and cons of each strategy. For example, the purification and immobilisation of an enzyme may provide greater stability, recovery and prevent unwanted side reactions, however, this can be costly and a readily available raw material may suffice for use as a crude preparation. Although both strategies, purification/immobilisation and crude preparation, were considered when developing the use of *Cs* peel much of the optimization was made through trial and error.

Previously the extraction and immobilisation of acetylesterase from *Cs* peel[182] was found to provide a selective esterase which could be recovered and re-used in subsequent reactions. However, with each stage of processing enzyme activity will inevitably be lost (25% in this case). Our own efforts to partially purify *Cs*WH through batch purification and filtration, as would be used for large scale protein extraction industrially, failed to provide adequate yields of the enzyme (Figure 6.11).

An alternative approach was required. Our studies with *Cs*WH and *Pe*WH had previously indicated that the enzymes were wall-bound proteins associated with the polysaccharide matrix. Hence, instead of using salt to liberate the hydrolases prior to purification and immobilisation it was potentially simpler (and cheaper) to use the peel directly with the biocatalyst postulated to be partially immobile.



Sample	Total activity (nkat/mg)	Purification (x fold)	Yield (%)
Crude salt extract	5120	-	-
Octyl purified	1690	2.9	33
DEAE purified	760	8.7	15

Figure 6.11 Batch purification of *Cs*WH. This simple purification procedure was used to mirror a large scale industrial purification process. Filtration replaced centrifugation and affinity purification of the protein was conducted in single step batches rather than through columns. The recovery of activity and fold purification of thioesterase activity toward MFTA is shown (bottom).

However, using fresh peel material meant the shelf life had to be considered, and in an effort to prevent bacterial contamination, *Cs* peel was freeze dried. The lypholyzed form of the *Cs*WH present in the peel preparation was found to be significantly stable for several months at -20°C (3 months: activity > 90%). Furthermore, drying the peel meant it could be powdered, greatly increasing the surface area for substrate-catalyst interactions in comparison to the fresh homogenised peel.

Further optimisation of the drying process included homogenising the peel in different buffers at varying pHs prior to freeze drying. It has been demonstrated previously that enzymes have the remarkable property of retaining active site conformation following drying, even when used in different solvents[185]. This property is described as "memory" and was investigated here to determine whether improved activities could be achieved, especially when optimizing activity in non-aqueous solvents or biphasic reactions. However, in this case it was found that simply drying the fresh homogenized peel without buffer was optimal for retaining activity (91 +/- 3 % retained following drying).

A final consideration was recovery of biocatalytic activity between reactions. Although it was postulated the enzyme was partially immobile, would this be adequate to recover significant activity for it to be re-used in subsequent reactions? As a means of determining this, 5 mg of dried peel was incubated (30°C) with 1 mM MFTA in 0.1 M phosphate buffer (1 ml reaction) for 1 minute and the specific activity determined through the Ellman's assay prior to recovery of *Cs* peel through centrifugation (5 min, 10,000 g). Recovered catalyst was resuspended in an identical reaction and activity again determined following 1 minute incubation. Repeated reactions and recovery illustrated that the enzyme was in fact partially bound, whereby activity was lost exponentially between subsequent reactions (Figure 6.12). However, a fraction remained tightly bound and accounted for a retention of 40-45% activity over 7 reactions. With the intention of enhancing the enzyme's affinity for the pectin matrix, fresh peel was homogenised and soaked in calcium (0.1 M) buffered solutions prior to freeze drying. It was postulated that Ca^{2+} ions

which are known to ionically bind to the -ve charges on pectin may act to chelate the enzyme to the matrix[186]. The addition of calcium however appeared to remove the partially bound fraction altogether (Figurer 6.12). This could be a consequence of calcium or chloride ions interfering with ionic binding to the matrix during homogenisation and leading to deactivation of the enzyme during freeze drying.

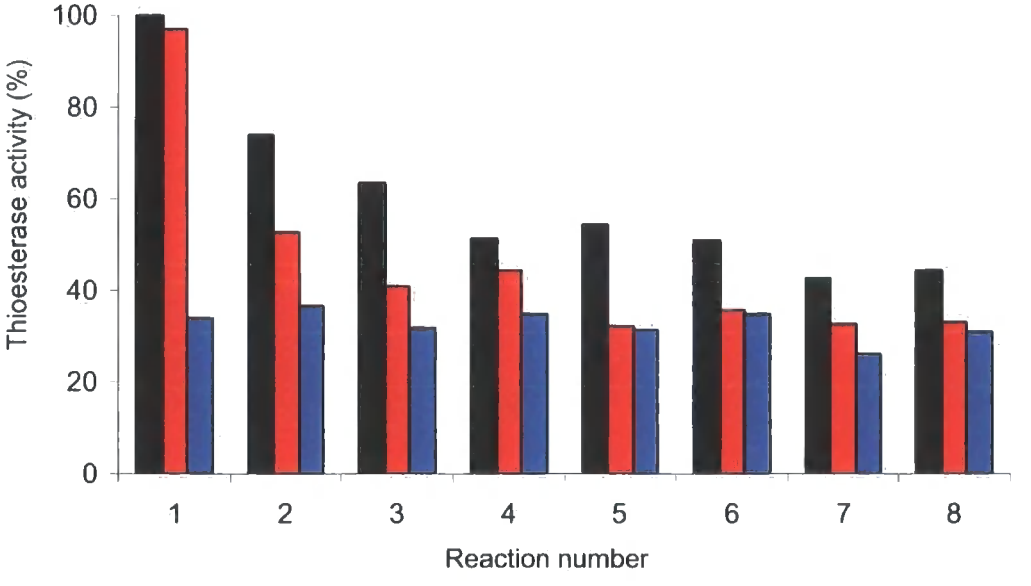


Figure 6.12 Recovery of *Cs* peel biocatalyst. 5 mg of dried peel (+/- Ca^{2+}) was incubated with 1mM MFTA in 0.1M phosphate buffer pH 7.2. Relative thioesterase activity toward MFTA was determined using the Ellman's assay prior to recovery of the catalyst through centrifugation and re-suspension in an identical reaction. Black bars; fresh dried peel material (no additives/ buffer), red bars; homogenisation in 0.1 M phosphate buffer pH 7.2 prior to drying and blue bars; homogenisation in 0.1 M phosphate buffer pH 7.2 + 0.1 M $CaCl_2$ prior to drying.

An interesting observation made when working on the optimization of *Cs* peel biocatalyst was that the water content of the reaction affected the properties of the catalytic material, with the pectin forming a gel at low water content (7% w/v). Furthermore, the gel could be set onto the surface of tubes and reaction vessels through rotating the container as the gel set. This may serve as an efficient system for conducting bioreactions on an industrial scale where substrates could be continually flowed over a flat bed of set gel. Through

warming the solution and increasing the water content the gelling process could be reversed. Such an approach may also prove useful for the preparation of further biocatalysts where they would essentially be retained in a gel between subsequent reactions and washes.

6.4.2 Productivity of *Cs* Biocatalyst

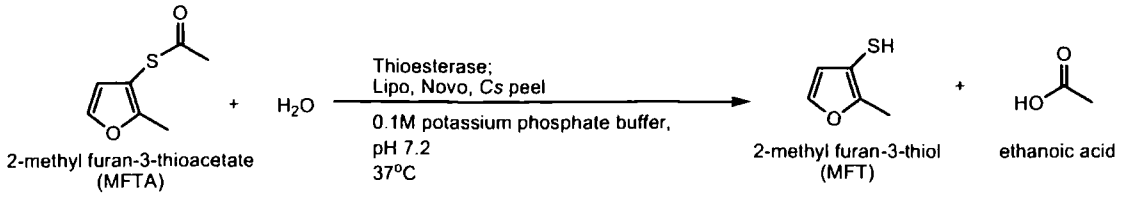
As a means of determining the potential use of *Cs* peel dried catalyst, it was important to assay the material against other commercially available biocatalysts. Two immobilised lipases were therefore chosen, Lipozyme (lipase from *Mucor miehei* immobilized on a macroporous ion-exchange resin) and Novozyme (lipase from *Candida antarctica* immobilized on acrylic resin) (Figure 6.13). Both optimized enzymes had been previously used for the biosynthesis of sulphur aroma chemicals[115]. Small lab scale bioreactions (50 ml) were set up in round bottomed flasks incubated at 37°C. Initially the thioesterase activity of each enzyme preparation (*Cs* peel, Lipozyme, Novozyme) was investigated where 100 mg of catalyst was incubated with each thioester substrate (10 mM) for 12 hours as a means of determining the hydrolytic conversion to thiol VOSC product (Figure 6.14). Cs peel was a superior biocatalyst toward MFTA, 3-(thioacetyl)hexylacetate and furfuryl thioacetate. Product yields of thiomenthone were however lower (47%).

Citrus sinesis; freeze dried Valencia peel

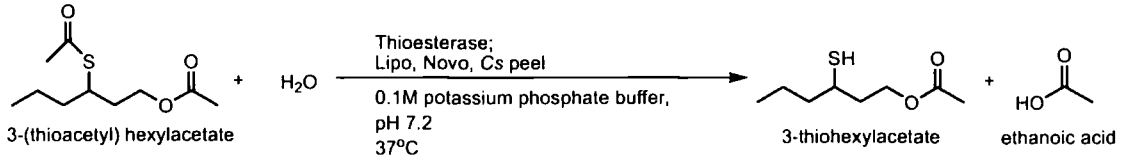
Lipozyme; immobilised lipase from *Mucor miehei*

Novozyme; immobilised lipase from *Candida antarctica*

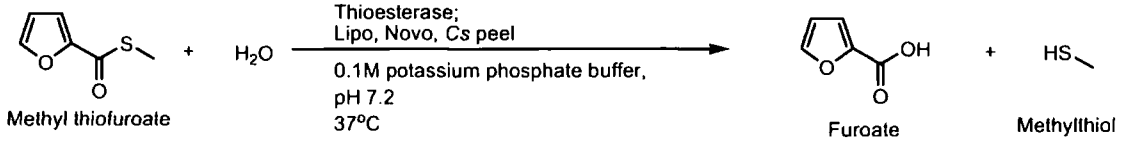
Figure 6.13 Biocatalysts used for the transformation of VOSCs.



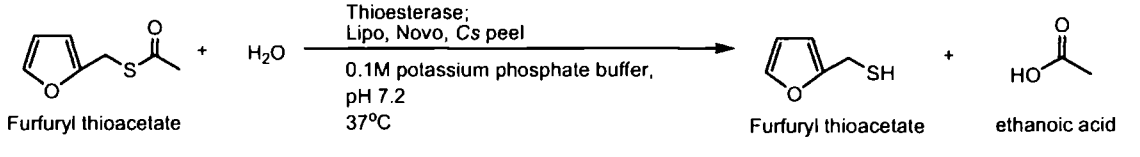
	Cs peel	Lipo	Novo
Product yield (%)	56	40	53



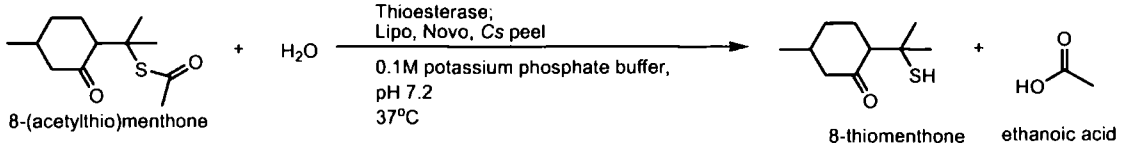
	Cs peel	Lipo	Novo
Product yield (%)	81	76	78



	Cs peel	Lipo	Novo
Product yield (%)	1	0.5	1



	Cs peel	Lipo	Novo
Product yield (%)	97	83	87



	Cs peel	Lipo	Novo
Product yield (%)	47	73	49

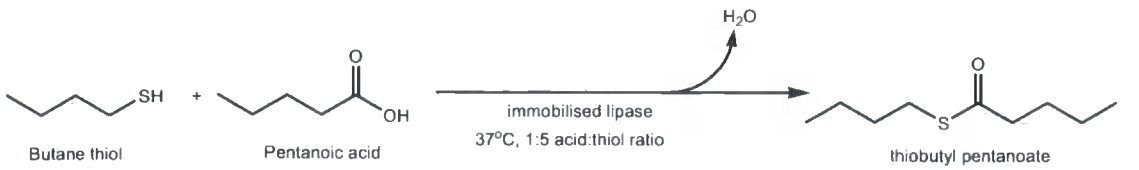
Figure 6.14 The hydrolysis of thioester VOSCs using *Citrus sinensis* (Cs) peel and the immobilised lipases, Lipozyme (Lipo) and Novozyme (Novo). 100 mg of catalyst was incubated with 10 mM substrate for 12 hours prior to quantification of thiol product using GC analysis.

A previous study had demonstrated that immobilised lipases in solvent free media could reverse the hydrolytic reaction and form thioesters from reaction media containing pure thiol and acid[115]. The immobilised lipases from *Mucor miehei* (Lipozyme) and *Candida antarctica* (Novozyme) were shown to produce short-chain thioesters in solvent-free medium (i.e. no water, only thiol and acid)[115]. A thioester yield of 40% was achieved for the condensation of pentanoic acid and butanethiol at 60°C, 1:5 (acid: thiol) ratio with the addition of molecular sieves to remove the water generated upon condensation. Although the ability to function in non-aqueous systems is characteristic of lipases it was interesting to re-run such bioreactions using the dried *Cs* peel.

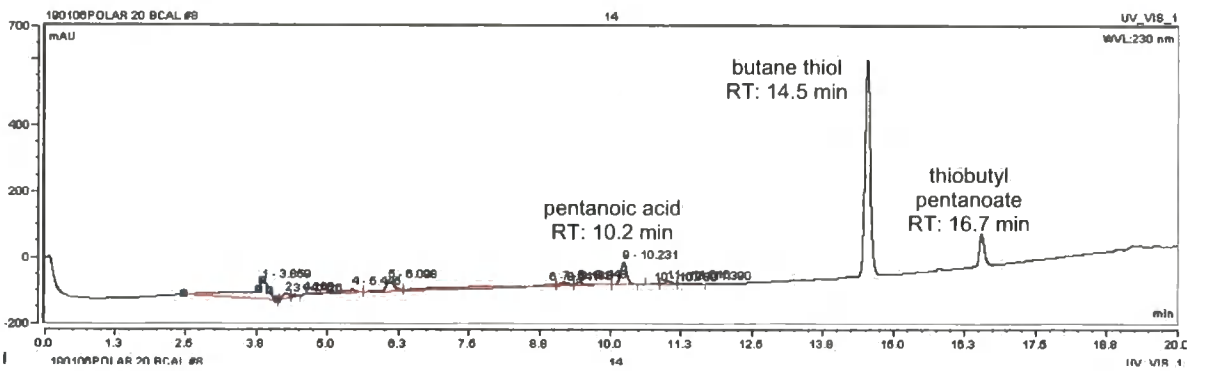
It was however important to ensure the dried peel enzyme could tolerate a non-aqueous environment. Therefore freeze dried peel (100 mg) was soaked in dichloromethane (DCM) with agitation for 1 hour before being recovered through centrifugation (10,000 g, 5 min) and thioesterase activity determined through the Ellman's assay. 35 % of activity remained after incubation in solvent (DCM) for 1 hour. Evidently a microscopic layer of water had remained bound to the enzymes following drying which enabled them to retain conformation and activity following contact with solvent. Although only 35% activity remained it was interesting to determine whether the dried enzyme could drive the natural hydrolytic reaction in reverse.

All bioreactions were run for 3 hours in round-bottomed flasks (50 ml) with 100 mg of each biocatalyst. Reaction products were analysed by HPLC (Methods 2.3.5) and quantified through calibration of peak area using standard compounds (Figure 6.15). Both the immobilised lipases, Lipozyme and Novozyme, generated equivalent amounts of thioester (18-20%) under such conditions. However, no product was observed when using *Cs* peel. Genetic modification of the enzyme active site may be considered when optimising this enzyme for non-aqueous environments.

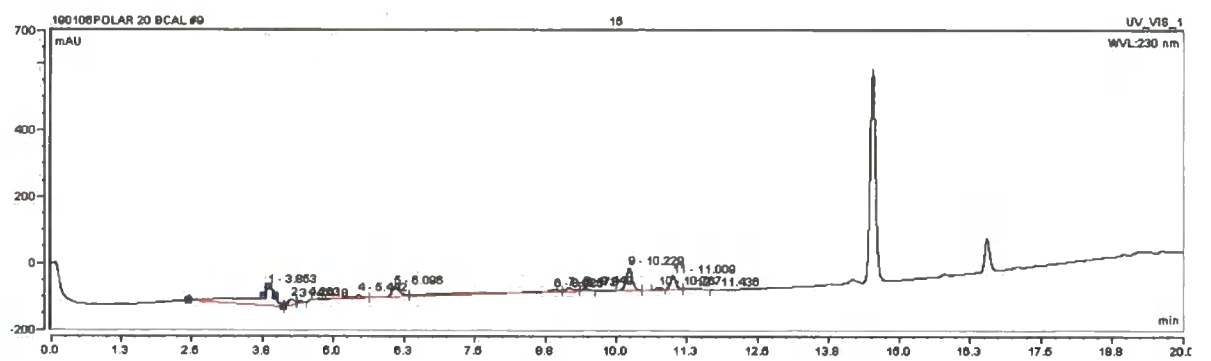
A.



B. Lipozyme; 19.4% product +/- 4.1



Novozyme; 18.2% product +/- 2.0



Cs peel; No product

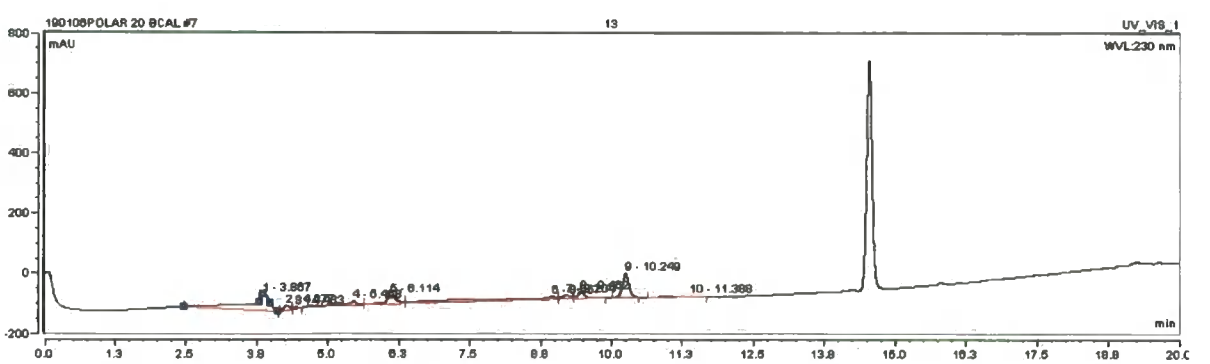


Figure 6.15 Reaction scheme illustrating the formation of thioesters in solvent free media using immobilised lipases (A.). HPLC traces showing reaction products following incubation of biocatalyst (100mg) with butane thiol and pentanoic acid (1:5 acid:thiol ratio, 50 ml reaction) for 3 hours at 37°C. A 10µl aliquot was diluted 1000 fold in DCM and dried prior to loading.

6.5 Discussion

Screening further fruits for thioesterase activity demonstrated that *Citrus*, and in particular orange, contained high levels of the desired activity and formed a useful source of thioesterases. The orange juicing industry is large ($2.2 billion US)[183], and produces quantities of waste peel which could be used as a source of biocatalyst.

Purification of MFTA thioesterase activity in *Citrus sinensis* enabled the identification of the enzyme which was found to be a member of the pectin acetylesterase family of carboxylester hydrolases. MALDI ToF proteomics combined with MS-MS sequencing provided convincing evidence the defined thioesterase activity was due to this family of carbohydrate binding esterases. Furthermore, the identification of *Cs*WH as a PAE supports the initial work on the thioesterase from passion fruit, *Pe*WH, where the enzyme was identified as a PAE homologue. It was not possible to use MALDI analysis to identify *Pe*WH as there were no matching fruit EST entries for this species within the non-redundant protein databases. However, the identification of the thioesterase *Cs*WH as a PAE in orange adds weight to the proposition the thioesterase *Pe*WH is a PAE homologue. The proteins were purified in the same manner against the same substrate and shared many properties (size: 43 kDa, pI: 9-9.5). Hence, it appears both fruit species contain wall bound acetylesterases with the ability to transform volatile organosulphur chemicals. As attempts to express putative PAEs in bacteria were unsuccessful, this was important in linking thioesterase activity in fruit to the proposed family of acetylesterase enzymes.

Freeze dried *Cs* peel was found to have many favourable properties (stability, recovery, turnover) for potential use in flavour and fragrance biosynthesis. This illustrates that biocatalysts can be sourced from dried plant materials, such as fruit, vegetables, leaves and root extracts which are abundant in certain activities. For example, freeze dried extracts from *Allium* species may provide a cheap and readily available source of CS-β-lyases. Although *Cs* peel has potential for further development as a biocatalyst, the stereoselective

properties of the enzyme remain to be determined. This is of particular interest for the bioproduction of flavourings which often have significantly different characteristics dependent upon optical form. Oxford Chemicals, with their GC analytical equipment for separating and identifying different enantiomeric forms of volatile flavourings plan to take this forward.

7. Discussion and Future Work

The current human population of 6.1 billion continues to rise (1.33% /annum), and is predicted to plateau at 11-12 billion over the coming 150-200 years[187]. This is putting an increasing pressure on natural resources and it has been proposed that the sustainability of the population has been reached[188]. However, economies strive to grow and frequently neglect the world around them through resculpting the landscape, releasing harmful gases into the atmosphere and toxins into rivers[187]. Overexploitation is threatening biodiversity and a change toward sustainable practices is required to reverse further damage. With large industrial companies the major polluters they have an urgent responsibility to act.

This work is a reflection of the change in direction being taken by today's chemical manufacturers, that of moving away from metal catalysts and non-renewable feedstock to more sustainable "clean" processes[189]. The increasing price of oil-derived feedstock makes such a transition a necessity[190]. In initiating this work for the UK flavourings manufacturer, Oxford Chemicals Ltd., we set out to develop a novel means of generating high-impact sulphur aroma chemicals through an enzymatic approach. A biocatalyst with thioesterase activity was chosen as a potentially versatile tool for transforming volatile organosulphur compounds (VOSCs). Based on their ability to act in a controlled stereospecific manner, there has previously been interest in identifying hydrolases that catalyse the cleavage of thioesters to release optically active VOSCs[113,114]. However, these studies investigated the use of microbial lipases and mammalian esterases where product yields and enatioselectivities were found to be inadequate for industrial manufacture. Little attention has focussed on the endogenous thioesterases (hydrolases) in plants, and in particular those from fruit, which we predicted would have evolved enhanced specificities in metabolising short-chain flavour volatiles. Exotic fruit, being unique in their ability to combine both sulphur metabolism and flavour biosynthesis for the generation of attractive and valuable VOSCs were therefore chosen in an initial screen for thioesterase enzymes.

7.1 Biochemical Considerations

When conducting this research it was postulated that VOSCs in tropical fruit may be sequestered as the less reactive thioester precursors. This was supported by the identification of thioester derivatives, including 3-(thioacetyl) hexylacetate identified in passion fruit[119] and 8-acetylthio-*p*-menthan-3-one which was reported in buchu plant extracts[120]. It was believed that, upon ripening, hydrolases would catalyse the liberation of VOSCs through thioester hydrolysis under developmental regulation.

Although such an activity was measured in purple passion fruit (*Passiflora edulis* Sims) it became apparent this was not accounted for by a specific thioesterase unique to tropical fruit. The purified enzyme, referred to as *Pe*WH, was a wall-bound hydrolase localised in the cortex (mesocarp) and was identified as a member of the pectinacetyl esterase family of enzymes (EC 3.1.1.6). This sub-family of α/β hydrolase fold proteins had been previously identified and partially characterised in plants due to their involvement in carbohydrate cell wall metabolism[158,150]. In addition, such hydrolases were found to be particularly abundant in the peel of fruits such as *Citrus*[167], with homologous proteins also identified in model plants, including rice and *Arabidopsis*. The purification and identification of a PAE with thioesterase activity (*Cs*WH) in orange (*Citrus sinensis*), a fruit which does not generate VOSCs, illustrates that this hydrolase if involved in flavour metabolism has a general hydrolytic function and is not specific to the liberation of VOSCs in this case. *Pe*WH did however demonstrate thioesterase activity toward 3-(thioacetyl)hexylacetate, the proposed aroma precursor in passion fruit, but whether this is the endogenous substrate for this enzyme is debatable. The activity determined for the partially purified protein was low (1.12 nkat/mg), 12-fold less than the mammalian esterase (PLE), but with greatest activity toward synthetic cyclic thioesters (e.g. MFTA, 149.38 nkat/mg) and carboxylester substrates (e.g. *p*-NA, 11.39 nkat/mg) it appears the enzyme is a somewhat versatile hydrolase.

This does not however discount the possibility that PAEs may serve a role in flavour liberation in fruit, and based on our observations with *Pe*WH and

*Cs*WH it is possible that pectin associated hydrolases present in fruit peel may serve a further role in regulating flavour release. Such a role in natural product metabolism may extend to many fruits. The endogenous functions of pectin acetylesterases are not well understood, though their activity in regulating the degree of pectin acetylation has been ascribed roles in softening fruit tissue[167] and in controlling cell growth and elongation[160]. Interestingly, in *Citrus* fruit peel, pectin acetylesterases were localised at high concentrations in oil vesicles where functions in wall modifications would seem improbable[167]. Furthermore, in apples, butyl esters are hydrolyzed to release butanol into the surrounding air by undefined esterases found at high concentrations in the cortex and peel[93]. In addition a recent study in the tropical snake fruit (*Salacca edulis*) found that pectin methylesterase activity was linked to the metabolism of volatile methyl esters[191]. The diversity of roles adopted by pectin hydrolases in fruit ripening[192,193] and their potential to act as bifunctional enzymes contributing to the biogenesis of volatile secondary metabolites is an interesting area for future study.

The fluorophosphonate suicide probe proved a powerful tool in confirming the presence of a serine hydrolase catalytic motif and may aid in characterizing the roles of further esterases involved in fruit maturation. The active site pockets of *Cs*WH and *Pe*WH appear constrained so as to accommodate predominantly short chain (C2) acetyl constituents as the acyl components in esters, and therefore defining the crystal structure of these plant proteins would be important in resolving the cause of such tight chemoselectivity. The abundance of acetylated (and methylated) flavourings in fruit[25] adds further weight to the involvement of cell wall hydrolases such as acetyl- and methylesterases in flavour metabolism. In ascertaining whether there is an endogenous role for PAEs in flavour and fragrance biosynthesis future studies may consider using crude enzyme preparations (or purified *Pe*WH/ *Cs*WH) in combination with the postulated substrates, pectin (methylated and/ or acetylated) and alcohol-CoAs. Such *in vitro* analysis of the effects of acetylesterases (or methylesterases) on the volatile spectrum generated by fruit may help in determining whether an endogenous role for these wall-bound hydrolases in flavour biosynthesis does in fact exist.

Volatiles in plants are also sequestered as glycosidically bound components, where the volatile aromas are released from the stable glycoconjugates through enzymatic hydrolysis[86,87]. PAEs and PMEs combine both hydrolase activity and an affinity for carbohydrate constituents, and hence it was postulated they may serve a role in releasing volatile secondary metabolites from wall bound components in the peel of fruit, either through direct hydrolysis or acyl- methyl-transferase activity (Figure 7.1A).

The potential involvement of *Pe*WH in such metabolism was investigated. The glycosidically bound fraction from the pulp of a ripe purple passion fruit was isolated through Amberlite XAD-2 adsorption and methanol elution, as described previously (Chassagne *et al.*, 1998)[88]. The extract was subsequently incubated with partially purified *Pe*WH (20 μg), (phenyl superose fraction) for 8 hours at 37°C (1 ml reaction volume). The volatile components liberated were partitioned into an equal volume of solvent (DCM), dried with sodium sulphate and concentrated 10-fold prior to GC analysis (Figure 7.1B). A comparison of the volatile spectra for enzyme assays and boiled controls indicated a slight perturbation in volatile generation whereby the addition of concentrated *Pe*WH led to an increase in the liberation of certain volatiles. However, in this case their identity could not be confirmed through GCMS as the sample was too complex. The reaction was also run with the addition of 3-mercaptohexan-1-ol (10 μM) as a means of investigating the potential transferase activity of *Pe*WH, however, no perturbation in the volatile spectrum was observed. Repeating such *in vitro* studies using specialised fibres for trapping volatiles in the headspace of sealed reaction vessels (e.g. solid phase microextration, SPME) in combination with a GCMS system set-up for such analysis would provide a more sensitive and quantitative assay. At the time of this work, our labs unfortunately did not have the analytical equipment for such studies. In addition, using activated substrates, such as acyl-CoA derivatives may be a requirement of the enzyme's transferase activity, however, this remains to be studied.

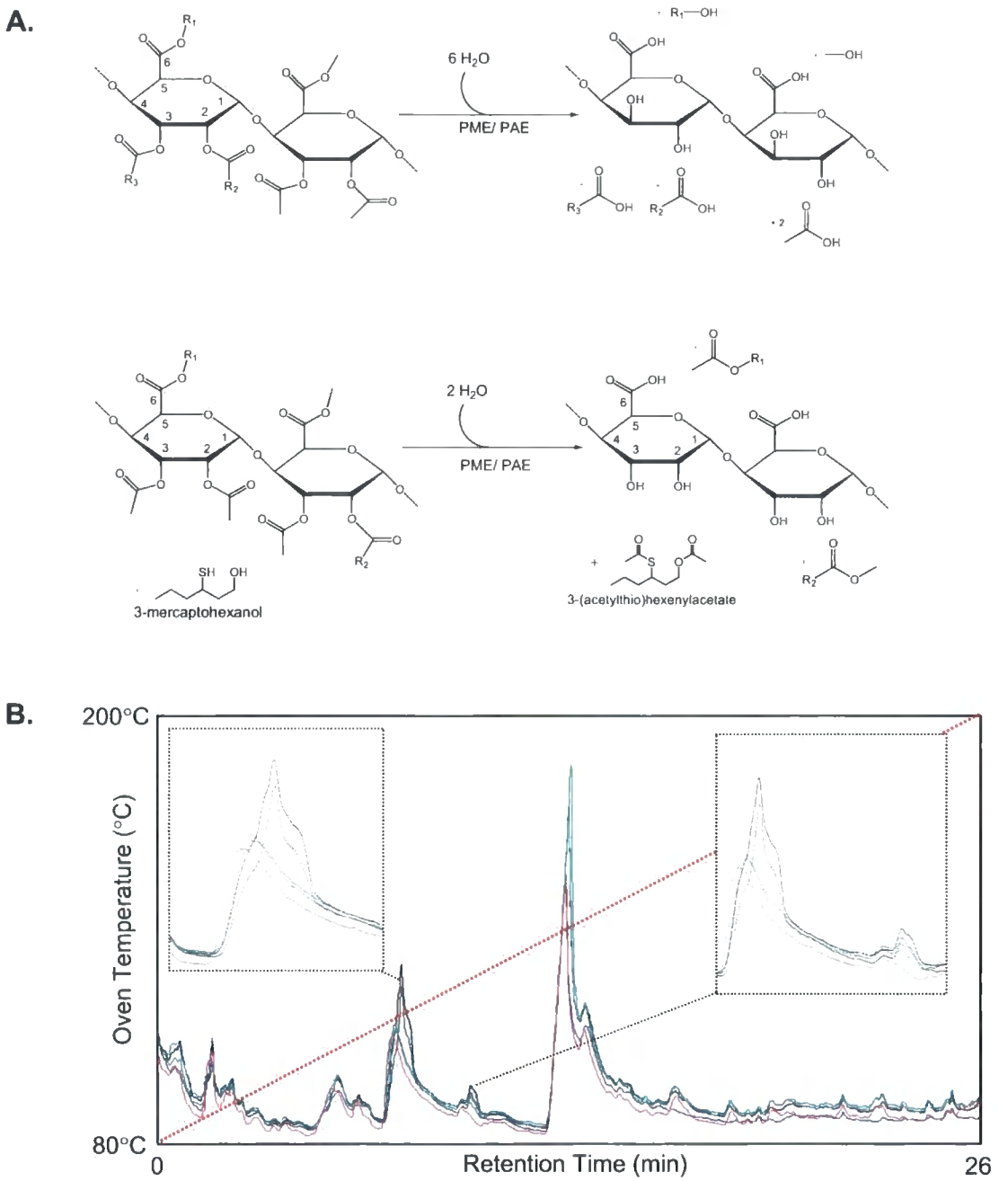


Figure 7.1 A A general model for the release of volatile flavourings from glycoconjugates (e.g. pectin, monosaccharides, disaccharides) by glycoside hydrolases. Top: The hydrolysis of side chain constituents generating volatile alcohols (C6 position) and acids (C2,3). Bottom: postulated acyl or methyl transfer between side chain constituents or free volatiles (e.g. 3-mercaptohexanol). **B.** The involvement of *Pe*WH in the liberation of volatile metabolites from glycoconjugates in passion fruit. 20 µg of partially purified *Pe*WH (phenyl superose fraction) in 0.1 M phosphate buffer, pH 7.2 was incubated with carbohydrate extracted from an equivalent of one passion fruit for 8 hours at 37°C (1 ml reaction volume). The GC spectrum of total volatiles following extraction into an equal volume of DCM (10-fold concentrate) is shown; blue and black lines show duplicate samples, green and turquoise lines show boiled controls and pink shows - enzyme.

The potential involvement of glycoside hydrolases in the release of flavourings from the cortex of fruit is depicted in figure 7.2 which illustrates a simplistic model for VOSC metabolism. Such cells have evolved to incorporate sulphur into volatile secondary metabolites through pathways that show similarity to plant xenobiotic detoxification processes[77], whereby glutathione is incorporated into short-chain electrophilic metabolites through the thiol group of cysteine. The incorporation of glutathione into xenobiotics, or in this case flavour precursors, renders them less toxic to the cell and facilitates their removal to the vacuole or cell wall[77]. The involvement of glutathione in the incorporation of sulphur into aromatic volatiles may be the origins of the detoxification system used by plants today as a means of tolerating synthetic xenobiotic chemicals, such as agrochemicals which were only developed during the 20th century.

7.2 Technological Applications

Although the endogenous roles of *Pe*WH and *Cs*WH in fruit aroma biosynthesis remain to be resolved, it is important to note that both enzymes are robust and have the desired catalytic function to be exploited as biocatalysts. The greater specific activities toward VOSC thioester precursors demonstrated by *Cs*WH made this the enzyme of choice for optimizing as a biocatalyst. Freeze-drying the peel of the fruit as a crude enzymatic mix was found to be the optimal source of the biocatalyst, where hydrolase content was stable ($t_{1/2}$: 7 days 22 hours, 30 °C) with an adequate turnover number toward VOSC thioesters (e.g. MFTA, K_{cat}: 7.8 S^{-1}). This material provides a promising and inexpensive biocatalyst for development by the industry in the future and highlights that biocatalysts from plants need not be sourced in purified form when dried materials may suffice.

An application for dried orange peel would be in the generation of optically active thiol and alcohol flavourings. Currently thiol aromas are industrially generated through the hydrolysis of synthetic thioester precursors, produced through the addition of sulphinic acid to unsaturated alkenes (Figure 7.3).

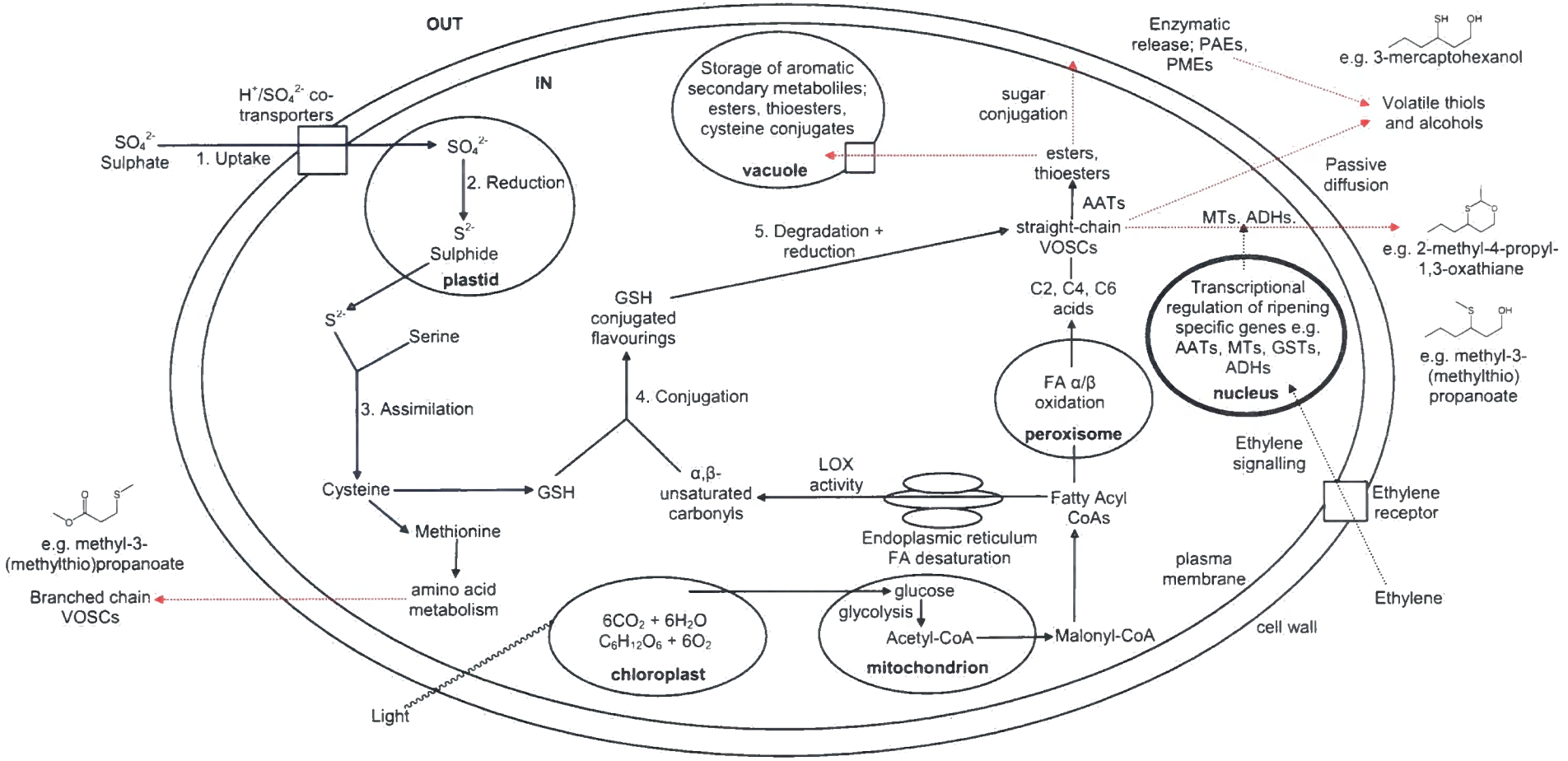


Figure 7.2 A hypothetical model for the generation of volatile organosulphur compounds (VOSCs) in tropical fruit. Metabolism illustrates sulphur uptake, reduction and assimilation (blue lines), the generation of core VOSC precursors (black arrows), the storage and release of diversified products (red dotted arrows) and the regulation of ripening specific genes through ethylene signalling (black dotted arrows). This is a simplistic model based upon defined sulphur metabolic pathways in plants.

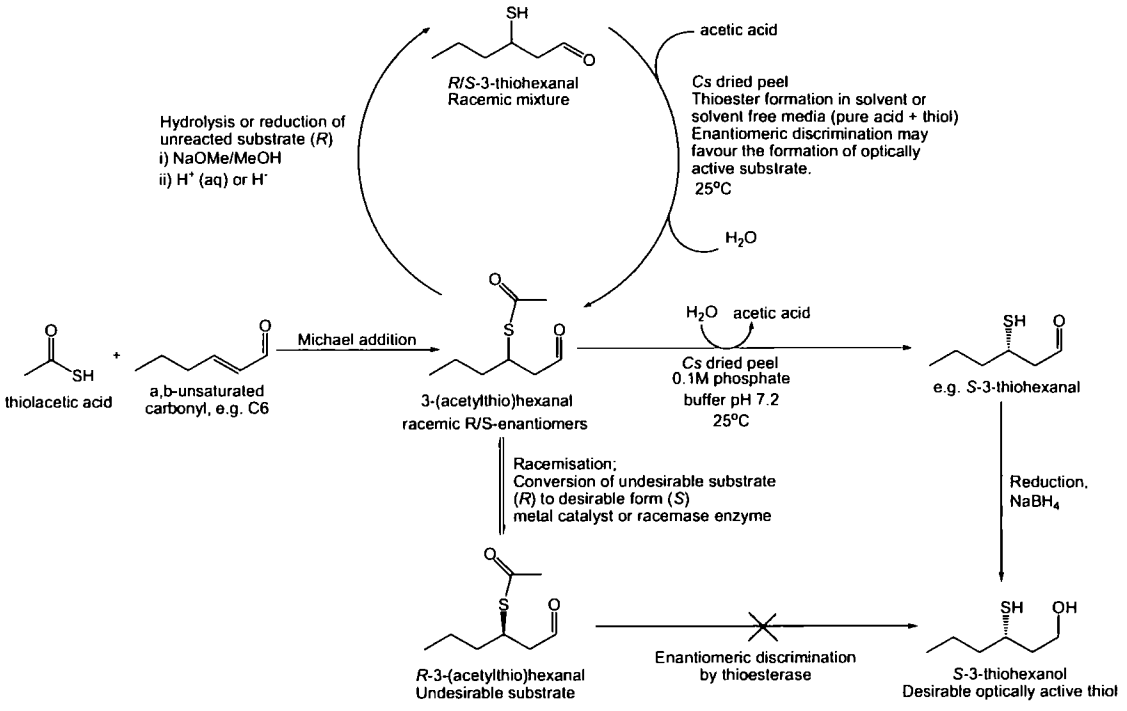


Figure 7.3 The application of *Citrus sinensis* (*Cs*) dried peel thioesterase preparation in bioreactions. The illustration shows hypothetical bioconversions leading to the formation of optically pure 3-thiohexan-1-ol (passion fruit mercaptan). However, through changing the α,β-unsaturated carbonyl starting material, different optically pure products could be generated. This illustrates how current synthetic chemistries could work alongside enzymatic approaches.

Such an approach is non-selective for specific enantiomers and a racemic mixture is produced. Using the identified biocatalyst, *Cs*WH, in crude freeze-dried form may promote the formation of specific enatiomers. However, this remains to be examined and will be taken further by Oxford Chemicals Ltd. It is hoped that stereoselective plant thioesterases may link in to further biosynthetic processes being developed by the company whereby they would form the final biosynthetic steps in generating optically pure VOSCs with improved flavour character in comparison to synthetic counterparts. Plant cell wall hydrolases form a novel source of biocatalysts to add to the limited collection of esterases commercially available. Engineering such enzymes through mutation or immobilisation may provide the protein with further desirable properties, such as a tolerance toward solvents or biphasic solutions, improved stability and recovery of enzyme. The potential to drive natural

hydrolysis of esters or thioesters in reverse using such selective catalysts would be a very worthwhile endeavour.

This research initiated a programme for the bioproduction of natural high-impact aroma chemicals and therefore in addition to isolating the desired thioesterase it was important to identify further potential biocatalysts in plants for future work. Passion fruit cDNA plasmid libraries were therefore constructed from both secretory gland and mesocarp tissues using the SMART cDNA library construction kit (BD Biosciences, Methods 2.7), with the purpose of identifying further fruit enzymes of interest. A random screen of each library (n=30), provided multiple cDNA sequences (Annex IV). Each sequence was screened against the databases using nucleotide BLAST and translated in each reading frame before screening against the non-redundant protein databases using protein BLAST. A total of 28 encoded proteins were identified in the mesocarp library and 23 in the gland library (Table 7.1). The majority of enzymes (82%) were those of 1° metabolism, including proteins responsible for cellular structure, signalling and catalysis in central metabolic reactions. In addition, some interesting enzymes of secondary metabolism were also identified including a glutathione-S-transferase (GST), an alcohol dehydrogenase (ADH) and an alcohol acyl transferase (AAT), all potentially involved in aroma biosynthesis in passion fruit. Interestingly, the gland library contained a greater collection of secondary metabolic enzymes (22%) compared to the mesocarp library (14%), a likely consequence of the specialised role these glandular structures are postulated to have in flavour metabolism. A larger screen of this library or a targeted approach using probes may yield further cDNA sequences encoding enzymes with application in VOSC manufacture.

Table 7.1 Clones identified through random screening of passion fruit cDNA libraries. 30 clones were selected at random from each library (M: mesocarp, G: gland) and plasmid inserts sequenced from the 5' end using the pTriplEx2Forward primer. cDNA sequences are catalogued in Annex IV. The identification of each clone through similarity to database sequences is shown and the significance of each hit highlighted through Score and Expect values (score: higher the better, measure of how close the match is, Expect: lower the better, the chance such an alignment could occur by chance). Partial 5' sequences may be full length inserts but require 3' sequencing to determine the full length code. Clones involved in 1° metabolism are listed above the blue dotted lines, those of 2° metabolism below.

Clone	Top BLAST hit Description	Species hit Accession	Score E value
M1	Glyceraldehyde-3-phosphate dehydrogenase. Catalyzes the 6th step of glycolysis. Full length.	*Solanum lycopersicum* AAB51592	298 $2e^{-80}$
M2	Phosphoenolpyruvate carboxykinase (PEPCK)-HprK superfamily. Catalyzes the diversion tricarboxylic acid cycle intermediates toward gluconeogenesis. Partial 3' sequence.	*Flaveria pringlei* BAB43909	288 $8e^{-78}$
M3	M20 dimer superfamily protein. Aspartyl aminopeptidase (putative). Protein metabolism. Internal sequence.	*Arabidopsis thaliana* NP200824	327 $6e^{-89}$
M4	Auxin-repressed protein, cellulose synthase like. Full sequence.	*Elaeagnus umbellata* AAC62104	186 $8e^{-47}$
M5	Octicosapeptide/ phox/ bemlp (PB1) domain containing protein. Cytoplasmic signalling, kinase activity. Full length	*Arabidopsis thaliana* NP190407	161 $4e^{-39}$
M6	Transcription initiation factor II F (β-subunit). Interacts with further transcription factors and RNA polymerase, required for accurate transcription. Full length.	*Vitis vinifera* CAN81117	223 $1e^{-57}$
M7	Frigida superfamily protein. Hydroxy proline rich, probably nuclear and required for the regulation of flowering time in late flowering phenotype. Partial 3' sequence.	*Arabidopsis thaliana* NP566709	344 $3e^{-94}$
M8	Ribosomal-L22 superfamily protein. core protein of the large ribosomal sub-unit. Full sequence.	*Populus trichocarpa* ABK94222	347 $2e^{-95}$
M9	Heat shock cognate protein 70. HSP-70. Shaperone protein family member, ATP dependent. Partial 5' sequence.	*Thellungiella h* AAS09825	491 $2e^{-138}$
M10	ER lumen protein-retaining receptor family protein. Partial 3' sequence.	*Oryza sativa* EAZ09872	187 $1e^{-47}$
M11	Arabinogalactan protein 10 (AGP10). Extracellular matrix protein. Full length.	*Arabidopsis thaliana* NP192642	44.3 $2e^{-4}$
M12	High mobility group 1 (HMG-1) protein. Contains a DNA binding domain involved in chromatin restructure and the recruitment of transcription factors. Full length.	*Lanavalia gladiata* BAA19156	167 $2e^{-41}$
M13	Histone H3.2. Maintenance of chromatin structure. Full length.	*Arabidopsis thaliana* NP001078516	273 $2e^{-73}$
M14	NAP-superfamily protein. Nucleosome assembly protein, moving histones into the nucleus, nucleosome assembly and chromatin fluidity, affects the transcription of many genes. Internal sequence.	*Zea mays* NP001105594	220 $2e^{-57}$

M15	DEAD/DEAH box helicase (putative). ABC-ATP-ase superfamily. Involved in RNA metabolism. Internal sequence.	*Arabidopsis thaliana* NP188490	340 $5e^{-93}$
M16	Heavy metal associated (HMA) domain superfamily protein. Transport of heavy metals, detoxification, copper chaperones. Internal sequence.	*Arabidopsis thaliana* NP568449	194 $1e^{-49}$
M17	Defender against cell death 1 (DAD1) superfamily protein. Integral membrane protein, causes apoptosis if mutated. Full length.	*Citrus unshiu* Q9ZWQ7	224 $1e^{-58}$
M18	Phosphate induced (Phi-1) superfamily protein. Partial 5' sequence.	*Vitis vinifera* CAN84160	249 $1e^{-65}$
M19	Domain of unknown function (DUF408). Similarity to myosin heavy chain protein. Partial 5' sequence.	*Arabidopsis thaliana* AAB61054	103 $2e^{-22}$
M20	Unknown protein. Internal sequence.	*Vitis vinifera* CAO46545	273 $1e^{-72}$
M21	Unknown protein. Full sequence.	*Vitis vinifera* CAO61141	145 $2e^{-34}$
M22	Unknown protein. Full length.	*Populus trichocarpa* ABK95804	308 $1e^{-83}$
M23	Unknown protein. Internal sequence.	*Vitis vinifera* CAO65959	319 $1e^{-86}$
M24	Unknown protein. Internal sequence.	*Vitis vinifera* CAO61935	129 $6e^{-30}$
M25	Glutathione-S-transferase (Phi class). Catalyses the conjugation of reduced glutathione to electrophilic centres on a range of substrates. Full length.	*Vitis vinifera* ABW34390	345 $9e^{-95}$
M26	AdoHcyclase superfamily protein. S-adenosyl-L-homocysteine hydrolase (hypothetical). Inhibits S-adenosyl-L-methionine-methyl transferases activity and modulates their activity. Full length	*Vitis vinifera* CAN68176	265 $2e^{-70}$
M27	Endochitinase I precursor. Chitin bindinding domain 1 (chtBD1). Binds N-acetylglucosamine, wound induced glyco hydrolase. Partial 5' sequence.	*Theobrama cacao* Q41596	111 $1e^{-24}$
M28	AdoHcyase superfamily protein. S-adenosyl-L-homocysteine hydrolase (hypothetical). Inhibits S-adenosyl-L-methionine-methyl transferases activity and modulates their activity. Partial 3' sequence.	*Medicago trunculata* ABD28441	177 $1e^{-44}$

Clone	Top BLAST hit Description	Species hit Accession	Score E value
G1	NADH dehydrogenase (sub-unit 7). Partial 3' sequence.	*Brassica napus* YP717100	298 $8e^{-81}$
G2	ATP-dependent CLP protease proteolytin sub-unit, classic ser, asp, his catalytic triad. Partial 3' sequence.	*Solanum bulbocastanum* YP538874	69.7 $4e^{-12}$
G3	CLP protease superfamily protein. Partial 3' sequence.	*Solanum bulbocastanum* YP538874	70.1 $1e^{-11}$
G4	ARF family protein. Ar11 ABC-ATPase superfamily protein. ADP ribosylation factor 3 ARF3/ARL1/ATARL1. GTP binding protein of RAS family, regulates vesicular traffic and actin remodelling. Full length.	*Arabidopsis thaliana* NP850057	358 $1e^{-98}$

G5	PP2Cc superfamily protein, protein phosphatase 2C (PP2C). Serine/ threonine phosphatase. Partial 3' sequence.	*Fagus sylvatica* CAB90634	281 $1e^{-75}$
G6	Protein phosphatase. Alpha-crystallin-Hsps superfamily protein. Partial 3' sequence. Heat shock protein believed to be ATP dependent, chaperone that prevents aggregation and aids in protein folding.	*Rubus idaeus* CAB61630	317 $4e^{-85}$
G7	Putative choline kinase. Catalyzes the formation phosphocholine from ATP and choline. Involved in glycine, serine and threonine metabolism. Internal sequence.	*Arabidopsis thaliana* NP177572	186 $4e^{-47}$
G8	CKL6/PAPK1 Casein kinase ADK1-like protein. Regulator of signal transduction. Partial 3' sequence.	*Arabidopsis thaliana* NP567812	76.3 $1e^{-13}$
G9	DNA-directed RNA polymerase I, II and III. 7kDa sub-unit (putative)	*Arabidopsis thaliana* NP198917	72.4 $2e^{-14}$
G10	PAPA-1 like family protein. Zinc finger (HIT type) coordinates zinc. Partial 3' sequence.	*Arabidopsis thaliana* NP176041	137 $2e^{-32}$
G11	HMA superfamily protein. Heavy metal associated domain-containing protein. Full length.	*Arabidopsis thaliana* NP197247	5.216 $9e^{-56}$
G12	Copper binding family protein. Partial 3' sequence.	*Arabidopsis thaliana* NP974830	76.6 $3e^{-14}$
G13	Contains similarity to myosin heavy chain protein. Partial 3' sequence.	*Arabidopsis thaliana* AAB61054	239 $1e^{-62}$
G14	Ribosomal L21-e superfamily proteins. 60 S ribosomal protein, L21 sub-unit. Partial 3' sequence.	*Oryza sativa* ABF93903	251 $1e^{-66}$
G15	Cyclophilin-like protein. Binds to cyclosporin A, facilitates protein folding and isomerises peptide bonds. Full length.	*Nicotiana tobacum* ABS30424	359 $5e^{-99}$
G16	40 S ribosomal protein 514. Ribosomal S-11 superfamily protein. Partial 3' sequence.	*Zea mays* P19950	118 $9e^{-27}$
G17	Unknown protein. Full length.	*Vitis vinefera* CAO16842	93.2 $3e^{-14}$
G18	Unknown protein, UPF0139 superfamily. Uncharacterised. Partial 3' sequence.	*Populus trichocarpa* ABK93624	179 $5e^{-45}$
G19	Cinnamyl alcohol dehydrogenase, putative. Oxidoreductase. Partial 3' sequence.	*Malus domestica* AAC06319	106 $3e^{-23}$
G20	Thaumatin superfamily protein. Natural sweetener, pathogen response protein believed to inhibit sporulation and hyphal growth in fungi. Partial 3' sequence.	*Populus trichocarpa* ABK96488	150 $2e^{-36}$
G21	Endoglucanase. Degradation of cellulose. Internal sequence.	*medicago trunculata* ABD33428.1	221 $2e^{-57}$
G22	1-amino cyclopropane-1-carboylate oxidase. Partial 3' sequence.	*Hera brasiliensis* AAP41850	97.1 $2e^{-20}$
G23	Alcohol acyl transferase (AAT). Catalyzes acyl group transfer to alcohols. Internal sequence.	*Hevea brasiliensis* AF429383	51.8 $2e^{-4}$

The random library screen and the targeted biochemical approach used for the isolation of *Cs*WH and *Pe*WH thioesterases illustrate different strategies for obtaining novel biocatalysts for flavour and fragrance biosynthesis. With a greater understanding of the metabolism of VOSCs in tropical fruit, such as passion fruit, and with the strategy for obtaining the enzymes involved, it is possible to build a model for the enzymatic production of sulphur high-impact flavourings (Figure 7.4). This "biofactory" encompasses both plant and microbial enzymes and illustrates how cheap and readily available feedstock, such as plant oils, animal keratin (hair, feathers, wool) and yeast extracts can be transformed into an array of distinctive flavourings. Importantly, the products derived from such processes are considered natural as feedstock is from a natural source, enzyme reaction conditions are mild and there is no use of synthetic catalysts or solvents[122,123]. Plant fruit cells have evolved over thousands of generations to produce specific flavour and fragrance secondary metabolites, and here we essentially plan to recreate these processes on the factory floor. Chemical manufacturers strive toward efficiency and sustainability. What better way of achieving this than to reproduce what nature has so elegantly designed?

Oxford Chemicals Ltd. has invested in fermentation equipment and is funding three further biochemists from Durham University to develop the biofactory. Their first enzymatically produced high-impact flavouring went on sale in February 2008.

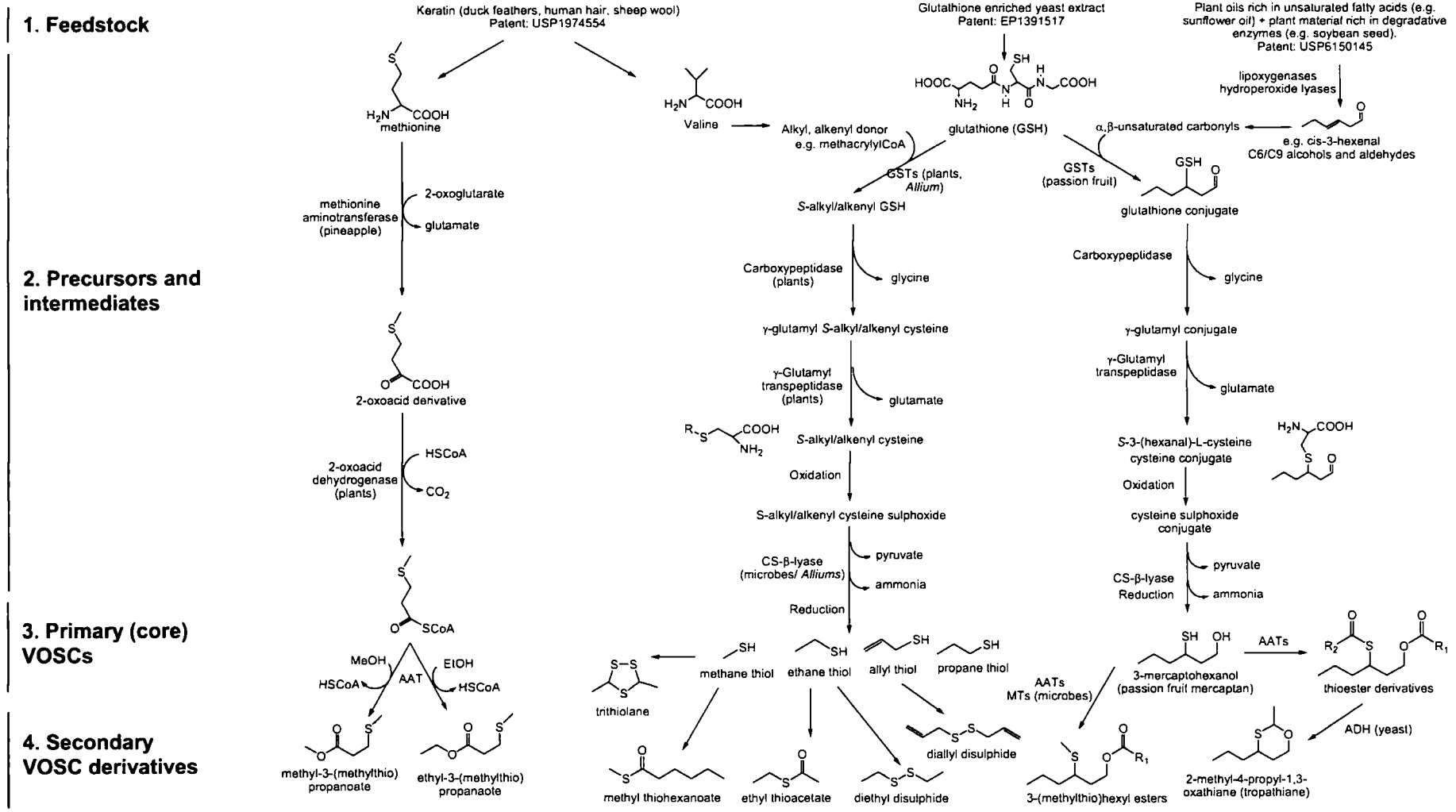


Figure 7.4 A hypothetical model for the bioproduction of high-impact flavourings. The generation of feedstock is covered by patents; US patent 6150145: Process for the production of degradation products of fatty acids, US patent 1974554: Process for the production of amino acids from substances containing keratin, European patent 1391517: Process for producing glutathione. The enzymatic steps leading to final products are currently hypothetical.

Supplementary Data

Annex I: *Pe*WH Protein Analysis

A. MALDI data for *Pe*WHa and *Pe*WHb

B. MS-MS data for *Pe*WHb

Annex II: Protein and DNA Sequence Analysis

A. Alignment of full length plant PAE sequences

B. Alignment of full length plant PAE sequences within each of the three distinct clusters of homology (clades 1-3)

C. Alignment of full length plant PAE cDNA sequences from sub-family 1.

D. Alignment of translated internal PAE sequences amplified from the leaf and mesocarp tissues of purple passion fruit.

E. Partial nucleotide and deduced amino acid sequence of the *Pe*PAE cDNA from passion fruit mesocarp tissue

Annex III: *Cs*WH Protein Analysis

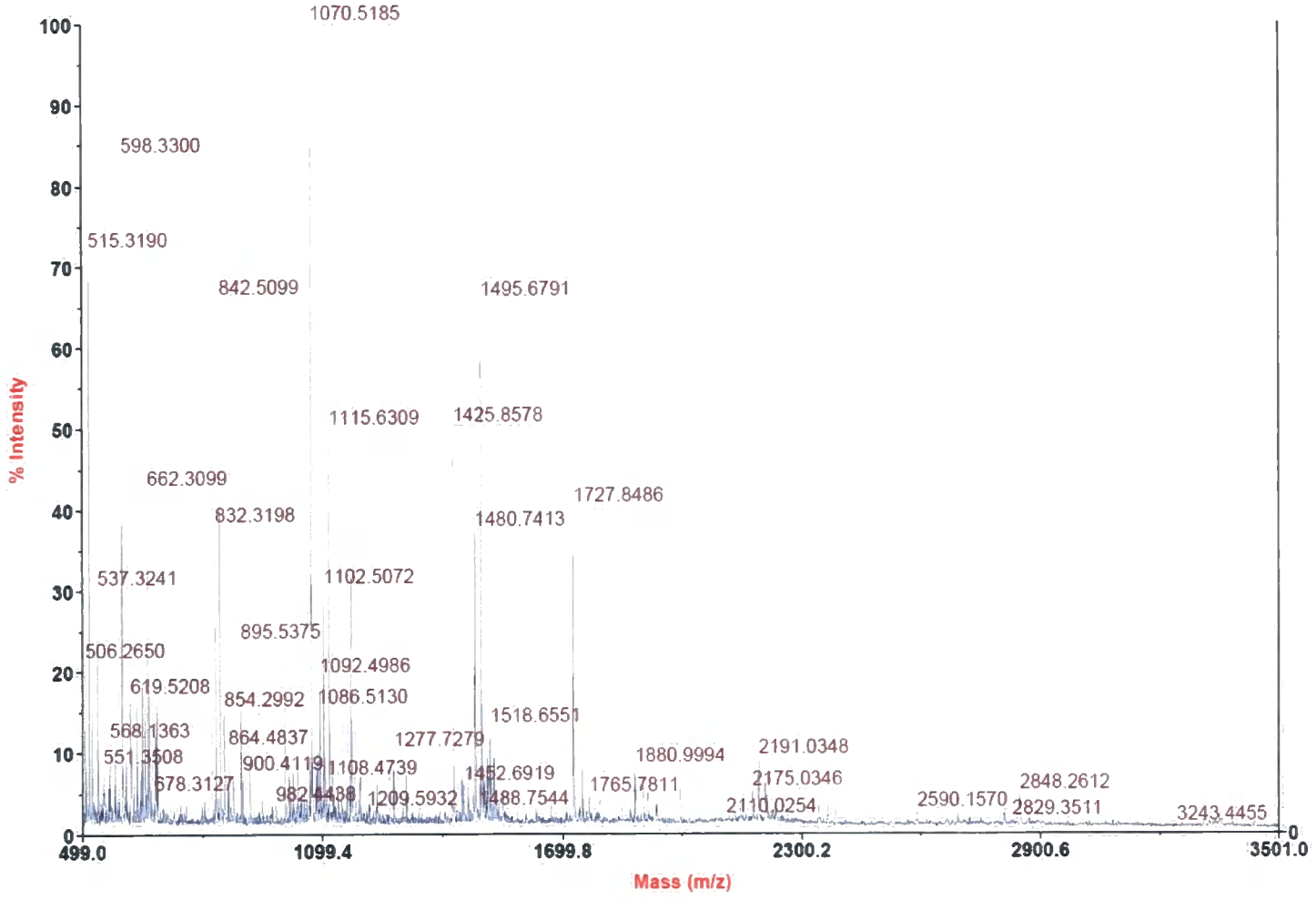
A. MALDI data for *Cs*WHa and *Cs*WHb.

B. Predicted MS-tryptic digest of constructed *Citrus sinesis* L. Osbeck PAE

C. MS-MS data for *Cs*WHa

Annex IV: Passion fruit cDNA sequences identified in random library sceen

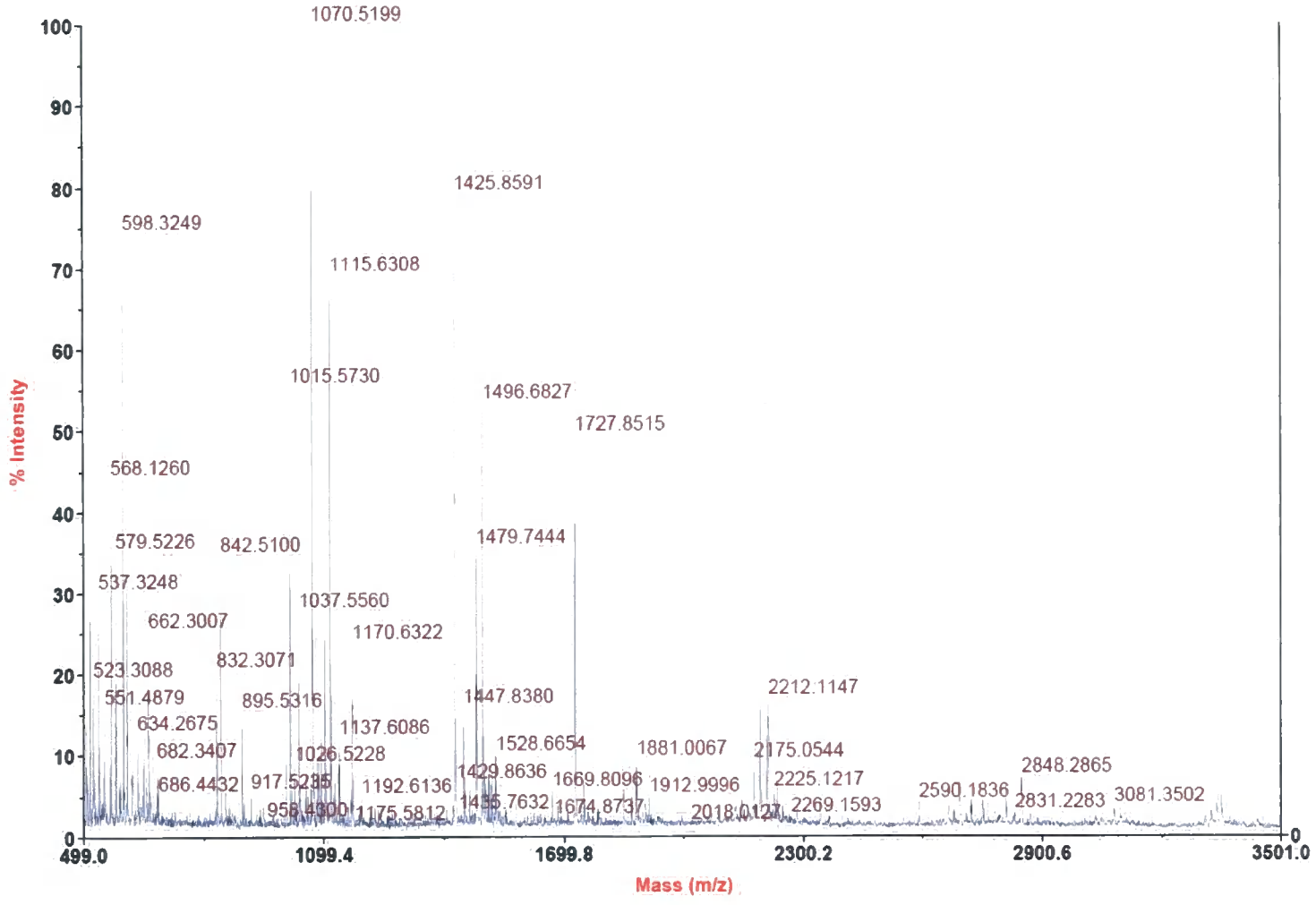
1.



Annex I *Pe*WH Protein Analysis.

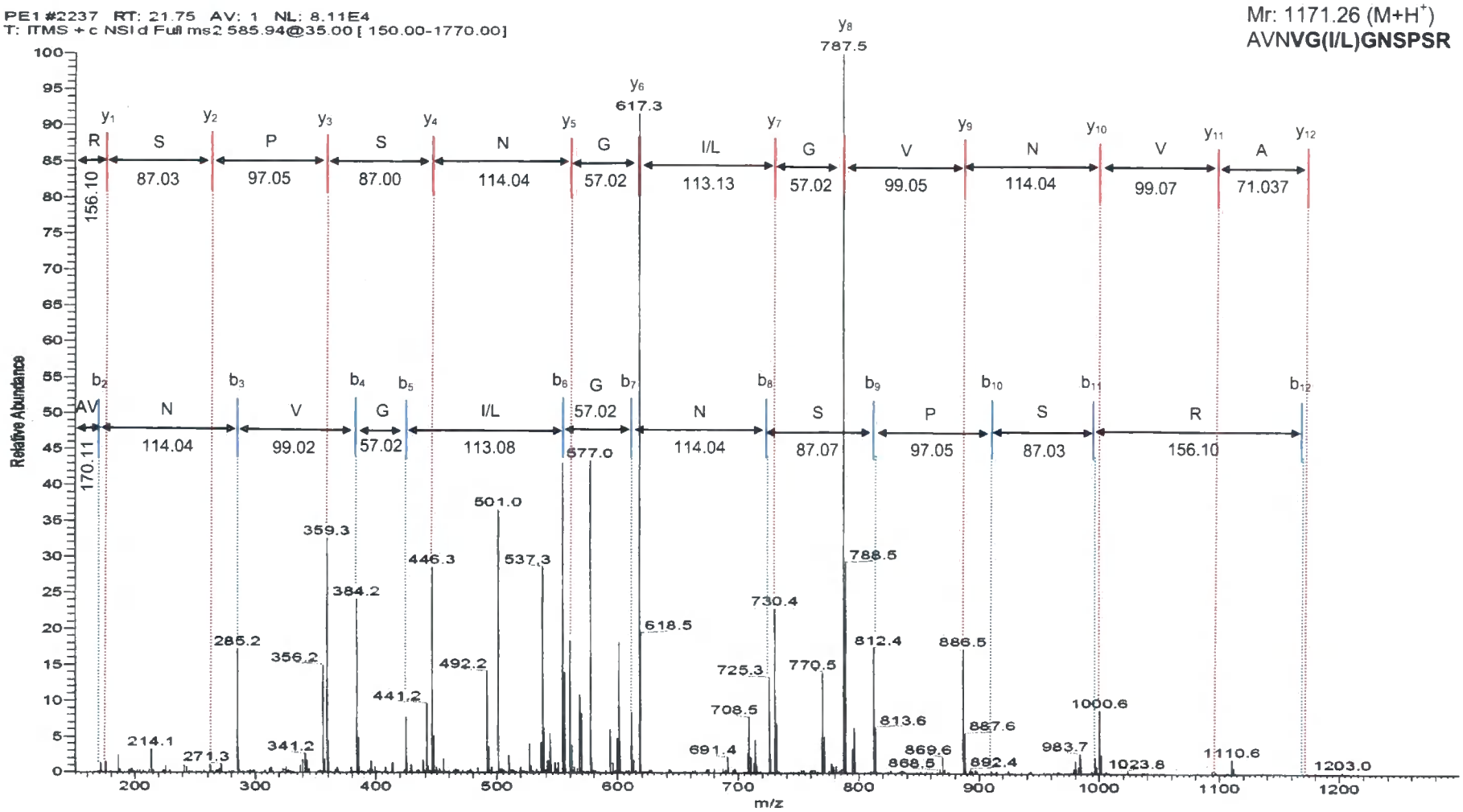
A. MALDI traces for *Pe*WHa (**1.**) and *Pe*WHb (**2.**). Monoisotopic peak lists for MALDI on *Pe*WHa and *Pe*WHb (**3.**).

2.

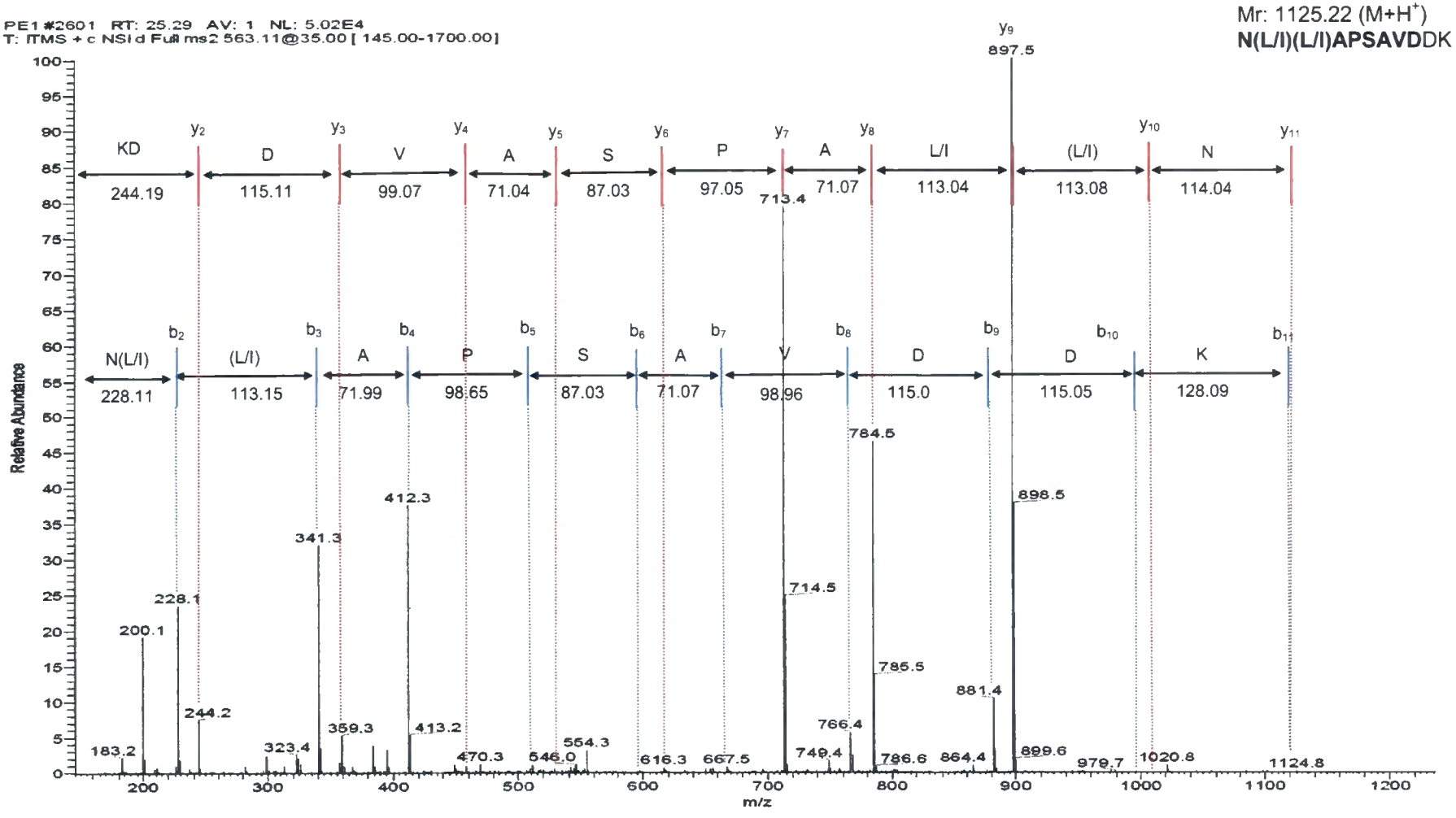

1070.5199
1425.8591
598.3249
1115.6308
1015.5730
1496.6827
1727.8515
568.1260
579.5226
842.5100
1479.7444
537.3248
1037.5560
662.3007
1170.6322
832.3071
523.3088
551.4879
895.5316
1447.8380
2212.1147
634.2675
1137.6086
1528.6654
1881.0067
2175.0544
682.3407
1026.5228
1429.8636
1669.8096
1912.9996
2225.1217
2848.2865
686.4432
917.5285
1192.6136
1455.7632
1674.8737
2018.0127
2269.1593
2590.1836
2831.2283
3081.3502
958.4800
1175.5812
% Intensity
Mass (m/z)
100
90
80
70
60
50
40
30
20
10
0
499.0
1099.4
1699.8
2300.2
2900.6
3501.0

3. Monoisotopic peak list for MALDI on *Pe*WHa and *Pe*WHb. The 8 most intense peaks from each 100Da window of the spectra are shown. The fragmentation pattern indicates they are the same protein with the same mass ion fragmentation. Possible post translational modifications must therefore account for the differences in properties between the soluble and wall-bound forms of the protein.

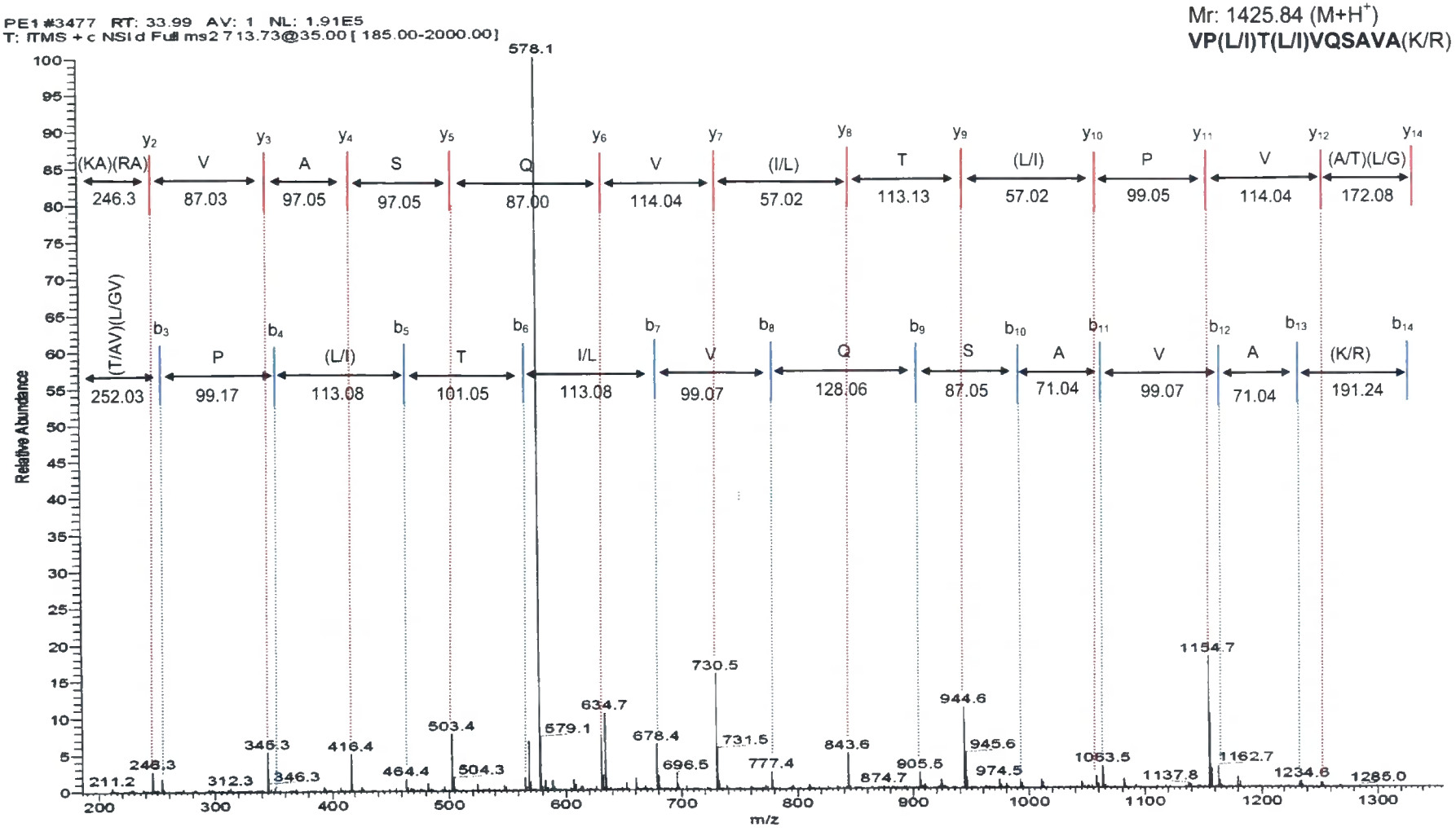
***Pe*WHa**			***Pe*WHb**		
506.2648	1167.5234	2190.0339	506.2529	1425.8591	2720.2815
515.3191	1170.6319	2206.0490	515.3127	1447.8376	2748.3483
523.3092	1179.6169	2211.1029	523.3086	1451.7013	2807.3283
537.3241	1192.6057	2342.1459	537.3249	1478.6793	2846.2993
565.3610	1214.6307	2692.3135	551.4870	1479.7443	2870.2789
568.1364	1232.6082	2846.2712	568.1260	1493.6997	3079.3578
579.5246	1235.5515	2868.2497	579.5227	1495.6817	3095.4533
598.3299	1277.7281		598.3250	1501.7306	3323.6935
619.5209	1299.7023		607.5548	1509.6681	3338.8006
634.2720	1307.6785		620.3104	1511.6799	3347.6923
647.5560	1308.6685		622.0340	1517.6719	
650.0457	1324.6800		634.2680	1527.6696	
662.3099	1425.8578		650.0396	1539.7046	
666.0140	1447.8254		662.3012	1549.6504	
682.3519	1462.7189		682.3401	1578.9457	
685.4352	1478.6614		685.4254	1669.8110	
700.2842	1479.7377		832.3076	1726.8542	
701.3705	1480.7414		834.3224	1741.7738	
758.4627	1493.7114		842.5101	1748.8420	
829.4638	1495.6790		854.2982	1758.8591	
832.3198	1501.7112		864.4865	1783.8668	
834.3239	1507.7314		895.5312	1846.8027	
842.5099	1511.6714		917.5238	1868.7984	
854.2993	1517.6557		948.4728	1879.9838	
864.4827	1527.6604		1005.4835	1901.9791	
876.2842	1533.6455		1015.5730	1911.9861	
895.5377	1539.6548		1026.5239	2143.0576	
900.4137	1549.5963		1037.5561	2145.5570	
917.5250	1708.8182		1070.5199	2174.0396	
948.4449	1726.8533		1083.5485	2190.0457	
1005.4865	1740.8245		1086.5175	2191.0581	
1015.5644	1748.8264		1092.5035	2206.0440	
1026.5215	1765.7822		1102.5031	2208.0330	
1037.5490	1791.7696		1105.5261	2211.1046	
1068.5211	1846.8022		1115.6309	2225.1275	
1070.5187	1868.7938		1124.6021	2230.0596	
1086.5133	1879.9992		1127.5626	2233.0999	
1092.4988	1901.9906		1137.6084	2341.1591	
1102.5072	1912.0050		1167.5373	2589.1603	
1115.6307	1933.9805		1170.6326	2663.3573	
1124.4556	1993.9963		1226.6569	2677.2802	
1137.6065	2174.0360		1424.8177	2691.3063	



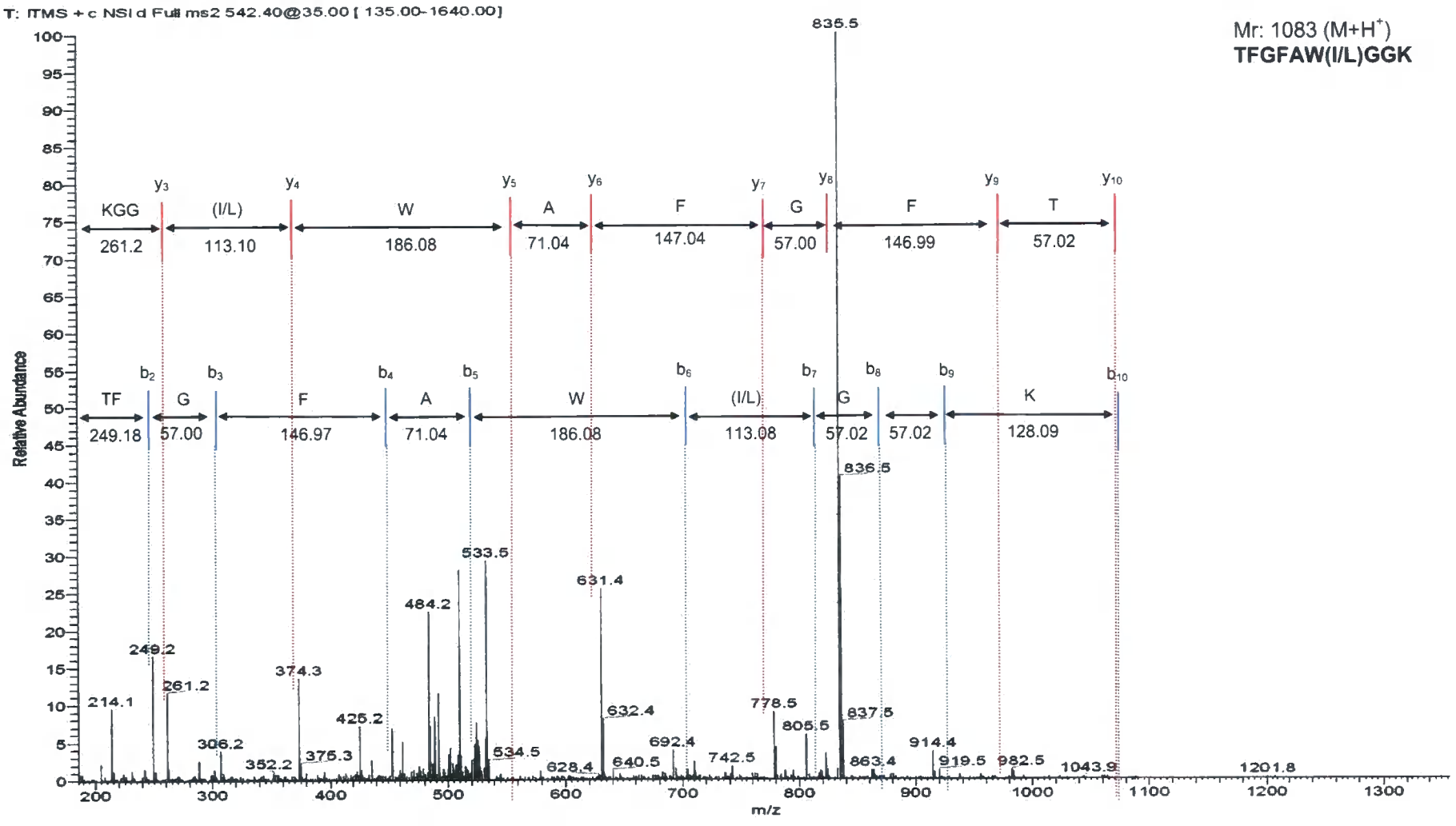
Annex I B MS-MS data for *Pe*WHb. Trace for ion m/z 586.



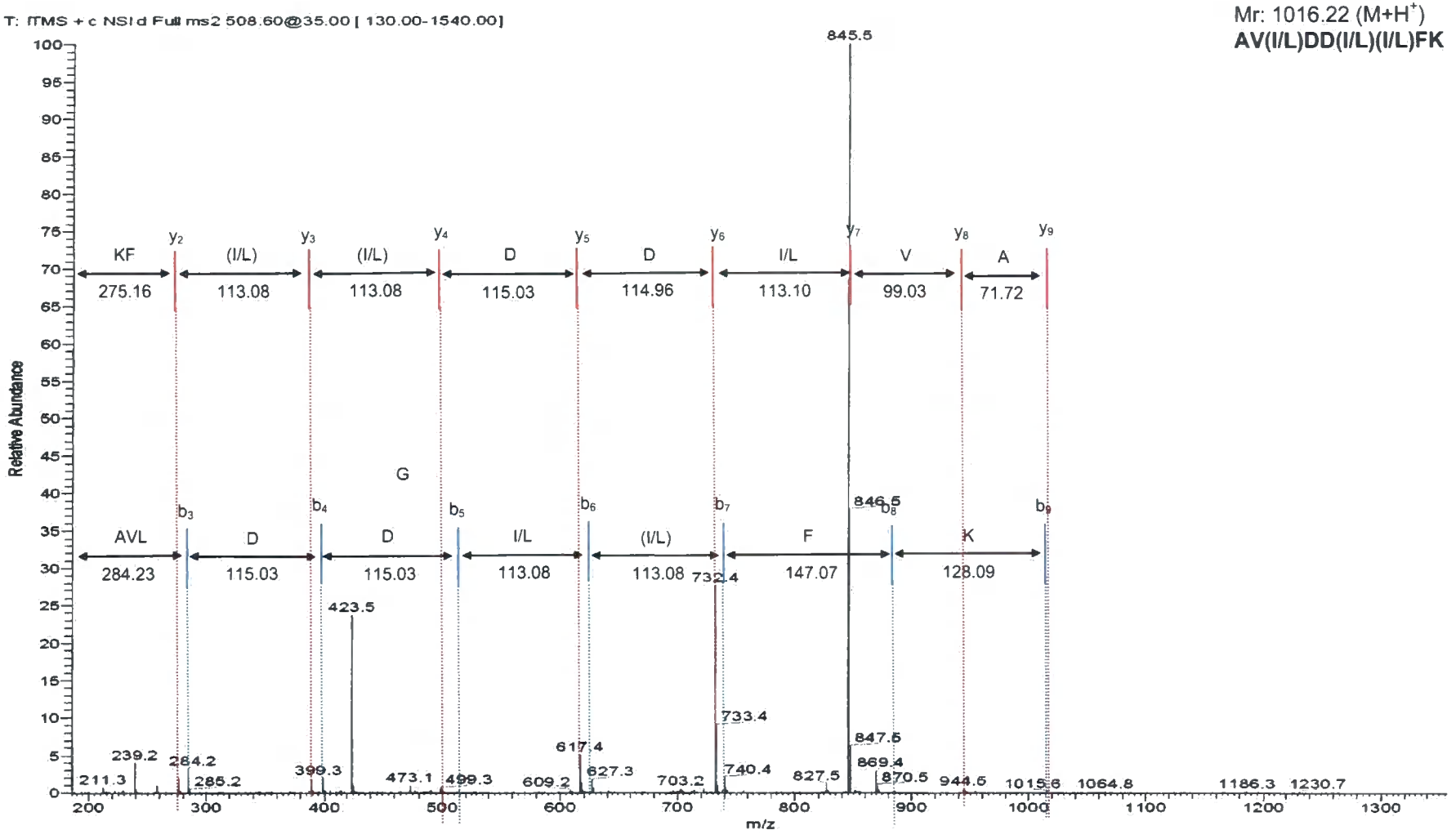
Trace for ion m/z 563. b and y ions are shown with predicted amino acid sequence. Amino acid sequence in bold was used to search the protein databases.



Trace for ion m/z 713. It is not possible to distinguish between leucine and isoleucine as they are isomers.



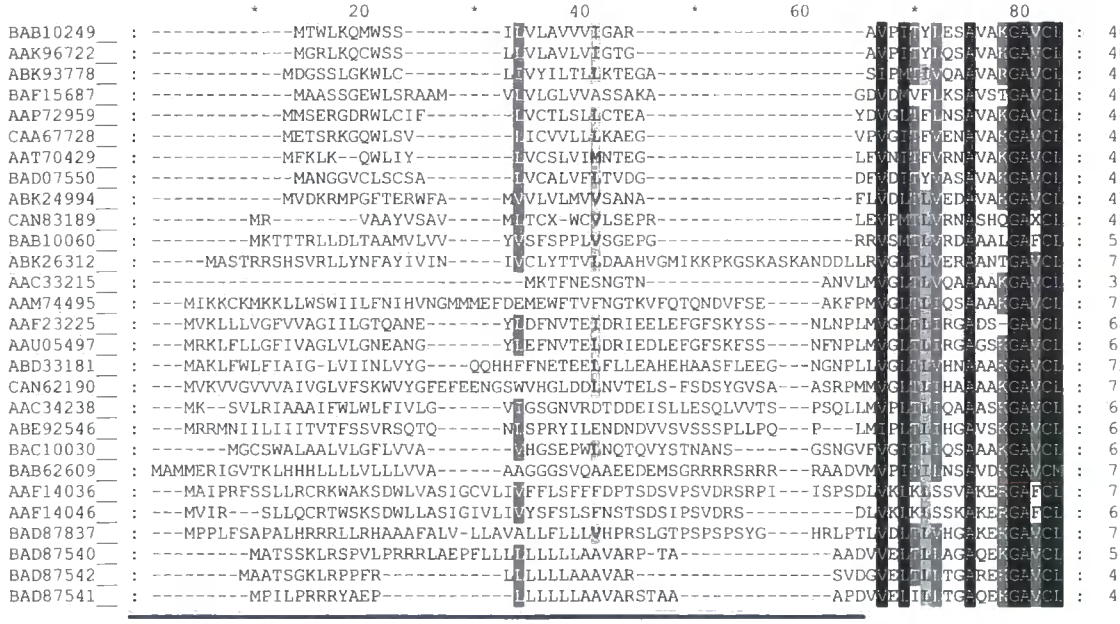
Trace for ion m/z 542. Only the amino acid sequence for accurately determined residues was used for further characterisation (i.e. the mass difference between both b and y ions confirmed identity).

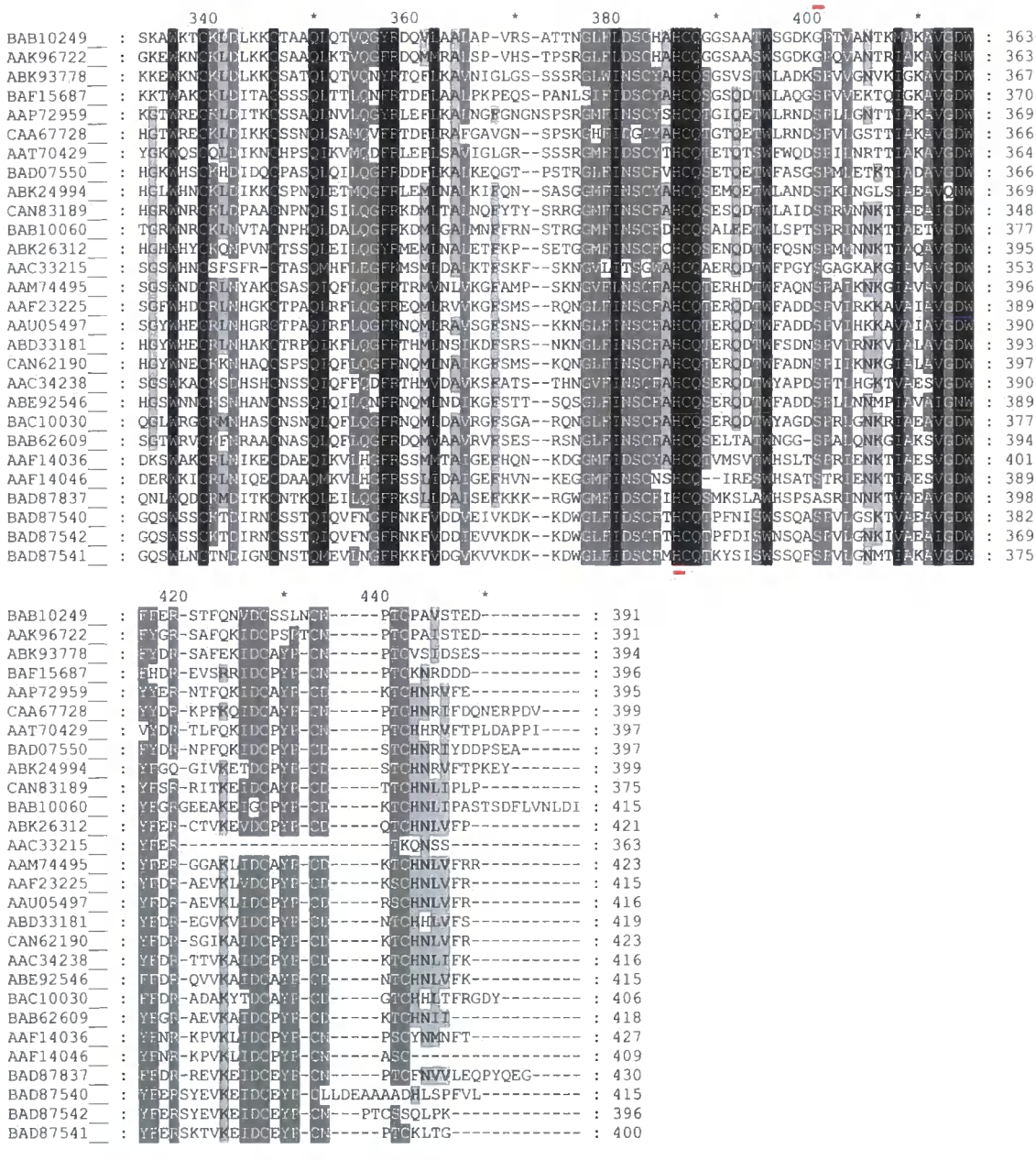


Trace for ion m/z 508. The amino acid sequence was determined using the automated *de novo* sequencing software DeNovoX (Thermo) and then checked by manual calculation for confirmation of the sequence.

Annex II Protein and DNA Sequence Analysis.

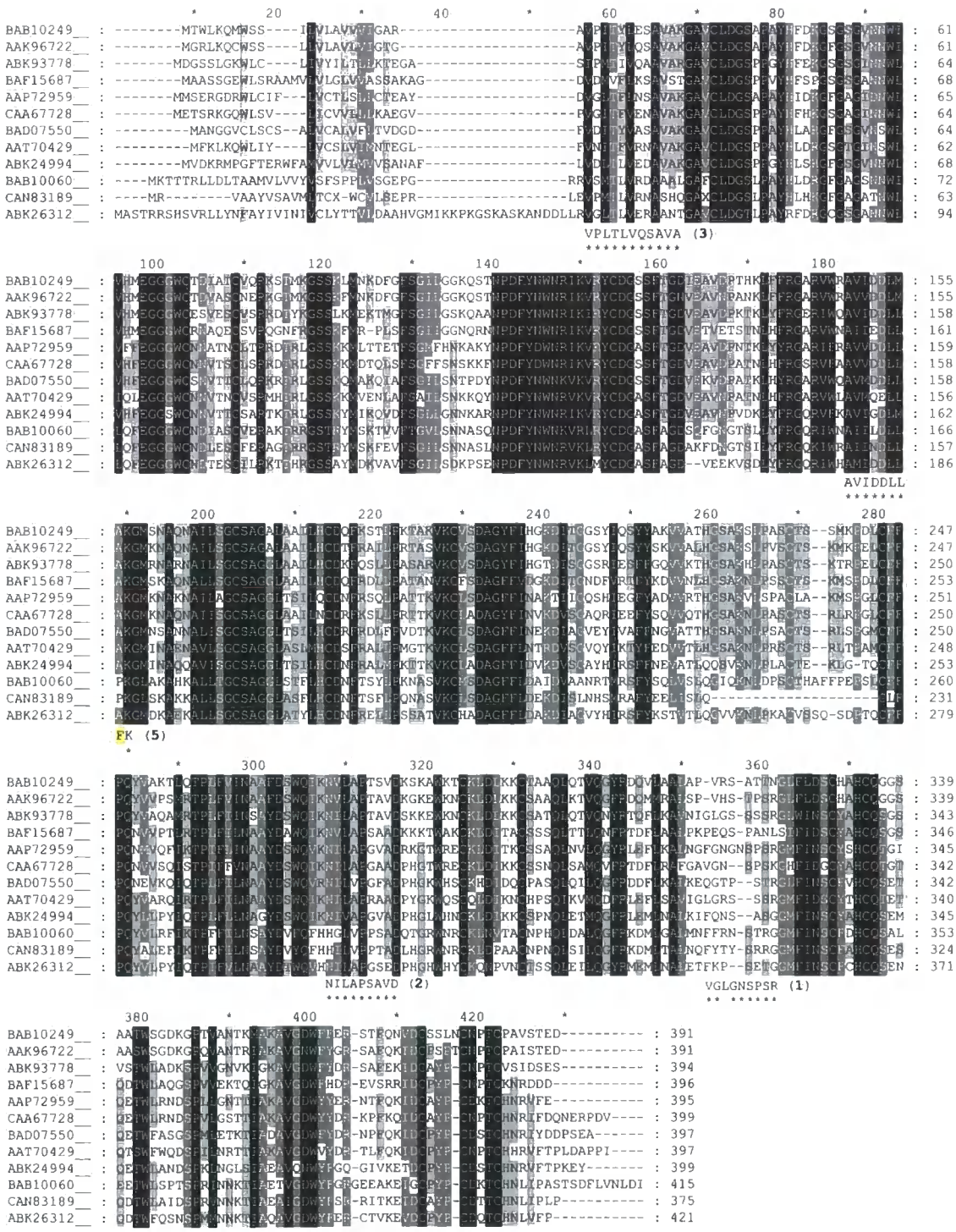
A. Alignment of full length plant PAE sequences from the carbohydrate active enzyme database (family 13 pectin acetylesterases). Each protein sequence is identified by its Genbank accession number. Black shading highlights complete sequence conservation, dark grey shading shows 80% or greater conservation and light grey shows 60% or greater conservation. The the GXSXG catalytic motif and aspartate and histidine residues which make up the catalytic triad are shown conserved across all sequences (underlined red). The predicted signal peptide sequence which targets the proteins to the cell wall is underlined in blue.



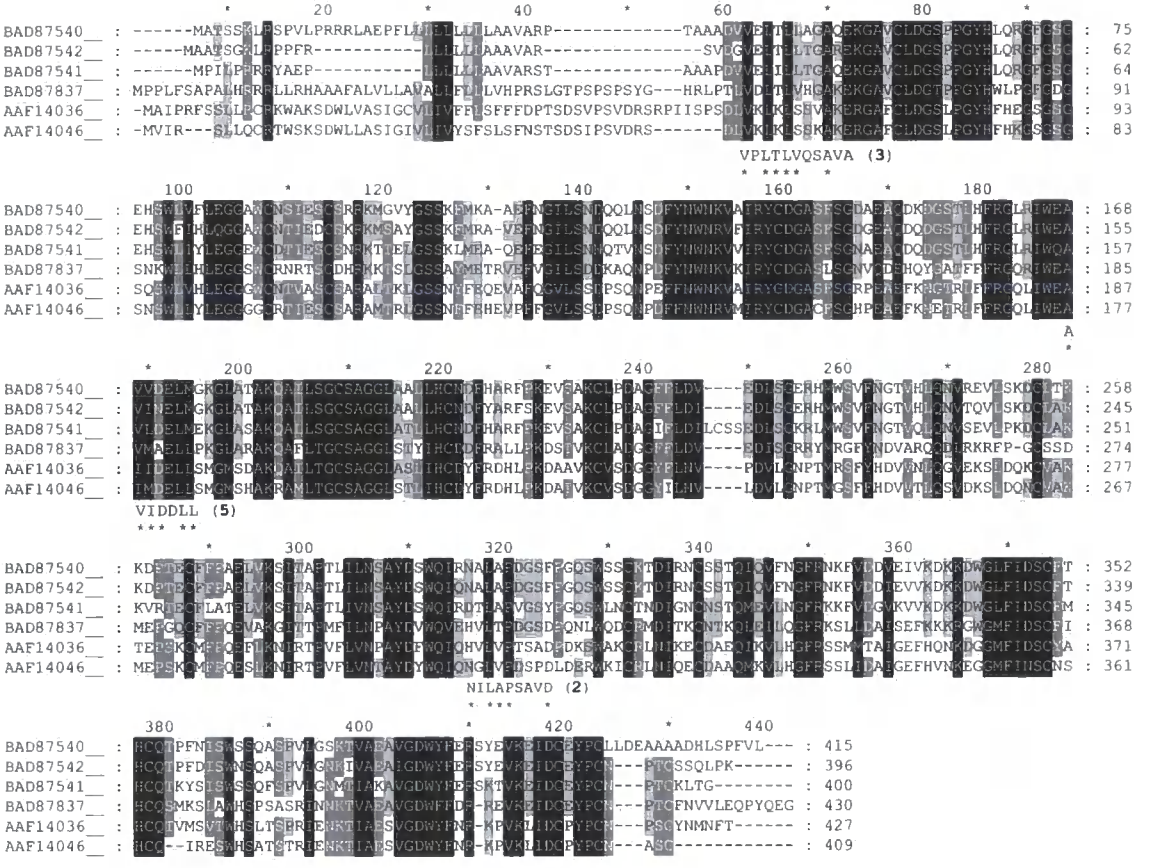


Annex II B. Alignment of full length plant PAE sequences within each of the three distinct clusters of homology (clades 1-3). Each protein sequence is identified by its Genbank accession number. Black shading highlights complete sequence conservation, dark grey shading shows 80% or greater conservation and light grey shows 60% or greater conservation. The alignment of peptides from *Pe*WH are shown with matches (*) and mismatches (space).

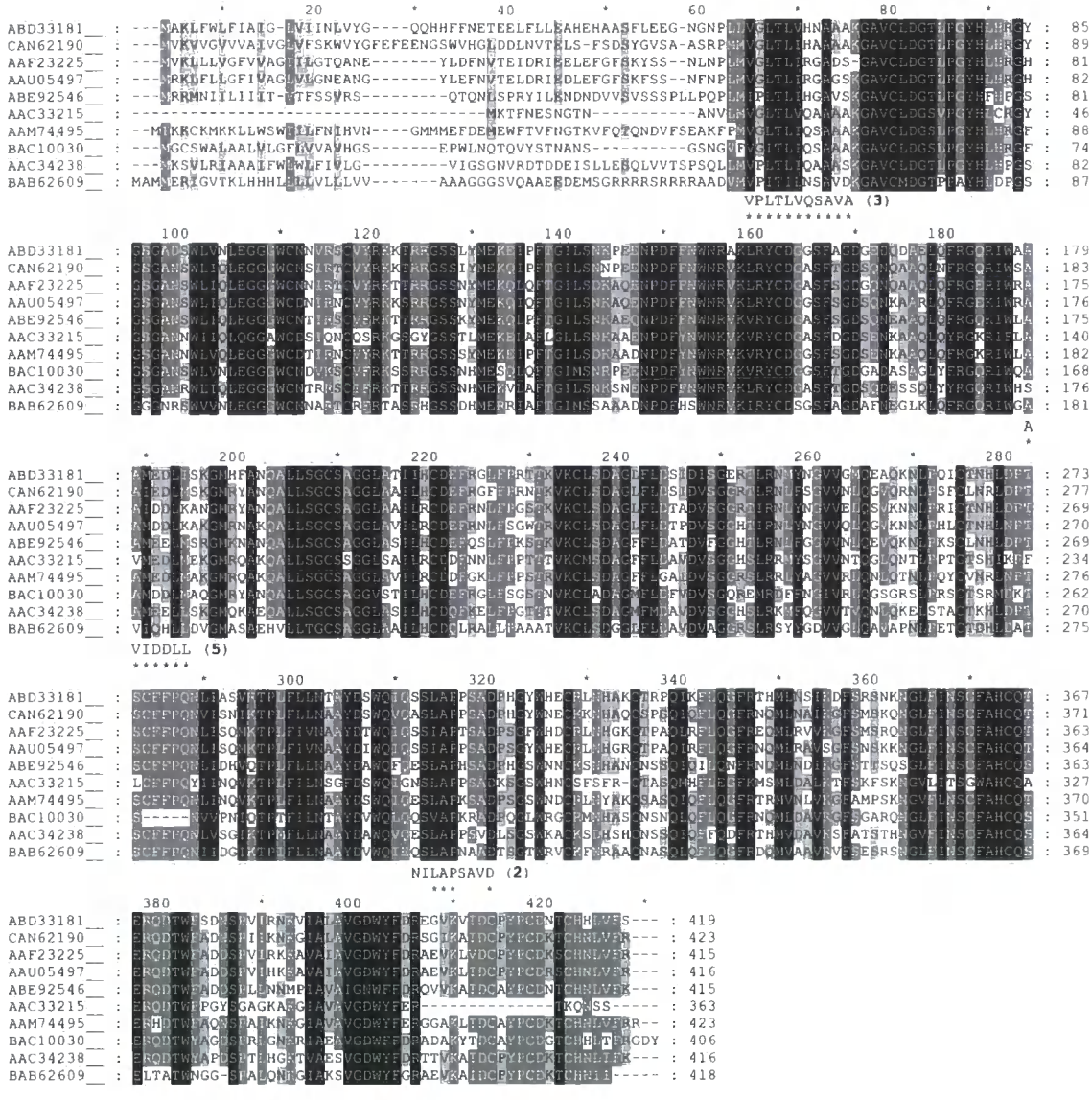
Clade 1:



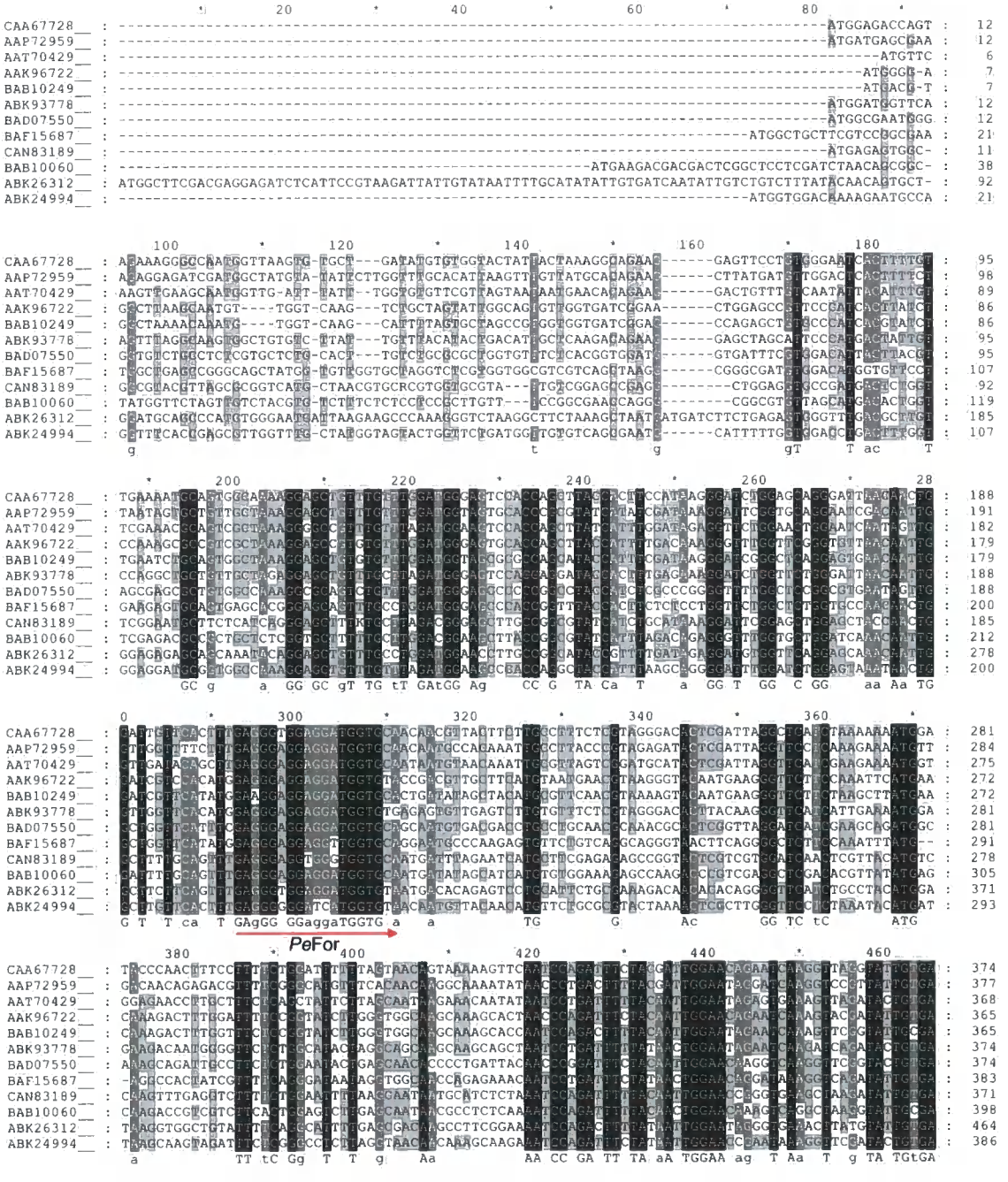
Clade 2:

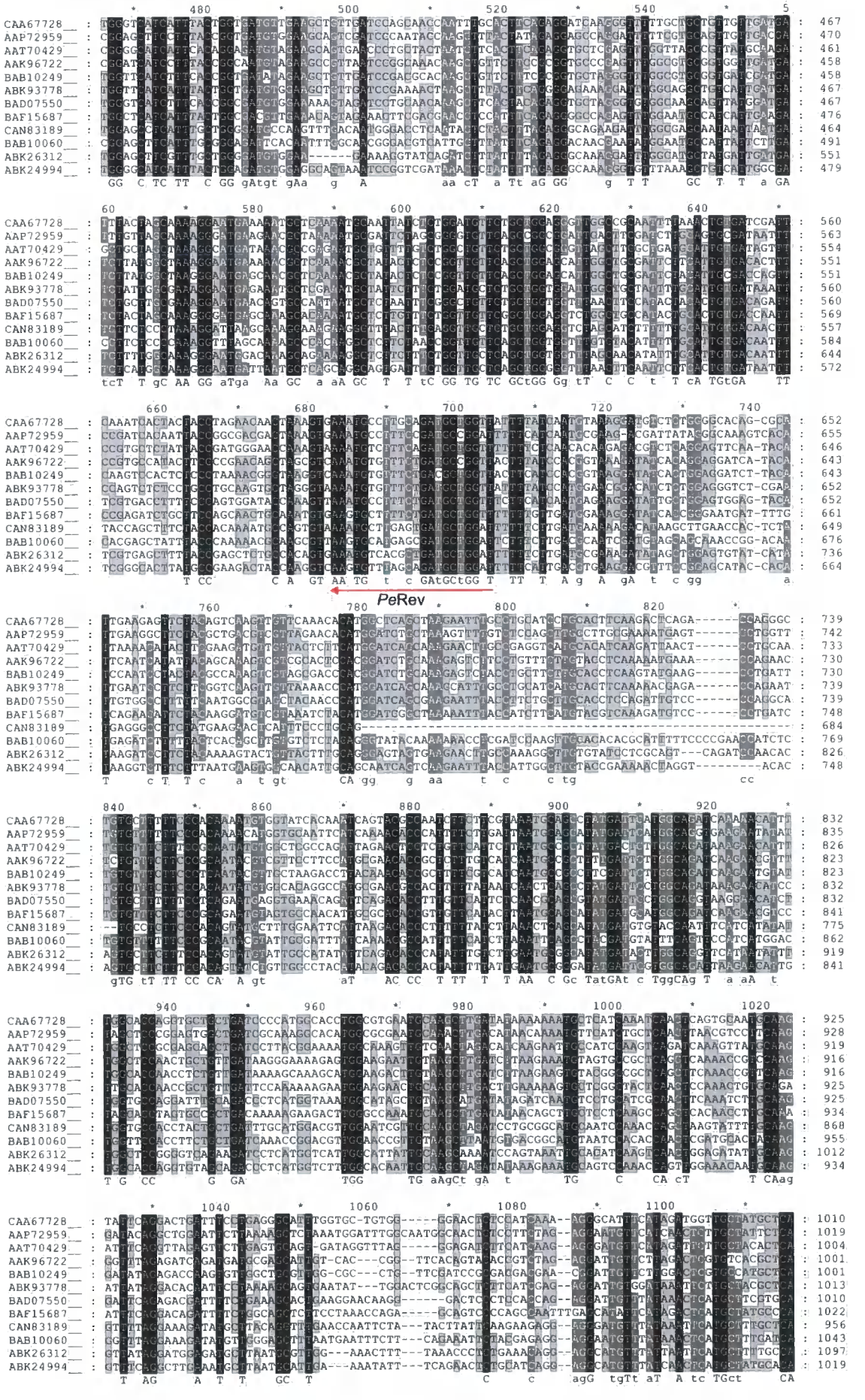


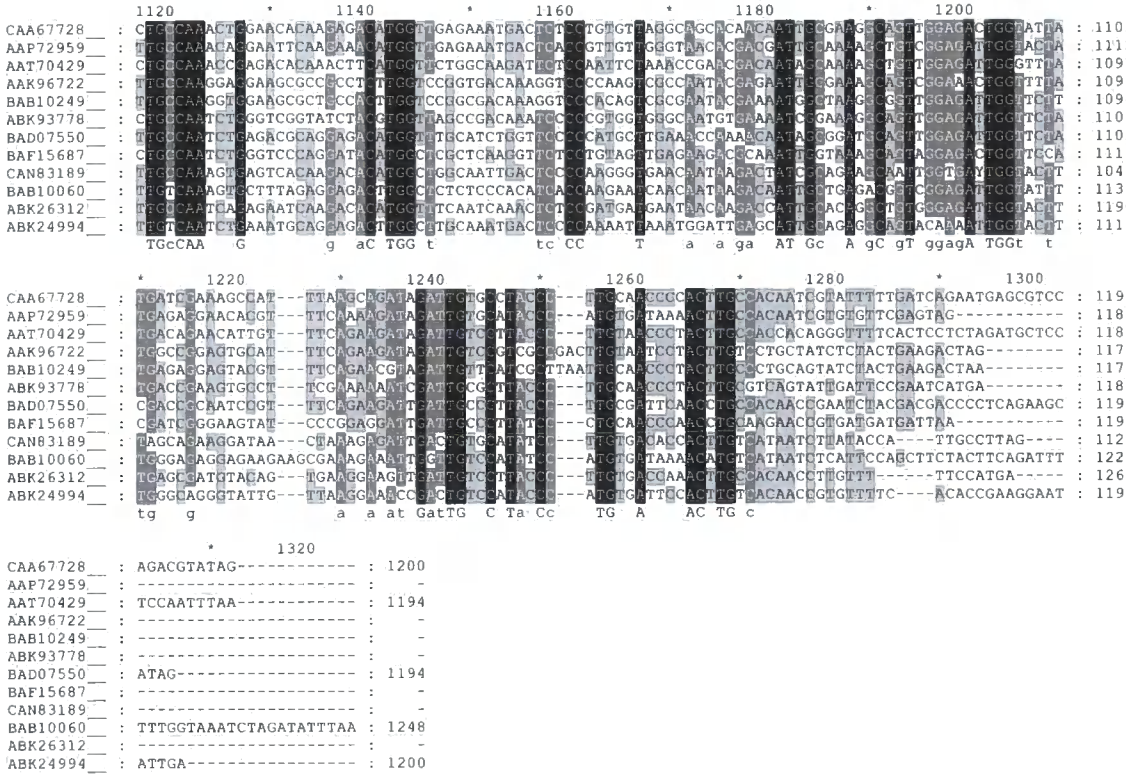
Clade 3:



Annex II C. Alignment of full length plant PAE cDNA sequences from sub-family 1. Each nucleotide sequence is identified by its Genbank accession number. Black shading highlights complete sequence conservation, dark grey shading shows 80% or greater conservation and light grey shows 60% or greater conservation. The consensus sequence is shown below the alignment. The optimal stretches of sequence used for designing degenerate primers *Pe*For and *Pe*Rev are underlined (red arrows).




PeRev



Annex II D. Alignment of translated internal PAE sequences amplified from the leaf and mesocarp (Meso) tissues of purple passion fruit. Conserved amino acid residues are marked (*), strong amino acid matches (:), weak amino acid matches (.) and mismatches are marked (space) as determined by CLUSTALW. The GXSXG catalytic motif is underlined (red). The predicted sequence over which *Pe*WH fragment **5** should match is shown.

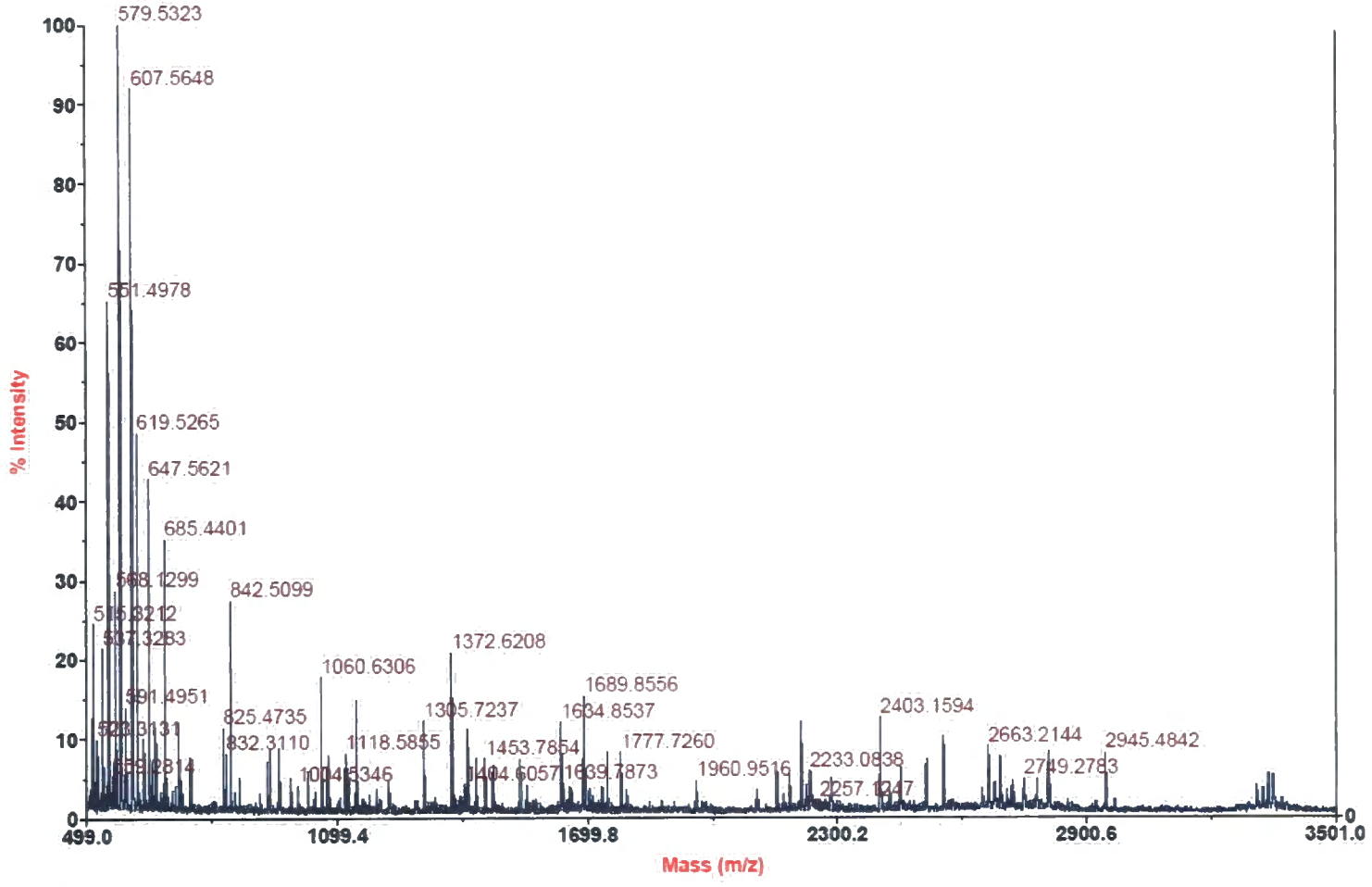
```
Leaf    EGGGWCNNVT TCLYRKNTHL GTSKLMGQPL AFSGILSNRG      40
Meso    EGGGWCNNVT NCVSRMHTRL GSSKKMVENL AFSAILSNKK
        ********** .*: * :*:* *:** * : * ***.****:

Leaf    TFNPDFYNWN RVKIRYCDGS SFTGDVQKVN PATNLHFRGA      80
Meso    QYNPDFYNWN RVKVRYCDGA SFTGDVEAVN PATNLHFRGA
         :******** ***:*****: ******: ** **********

Leaf    RIWLAVVQEL LAKGMKNAEN ALLSGCSAGG LASILHCDSF     120
Meso    RVWLAVMQEL LAKGLINAEN AVLSGCSAGG LTSLMHCDSF
        *:****:*** ****: **** *:********* *:*::*****
            AVIDDL LFK (5)

Leaf    RALLPMGAKV KCLSDAG                              137
Meso    RALLPMGTKV KCLSDAG
        *******:** *******
```

Annex II E. Partial nucleotide and deduced amino acid sequence of the *Pe*PAE cDNA from passion fruit mesocarp tissue. The GXSXG serine hydrolase catalytic motif and the histidine and aspartate residues which make up the catalytic triad are underlined (red). Two potential *N*-glycosylation sites are highlighted (grey). The regions used for primer design are underlined (black arrows). The partial sequence shares 87 % nucleotide and 94 % amino acid sequence with the *Pe*PAE from leaf over the corresponding amino acid and nucleotide sequence.



Annex III *Cs*WH Protein Analysis

A. MALDI ion trace for *Cs*WHb (**1.**), and monoisotopic ion peak lists for *Cs*WHa and *Cs*WHb (**2.**).

2. Monoisotopic ion peak list for MALDI on *Cs*WHa and *Cs*WHb. The 8 most intense peaks from each 100Da window of the spectra are shown. The fragmentation pattern indicates they are the same protein with the same mass ion fragmentation. Masses in bold are identical fragments from each protein, excluding values below 1000 (possible trypsin fragmentation).

***Cs*WHa**			***Cs*WHb**	
515.3151	1429.6218	2255.1758	512.3264	**1733.8050**
523.3062	**1453.7718**	**2283.1754**	515.3212	1744.8261
524.1267	1552.8084	2331.1401	537.3283	**1777.7264**
537.3243	1574.7935	2347.1160	551.4979	**1959.9419**
551.3402	1616.7641	2401.1564	568.1298	**2153.0082**
565.3591	1628.7091	**2402.1571**	579.5325	2183.2633
568.1242	**1634.8532**	2418.0423	580.5346	**2211.1046**
579.5160	1638.7842	**2424.1292**	591.4950	**2233.0835**
607.5511	1656.8495	**2451.2135**	607.5648	**2283.1600**
634.2722	1684.8375	2468.1908	607.7429	**2402.1644**
642.4281	**1686.8081**	2473.1703	619.5265	**2424.1177**
656.0525	**1689.8395**	2483.1619	635.5957	**2451.2244**
662.3029	1700.8268	2509.2595	647.5622	**2513.3211**
663.4272	**1702.8031**	**2513.3166**	662.3150	**2554.1296**
664.3808	1716.7976	2553.1633	663.4374	2648.1254
685.4371	**1733.8026**	**2554.1318**	685.4402	**2662.2403**
710.3350	1746.8645	2568.1444	719.5666	2663.2156
726.5240	1755.7733	2571.2544	726.5224	2678.2256
831.4058	**1777.7198**	2576.1487	747.6069	**2691.2621**
832.3088	1791.7547	2587.1407	825.4734	**2720.2352**
832.4937	1846.9407	2600.1791	832.3114	**2748.2633**
842.5101	1876.8212	2631.2298	842.5100	**2807.2953**
955.5694	1890.8455	2646.1326	931.7913	2944.5102
1004.5130	1907.0572	2660.2185	937.5307	3310.6907
1028.4926	**1959.9501**	**2662.2089**	959.8303	**3324.6818**
1042.5068	1973.9567	2676.2129	**1028.5106**	**3338.6851**
1045.5584	1991.9645	2679.2080	1060.6305	**3348.6160**
1104.6187	2078.0344	**2691.3200**	1076.5854	3350.5190
1144.5141	2104.9984	2706.1706	1118.5859	
1166.4717	2107.3593	2708.3623	1124.5847	
1176.5042	2150.0443	2719.3031	**1144.5200**	
1201.5830	**2153.0111**	**2720.2658**	1305.7239	
1328.6321	2167.0065	**2748.3551**	**1372.6208**	
1370.5987	2169.0130	**2807.2971**	1411.8054	
1371.5921	2175.0035	3309.6949	1430.7616	
1372.6095	2196.0004	**3324.7354**	**1453.7847**	
1376.6257	2210.0264	**3338.7007**	1471.7139	
1386.5995	**2211.1046**	3347.6522	1536.8502	
1388.6036	2222.2123	**3348.5298**	**1634.8536**	
1394.5930	**2233.0895**	3353.7153	**1686.8143**	
1404.6009	2245.1759		**1689.8554**	
1413.6414	2252.0875		**1702.8073**	

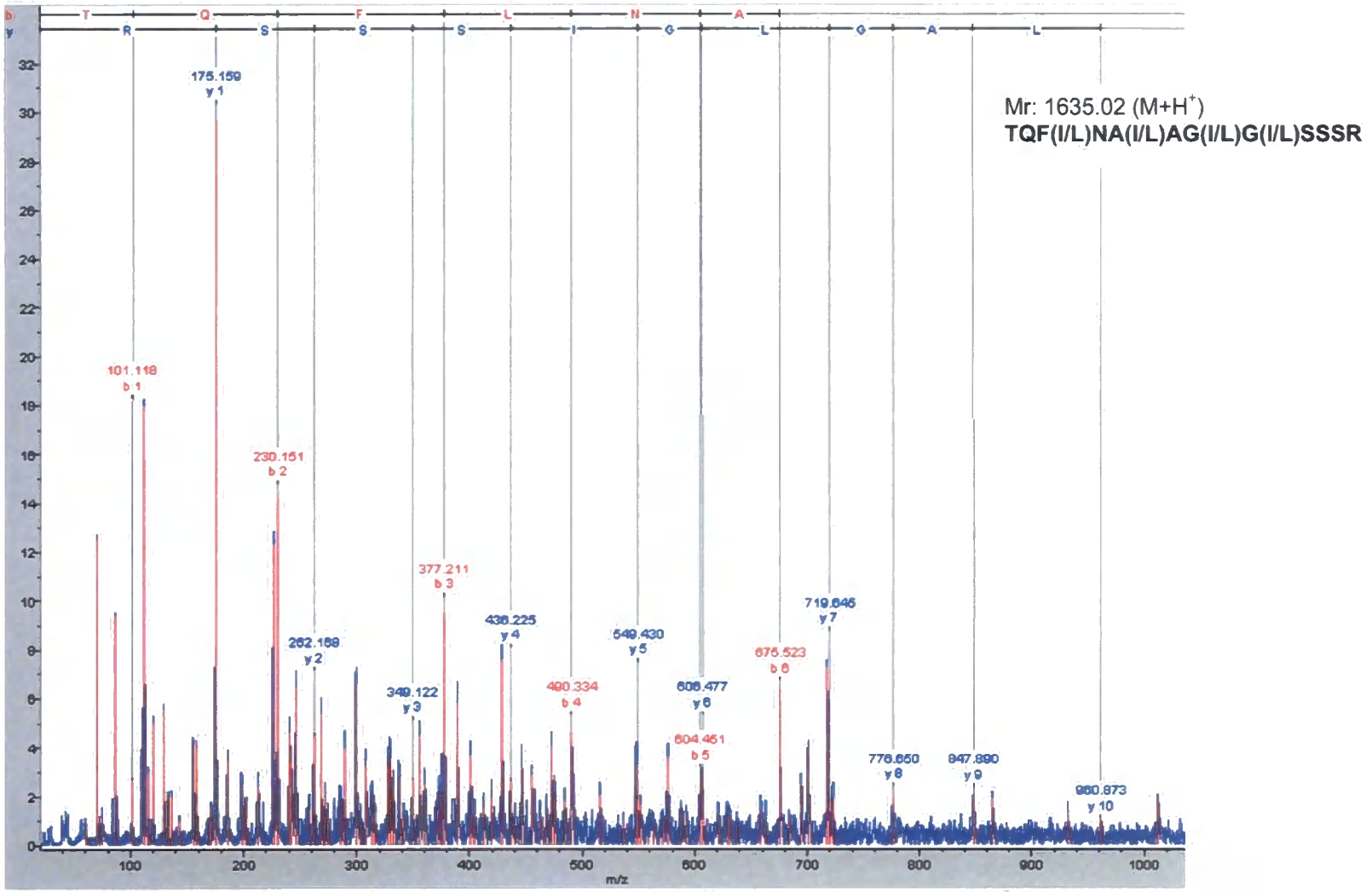
Annex III B. Predicted MS-tryptic digest of constructed *Citrus sinesis* L. Osbeck PAE using the prospector software (www.prospector.ucsf.edu). Considered modifications: peptide N-terminal glutamine to pyro-glutamic acid, oxidation of methionine, protein N-terminus acetylation. Missed cleavages: 1. Mass values 800-3500. minimum peptide length: 5. 20 predicted ion fragments matched with experimental ion masses (red bold), accounting for 76% coverage of the protein.

pI of Protein: 8.5
Protein MW: 43028

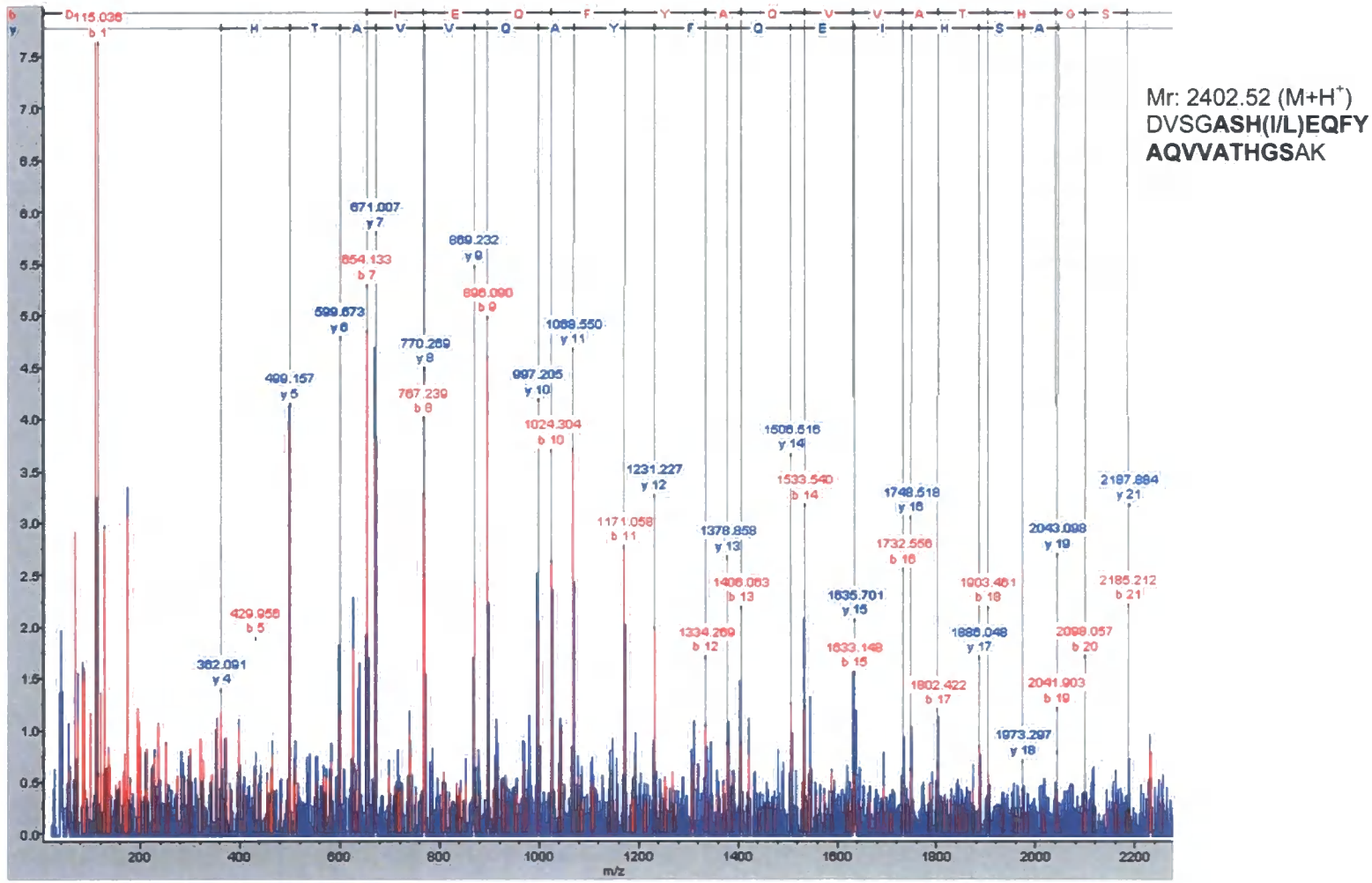
```
1    ESFXYVHSXX LLKADGFNVG ITYVENAVVK GAVCLDGSPP AYHFDKGFGA GINNWLVHIE
61   GGGWCNNVTT CLERKKTRLG SSKQMVKVVA FSGMLSNKQK FNPDFYNWNR IKVRYCDGAS
121  FTGDVEAVNP ANNLHFRGAR VFQAVMEDLM AKGMKNAQNA VLSGCSAGGL TSILHCDNFR
181  ALFPVGTKVK CFADAGYFIN AKDVSGASHI EQFYAQVVAT HGSAKHLPAS CTSRLSPGLC
241  FFPQYMARQI TTPLFIINAA YDSWQIKNIL APGVADPHGT WHSCKLDINN CSPTQLQTMQ
301  SFRTQFLNAL AGLGISSSRG MFIDACYAHC QTEMQETWLR TDSPVLGKMS IAKAVGDWYY
361  DRSPFQKIDC AYPCNPTCHN RVFDSNVHSE V        391
```

m/z (mi)	m/z (av)	Modifications	Start	End	Missed Cleavages	Sequence
816.4462	816.9351		341	348	0	TDSPVLGK
832.4927	833.0244		181	188	0	**ALFPVGTK**
977.5448	978.2059		79	87	1	LGSSKQMVK
993.5397	994.2053	1Met-ox	79	87	1	LGSSKQMVK
1042.5104	1043.1972		226	234	0	**HLPASCTSR**
1059.6561	1060.3330		181	190	1	ALFPVGTKVK
1132.5269	1133.2100		382	391	0	VFDSNVHSEV
1144.5058	1145.2238		354	362	0	**AVGDWYYDR**
1152.6082	1153.3942		88	98	0	VVAFSGMLSNK
1168.6031	1169.3936	1Met-ox	88	98	0	VVAFSGMLSNK
1346.7348	1347.6265		341	353	1	TDSPVLGKMSIAK
1362.7297	1363.6259	1Met-ox	341	353	1	TDSPVLGKMSIAK
1372.6069	1373.4791		101	110	0	**FNPDFYNWNR**
1376.6247	1377.5234		191	202	0	**CFADAGYFINAK**
1408.7617	1409.7010		88	100	1	VVAFSGMLSNKQK
1413.6752	1414.6929	2Met-ox	141	152	0	**VFQAVMEDLMAK**
1424.7566	1425.7003	1Met-ox	88	100	1	VVAFSGMLSNKQK
1546.7723	1547.8273		189	202	1	VKCFADAGYFINAK
1561.8988	1562.9044		1	13	0	ESFXYVHSXXLLK
1613.7859	1614.8147		101	112	1	FNPDFYNWNRIK
1621.8440	1623.0015	pyroGlu	84	98	1	QMVKVVAFSGMLSNK
1628.7605	1629.7858		99	110	1	**QKFNPDFYNWNR**
1634.8860	1635.8720		304	319	0	**TQFLNALAGLGISSSR**
1637.8390	1639.0009	pyroGlu1Met-ox	84	98	1	QMVKVVAFSGMLSNK
1638.8706	1640.0322		84	98	1	**QMVKVVAFSGMLSNK**
1653.8339	1655.0003	pyroGlu2Met-ox	84	98	1	QMVKVVAFSGMLSNK
1654.8655	1656.0315	1Met-ox	84	98	1	QMVKVVAFSGMLSNK
1665.8451	1667.0143		138	152	1	GARVFQAVMEDLMAK
1670.8604	1672.0309	2Met-ox	84	98	1	QMVKVVAFSGMLSNK
1674.7945	1675.9152		349	362	1	MSIAKAVGDWYYDR
1676.7737	1677.8876		31	46	0	GAVCLDGSPPAYHFDK
1681.8400	1683.0137	1Met-ox	138	152	1	GARVFQAVMEDLMAK
1690.7894	1691.9146	1Met-ox	349	362	1	MSIAKAVGDWYYDR
1697.8349	1699.0131	2Met-ox	138	152	1	GARVFQAVMEDLMAK
1697.8423	1699.1194		141	155	1	VFQAVMEDLMAKGMK
1700.8293	1702.0632		235	248	0	**LSPGLCFFPQYMAR**
1713.8372	1715.1188	1Met-ox	141	155	1	VFQAVMEDLMAKGMK
1716.8242	1718.0626	1Met-ox	235	248	0	**LSPGLCFFPQYMAR**
1729.8322	1731.1181	2Met-ox	141	155	1	VFQAVMEDLMAKGMK
1731.8125	1732.9045		354	367	1	AVGDWYYDRSPFQK
1745.8271	1747.1175	3Met-ox	141	155	1	VFQAVMEDLMAKGMK
1777.7157	1779.0856		368	381	0	IDCAYPCNPTCHNR
1795.9225	1797.0302		14	30	0	ADGFNVGITYVENAVVK
1973.9656	1975.2507		268	285	0	**NILAPGVADPHGTWHSCK**

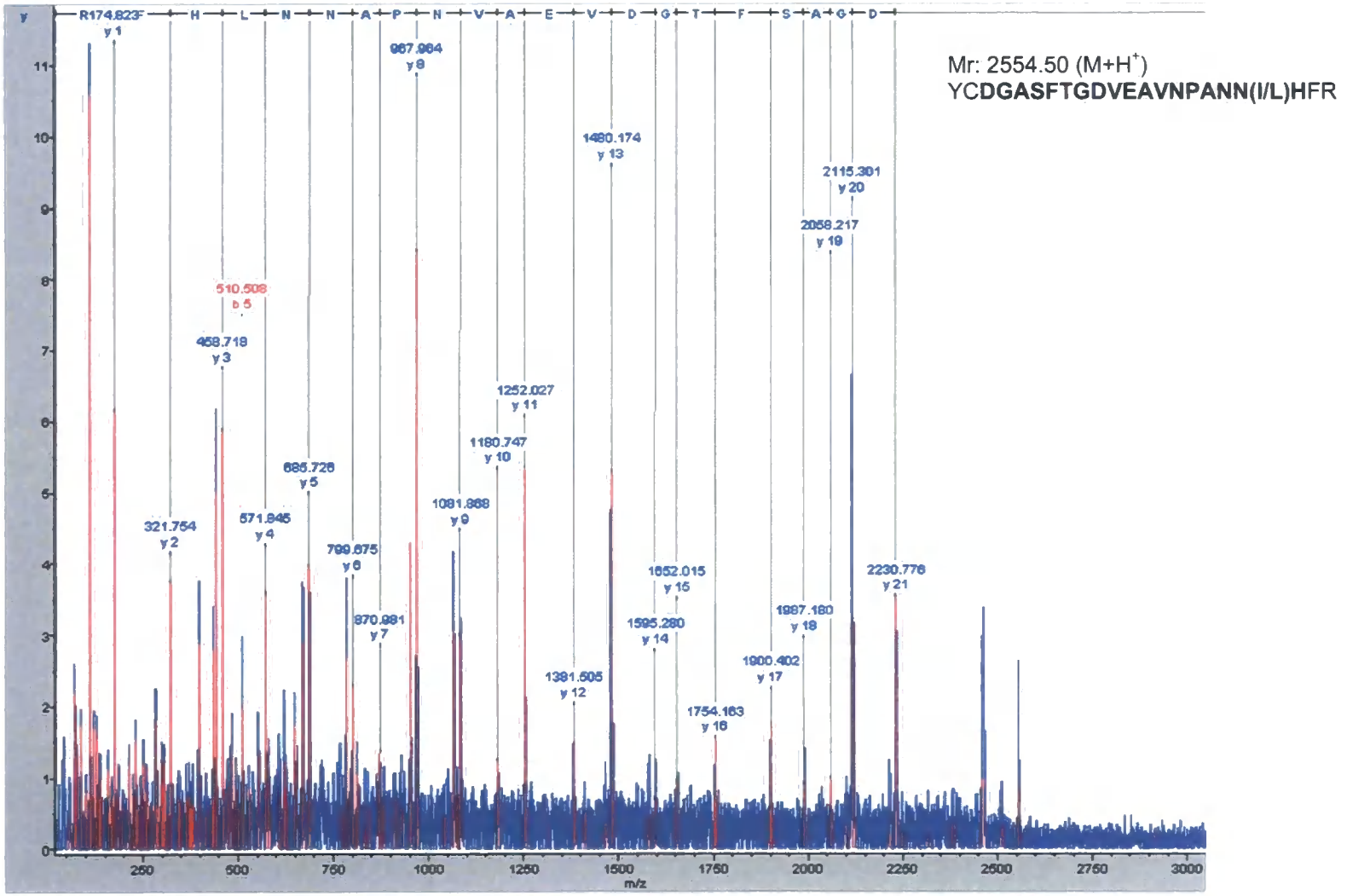
2167.0276	2168.4710		286	303	0	**LDINNCSPTQLQTMQSFR**
2193.9627	2195.5209		363	381	1	SPFQKIDCAYPCNPTCHNR
2205.1590	2206.5624	pyroGlu	249	267	0	QITTPLFIINAAYDSWQIK
2222.1856	2223.5930		249	267	0	**QITTPLFIINAAYDSWQIK**
2402.1735	2403.6296		203	225	0	**DVSGASHIEQFYAQVVATHGSAK**
2533.0767	2534.9341		320	340	0	GMFIDACYAHCQTEMQETWLR
2548.2031	2549.8644		156	180	0	**NAQNAVLSGCSAGGLTSILHCDNFR**
2568.1578	2569.7893		115	137	0	**YCDGASFTGDVEAVNPANNLHFR**
2582.2465	2584.0780		226	248	1	HLPASCTSRLSPGLCFFPQYMAR
2598.2414	2600.0774	1Met-ox	226	248	1	HLPASCTSRLSPGLCFFPQYMAR
2691.1464	2693.0926	1Met-ox	320	340	0	**GMFIDACYAHCQTEMQETWLR**
2720.1650	2722.0275		368	391	1	IDCAYPCNPTCHNRVFDSNVHSEV
2752.2896	2754.0315		113	137	1	VRYCDGASFTGDVEAVNPANNLHFR
2781.2798	2783.0296		115	140	1	YCDGASFTGDVEAVNPANNLHFRGAR
2864.3600	2866.2896		153	180	1	GMKNAQNAVLSGCSAGGLTSILHCDN FR
2880.3549	2882.2890	1Met-ox	153	180	1	GMKNAQNAVLSGCSAGGLTSILHCDN FR
3017.4145	3019.4114		47	74	0	GFGAGINNWLVHIEGGGWCNNVTTCL ER
3145.5095	3147.5866		47	75	1	GFGAGINNWLVHIEGGGWCNNVTTCL ERK
3330.5050	3332.8464		320	348	1	GMFIDACYAHCQTEMQETWLRTDSPV LGK
3338.8035	3340.9118		1	30	1	**ESFXYVHSXXLLKADGFNVGITYVEN AVVK**
3346.5000	3348.8458	1Met-ox	320	348	1	GMFIDACYAHCQTEMQETWLRTDSPV LGK
3354.6284	3356.7242		203	234	1	DVSGASHIEQFYAQVVATHGSAKHLP ASCTSR
3361.6780	3363.8661		156	188	1	NAQNAVLSGCSAGGLTSILHCDNFRA LFPVGTK
3362.4949	3364.8451	2Met-ox	320	348	1	GMFIDACYAHCQTEMQETWLRTDSPV LGK
3453.6784	3455.8951		14	46	1	ADGFNVGITYVENAVVKGAVCLDGSP PAYHFDK
3702.7646	3705.1256		191	225	1	CFADAGYFINAKDVSGASHIEQFYAQ VVATHGSAK
3711.8581	3714.2403		286	319	1	LDINNCSPTQLQTMQSFRTQFLNALA GLGISSSR
3727.8531	3730.2397	1Met-ox	286	319	1	LDINNCSPTQLQTMQSFRTQFLNALA GLGISSSR
3832.9594	3835.5537		235	267	1	LSPGLCFFPQYMARQITTPLFIINAA YDSWQIK
3848.9543	3851.5530	1Met-ox	235	267	1	LSPGLCFFPQYMARQITTPLFIINAA YDSWQIK
3979.9000	3982.5392		268	303	1	NILAPGVADPHGTWHSCKLDINNCSP TQLQTMQSFR
3995.8949	3998.5386	1Met-ox	268	303	1	NILAPGVADPHGTWHSCKLDINNCSP TQLQTMQSFR



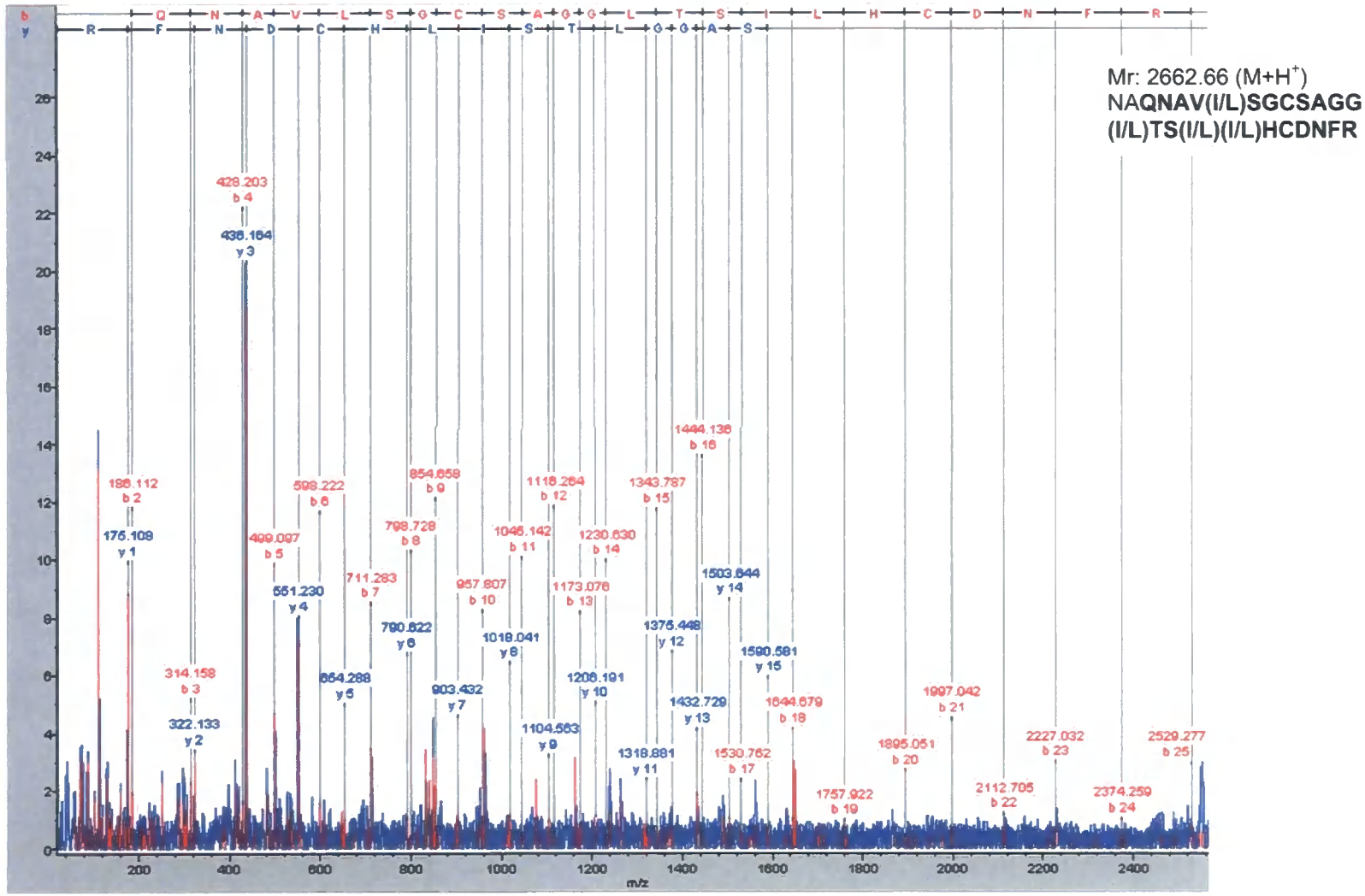
Annex III C. MS-MS data for *Cs*WHa. Trace for ion 1635. b and y ions are shown with predicted amino acid sequence.



MS-MS trace for ion 2402. The amino acid sequence in bold was determined with considerable certainty (i.e. b and y ion overlap).



MS-MS trace for ion 2554.



MS-MS trace for ion 2662. Software: FlexAnalysis 2.4 and Biotools 3 (Bruker Daltonics).

Annex IV Passion fruit cDNA sequences identified in random library sceen. 30 clones were selected at random from each library (M: mesocarp, G: gland) and plasmid inserts sequenced from the 5' end using the pTriplEx2Forward primer. Non-coding 3' sequence is highlighted in lower case.

M1
ATGACCACCGTCCACTCAATCACTGCAACTCAAAAGACTGTTGATGGGCCATCAATGAAGGATTGGGGAGGTGGAAGAGCTGCTTCA
TTCAACATCATTCCTAGCAGCACTGGAGCTGCTAAGGCTGTTGGGAAGGTTCTACCTGCGTTGAATGGGAAACTCACTGGAATGGCT
TTCCGTGTTCCTACTGTGGATGTCTCCGTAGTTGACCTCACCGTCCGACTTGAGAAGGGAGCTAGCTACGATGAAGTTAAAGCTGCT
ATCAAAGAGGAGTCTGAGGGAAAGTTGAAGGGAATCTTGGGCTACACCGAAGATGATGTGGTTTCCACCGACTTTGTCGGTGATAGC
AGGTCAAGCATCTTTGATGCTAAAGCTGGAATTGCTTTGAGCAAGGATTTTTTGAAACTTATCGCCTGGTACGACAACGAATGGGGT
TACAGCACACGTGTGGTCGACTTGATTGTCTACATTGCCTCCGTTTCGAAGTGAgcacttgttatcgtctttccttccacctggtta
tggtatgagctgaatgaaacagtaaaataaaattgctgcgggacagtggtgtttttaataacaatttaaagcgagttgtctttagtt
tcttaggtttgggttctcctgtttctcggggatgctgaaacccctcgaacttttatttttattttccgtcaattttctctgnnctgnt
actaatcatatcccttctttagcgtccgaagtatgttcccccacacaantancttacntaanaatnatccntngttggcccgccctc
gggcccnttcnantttaaactcgagccaagcttttgcatgcggccccaattcgagctccttgggccaattcccc
M2
GAGGATGGAATTAAAGAGCCACAGGCAACATTCTCGGCCTGCTTTGGTGCAGCTTTTATAATGATGCATCCCACCAGGTATGCAGCT
ATGCTCTCCGAAAAGATGCAGAAACACGGAGCGACGGGATGGCTTGTCAATACTGGCTGGTCAGGCGGAAGGTATGGATCAGGCAGC
CGTATCAAGCTTGCGTACACTCGGAAGATTATCGATGCCATACATTCTGGGAGTCTTTTGAATGCCAACTACAAAAAGACTGAAGTG
TTTGGACTTGAGATCCCGACTGAGATTGACGGGGTGCCGTCTGAAATCCTTGATCCTATAAATACTTGGTCGAACAAGAATGCTTAC
AACAACACCCTGTTGAAGTTGGCTGGATTGTTCAAGAACAACTTTGAGGTGTTCACAACCCACAAAATTGGAAAAGACAACCAGCTG
ACGGAGGAGATCCTTGCAGCGGGTCCTAATTTCTAAtcaaagtgtGctgatattgcaagagagtgtgtacc
M3
GGGAGGGTGATTGTGANAGAAGAAAAGGGTGGNTCTGTTTCTTATGCTGCNCCGGCTGGATCAGAATTGGAGAACCCATTATGCGTG
TCCCAACCCTTGCAATTCACTTGAACAGGACTGTTAACAGTGAANGGTTTAAAGTGAACACCGAGACTCATCTTGTTCCTATTTTGG
CAACGTCAATTANGGCCGAACTGACTANAGTGGTTGCTGAAANTGCTCCNCTTGGAANTGATGATCAGTCTGATGGAAAGAANATTA
ACANGGANACCNGCTCAGAACNTCNCTCTATTCTTCTTCAAATGATTGNNAGTCANCTTGGCTGTGAACCANATGATATATGTGATT
TCGAATTGCNNGCTTGNGACACTCAACGAAGTACTATAGGTGGTGCTGCANTGGAATTTGTTTACTCTGGAAGGCTTGATAATCTCT
GCATGTCATTTTGCTCTTTGAAGGCCCTANTAGATGCTACATCTTCTGAAAGGGATCTTGACGAGGAAGCTGGTGTCANAATGGTTG
CATTATTTGATCATGAGGAGGTGGGATCTGATTCANCACNAGGAGCCGGGTCTCCTGTAATGTTCGATGCTCTGTCACCAATCNCNN
GTTCCTTCNNCTCAGATTCTAAGCTGCTGCAAAAAGCCATCCAGAGGAGTTTCCTTGTCTCTGCCGACATGGCACATGCAATACATC
CTAATTATCCGGACAAACACGAGGATAATCATCNGCCCACGATGCACANAGGACTTGTAATAAANCNTAATGCNANTCAACGCTATG
CNACCNATGCTGTTACTGCCTTCATATTTGGGGAGATANCAANAAANCNTANCCTTCCCNCCCCAANATTTTGTGGTCCGAAACNAC
NTGCCTTGNGGTTCACAATTGGNCCTATTC
M4
ATGGTTTTGCTAGAGAAGTTGTGGGATGATGTTGTTGCTGGACCTCAACCTGAGCGTGGCCTTGGCAAGCTCAGAAAGATCAGCACA
AAATCTTTGAACATCAAAGATGTAGAGGTTGGAGAAGGAAGCAACTACCAGAGGTCTCTCTCCATGCCAGCAAGTCCGGCCACACCA
GGGACCCCGGGAACACCCGCCACGCCCTTGTCAGCGCGCAAGGAAAATGTGTGGAGAAGCGTGTTTAATCCTGGCAGCAACCTTGCC
ACCAGGAGTATCGGAGCTCAGGTGTTTGACAAGCCTGCACACCCTCAGGCCCCTACTGTCTATGACTGGCTTTACAGCGGCGAGTCC
AGGAGCCAGCACCGCTGAagtggcaagccaggttgagctgaggtgccatgtgtaaataatgtggttttggttctgtctatgctattt
gtcccctctgttcaatgggtgctatgactactacgtttttcttcaacatgccatcaacaaagatgactagtttccttggagcttttg
gagtgttcaataattctatggtgttaaatatttgtcccagatagcttgatctgtgttgatcctgatctgcagtggacaagtgtgatg
agtctagtatagtatatctagtacaaataatctgtattcggtcctaccgaccgaccgtccttaatgtagtttatccttcctttttctc
c
M5
ATGGCGAGTGATTGTGAATCTAAGAGGAGCTCTGAGACTGTGAAGAAGTTTCTGTGCAGTTACGGCGGTAAAATTCTTCCTCGCTCC
GGTGACGGCAAGCTCCGTTACGTTGGTGGAGTAACTCGTGTTCTCGCCGTCGATCGTTCCGTATCGTTTGCTGAACTGATGGTGAAG
CTTGGTGAGTTCTGTGGATACTCCGTGGATTTGAAGTGTCAATTACCTGATGGAGATTTGGAGACGTTGATTTCTGTAAAATCTGAT
GAAGAGTTGGTGAATTTGATTGAAGAATACGATCGATCCTCCCCTGGTTCAAAAATCAGAGCCGTATTAACCCCGCCGAAATCGCTG
AAGTCGGTTTCTCCTCTCCCGTCGTCTCCGACGAGTGTTAACTATATTACAACCAGATCTCCGAGCAGCTCCCCCGATTACCGTCGA
TACCGGAATGATTCCCCGCCGCTAGATTATTCCTTCGGCAGGCTTCTGAATGATTCCCGTCGTGGAATGGGGAATCACAAAGCTCTG
TGCTGCGGTCCTCGCTGTTGCAACTACTACTGGTGATCGATCGGATGTGCTTGAGACTTAAAAAGGCCCCAAAAAATCTANGGGTAT
TTCNCCAGATAGAATAGATTAGCTGGATAAGAAATAAcgaacaaaagaagtatattgggatgatttccc
M6
ATGGACTTTTTATTCAGAGGTCTGAATGAGGAATCTGCAAACTCTCATGTTGACATTCATAGGTGTCCCTTTTTGAGGAACATCAAT
GAGCCAACCAACTTCTCTTTCTCCTCTTCCATGCCTTTTCCTATGCCTGTACGGGCAGGAAAAGGTCCGATTTTTGAAGATGGTCCT
AATTTTGATATGGCATTTAGGTTATTCCATGGAAATGATGGAGTTGTACCGCTTTCTGAAAGATCACTTTCGTTTTCTGAGAAAAAA
CCCCAAGCAACCGCACCTGCGTTTAATCCTTTAGCTGCCAAGGCAGCCACCATCAGTCTTTCATCTTTTGGACTTGGTGGTTTCTTC
GGCTTTGACTCATTTTCTAAGAAGCGGCTGAATGAAAAAAGAAATGGCAAGTCGTCCAATAAAGGATCGTCTTCCAAGGGAGGAGGC
CCTAATCACGAGGCATTGAGCGATGAATGGCTGAAAAATGGGAACTGTCCCATTGCCAAATCTTTTCNGGCAGTTAGCCATGTCTTG
GCACTTGTTGGCAAANGCTCTCCAACCCCCCTCAAGGNATGGAAACTGAAAGTGGCCCACCTGCANTGGGTTGCAAGCAAGAGCNAC
CCATATCCGCGAACTGCTTTTNCNAAGAAACTTCCGGCCTCANCCTTTGGCCTGNAAAAAGTGGTTTGTGGAAAAGGAATGGCTTNG
GGCATGGCAGNCNAAACATACCGTTAGggaaanttggagannaacatacccgaggaaaattttcanncgcccatnggttccnctgca
agg
M7
GGTTCTTTCGGGGATTGTGAAGAAGGAGGATGTGAAACTTGTACCGGCAAGCTTGTAGTTGGGTCCGCTTGGCGTAAACAGATGCCT
AAGCTTGCTGTTTCGCTCGGTTTGGGTGATGAAATGCCCGATATGATTAAAGAATTAATTAGTAGGGGACAGCAGCTTGATGCGGTG
CACTTTACTTATGAAGTTGGACTTGTGGATATGTTTCCTCCAGTCCCATTGCTAAAATCCTTTTTGAAGGATGCAAGGAAGGCTGCA
GCTTCTATTATGGAGGATCCTGATAATAGTGGTCGAGCTGCGGGTCTTGCTGCTCGTAAGGAGCAGTCAGCACTCAGGGCTGTCATC
AAATGCATTGAAGAATACAAATTAGAGGCTCTGTTCCCTCCAGAGAACCTTAAGAAACGACTTGAACAGCTAGAGAAAGCCAAGACA
GAGAAGAAAAGACCTGCCGCAGTACCTGCCAACAAACGAACCCGTGCCAGCAATGGTGGCCCAATGCCTCCAGCTAAGGCTGGTCGT
TTGACAAATGCATATGTTTCATCTTTCCCTGCGCCTCCCACGTTTGTCAGGTCTCCTTCACAAACACAATACATGCCAGCACCACCA
CCATATCCTTCTCCCCCCGCAGTTTATGGAAGCAGGAGCCCACCATCCCCATATGCCTACTCACCGGAGGCTGTACCTGCTCCCATG

GCCGGATCGTATCCTTTGAATTATCAGGCATATGGAGGCTATGGCAATGGTTTTGCACCAGCTTATCAGCAGGCTTACTACCGATAG
acatgacaactcaatgatggtaaccactgatatgatgaccccaatctcattttgctttcccanctggtttgttttgtaancactaac
cgttttaacnggtagcaatgtcgggtgtgtttatttataactgtcctacttcttgcng
M8
ATGGTGAAGTATTCGAGAGAGCCAGACAATCCCACCAAGTCCTGCAAAGCCAGGGGCTCTGATCTCCGTGTTCACTTCAAGAACACC
AGAGAAACTGCCCATGCACTTCGCAAGTTGCCTTTGGTCAAGGCTAAAAGGTACCTGGAAGATGTTATGGCCCACAAGCAAGCTATT
CCATTCCGCCGTTTCTGTCGTGGTGTTGGGCGAACTGCTCAGGCAAAGAACCGTCATTCAAATGGACAAGGTCGTTGGCCTGCTAAA
TCTGCAAAATTCATTCTTGATTTGCTAAAGAATGCTGAAAGTAATGCTGAGGTCAAAGGTTTGGATGTGGATGCACTCTACATATCT
CATATCCAAGTTAACCAAGCACAAAAACAACGTCGCCGCACATACCGTGCTCATGGAAGAATCAATCCTTACATGTCCTCCCCGTGC
CACATTGAGTTAACTTTGTCAGAAAAAGAAGAGGCTGTGCAGAAAGAGCCCGAGACTCAGATTGCTCCCAGGAAACAAAAGGGTAAT
GCTCTACAGAGCGGTGCTTCCTCTTAAatttgagcagcagatcaatgctggataaaaaaggatccnccgactgaaaacatttttgt
ttttccgtacctgnttnnaaccaccaatttgaacctggttctngggaataaatggggttttttttttttcccccttttggccnt
ttanatggggaaatttggagantccttnttt
M9
ATGGCCGGTAAAGGAGAAGGACCAGCCATCGGAATCGATCTCGGCACAACATACCCCTGCGTCGGAGTATGGCAGCACGACCGAGTG
GAGATCATCGCCAACGACCAGGGAAACCGTACGACGCCTTCTTATGTTGCTTTCACCGATAGCGAGCGTCTTATCGGCGATGCTGCC
AAGAACCAGGTCGCCATGAATCCCATCAACACCGTCTTCGACGCGAAGAGGTTGATTGGTAGGAGATTCAGTGATTCCTCAGTACAG
AGTGACATCAAGTTGTGGCCATTCAAGGTCATTCCTGGACCTGCTGACAAGCCCATGATTATTGTCACCTACAAGGGTGAGGATAAG
CAATTTGCAGCGGAAGAAATCTCTTCCATGGTTCTCATTAAGATGCGTGAGATAGCCGAGGCTTACTTGGGCACTACTATTAAGAAT
GCAGTGGTCACTGTCCCTGCCTATTTCAACGACTCCCAAAGGCAGGCCACCAAGGATGCCGGTGTGATTGCTGGCCTTAATGTCATG
CGTATCATCAACGAGCCCACTGCAGCTGCCACTGCTTATGGTCTTGACAAGAAGGCAACCAGTGTCGGTGAGAAGAACGTCTTGATC
TTCGATTTGGGTGGAGGCACTTTTGATGTCTCTCTTCTTACCATCGAAGAGGGTATCTTTGAGGTTAAGGCCACAGCTGGTGACACC
CATCTTGGAGGTGAGGACTTTGACNACNGAATGGTGAACCNTTTTGTTCNGGAATTC
M10
GGGAGGCTGTTTCAGTATTGCCTCAACTCCGAGTCATGCAGAACGCAAAGATTGTTGAACCATTCACTGCTCATTATGTTTTTGCGC
TTGGAGTAGCTAGGTTCTTTAGCTGTGCTCATTGGGTTCTCCAGCTGTTAGATAGTCGGGGACTTTTATTGACAGCACTGGGTTATG
GATTGTGGCCTCCCATGGTCCTCCTCTCAGAAATAGTCCAGACATTCATTCTTGCAGATTTTTGCTACTACTATGTCAAGAGTATCG
CTGCTGGGCAACTTGTTCTGCGACTCCCCTCAGGAGTGGTGTAAggtcaaaaacatgaaggcagatatggcttacttcaccttccac
acgccaacttctcaaggctagctgtacacgcctgagctgccacttaagagttactgtagcagctagccatcaaaggctgcaggtgcc
atggcattttaaagtgtataaaaaccaatctagtgtactccttaatttgacttggaaaatttgtcttccaccggttttagtaccatt
ctgcaaagacttcttatttcattacaaaatccgtacaaagtctcgagtttttt
M11
ATGGCTTACAAGACCATAGTCGCACTGATTCTCTTGTCTCACCTGGCCAGTTCGGCTCTCGCTCAGGCTCCTGGAGCTGCGCCGACC
AAGTCCCCTCTTTCTCCTGCACCAGCTCCGAAAAAGGCCGCAACTCCACCTCCGGCACCAGCTCAGGCTCCCGCTCTCGCACCACCC
GTCTCTGCTCCTGCAGACGCTCCCAGCGGCGTACCATCCCCTGCGCCGACCAGCTCGCCACCATCTCCGGTCGCTCCTACAACCGGT
CCTTCTGCATCTACTACTCCAGGCTCTGTGGCCTCTACTCCAGGCTTGGCGCCAACCAGTACTCCTCCTTCAGGCGGTGCCTTCGTG
AACGGTGCAGCTGTTTTCAGCACTCTGGTCGGAGCTGCGTCCGCGTTGTCTTTTTTATTGCTGTAGatctgagcagtatcgggcttt
aggttttgattttactaggatattgagtgtttgatttatttttttgtttcttgtccatggacggtggtggtggtgacttttcgata
tgttgacggttccgatttgatatcgctctcttaggtagatcattgttatcgctatttttacctc
M12
ATGAAGGGCGGAAAATCGAAGGCCGGGATCAAAGTTACCGATGCTAAGCTAAAGAGGAAAGGTGCTGGAGCCAGCACGAAAGCGGCC
AAAAAGGCTGCAAAGGACCCTAACAAGCCAAAGAGGCCTCCTAGTGCTTTCTTCGTTTTCATGGAGGAGTTTAGGCAGGAGTTCAAG
GCTAAGCATCCTACCAATAAATCTGTAGCTGCTGTTGGTAAAGCTGCTGGTGATAAGTGGAAGTCATTGTCAGAGGCTGAGAAGGCT
CCTTACCAAGCTAAGGCGGAGAAGAGAAAGGCTGACTACAACAAGAATTTAGCAGCATACAACAAAAAATTGGCTGATGGAGGCCCT
GCTGATGATGAATCTGACAAGTCAAAATCAGAAGTGAATGATGAGGATGATGAAGAGGAGAGCGGTGGAGAGGAGGAAGAGGATGAC
GAGTAAttgactagaactgcaaagggtgcaaggggccaaaggagcacgttgatgtagattgcttaggatgttgtacttgcatttcta
attcaaatagtttatctggccatttttagtttaggttaggattaaaacaaagattagttttgggnaacacttcctgaagctcatatgg
gtaatgtacttaatgtctttttttaaaggcacaaanatnntnaangannnnnanntnccnttgtcggccncccctcggcccc
M13
ATGGCCCGTACCAAGCAAACTGCTCGTAAGTCTACAGGAGGAAAGGCACCAAGGAAGCAGCTCGCCACCAAGGCTGCCCGTAAGTCT
GCCCCTACCACCGGTGGGGTTAAGAAGCCCCATCGCTACCGTCCAGGGACTGTTGTTCTTCGTGAAATCCGCAAGTATCAGAAGAGC
ACCGAGCTGTTGATCAGGAAACTTCCGTTCCAGAGACTTGTTCGTGAAATTGCGCAGGATTTCAAGACTGATTTGCGTTTCCAGAGC
CATGCTGTGCTGGCCCTGCAAGAGGCAGCCGAGGCGTACTTGGTGGGTTTGTTCGAGGACACTAACCTCTGCGCTATTCATGCCAAG
CGTGTCACTATAATGCCCAAGGATATCCAGTTGGCCAGGAGGATCAGGGGCGAGCGCGCTTGAggtttataataatgctgtgactat
ttcttttggtcctttcccttgttgggggttatcagtcggatgcgtctagcaagatagattagatagatagagatacttctggtcaat
tagctatgttgggtgctgacggtttgtaccaatctataaattcggaattttcttgtaatttatattagctctgctgttttgtttgaa
cacgacctatttttg
M14
GGATCCTAAAGGGATTTAAGCCTTGAGTTCTTCTTTGACTCGAACCCATACCTTCAAGAATTCTGTCCTTACAAAGACATATCACAT
GATTGATGAAGATGAGCCTATTCTCGAGAAAGCCATTGGGACGGAGATTGAGTGGTACCCAGGGAAATGCCTAACGCAAAAGCTTCT
TAAGAAGAAGCCTAAGAAGGGATCAAAGAGTGCCAAGCCAATTACAAAAACCGAGGAGTGTGAAAGTTTTTTCAATTTCTTTAATCC
ACCTCAAGTCCCGGAGGATGATGAGGATATTGATGATGACACCGCGGAAGAACTCCAGAATGAGATGGAACAAGATTATGACGTCGG
GTCTACCATTCGAGACAAGATCATTCCTCATGCTGTGTCATGGTATACCGGGGAGGCCATGCCTGGAGATGATCTTGATATTGATGA
TGAGGATGACGACGACGATGATATTGATGATGATGANGATGANGATGATG
M15
GTATAATGTTCGTTTTGCACCTCGCAATGGAACGGGCGTCGTTGTTATTTGCCCCACAAGAGAGCTTGCCATACAGACACATGCTGT
GGCAAAGGACCTTCTCAAGTATCACTCCCAGACTCTTGGCTTGGTTATTGGAGGCTCAGCTAGAAGAGGAGAAGCAGAACGCATAGT
CAAAGGAGTGAACTTATTAGTTGCTACCCCTGGTCGACTTCTTGATCATTTACAAAATACCAAAGGGTTTATTTATAAAAATTTAAA
GTGCCTCATGATTGATGAAGCAGATAGGATCTTGGAGGCCAATTTTGAAGAAGAAATGACACAAATTATCAAGCTTCTACCCAAGAT
GAGGCAAACTGCTTTATTTTCAGCTACTCAAACTAAAAAGGTTGAGGACCTTGCCCGCTTGTCGTTTCAGACAACTCCTGTCTATAT
TGATGTGGATGATGGAAGAAGAAAGGTCACCAATGAGGATCTGCAGCAAGGCTATTGTGTTATTCCAAGTTCCAAGAGATTTATTCT
TCTCTACTCTTTCTTGAAGAGAAACTTGTCAAAGAAAGTTATGGTCTTCTTCTCTTCATGCAACTCGGTCAAAATTTCATTCTGATT
TACTCAGATACATTCAGATTGAATGCTTTGACATCCATGGAAAACAAAAGCAGCAGAAACGGACAAACTACATTTTTTGATTTCTGT
AAAGCAGATAAAGGGGATCTTGCTGTGTACTGATGTTGCTGCCCCGTGGACTCNATATTTCCTGTTGGNGGACTGGGATTGTGCAGT
ATGATCCCTCCCCGGATGNAACCCAAAGGAATACATTCACAGGANTCCGGCCCGGACAGCNTCGTGGGGGGAAGGTGCCAAAAAGGG
AAATGCGCTTCTCTTT

M16
GTTTGAGGTGTGGAAGGAGTCACAATAGATATGGCTCAAAACCAGGCTAACCATAAAGGGAATAGTTGAACCTCAAGCAGTTAGTAA
CAAAATTATGAAGAAAACCAAGAGAAGAGCCAAAGTTCTATCTCCATTGCCTGAAACCGAGGGCGCGCCAATGCCCGAAGTGGTCAC
TTCACAGGTTAGTGGAGCAACGACTGTGGAGCTTGACGTGAACATGCACTGTGAGGCCTGTGCTGGGCAACTCAAGAGAAAGATACT
CAAAATGAGAGGAGTGCAAACAGCTGTGACAGATCCAAGTGCAGGCAAAGTCACCGTTACTGGGACAATGGACGCGAACAAGCTGGT
GGACTATGTCTATAGGCGGACAAAGAAGCAAGCCCGGATTGTACCACAGCCCGAGCCTGAAGCAGAAAAGCCAGAAGAAAAACCAGC
CGAGGAACCGAAACCT
M17
ATGGCGAGATCGACTGGCAAGGATGCGCGAGCTCTGTTCCATTCTCTTCGTTCTGCCTATGCTTCCACACCCACTTCCCTCAAGATC
ATCGATCTGTACGTGGGTTTTGCGGTTTTCACTGCCCTGATTCAGGTGGTTTACATGGCGATTGTTGGGTCATTCCCATTCAACTCG
TTTCTCTCAGGAATACTTTCCTCCGTTGGGACTGCAGTGCTCGGTGTCTGCCTCCGAATCCAAGTGAACAAAGAAAACAAGGAATTC
AAGGATCTACCACCAGAACGTGCTTTTGCTGATTTTGTGCTCTGCAATCTGGTTCTGCATTTGGTGATCATGAATTTCCTAGGTTAA
gctgatgcaatgctttaccttatatgtaacttttcagacagttttgttatgtcattttatgccatcaggcttaatttcccaaattag
gacaagtatctacaaaccggttttgactgtattcatgaatttgttttcctatcanggaaacaaatatcctcc
M18
ATGGCTTCTTCTGTTTGCTCCCATCTTCATCTTCTCCTACAAGTTTCTCTTTTGCTCTGTCTGGTTCATTTGAGTTCTGCTTCAAGA
GAACTAGCTGAGTCAGACCCAACCCAACAACCATTTCTGTTCCAATACCACAACGGTCCTCTTCTAACTGGTGACATCTCAATTAAC
TTGATCTGGTACGGCAAGTTCAAACCTTCTCAACGTGCTATAGTCACGGATTTCGTTGCTTCAATTTCTTCTTCTAGACCGCCAGCA
ACTTCAGCCCAACCCTCCGTTGCCAAATGGTGGAAAGGTATAGCAAAATTCTACCACCTGGCTAAGTCCAAGAAGTCATCTTCTCTT
CTTCTGTCCTTGGGAACACAGATCCTGGAGGAGAACTATCCATTGGGGAAGTCCCTTTCAAACAAGCAAATCGTGCAATTAGCATCA
AAGGGTGGTCAAAAAAATGCCATCAATGTTGTCTTGACAGCTTCTGATGTTGTGGTCGAAGGGTTCTGCTCCAGTAAGTGTGGAACT
CACGGGTCTGCTGCAAGTCGCACCGGCAAGAACTCCAAATTCGCTTACGTTTGGGTGGGTAACTCAGAGATTCAATGCCCTGGTCAA
TGGCGCATGGCCATTCCACCAGCCAAATCTATGGGACCACANGAGTCCACCATTGGGTTGGCACCAAAACAACGAATGGGGGGTAAT
NGGAATGGNNATGAATTATTAAAACCCTGGGGGTAGGNCCNTTTTTGGGCANGGGNAACCGGCCCACCCAAANCCCAATTTGGGGAA
AAATGGGCTTANTTTTCCANGGGGTNCCGAAAAAAAAGGCACCT
M19
ATGGCGAATAATAATCGATCAGTCTCCACAAGCGAAGTCGTTTACAGATTGCAGCTAGAGCTTCTCGATGGGATAAAGAGCGAGGCC
CAGCTCTTTGCCGCTGGAGCCATGCTGTCGCTCAGGGACTACCAAGATGTGGTGATCGAGCGTACCATCGCCAACCTCTGCGGTTAC
CCGCTGTGTCCTAACTCCCTGCCCCCTGAGAGTAACCGGAAGGGACACTACCGTATCTGGCGCAGTGAGCATGAGATCTCTGATTTG
CACGAGGTTTACCTGTACTGTTCCCCGACTTGCCTGTATAAGAGTAAATCGAGCTTTAAAAAGATGAGGAGCCCGGTTGAGGTTGGG
GAATTGAGTTTGGAGAGTG
M20
GGATGGGTTTTTACATGGAAGAAAGATGGATAAAAAAGTGGAAGAAGATGGAACATTTCATTTCTGTCAATGCCAGTTTATATCAAG
AGATGGAAATGCTTGCAGATCTGGAACAGACTGTGAGGAGAATGAAGTGTAATGATCCCGAGCCTGATAATTTACTTGACTATCAAA
AGAAGCTTGTCTGGAAACGGCAGGAAGTGAAGAATCTTCGGGATATGTCCCTTTGGAACACGACTTATGATTATACTATTCGGCTTC
TAGCGAGATCTTTATTCACAATATACTTCAGAATCAATCATGTTTTTGGAATTGCTCCGATGATTAATCCTGGGCACTCTAAAAATC
TGAATTCTGATTATATACATCGCAGCCAGTCAGTGTCTGCATTACTGCAGTCTTCGGTTCATCCATCTAAGAATAGTAGCATTCCCA
GATTTTCCTCTGGCCCCCTTGGGAAGTTTGCCACAAATTGTGATCCAAACTCCAGAACTAATAACACCACCAATTTCCATTCAGGTC
CTCTTGGTGGCTCAACCACAAAGACTGGTCCTATCACTGGAAAGCATAGGAACTTTAACTTCTTTTCAGGCCCGCTTGGGAGACCCA
CAACAAAGTCAGGTCCAATTTCTGGAACAAGTAAAAGCAGCAAGAAGATGTGGCAAATCCCTCAATCATCCACTCTTCTGGGAAAGA
AATCACATTCGAAGCAGAATCGGNTGACTCAAGTAGGACCTTTTAAAGGAGCATGGGAGCTTCCAATAGCTTCACCTGGTGGCCAAT
GGTTATTTNAATTCAACTGGANGTCCATTCCCAANTTCTCCTCCGGATCCANGGAAGGCCNTNTGGGACTACATGCCTCATTCTACT
CCATCAATCCTCCGNTCCNAANCANAAAATTTTTGGAAAGGGNGTTCC
M21
ATGTCAGGTTGGGACTCCCCTGTCTCAGATCCTAAATCAGGGAAGTATAGAAAGAGTTGGTCATTCACCAAGGAGGAGATAGAATCT
TACTGGAAACTGAAGAAGAAAATCGAGGAGGAACATCTGAAGGCTATTTCCACTCCATCTGATGGCACCAAGGGTGGCCAGGAGGAG
GATGGTGGGGCAAAGTTACAGAGATCAAGCTCCGTGCCTGCTGCCGAAAGCAAGACGGGCTTCATGGACAAGGAAACCAAAGCAAGT
CTGCAACAAATCATGAAGAAAAGTGACTGGTGGACCAGCAGCAGCTGGGCGTTTCTGAATGAACCTCCAGTACTGGACCGTCCTTCC
AGCACTTACACATCACAGTTTCACATTGCTAGTATCTCCAACTCAAAAGCAGACCGTGAAATTGCCACCTAAttctgaattcacgcc
gccacacgcaaacttataatcaatagtcttcagtatgttttatataaatatccccggaacagattttgctccgagtttaccacttgc
gtaggatgggtgtgaacatgaatcctggatttagttttgggtgtggaatgctgtacatattcgactagaatagatcctagctgtggg
tactctctgtctgtgtgctctgatattgccgtgcgagactttgattcttgnatgtggnattaactgganttgtatattccatagaaa
ttcccaacttagctnagttcaaggt
M22
ATGGACAAAGGAAAAGCGGTGATGGGATCGGCGCGTAGATGGGCCGTCGAATTCACTGATAATTCCAGCGCTCCTTCTTCTCGCGAC
ATATCTGATCCTCCCGGCTTCTCCCGCGCCTCACTAGATCAGGACGATTCGACGCTGACCAAGCAGAAGAAGGATGCGGAGGCTACT
TGGAAAGCTCAGAAAGCTTGGGAAGTTGCACAAGCGCCTTTCAAGAATTTGTTTATGATGGGTTTTATGATGTGGATGGCTGGAAAT
ACAGTGCATCTGTTTAGCATTGGCATTACATTTTCAGCTCTTTGGCAGCCAATAAGTGCTCTTCAAGGGGTTGGAAAGGTTTNCGAG
CCTTATAAAGACAAGAAAGTGGACCTTCTTGGCCCTAAATTACTCTTTATTGCCCTTAACTTCNGAGGATTANCTCTTGGTATCTGG
AAGCNGAACGCGTTGGGTCTTCTNCCNACACACGTATCTGACTGGGTTTCATCCTTACCCCCTGCTCNNGANGTGGAATATTCTGGT
GGAGGCATTCCTCTGCACTAGtggtcctttctgatatatcatgattcttanatgttantttgccctgttnnncnttatctgtgcnt
gtcncanccntgnacnatncnctgncnnggccntttt
M23
GGGGGACAAGAATCAAAGAGACAAACAAACATTACACCTTCAGCCTCAAAGACAAAGCTGCCACAAATCATCACCATGCTATCTTCT
ATCGCTCCCTCACCATGCCAGGATCCTCCAAGCGCCTACCTGCAAACCCTTTTAGCCTCGGCTCGACCCTTCCTTCGAGGAGAATTG
GAATCCATAGACACAAACCTGCCTTCTCTTGTCTCCGTCTTGCGTTCTGTCGGTGCGGGAGAGTGCTGGCACAGGCACGGTAGCTTT
CTCGATCACCTGGTTGACATCTACCGCATCCTCAAGATATGGAGAGCACCGGACTCTGTCTGTCTCTGTGGACTGTTTCATTCTGCT
TATTCCAACTCTTACGTTAATCTTGCCATCTTTGATCCCAACACTGGCCGCGATGTCGTTCGGAATCATGTTGGCGAGGCTGCGGAG
AGGCTGATTCACTTGTTCTGCATTGTTCCTCGCCAACCCCTTATCCATGATGACCTTTTGTTCAAGTATTCGGATGAGGAACTCGTC
GATCACCTCAAGGTTTCTGAGTTGTCGTTGAGGAATGCAAAGGAGAAAGGGTTGTTTAATGGGGAAGAACCCTGGAGAATGAAACTG
TCCTCTCTTTTGCCTGCTGCTGGAATCACTGTGAAGCATATCAAGACTGGAGAGGANGTCTTGCTCACGAGGANGATGGTGGNTACN
TTT
M24
GGGAAAAGAACGAGAAGAGTAGAAAGCGGGAAGATGAAAGTGACATGGAAGAAGACAGACAACAACAAGAACAAGAAGCGGTCTTTG
TCGGCCGTATCAAAGTACCCTAACCTCCCTTTTGATTCGGAACAGCAAGATCAACCTGATACTGCCCCCGACAGTGAGCACCAGCAA
GCTTTTGTACCTACCGACCAACTGGGTTCCGCGTCCGATTCCCAACTTGCGCAGTCCTTTCAAGCCCAAGGCGACNAGCTTGCGGAG
GATGGGAAATATCNTGAAGCGCTTGGTAAATGGGAANCTGCTCTCCNTTTGATGCCCAAAAATGCANTCTTACGTGAACAANAGGCN

CAAGTTTTGCTTGAGATTGGAGATGCATGGAATGCATTGAANGCANCAACTCGCTTGGATCGCTC
M25
ATGGGGGTAAAGGTGTATGGTCCAGCCTACGCTTCTCCGAAGCGGGTGCTCGCATGTCTTTTGGAGAAGGGAGTCGAGTTCAAGGTT
GTGCCCCTTGATATTATCAAGGGTGAGAACAAAAACCCAGACTTCCTCAAGTTGCAGCCTTTTGGATGGGTTCCTGTAATTCAAGAC
GGAGATTACACTCTATATGAGTCGAGAGCTATTATACGGTACTATGCTGAAAAATACAAAAGTCAGGGGACTGATTTGCTGGGGAAG
ACAATAGAAGAAAGGGGCCTTGTCGAGCAATGGCTAGAGGTGGAGGCACATAACTATCATCCAGCAGTTTATGACTTGACCCTTCAC
ATTTTGTTTGCCTCCTCAATGGGGTTTGAGCCAGACCAAAAGGTCATAGAAGAGAGTGAGCAGAAGCTTAGTAAGGTGCTGGACATT
TATGAGGACAGGCTCTCCAAATCCAAATACTTGGCTGGGGATTTCTTCAGTCTTGCTGACCTTAGTCACCTCCCATTTACTCAGTAC
CTGGTCGGTCCGATGAAGAAGGAGCATTTCATAAGGGACCGAAAACACGTCAGCGCGTGGTGGGATGATATCAGCAACAGACCTTCA
TGGAAGAAGGTGCTTCAACTCTAGaatcgtattagcacagagacccatcatggaagaagctccttcactcaataatcatattagtgt
gttttgtgagtcggaattttgtaatggnttccttggccanaagggatgcttgtaaccnttanctgaagttggtttatgttttattgt
M26
ATGGCAAGCGTCTCTGTACTCCCTGAACCTGCAAGAAGGCTTGAAGGTAAAGCGGCATTAATCACTGGTGGGTCCCGAGGTATTGGT
GAGTGCACAGCGAGACTTTTCTTCAAACATGGTGCCAAAGTGGTGATCGCAGACATCCTAGATGATTTGGGCCACGCTGTTTGCAAG
GAAATCAACTCGGAATCTTTTTCCTATGTCCACTGCGATGTGAGCAAAGAAGCAGACGTGGAGAATGCTGTGAACACAGCTGTTTCA
AGGTATGGGAAGCTGGACGTCATGTTCAACAATGCTGGTATTGTAGGCGTGGTCAAGACCAACATACTGGATATCACCCAGGACGAA
TTCGAGCAAGTCATCAAGGTTAATGTAACAGGTACATTCTTGGGCACAAAACACGCAGCCCGAGTGATGATCCCGAATCGCAAGGGC
AGCATAATCTCAGCCTGTAGCATCTGCGCTAGTGTTGGTGGGTTTGCTTCACACGCCTACACCAGCTCAAAGCATGGGGTGTTGGGG
ATAGTGAGAAGCACTGCCGTGGAGCTTGGGAACCACNGCATTCGTGTGGAATGCTGTCTCCCGTACATTGTCGAGAGTGACATGTCA
AAGCATTTCCTGAAACTTGGACNNATGAAAGGGTTTTAAAAACGTGGTACCTCCCAATCCTTNAAAGGGAAGGGTCCNTTCANGGCT
CCAAGGAATGTGGGCCNAATGCTTGTNCCTGGNTTCNTTGGGCCAAAGCGGAATGNATTGCCNAAANGTAATTGTTAAACCNGGGCC
CATTAAACCCTTTTTCCGGTCCAAT
M27
ATGGGGATTCGCGCCTTTATTGTTACGTCTCTAGTATTCTCGCTGTTAGTCGGAGCCTTAGCTGAGCAATGTGGAAGTCAAGCGGGT
GGTGCCGAGTGCCCAGGTGGCCTCTGTTGTAGCCAGTGGGGTTGGTGCGGCACCACAGATGAATATTGTTGTACTAGTAAAGGCTGC
CAAAGCAACTGTCGTGCGTGTGCTGCTATTGGTGGCGGTGCTGATGGTTATCTGGACAGCATCATCTCAAAATCAACATTTAAAAAG
ATGTTTCCGCAAAGGTACCCCTACGAAGCTTTCATCCAAGCTGCCAAGGCATTTCCTGATGGCTTTCCCATCCAAGGGGAGATAGCT
GCTTGACTGTTACAATCTTGGACTTGGTANACACCATCNATCTTATCAATCTAATANGGTGTTGTGAACCTTTATCA
M28
GGGGGAAGTTGAAGAATGTATCAAATGAATGGGCAAAAGGAGTTCTAGGTGATGCTGAAAACCTTACGGAAGAGAGAGTGGATGAGG
TTTTGAGGGAGTATCTACGAGATCTCAAAGAGGGTTCTCTGAAGGCCAAAGGCTGGCCTCCTGTTGCATACATACTCTCAAAAGCTG
CTATGAATGCCTACACCAGGATTCTTGCGAAGAAGTACCCGAATTTCTGTATCAATTCTGTTTGCCCTGGCTATGTCAGTACAGACA
TCAACTTTCATACTGGAAAATTATCCGTTGAAGAAGGTGCTGAAGGTCCAGTGAAGCTAGCATTGTCTCCTGATGACAGTCTTTCAG
GCCGGTTCTTTGTTGAGAAGGAAGAATCACCTTTTTGAttggattgtatctctttcttcatgtgaagcctataatatatttggaaaa
tcggtaccaaacaagtgtcttcaggaataaaagtcatcaagcagtccaaatacatgtcagagtccataatttttcttagttttgatgt
ataagaattgcgctttacatgttgtatttatattancaaacacgtggtttgatcnaacatnnccantnaaangnatataannannct
tgtcggccgcctcggcccactcgactctagactcgaccaagcttttgcntgcgggccgcanttcnnagctcncttggccatttcgcc
ctatagtgngtcggnatnacaattcnctgnccgccgntttccancggncgtgaccgnntanacccngnncgttnacccnacttantc
gccttnnnngnnnctcccccttttccccacctgggntantnanccgaatanggncctncnccttnntnccccnttncnccncngntng
gccccncctngncgggcnantcgnntcntntancncnntacnnttttnngacanttccccncnnnatttntngttcaaaacangccn
ntntnnancccgcntngtctaaacca

G1
GGTACCAGAGGAGATTGTTATGATCGTTACTGTATCCGTATTGAAGAGATGCGACAAAGTGTTCGGATCATTGTGCAATGTCTTAAT
AAAATGCCTAGTGGCATGATCAAAGCCGATGATCGTAAGCTATGTCCTCCATCACGATGTCGAATGAAACTATCCATGGAATCCTTA
ATTCACCATTTCGAACTTTATACAGAAGGTTTTTCCGTACCAGCTTCTTCTACCTATACCGCAGTTGAAGCACCTAAAGGAGAATTT
GGTGTCTTTCTGGTCAGTAATGGAAGCAATCGTCCTTACCGTTGTAAAATAAGAGCACCTGGCTTTGCCCATTTACAAGGACTCGAT
TTTATGTCCAAACATCACATGCTAGCAGATGTGGTTACCATCATAGGTACTCAAGATATTGTGTTTGGAGAGGTGGATTAGcatang
actattcanttgttcgnatca
G2
GNNGACGACGACGCCCGACGACAACGACGACAACGACGTACGACGACGCATAACGTATGACAACAACGAAAACATCCAATTTTTGAG
GGAAATGGACGAACTAATGATCATTCAGGAAGACATCGCCAAAATTTATGCAGAAAGAACGAATCAACCTTTAGGGGTCATACTCAA
AGACCTGCAAAGCGATTCTTTTATGTCAGCAACAGAAGCCCAAGCTTATGGAATTGTTGATTTTGTAGCACACTGGGAGCGACGAGT
TAGAAAATAActggaatttcatacaaatcgtggttggggttgtttcatattttatattaacattctattaatttgaatagaatgtta
atatagaaataatccctaatcctccgcataatcatccggttaggagcgacctaaaccaacccattatgcntatgattcatcatgcca
actgttaaacaacttattaggaaggcaagacagccaatcagaaatgtcaacnaaaacccccgcgcttagagggtgtcctcagcgccga
ggaatatgtactcgggtgtatgtgcgactc
G3
GGANGACGAACGACGCCGACGACAACGACGACAACGACGACGACGACGCATTAACGATGACAACAACGAAAACATCCAATTTTTGAG
GGAAATGGACGAACTAATGATCATTCAGGAAGACATCGCCAAAATTTATGCAGAAAGAACGAATCAACCTTTAGGGGTCATACTCAA
AGACCTGCAAAGCGATTCTTTTATGTCAGCAACAGAAGCCCAAGCTTATGGAATTGTTGATTTTGTAGCACACTGGGAGCGACGAGT
TAGAAAATAActggaatttcatacaaatcgtggttggggttgtttcatattttatattaacattctattaatttgaatagaatgtta
atatagaaataatccctaatcctccgcataatcatccggttaggagcgacctaaaccaacccattatgcatatgattcatcatgcca
actgttaaacaacttattaggaaggcaagacagccaatcagaaatgtcaacaaaaccccgcgcttagagggtgtcctcagcgccga
ggaatatgtactcgggtgtatgtgcgactcgttcggattgcggattggaacaaaagaaaggaaacgaattttccactattcgcgatc
aataccgatcaataggatagaatggtgaaaacccttcactatctaatatctaaaacactatctaaaaaaaaaagatctaccatatc
cttttattgtgtatcaaatttatggtttctattantgaaaattcnatcntctcantttcnnatttanttgtgaaagattccccatt
ggtagcnaatgattanccgttagcaggggaaatcttanctaggaaaaggccnaanatttntgcctctctcttgcanacgggct
G4
ATGGGGATTATTTTTACTCGAATGTTCTCTTCTCTTTTCGGCAACAAAGAAGCTCGAATCCTGGTCCTCGGCCTCGATAACGCCGGC
AAAACTACCATTCTCTATCGCCTTCAAATGGGTGAGGTGGTGTCCACGATCCCAACAATTGGATTTAATGTGGAAACTGTAGAGTAC
AACAACATAAAGTTCCAAGTCTGGGATTTAGGTGGGCAGACCACAATCAGGCCATATTGGAGATGCTATTTTCCAAATACTCAGGCT
GTAATTTATGTGGTTGATTCAAGTGACACTGAGAGAATTAGTGTAGCTAAAGAAGAATTTCACACAATTTTGGAGGAGGATGAACTG
AAAGGGGCGGTGGTCCTTATCTTTGCAAACAAGCAGGATCTCCCTGGTGCACTAGATGATGCTGCCGTGACCGAGGCCTTGGAGTTG
CACAAGATCAAAAATCGTCAGTGGGCAATCTTCAAAACCTGTGCAGTTAAGGGTGAAGGCCTCTTTGAGGGTTTGGACTGGCTGAGT
AATACTCTAAAGTCCGTGGGTGCCTAAatttgcttagctttatatgcacgtcgctgtggtgataagttatattacattgattgcact
gaacacttttttctttcttcgtaatatttttcatttttcttgggtgggttatagccttctnaatttattggaattttagcctttaagt

ntatnatantttgatgaatgtgaatttccaaaaatgaatttttgggggggaatnaaaaaccaagatgtatgttcntgg
G5
GCAGGATCCATTGGCGGATGTATATTTGAAAAAGGCTGAGTTCAACAGGGAGCCTTTATATGCCGAGTTTCGGCTTAGGGAACCAAT
GAGAAAGCCTATATTAAGCTCGGAACCTTCGATTTCGGTGCATCAGCTACAGCCACATGATCTATTTGTCATATTTGCATCCGATGG
CCTGTGGGAGCATGTAAGCAATCAAGAGGCTGTTGACATAGTTCAAAGCCATCCACGCTCGGGAAGTGCGAGGAGGCTGGTGAAAAC
TGCATTGCAAGAAGCAGCAAAGAAGAGAGAAATGAGGTATTCAGACTTGAAGAAGATCGATCGTGGCGTCAGACGGCATTTCCACGA
TGATATCACTGTGATTGTTGTGTTTCTCGACCCGTATCTTGTTTCCGGCTCTAGCTCAGTTAAATGCTCTAATATATCTGTAAGAGG
GGGTGGAATCAGCTTGCGCCCCAGCACTCTGGCTCCTTGTGCCACACCAACCGAAGCCGGAAGTACCTGAtgaagctgtttctcatg
ttgttaatctctcttgtattgttctcttgtgaataccaggtagtttctgaaggggacaaaacagaggagggagtattctattctttt
aatatacaatgctacttttttaacangggactggttcacggttttaaggcgacaaaagggctattgtaattatttcttgagaaaagtc
tctactctgtacatttttccatcttctccgcgcactcagttttgtaaaancgttnccttgccaaatgaagtccctgggttcctgcct
gctc
G6
AGAAAGAGGTAGTAGTGACTATATTTGCAAAGGGTGTACCAGCGAGCAGTGTTTCAATTGACTTTGGCGAACAAATTCTAAGTGTTA
GCATTGACATCCCTGGTGAAGTGCCTACACACTTTCAACCTCGTTTATTTGCAAAGATTGTCCCTGACAAGTGCAAGTTCAGCATCT
TGTCAACTAAAGTTGAAATCCGCCTTGCAAAAGCTGAACCCATACACTGGACATCTCTTGAATACATGGAGGACACTGCGGTTGTAC
AGAGGGCTGTTGTGTCATCTGATGCTGGGGCTCAAAAGCCTGTTTATCCCTCATCAAAGCCAAAAAGGGTAGATTGGGATAAGATTG
AAGCCCAAGTGAAGAAGGAGGAAAAGGATGAAAAGCTAGATGGTGATGCAGCTTTGAACAAATTTTTCCGTGAGATATACCATGATG
CTGATGAAGACACTAGAAGAGCCATGCAAAAATCTTTTGTGGAGTCAAATGGGACGGTGCTCTCAACGAACTGGAAAGAAGTGGGAA
ATAAGAAGGTGCAAGGAAGTCCTCCCGATGGTATGGAAATGAAGAAATGGGAGATTTAGaagtcttgacatatcataactcgctaca
cggttttgttctatggatgtcgtggagtttccctgctgttatcaatgttgtgtttgattgtctggtggaattgtgttcttgcacagc
tatgctacttttgtaaagttgcatttcaaacancatagctcctttgtttgaatatatgtactggatgtcaacatggtcgaatctttg
acgtgtgaactgggtctatgtctgcantttgtnccccctccatttactttgcatggctggaagggcttttgaatttttaattttatg
ggatcantttntg
G7
GGCAAAGGATATGCTTAAATCATTGGCATCAAAGTGGGATGATGTTGTTGATGCAAACGCACTCCAGGTGATCCCTCTTAAGGGTGC
TATGACCAATGAGGTATTCCAAATAAAATGGCCTGGAAAGTCCAAGGACATGCCACGGAAGGTTTTAGTTAGGATTTATGGTGAGGG
TGTGGAGGTGTTCTTCGACCGATGCAATGAGATACAGACGTTTGAGTTTATGTCAAAACAAGGGCAAGGTCCTCGCCTTCTAGGCCG
TTTCTCCAATGGACGAATTGAAGAGTTCATCCATGCACGGACTTTATCAGCATCTGATTTACATGATCCAGAGATATCTGCTCTTAT
AGCAAC
G8
CGGGAGGAGATGCAAAGATCCTCGGGTACTCTAGAAGCATTTTCCACGAAAAAGATCTCAAGTCCACACGATTCTTCCAGAAATAGG
ACAGATAACCAACCTGAGCCAGAAAAGGGCCACAGTTCTTCTCATTATGGCAGCAGTGCAAGAAAAGCTGTCATCTCAAAGAGTGGA
CCGAACTCCTCTGGCGAGCCTAGTGACAATCGTAACAGCAGCCAAATATTCTCAAGTGGTGGCCGTCTGATTTCTTCTGCACAAAGA
GTCCAACCTGTGTATGAGCCTAAATCATCTCATAGTCGTGGCGCAGCTGCTAGAGGCACTCGCGAGGAGGCCTTCCGGAGCTTCGAG
TTCCTCTCCATCAGGAAATGAacgaggaattggagcttgtgagaagctatcgacttacataaccattaggtaaaattcttttcaacg
tgtagaaatggaacgcgaccgggtacttgggatactggtggttattcaacgctcgtgtcaaatcacttcaattccatgaagaacctc
tggcaccttcttggttccctttatatttcctacaagagtttcggaccagaatgctgtgtagtaactcagagagagacgcgtagagaa
attatatttgcatgcaggcagatatgaccacgttttcaattcatcctggtgggcgtctcttcttgtttg
G9
GGGACTCAAGGNACGGGAGNTCAAGAAGAGCGTAGCTAGTAAGGGAGAAGCGACANCCCGTGTCATGGTACCCTCCGCCAGCAAGAA
CCAGNTCAGTTACATCTGCTGGAGATTGTGGCATGGAGAACACAATGAAACCGGGGGATGTGATCCAGTGCAGAGAGTGCGGTTACC
GCATCCTCTACAAGAAGCGTACTCGCCGAATTGTCCAGTACGAGGCACGCTAGaggaataaacgatgatatgtgctttctgcatatg
gtgtccgatgcctaaaagcattacaaatagttgatgtgtttggttttctgtt
G10
CAGATTCCACCCGGAAGAAGCGAGAAGACAAACTGNAAGAAACGGCAGGAAGAACTGGCGCAGGAGAAGGCTGCAAATGCTTTTGTA
CTTCCATCTGACCATGTTCGATGGGTAATAGGGCCCTCGGGAACGGTCGTAACGTTCCCCAACGAGATGGGCTTCCCTAGCATATTT
GACTCCGGGACTTGCAGCTACCCTCCTCCGCGCGAGAAATGCGCTGCTCCATCCTGTAATAATCCATACAAGTACCGGGATTCAAAA
TCAAAGCTGCCTTTGTGCAGTCTTCAGTGCTACAAGGTAATAAATGAAAAGGTGGAGCCCCTGACTGCTTGCTAAgttcaagggcaa
ccagtgcccaaggcctgataagagatgggaataagggtgtgacaaggcggtgtgtatgtaatgcaataatgattggtgttgtttcca
gctactctgggcatggagaaattgttggcttagcagcataatagctatacgtcgagtgtatgcaaatactgaattatgtatttgtat
tagttttttttttatatatgtaataaattttgatgtaagctcgttttattaaccagatgaaagtgcagtctctctaat
gttgaactaaggcttcctacatccgaataatacttattagttcggcgtgtaanc
G11
ATGGGTGCTCTTGATTACCTCTCAAACTTTTGCACTGTCACCAGCACAAGAAGCAAACGCAAAGCAATGCAGACAGTTGAAATCAAA
GTGAAAATGGATTGTGATGGCTGTGAAAGAAGAGTGAAACATGCAGTTACCACCATGAAAGGTGTAAAGACTGTGGAAGTGAATCGT
AAACAAAGCCGGGTAGTAGTGAGTGGATATGTTGAACCAAACAAGGTGTTGAAGAGAGTGAGGAGAACAGGCNAGAGAGCTGAGTTT
TGGCCATATATACCGCAACATCTAGTTTACTATCCTTATGCTGCCGGTGCTTATGACAAAAGGGCACCTGCTGGTTATGTCCGCAAT
GTCGTTCAAGCTTATCCAGCTTCGACTGCACCTGAGGACAACATGATCTCTCTATTCAGTGATGATAATGTGAATGCATGCTCCGTC
ATGTAAttgaccaggttaattaatgtaatccatgcaggcatgttcT
G12
GGCGAGAGGAGCTAAAGAGGGTGAAGAAGAGCCCAAGAAGGAAGANGTAGTAGTGGAGGAAAATATCATCAATGTTGAAGAAGAAAA
TATGAAGAGGACGATGTATTATTACCAGCCGCTATATGTGATTGAGAGAATCCCACCGCCACAAATCTTCAGTGATGAGAATCCCAA
TGCATGTTGTATTACATAActacgtaatcatcatgatatgaaaatatgctattatgttgtgtatcatttatgtaatcatattctgta
tgggaataaaaagtcggtgcatcgtatatatatatatagtgttacctacttataagtactgtagtctacttgtatataataagattc
cttttcttaaataac
G13
TATGGAGGATGCTAATATCCTTGAAGAAGAAAATGCACCTCTTGAGCATCCAACAAAGCCTGGTATTTCACGTTCAGAGCATTTTGA
CATGGATAATACATGGTACGATGCCGCACCTGATGGATTCAGCTTGACGTTATCTCCTTTTGCAACAATGTGGATGGCGCTCTTTGC
ATGGGTAACCTCATCTTCTTTGGCTTATATATATGGGAGGGATGAAAGTTCTCACGAAGATTACTCAGAGGTTAATGGGAGAGAATA
TCCTCACAATATTGTATTGAGTGACGGTCGTTCTAATGAAATTAAGCAGTCGGTTGCTGAATGTCTTAGTCGAGCTTTTCCTGGCCT
TGTCTCTGAACTCAGGATGGGAACACCATTATCAACTTTGGAGCAAAGAACGGCCTGCTTGTTGGATACAATGTCATTTGTTGATGC
AATCCCAGCATTCAAACTGAAACAATGGCAAATCATAGCACTTCTCTTCATTGAAGCTCTATCTGTTTCAAGGTTACCTGGGCTTGT
TTCTTACATGACCAATAGGAGAATGGTCCTTCACAAGTTGTTGGATGGTGCTCACATAACTGCAGAAGAATACGAGACCATGAAGGA
TCTATTGATACCCCTTGGCCGGGCACCTGATTTTTCGCCGCAGAGTGGAGCCTAAtttggtccaaacgtttctccnagtaagcatgc
aggaaatgtttgcgaccacatattaccctgatccttgaaacggtagnttgattcgagtcagaattttgggaacctgtgntttgttgg
tntatcgacataactgaaaacagtggatgtgtatgtaacttataaactctcggnttgcaataccaatttggcatggatccctnctgt

ttg
G14
GATTGGCGACTATGTTGACATCAAGGTTAACGGCGCCGTCCACAAGGGTATGCCTCACAAGTTCTACCATGGACGTACCGGTCGTGT
TTGGAACGTCACCAAGCGCGCCGTCGGCGTCGAGGTCGACAAGCAGGTGGGTAACAGGATCATCAGGAAGAGGATCCATGTACGCAT
AGAGCATCTACAGCCTTCGAGGTGCACTGAGGAGTTCCGTCTGAGGAAGAAGAAGAATGATGAGCTAAAGGCAGAAGCAAAGACAAG
AGGTGAAGTTATCAGCACCAAGAGACAGCCAGAAGGACCAAAACCTGGATTTATGGTGGAAGGTGCAACATTAGAAACTGTGACTCC
CATTCCCTACGATGTCGTGAACGATCTCAAGGGTGGTTACTAGgagtctcatgtttttgttctttttcaatgctactttgtgcgtgt
gtttaaacttattggtatggattagtacttttggattctgaagatatataatcggagaattattcccc
G15
ATGGCAAGATCGAGGCATCTCGCTATTACTCTGCTGCTATGGGCTTTGGCTCTTCTTGGAACCGTAACTCTCAGCCAGGCAAAGAAG
TCGAAGGAAGATTTGAAAGAGATCACGCATAAGGTCTATTTTGACGTTGAGATTGATGGAAAACCAGCTGGGCGGGTTACAATGGGT
CTCTTTGGAAAGGCGGTTCCTAAAACTGCGGAAAATTTTCGTGCATTGTGCACAGGGGAAAAGGGCATTGGAAAGAGTGGAAAACCC
CTTCATTACAAGGGGAGCAAATTCCACCGGATTATTCCCAGCTTTATGATACAAGGAGGTGATTTTACACTAGGCGATGGAAGAGGA
GGAGAATCAATTTATGGAGAAAAATTTGCTGACGAGAATTTTGAGTTGAAGCATACTGGACCTGGGCTTTTATCCATGGCAAATGCT
GGTAGAGATACTAATGGTTCACAATTCTTTATCACCACCGTGACAACTTCCTGGCTTGACGGAAAACACGTTGTGTTTGGGAAGGTG
ATATCCGGTATGGACGTGATTTACAAGATTGAAGCCGAAGGCAAGCAGAGTGGCACGCCTAAGAGCAAGGTCATGATTGCAGACAGT
GGGGAACTGCCCCTGTAGtcttttccttcttttttncccaaactaaaaatcttcagcnaacaagttcagatcttttgctggagacca
aattattcccccatatttagcaatgccaaccnttttttctttnnggt
G16
AAGGAGCTTGGTGTCACTGCTCTTCATATTAAGCTCCGAGCTACAGGTGGCAATAAGACTAAAACCCCTGGTCCTGGTGCACAGTCT
GCTCTTCGGGCTCTTGCTCGTGCTGGGATGAAAATTGGCCGCATAGAGGATGTGACGCCCATTCCAACTGATAGCACTCGTAGAAAG
GGTGGTAGAAGAGGAAGAAGGCTGTAAGCTTTTGTGTCTTCCCTGTTGGTACGGGTCTGCTCAGGTTTTATTCTCCATTGTGGGATA
AGGCAAGGAGTACCTGCTGTCATTGGGTTTTACTTGAAAGCTATTTACCTCTTTCGGAACTGTTGTTTTAGCTTTTTTATTAAAATG
ACTTGGCCGTGGTTTGTCTTAAAACTGTTGTAGAAGAGATTTTGGATACTTTAATTTTGTGAcattcttcc
G17
ATGGCGAAGAGACAGTTGCTTGACCCGCGNTAGCAACCGGAGCCGCCGNTGGGGGGTGCTGTCGGCGGCGCTGNTGCTGTTTGGGGA
ACTTATGAGGCTATTAGGTATAAGGTGCCAGGTCTTATGAAAGTGAGGTATATCGGACAAACTACGCTTGGCACCGCCGCGGTTTTT
GGGCTCTTTTTGGGAGCCGGAAGCCTAATACATTGTGGAAAGTCTTACTGActgtttatccttttgttgtctttaaaaattcgtact
gtttgtgacatttcaagagtattcaatccaaatccagcatctgtttcc
G18
ATGTCGTCGCACAGCAGCAACGCGTCGTCTGCTAACGATCCGCGGCTGCCTTCGGCGGCGAAACGTTATGTAGCGCCGATGGTGTCG
CCGCAGAATCTTCCGGTCGACTACTCTGGATTCATCGCCGTCGTCTTCGGCGTCGCTGGGATGATGTTTAGGTACAAACTGTGTTCT
TGGCTTGCAATCATATTCTGCGCACAATCGCTTGCGAACATGAGAAACATAGAGAATGATCTTAAGCAGATCTCCATGGCTTCTATG
TTTGCAATCATGGGCTTGGTGACAAACTACTTGGGTGTCGGCCGAGCTAATCCAAAAACTTGAgattccttcaatcagacattagaa
ggaaacactggccgcacgatgttaatctaacgatggcattgggttcggcaaagacgggtgataacattttaaaaacttttcatgatg
agatttgatgcaacgatttagaacatgatgtgaaaacttcttttggaatataaaattatgtagcttagattgctgaactttcaactt
gatttgtactgcaatttgtagaattatttcatatggccatctcgtttgacctttttggcttcttttatccgtatgtttgtct
G19
GGGTTCTGCACGAGCTTTATCCTGATTTGAAGCTTCCTGAAAAGTGTGCGGACGACAAGCCTTATGTGCCAACATATCAGGTATCTA
AAGAAAAAGCCAAGAGCTTGGGCATTGACTTCATTCCCCTGGAGGCGAGCATCAAGGAAACCGTTGAAAGCTTGAAGGAGAAAGGAT
TTTTCAGTTTCTGAttcctattctgttctcctgattaccaaaaataaaaatgatgttggatgcgctttgtccttcaacctggttctt
gtgttttcttagaatcaccgtgttaacgatgatgtaacgtttgttttgcatgcatgtgatggcattgctgaagttatcagaaatttg
aaactaataatcattactggataattctccc
G20
GTACCAGGGAGGGTGCCAAAACCCATGCACTGTATTCAAGACTAATGAGTACTGCTGCACTCAAGGGCAAGGAAGCTGTAGACCGAC
CAATTACTCGAGGTTTTTCAAGGATAGATGCCCTACTTCTTACAGTTACCCTCAAGATGACCCTTCAAGCACATTCACCTGCCCCGG
TGGGACTAACTACAGAGTTGTGTTCTGTCCTCGGGGGTCTCCTCATTTTCCTTTGGAGATGGTTGGAGAGAAGAATGTTGAGTAAtt
aatcaaggacacggtggatggatgcaagcatgtgtgttttgcgaggccttgaaaacaagctttgcaacaataaaggcctagttatgg
tttggagttgtagatccttttttgttcccttttgcgaacaaaacatcagttgtgttttgttcatgcctgttactgtactcagttgttg
tgatggaaacttcagtgaaataaagacaacttaaatcgtattacaattcactggccgtcgttttacaacgtcgtgactgggaaaac
cctggcgttacccaacttaatcgccttgcagcacatcccccttcgccagctggcgtaatagcgaanagccccccgat
G21
GGCCAACAGACCATTGCCAACAACCAGGNTAACAAAAAGAAAATGAAGACACAACTTTGCTCAAGTGCTGCCTTCCTTACATTTCTT
TTCCTCGTCTCAATGTGTTCGAGTATAGAAGCTCAGACGTGCAGGCCCAGTGGCAATATCAGAGGGAAAACGCCTCCAGCAAACCAA
TGTAACCGAGAGAACGATTCTGATTGCTGTGTGGATGGCAAGTTGTATCCCATTTACAAGTGCTCACCCTCAGTTTCCAGCCGTACA
AAGGCAACACTGACGCTGAACAGCTTCCAGAAAGGTGGGGATGGTGGAGGACCATCAGAATGCGACAACCAATACCACTCGGATGAC
ACCCCTGTTGTGGCACTCTCAACCGGATGGTTCAACCACAGGAGCAGATGCTTGAACTATATTAACATCTATGGTAATGGAAAGAGT
GTCAAGGCCATGGTAGTCNATGAGTGTGATTCTNCAATGGGTTGTGATTCTGATCATGATTACCNGCCTCC
G22
GGGAGGGATGCCGTGATCTACCCAGCACCAGCCTCTGGTGGAGAAAGAAGCAGAGGAGAAAAAGCAACTGTATCCGAAATTTGTGTT
CGAGGACTATATGAAGATCTATGCTGGACTAAAATTTCATGCCAAGGAACCAAGATTTGAAGCGTTGAAAGCCGCTGAAAATAATGC
CAATTCAGGTCCAATAGCCACCGCTTAAgctcggctttagtacaagaggagcatgtgtttggtgtgTtattgtctctatagtgactt
aaaagagaatcatcacgaaggacacatacatatagagctgccttgaaaaacattggttcgctcttgtgagtgtcgatatagttattg
tgtaagaaataaatgcgattgttatatgataaataagtttgtttcatagc
G23
TGAAAGGAGGAGGATGGGCATCTGCCCTTGTTGTTGAGCTTTTGTAAGCCCTTTTTCAGAGCCAGCGTTTCGTGCAAAGCTCAAAGA
TCCTGCTTCCTTCATCACCGTACATGCAAATTCCAAATGTATTACACACCAGACGGGAAGGGTGGACCGAAATTTAGCTACATAATT
GTTAGGTTTTAGTCTTAGATACTGTGAAAGTGAAGTTTTCTTAGATCATTCAGCAAT

References

1. Silverstein, R.M., Rodin, J.O., Wood, D.L. 1966. Sex attractants in frass produced by male *Ips confusus* in *Ponderosa* pine. Science 154: 509-510.
2. Hendry, L.B., Piatek, B., Browne, L.E., Wood., D.L., Byers, J.A., Fish, R.H., Hicks, R.A. 1980. *In vivo* conversion of a labelled host plant chemical to pheromones of the bark beetle *Ips paraconfusus*. Nature 284 (3): 485.
3. Dudareva, N., Murfitt, L.M., Mann, C.J., Gorenstein, N., Kolosova, N., Kish, C.M., Bonham, C., Wood, K. 2000. Developmental regulation of methyl benzoate biosynthesis and emission in snapdragon flowers. The Plant Cell 12: 949-961.
4. Mottram, D.S., Mottram, H.R. 2002. An overview of the contribution of sulfur-containing compounds to the aroma in heated foods. In "Heteroatomic Aroma Compounds" Reineccius, G.A., Reineccius, T.A. Eds., American Chemical Society. p. 73-92.
5. Knudsen, J.T., Tollsten, L., Bergstrom, G. 1993. Floral scents - a checklist of volatile compounds isolated by head-space techniques. Phytochemistry 33: 253-280.
6. Dudareva, N., Pichersky, E., Gershenzon, J. 2004. Biochemistry of plant volatiles. Plant Physiology 135: 1893-1902.
7. Dewis, M.L. 2005. Molecules of taste and sensation. In "Chemistry and Technology of Flavors and Fragrances" Rowe. D.J. Ed., Blackwell Publishing Limited. p. 199-243.
8. Rowe, D.J. 2005. Introduction. In "Chemistry and Technology of Flavors and Fragrances" Rowe. D.J. Ed., Blackwell Publishing Limited. p. 1-11.
9. Turin, L. 2005. Rational Odorant Design. In "Chemistry and Technology of Flavors and Fragrances" Rowe. D.J. Ed., Blackwell Publishing Limited. p. 261-273.
10. Zaninotto, P., Wardle, H., Stamataakis, E., Mindell, J., Head, J. 2006. Forecasting obesity to 2010. Department of Health Report.
11. Tackling Obesity in England, 2001. National Audit Office Report.
12. Margetts, J. 2005. Aroma chemicals V: Natural aroma chemicals. In "Chemistry and Technology of Flavors and Fragrances" Rowe. D.J. Ed., Blackwell Publishing Limited. p. 169-198.
13. Johansson, D.J.A., Azar, C. 2007. A scenario based analysis of land competition between food and bioenergy production in the US. Climatic Change 82(3): 267-291.
14. Saitta, M., Di Bella, G., Salvo, F., Lo Curto, S., Dugo, G. 2000 Organochloride pesticide residues in Italian citrus essential oils 1991-1996. Journal of Agricultural and Food Chemistry 48(3): 797-801.
15. Krings, U., Berger, R.G., 1998. Biotechnological production of flavours and fragrances. Applied Microbiology and Biotechnology 49: 1-8.
16. Agricultural Commodities: Section 3. Tropical fruit, 2002. Food and Agriculture Organization of the United Nations Report.
17. Maróstica, M.R., Pastore, G.M. 2007. Tropical fruit flavour. In "Flavours and Fragrances" Berger, R.G. Ed., Springer Berlin Heidelberg. p. 189-202.
18. Goff, S.A., Klee, H.J. 2006. Plant volatile compounds: Sensory cues for health and nutritional value? Science 311: 815-819.

19. Ridley, M. 1994. The Red Queen: Sex and Evolution of Human Nature. Penguin Books. p. 63
20. Gatehouse, J.A. 2002. Plant resistance towards insect herbivores: a dynamic interaction. New Phytologist 156: 145-169.
21. Shibamoto, T., Tang, S. 1990. The flavour of fruit. In "Food Flavours" Morton, I.D., MacLeod, A.J. Eds., Elsevier: Amsterdam, The Netherlands. p. 221-234.
22. Jordan, M.J., Goodner, K.L., Shaw, P.E. 2002. Characterization of the aromatic profile in aqueous essence and fruit juice of yellow passion fruit (*Passiflora edulis* Sims F. *Flavicarpa degner*) by GC-MS and GC/O. Journal of Agricultural and Food Chemistry 50: 1523-1528.
23. Parliment, T.H. 1972. Some volatile constituents of passion fruit. Journal of Agricultural and Food Chemistry 20 (5): 1043-1045.
24. Engel, K.H., Tressel, R. 1991. Identification of new sulfur-containing volatiles in passion fruits (*Passiflora edulis* f. *flavicarpa*). Journal of Agricultural and Food Chemistry 39: 2249-2252.
25. Werkhoff, P., Guntert, M., Krammer, G., Sommer, H., Kaulen, J. 1998. Vacuum headspace method in aroma research: flavor chemistry of yellow passion fruits. Journal of Agricultural and Food Chemistry 46:1076-1093.
26. Singer, G., Heusinger, G., Frohlich, O., Schreier, P., Mosandl, A. 1986. Chirality evaluation of 2-methyl-4-propyl-1,3-oxathiane from yellow passion fruit. Journal of Agricultural and Food Chemistry 34: 1029-1033.
27. Weber, B., Maas, B., Mosandl, A. 1995. Stereoisomeric flavour compounds. 72. Stereoisomeric distribution of some chiral sulfur-containing trace components of passion fruits. Journal of Agricultural and Food Chemistry 43: 2438-2441.
28. Brat, P., Brillouet, J.M. Reynes, M., Cogat, P.O., Olle, D. 2000. Free volatile components of passion fruit puree obtained by flash vacuum-expansion. Journal of Agricultural and Food Chemistry 48: 6210-6214.
29. Chen, C.C., Kuo, M.C., Hwang, L.S., Wu, J.S.B., Wu, C.M. 1982. Headspace components of passion fruit juice. Journal of Agricultural and Food Chemistry 30: 1211-1215.
30. Baldry, J., Dougan, J., Howrad, G.E. 1972. Volatile flavouring constituents of durian. Phytochemistry 11: 2081-2084.
31. Moser, R., Duvel, D., Greve, R. 1980. Volatile constituents and fatty acid composition of lipids in *Durio zibethinus*. Phytochemistry 19: 79-81.
32. Weenen, H., Koolhaas, W.E., Apriyantono, A. 1996. Sulfur-containing volatiles of durian fruits (*Durio zibethinus* Murr.). Journal of Agricultural and Food Chemistry 44: 3291-3293.
33. Pino, J.A., Mesa, J., Munoz, J., Pillar Marti, M., Marbot, R. 2005. Volatile components from mango (*Magnifera indica* L.) cultivars. Journal of Agricultural and Food Chemistry 53: 2213-2223.
34. Dewis, M.L., Kendrick, L. 2002. Creation of flavours and the synthesis of raw materials inspired by nature. In "Advances in Flavours and Fragrances-From Sensation to the Synthesis". Swift, K.A.D. Ed., The Royal Society of Chemistry, Cambridge, UK. P.147-160.
35. Umano, K., Hagi, Y., Nakahara, K., Shoji, A., Shibamoto, T. 1992. Volatile constituents of green and ripened pineapple (*Ananas comosus* L. Merr.). Journal of Agricultural and Food Chemistry 40: 599-603.

36. Rodin, J.O., Himel, C.M., Silverstein, R.M., Leeper, R.W., Gortner, W.A. 1965. Volatile flavour and aroma components of pineapple. 1. Isolation and tentative identification of 2,5-dimethyl-3[2H]-furanone. Journal of Food Science 30: 280-284.
37. Wu, P., Kuo, M.C., Hartman, T.G., Rosen, T.T., Ho, C.T. 1991. Free and glycosidically bound aroma compounds in pineapple (*Ananas comosus* L. Merr.). Journal of Agricultural and Food Chemistry 39: 170-172.
38. Takeoka, G.R., Buttery, R.G., Teranishi, R., Flath, R.A., Guntert, M. 1991. Identification of additional pineapple volatiles. Journal of Agricultural and Food Chemistry 39: 1848-1851.
39. Berger, R.G., Drawert, F., Kollmannsberger, H., Nitz, S., Schraufstetter, B. 1985. Novel volatiles in pineapple fruit and their sensory properties. Journal of Agricultural and Food Chemistry 33: 232-235.
40. Rowe, D., 1999. Aroma chemicals for the sweet field. Perfumer and Flavourist 24 (6): 36-41
41. Morton, I.D., MacLeod, A.J. 1990. The flavour of fruits (Part C). In "Food Flavours". Morton, I.D., MacLeod, A.J. Eds., Elsevier: Amsterdam, The Netherlands. 360 pages.
42. Parliment, T., Croteau, R. Eds. 1986. Biogeneration of Flavours. American Chemical Society, Washington, DC. 317 pages.
43. Lalel, H.J.D., Singh, Z., Tan, S.C. 2003. Aroma volatile production during ripening of "kensington pride" mango. Postharvest Biology and technology 27: 323-336.
44. Song, J., Bangerth, F. 2003. Fatty acids as precursors for aroma volatile biosynthesis in pre-climacteric and climacteric apple fruit. Postharvest Biology
45. Rowan, D.D., Lane, H.P., Allen, J.M., Fielder, S., Hunt, M.B. 1996. Biosynthesis of 2-methyl-2-butenyl, and 2-methylbutanoate esters in red delicious and granny smith apples using deuterium-labelled substrates. Journal of Agricultural and Food Chemistry 44: 3276-3285.
46. Rowan, D.D., Allen, J.M., Fielder, S., Hunt, M.B. 1999. Biosynthesis of straight-chain ester volatiles in red delicious and granny smith apples using deuterium-labelled precursors. Journal of Agricultural and Food Chemistry 47: 2553-2562.
47. Pechous, S.W., Whitaker, B.D. 2004. Cloning and functional expression of an (E,E)-α-farnesene synthase cDNA from peel tissue of apple fruit. Planta 219: 84-89.
48. Manriquez, D., El-Sharkawy, I., Flores, F.B., El-Yahyaoui, F., Regad, F., Bouzayen, M., Latche, A., Pech, J. 2006. Two highly divergent alcohol dehydrogenases of melon exhibit fruit ripening-specific expression and distinct biochemical characteristics. Plant Molecular Biology 61: 675-685.
49. Ringer, K.L., Davis, E.M., Croteau, R. 2005. Monoterpene metabolism. Cloning, expression, and characterization of (-)-isopiperitenol/ (-)-carveol dehydrogenase of peppermint and spearmint. Plant Physiology 137: 863-872.
50. Devitt, L.C., Sawbridge, T., Holton, T.A., Mitchelson, K., Dietzgen, R.G. 2006. Discovery of genes associated with fruit ripening in *Carica papaya* using expressed sequence tags. Plant Science 170: 356-363.
51. Schaffer, R.J., Friel, E.N., Souleyre, E.J.F., Bolitho, K., Thodey, K., Ledger, S., Bowen, J.H., Ma, J., Nain, B., Cohen, D., Gleave, A.P., Crowhurst, R.N., Janssen, B.J., Yao, J., Newcomb, R.D. 2007. A genomics approach reveals

that aroma production in apple is controlled by ethylene predominantly at the final step in each biosynthetic pathway. Plant Physiology 144: 1899-1912.
52. Aharoni, A., Keizer, L.C.P., Bouwmeester, H.J., Sun, Z., Alvarez-Huerta, M., Verhoeven, H.A., Blaas, J., van Houwelingen, A.M.M.L., De Vos, R.C.H., van der Voet, H., Jamsen, R.C., Guis, M., Mol, J., Davis, R.W., Schena, M., van Tunen, A.J., O'Connell, A.P. 2000. Identification of the SAAT gene involved in strawberry flavour biogenesis by use of DNA microarray. Plant Cell 12: 647-661.
53. Fellman, J.K., Miller, T.W., Mattinson, D.S., Mattheis, J.P. 2000. Factors that influence biosynthesis of volatile flavor compounds in apple fruits. Horticultural Science 35: 1026-1033.
54. Galliard, T., Matthew, J.A. 1977. Lipoxygenase-mediated cleavage of fatty acids to carbonyl fragments in tomato fruits. Phytochemistry 16: 339-343.
55. Baysal, T., Demirdoven, A. 2007. Lipoxygenase in fruits and vegetables: A review. Enzyme and Microbial Technology 40: 491-496.
56. Qi, Q., Sparace, K.F., Sparace, S.A., 1995. The utilization of glycolytic intermediates as precursors for fatty acid biosynthesis by pea root plastids. Plant Physiology 107:413-419.
57. Spiers, J., Lee, E., Holt, K., Yong-Duk, K., Scott, N.S., Loveys, B., Schuch, W. 1998. Genetic manipulation of alcohol dehydrogenase levels in ripening tomato fruit affects the balance of some flavour aldehydes and alcohols. Plant Physiology 117: 1047-1058.
58. Longhurst, T.J., Tung, H.F., Brady, C.J. 1990. Developmental regulation of the expression of alcohol dehydrogenase in ripening tomato fruits. Journal of Food Biochemistry 14: 421-433.
59. Souleyre, E.J.F., Greenwood, D.R., Friel, E.N., Karunairetnam, S., Newcomb, R.D. 2005. An alcohol acyl transferase from apple (cv. Royal Gala), MpAAT1, produces esters involved in apple fruit flavour. FEBS Journal 272: 3132-3144.
60. Delfilippi, B.G., Kader, A.A., Dandekar, A.M. 2005. Apple aroma: alcohol acyltransferase, a rate limiting step for ester biosynthesis, is regulated by ethylene. Plant Science 168: 1199-1210.
61. Harada, M., Ueda, Y., Iwata, T. 1985. Purification and some properties of alcohol acyltransferase from banana fruit. Plant Cell Physiology 1067-1074.
62. Yahyaoui, F.E.L., Wongs-Aree, C., Latche, A., Hackett, R., Grierson, D., Pech, J.C. 2002. Molecular and biochemical characteristics of a gene encoding an alcohol acyl-transferase involved in the generation of aroma volatile esters during melon ripening. The European Journal of Biochemistry 269: 2359-2366.
63. El-Sharkawy, I., Manriquez, D., Flores, F.B., Regad, F., Bouzayen, M., Latche, A., Pech, J.C. 2005. Functional characterization of a melon alcohol acyl-transferase gene family involved in the biosynthesis of ester volatiles. Identification of the crucial role of threonine residue for enzyme activity. Plant Molecular Biology 59: 345-362.
64. Perez, A.G., Sanz, C., Olias, J.M. 1993. Partial purification and some properties of alcohol acyltransferase from strawberry fruits. Journal of Agricultural and Food Chemistry 41: 1462-1466.
65. Rodriguez-Concepcion, M., Boronat, A. 2002. Elucidation of the methylerythritol phosphate pathway for isoprenoid biosynthesis in bacteria and

plastids. A metabolic milestone achieved through genomics. Plant Physiology 130: 1079-1089.
66. Tholl, D. 2006. Terpene synthases and the regulation, diversity and biological roles of terpene metabolism. Current Opinion in Plant Biology 9: 297-304.
67. Simkin, A.J., Schwartz, S.H., Auldridge, M., Taylor, M.G., Klee, H.J. 2004. The tomato carotenoid cleavage dioxygenase 1 genes contribute to the formation of the flavor volatiles β-ionone, pseudoionone, and geranylacetone. The Plant Journal 40: 882-892.
68. Bolwell, G.P., Bozak, K., Zimmerlin, A. 1994. Plant Cytochome P450. Phytochemistry 37(6): 1491-1506.
69. Bouwmeester, H.J., Gershenzon, J., Konings, M.C.J.M., Croteau, R. 1998. Biosynthesis of the monoterpenes limonene and carvone in the fruit of caraway. Plant Physiology 117: 901-912.
70. Wyllie, S.G., Fellman, J.K. 2000. Formation of volatile branched chain esters in bananas (*Musa sapientum* L.). Journal of Agricultural and Food Chemistry 48: 3493-3496.
71. Myers, M.J., Issenberg, P., Wick, E.L. 1970. L-leucine as a precursor of *iso*amyl alcohol and *iso*amyl acetate, volatile aroma constituents of banana fruit discs. Phytochemistry 9: 1693-1700.
72. Perez, A.G., Rios, J.J., Sanz, C., Olias, J.M. 1992. Aroma components and free amino acids in strawberry variety chandler during ripening. Journal of Agricultural and Food Chemistry 40: 2232-2235.
73. Perez, A.G., Olias, R., Luaces, P., Sanz, C. 2002. Biosynthesis of strawberry aroma compounds through amino acid metabolism. Journal of Agricultural and Food Chemistry 50: 4037-4042.
74. Tominaga, Y., Dubourdieu, D. 2000. Identification of cysteinylated aroma precursors of certain volatile thiols in passion fruit juice. Journal of Agricultural and Food Chemistry 48: 2874-2876.
75. Leustek, T., Martin, M.N., Bick, J.A., Davies, J.P. 2000. Pathways and regulation of sulfur metabolism revealed through molecular and genetic studies. Annual review of Plant Physiology and Plant Molecular Biology 51: 141-165.
76. Noctor, G., Arisi, A.C.M., Jouanin, L., Kunert, K.J., Rennenberg, H., Foyer, C.H. 1998. Glutathione: biosynthesis, metabolism and relationship to stress tolerance explored in transformed plants. Journal of Experimental Botany 49(321): 623-647.
77. Edwards, R., Dixon, D. 2005. Glutathione transferases and gamma-glutamyl transpeptidases. Methods In Enzymology 401: 169-186.
78. Rausch, T., Wachter, A. 2005. Sulfur metabolism: a versatile platform for launching defence operations. Trends in Plant Science 10 (10): 503-509.
79. Halkier, B.A., Gershenzon, J., 2006. Biology and biochemistry of gucosinolates. The Annual Review of Plant Biology 57: 303-333.
80. Lancaster, J.E., Kelly, K.E. 1983. Quantitative-analysis of the *S*-alk(en)yl-L-cysteine sulfoxides in onion (*Allium cepa* L). Journal of the Science of Food and Agriculture 34(11): 1229-1235.
81. Jones, M.G., Hughes, J., Tregova, A., Milne, J., Tomsett, A.B., Collin, H.A. 2004. Biosynthesis of the flavour precursors of onion and garlic. Journal of Experimental Botany 55 (404): 1903-1918.

82. Lancaster, J.E., Shaw, M.L. 1989. γ-Glutamyl peptides in the biosynthesis of *S*-alk(en)yl-L-cysteine sulphoxides (flavour precursors) in *Allium*. Phytochemistry 28(2): 455-460.
83. Tominaga, T., Furrer, A., Hendry, R., Dubourdieu, D. 1998. Identification of new volatile thiols in the aroma of *Vitis vinifera* L. var. Sauvignon blanc wines. Flavour and Fragrance Journal 13: 159-162.
84. Tominaga, T., Peyrot des Gachons, C., Dubourdieu, D. A new type of flavor precursors in *Vitis vinifera* L. cv. Sauvignon blanc: *S*-cysteine conjugates. Journal of Agricultural and Food Chemistry 46:5215-5219.
85. Jocelyn, P.C. 1972. Chapter 1 Introduction. In "Biochemistry of the SH Group" Jocelyn, P.C. Ed., Academic press Inc., London. p.15-19.
86. Lalel, H.J.D., Singh, Z., Tan, S.C. 2003. Glycosidically-bound aroma volatile compounds in the skin and pulp of "Kensigton Pride" mango at different stages of maturity. Postharvest Biology and Biotechnology 29: 205-218.
87. Engel, K.H., Tressl, R. 1983. Formation of aroma components from nonvolatile precursors in passion fruit. Journal of Agricultural and Food Chemistry 31: 998-1002.
88. Chassagne, D., Crouzet, J., Bayonove, C.L., Baumes, R.L. 1998. Identification of passion fruit glycosides by gas chromatography/ mass spectrometry. Journal of Agricultural and Food Chemistry 46:4352-4357.
89. Parada, F., Duque, C., Fujimoto, Y. 2000. Free and bound volatile composition and characterization of some glucoconjugates as aroma precursors as aroma precursors in melon de olor fruit pulp (*Sicana odorifera*). Journal of Agricultural and Food Chemistry 48: 6200-6204.
90. Genard, M., Gouble, B. 2005. ETHY. A theory of fruit climacteric ethylene emission. Plant Physiology 139:531-545.
91. Gershater, M.C., Edwards, R. 2007. Regulating biological activity in plants with carboxylesterases. Plant Science 173: 579-588.
92. Wheelock, C.E., Shan, G., Ottea, J. 2005. Overview of carboxylesterases and their role in the metabolism of insecticides. Journal of Pesticide Science 30(2): 75-83.
93. Bartley, I.M., Stevens, W.H. 1981. Carboxylic ester hydrolases of apple. Journal of Experimental Botany 32 (129): 741-751.
94. Goodenough, P.W., Entwistle, T.G. 1982. The hydrodynamic properties and kinetic constants with natural substrates of the esterase from *Malus pumila* fruit. European Journal of Biochemistry 127: 145-149.
95. Faber, K. 2004. Introduction and background information. In "Biotransformations in Organic Chemistry". Faber, K. Ed. Fifth Edition, Springer, Germany. p.1-28.
96. Serra, S., Fuganti, C., Brenna, E. 2005. Biocatalytic preparation of natural flavours and fragrances. Trends in Biochemistry 23 (4): 193-198.
97. Vandamme, E.J. 2003. Bioflavours and Fragrances via fungi and their enzymes. Fungal Diversity 13: 153-166.
98. Krings, U., Hinz, M., Berger, R.G. 1996. Degradation of [^{2}H]phenylalanine by the basidomycete *Ichnoderma benzoinum*. Journal of Biotechnology 51: 123-129.
99. Rao, S.R., Ravishankar, G.A. 2000. Vanilla flavour: Production by conventional and biotechnological routes. Journal of the Science of Food and Agriculture 80 (3): 289-304.

100. Janssens, L., Depooter, H.L., De Mey, L., Vandamme, E.J., Schamp, N.M. 1989. Fusel oil as a precursor for the microbial production of fruity flavours. Mededelungen Faculteit Landbouwwetenzhappen Rijksuniversiteit, Gent 54 (4a):1387-1391.
101. Romagnoli, L.G., Knorr, D. 1988. Effects of ferulic acid treatment on growth and flavour development of cultured *Vanilla planifolia* cells. Food Biotechnology 2: 93-104.
102. Westcott, R.J., Chetham, P.S.J., Barraclough, A.J. Use of organized viable, *Vanilla* plant aerial roots for the production of vanillin. Phytochemistry 35: 135-138.
103. Faber, K. Ed. 2004. Chapter 2.1 Hydrolytic Reactions. In "Biotransformations in Organic Chemistry". Faber, K. Ed. Fifth Edition, Springer, Germany. p.29-176.
104. Gubicza, L., Kabiri-Badr, A., Keoves, E., Belafi-Bako, K. 2000. Large-scale enzymatic production of natural flavour esters in organic solvent with continuous water removal. Journal of Biotechnology 84: 193-196.
105. Hamid, N.S.A., Zen, H.B., Tein, O.B., Halifah, Y.M., Saari, N., Bakar, F.A. 2004. Screening and identification of extracellular lipase-producing thermophilic bacteria from a Malaysian hot spring. World Journal of Microbiology and Biotechnology 19 (9): 961-968.
106. Karra-Chaabouni, M., Ghamgui, H., Bezzine, S., Rekik, A., Gargouri, Y. Production of flavour esters by immobilized *Staphylococcus simulans* lipase in a solvent-free system. Process Biochemistry 41 (7): 1692-1698.
107. Manjon, A., Iborra, J.L., Arocas, A. 1991. Short-chain flavour ester synthesis by immobilized lipase in organic media. Biotechnology Letters 13 (5): 339-344.
108. Carta, G., Gainer, J.L., Benton, A.H. 1991. Enzymatic-synthesis of esters using an immobilized lipase. Biotechnology and Bioengineering 37 (11): 1004-1009.
109. Athawale, V., Manjrekar, N., Athawale, M. 2001. Enzymatic synthesis of chiral menthyl methacrylate monomer by pseudomonas cepacia lipase catalysed resolution of (+/-)-menthol. Journal of Molecular Catalysis B: Enzymatic 16: 169-173.
110. Pamies, O., Backvall, J.E. 2002. Enzymatic kinetic resolution and chemoenzymatic dynamic kinetic resolution of δ-hydroxy esters. An efficient route to chiral δ-lactones. Journal of Organic Chemistry 67: 1261-1265.
111. Lagarde, D., Nguyen, H.K., Ravot, G., Wahler, D., Reymond, J.L., Hills, G., Veit, T., Lefevre, F. 2002. High-throughput screening of thermostable esterases for industrial bioconversions. Organic Process Research and Development 6: 441-445.
112. van Gemert, L.J. 1999. Compilation of odour thresholds in air and water. Boelens Aroma Chemical Information Service (BACIS), The Netherlands.
113. Wakabayashi, H., Wakabayashi, M., Eisebnreich, W., Engel, K.H. 2003. Stereoselectivity of the generation of 3-mercaptohexanal and 3-mercaptohexanol by lipase-catalyzed hydrolysis of 3-acetylthioesters. Journal of Agricultural and Food Chemistry 51: 4349-4355.
114. Rhylid, R.B., Matthey-Doret, W., Blank, I., Fay, L.B., Juillerat, M.A. 2002. Lipase-assisted generation of 2-methyl-3-furanthiol and 2-furfurylthiol from thioacetates. Journal of Agricultural and Food Chemistry 50: 4087-4090.

115. Cavaille-Lefebvre, D., Combes, D. 1997. Lipase synthesis of short-chain flavour thioesters in solvent-free medium. Biocatalysis and Biotransformation 15: 265-279.
116. Wakabayashi, H., Wakabayashi, M., Eisebnreich, W., Engel, K.H. 2004. Stereochemical course of the generation of 3-mercaptohexanal and 3-mercaptohexanol by β-lyase-catalysed cleavage of cysteine conjugates. Journal of Agricultural and Food Chemistry 52: 110-116.
117. Wakabayashi, H., Wakabayashi, M., Eisebnreich, W., Engel, K.H. 2002. Stereoselectivity of the β-lyase-catalysed cleavage of S-cysteine conjugates of pulegone. European Food Research and Technology 215: 287-292.
118. Jameson, S.B. 2001. Organosulphur chemistry- Aroma chemicals and beyond part II. Speciality Chemicals 21: 17-24.
119. Jameson, S. B. 2005. Sulfur compounds. In "Chemistry and Technology of Flavours and Fragrances". Rowe, D., Ed. Blackwell Publishing Limited. p.116-142.
120. Posthumus, M.A., Van Beek, T.A., Collins, N.F., Graven, E.H. 1996. Chemical composition of the essential oils of *Agathosma betulina*, *A. crenulata* and an *A. betulina x crenulata* hybrid (Buchu). Journal of Essential Oil Research 8: 223-228.
121. Jones, A., Davies, H.M., Voelker, T.A. 1995. Palmitoyl-Acyl Carrier Protein (ACP) thioesterase and the evolutionary origin of plant acyl-ACP thioesterases. The Plant Cell 7: 359-371.
122. The Council of the European Communities, 1988. Council Directive 88/388/EEC of 22nd June 1988.
123. US Code of Federal Regulations, 1985. 21, 101.22a.3. Food and Drug Administration, Washington, D.C.
124. Bradford, M.M., Williams, W.L. 1976. New, rapid, sensitive method for protein determination. Federation Proceedings 35 (3): 274-274.
125. Gershater, M.C., Cummins, I., Edwards, R. 2007. Role of carboxylesterase in herbicide bioactivation in *Arabidopsis thaliana*. The Journal of Biological Chemistry 282 (29): 21460-21466.
126. Thompson, J.D., Higgins, D.G. Gibson, T.J. 1994. CLUSTAl-W Improving the sensitivity of progressive multiple alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Research 22 (22): 4673-4680.
127. Barnes, W.M. 1994. PCR amplification of up to 35 kb DNA with high-fidelity and high-yield from Lambda-bacteriophage templates. Proceedings of the National Academy of Sciences of the United States of America 91 (6): 2216-2220.
128. Zuker, M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Research 31 (13): 3406-3415.
129. Jocelyn, P.C. 1972. Chapter 6 Assay of thiols and disulphides. In "Biochemistry of the SH Group" Jocelyn, P.C. Ed., Academic press Inc., London. p.137-159.
130. Woodward, G.E., Fry, E.G. 1932. The determination of blood glutathione. The Journal of Biological Chemistry 97(2): 465-482.
131. Benesch, R., Benesch, R.E. 1957. determination of sulfhydryl and disulfide groups by specific proton displacement. Biochimica et Biophysica Acta 23(3): 643-644.

132. Roberts, E., Rouser, G. 1958. Spectrophotometric Assay for reaction of *N*-ethylmaleimide with sulfhydryl Groups. Analytical Chemistry 30(7): 1291-1292.
133. Alexander, N.M. 1958. Spectrophotometric Assay for sulfhydryl groups using *N*-ethylmaleimide. Analytical Chemistry 30(7): 1292-1294.
134. Boyer, P.D. 1954. Spectrophotometric study of the reaction of protein sulfhydryl groups with organic mercurials. Journal of the American Chemical Society 76: 4331-4337.
135. Newton, G.L., Dorian, R., Fahey, R.C. 1981. Analysis of biological thiols: Derivatization with monobromobimane and separation by reverse-phase high-performance liquid chromatography. Analytical Biochemistry: 114: 383-397.
136. Ellman, G.L. 1959. Tissue sulfhydryl groups. Archives of Biochemistry and Biophysics 82: 70-77.
137. Vignais, P.V., Zabin, I. 1958. Synthesis and properties of palmityl adenylate, palmityl coenzyme-A, and palmityl glutathione. Biochimica et Biophysica Acta 29(2): 263-269.
138. Sanner, T., Pihl, A. 1962. Alleged splitting of thioesters by *P*-chloromercuribenzoate. Biochimica et Biophysica Acta 62(1): 171.
139. Fahey, R.C., Newton, G.L., Dorian, R., Kosower, E.M. 1981. Analysis of biological thiols: Quantitative determination of thiols at the picomole level based upon derivatization with monobromobimanes and separation by cation-exchange chromatography. Analytical Biochemistry 2(1): 357-365.
140. Passerini, R.C. 1960. In "Organic Sulphur Compounds" Kharasch, N. Ed., Pergamon, New York. P. 57.
141. Riddles, P.W., Blakeley, R.L., Zerner, B. 1979. Ellman's reagent – 5,5'-dithiobis(2-nitrobenzoic acid) re-examination. Analytical Biochemistry 94: 75-81.
142. Lelievre, J.M., Latche, A., Jones, B., Bouzayen, M., Pech, J.C. 1997. Ethylene and fruit ripening. Physiologia Plantarum 101: 727-739.
143. Cummins, I., Edwards, R. 2004. Purification and cloning of an esterase from the weed black-grass (*Alopecurus myosuroides*), which bioactivates aryloxyphenoxypropionate herbicides. The Plant Journal 39: 894-904.
144. Cao, L. 2005. Immobilised enzymes: science or art? Current Opinion in Chemical Biology 9: 217-226.
145. Tapp, E.J., Brassington, D. 2007. Detecting volatiles with Ellman's reagent. Speciality Chemicals 27(9): 38-39.
146. Brummell, D.A., Harpster, M.H. 2001. Cell wall metabolism in fruit softening and quality and its manipulation in transgenic plants. Plant Molecular Biology 47:311-340.
147. Downs, C.G., Brady, C.J., Gooley, A. 1992. Exopolygalacturonase protein accumulates late in peach fruit ripening. Physiologia Plantarum 85(2): 133-140.
148. Gerardi, C. Blando, F., Santino, A., Zacheo, G. 2001. Purification and characterisation of a β-glucosidase abundantly expressed in ripe sweet cherry (*Prunus avium* L.) fruit. Plant Science 160: 795-805.
149. Peumans, W.J., Barre, A., Derycke, V., Rouge, P., Zhang, W., May, G.D., Delcour, J.A., Van Leuven, F., Van Damme, E.J.M. 2000. Purification, characterization and structural analysis of an abundant β-1,3-glucanase from banana fruit. European Journal of Biochemistry 267: 1188-1195.

150. Williamson, G. 1991. Purification and characterization of pectin acetylesterase from orange peel. Phytochemistry 30 (2): 445-449.
151. McCollum, T.G., Doostdar, H., Mayer, R.T., McDonald, R.E. 1997. Characterization of chitinases and β-1,3-glucanases in grapefruit flavedo during fruit development. Physiologia Plantarum 99: 486-494.
152. Glover, H., Brady, C. 1994. Purification of three pectin esterases from ripe peach fruit. Phytochemistry 37: 949-955.
153 Janson, J.C, Ryden, L. Eds. 1989. Protein Purification: Principles, high resolution methods and applications. Wiley. 679 pages.
154. Scopes, R.K. Ed. 1994. Protein Purification: Principles and Practice. Springer, New York, USA. 375 pages.
155. Hoffmann, E.D., Stroobant, V. 2002. Mass Spectrometry: Principles and Applications. 2nd Edition. Chichester, New York, Wiley. 393 pages.
156 Li, S.C., Han, J.W., Chen, K.C., Chen, C.S. 2001. Purification and characterization of isoforms of β-galactosidases in mung bean seedlings. Phytochemistry 57: 349-359.
157. Chiang, J.Y., Shu, S.W., Ko, C.W., Wang, C.S. 2006. Biochemical characterization of a pollen-specific cDNA encoding polygalacturonase in *Lilium longiflorum*. Plant Science 170: 433-440.
158. Bordenave, M., Goldberg, R., Huet, J.C., Pernollet, J.C. 1995. A novel protein from mung bean hypocotyl cell walls with acetyl esterase activity. Phytochemistry 38 (2): 315-319.
159. Jansen, E.F., Jang, R., Mac Donnell, L.R. 1947. Citrus acetylesterase. Archives of Biochemistry 15 (3): 415-431.
160. Breton, C., Bordenave, M., Richard. L., Pernollet, J.C., Huet, J.C., Perez. S., Goldberg, R. 1996. PCR cloning and expression of a cDNA encoding a pectinacetylesterase from *Vigna radiate* L. Federation of European Biochemical Societies 388: 139-142.
161. Coutinho, P.M., Henrissat, B. 1999. Carbohydrate-active enzymes: an integrated approach. In "Recent advances in carbohydrate engineering". Gilbert, H.J., Davies, G.L., Svensson, B., Henrissat, B. Eds. Royal Society of Chemistry p. 3-12.
162. Davies, G.J., Gloster, T.M., Henrissat, B. 2005. Recent structural insights into the expanding world of carbohydrate-active enzymes. Current opinion in structural biology 15: 637-645.
163. Ooms, A.J.J., Breebaart-Hansen, J.C.A.E., Ceulen, B.I. 1996. The reaction of organophosphorous compounds with hydrolytic enzymes II. The inhibition of citrus acetylesterase. Biochemical pharmacology 15: 17-30.
164. Cummins, I., McAuley, K., Fordham-Skelton, A., Schwoerer, R., Steel, P.G., Davis, B.G., Edwards, R. 2006. Unique regulation of the active site of the serine esterase S-formylglutathione hydrolase. Journal of Molecular Biology 359: 422-432.
165. Molgard, A., Kauppinen, S., Larsen, S. 2000. Rhamnogalacturonan acetylesterase elucidates the structure and function of a new family of hydrolases. Structure with Folding and Design 8 (4): 373-383.
166. Fischer, R.L., Bennett, A.B. 1991. Role of cell wall hydrolases in fruit ripening. Annual Review of Plant Physiology and Plant Molecular Biology 42: 675-703.

167. Nielsen, J.E., Christensen, T.M.I.E. 2002. Distribution of pectin methyl esterase and acetylesterase in the genus *Citrus* visualized by tissue prints and chromatography. Plant Science 162: 799-807.
168. Crosby, C. 1992. Chirality in Industry – an overview. In "Chirality in Industry". Collins, A.N., Sheldrake, G.N., Crosby, J. Eds. Wiley, Chicester, p. 1-66.
169. Faber, K. 2004. Appendix 5.3 Suppliers of Enzymes. In "Biotransformations in Organic Chemistry". Faber, K. Ed. Fifth Edition, Springer, Germany, p.431.
170. Strader, M.B., Tabb, D.L., Hervey, W.J., Pan, C., Hurst, G.B. 2006. Efficient and specific trypsin digestion of microgram to nanogram quantities of proteins in organic-aqueous solvent systems. Analytical Chemistry 78: 125-134.
171. Felsentein, J. 1989. PHYLIP-phylogeny inference package. Cladistics 5: 164–166.
172. López-Gómez, R., Gómez-Lim, M.A. 1992. Changes in mRNA and. protein synthesis during ripening in mango fruit. Journal of Plant Physiology 141: 82-87.
173. Chomczynski, P., Sacchi, N. 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Analytical Biochemistry 162: 156-159.
174. Mathews, D.H., Sabina, J., Zuker, M., Turner, D.H. 1999. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. Journal of Molecular Biology 288: 911-940.
175. Dixon, D.P., Lapthorn, A., Madesis, P., Mudd, E.A., Day, A. and Edwards, R. 2008. Binding and glutathione conjugation of porphyrinogens by plant glutathione transferases. Journal of Biological Chemistry 283: 20268-20276.
176. Tartof, K.D., Hobbs, C.A. 1987. Improved media for growing plasmid and cosmid clones. Focus 9:12.
177. Dixon, D.P., Edwards, R. 2008. Functional screening of the *Arabidopsis* glutathione transferase super-family for ligand binding *in vivo. In press.*
178. Ollis, D.L., Cheah, E., Cygler, M., Dijkstra, B., Frolow, F., Franken, S.M., Harel, M., Remington, S.J., Silman, I., Schrag, J., Sussman, J.L., Verschueren, K.H.G., Goldman, A. 1992. The α/β hydrolase fold. Protein Engineering 5(3): 197-211.
179. Jones, D.T. 1999. Protein secondary structure prediction based on position-specific scoring matrices. Journal of Molecular Biology 292 (2): 195-202.
180. Williamson, G., Kroon, P.A., Faulds, C.B. 1998. Hairy plant polysaccarides: a close shave with microbial esterases. Microbiology 144: 2011-2023.
181. Industry Research Report: Juices. 2008. Global Industry Guides, Datamonitor. p. 1-157.
182. Pasta, P., Verga, R., Zambianchi, F., Daminati, M. 2004. Acetyl esterase from Mediterranean oranges: partial purification, immobilisation and biotransformations. Biocatalysis and Biotransformations 22 (3):221-224.
183. Industry Research Report: World Markets and Trade - Orange Juice. 2007. United States Department of Agriculture.

184. Waldmann, H., Heuser, A., Reidel, A. 1994. Selective enzymatic deprotection of hydroxy- and amino groups in carbohydrates and nucleosides. Synlett 1: 65-67.
185. Faber, K. Ed. 2004. Biotransformations in Organic Chemistry. Fifth Edition, Springer, Germany. 454 pages. Chapter 3 Special techniques. p. 334-383.
186. Willats, W.G.T., McCartney, L., Mackie, W., Knox, J.P. 2001. Pectin: cell biology and prospects for functional analysis. Plant Molecular Biology 47: 9-27.
187. Nentwig, W. 1999. The importance of human ecology at the threshold of the next millenium: How can population growth be stopped? Natur Wissenschaften 86: 411-421.
188. Bartlett, A.A. 1997. Reflections on sustainability, population growth, and the environment – revisited. Renewable Resources Journal 15 (4): 6-23.
189. Brassington, D., Edwards, R. 2006. Learning from nature: High impact aroma chemistry. Perfumer and Flavourist 31 (9): 30-32.
190. Martin, J. 2006. Price considerations for use of soy feedstocks for the chemicals industry. Industrial Biotechnology 2 (2): 92-95.
191. Supriyadi Suzuki, M., Wu, S., Tomita, N., Fujita, A., Watanabe, N. 2003. Biogenesis of volatile methyl esters in snake fruit (*Salacca edulis*, Reinw) cv. Pondoh. Bioscience, Biotechnology and Biochemistry 67: 1267-1271.
192. Fischer, R.L., Bennett, A.B. 1991. Role of cell wall hydrolases in fruit ripening. Annual review of plant physiology and plant molecular biology 42: 675-703.
193. Pelloux, J., Rustérucci, C., Mellerowicz, E.J. 2007. New insights into pectin methylesterase structure and function. Trends in Plant Science 12: 267-277.