Elucidation of Insect Olfaction in the Pea Aphid, *Acyrthosiphon pisum*

Cassie Marie Sims

A thesis submitted to the University of Nottingham

for the degree of Doctor of Philosophy

2020

ROTHAMSTED RESEARCH

University of Nottingham
UK | CHINA | MALAYSIA

Declaration

I hereby declare that I have composed this thesis. The work of this thesis is a record of my own work.

Cassie Marie Sims

Contents

Acknowledgements

Firstly, I would like to thank my supervisors Mike, Neil and Rob, with extra special thanks to David, for their guidance, corrections, criticisms, and advice. I suspect I am not the easiest person to have as a PhD student, so thanks for putting up with me. They are of course, not the only ones who helped with the science, so thanks to Jedd for endless help and teaching me everything about proteins, Suzanne, John C., Dan B., Ian B., Vanderson, Bruna, and all my other lab buddies over the years. Also thank you to Jing-Jiang for providing the plasmids to kick start this project and to David Jones for invaluable advice on fluorescence assays, NMR and OBPs in general.

Thanks to the teams who kept me sane, our pub quiz team 'We thought we were speed dating', who kept on despite the global pandemic, and my roller derby league, Hertfordshire Roller Derby. A special thanks goes to my roller derby 'mum', Hel-Razor, who I am sure will show up at my house randomly for a skate no matter where in the world I move to. I also cannot go without thanking the NHS. Of course, I am the idiot who breaks their ankle roller skating during their PhD, but without being too dramatic, without the NHS I would not have made it this far, so thank you free healthcare!

Finally, a PhD is not possible without the love and support of your friends and family and all those who fit into the blurred line in between. I am not one for soppy remarks, but if there ever were a time here it is. Thanks to my Mum and Dad, for supporting me through their endless confusion about what I actually do. A huge thanks to Amma and Darja, the two best buds I have picked up over these last 4 years who have been complete rocks and whom I know I will be in my life for a very long time going forward. Thank you to Donna for being an amazing friend and for never judging my infinite vending machine trips. A big thanks goes to my housemates dealt with my singing, sarcasm and endless moaning for years, Matt and Dan. Thank you to my fellow Keelite, ultimate theme park buddy and ranting queen Krystal for supporting me and keeping me inundated with funny and inspiring content throughout these trying times. Not everyone is getting a mention because it would be a thesis in itself, but to all of the other poor innocent friends, old and new, who had to put up with me over the last few stressful years, thanks for the chocolate, Pepsi Max, Pav drinks, PhD coffees, Zoom chats, murder mystery parties and everything else.

Abbreviations

$[\alpha]_D$	specific rotation
ANS	1-anilinonaphthalene- 8-sulfonic acid
c	concentration
d	doublet
D_2O	deuterated water
DCM	dichloromethane
de	diastereomeric excess
DMSO	dimethyl sulfoxide
DTT	dithiothreitol
EDG	electron donating group
ESI	electrospray ionisation
EtOAc	ethyl acetate
EWG	electron withdrawing group
FPLC	fast protein liquid chromatography
FRET	Förster resonance transfer
GPCR	G-coupled protein receptor
h	hour
H_2O	water
HOMO	highest occupied molecular orbital
HPLC	high performance liquid chromatography
HRMS	high resolution mass spectrometry
Hz	Hertz
IEDDA	Inverse electron demand Diels-Alder
IPA	isopropyl alcohol
IPTG	isopropyl-β-D-thiogalactopyranoside
IR	infrared
J	coupling constant
K_D	dissociation constant
kDa	Kilo Daltons
LB	Luria-Bertoni
LUMO	lowest unoccupied molecular orbital

m	multiplet
M	molar
Me	methyl
MeOH	methanol
$MgSO_4$	magnesium sulfate
min	minute(s)
mmol	milli moles
MS	mass spectrometry
m/z	mass to charge ration
$NaBH_4$	sodium borohydride
NPN	*N*-Phenylnaphthalen-1-amine
NMR	nuclear magnetic resonance
OBP	odorant-binding protein
OD	optical density
OR	odorant receptor
ORN	odorant receptor neuron
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
Pet. Ether	petroleum ether 40-60
q	quadruplet
RT	room temperature
s	singlet
SDS	sodium dodecyl sulfate
STD	saturation transfer difference
t	triplet
TEMED	N,N,N',N'-tetramethylethane-1,2-diamine
TFA	trifluoroacetic acid
TLC	thin layer chromatography
TOF	time-of-flight
V	volts

Abstract

Aphids (Homoptera; Aphididae) are major pests on arable and horticultural crops. Aphids communicate using olfaction, employing semiochemicals such as the alarm pheromone, (*E*)-β-farnesene, and the sex pheromone components, (4*aS*,7*S*,7*aR*)-nepetalactone and (1*R*,4*aS*,7*S*,7a*R*)-nepetalactol. An understanding of olfaction in aphids, specifically molecular recognition, and discrimination of odorants, can help when designing novel protection tools. However, little is understood about the aphid olfactory system, specifically, the role of odorant-binding proteins (OBPs), small, soluble proteins that are found in high concentrations in sensory organs (antennae). In this thesis, the molecular recognition of the sex pheromone in the pea aphid, *Acyrthosiphon pisum* was investigated by exploring the roles of olfactory proteins such as OBPs and odorant receptors (ORs)

Initially, aphid sex pheromone components, (4*aS*,7*S*,7*aR*)-nepetalactone and (1*R*,4*aS*,7*S*,7a*R*)-nepetalactol, and their respective non-naturally occurring enantiomers, (4*aR*,7*R*,7*aS*)-nepetalactone and (1*S*,4*aR*,7*R*,7a*S*)-nepetalactol were synthesised, with the synthesis optimised at various points, giving a final yield of approximately 1%. *A. pisum* OBPs were also expressed on a litre-scale and purified in large quantities (> 1mg). Binding activity between *A. pisum* OBPs and sex pheromone components was screened using *in silico* methods. These screens revealed OBP6 as a strong candidate for binding of the sex pheromone components due to low K_Ds, and OBP9 as a protein with little binding activity with the ligands of interest. Furthermore, potential strong-binding analogues of nepetalactol and nepetalactone were predicted through *in silico* methods. These analogues have the potential to act as olfactory disrupters and may be a valuable pest control method in future. *In silico* work also predicted the binding sites for electrophysiologically active molecules in ORs, whilst predicting the potential sex pheromone binding OR in aphids.

After successful *in silico* screens, a range of fluorescence binding assays were performed with OBP6 and OBP9 and the sex pheromone components. The fluorescence assays used a non-traditional approach, observing intrinsic fluorescence of the tryptophan residues both alone and in the presence a range of fluorescent probes. These assays confirmed the predicted results of the *in silico* screens. OBP6 was seen to bind sex pheromone components, with a particularly low K_D of 1.30 ± 0.60 μM observed for the interaction between OBP6 and (4*aS*,7*S*,7*aR*)-nepetalactone. Due to OBP6's high expression in the aphid antennae, these results suggest it may play a role in sex pheromone perception.

To further explore the interaction between OBPs and the sex pheromone components, other methods, including mass spectrometry, nuclear magnetic resonance (NMR) and gas-chromatography (GC), were employed. Mass spectrometry experiments were unsuccessful, but Saturation Transfer Difference (STD) NMR and a biphasic GC assay both gave binding results for OBP6 consistent with the *in silico* and fluorescence experiments. The STD-NMR assay also allowed for epitope mapping of the binding interactions between the sex pheromone components and OBP6, which aligned with the predicted conformations from *in silico* screens.

Overall, OBP6 was shown to bind the sex pheromone component (4*aS*,7*S*,7*aR*)-nepetalactone with a higher affinity than other ligands, including the alarm pheromone (*E*)-β-farnesene. The results of this work suggest a role for OBP6 in sex pheromone binding. All assays focused on the enantiomeric discrimination potential of OBP6. Though no clear discrimination was observed in this study, levels of discrimination could be seen in the fluorescence assays. Further work is required to fully elucidate OBP6's potential role.

Future work will focus on the exploration of OBP6's role in aphid olfaction, from behavioural studies to further binding assays and structural determination. Additionally, results from this thesis may be adapted for use in the development of novel pest control tools.

1. Introduction

1.1 THE PEA APHID, *ACYRTHOSIPHON PISUM*

1.1.1 Aphid General Biology

Aphids (Homoptera: Aphididae) are small, soft-bodied insects which are distributed worldwide, though most prevalent in temperate regions. They are pests of many agricultural and horticultural crops. There exists a diverse range of aphid species, many of which are capable of migration over great distances.[1]

In 2010, the pea aphid, *Acyrthosiphon pisum* (Figure 1.1), was the first hemipteran insect to have its genome fully sequenced.[2] The pea aphid was chosen as a model organism due to its complex life cycle, symbiosis with bacteria, polyphenism and interaction with host plants.[3] Furthermore, its relatively larger size and ease of culturing make it ideal to study experimentally in controlled laboratory conditions.[3]

Figure 1.1: The pea aphid, *Acyrthosiphon pisum.* Image from Rothamsted Visual Communications Unit.

1.1.2 Aphid Life History

The life cycle of aphids is complex and seasonal (Figure 1.2). Parthenogenetic (asexual) reproduction usually occurs throughout the summer[4,5], until a reduction in daylight hours causes the production of winged males and females (gynoparae). The gynoparae migrate to the winter host, producing wingless, sexual females (oviparae) that release a sex pheromone to attract males, and subsequently reproduce sexually.[4–6] The females lay cold-resistant, overwintering eggs, from which the stem mother, or fundatrix, will hatch initiating parthenogenesis the following year.[2]

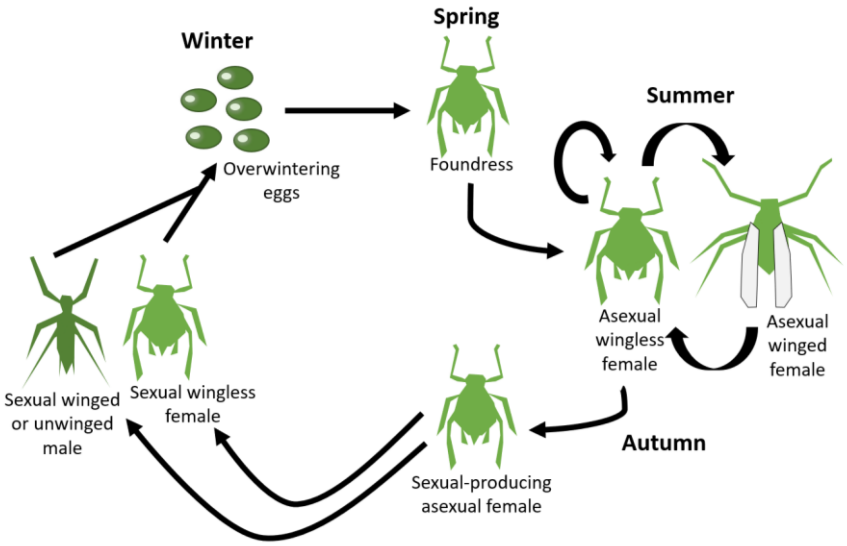


Figure 1.2: The life cycle of the pea aphid, *Acythrosiphon pisum*. Image taken adapted from[3]

Aphids can either spend their entire life cycle on one host (autocecious) or aternate between hosts (heteroecious).[7] The pea aphid is autocecious and generally lives on legume plants, such as the broad bean, *Vicia faba.*[8,9]

Male aphids are only produced as a result of seasonal changes initiating the sexual reproduction phase of the life cycle.[4] In female aphids the winged phenotype is an environmentally determined polyphenism, used for colonisation purposes, wheras in males wing polyphenism is genetically determined and both winged and unwinged male aphids can exist.[10] Although there is an associated fitness cost with developing wings, winged males have higher reproductive success.[10]

1.1.3 Aphids as a Pest

Aphids are a prevalent pest species affecting many economically important crops globally, including cereals, sugar beet, oilseeds, field beans, peas and potatoes.[11] Around 25% of plants species are hosts to aphids,[12] though it is difficult to give an estimate of the economic losses due to aphid-caused damage due to the range of both direct and indirect consequences of aphid infestation.[13]

The majority of crop damage is as a direct result of phloem-sap feeding, during which essential nutrients are removed from the plant.[14] Many aphid species also instigate indirect damage, acting as vectors for the transmission of lutoviruses, such as the barley yellow dwarf virus (BYDV), and enamoviruses, such as the pea enation mosaic virus (PEMV).[14,15] Aphids may also stimulate the development of saprophytic fungi by secretion of honeydew, which in turn blocks light for photosynthetic activity.[16]

The success of aphids as pests is largely due to their ability to exploit multiple host plant species throughout the year, as well as their fast-reproductive expansions resulting from parthenogenesis. Aphid population dynamics can be linked to environmental factors including the weather and general climate – the aphid life cycle is dependent on temperature, in addition to other climate-affected factors. An increase in temperature of only a few degrees will lead to earlier flight times of aphids and a significant change in their population dynamics.[9] A major concern relating to aphids as a pest is the effect that climate change will have on aphid populations, and the increased need for more efficient control.[9]

1.1.4 Management of Aphids

Aphids can be managed by a variety of strategies, with effective management often being achieved by employing a combination of chemical and biological control. Aphid populations are mainly managed by the use of pesticides, including pyrethroids **1**, carbamates **2**, neonicotinoids **3** and organophosphates **4** (Figure 1.3).[17–20] However, a decrease in pesticide usage is encouraged due to potential off-target effects, an increase in resistance occurrence within pest insect populations and potential negative environmental effects.[17]

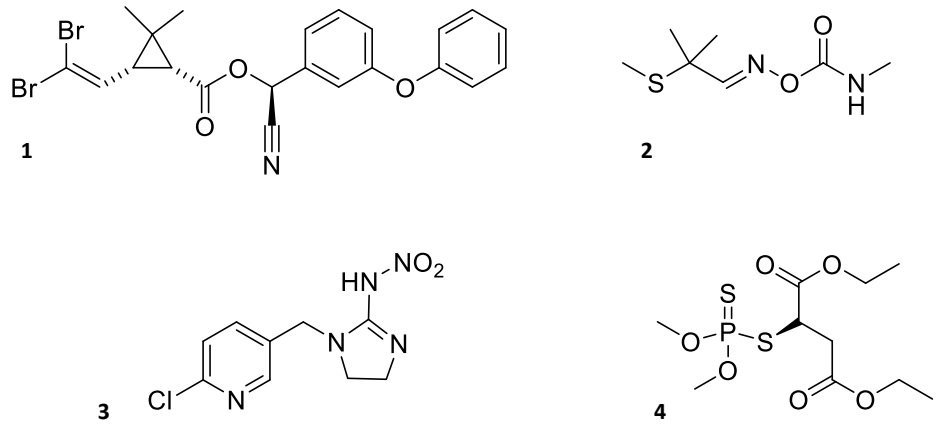


Figure 1.3: Examples of pesticides used in aphid management: pyrethroid, *deltamethrin* **1**, carbamate *aldicarb* **2**, neonicotinoid, *imidacloprid* **3**, and organophosphate, *(R)-(+)-malathion* **4**.

Insecticide resistance can arise from mutations to the target proteins.[20] Due to the high reproductive rate of aphids and significant survival advantage of the mutation, resistance spreads and evolves quickly amongst a given population.[20–22] This target-site specific resistance mechanism is generally known as 'knockdown resistance' or *kdr*[20] – for example, peach-potato aphids, *Myzus persicae*, with a kdr mutation in transmembrane segment IIS6 of a sodium channel, the target site of pyrethroid insecticides, have a 35-fold resistance to the pyrethroid deltamethrin.[20]

In addition to target site resistance, other mechanisms of resistance exist in insects. The main alternative insecticide resistance mechanism is metabolic resistance; the production of metabolic enzymes capable of breaking down insecticidal chemicals, such as the overproduction of carboxylesterases in aphids, leads to broad spectrum organophosphate resistance.[20,22] Other mechanisms include the existence of efflux pumps, which transport toxic substances out of cells, and behavioural or phenotypic changes in insects that make them less susceptible to pesticide exposure.[23,24]

One of the main alternatives or additions to synthetic chemical control is the use of biological control, or biocontrol, including the use of natural enemies (Figure 1.14).[25,26] Despite the broad spectrum of natural enemies that exist for aphids, agricultural intensification has resulted in a decrease in biodiversity and loss of suitable habitats for many of their natural enemy species.[26] By planting diverse field margins or specifically recruiting these species, aphid populations can be kept below an economically important threshold.[26–28]

Figure 1.4: A natural enemy, the aphid parasitoid of *Aphidius* species, seen laying an egg into a nymphal aphid. Image from Rothamsted Visual Communications Unit.

Biocontrol of aphids can also be achieved by the introduction of entomopathogenic fungi, such as *Verticillium lecanii.*[29] The fungus can be applied by spraying spores or planting infected plants nearby, and could be used in combination with other bio-control agents.[29]

Knowledge of the phytochemistry of particular crops can be used – the natural chemistries of plants may cause them to be repellent or attractive to aphids.[30] The careful planting of repellent and attractive plants surrounding commercial crops can achieve an effective 'push-pull' pest control management.[30] Furthermore, By exploiting the natural chemical ecology of aphids, lures (combined with traps) or repellents can be designed.[30]

1.2 INSECT CHEMICAL ECOLOGY

1.2.1 Semiochemicals

A semiochemical is a compound that is secreted by organisms, which modifies the behaviour and/or development of another organism.[31,32] Semiochemicals are categorised into intraspecific semiochemicals, *pheromones*, a compound or group of compounds that are released by an organism and induce a response in an individual of the same species, and interspecific semiochemicals, *allelochemicals*, which stimulate organisms of different species.[31,32] Pheromones are critical for communication between insects of the same species; these may include sex pheromones, aggregation pheromones, and alarm pheromones.[31]

Pheromones can be further categorised into *releasers,* pheromones which induce an immediate behavioural change, and *primers*, pheromones which initiate a complex set of physiological or developmental changes, but may result in no immediate behavioural change.[32]

In addition, allelochemicals can also be further categorised depending on whether the emitter or receiver is the beneficiary (Table 1.1).

Table 1.1: Allelochemical categorisations, based on definitions by Nordlund & Lewis, 1975.[32] A "**+**" indicates a positive effect and "**-**" indicates a neutral or negative effect

*where the emitter is non-living material

Allelochemical Categorisation	**Beneficiary**	
	Emitter	**Receiver**
Allomone	**+**	**-**
Kairomone	**-**	**+**
Synomone	**+**	**+**
Apneumone*	**-**	**+**

Semiochemicals are used in host location, mating and enemy warning systems. Though semiochemicals are widely employed by insects for communication, communication chemistry and potential semiochemicals have been identified in many other organisms including mammals, birds and fish.[33–35]

Aphid species, including the pea aphid, *Acyrthosiphon pisum,* have been shown to employ a sex pheromone and an alarm pheromone, in addition to a range of allelochemistry generally utilised for host location.[36]

1.2.2 Aphid Sex Pheromone

For sexual reproduction, which occurs between wingless sexual females (*oviparae*) and winged sexual males, aphids utilise a sex pheromone, generally consisting of two components (1*R*,4a*S*,7*S*,7a*R*)-nepetalactol **5** and (4a*S*,7*S*,7a*R*)-nepetalactone **6** (Figure 1.5).[37] Some aphids, such as the rosy apple aphid *Dysaphis plantaginea* and the damson-hop aphid *Phodrodon humuli*, also employ other isomers including (1*S*,2*R*,3*S*)-dolichodial **7**, (1*S*,4a*R*,7S*,7a*S)-nepetalactol **8** and (1*R*,4a*R*,7S*,7a*S)-nepetalactol **9**.[36,38]

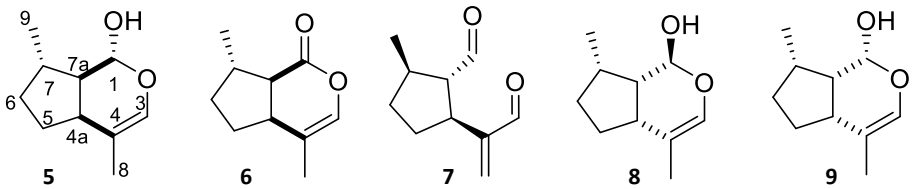


Figure 1.5: The chemical structures of (1*R*,4a*S*,7*S*,7a*R*)-nepetalactol **5**, (4a*S*,7*S*,7a*R*)-nepetalactone **6**, (1*S*,2*R*,5*S*)-dolichodial **7,** (1*S*,4a*R*,7S*,7a*S)-nepetalactol **8** and (1*R*,4a*R*,7S*,7a*S)-nepetalactol **9.**

Although the aphid sex pheromone components are ubiquitous across most species, the ratio between **5** and **6** is species-specific.[6] For example, for the pea aphid, *Acyrthosiphon pisum,* the sex pheromone consists of **5** and **6** in a 1:1 ratio, whereas the black-bean aphid, *Aphis fabae,* uses a very high ratio of **6** to **5** (Table 1.2).[6,39]

Table 1.2: Aphid species with known sex pheromone and ratio of nepetalactol to nepetalactone found in each. *Nepetalactol diastereoisomers

Common Name	Species Name	Ratio (5:6)
Pea aphid	*Acyrthosiphon pisum*	1:1[39]
Black bean aphid	*Aphis fabae*	1:29[39]
Spiraea aphid	*Aphis spiraecola*	1:2[40]
Leaf-curling plum aphid	*Brachycaudus helichrysi*	1:2.6[41]
Mealy cabbage aphid	*Brevicoryne brassicae*	0:1[42]
Rosy apple aphid	*Dysaphis plantaginea*	3.7:1 – 3.3:1 (age effect)[43]
Mealy plum aphid	*Hyalopterus pruni*	1:2.5[41]
Potato aphid	*Macrosiphum euphorbiae*	4:1-2:1 (age effect)[44]
Peach-potato aphid	*Myzus persicae*	1.5:1[39]
Damson-hop aphid	*Phorodon humuli*	1*:0[38]
Bird cherry-oat aphid	*Rhopalosiphum padi*	1:0[45]
Grain Aphid	*Sitobion avenae*	trace:1[46]
Black-berry cereal aphid	*Sitobion fragariae*	0:1[47]
Peach aphid	*Tuberocephalus momonis*	1:4[48]

Before the chemical identification of the aphid sex pheromone, aphids were thought by many entomologists to produce a close range aphrodisiac.[49] First evidence for an aphid sex pheromone was published by Pettersson in 1970[50], where it was demonstrated that male aphids of *Schizaphis* species responded to a volatile odour produced by sexual female aphids. The first characterisation of the pheromone was by Dawson *et al.* in 1987[37], using the vetch aphid *Megoura viciae* Buckto. The sex pheromone was found to comprise of **5** and **6** (Figure 1.15).[37] Initially, the pheromone components were identified by the air entrainment of the hind legs (tibiae) of oviparae, and analysis of the volatile compounds using coupled gas chromatography mass spectrometry (GC-MS).[36,37]

Identification of the pheromone was confirmed by comparison with synthetic authentic standards (Scheme 1.1) and plant-extracted material, using GC-MS and ^{1}H and ^{13}C nuclear magnetic resonance (NMR) spectroscopy[36]. The stereochemistry of the lactol was determined using the chiral derivation agent Mosher's acid **10** [(S)-α-methoxy-α-(trifluoromethyl) phenylacetic acid] (Figure 1.6) *via* NMR, and X-ray crystallography.[36]

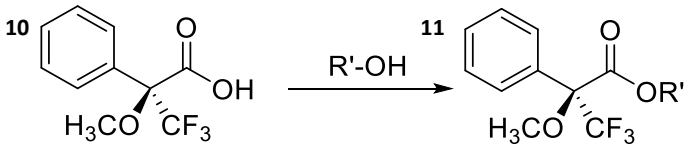
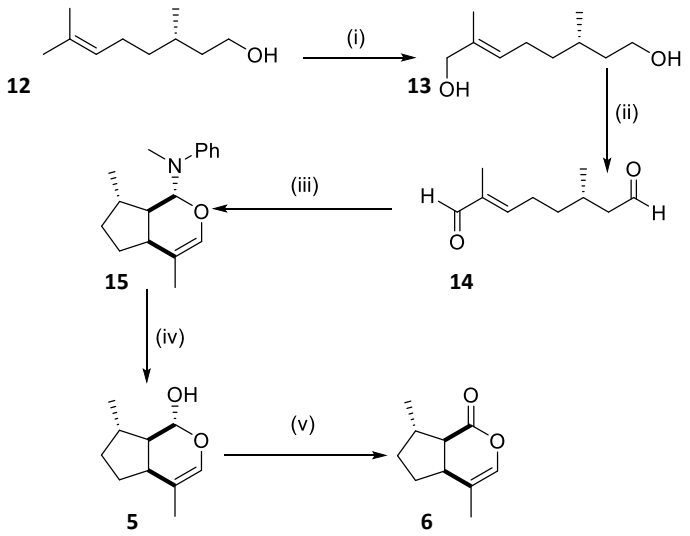


Figure 1.6: The structure of chiral shift agent, Mosher's acid, [(S)-α-methoxy-α-(trifluoromethyl) phenylacetic acid] **10** and its conversion to Mosher's ester **11** with an alcohol.

Behavioural responses, alongside electrophysiological studies, further confirmed the function of **5** and **6** as a sex pheromone.[37] Electrophysiological studies involve connecting sharp, tungsten microelectrodes to the whole antenna or to single cells. Compounds are tested for activity by detecting minute electrophysiological responses, giving either an electroantennogram (EAG) or single-cell recording (SCR).[37] This technique can also be coupled with gas chromatography to identify active compounds from complex mixtures.

Each stereoisomer of **5** and **6** can be synthesised *via* a route developed by Dawson *et al.*[51] (Scheme 1.1), a modification of previous work by Schreiber *et al.*[52]. Commercially available enantiomerically pure (*S*)-citronellol **12** is used as the starting material, which is oxidised *via* an allylic oxidation with (catalytic) selenium dioxide at the C-1 position to give a diol **13**. This diol is further oxidised to a dialdehyde **14**, then cyclised in an enamine-mediated, stereo-controlled [4+2] Diels-Alder like cycloaddition to give a bicyclic product **15**.[51,52] The phenylaniline group is hydrolysed with *p*-toluenesulfonic acid to yield the nepetalactol product **5** and a final oxidative step provides the nepetalactone product **6**.



Scheme 1.1: The synthetic route to **5** and **6** from (*S)*-citronellol **12**, based upon the route by Dawson *et al.*[37] i) SeO_2, *t*-BuOOH, DCM; (ii) 1) $(COCl)_2$, CH_2Cl_2, DMSO, -78°C 2) Et_3N, CH_2Cl_2; (iii) PhNMe, Et_2O, 4Å sieves; (iv) TsOH, THF, H_2O; (v) Ag_2CO_3, Celite, PhMe, 120°C

The lactone **6** also occurs naturally in a variety of plant species, including the catmint, *Nepeta cataria*, and can be obtained as an oil in high yields following an extraction from the plant.[6] The oil is isolated by steam distillation, where steam is applied alongside a solvent (generally cyclohexane) to plant material, the mixture then distilled, followed by separation of the organic layer and removal of the solvent.[6] The resultant oil contains **6** and other plant-produced compounds which can be further purified using flash column chromatography or HPLC.[6] Enantiomerically pure **6** can be reduced into **5** in a stereoselective reduction with sodium borohydride (Figure 1.7).

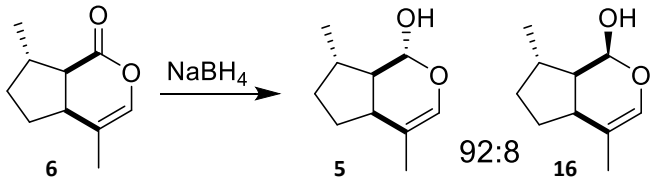


Figure 1.7: Stereoselective reduction of (4a*S*,7*S*,7a*R*)-nepetalactone **6** to (*1R,*4a*S*,7*S*,7a*R*)-nepetalactol **5** and (1*S,*4a*S*,7*S*,7*aR*)-nepetalactol **16** respectively. Adapted from [6].

1.2.3 Aphid Alarm Pheromone

Upon attack by a natural enemy, such as a predator or parasitoid wasp, aphids release an alarm pheromone, which causes nearby aphids to cease feeding and disperse.[36,53,54] In most species, this pheromone is the sesquiterpene (*E*)-β-farnesene **17** (EBF) (Figure 1.8).[36,53] Some aphid species have also been shown to release small amounts of other compounds, such as (1*E*,5*E*,8*S*)-germacrene A **18** and (1*S*,5*S*)-α-pinene **19**.[53,55,56]

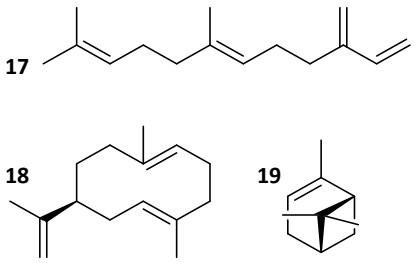


Figure 1.8: The sesquiterpene aphid alarm pheromone, (*E*)-β-farnesene **17** and other aphid alarm pheromone components, (1*E*,5*E*,8*S*)-germacrene A **18** and (1*S*,5*S*)-α-pinene **19** [36,53,55,56]

EBF **17** is a terpenoid comprised of three isoprene (C5) units (Figure 1.9).[57] Biosynthesis of terpenoid compounds is believed to occur exclusively by the mevalonate pathway in animals, though the mechanism for biosynthesis of EBF **17** in aphids has not yet been fully elucidated.[57]

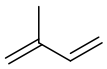


Figure 1.9: An isoprenoid unit **20**.

Chemical synthesis of EBF **17** can be achieved by a one-step reaction of (*E,E*)-farnesyl chloride with 18-crown-6 ether and potassium *tert*-butoxide (Figure 1.10).[58]

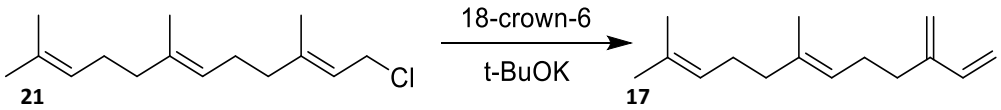


Figure 1.10: Synthesis of (*E*)-β-farnesene **17** from (*E,E*)-farnesyl chloride **21.**

The alarm pheromone has a repellent effect on aphid populations, making it a suitable candidate for aphid management applications. EBF **17** can also be used to recruit natural enemies of the aphid.

1.2.4 Aphid Allelochemicals

Allelochemicals are compounds that when emitted by an organism of one species affect the behaviour and development of an organism of another species.[32] When seeking a suitable host, or avoiding an unsuitable one, aphids may exploit allelochemicals or blends of allelochemicals emitted by plants.[36,59] Plant-aphid interactions are affected by attractive and repellent compounds, and the presence of these can prove critical in aphid host preference or natural resistance of the plant to insects.[60]

Natural attractiveness or repellence to aphids can be observed in many plant species. For example, Brassicaceae (cabbage) plants produce a class of chemicals known as glucosinolates which can generate toxic isothiocyanates when the plant is attacked.[36] A range of these chemicals, including allyl **22**, 3-butenyl **23** and 4-pentenyl isothiocyanate **24** (Figure 1.11), are used by some aphid species (cabbage aphid, *Brevicoryne brassicae*; turnip aphid, *Lipaphis erysimi*) to locate a host.[36] However, other species, such as *A. fabae*, are repelled by isothiocyanates, and their presence will render a plant unsuitable as a host.[36]

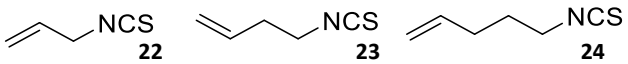


Figure 1.11: Semiochemicals emitted by Brassicaceae plants when attacked by feeding herbivores; allyl **22**, 3-butenyl **23** and 4-pentenyl isothiocyanate **24**.[36]

Plants may also have induced responses following attack by pests, often resulting in the release of semiochemicals.[61] Induced responses are complex – different pest types, insect feeding types and species can all play a role in altered plant defence responses.[61] Furthermore, below-ground attacks can induce above ground signalling, and *vice versa.*[61]

Semiochemicals released by plants upon attack may be repellent to aphids. These include (*S*)-germacrene D **25**, (1*R*,4*E*,9*S*)-caryophyllene **26** and (*E,E*)-4,8,12-trimethyl-1,3,7,11-tridecatetraene **27**, identified from the host bean plant *V. faba.*[36,59] Other compounds, such as (*E*)-ocimene **28** and 6-methyl-5-hepten-2-one **29** (Figure 1.12), may be attractive to aphid natural enemies (parasitic wasps and lacewings).[36,62] Finally, some plant-produced semiochemicals can also subsequently induce a defense response in an intact plant, such as in the case of (*Z*)-jasmone **30** (Figure 1.12).[36]

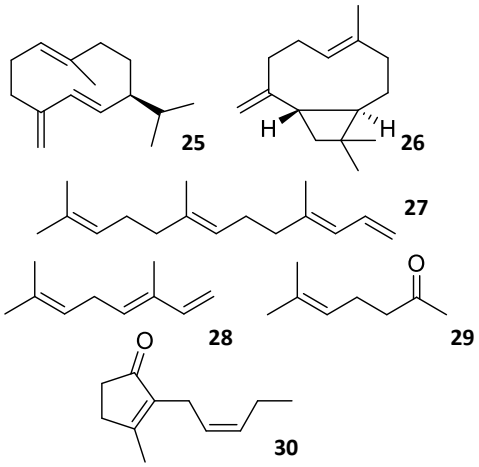


Figure 1.12: The structure of various plant semiochemicals involved in repelling aphids. (*S*)-germacrene D **25**, (1*R*,4*E*,9*S*)-caryophyllene **26**, (*E,E*)-4,8,12-trimethyl-1,3,7,11-tridecatetraene **27** (*E*)-ocimene **28** and 6-methyl-5-hepten-2-one **29** and (*Z*)-jasmone **30**.[36,62]

1.2.5 Practical Deployment of Semiochemicals

Semiochemicals have the potential to be used in pest management, by using mating disruption, pheromone traps, push-pull strategies and recruitment of natural enemies. Synthetic sex pheromone components have been used to catch male aphids and recruit foraging parasitic wasps, *Aphidius ervi* Haliday and *Praon barbatum* Mackauer, in the field.[27,28,63] The synthetic sex pheromone components have also been found to attract other aphid natural predators, such as lacewings (*Chrysopa cognata).*[64,65]

The alarm pheromone **17** has been shown to be repellent to aphids in behavioural studies, whilst attractive to natural enemy predators and parasitoids. A hexaploid commercial variety of wheat, *Triticum aestivum* cv. Cadenza (Poaceae), was genetically engineered to biosynthesise and release EBF **17**.[66] Evaluation in field trials showed the transformed wheat variety was not significantly different from non-transformed varieties in managing aphids (assessed by aphid numbers and number of parasitized aphids), although controlled environment studies had demonstrated its effectiveness.[66] The lack of field effectiveness is

likely due to the release rate of the alarm pheromone from the plant being consistent and steady, in contrast to a natural quick burst of pheromone produced by aphids.[66]

Plant derived semiochemicals also have practical applications; (*Z*)-jasmone **30** has been found to be effective in reducing aphid numbers, attracting parasitic wasps and inducing the production of repellent volatiles in crops such as wheat, cotton and sweet peppers.[36] (*S*)-Germacrene D **25** has been identified as a potent repellent for aphids, however, has little potential for commercial application in crop protection due to its chemical instability and cost of production.[36,67] More stable analogues, which have comparable behavioural activity, could be developed, particularly using modified terpene synthases and unnatural substrates.[36,68]

Further advancement of the use of semiochemicals in pest management could stem from a deeper understanding of molecular recognition processes in the olfactory systems of pests. This may lead to '*ab initio*' design of ligands, which may have similar behavioural effects as other olfactory ligands, but better prospects for commercial producibility.[36] Currently, understanding of the olfactory system is limited, though two major groups of proteins are involved – olfactory receptors (ORs) and odorant binding proteins (OBPs) – both of which could provide potential pest management targets.[36,69]

1.3 INSECT OLFACTION

1.3.1 History of Insect Olfaction

Prior to the identification of the first pheromones, little was understood regarding insect olfaction with the majority based on anecdotal evidence.[70] The first insect pheromone to be characterised was (10*E*,12*Z*)-hexadec-10,12-dien-1-ol, or bombykol **31**, identified from the silkworm moth, *Bombyx mori* (Figure 1.13).[70–73] Due to the large physical size of this insect, a significant quantity of pheromone could be extracted – in total, 5 x 10^5 glands needed to be extracted to obtain sufficient quantities for elucidation of the chemical structure.[70] Initially, structural identification was achieved using infrared (IR) and ultraviolet (UV) spectroscopy, but was later confirmed with NMR and GC.[70–72]

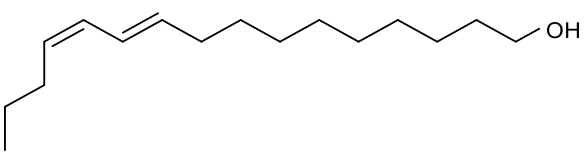
31 

Figure 1.13: The chemical structure of (10*E*,12*Z*)-hexadeca-10,12-dien-1-ol, or bombykol **31**, the first chemically identified insect pheromone, and sex pheromone of the silkworm moth, *Bombyx mori*[73].

The most significant advances in the field of olfaction research were founded in developments in analytical chemistry.[70] Introduction of techniques for identification of complex chemical structures, particularly NMR spectroscopy, and detection of low-levels of compounds, such as gas-chromatography, have vastly expanded the known chemical library of pheromones and other insect-related semiochemicals.[70] The identification of the sex pheromone previously discussed of the pea aphid, *A. pisum*, relied on such techniques.[6,36,51]

In recent years, significant developments have been made in the study of odorant perception, which has been furthered by work in genomics. With full genome sequences now available for many insect species, including model organisms *A. pisum*[2] and *Drosophilia melanogaster*[74], many genes for odorant specific receptors and binding proteins have been identified. Furthermore, by using advanced molecular biology techniques, the function of these proteins and receptors can be more extensively studied.[75–77]

1.3.2 Antennal Anatomy

The main olfactory organs in insects are the antennae, although some insects possess additional olfactory organs, such as maxillary palps.[78,79] The aphid antennae comprise of six segments, numbered from the base [including a *scape* (1st), *pedecil* (2nd) and four *flagella* (3rd-6th)].[75] Small, hair-like structures known as *sensilla* (Figure 1.14) cover the antennae.[75,76,79]

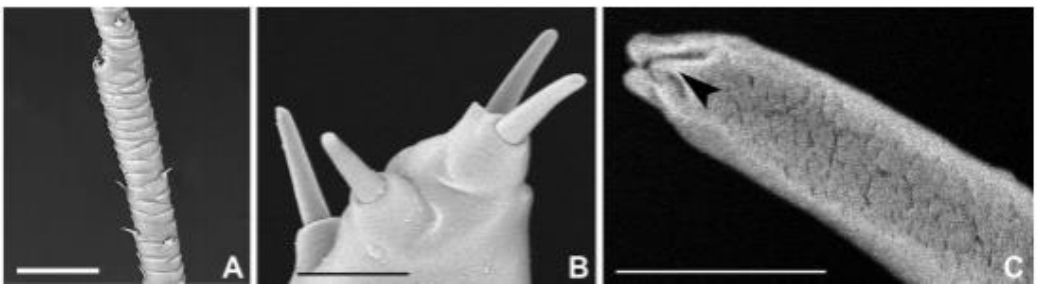


Figure 1.14: Type II trichoid olfactory sensilla from *A. pisum*. Images taken directly from De Biaso *et al.*, 2014.[75]

Olfactory sensilla are perforated with pores, through which odorant molecules diffuse[76]; each sensilla can be uniporous (one pore) or multiparous (many pores).[75] Sensilla contain an aqueous fluid, or sensillium lymph, in which olfactory receptor neurons are bathed, and where odorant-binding proteins are located.[76] Specific odorant receptors have been mapped to specific sensilla types.[76,79]

Diversity in function has been observed across sensilla. There are a few morphological classes of sensilla found in most insects; *basionic,* both large and small, *trichoid, placoid* and *coeloconic*.[76,78,79] Each morphological class is responsible for a different function; in *Drosophilia,* the basionic sensilla is responsible for fruit odours, trichoid for pheromones, and coeloconic for organic acid and amine based odours.[76]

In the pea aphid, *A. pisum*, trichoid sensilla of two types are found on the antenna. Type II sensilla, found mainly on the tip of the antenna, are short and blunt, possessing only a single 'apical pore'[75]. Type I sensilla are found across the entire antenna, and do not have pores on their tip.[75] In all life stages of the aphid, on the 5th and 6th antennal segments, exists the primary *rhinaria* ('sunken pits'). Placoid and coeloconic sensillium are also present. Electroantennography from these areas show detection of leaf volatiles and the alarm pheromone, (*E*)-β-farnesene **17**.[75] Secondary rhinaria are found in a higher abundance in male aphids and are sensitive to sex pheromone components and plant volatiles.[75] Sensilla have been found on other parts of *A. pisum*, including the hind legs and mouthparts, but their function in these areas is not yet fully described.[75]

1.3.3 Aphid Olfactory Neuroanatomy

Olfactory neuroanatomy is a complex and well-studied field. Insect olfaction begins in the sensilla of the antenna; antennal neuroanatomy mainly consists of olfactory receptor neurons (ORNs) expressing ORs.[77–80] In *Mammalia*, ORNs are bipolar, allowing dendrites to give rise to numerous specialised cilia and providing a large receptive surface for the binding of odours to ORs.[80] Removal of these specialised cilia in mammals removes associated olfactory responses.[80] ORs are activated generating an action potential, which travels along an odorant receptor neuron (ORN) that glomerulates and converges in the brain.[79–82]

The aphid brain is simple, with a well-defined central body consisting of a protocerebral bridge and lateral accessory lobes.[83] The protocerebrum includes optical lobes, the central complex and often ill-defined mushroom bodies, and the deutocerebrum contains both the antennal and dorsal lobes.[83,84]

Figure 1.15: The general structure of a hemipteran brain. The main brain sections that are critical in olfactory processing are defined: the central body (purple), mushroom bodies (green) and antennal lobes (blue). Diagram adapted from [83]

Hemipteran central nervous systems have highly fused ganglia, with antennal nerves entering the brain ventrally and branching into glomeruli in the antennal lobe.[83,84] The antennal lobes of aphids are small, spherical structures between 35 µm and 45 µm in diameter.[83] Around 25-40 glomeruli can be visualised in aphid antennal lobes, though they are often poorly delineated.[81,83] From the antennal lobe, olfactory processing occurs either in the mushroom body, thought to be linked to learned behaviours, or the lateral horn, for innate behaviours.[85,86] Mushroom bodies are an insect-specific brain region critical for olfactory output processing, and are usually well developed in species with complex social behaviour or strong capacity for olfactory learning and memory.[83,85] Low development of the mushroom bodies can be seen in aphid species, which is consistent with their simple behaviour and social interactions.[83,85] Little is understood about the specific roles of the lateral horn and mushroom body in insect olfaction, including neuronal organisation and how signals are coded and communicated between the two brain regions.[85,86]

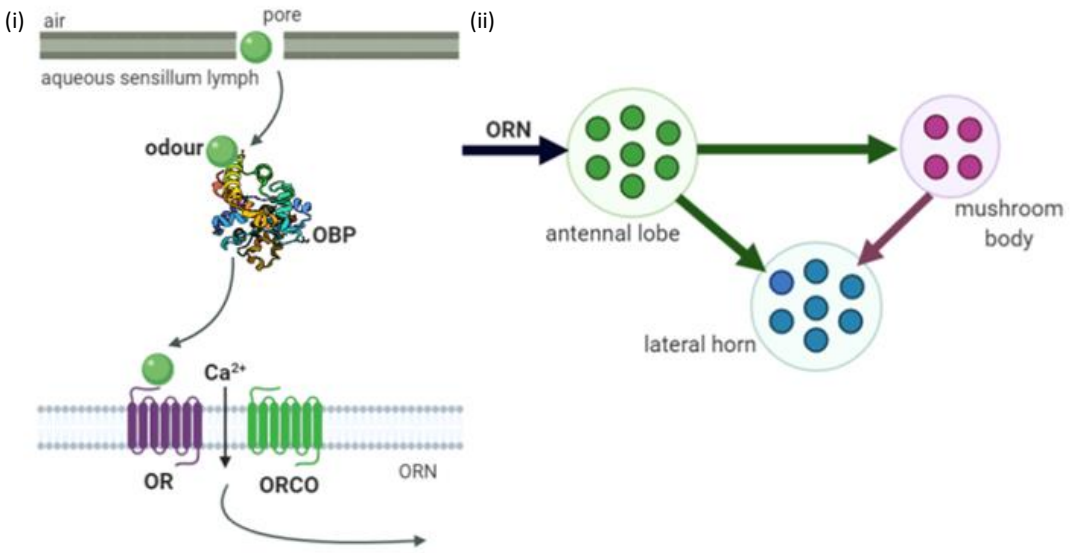


Figure 1.16: General olfactory processing in insects. (i) Within the antennal sensilla, odorant-binding proteins (OBPs) play a role in allowing odorants to activate odorant receptors (ORs) which are co-expressed with the olfactory receptor co-receptor (ORCO); (ii) Once activated, an action potential travels along the odorant receptor neuron (ORN) to the antennal lobe. Further processing occurs in either the lateral horn or the mushroom body.[79–81] Graphic made using BioRender.

1.3.4 Odorant Receptors

Odorant receptors (ORs) are a highly conserved group of G-coupled protein receptors (GPCRs) found on the surface of odorant receptor neurons, and are involved in the perception of odours.[80] GPCRs are defined as a large, diverse family of seven transmembrane (7-TM) or heptahelical proteins with an extracellular N-terminus and intracellular C-

terminus.[80,87] Generally, GPCRs are involved in signalling pathways *via* a secondary messenger system mediated by cyclic adenosine monophosphate (cAMP).[80,87]

Though GPCRs and subsequently ORs are a diverse family of proteins found across a variety of species, insect ORs are distinctive from other GPCRs and may not be categorised as true GPCRs.[82] Insect ORs possess an inverse heptahelical topology – the N-terminus is unusually located in the intracellular section of the transmembrane protein, with the C-terminus found extracellularly.[82,88] This suggests that insect ORs are a unique protein family, different from all other chemosensory receptors.[82] Insect ORs also differ in signal transduction and they have been proposed to form a unique class of heteromeric cation channels.[88–90]

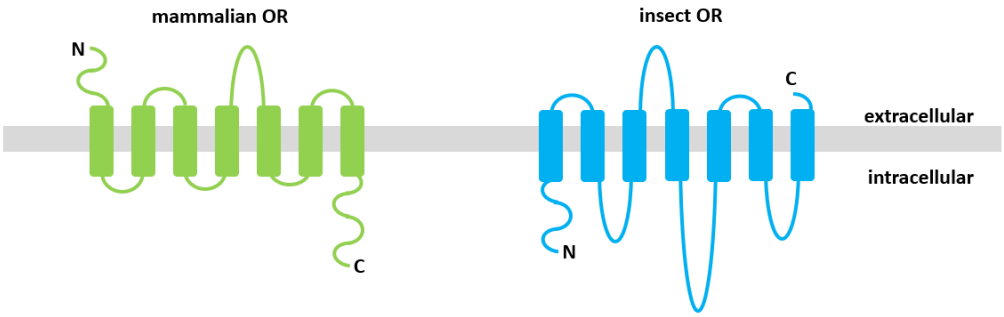


Figure 1.17: Mammalian odorant receptors (OR) in contrast to insect ORs. Adapted from [82].

ORs are a much larger and more diverse group of receptors in mammals, however, most ORs found in insects are co-expressed with another OR, known as *ORCO* (odorant receptor co-receptor, identified as OR83b in *D. melanogaster*).[82,88] Orco is structurally similar to other insect ORs, however, it is highly conserved across different insect species – only 20% conservation is seen between ORs, however approximately 60% shared identity can be seen for ORCO.[88] The structure of ORCO was solved in 2018 by Butterwick *et al.* using cryo-EM techniques, leading to large advances in knowledge of the assembly and functionality of ORCO-OR tetramers, including their properties as an ion-channel for olfactory signal transduction.[88] Structurally, ORCO can be divided in to two domains including four loose peripheral transmembrane domains and a single central anchor domain.[88] There is no known mammalian ortholog of ORCO.[82]

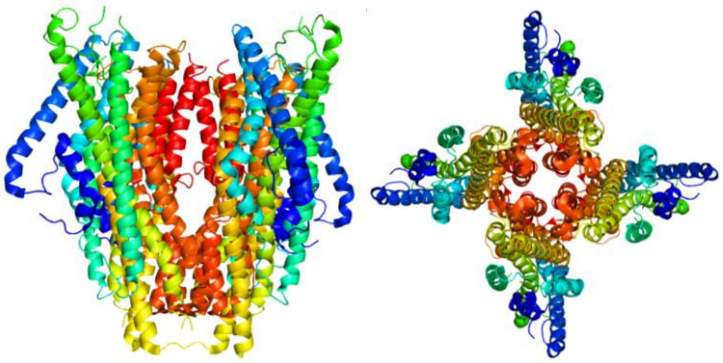


Figure 1.18: The cryo-EM solved structure of the olfactory receptor co-receptor (ORCO) from the parasitic fig wasp, *Apocrypta bakeri*.[88] The Protein Data Bank (PDB) unique identifier is 6C70. Image made in PyMol.[91]

ORCO-deficient mutants show significant loss of olfactory function and related behaviours, indicating that orco must play an important role in olfactory signal transduction.[92–94] Antennal lobe glomeruli were also reduced in some ORCO-deficient mutants.[94]

ORs generally bind multiple active odorants with varying specificity.[95] This contrasts with the seemingly incredibly specific behavioural activity of specific odours. In mammals, odour perception has shown to be combinatorially coded – multiple ORs are activated in unique combinations as a response to a specific odour.[95] Little is known about odour coding in insects, however, evidence for distinctive neuronal perception between attractive and repellent odours has been identified.[86]

OR binding activity is generally investigated using electrophysiological techniques, where the receptor is expressed in a membrane, exposed to a ligand and electrophysiological responses measured. Other techniques that can be used include calcium imaging.[78] Unfortunately, ORs are difficult to study, as membrane bound proteins are notoriously hard to express and a lack of available solved 3-dimensional structures reduces options for homology modelling and *in silico* studies.[96]

Genes encoding chemosensory receptors have been identified in aphid species, including 79 OR-candidate genes identified in the pea aphid *A. pisum.*[2,97,98] There have been limited studies into the role of these ORs in aphids. *A. pisum* OR1 has been identified as an odorant receptor co-receptor (ORCO), similar to OR83b in *D. melanogaster.*[98] Other *A. pisum* odorant receptors have been identified; reduction in response to the alarm pheromone, EBF **17**, is lost when *A. pisum* OR5, co-expressed with Orco, gene expression is knocked down by interference RNA (RNAi).[99] Furthermore, *A. pisum* OR4 has been shown, through electrophysiological assays, to be broadly tuned to a series of volatiles.[100]

In addition to ORs, insects possess a unique group of olfactory proteins. Odorant-binding proteins (OBPs) are a group of proteins found in high abundance in the sensillum lymph, the aqueous fluid that can be found within the sensilla of the antennae.[101] These proteins are highly structurally conserved across insects and are thought to play a role in olfaction.

1.4 ODORANT-BINDING PROTEINS

1.4.1 General Structure and Classification

OBPs are a diverse group of small water-soluble proteins, abundant in the aqueous fluid surrounding ORNs in the sensillum (sensillum lymph) of insect antenna.[76,101–103] The mRNA encoding for OBPs are the most abundant mRNA found in the antenna, indicating that they are expected to play an important role in insect olfaction.[76]

The first OBP to be identified in invertebrates was from an extract of the large moth, *Antheraea Polyphemus* Cramer, using a radioactively-labelled pheromone in ligand-binding experiments; OBPs were discovered in vertebrates at a similar time.[101] Initially, OBPs could be categorised by their six conserved cysteines, which results in a similar 3D structure despite diverging amino acid sequences.[101] However, further research showed a greater diversity of 3D structure, and three distinct categories have now been defined; classic OBPs (possessing 6 highly conserved cysteines that form disulphide bridges), Plus-C OBPs (possessing 8 conserved cysteines and one conserved proline) and Atypical (possessing 9 or 10 conserved cysteines).[103] The conserved cysteines and multiple disulphide bridges leads to high thermal stability in these proteins. OBPs range in size from approximately 110 to 240 residues, usually resulting in proteins of 10-25 kDa in size.[102]

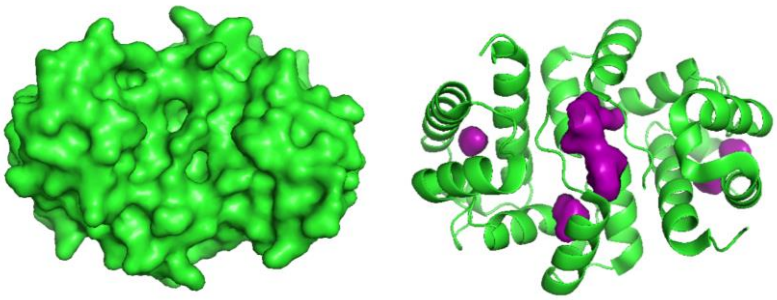


Figure 1.19: The structure of OBP3 from the vetch aphid, *Megoura viciae,* solved by X-ray crystallography. Shown as the surface of the protein (L) and with only the binding pocket displayed (R, binding pocket displayed in purple). Image generated using PyMol.[91] The Protein Data Bank (PDB) unique identifier is 4Z39.[104]

Though thermally stable, OBPs are flexible globular proteins that may occur in multiple conformations. Alternate conformations may arise as a result of conditions such as pH change or binding activity. Pheromone-binding proteins (PBPs) found in *Bombyx mori* and other insect species, show pH-dependent conformational changes.[71,105–110] This change is often associated with the binding and release of a ligand.[105] Furthermore, OBPs and PBPs have been shown to have adaptable binding pockets, that can sometimes swell to accommodate multiple ligand molecules.[111]

In addition to OBPs, another family of proteins, known as chemosensory proteins or CSPs, have been described. CSPs shown similar binding activity to OBPs, but no sequence similarities, and only share 4 conserved cysteines.[111]

1.4.2 Suggested Olfactory Role

Olfactory ligands need to be recognised amongst a plethora of background signals; olfaction is an external system, in contrast to hormones or other internal chemical communication. This implies that olfactory proteins, potentially including OBPs, must be more sophisticated in their discrimination ability in comparison to other recognition proteins.[68] Many olfactory ligands differ significantly in behavioural activity, but only by minute chemical structural difference.[68] Understanding the role of OBPs in olfactory perception is critical to understand how insects can discriminate between olfactory ligands with extreme precision.[101,112]

Numerous roles have been proposed for OBPs *via* studies on a range of insect species.[101,112,113] Many insect odorant molecules are highly lipophilic and are poorly soluble in aqueous solutions, which indicates these proteins may play a role in solubilising or transporting the ligands to the odorant receptors in the aqueous sensillium lymph.[101,112]

OBPs have been shown to reversibly bind behaviourally active olfactory ligands, suggesting they play a role in olfactory perception.[69,102,114–119] One of the most widely studied OBPs is found in *Drosophilia melanogaster,* and known as LUSH.[120,121] LUSH is thought to play a role in recognition and response to the male sex pheromone *(Z)*-vaccenyl acetate (VA), and has been show to bind to VA *in vitro*, as well as other insect pheromones, short chain alcohols and phthalates.[76] *Bombyx mori* OBPs and PBPs are capable of discriminating and binding *B. mori* sex pheromone components, such as bombykol[115,122,123], and OBPs in aphids have been shown to discriminate the alarm pheromone.[102,104,124,125]

Developments in molecular techniques, specifically the use of knockdown *via* RNA interference (RNAi) and CRISPR/Cas9 knockout techniques, have led to advances in understanding of the function of OBPs.[126] Expression of the gene encoding for a specific OBP

can be suppressed, resulting in lower levels of gene expression, or the DNA encoding for the gene can be knocked out, removing expression of the gene entirely.[126] Deletion of LUSH in *D. melanogaster*, PBP3 in *Spodoptera litura,* the tobacco cutworm, and PBP1 and 3 in *Chilo suppressalis,* the striped rice stem borer, results in significant reductions in electrophysiological responses – measured by observing firing impulses in odorant receptor neurons or ORNs – to their respective binding ligands.[76,127–129] Furthermore, DmOR69d neurons, thought to be associated with VA perception, were found to be spontaneously firing in the absence of LUSH.[77] OBP-deficient mutants may also show behavioural defects; by knocking down either OBP3 or OBP7 in *A. pisum,* the repellent behaviour of (*E*)-β-farnesene is lost.[99] In *Drosophilia sechillia,* knocking out genes for OBP57d and OBP75e altered behavioural response to hexanoic acid and octanoic acid, volatile organic compounds from ripe fruit of their host plant *Morinda citrifolia.*[130] However, other mutants may have different responses – some literature shows that olfactory neuron responses are still functional after deletion of LUSH, and LUSH-deficient mutants do not show any behavioural defects.[76,77] The existence of such contradicting studies only furthers the need for an understanding of the role of OBPs.

Insects are also often capable of behaviourally and physiologically discriminating between structurally similar compounds, including those with only small modifications or stereochemical differences, such as the sex pheromone components of the aphids.[37] OBPs may assist in this process, though some evidence suggests they entirely dictate this recognition, binding as an OBP-ligand complex to the OR. Supporting evidence for this hypothesis would include the occurrence of a conformational change, induced in OBPs when biologically active ligands bind. *D. melanogaster* OBP LUSH possesses a salt bridge between Lys87 and Asp118 that is only present in the *apo* (unbound) structure (non-VA-bound 3D-structure). When this salt bridge was disrupted in LUSH mutants, the *DmOR69d* neurons were activated in the absence of VA.[77] This suggests that LUSH is conformationally activated by VA, and in turn activates ORNs, or that a VA/LUSH complex interacts with ORNs.[77] OBPs and PBPs from different insect species also possesses comparable salt bridges – *B. mori* possesses a salt bridge between Lys89 and Glu125 that is structurally analogous to LUSH.[115]

Current evidence for a conformation change or conformation activation of OBPs generally indicates a C-terminal folding domain.[105,115,117,131] This C-terminus folding may be dictated by pH changes, where the acidic-residue rich C-terminus loses negative change at a low pH and forms an additional α-helix.[131] This additional helix can then enter the binding pocket and displace any ligand present.[131] This conformation activation has been observed in a range of

species and could be responsible for interactions at the OR, where the OBP may expel the ligand for OR-binding, or the OBP may bind itself in a protein-protein interaction.[107–110] Some OBPs have also been shown to dimerise, and it is possible this dimerization could be disrupted by conformational changes, suggesting alternative ligand binding and release mechanisms.[132]

1.4.3 Alternative Roles of OBPs and CSPs

Genomic and proteomic studies generally show that many OBPs or CSPs are not found within the primary sensory organs (antennae) and are expressed in a wide diversity of spatial patterns, suggesting they possess roles beyond olfaction. For the honeybee, *Apis mellifera*, only 12 of 21 identified OBPs and two of six CSPs have been identified within the antennae.[76,113,133] Though some OBPs are expressed in the gustatory (taste) system of some species, as well as in larval chemosensory organs (*Drosophilia*), many OBPs and CSPs may possess entirely non-chemosensory roles.[76,113,134]

OBPs and PBPs have been found in the pheromonal or ejaculatory glands of insects.[113,135,136] For example, in *D. melanogaster* a CSP exists in the ejaculatory bulb, the organ that produces the male pheromone vaccenyl acetate.[136] Transcriptome analysis has identified similar CSPs in the pheromonal glands of many species.[135] Furthermore, putative pheromones may be transferred between sexes during copulation – a CSP identified in male reproductive organs in the *Locusta migratora s*pecies was found in the sexual organs of females after mating.[137]

CSPs have also been linked to development; CSPs have been shown to be upregulated when limb regeneration occurs, during embryo development and in cuticle development.[113] CSPs are sometimes specifically expressed in sexual organs (ovaries and eggs), and are thought to play a role in embryo development – in the honeybee, CSP5 was found in embryos and queen ovaries.[138]

In addition to pheromone transport and development, OBPs and CSPs play a role in a variety of other biological processes. These include anti-inflammatory action in disease carrying insects, humidity sensing in *Drosophila* and other roles in nutrition, vision, migration and insecticide resistance.[113,134,139]

1.4.4 Aphid OBPs

Advances in genomics have allowed research into specific aphid OBPs, particularly the pea aphid, *A. pisum*. In 2010, the genome for *A. pisum* was fully sequenced; the first hemimetabolous insect genome sequence to be published.[2,103] Using bioinformatic

techniques, 15 genes encoding for putative OBPs have been identified in the pea aphid, specifically 13 Classic OBPs and 2 Plus-Cs (OBP5 and 6).[75,103] Though OBP sequences are divergent, there is a high similarity between orthologs within a range of aphid species; 10 of the OBPs are highly conserved across aphid species.[2,103] The pea aphid genome also contains 13 genes encoding for putative chemosensory proteins (CSPs).[2]

A. pisum OBP3 and 7 (Figure 1.20) have been shown to play a role in the detection of (*E*)-β-farnesene **17**, the alarm pheromone in both fluorescence binding studies[102,125,140] and RNAi mediated knockout behavioural studies[99] (both conducted in *A. pisum*). The two OBPs are presumed to be linked to the repellency of the alarm pheromone.[102] Presently, no aphid OBP has been shown to bind either of the aphid sex pheromone components.

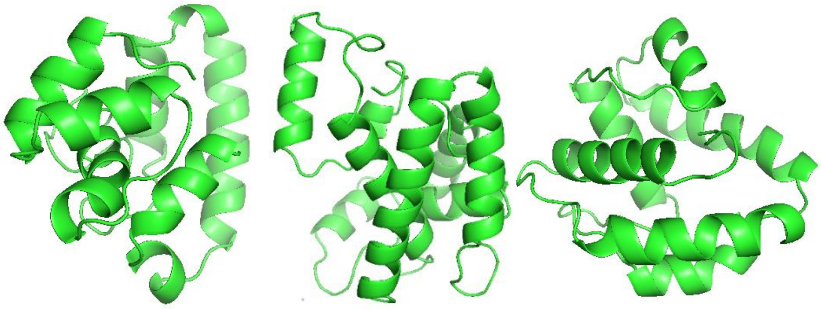


Figure 1.20: *Acyrthosiphon pisum* odorant-binding protein computer-generated models. L-R OBP3, OBP6 and OBP7. Generated using i-TASSER online server[141] and PyMOL.[91]

In *A. pisum* OBP6 and 7 have been found to have the highest transcript levels in the antenna, and OBP1 and 3 in the terminal region of the abdomen.[75] OBP6 also has higher transcript levels in winged morphs than unwinged, and is found in type II trichoid sensilla, suggesting a role in sexual reproduction.[75,103]

Generally, very little is known about the function of aphid OBPs, except for in the discrimination of the alarm pheromone (Table 1.3). No binding studies have been conducted with the sex pheromone components, and there is limited available structural data. An understanding of sex pheromone perception and the role of OBPs in aphids could be extended to other insects and support hypotheses surrounding insect olfaction. As aphids are a prevalent pest in agriculture, knowledge of their olfactory system could provide new pest management and monitoring strategies.

Table 1.3: Summary of literature reporting current knowledge pf aphid odorant-binding protein function. OBP1,3,6,7 and 8 are the only OBPs with transcriptome analysis or tested activity.

Aphid OBP	Highest Transcript Levels[75,103]	OBP Classification[103]	Reported Activity[99,102,104,125]
OBP1	Terminal region of body; 4th instar/unwinged adult stage	Classic	No known
OBP3	Terminal region of body; 4th instar/unwinged adult stage	Classic	Binds alarm pheromone, (E)-β-farnesene **17**
OBP6	Antennae; winged adults; type II trichoid sensilla	Plus-C	No known
OBP7	Antennae; 4th instar/unwinged adult stage	Classic	Binds alarm pheromone, (E)-β-farnesene **17**
OBP8	Head; all instars	Classic	No known

1.4.5 OBP Applications in Pest Management

Resistance to pesticides in insects is becoming significantly more prevalent and persistent within populations, leading to the need to develop smarter pest management solutions.[20] Ligand-specific proteins may be used in the discovery and confirmation of novel olfactory ligands for pest management.[30,68] By screening synthetic ligands with olfactory proteins, their binding affinity can be calculated, and potential biological activity deduced. Synthetic ligands could provide future solutions for pest management, as they may be cheaper to make or more field-stable than naturally occurring olfactory ligands.[68]

Additionally, advances in RNAi and gene-editing could lead to management solutions. It has been shown that RNAi can be used to disrupt to expression of OBPs or ORs in aphids and other insect species.[142] Disruption of the production of these proteins could be used for management of insects – disruption of olfactory protein expression would lead to consequential disruption of communication. This may interfere with an insect's ability to avoid predation or to sexually reproduce.

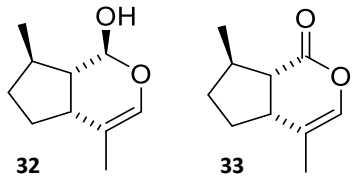
In order to elicit improved management of pest species, more precise and specific management and application is required. However, for specific control, accurate and detailed monitoring is also required. Many advances have been made in the field of sensors, specifically in agriculture, in recent years, and by furthering our understanding of the volatile profile of plants infested with pests/pathogens, and of the pests/pathogens themselves, sensing technology can be aimed toward the detection of these compounds.[143] This may include the generation of general sensor, which detects an increase in volatiles. The most desirable sensing technologies would be those that provide an indication to specific

compounds related to specific pests/pathogens. With regards to insect pests, a wide library of compounds that related to particular species and their interactions have been identified. For example, detection of the aphid semiochemicals would be useful as a monitoring tool for aphid numbers in a field.[102]

Sensing technology is now incorporating the use of both natural and synthetic peptides, alongside larger proteins, to assist with detection of specific compounds. If OBPs show specificity for compounds used in insect communication, these proteins could be used to develop provide insect-specific detection systems.[144]

1.5 AIMS AND OBJECTIVES

This thesis aims to investigate the molecular mechanisms of insect olfaction by evaluating the structure and function of odorant-binding proteins (OBPs) found in the model organism the pea aphid, *A. pisum*. Specifically, the project will focus on the perception of the sex pheromone components, (1*R*,4a*S*,7*S*,7a*R*)-nepetalactol **5** and (4a*S*,7*S*,7a*R*)-nepetalactone **6**, and their respective non-naturally occurring enantiomers, (1*S*,4a*R*,7*R*,)-nepetalactol **32** and (4a*R*,7*R*,7a*S*)-nepetalactone **33** (Figure 1.21).



32 **33**

Figure 1.21: The non-naturally occurring enantiomers, (1*S*,4a*R*,7*R*,7a*S*)-nepetalactol **32** and (4a*R*,7*R*,7a*S*)-nepetalactone **33**, of the aphid sex pheromone components.

OBPs are thought to play a role in olfaction, though exactly how they are involved in the perception of odorants is unknown. One main hypothesis surrounding the role of OBPs in insect olfaction will be investigated: OBPs are involved in olfactory perception and play a discriminatory role, potentially by directly activating odorant receptors (ORs) with an OBP-semiochemical complex.

This hypothesis will be tested using a range of multi-disciplinary techniques including well-established binding studies, such as fluorescence and *in silico* modelling, in addition to alternative chemical biology techniques, such as mass spectrometry and nuclear magnetic resonance (NMR) spectroscopy (Figure 1.22).

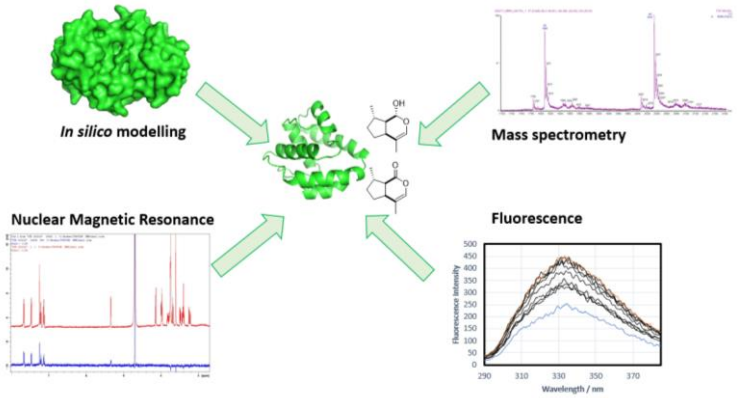


Figure 1.22: A multidisciplinary approach will be taken to study aphid OBP-sex pheromone interactions and determine the role of OBPs in insect olfaction.

In addition to these hypotheses, evidence gathered will assist in the development of a complete model of insect olfaction. This may include the observation and identification of conformational changes related to ligand binding or changes in conditions, such as pH.

Studies are being conducted in the pea aphid, *A. pisum*, due to the widely available genomic and behavioural data. The pea aphid is also a problematic pest worldwide, and the work of this thesis may be used to design smarter pest management and monitoring strategies. Furthermore, aphids are known, from electrophysiological and behavioural data, to be capable of discriminating between the two enantiomers of the sex pheromone **5/6** and **32/33**. Observing enantiomerically specific OBP-sex pheromone binding may be indicative of a discrimination function of OBPs.

In this thesis, both sex pheromone components **5** and **6** (and respective non-naturally occurring enantiomers **32** and **33**), have been synthesised and purified, along with *A. pisum* OBPs. *In* silico modelling was utilised to screen potential protein-ligand interactions, followed by *in vitro* – fluorescence binding assays, NMR and mass spectrometry based-assays for confirmation.

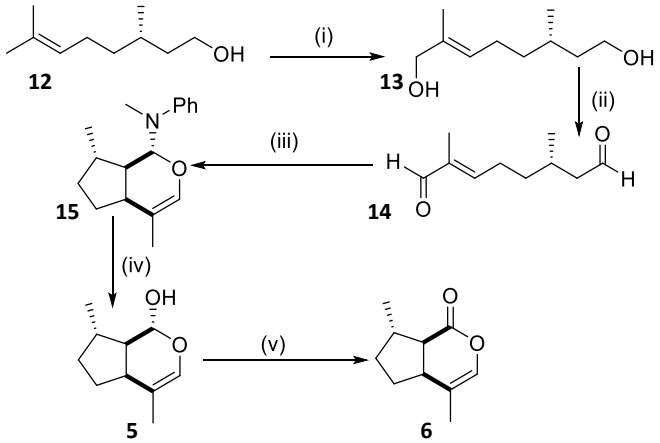
2. Synthesis and Purification of Aphid Sex Pheromone Components

2.1 INTRODUCTION

Aphids communicate using a sex pheromone consisting of two main components (4a*S*,7*S*,7a*R*)-nepetalactone **6** and (*1R,*4a*S*,7*S*,7a*R*)-nepetalactol **5**.[37] Aphids are capable of discrimination between the sex pheromone and it's respective enantiomers, (4a*R*,7*R*,7a*S*)-nepetalactone **33** and (*1S,*4a*R*,7*R*,7a*S*)-nepetalactol **32**, both of which are behaviourally and electrophysiologically inactive.[145] This discrimination ability is not fully understood, though olfactory proteins located in the antenna (odorant receptors and odorant-binding proteins) are hypothesised to play a role in discrimination of these enantiomers.

To fully understand the role olfactory proteins may play in this discrimination process, binding studies were performed with both the sex pheromone components and their respective enantiomers. High enantiomeric purity is required for behavioural efficacy of the pheromone, however, due to the multiple stereocentres present in (4a*S*,7*S*,7a*R*)-nepetalactone **6** and (*1R,*4a*S*,7*S*,7a*R*)-nepetalactol **5**, synthesis is both challenging and expensive.[6]

Although the synthesis of the naturally occurring enantiomers of nepetalactol and nepetalactone has been published (Scheme 2.1), they are readily available via steam distillation of the catmint plant, *Nepeta cataria*, which gives an oil containing the nepetalactone.[6] (4a*S*,7*S*,7a*R*)-Nepetalactone **6** can subsequently be reduced with sodium borohydride to yield (1*R*,4a*S*,7*S*,7a*R*)-nepetalactol **5**.[6]



Scheme 2.1: The devised synthetic route, based on the previously published synthetic route to nepetalactol and nepetalactone from (*S)*-citronellol **12**, by Dawson *et al.*[51] i) SeO_2, *t*-BuOOH, DCM; (ii) 1) $(COCl)_2$, CH_2Cl_2, DMSO, -78°C 2) Et_3N, CH_2Cl_2; (iii) PhNMe, Et_2O, 4Å sieves; (iv) TsOH, THF, H_2O; (v) Ag_2CO_3, celite, PhMe, 120°C

Citronellol is oxidised to the corresponding 1,8-dialdehyde over two steps (Scheme 2.1). The dialdehyde follows an intramolecular enamine-mediated [4+2] Diels-Alder-like cycloadditon.[52] This cycloaddition results in the desired bicyclic structure, which can be hydrolysed to nepetalactol and further oxidised to nepetalactone.[51] Though the synthetic route has previously been reported, the published yield is very low (< 1% total).[51] This provided the opportunity for optimisation of reaction conditions; for example, there are a range of gentle, high-yielding oxidations available to convert the diol to the dialdehyde. Furthermore, there has been no investigation of substituent effects on the aniline in the cyclisation step.

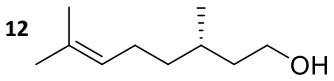
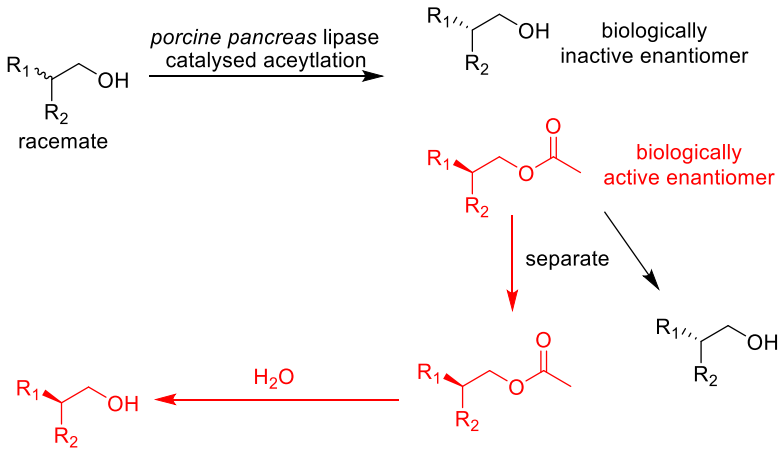


Figure 2.1: (*S*)-Citronellol **12**, the starting material for synthesis of (4a*S*,7*S*,7a*R*)-nepetalactone **6** and (*1R,*4a*S*,7*S*,7a*R*)-nepetalactol **5**.

Synthesis starting with (*S*)-citronellol results in the naturally occurring (4a*S*,7*S*,7a*R*)-nepetalactone **6** and (*1R,*4a*S*,7*S*,7a*R*)-nepetalactol **5**, whereas using (*R*)-citronellol yields the non-naturally occurring enantiomers, (4a*R*,7*R*,7a*S*)-nepetalactone **33** and (1*S*,4a*R*,7*R*,7a*S*)-nepetalactol **32**. Generally, enantiomerically pure starting materials are higher in cost when compared with the corresponding racemate. In an effort to reduce costs, racemic resolutions of a racemic mixture can be used to separate the two different stereoisomers. The resolution of racemic alcohols, such as citronellol, generally involve enantioselective acetylation of a specific enantiomer typically catalysed by lipase enzymes (Scheme 2.2).[146] The acetylated

enantiomer can then be chromatographically separated and hydrolysed back to the alcohol if required. A commonly utilised enzyme for his acetylation is lipase from porcine pancreas.[146]



Scheme 2.2: Schematic example of racemic resolution by acetylation of a racemic alcohol with lipase from *porcine pancreas*.

The aim of this chapter is to synthesise both aphid sex pheromone components, (4a*S*,7*S*,7a*R*)-nepetalactone **6** and (*1R*,4a*S*,7*S*,7a*R*)-nepetalactol **5**, and their respective non-naturally occurring enantiomers, (4a*R*,7*R*,7a*S*)-nepetalactone **33** and (1*S*,4a*R*,7*R*,7a*S*)-nepetalactol **32**. Resolution of racemic (*R/S*)-citronellol to (*R*)-citronellol will be attempted, and the synthesis of the non-naturally occurring enantiomer optimised, specifically the oxidation and cyclisation steps.

2.2 RESULTS & DISCUSSION

2.2.1 Synthesis and Purification of (1*R*,4a*S*,7*S*,7a*R*)-Nepetalactol and (4a*S*,7*S*,7a*R*)-Nepetalactone

2.2.1.1 Purification of (4aS,7S,7aR)-nepetalactone from Nepeta cataria oil

The naturally occurring enantiomer of nepetalactone, (4a*S*,7*S*,7a*R*)-nepetalactone **6**, can be isolated from the steam-distilled oil of the catmint plant, *Nepeta cataria.*[6] The oil consists of two major components, approximately 25% (4a*S*,7*S*,7a*R*)-nepetalactone **6** and 50% caryophyllene, with the remainder consisting of various other plant-produced compounds.

N. cataria steam-distilled oil was initially purified on silica gel with ethyl acetate and petroleum ether in a 3:2 ratio *via* flash column chromatography. The resultant (4a*S*,7*S*,7a*R*)-nepetalactone **6** was purified further for binding studies *via* reverse-phase high-performance liquid chromatography (HPLC), using an acetonitrile/water solvent system and an ACE HQ

C18 column. Fractions containing the product were collected and concentrated *in vacuo.* Unfortunately, the majority of the product was lost during the concentration process resulting from the vigorous conditions. The volatile samples were in water, which requires a higher temperature and higher vacuum, therefore lower pressure, evaporation than other solvents, leading to evaporation of the product.

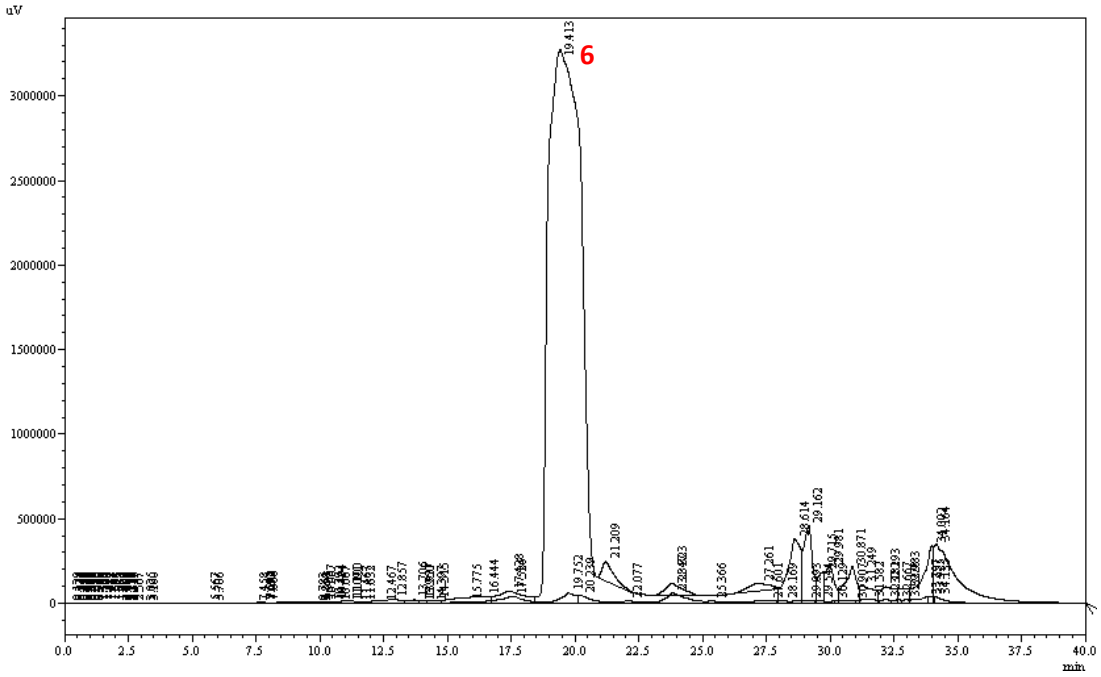


Figure 2.2: High-performance liquid chromatography (HPLC) traces of purified of (4aS,7S,7aR)-nepetalactone **6** and various impurities in reverse-phase (acetonitrile and water).

Normal-phase HPLC, using volatile solvents petroleum ether and ethyl acetate, was attempted to overcome this issue. As these solvents are easier to remove *in vacuo*, the final product of (4a*S*,7*S*,7a*R*)-nepetalactone **6** was successfully isolated with a final yield of 17%.

2.2.1.2 Reduction to (1R,4aS,7S,7aR)-nepetalactol

(*1R,*4a*S*,7*S*,7a*R*)-Nepetalactol **5** was synthesised successfully by a sodium borohydride reduction of (4a*S*,7*S*,7a*R*)-nepetalactone **6**. Stereoselective reduction is achieved as the hydride generally attacks the least hindered face of the lactone. This mechanism is a stereoselective reaction – the literature precedent for this reaction shows the reduction is stereospecific, yielding a 92:8 ratio of diastereomers.[6,147] The major diastereoisomer obtained in this reaction is the opposite to the expected result, and is a result of an attack by the hydride at the *most* hindered face, a phenomenon previously reported with the synthesis of some iridoids.[147]

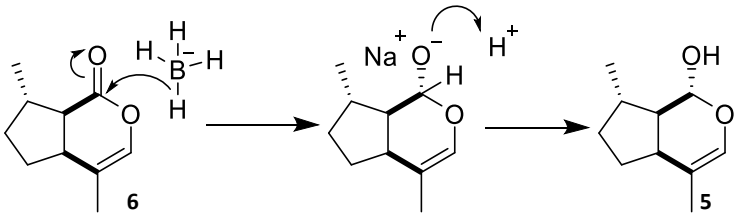


Figure 2.3: The mechanism of sodium borohydride reduction of nepetalactone **6** to nepetalactol **5.**

A diastereomeric ratio of 90:10, with a diastereomeric excess (de) of 80%, was achieved in this reaction as determined by NMR spectroscopy (Figure 2.4). (*1R,*4a*S*,7*S*,7a*R*)-Nepetalactol **5** was purified for future use, as nepetalactone (2.2.1.1), with normal-phase HPLC.

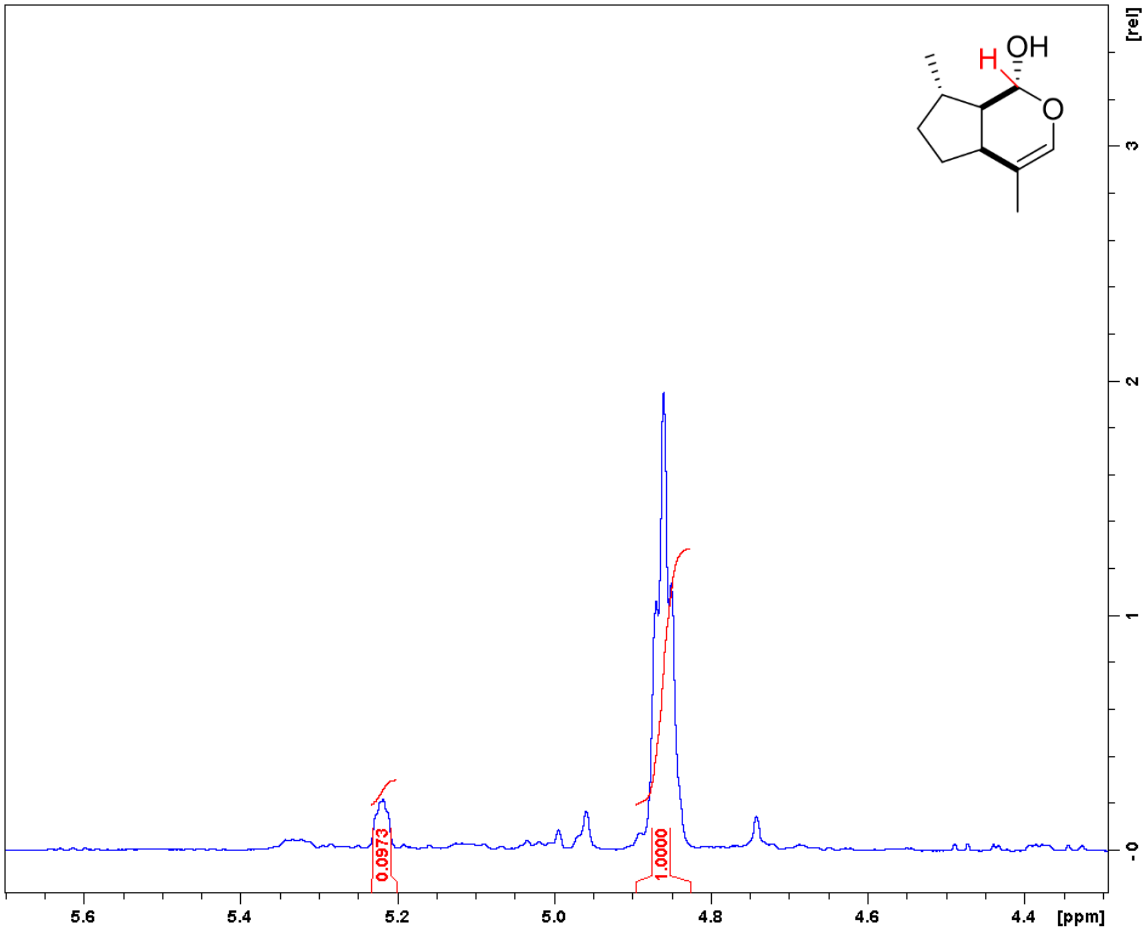


Figure 2.4: ^{1}H-NMR showing the proton adjacent to the alcohol in nepetalactol, for which diastereomeric differences can be clearly observed. The peak at 4.86 ppm represents the major diastereomer, whereas the peak at 5.22 ppm shows the minor diastereomer.

2.2.2 Resolution of (*R/*S)-Citronellol to give (*R*)-Citronellol

The non-naturally occurring enantiomers of nepetalactol and nepetalactone can be synthesised from enantiomerically pure (*R*)-citronellol. Enantiomerically pure materials are expensive, therefore resolution of the comparatively inexpensive racemate can be a useful

source of enantiomerically enriched starting materials; racemic citronellol costs £0.17/g and (*S*)-citronellol £10.00/g, in comparison to (*R*)-citronellol at £47.30/g (Sigma-Aldrich, 2019).

Resolution of racemic (*R/S*)-citronellol was initially attempted using a stereoselective, enzyme mediated acetylation. Alcohols, including citronellol, can be stereoselectivley acetylated by an enzyme, such as lipase from *Porcine pancreas*. This results in two products; one acetylated enantiomer, and the other enantiomer remaining as the unacetylated alcohol. If the desired enantiomer is acetylated by the enzyme, a simple ester hydrolysis can be used to convert back to the desired alcohol. In order to determine if the resolution is a success, a suitable analytical technique capable of distinguishing the enantiomers is required. Typically, chiral-GC is used, whereby a homochiral stationary phase (e.g. β-cyclodextrin) is used to afford separation of the different stereochemistries.

Initially, the acetylated forms of racemic citronellol were synthesised to ensure the two acetylated enantiomers could be separated using chiral-GC. Acetylation of (*R/S*)-citronellol to citronellyl acetate with acetic anhydride and pyridine was successful. However, only a single peak was observed when the racemic mixture was analysed using chiral gas chromatography, indicating that no separation of the enantiomers was achieved (Figure 2.5). Even after optimisation of the chiral-GC oven conditions, no separation was observed.

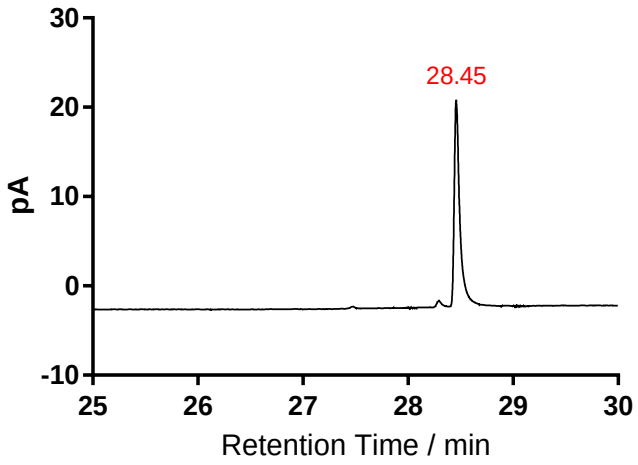


Figure 2.5: Chiral gas chromatograph of citronellyl acetate. Only one peak, at 28.45 min, can be observed; the two enantiomers are not resolved.

As the two enantiomers of (*R/S*)-citronellyl acetate could not be resolved, alternative acetylated products were produced. Acetylation to the butyrate with butyryl chloride and to the trifluoroacetate with trifluoroacetic anhydride were both attempted to investigate their potentials for chiral GC. The synthesis of the acetylated forms of citronellol were successful,

however chiral separation with gas chromatography for either alternative acetylated product could not be achieved (Figure 2.6).

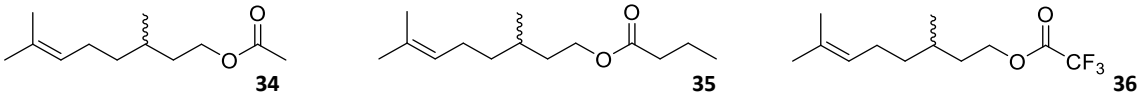


Figure 2.6: The acetylated forms of (*R/S*)-citronellol synthesised in racemic resolution attempts; citronellyl acetate **34**, citronellyl butyrate **35** and citronellyl a2,2,2-trifluoroacetate **36**.

Other analytical methods, such as nuclear magnetic resonance (NMR) spectroscopy may prove more sensitive in differentiating between diastereoisomers. Future racemic resolution attempts could involve coupling with Mosher's acid chloride to make diastereoisomers, and subsequent enantiomeric resolution *via* NMR. However, in the interest of time, this was not pursued.

2.2.3 Synthesis of (1*S*,4a*R*,7*R*,7a*S*)-Nepetalactol from (*R*)-Citronellol

2.2.3.1 Allylic oxidation of (R)-citronellol to (R)-8-hydroxycitronellol

With the inability to prepare (*R*)-citronellol by chiral resolution, it was decided to purchase the expensive commercially available material. The first step in the synthesis of (1*S*,4a*R*,7*R*,7a*S*)-nepetalactol **32** from (*R*)-citronellol is the allylic oxidation of (*R*)-citronellol to (*R*)-8-hydroxycitronellol and (*R*)-8-oxocitronellol using selenium dioxide. Oxidation with selenium dioxide, also known as a Riley oxidation, can be used to perform multiple types of oxidations, including allylic oxidations (Figure 2.7)[148]. This reaction is particularly slow allowing for easy isolation of the alcohol product. Production of (*R*)-8-oxocitronellol can occur by a second oxidation step when the reaction is left for a longer period of time, producing an aldehyde-hydrate at the 8 position, followed by an elimination of water to generate the aldehyde. Generally, the ratio of aldehyde to alcohol increases over time as further oxidations occur; this was observed with the oxidation of citronellol.

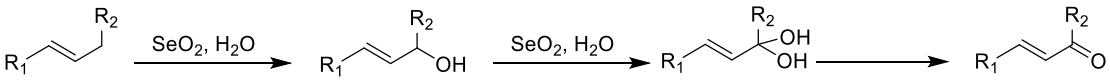
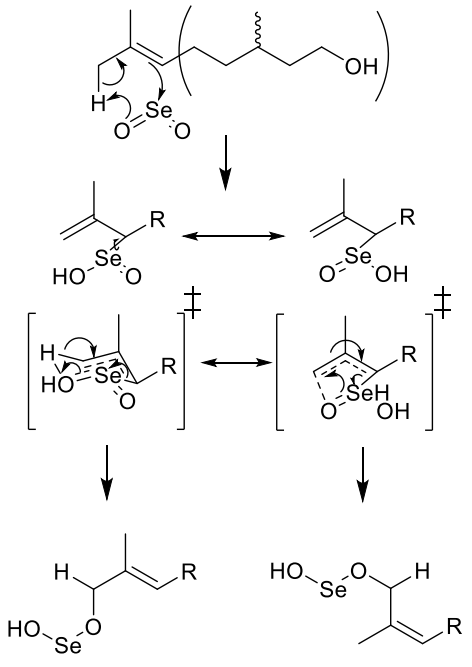


Figure 2.7: Allylic oxidation with selenium dioxide to produce the alcohol. A second oxidation step can result in the production of an aldehyde.

With citronellol, the oxidation occurs at the 8-carbon only in the *trans* position due to the chair transition state conformation adopted during the mechanism (Figure 2.8 [i]). π-Electrons of the alkene attack the selenium atom, forming a selenic acid-containing intermediate that undergoes a [2,3]-sigmatropic rearrangement to give an allylic selenic acid. The *trans*-selectivity is driven by the stability of this transition state, which can exist in one

of two conformations (Figure 2.8 [ii]), the more stable is a chair conformation yielding the trans geometry.

(i)



ii)

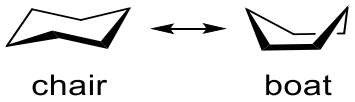


Figure 2.8: (i) The mechanism of the allylic oxidation of citronellol with selenium dioxide, and the two possible transition states by which it occurs[149]; (ii) Boat and chair conformations of 6-membered ring structures, such as cyclohexane

The allylic selenic acid is hydrolysed with *tert*-butyl hydroperoxide in water to yield the final alcohol product. Allylic oxidation with selenium dioxide is a slow and low-yielding step – the highest observed yield was around 35% - however, there are no viable alternatives for this synthetic step.

2.2.3.2 Oxidation of (R)-8-hydroxycitronellol and (R)-8-oxocitronellol to (R)-8-oxocitronellal

The second synthetic step involves the generation of a dialdehyde *via* oxidation of the alcohol functional groups at position 1 and 8. This oxidation can be performed by a multitude of soft-oxidation methods required to selectively yield the desired aldehyde without over-oxidation to the corresponding carboxylic acid. Initially, oxidation using tetrapropylammonium perruthenate (TPAP)/*N*-methylmorpholine *N*-oxide (NMO) and 2-iodoxybenzoic acid (IBX)

were attempted. However, both gave a low yield, and the use of IBX suffered from practical difficulties including photodegredation and poor solubility in organic solvents (Table 2.1).

A Swern oxidation provided higher yield overall (crude yield 77% vs. 16% with other methods). Dess-Martin oxidation was also investigated, which gave a reasonable yield after a reaction time of only one hour, without the undesirable by-products associated with the Swern oxidation. However, for this work, a combination of Swern and Dess-Martin oxidations were used. The dialdehyde product of this reaction step was particularly unstable, and the next reaction step was performed immediately.

Table 2.1: Comparison of methodologies utilised for the oxidation of (*R*)-8-hydroxycitronellol and (*R*)-8-oxocitronellol to (*R*)-8-oxocitronellal

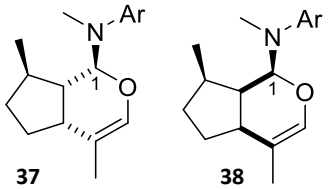
Oxidation Method	Yield	Time	Reaction Conditions
TPAP/NMO	29%	1-2 h	Anhydrous, RT
IBX	16%	3 h	RT
Dess-Martin	54%	1 h	Anhydrous, RT
Swern	77% (crude)	1 h	Anhydrous, -78°C, inert atmosphere

2.2.3.3 Cyclisation of (R)-8-oxocitronellal to the N-methyl aniline adduct of nepetalactol

The next step in the synthesis of nepetalactol involves the conversion of the 1,8-dialdehyde (*R*)-8-oxo-citronellol to the cyclised *N*-methylaniline adduct *via* an intramolecular [4+2] cyclisation. This step is based on the work of Schreiber *et al*[52], and gives a mixture of the two possible diastereoisomers.

The use of *N*-methylaniline as the secondary amine in the reaction, as described in the initial literature, gives a successful cyclisation with a diastereomeric excess of 72%. This excess can be measured by observing the proton NMR peak corresponding of the proton adjacent to the aniline, using the peak integration value to calculate the excess percentage of this diastereoisomer over the minor diastereoisomer (Figure 2.9 (i) – 1-H). The highlighted proton in the ^{1}H-NMR spectra appears as a doublet, with visible separation of the major and minor diastereoisomers at 4.71 ppm and 4.81 ppm respectively (Figure 2.9 (ii)).

(i)



37 38

(ii)

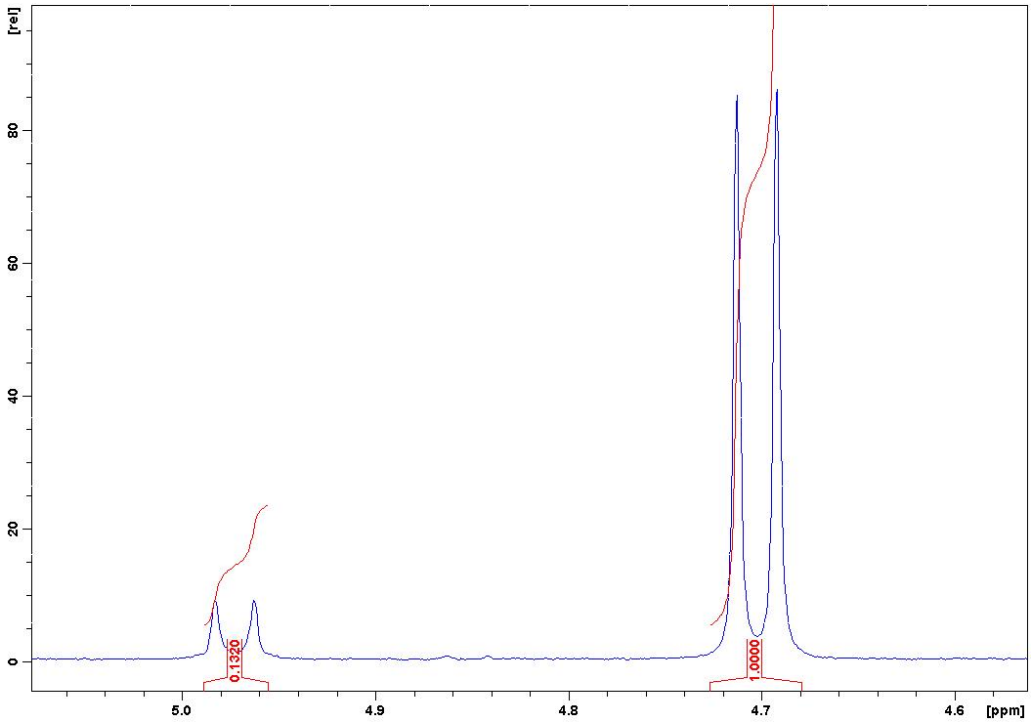


Figure 2.9: (i) The major (**37**) and minor (**38**) diastereoisomers of *N*-4,7-trimethyl-N-phenyl-1,4a,5,6,7,7a-hexahydrocyclopenta[c]pyran-1-amine; (ii) ^{1}H-NMR showing the proton on carbon 1, for which diastereomeric differences can be clearly observed. The peak at 4.71 ppm represents the major diastereomer, whereas the peak at 4.81 ppm shows the minor diastereomer. At 4.85 ppm, a potential third diastereoisomer can also be observed.

To improve the diastereoselectivity, substituent effects of the aniline moiety on the cyclisation were investigated. When Schreiber first described this cyclisation in 1986, little was known about substituent effects on the reactivity of the aniline.[52] However, modification of the aniline with different substituents should affect the electronics of the reaction and was expected to alter the diastereoselectivity and reactivity.

Initially, an electron-donating methoxy (-OMe) group was added to the aniline ring to determine the effect on the diastereomeric ratio. The reaction with *N*-methyl, 4-methoxyaniline consumed all the starting material after 16 hours, as with standard *N*-methylaniline. Unfortunately, during the flash column chromatography on silica gel, the final product decomposed. As a result, the pure product could not be isolated. However, crude proton NMR indicated a successful cyclisation with a diastereomeric excess (de) of 65%, a

deterioration of the initial 72%. However, when the reaction was left for a further 16 hours (32 hours in total), a crude yield of 61% was obtained with a de of 72% as determined from the crude NMR (Table 2.2).

To further investigate electronic effects, an electron-withdrawing nitro (-NO_2) group was included at the 4-position by using *N*-methyl-4-nitroaniline as the secondary amine. No reaction had occurred after 16 hours, though after 32 hours product began to appear on the TLC. The reaction was left for a total of 72 hours (Table 2.2). From crude NMR analysis, only a small amount of the mixture had cyclised (≈ 14%). The product could not be isolated due to the low yield and insolubility of the final product. However, only one major diastereoisomer was observed in the crude NMR – this suggests only one isomer was produced during the reaction. However, as such low amounts of product were produced overall, the minor diastereoisomer may have been produced in negligible amounts and was therefore not visible in the NMR spectra.

The final substituent investigated was a chloro (-Cl) group on the 4-position *via* a reaction with *N*-methyl- 4-chloroaniline as the secondary amine. This cyclisation was successful, with comparable yields to the standard *N*-methylaniline. However, a significant improvement (72% to 81%) was observed in the diastereomeric ratio (Table 2.2). Chlorine behaves electronically as an electron withdrawing group but has significantly less electronic effect compared to the nitro group. We would therefore expect to observe similar results compared with the nitroaniline, but at the sacrifice of diastereoselectivity, as observed in this reaction. In subsequent synthesis of nepetalactol and nepetalactone, 4-chloro-*N*-methylaniline should provide improved diastereomeric selectivity at no cost to yield, versus *N*-methylaniline.

Table 2.2: Comparison of substituent effects on the [4+2] cycloaddition with the N-methyl aniline. The effect of substituents on the 4-position have been investigated.

*Calculated from crude nuclear magnetic resonance (NMR) spectroscopy

***N*-methyl aniline R-group at 4-position**	**Yield**	**Time**	**Diastereomeric Excess**
H	47%	16 h	72%
NO_2	Crude 14% *	72 h	100%
OMe	Crude 39%	16 h	65%
	Crude 61%	32 h	72%
Cl	45%	16 h	81%

There are a few observations to be made from these reactions; the substituent on the aniline confers changes in diastereoselectivity and yield of the reaction, and these appear to be inversely affected; a substituent change that causes an increase in yield diastereoselectivity is sacrificed. For this full synthesis, the most favourable reaction would be one in which there is a balance of yield and diastereoselectivity. The observation of these results, specifically the differences between the electron donating and withdrawing substituents, may provide evidence relating to the mechanism of this reaction. Finally, the changing yield and diastereomeric excess observed with the methoxy substituent when the reaction is left for an extra 16 hours may suggest that the reaction is reversible.

The initial step in the mechanism is the generation of an enamine at the 8-position of the dialdehyde, followed by cyclisation. The mechanism for the cyclisation step is generally described as a Diels-Alder like [4+2] intramolecular cycloaddition between a diene (such as the enal) and dienophile (such as the enamine alkene, adjacent to the aniline group).[51] In a Diels-Alder reaction mechanism, the highest occupied molecular orbital (HOMO) of the diene interacts with the lowest unoccupied molecular orbital (LUMO) of the dienophile (Figure 2.10). The energy level of the HOMO can be increased by the presence of an adjacent electron donating group, whereas the energy of the LUMO can be decreased with the presence of an adjacent electron withdrawing group, thereby reducing the difference in energy levels promoting a more energetically favourable reaction. (Figure 2.10).

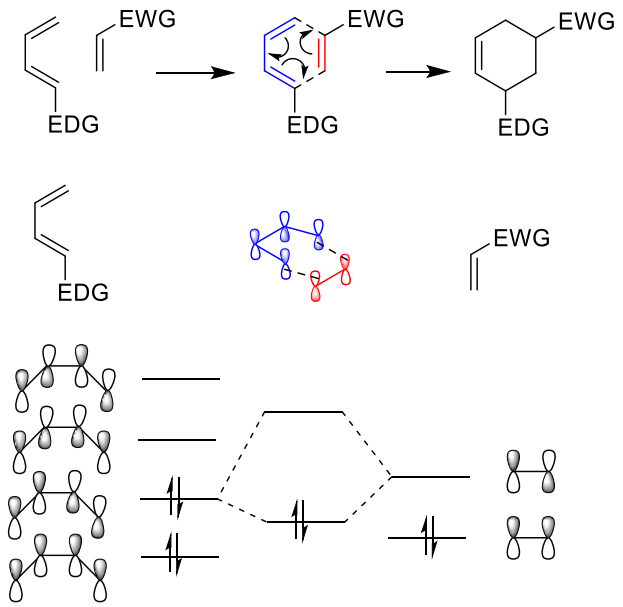


Figure 2.10: A typical Diels-Alder reaction. A pericyclic reaction occurs where the HOMO of the diene interacts with the LUMO of the dienophile. The presence of an electron-donating group on the diene (HOMO) and electron withdrawing group on the dienophile (LUMO) lowers the energy gap between the two molecular orbitals, resulting in a more energetically favourable reaction.

A classical Diels-Alder reaction is also endo-directing mechanism. The cycloaddition mechanism must obey Woodward-Hoffman general selection rules – for a three-component electronic system, as seen here, all components must interact in a suprafacial manner resulting in a boat-like transition state (Figure 2.11). A chair-like transition state is thermally disallowed by the Woodward-Hoffman rules. Furthermore, stereochemistry of the reactants is retained, following the Alder-Stein principle – the aniline group is *trans* in geometry to the rest of the molecule before cyclisation, resulting in a bent final molecule with the cyclopentane group on an opposing face to the aniline. An electron withdrawing group is favoured in the *endo* position as the orbitals of both the HOMO and LUMO are in the best position to interact, with an overlap of molecular forming a boding interaction.[150,151] Electron donating groups are subsequently less favoured in the endo position than in the *exo.*[150] Although the *endo* product is usually more sterically hindered, resulting in the *exo* product being the thermodynamic product of a Diels-Alder reaction, the interaction between overlapping orbitals results in the *endo* product generally being the major product (Figure 2.11).[150] If the reaction is more reversible, then the more stable *exo* product is formed more over time. Substituent effects and subsequent electronics can affect the energetic favourability of the product, depending on their geometry.

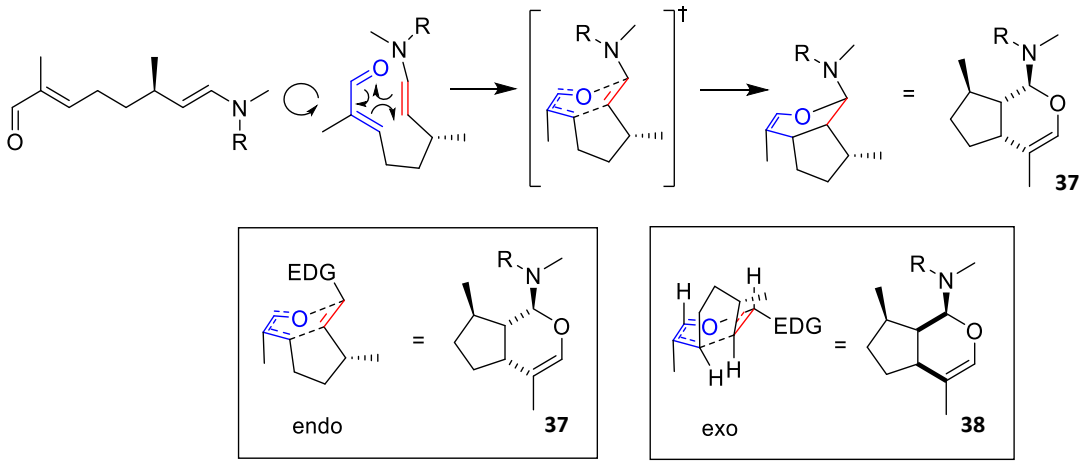


Figure 2.11: The boat-like transition states of the cycloaddition reaction mechanism, as required by Woodward-Hoffman rules. Stereochemistry is retained, as the Alder-Stein principle.

In addition to the classical Diels-Alder reaction, there exists an inverse electron demand Diels-Alder reaction (IEDDA) mechanism.[151,152] In an IEDDA mechanism the diene would provide the LUMO and the dienophile would provide the HOMO (Figure 2.12).[152] This type of Diels-Alder is often seen when a heteroatom (in this case, the oxygen of the enal) is involved in the cycloaddition step. An electron donating group on the dienophile, now the HOMO, decreases the energy barrier and allows a more energetically favourable reaction and therefore predicted higher yields (Figure 2.11). Inverse electron demand Diels-Alder reactions are still *endo*-directing, however, due to the differences in the electronics of the HOMO and LUMO, electron donating groups are more favourable in the *endo*-position in relation to the ring relative to electron withdrawing groups, the inverse of a classic Diels-Alder (Figure 2.11).[151]

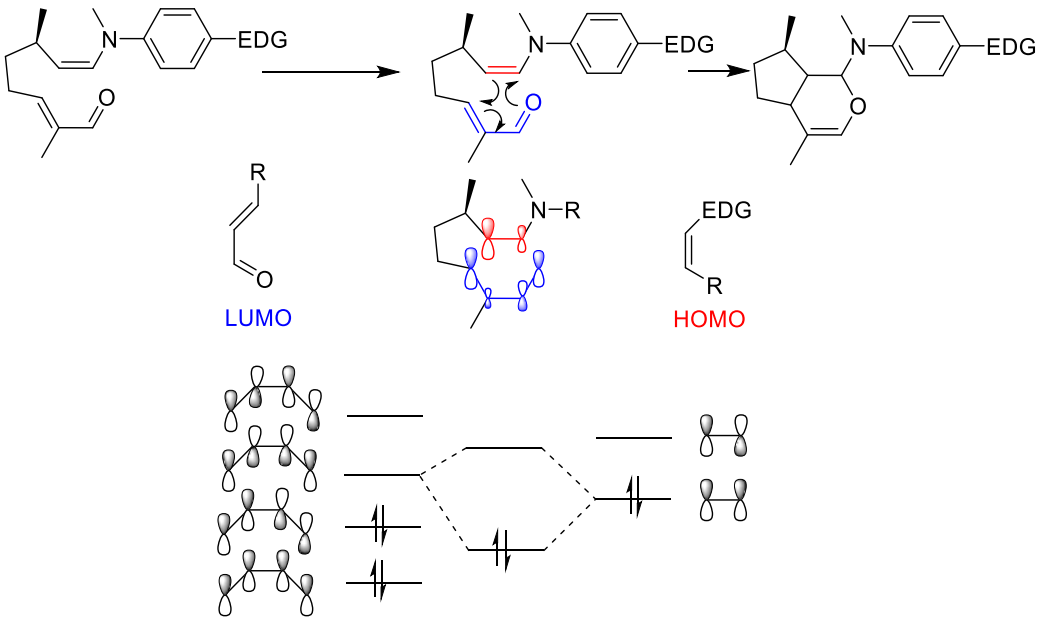


Figure 2.12: The intramolecular inverse electron demand Diels-Alder reaction. A pericyclic reaction occurs where the HOMO of the dienophile interacts with the LUMO of the diene. The presence of an electron-donating group on the dienophile (HOMO) lowers the energy gap between the two molecular orbitals, resulting in a more favourable reaction.

In this specific intramolecular cyclisation, the electronics of the dienophile were modified by altering substituents of the aniline. It was observed that an electron donating effect, such as the addition of a methoxy group, appeared to result in a preliminary increase in yield (after 32 hours). Of the three substituents trialled, the stronger the electron donating effect, with a chloro group giving a weaker electron withdrawing effect than a nitro group, the lower crude yield of the reaction. This is a contrast to a typical Diels-Alder reaction (Figure 2.10) and more reflecting of an IEDDA (Figure 2.12). Furthermore, the major diastereoisomer observed in this reaction is the *endo* product. A change in diastereomeric ratio occurs, with an increase of major diastereoisomer, when the dienophile is adjusted to be more electron donating. These results may suggest that this reaction occurs *via* an IEDDA reaction, though further investigation is required for confirmation.

In addition to the observed effects, the change in diastereomeric ratio with reaction time seen with in the reaction with *N*-methyl-4-methoxyaniline may suggest the cyclisation is reversible. To further investigate this phenomenon, NMR-observed reactions between the dialdehyde with *N*-methylaniline and 4-methoxy-*N*-methylaniline were conducted (Figure 2.13). The two compounds were combined in deuterated chloroform and allowed to react, with proton NMRs being taken at one-hour intervals for 16 hours.

As expected, 4-methoxy-*N*-methylaniline reacted much quicker than *N*-methylaniline – the product was visible in the initial NMR after 60 mins while the *N*-methylaniline adduct was only observed after 4 hours. Both reactions occurred less favourably in chloroform, with a maximum estimated yield between 25 and 30%. In future experiments, alternative solvents such as deuterated diethyl ether, may be explored to better replicate lab synthetic conditions.

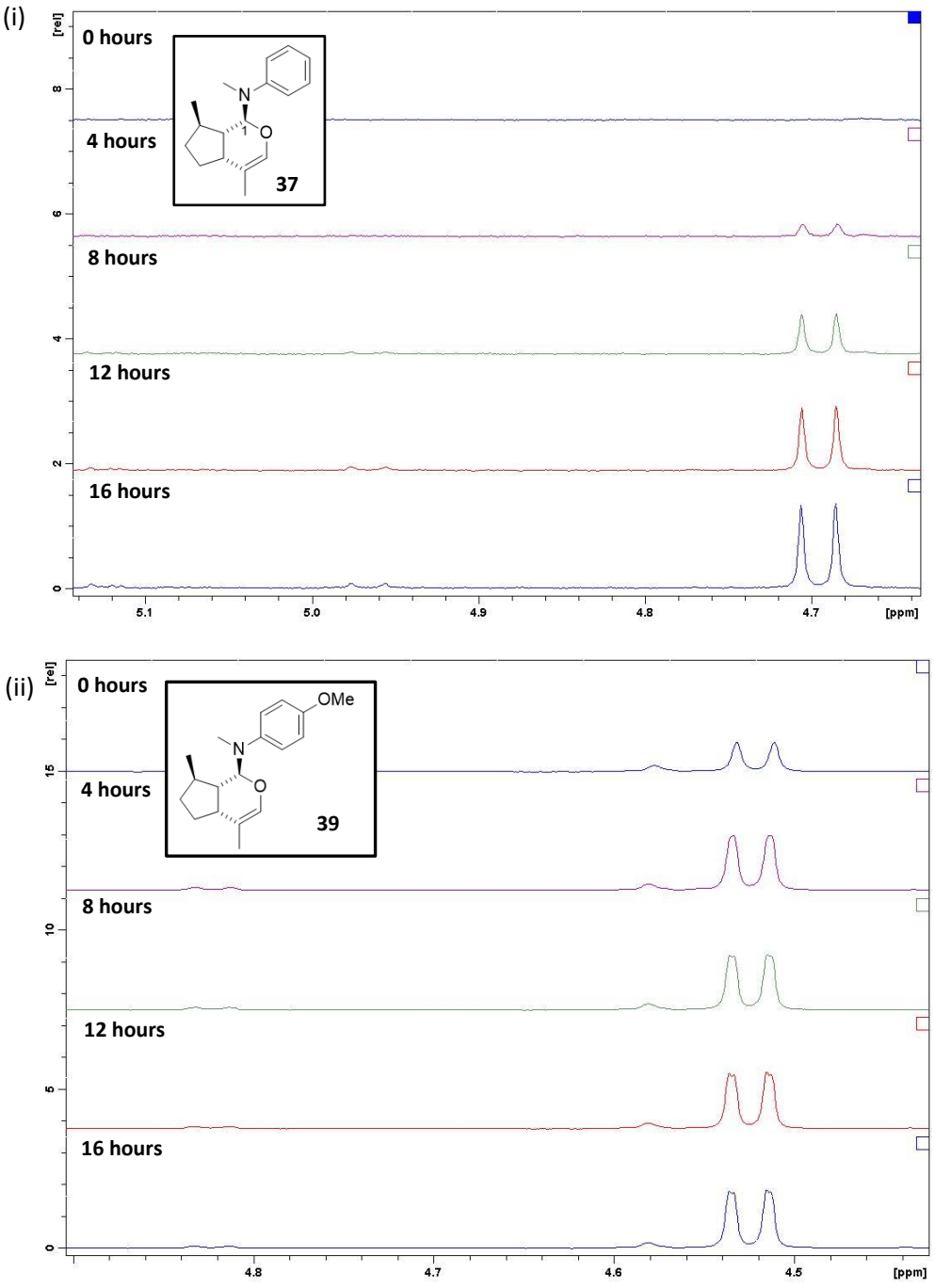


Figure 2.13: ^{1}H NMR of the production of (i) N,4,7-trimethyl-N-phenyl-1,4a,5,6,7,7a-

hexahydrocyclopenta[c]pyran-1-amine **37** and (ii) N,4,7-trimethyl-N-(4-methoxyphenyl)-1,4a,5,6,7,7a-hexahydrocyclopenta[c]pyran-1-amine **39** showing the 1-H proton, for which diastereomeric differences can be observed. Spectra are shown (top to bottom) for 0 hours, 4 hours, 8 hours, 12 hours and 16 hours. The peaks at (i) 4.71 and (ii) 4.51 ppm represent the major diastereomer, whereas the peaks at (i) 4.82 and (ii) 4.83 ppm show the minor diastereomer, respectively. Some impurities can also be seen.

From the crude NMRs, the diastereomeric excess did not appear to change throughout the 16-hour observed period. However, the slower reaction in the NMR gave comparably higher diastereomeric excess - around 80% versus 72% for the *N*-methylaniline product and around 92% versus 72% for the 4-methoxy, *N*-methylaniline product. These results demonstrate the susceptibility of the diastereomeric ratio to solvent effects, and future work could evaluate using chloroform as the solvent for the reaction to achieve a higher diastereoselectivity.

These initially optimisation studies highlighted that both standard *N*-methylaniline and its 4-chloro substituted version are practical options for the synthesis of aphid sex pheromone components. Indeed, by careful control of the reaction conditions (e.g. solvent, temperature), this preliminary investigation has indicated that it should be possible to greatly improve the yield and diastereoselectivity of these reactions.

2.2.3.4 Hydrolysis of the N-methyl aniline adduct to (1S,4aR,7R,7aS)-nepetalactol

To convert the bicyclic aniline adduct to (1*S*,4a*R*,7*R*,7a*S*)-nepetalactol **32**, a stereospecific hydrolysis with *para*-toluensulfonic acid (*p*-TsOH) was performed. The reaction occurs via a stereospecific mechanism by the formation of a double bond to the positively charge oxygen in the lactol ring (Figure 2.14). The final reaction step gives a by-product of *N*-methylaniline, which proved difficult to remove *via* flash column chromatography. The final crude mixture, therefore, needed to be purified using HPLC.

As in 2.1.1 with (4a*S*,7*S*,7aR)-nepetalactone **6**, reverse-phase HPLC was not successful. Normal-phase HPLC was used and (1*S*,4a*R*,7*R*,7a*S*)-nepetalactol **32** was successfully purified.

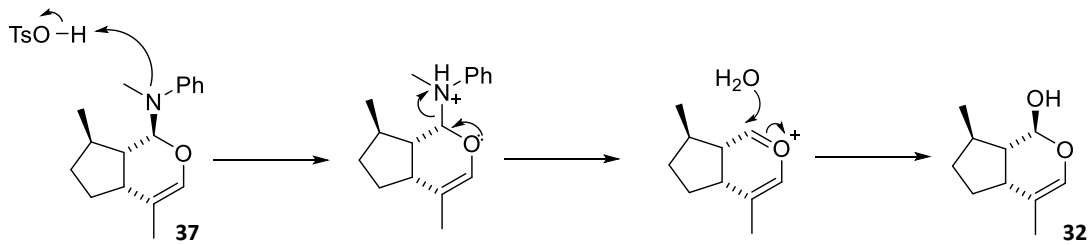


Figure 2.14: The mechanism of stereospecific hydrolysis with *p*-toluenesulfonic acid

2.2.4 Oxidation of (1*S*,4a*R*,7*R*,7a*S*)-Nepetalactol to (4a*R*,7*R*,7a*S*)-Nepetalactone

The final step in the synthesis of the nepetalactone from the nepetalactol can be achieved by a Fétizon's Oxidation with silver(I) carbonate adsorbed onto the surface of celite (Figure 2.15). In this oxidation, both the alcoholic oxygen and the neighbouring proton undergo a single electron oxidation by silver(I). The carbonate ion deprotonates the generated carbonyl, which is then protonated and water eliminated.

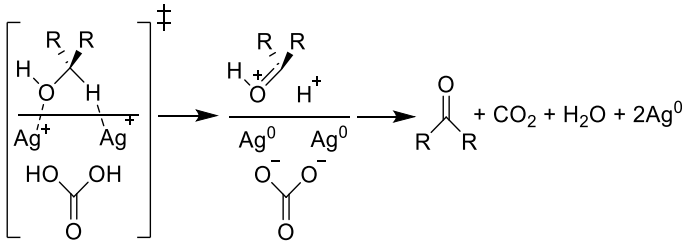


Figure 2.15: The mechanism of Fétizon's Oxidation of an alcohol with silver (I) carbonate on celite

The final synthetic product of (4a*R*,7*R*,7a*S*)-nepetalactone **33** was successfully produced and purified using normal-phase HPLC.

2.3 CONCLUSIONS

In conclusion, isolation of (4a*S*,7*S*,7a*R*)-nepetalactone **6** and synthesis of (1*R*,4a*S*,7*R*,7a*R*)-nepetalactol **5**, (1*S*,4a*R*,7*R*,7a*S*)-nepetalactol **32** and (4a*R*,7*R*,7a*S*)-nepetalactone **33** were successful.

Racemic resolution of (*R/S*)-citronellol was unsuccessful, though other options could be explored in future if large amount of the non-naturally occurring enantiomers are required. The synthetic route reported Dawson *et al.* was optimised, initially by investigating a selection of modern oxidation protocols for the diol oxidation to the corresponding 1,8-dialdehyde. Swern and Dess-Martin oxidations were found to be the most practically viable and highest yielding of the approaches investigated.

In the [4+2] cycloaddition step, diastereomeric ratio and reactivity differences were observed when different substituents were introduced onto the aniline ring reagent. The effects of electronic changes suggest that this cyclisation reaction may be an inverse electron demand Diels-Alder reaction, as opposed to a classical Diels-Alder. The preliminary optimisation of this cyclisation step carried out here could be further optimised in future work to increase the yield and desired diastereoselectivity.

Final synthesis and purification of the non-natural enantiomer was successful. Produced compounds from this chapter will be used protein-ligand binding studies in future chapters.

3. Overexpression, Purification and Characterisation of *A. pisum* Odorant-Binding Proteins

3.1 INTRODUCTION

Odorant-binding proteins (OBPs) are small, soluble proteins found in the antennae of insects and are hypothesised to possess an olfactory role.[112] Though the binding activity of some OBPs has been studied, their specific role in the olfactory systems of insects is yet to be understood. Before binding studies can be conducted large amounts of pure OBPs need to be produced. This can be achieved by generating recombinant bacterial lines *via* the transformation of *Escherichia coli* BL21(DE3) cells with DNA in the form of plasmids. The proteins may then be produced by over-expression of the gene in the recombinant *E. coli* and purification of the protein from cells.[153]

Competent cells are bacterial cells specifically primed for transformation; competency arises from alterations to the cell wall that allow DNA to pass through more easily through a process involving calcium chloride and heat shock.[154] Plasmids can be specifically designed to contain the gene relating to the protein of interest alongside other desired properties. Such desirable properties can include an antibiotic resistance cassette to allow for careful selection of successfully transformed *E. coli* colonies, or the addition of a specific promotor to express the gene and histidine-tag for later purification of the recombinant protein.[154]

Initially, during purifications steps, protein is released from the *E. coli* cells by sonication and the use of a surfactant. The protein can then be purified using nickel affinity chromatography (if a polyhistidine tag (His-Tag) is present) and size-exclusion chromatography, to separate the desired protein from other *E. coli* produced proteins.[153] Once purified, proteins are analysed using gel electrophoresis and mass spectrometry techniques. Gel electrophoresis, specifically sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) separates proteins by their molecular size using an electrical current. Mass spectrometry provides an accurate molecular weight of the recombinant protein to confirm its structure.

This chapter reports the transformation of competent *E. coli* to produce a range of *Acyrthosiphon pisum* odorant-binding proteins. The proteins were expressed on a large scale and characterised using gel electrophoresis and mass spectrometry.

3.2 RESULTS & DISCUSSION

3.2.1 Transformation of Competent *E. coli* BL21(DE3) Cells

Plasmids containing genes for *Acyrthosiphon pisum* odorant-binding protein (OBP1,2,3,4,5,6,7,8,9,10, 11) were separately transformed into E. coli BL21 (DE3) chemically competent cells. Successful transformation was determined by an agarose gel of PCR products from clonal cultures (Figure 3.1). Plasmids containing OBP4 and OBP11 could not be transformed successfully. It can be concluded that the plasmids were not up taken by the competent cells as no PCR products were seen for these transformations.

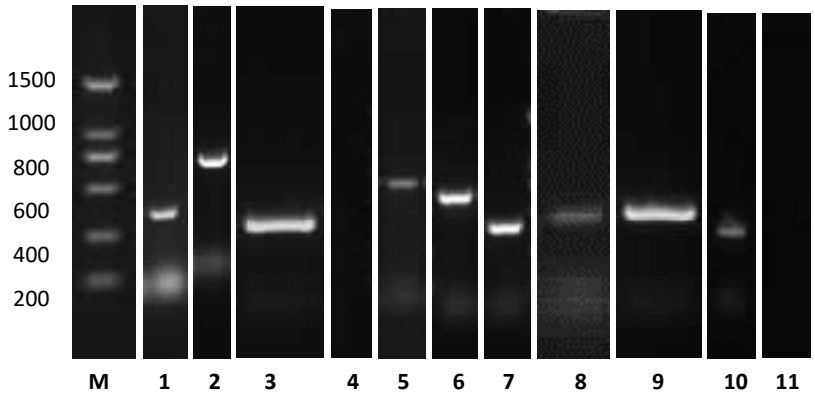


Figure 3.1: *A. pisum* odorant-binding protein PCR products from transformations. L-R: OBP1-His_6, OBP2-His_6, OBP3-His_6, OBP4-His_6, OBP5-His_6, OBP6-His_6, OBP7-His_6, OBP8-His_6, OBP9-His_6, OBP10-His_6 and OBP11-His_6. Size of DNA Marker (M; size in bp).

The plasmids, kindly donated by Jing-Jiang Zhou and colleagues, contained either an ampicillin (pET15b or pET45b) or kanamycin (pNIC-BSA28a) resistance cassette (Appendix Figure A.1, and a hexa-histidine (His_6) tag with a plasmid specific restriction enzyme site for future removal (Table 3.1, Table 8.6).

Table 3.1: Genes transformed into competent *E. coli.* All transformations were successful unless otherwise stated.

***A. Pisum* Protein**	**Gene Length (bp)**	**Vector**	**Resistance**
OBP1-His_6	480	pET15b	Ampicillin
OBP2-His_6	741	pET45b	Ampicillin
OBP3-His_6	423	pNIC28-Bsa4	Ampicillin
OBP4-His_6	579	Not successfully transformed	
OBP5-His_6	648	pNIC28-Bsa4	Kanamycin
OBP6-His_6	657	pET45b	Ampicillin
OBP7-His_6	444	pET45b	Ampicillin
OBP8-His_6	498	pNIC28-Bsa4	Kanamycin
OBP9-His_6	492	pET45b	Ampicillin
OBP10-His_6	426	pNIC28-Bsa4	Kanamycin
OBP11-His_6	414	Not successfully transformed	

Successfully transformed *E. coli* cells were grown in colonies on a small scale in LB media and glycerol stock cultures produced for use in all future protein expression and purification. Further transformations of OBP4-His_6 and OBP11-His_6 to OBP15-His_6 were not performed due to limited availability of cloned plasmids.

3.2.2 Small Scale Expression Test

Before large scale expression for protein purification could be performed, small scale expression optimisations were performed to assess the recombinant bacteria's expression capability and growth viability. Cultures were grown on a small scale (10 mL) in appropriate antibiotic-containing media and the growth monitored carefully by observing changes in OD_{600} (cell density) over a series of 5 hours (Figure 3.3).

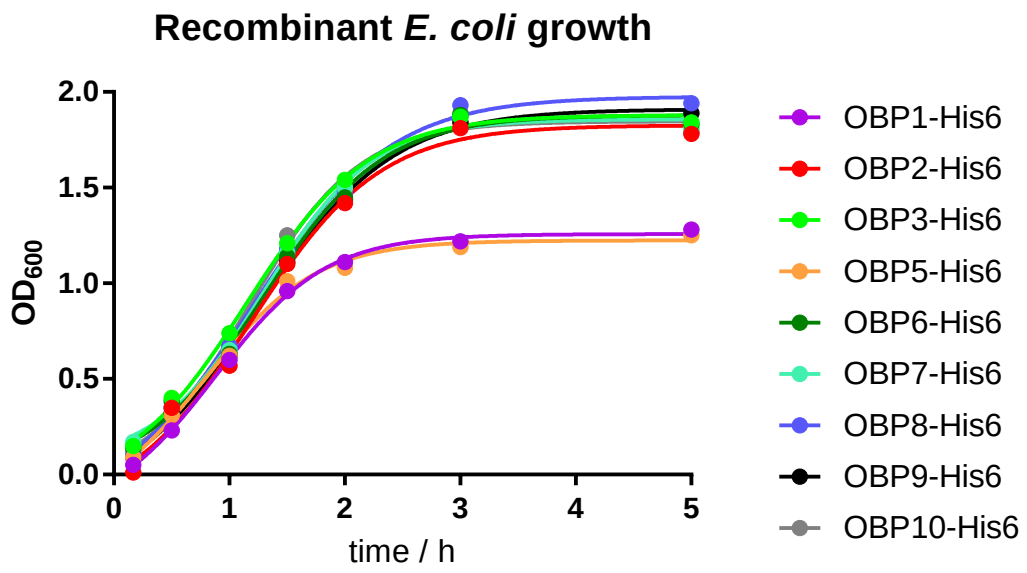


Figure 3.3 Curves showing the growth (by OD_{600}) of generated recombinant *E. coli* strains over time.

Strains of *E. coli* containing plasmids for OBP2-His_6,3-His_6,6-His_6,7-His_6,8-His_6,9-His_6, and 10-His_6 plateaued at an OD_{600} of approximately 1.9 after three hours. Two strains, OBP1-His_6 and 5-His_6, plateaued at only OD_{600} 1.25 after only two hours. The proteins produced by these strains (OBP1 and 5) may have potentially cytotoxic properties and are therefore not tolerated by *E. coli* and restricting growth. This was not investigated further, though in future more tolerable *E. coli* strains or alternative expressions systems could be used to produce OBP1-His_6 and OBP5-His_6.

A *lac operon* was encoded in the plasmid to allow for synthetic induction of protein expression. This promotor can be activated by the addition of isopropyl-β-D-thiogalactoside (IPTG), a lactose mimic, when *E. coli* growth is proceeding at an optimal rate, determined by measuring an OD_{600} value of 0.5-6 (Figure 3.3; Figure 3.4).

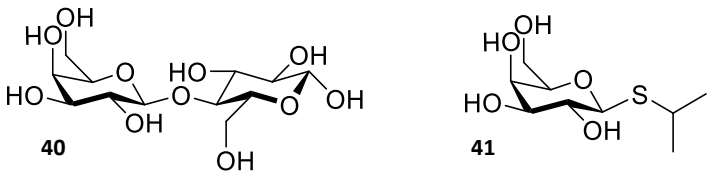
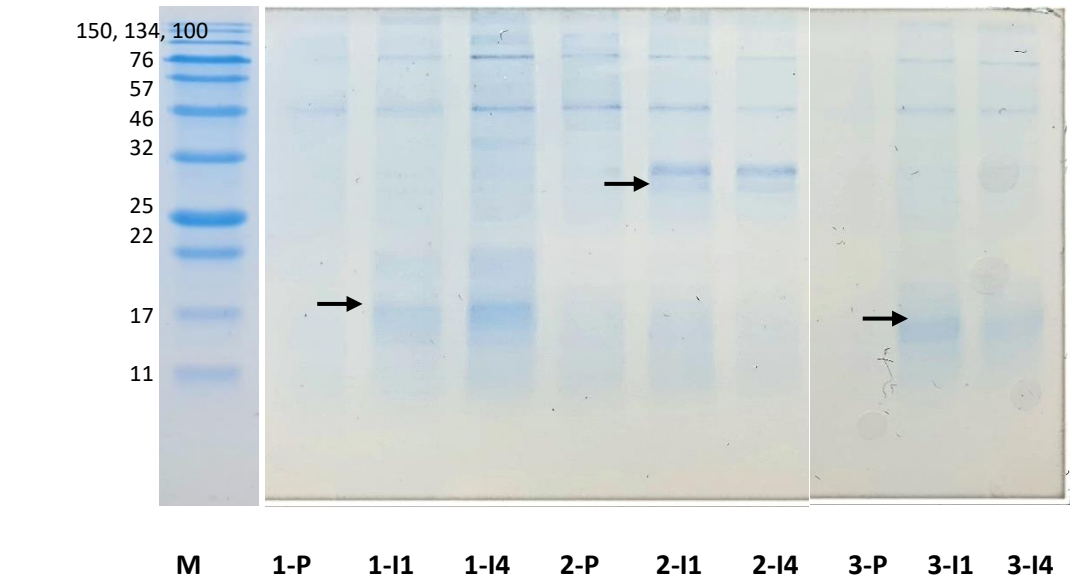
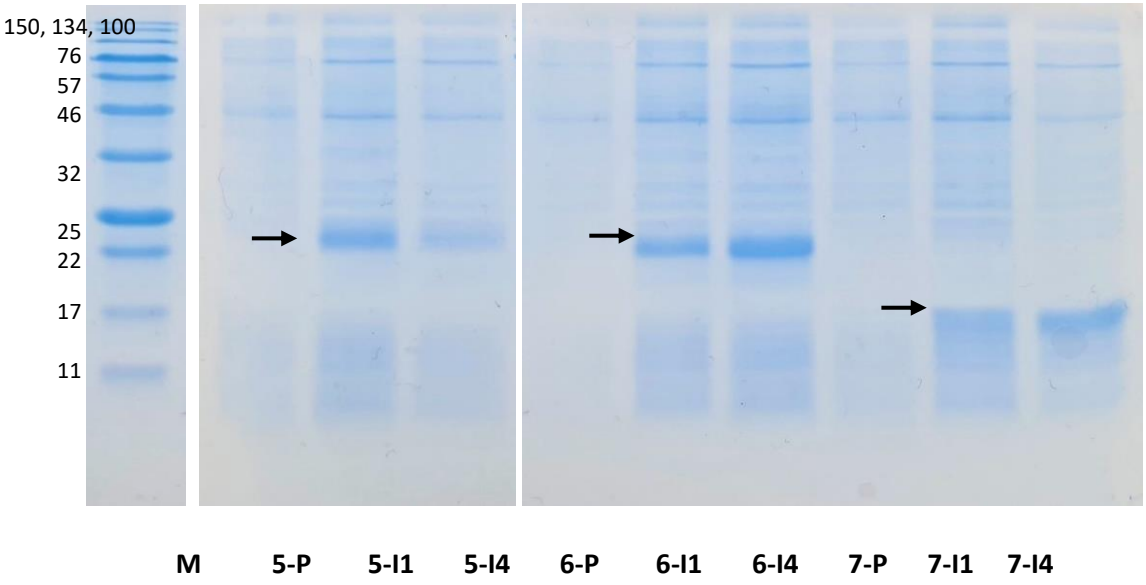


Figure 3.4: The chemical structure of lactose **40** and the lactose mimic isopropyl-β-thiogalactoside (IPTG) **41**; used to induce bacterial expression *via* activation of the *lac operon* promotor.

Samples of the small-scale expressions were collected pre-induction, in addition to one and four hour post-induction. A crude (SDS-PAGE) was performed to evaluate expression of the proteins (Figure 3.5).


150, 134, 100
76
57
46
32
25
22
17
11
M
1-P
1-I1
1-I4
2-P
2-I1
2-I4
3-P
3-I1
3-I4


150, 134, 100
76
57
46
32
25
22
17
11
M
5-P
5-I1
5-I4
6-P
6-I1
6-I4
7-P
7-I1
7-I4

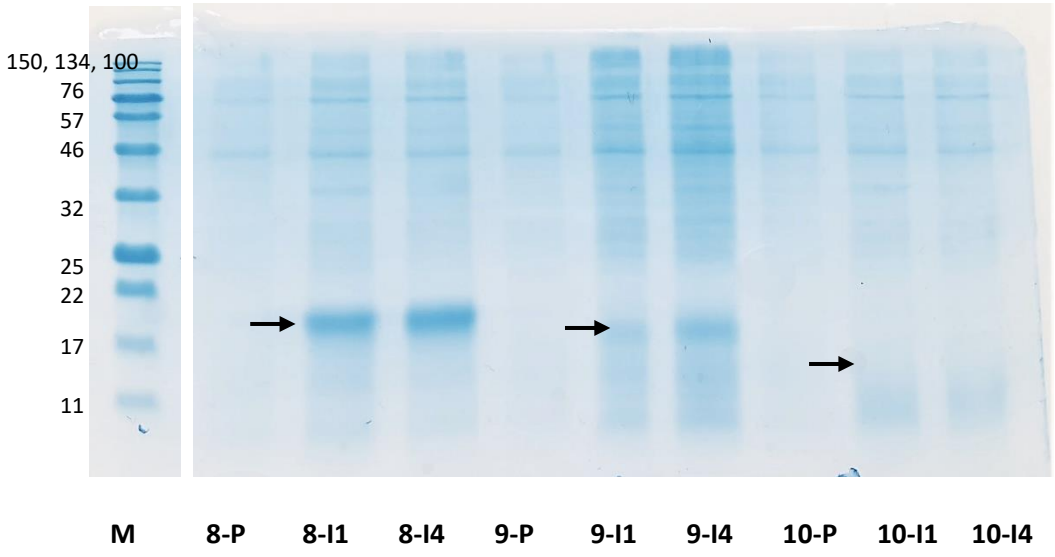

150, 134, 100
76
57
46
32
25
22
17
11
M
8-P
8-I1
8-I4
9-P
9-I1
9-I4
10-P
10-I1
10-I4

Figure 3.5: Expression and purification of OBP1-His_6, OBP2-His_6, OBP3-His_6, OBP5-His_6, OBP6-His_6, OBP7-His_6, OBP8-His_6, OBP9-His_6 and OBP10-His_6 from *A. pisum*. Electrophoretic analysis (SDS-PAGE) of crude bacterial pellets before (P), after 1-hour induction (I1) and after 4 hour induction (I4) with IPTG. Molecular weight (kDa) of markers (M) from the top is 150, 134, 100, 76, 57, 46, 32, 25, 22, 17, 11. Arrows represent the expected molecular mass of the proteins.

Generally, expression of the desired proteins was seen when the *E. coli* cultures were induced, however, relative levels of expression of the desired protein varied (Table 3.2). Protein expression was estimated by the intensity of equally stained SDS-PAGE bands (Table 3.2). As the main proteins of interest for future experiments expressed well in medium-high concentrations, the lower expressing strains were not further optimised. If these proteins were to be used in future, optimisation of the expression may be achieved, potentially *via* the use of a different competent cell line or vector construct.

Table 3.2: Size and relative expression levels (from SDS-PAGE gels) of expressed proteins. Proteins were expressed on a small scale and induced with IPTG for 3 hours.

*Expression levels were estimated from SDS-PAGE gel staining

***A. Pisum* Protein**	**Size (kDa)**	**Relative Expression***
OBP1-His_6	17.7	Medium
OBP2-His_6	26.7	Low/None
OBP3-His_6	16.0	Medium
OBP5-His_6	24.8	High
OBP6-His_6	25.3	High
OBP7-His_6	16.8	High
OBP8-His_6	18.8	High
OBP9-His_6	18.4	Medium
OBP10-His_6	15.9	Low/None

After successful small-scale optimisation, proteins of interest were expressed in large-scale experiments (< 1L) to produce large quantities of protein.

3.2.3 Large Scale Expression and Purification

OBP6, 7 and 9 were expressed and purified on large scales (1-2 L). Transformed recombinant *E. coli* BL21(DE3) cells were grown in appropriate antibiotic-containing media and induced with IPTG to produce the desired protein. The cells were harvested by centrifugation, producing a pellet.

The resultant cell pellet was sonicated with Triton X-100 in order to rupture cells and release the desired protein. OBPs are a promiscuous class of proteins with multiple cysteines (usually 6-8) that need to form disulphide bonds to result in the correct 3-dimensional structure.[101] It has previously been reported that these properties of OBPs can result in *E. coli* fatty acids being found in their binding pockets after purification, or resulting in the usually soluble protein being found in insoluble inclusion bodies.[153] To ensure the protein remained in the soluble fraction with a clear binding pocket, an unfolding/refolding step was added.[153] Unfolding was achieved using the denaturants urea and dithiothreitol (DTT) followed by refolding using glutathione (reduced and oxidised) in a rapid dilution. The glutathione reforms the disulphide bonds between cysteine residues, which subsequently form the structurally critical α-helix structures in OBPs.

The final soluble fraction was purified using a nickel-affinity column. The hexa-histidine tag associates with nickel ions bound to the stationary phase of the column and are subsequently displaced with the higher affinity ligand imidazole. Additional purification procedures including size-exclusion chromatography was performed as and when required (Figure 3.6).

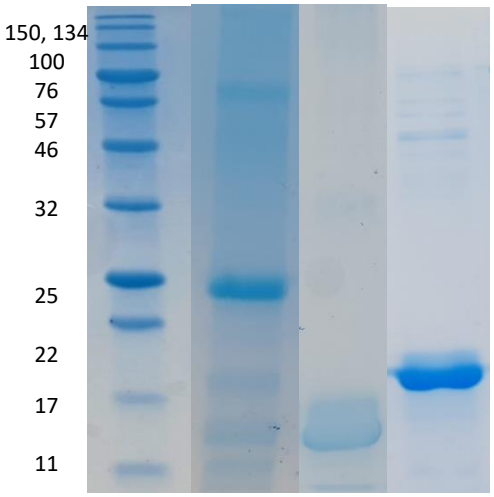


Figure 3.6: SDS-PAGE gels displaying protein fractions after one round of nickel affinity purification. From L-R protein marker,OBP6-His_6, OBP7-His_6 and OBP9-His_6. Molecular weight (kDa) of markers from the top is 150, 134, 100, 76, 57, 46, 32, 25, 22, 17, 11.

Protein yields were calculated based on an average of four purifications (Table 3.3).

Table 3.3: Average yields for large scale purification of *A. pisum* odorant-binding proteins, measured post-purification.

***A. Pisum* Protein**	**Average yield after 3 hours over-expression (mg L^{-1})**
OBP6-His_6	1.72 ± 0.08
OBP7-His_6	1.48 ± 0.05
OBP9-His_6	1.64 ± 0.04

3.2.4 Cleavage of Hexa-Histidine (His_6) Tag

The N-terminus His_6 tag of the refolded protein was removed using a protease cleavage enzyme, either enterokinase or TEV (tomato etch virus protease) depending on the specific cleavage site included. Only proteins of interest for future experiments (Chapters 4-6), OBP6,7 and 9, were purified by removal of the His_6 tag successfully.

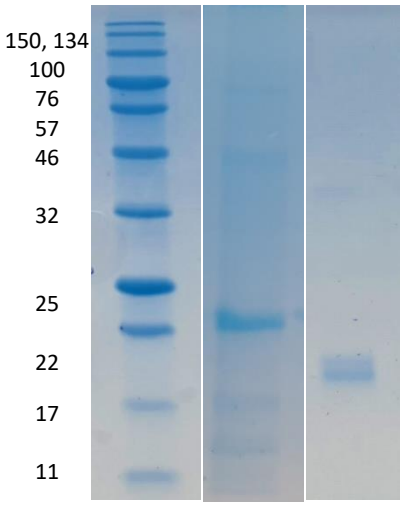


Figure 3.6: Cleavage of the hexa-histidine (His_6) tag of OBP6, as shown by SDS-PAGE gels. From L-R protein marker OBP6-His_6 after nickel-afinity purification and OBP6 after cleavage. Molecular weight (kDa) of markers (M) from the top is 150, 134, 100, 76, 57, 46, 32, 25, 22, 17, 11.

Tag removal of OBP6-His_6 was slow. This resulted in the undesirable cleavage of the protein into three fragments (Figure 3.7) that was confirmed by mass spectrometry (Section 3.2.6). The protein appeared to maintain some native structure on the gel, including a suspected dimer (Figure 3.7, B), however after extensive denaturation with DTT and heat, extra fragments could be seen (Figure 3.7, D and E). This issue with cleavage is likely due to an embedding of the His_6 enterokinase cleavage site into OBP6's 3D structure rendering it inaccessible to the cleavage enzyme.

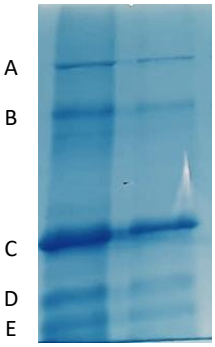


Figure 3.7: Cleaved OBP6, as shown by SDS-PAGE gels. The sample contains an impurity (A), suspected dimerised OBP6 (B), OBP6 (C) and two smaller fragments of peptide (D and E).

To rectify this issue, a thrombin cleavage site was inserted into the plasmid between the enterokinase site and the protein sequence (Figure 3.8). A Q5 Site-Directed Mutagenesis kit was used to insert the thrombin cleavage site. The mutated plasmids were then transformed into BL21(DE3) *E. coli* (as 3.2.1) and sequenced. In total, ten colonies were selected; one had fully successful insertion of the thrombin cleavage site, one had partial mutagenesis and eight had no mutagenesis.

```
OBP6      : MAHHHHHHVGTGSNDDDDKSPDP------AGYDRTWILRQKRGTNDDECRTLLPSPEKKLP
OBP6-SDM  : MAHHHHHHVGTGSNDDDDKSPDPLVPRGSAGYDRTWILRQKRGTNDDECRTLLPSPEKKLP
```

Figure 3.8: The original sequence with His_6 tag and enterokinase cleavage site (OBP6) and the inserted thrombin cleavage site LVPR/GS in the mutated sequence (OBP6-SDM). Only the first 50 residues are shown.

A small-scale expression test was performed of the successfully mutated line, as in 3.2.2.

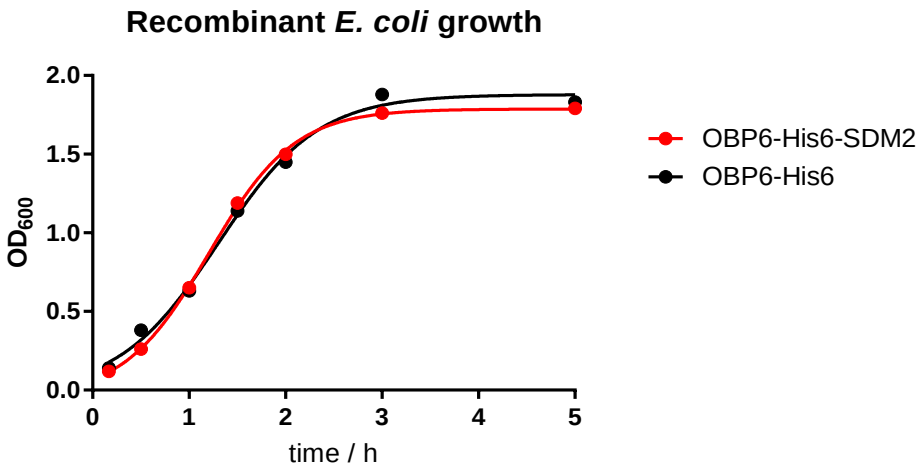


Figure 3.9 Curves showing the growth (by OD_{600}) of generated recombinant *E. coli* strains of the mutated OBP6 (OBP6-His6-SDM2) and non-mutated OBP6 (OBP6-His6) over time.

The mutated cell line reached an OD_{600} of 0.5 after 0.9 hours, later reaching a plateau at approximately 2 hours with an OD_{600} of 1.9. The growth curve for the mutated strain was comparable to that of OBP6 and the other OBPs. The successfully mutated cell line was subsequently grown on a large scale and OBP6-His_6 with the thrombin cleavage site purified. A cleavage with thrombin protease enzyme was performed, resulting in successful production of purified OBP6 (Figure 3.10).

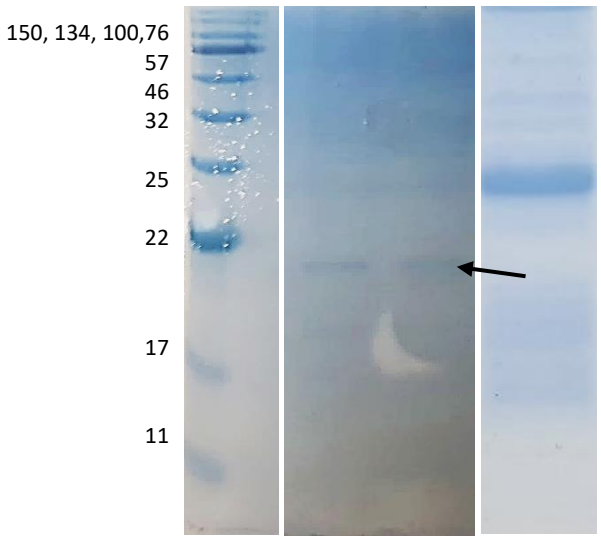


Figure 3.10: Cleavage of the hexa-histidine (His_6) tag of OBP6, as shown by SDS-PAGE gels. From L-R protein marker (standard from other gel for clarity), protein marker, OBP6 after cleavage and nickel affinity purification and OBP-His_6 after nickel-afinity purification. Molecular weight (kDa) of markers (M) from the top is 150, 134, 100, 76, 57, 46, 32, 25, 22, 17, 11. Arrow indicates expected molecular mass of the proteins.

3.2.5 Fast-Protein Liquid Chromatography (FPLC)

To achieve higher levels of purity, OBPs were purified and buffer exchanged into Tris-buffered saline (TBS) buffer using fast-protein liquid chromatography (FPLC) and a Superdex size exclusion column. Proteins were separated by molecular weight and collected using an autosampler (Figure 3.11).

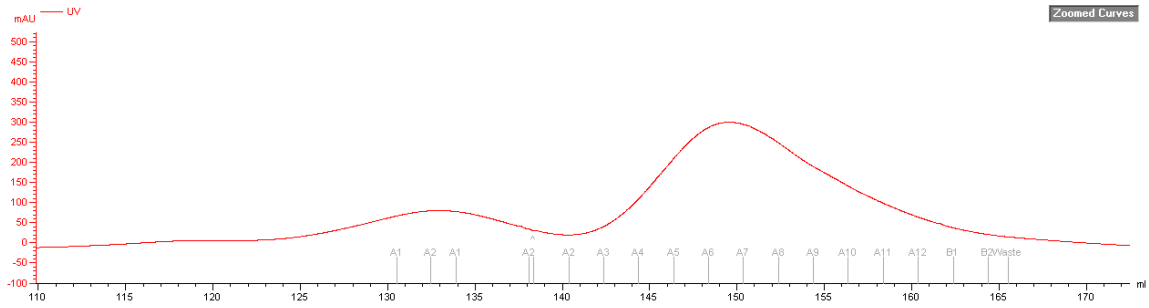


Figure 3.11: Fast-protein liquid-chromatography trace of standard proteins separated by size exclusion chromatography. Two proteins can be seen, α-lactabumin from bovine milk (14 kDa) and carbonic anhydrase from bovine erythrocytes (30 kDa).

3.2.6 Protein Mass Spectrometry

Mass spectrometry techniques can be used to study protein structure and activity. The development of 'soft' ionisation techniques, such as electrospray ionization (ESI) and matrix-assisted laser desorption (MALDI) has led to advances in protein mass spectrometry.

In ESI mass spectrometry (ESI-MS), proteins in a solvent are sprayed as charged droplets of about 10 µm in size, which then rapidly reduce to small solvated macro-ions or nanodroplets (Figure 3.12). The nanodroplets desolvate by ion evaporation resulting in a desolvated macro ion. Finally, the ion travels through a mass spectrometer and is selected and detected[155].

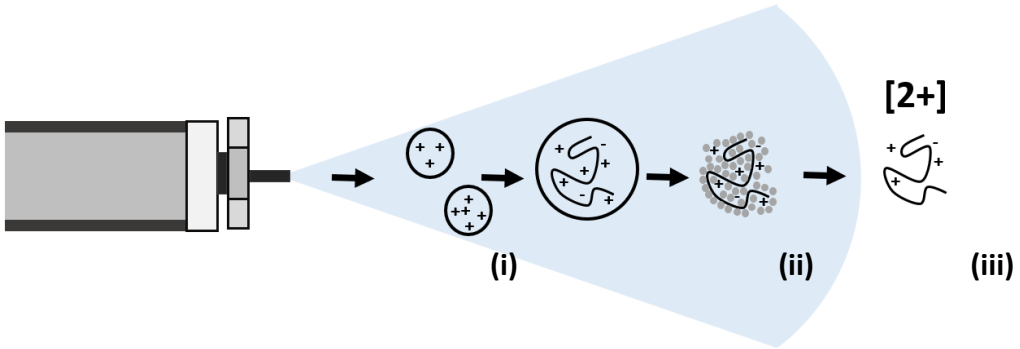


Figure 3.12: Electrospray ionisation (ESI) for ESI mass spectrometry. Proteins are sprayed as charged droplets (i), which then rapidly reduce to solvated macro ions (iii). The solvated macro ion is desolvated (iii) and travels through the mass spectrometer for detection. Image adapted from [155].

Mass spectra can be obtained for entire native proteins and denatured proteins by altering conditions.[155] By using soft ionisation techniques, mass spectra can be obtained for entire native proteins and denatured proteins.[155] Using acidic conditions, the mass spectra of a denatured spectra displaying a variety of positive charge states can be obtained (Figure 3.13). The charge states follow a conserved 'envelope' pattern and can provide data relating protein kinetics and structure. By observing the mass spectra of proteins, we can not only

confirm to synthesis of the correct protein, but also investigate how different conditions may affect a proteins overall structure and stability.

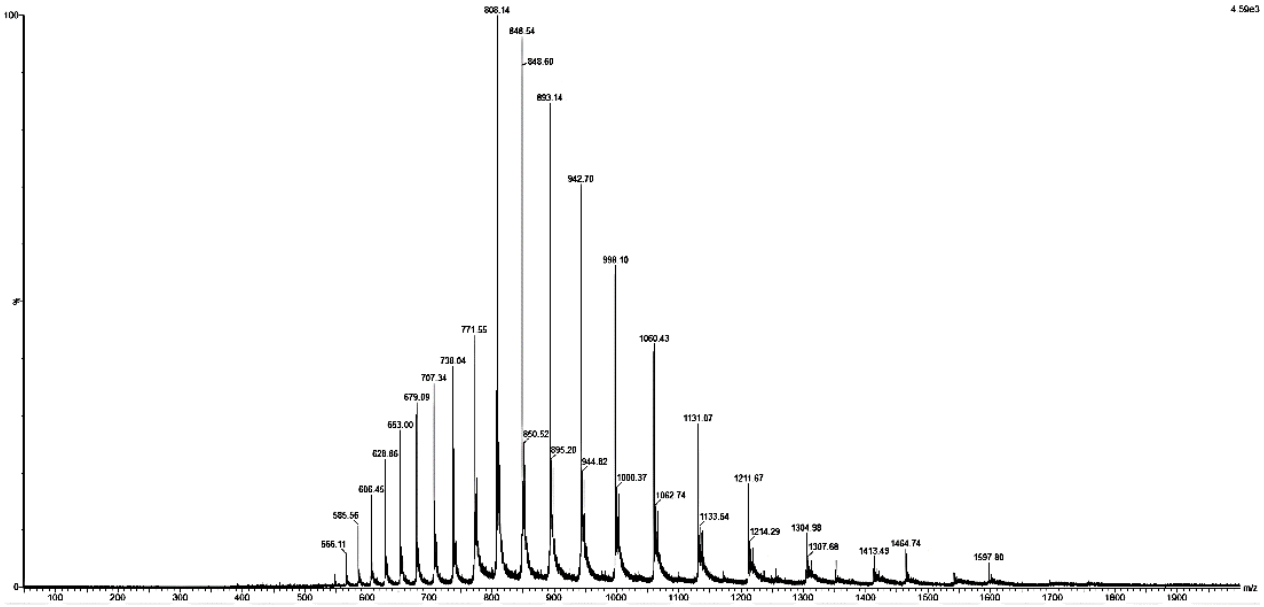


Figure 3.13: The ESI mass-spectra of *apo*-myoglobin showing a variety of charge states.

Mass spectrometry of denatured protein can be used to perform structural and functional assays, however, it is most useful in determining an accurate mass for a protein. To confirm efficient cleavage of the His-tag and determination of the structure, denatured mass spectrometry of the recombinant proteins was performed using a Waters QTof spectrometer. Samples were concentrated using a ZipTip® into an eluent buffer (80% acetonitrile, 0.1% formic acid). Formic acid was added to the sample to decreases the pH resulting in the denaturation of the protein.

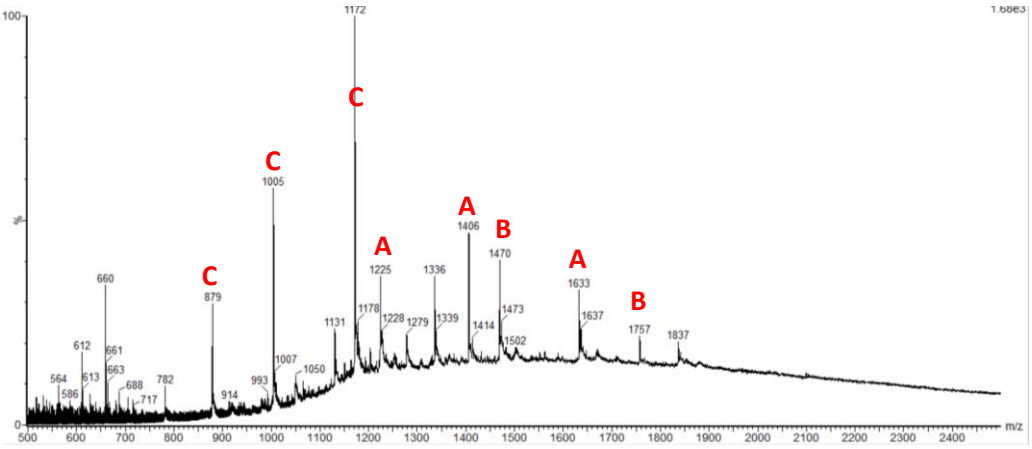
As proteins are very large molecules with many sites for ionisation, a wide variety of charge states are typically observed in electrospray mass spectra. The mass spectra recorded for purified recombinant odorant-binding proteins show an expected charge envelope shape, from which charge states could be identified at the overall mass estimated using ESIProt[156](Table 3.4). The mass spectra of OBP6-His_6, OBP6, OBP7-His_6, OBP7, OBP9-His_6 and OBP9 have been successfully obtained (Figure 3.14).

Table 3.4: Mass spectrometry data for purified *A. pisum* OBPs.

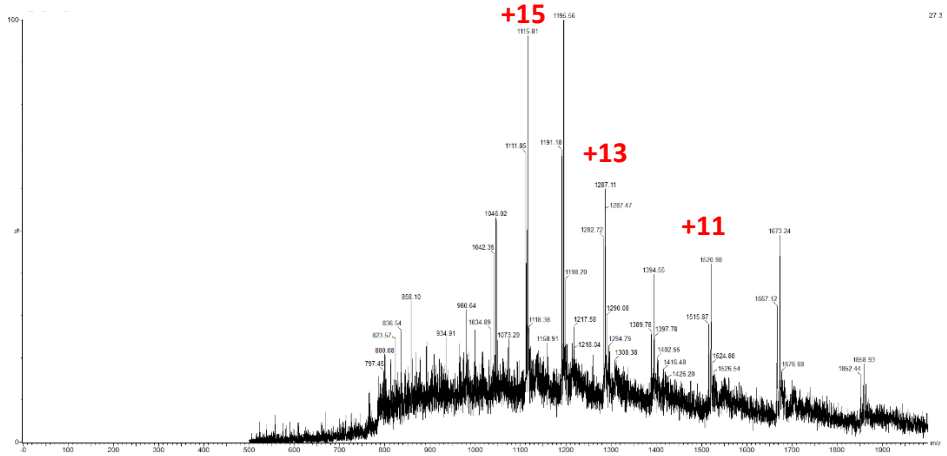
* the average Mr was calculated using the ProtParam website.[157]

***A. Pisum* Protein**	**Predicted Mass (average Mr*)**	**Observed Mass**
OBP6-His_6 (with enterokinase site)	25299.02	25161.47 ± 4.56
OBP6 (with enterokinase site)	22772.42	Fragment A: 14687.91 ± 2.52 Fragment B: 9814.28 ± 21.54 Fragment C: 7025.95 ± 2.00
OBP7-His_6	16798.38	16721.27 ± 3.63
OBP7	14254.79	14118.14 ± 2.76
OBP9-His_6	18421.05	18263.57 ± 0.18
OBP9	15895.48	16262.91 ± 3.67

(i)



(ii)



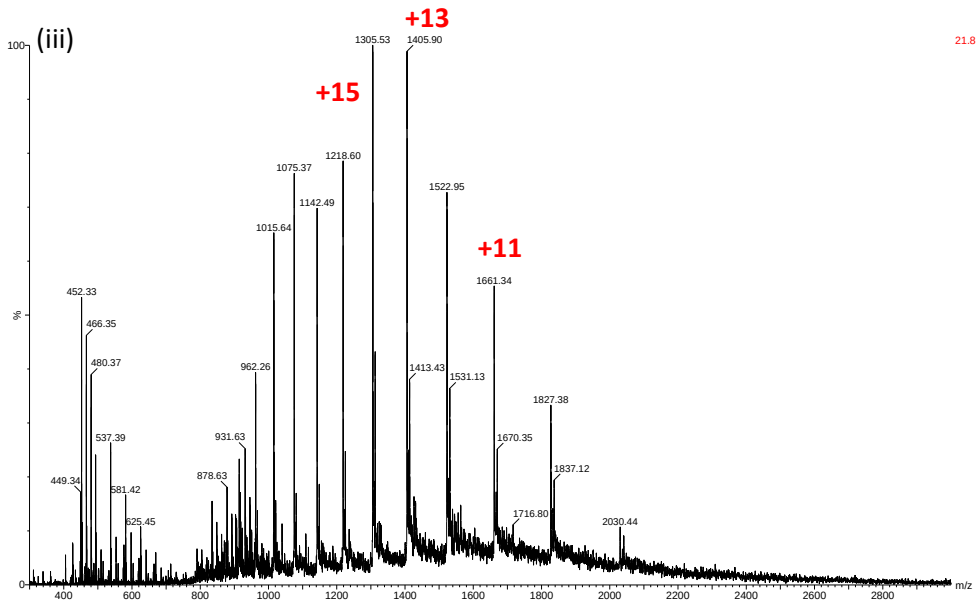


Figure 3.14: The ESI-mass-spectra of denatured *A. pisum* (i) OBP6, fragmented, (ii) OBP7 and (iii) OBP9.

As previously discussed, the initial construct of OBP6 was over-cleaved resulting in multiple fragments being observed on the SDS-PAGE gels (Figure 3.7) and in the denatured mass spectra (Figure 3.10). The enterokinase cleavage site occurs at DDDK/S (between the lysine and serine). From the size of the fragments, and knowledge of the cleavage site, we can assume that OBP6 was cleaved into 3 fragments of OBP6 and the corresponding His_6 site. There are multiple similar cleavage sites within OBP6, and these were used to predict the sequence of the fragments (Table 3.5).

Table 3.5: Predicted sequences of the over-cleaved sequences.

* the average Mr was calculated using the ProtParam website.[157]

Fragment	Predicted Sequence	Predicted Mass*	Found Mass (Table 3.4)
His_6 tag	MAHHHHHHVG TGSNDDDDK	2147.19	NA
A	PETKEYKEMA HGKEPPCLFQ CIFMQSGLTT SDGKLNEDAI TKKMSEGINN DEKWKSIWQN SLNKCFDDVK QEDKKQILIM NTPAGRLMKC FLRDMYMSCP KNVWVESSEC LSMKDLVQKC PEMPPPVFK	14933.39	14687.91 ± 2.52
B	MAHHHHHHVG TGSNDDDDKS PDPAGYDRTW ILRQKRGTND DECRTLLPSP EKKLPSCCQM PNILPNLDST WEKCFETFKQ FKD	9619.67	9814.28 ± 21.54
C	KQILIMNTPA GRLMKCFLRD MYMSCPKNVW VESSECLSMK DLVQKCPEMP PPVFKSPPKL I	7029.58	7025.95 ± 2.00

For OBP7, OBP9 and OBP6, experimentally observed masses in accordance with predicted values. Some disparity was seen between the predicted mass and the calculated mass. This may be due to small protein modifications or the presence of adducts. For OBP7 and OBP9 (Figure 3.14(ii) and (iii)), two species can be observed in the mass spectra. The second species is around 58 Da smaller for OBP7 and 98 Da larger for OBP9, due to unidentified adducts or modifications. The predicted mass of the protein was generated using ProtParam, which gives the average mass for the sequence provided; this may be divergent from the actual mass of the protein.[157,158] Furthermore, adducts may form with solvents, salts or other molecules that are introduced during the process, or artefacts from protein purification. Unfortunately, these discrepancies in mass could not be fully explained experimentally, though small inaccuracies are expected and the proteins were used in following experiments.

3.3 CONCLUSIONS

In conclusion, *A. pisum* odorant-binding proteins OBP1,2,3,5,6,7,8,9 and 10 were successfully over expressed and purified from *E. coli* expression cell lines. Each protein was expressed with a His_6 tag for purification ease and OBP6,7 and 9 were produced on a large scale. OBP6 required the insertion of a thrombin cleavage site due to over-cleaving by enterokinase. Future work may include optimisation of protein over-expression and purification and optimisation of transformation of *E. coli* with plasmids containing genes for OBP11 to OBP15 and OBP4 (with encoded His_6 tags). Analysis with gel electrophoresis and mass spectrometry showed successful production and cleavage of the His_6-tag of each of the proteins.

Chapter 4: *In Silico* Structural and Ligand-Docking Studies of Aphid Olfactory Proteins

4.1 INTRODUCTION

Molecular modelling techniques were first developed in the 1970s in order to speed up drug discovery; the structure of target proteins can either be predicted by *in silico* by homology modelling or solved *via* range of structural studies, and their interactions with various ligands observed. Computer aided drug design (CADD) is now a vast field of research, and many pharmaceutical drug discoveries have involved heavy use of computational methods.[159–161]

Proteins can be modelled using only their amino acid sequence *via* a process known as homology modelling. Homology modelling utilises available 3-dimensional (3D) protein structures, and finds similar or matching motifs, generating a model of a desired protein.[160] Further computational analysis can then be undertaken, such as the minimisation of amino acid geometries, where energy minimisations are made to remove sterically unfavourable interactions and generate the most favourable conformation of the molecule.[162] The energetic favourability of a generated model is typically checked by a Ramachandran plot – a plot of the torsional angles - phi (ϕ) and psi (ψ) of amino acids in the protein. A Ramachandran plot can also be used to assess different conformations of a protein.[162,163]

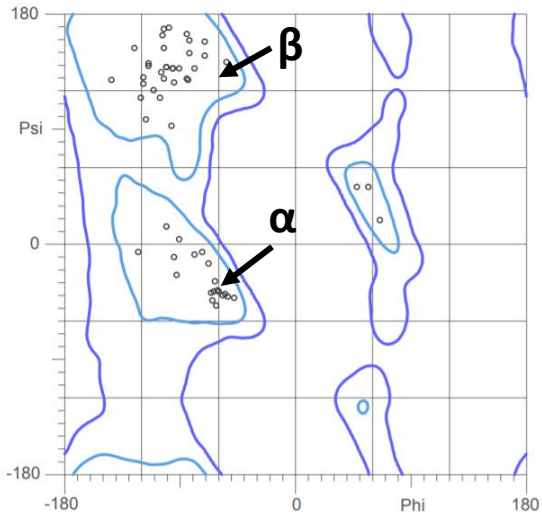


Figure 4.1: An example Ramachandran plot for human ubiquitin (PDB code 1ubq). Each point represents a residue, with the axes displaying the angles of the residues. Favourable combinations of psi and phi fall within the displayed zones (blue). Regions for alpha helices (α) and beta pleated sheets (β) are labelled.

Computational methods are also used to assess protein-ligand interactions. When a protein binds to a ligand favourably, a lowering in energy will be observed. *In silico* modelling studies sample ligands and proteins for the most favourable interactions using stochastic methods (algorithms), determining energetic data with scoring functions.[159]

Docking can be rigid, semi-flexible or flexible. Rigid docking is the simplest form of docking model, where no flexibility of the protein or ligand molecule is included, and is comparable to the 'lock and key' model of protein-ligand interactions.[159] Semi-flexible is most commonly used; the ligand is considered as flexible and the protein target rigid. Both rigid and semi-flexible methods neglect the conformational changes that a protein can undergo and assume that the fixed conformation provided to the model corresponds to the conformation responsible for binding the protein. This assumption is not always true, and a flexible docking approach may be used to provide a more realistic model[159]. Flexible docking assumes both ligand and protein are fully-flexible objects – unfortunately this results in large numbers of potential interactions, and often a substantial amount of computing power is required to perform these calculations.[159]

To determine the most favourable interactions, stochastic methods – algorithms that change randomly, not systematically – are used. This includes Tabu searching (searching based on previously sampled zones), swarm optimisation, and more commonly evolutionary algorithms and the Metropolis Monte Carlo algorithm. The Monte Carlo algorithm is an example of a more randomised algorithm, with random changes being assessed in energetic favourability at each stage.[159] Genetic algorithms (GA), including the Lamarckian algorithm, use evolutionary theory and ideas based in genetics.[164] In genetic algorithms protein-ligand interactions are defined by a set of variables which are assigned to a gene, and the associated fitness of that 'gene' is related to the total energy of the protein-ligand interaction.[164] As with traditional genetics, evolution of the variables is allowed to occur, with 'mutations' – random large changes to the variables – and crossovers also occurring.[164] The Lamarckian GA uses the specific principles of Lamarckian genetics to determine the most favourable protein-ligand interaction.[164]

There are a variety of scoring functions available to determine energetic values for ligand-docking studies. Force field-based functions are commonly used, where the approximate energy of the interaction is combined with both intra- and inter- molecular components. Other scoring functions include empirical, where various empirical energies such as van der Waals forces, hydrogen bonds and desolvation energy are summed, and knowledge based,

which uses a database of previously described contacts to explore more statistically favourable interactions.[159]

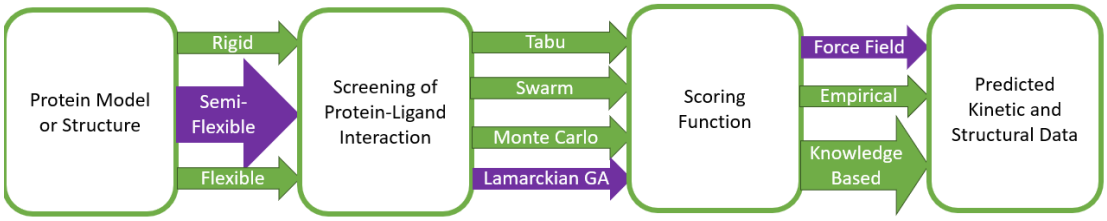


Figure 4.2: Process of ligand docking. The purple arrows represent the methods chosen for docking insect odorant-binding proteins and odorant receptors in this chapter.

For the modelling of insect odorant-binding proteins (OBPs), the generation of homology models is a viable method. Many protein structures for insect OBPs are already published, and OBPs are known for their high structural conservation. For OBP-ligand docking, use of AutoDock allows a semi-flexible sample approach, with a force field scoring function. A variety of algorithms can be used in AutoDock, including Lamarckian GA. The main drawback of this type of ligand-protein interaction assessment is the semi-flexible docking approach, where some OBPs have been thought to show conformational changes under different conditions or when ligands bind.[117]

Previously, modelling of insect odorant receptors (ORs) was limited to advance computational techniques, specifically observing amino acid evolutionary couplings.[165] ORs are a challenging protein to express and purify, therefore, only recent advancement in techniques have led to structural determination.[165] In 2018, Butterwick *et al.* published the first insect OR structure, of the odorant receptor co-receptor (ORCO) of *Apocrypta bakeri.*[88] The publishing of this structure allows for comparative modelling of ORs to be performed and broadens the scope for molecular modelling studies of insect ORs.[166]

For more complex and realistic modelling, molecular dynamics (MD) simulations can be performed. MD simulations provide detailed statistical analysis of computational models and minimise energy values.[159,167] The simulation solves Newton's classical mechanical equations for a system of numerous atoms whilst monitoring physical properties, including temperature and pressure, until an equilibrium state is eventually reached.[167–169] Programs such as GROMACS can be used to minimise model structure, model realistic environments including lipid bilayer (membrane) embedding, additions of solvents and neutralisation with ions. The stability of protein structures and their behaviours over time can be calculated and observed.[168,169]

Molecular dynamics simulations have limitations. Firstly, they obey classical Newtonian mechanics. Though classical mechanics are generally correct for atoms at average temperatures, some atoms, such as hydrogen, can exhibit quantum behaviour.[167–169] Furthermore, MD simulations do not consider electronic motions, and assume all electrons are in the ground state.[167–169] The force fields that are used in the process are also approximate, and long-range interactions cannot be calculated.[167–169]

The overall aims of this chapter are to generate homology models of aphid olfactory proteins, including OBPs and ORs. These will then be screened using docking analysis and the interactions assessed for their stability and likeliness. Simple semi-flexible ligand docking methodologies will be used. Successfully generated binding pairs will be analysed further and provide framework for future *in vitro* assays.

4.2 RESULTS & DISCUSSION

4.1.1 Sequence Alignment of Aphid OBPs

Aphid OBPs are phylogenetically similar to other OBPs and are classified based on conservation. For example, OBP6 of the pea aphid, *Acyrthosiphon pisum*, is more closely related to OBP6 of the peach-potato aphid, *Myzus persicae*, than OBP7 of *A. pisum*. This is true for all *A. pisum* OBPs, with OBP3 and OBP11 being the most divergent proteins (Figure 4.3).[103]

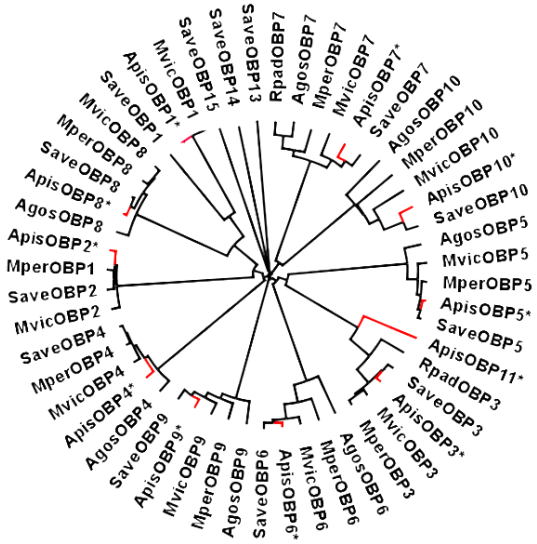


Figure 4.3: Phylogenetic analysis of odorant binding proteins (OBPs) from *Acyrthosiphon pisum*, *Myzus persicae, Sitobian avenae, Aphis gossypi* and *Megoura vicae.* OBPs from *A. pisum* are annotated with an * and a red line. Generated in FigTree v1.4.3 from multiple sequence alignment.

Each individual OBP is highly conserved across species; *A. pisum* OBP6 is 79-96% conserved across the five aphid species for which it has been characterised (Figure 4.4). This high sequence conservation allows functional comparisons to be made between the proteins.

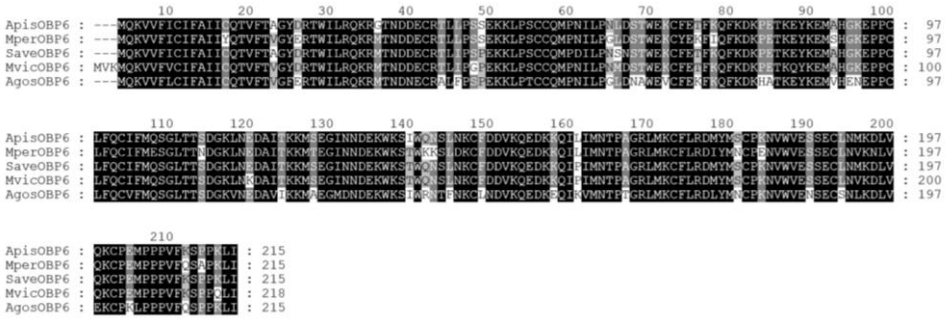


Figure 4.4: OBP6 conservation across five characterised aphid species.

4.2.2 Homology Modelling of *A. pisum* OBPs

3-Dimensional protein models of *A. pisum* OBPs were generated using iTASSER[141], which utilities a homology modelling technique by looking for structural motifs found in previously crystallised proteins, then visualised in PyMol.[91] Generated models were initially minimised using the Yasara minimisation server[170] and checked for energetic favourability using Ramachandran plots in PROCHECK[171], discrete optimised protein energy (DOPE) score in Python 3.2 with MODELLER 3.2[172] and Z-scores by ProSA.[173]

Table 4.1: Ramachandran data, discrete optimised protein energy (DOPE) scores and Z-scores from generated homology models of *A. pisum* odorant binding proteins.

	DOPE score	**Ramachandran Data**			
***A. pisum* Protein Model**		**% Residues Favoured**	**% Residues Allowed**	**% Residues Disallowed**	**Z-Score**
OBP1	-17409.66	81.2	14.7	4.0	-6.31
OBP2	-20577.94	70.1	27.5	2.5	-4.80
OBP3	-16512.96	89.9	9.3	0.8	-5.57
OBP4	-17184.65	82.8	14.6	2.5	-5.73
OBP5	-18264.56	80.1	18.3	1.7	-5.54
OBP6	-18156.41	84.3	13.5	2.1	-7.35
OBP7	-12823.21	90.9	9.1	0	-6.34
OBP8	-13678.49	85.8	11.8	2.4	-5.92
OBP9	-16153.08	89.2	10.0	0.8	-6.85
OBP10	-16104.70	78.7	18.8	2.4	-6.89
OBP11	-16330.00	88.4	10.9	0.8	-5.04
Vetch Aphid OBP3 (4Z39)	-15609.98	96.4	3.1	0.4	-5.63

Homology models were successfully generated for OBP1-11 (Figure 4.5). Overall, analysis of models was favourable, with DOPE scores and Z-scores comparable to published aphid OBP crystal structures, and Ramachandran analyses with few disallowed residues.

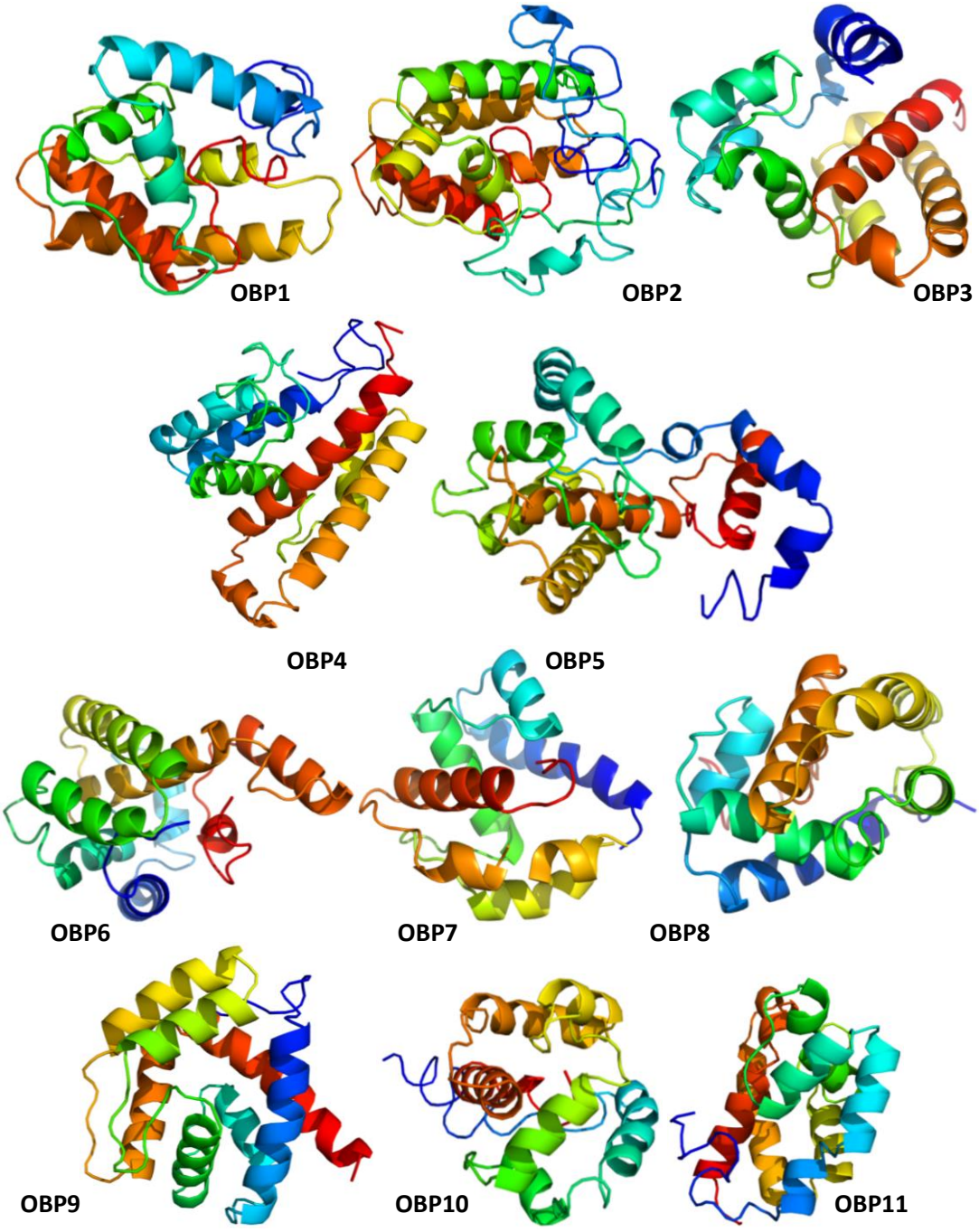


Figure 4.5: Homology models of *A. pisum* odorant binding proteins (OBP1-OBP11)

Initial assessment was made of the homology models – most had a clearly defined globular pocket which may provide space for binding. Crystal structures of OBP3 and OBP7 from the blackcurrant-lettuce aphid, *Nasonovia ribis-nigri,* have been previously published and their binding activity with the aphid alarm pheromone, (*E*)-β-farnesene **17** (EBF), described.[104] The

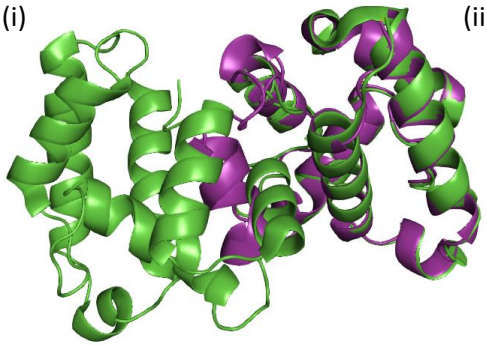
homology models of OBP3 and OBP7 were highly conserved against previously published crystal structures of the two OBPs, and OBP3 displays a clear binding pocket that has a long tubular shape (Figure 4.6). This shape of binding pocket matches the shape of EBF **17**, and could provide a perfectly shaped binding pocket.

(ii)

```
                10        20        30        40        50        60        70        80        90        100
ApisOBP3 : MISSTFYITLVFGIAMLISCGHGRFTTEQIDYYGKACNASEDDLVVVKSYKVPTTETGKCLMKCMITKLGLLNDDGSYNKTGMEAGLKKYWSEWSTEKIE : 100
NribOBP3 : ----------------------ASMRFTTEQIDYYGKACNASEDDLVVVKSYKVPSTETGKCLMKCMITKLGLLNDDGSYNKTGMEAGLKKYWSEWSTEKIE :  80

                110       120       130       140
ApisOBP3 : SINNKCYEEALLVSKEVVATCNYSYTVMACLNKQLDLDKST : 141
NribOBP3 : TINNKCYEEALLVSKEVVATCNYSYTVMACLNKQLDLDKST : 121
```

(i)



(iii)

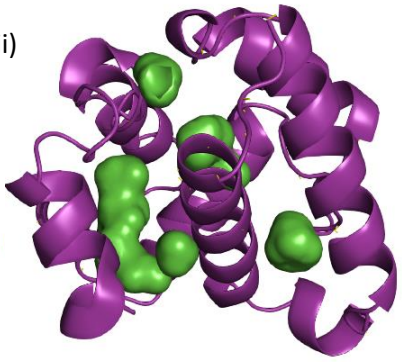


Figure 4.6: (i) Sequence alignment of A. pisum OBP3 with OBP3 from blackcurrant-lettuce aphid, *Nasonovia ribis-nigri* (ii) The homology model of *A. pisum* OBP3 (purple) aligned to OBP3 from *N. ribis-nigri* crystal structure (green, PDB identifier 4Z39). (iii) The structure of *A. pisum* OBP3 (purple) with the binding pockets displayed (green).

Other homology models provide clues about the activity of a protein. Firstly, OBP9 had a binding pocket, which could provide space for a variety of ligands to bind. OBP6 showed several binding pockets, in addition to similar structural motifs to OBPs from other insect species (Figure 4.7). For example, the Drosophila pheromone-binding OBP 'LUSH' was seen to have a structurally crucial salt bridge between Lys87 and Asp118.[76,115] When disrupted, the subsequent protein lost functionality.[76] This salt-bridge has also been found in other insect binding proteins, including pheromone binding proteins (PBPs) from *Bombyx mori.*[115] In OBP6, Lys144 and Asp174 could present a similar salt-bridge candidate; the homology model predicts them to be closely orientated.

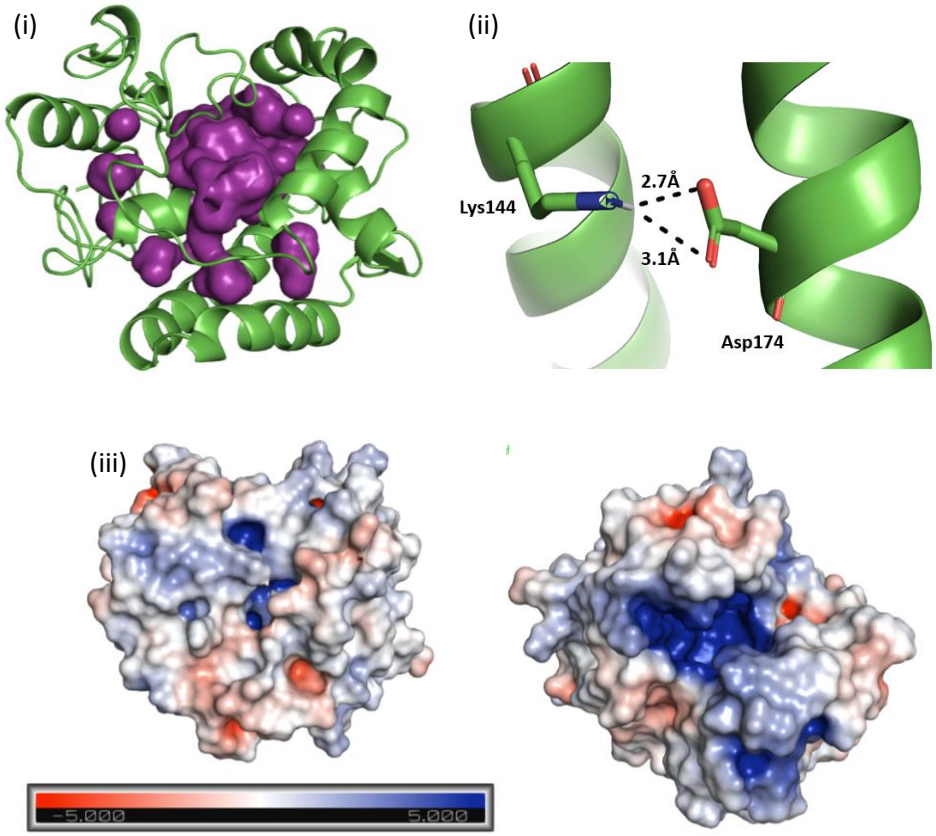


Figure 4.7: (i) Predicted structure of OBP6 (green) with potential binding pockets, visualised using PyMol's suface and cavity functions, shown in purple; (ii) Potential salt bridge between Lys144 and Asp174; (iii) Predicted structure of OBP6 with the Adaptive Poisson-Boltzmann Solver (APBS), calculated using the APBS PyMol plugin, electrostatic map displayed. Two potential binding pockets can be clearly seen in blue.

OBP6 is of paticular interest due to it's high expression levels in aphid antennae. The prescence of the conserved salt-bridge suggests it may play a role in pheromone binding.

4.2.3 Ligand-docking Studies of *A. pisum* OBPs using AutoDock

To identify potential protein-ligand interactions, homology models were screened in AutoDock for their interactions with specific semiochemicals. All modelled *A. pisum* OBPs (OBP1-11) were screened, with a wide range of aphid semiochemicals and allelochemicals being chosen. OBP3 and 7 are known to bind the alarm pheromone **17**, (EBF), but no other OBP-ligand interactions have been described in aphids as yet.[102,125,140] OBPs 6-9 showed the most promising results (Table 4.2), whereas OBPs 1-5 and 10-11 showed very few successful low energy binding interactions (Appendix Table A.1, Table A.2).

Table 4.2: Results from Autodock ligand-screening. The aphid sex pheromone components and respective enantiomers were tested, along with the aphid alarm pheromone. The values highlighted in **bold** represent the lowest energy interaction for that ligand. Results displayed to two significant figures.

NA = no favourable docking conformations were found in the screening

Ligand	***A. Pisum* OBP**							
	OBP6		**OBP7**		**OBP8**		**OBP9**	
	Binding energy (kcal mol^{-1})	**K_i (µM)**	**Binding energy (kcal mol^{-1})**	**K_i (µM)**	**Binding energy (kcal mol^{-1})**	**K_i (µM)**	**Binding energy (kcal mol^{-1})**	**K_i (µM)**
(1*R*,4a*S*,7*S*,7a*R*)-nepetalactol **5**	**-7.7**	**2.4**	-6.0	41	-5.9	48	-5.3	130
(4a*S*,7*S*,7a*R*)-nepetalactone **6**	**-7.5**	**3.1**	-6.0	37	-6.2	31	-5.8	57
(1*S*,4a*R*,7*R*,7a*S*)-nepetalactol **32**	**-7.7**	**2.3**	-6.00	40	-6.4	21	-5.5	97
(4a*R*,7*R*,7a*S*)-nepetalactone **33**	**-7.6**	**2.7**	-6.0	42	-6.3	23	-5.7	66
(*E*)-β-farnesene **17**	-6.7	12	**-6.8**	**11**	-6.0	39	-5.1	190
(*S*)-germacrene D **25**	**-3.1**	**5.1**	-6.7	13	-7.3	4.3	-6.1	36
(1*R*,4*E*,9*S*)-Caryophyllene **26**	-6.7	12	NA	NA	**-7.5**	**3.4**	-6.4	20
Myrcene	**-6.3**	**22**	NA	NA	-4.7	360	-4.4	580
(*E*)-Ocimene **28**	**-6.4**	**19**	NA	NA	-5.0	230	-4.5	490
(4*R*)-linalool	**-6.6**	**14**	NA	NA	-4.8	290	-4.6	410
(4*S*)-linalool	**-6. 6**	**16**	NA	NA	-4.8	290	-4.7	340
(*Z*)-jasmone **30**	**-7.7**	**2.3**	NA	NA	-6.0	41	**-7.7**	**2.3**

Molecular docking tools generate predicted binding energy and K_i. A lower binding energy and K_i is indicative of favourable binding activity. OBP6 shows potential low-energy interactions with sex pheromone components, both naturally occurring enantiomers and non-naturally occurring enantiomers (Figure 4.8). The molecular docking of OBP6 with both enantiomers of (1*R*,4a*S*,7*S*,7a*R*)-nepetalactol **5** and (1*R*,4a*S*,7*S*,7a*R*)-nepetalactone **6** gives a significantly stronger binding complex than with EBF **17**; an average (-7.6 kcal mol^{-1} versus -

6.7 kcal mol^{-1}). This is also reflected in the K_i values, the lowest K_i for sex pheromone being 2.3 µM and the K_i for EBF **17** predicted at a much higher 12µM.

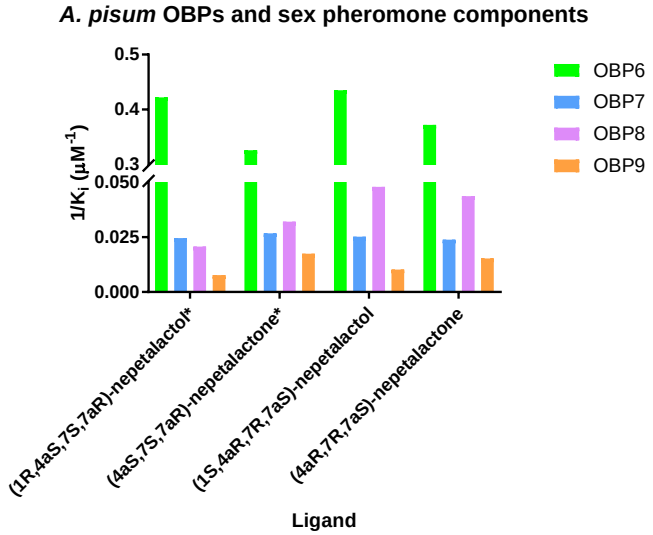


Figure 4.8: Predicted binding interactions (shown as 1/K_i) of the aphid sex pheromone components with OBPs 6-9. The natural enantiomers are indicated with a *.

Other protein-ligand interactions were observed. OBP6 interacted strongly with ligands other than the sex pheromone components, including a 2.3 µM interaction with allelochemical (*Z*)-jasmone **30**. OBP7 showed a stronger binding complex with **17**, than other ligands; this has been previously demonstrated *in vitro* with OBPs from *M. viciae*.[140] OBP8 showed a strong binding interaction with allelochemical (1*R*,4*E*,9*S*)-caryophyllene **26**. All interactions merely serve as predictions and do not necessarily correspond to biological activity; however, this data provides a route to targeted *in vitro* assays.

As the main focus of this project is to investigate sex pheromone-OBP interactions, OBP6 is a promising candidate for further studies. OBP6 is highly expressed in the antennae, specifically the type II trichoid sensilla of winged adult aphids, and has previously been hypothesised to play a role in sex pheromone perception.[75,103]

4.2.4 Binding Site Assignments of *A. pisum* OBPs

Multiple protein-ligand interactions were identified by docking homology models with AutoDock. The most significant finding was the hypothesised interaction between OBP6 and sex pheromone components. OBP6 and sex pheromone components [(1*R*,4a*S*,7*S*,7a*R*)-nepetalactol **5** and (4a*S*,7*S*,7a*R*)-nepetalactone **6**] interacted with lower predicted binding energy (-7.7 and -7.5 kcal mol^{-1} respectively) and significant lower K_i values (2.4 and 3.1 uM) than all other OBP-ligand interactions (Figure 4.9).

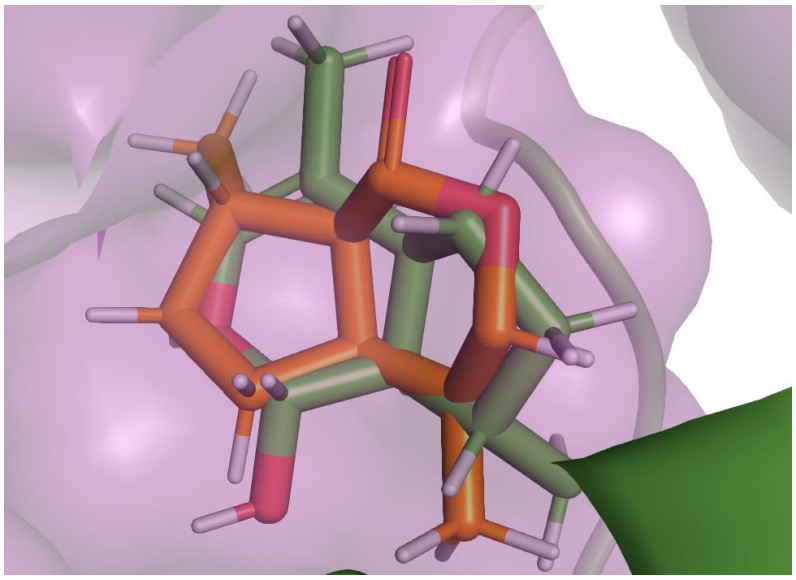


Figure 4.9: (1*R*,4a*S*,7*S*,7a*R*)-nepetalactol **5** (green) and (4a*S*,7*S*,7a*R*)-nepetalactone **6** (orange) in the binding pocket of OBP6 (purple).

Surprisingly, the non-naturally occurring enantiomers of the sex pheromone components, known to be behaviourally and electrophysiologically inactive, were predicted to bind with similar energy as the naturally occurring, biologically active enantiomers.[145] To further investigate this interaction and unexpected result, predicted docking conformations were analysed for potential contacts, including hydrogen bonds. Two major interactions, a hydrogen bond between the back-bone amine of Phe208 and the oxygen in the alcohol of (1*R*,4a*S*,7*S*,7a*R*)-nepetalactol **5**, in addition to a hydrogen bond between the phenol oxygen of Tyr176 and the alcohol hydrogen of (1*R*,4a*S*,7*S*,7a*R*)-nepetalactol **5**, were identified (Figure 4.10). The interaction with Phe208 could be critical in the binding of **5** to OBP6. Despite similar binding energies and binding constants being predicted, the same interaction was not observed with the enantiomer (1*S*,4a*R*,7*R*,7a*S*)-nepetalactol **32** (Figure 4.10). The biologically inactive enantiomer bound in an almost reverse conformation to the active enantiomer, and a unique hydrogen bond between the hydroxy group of the lactol and the amine of Lys169 could be observed (Figure 4.10). The distance of 2.7Å is fairly large for a hydrogen bond.

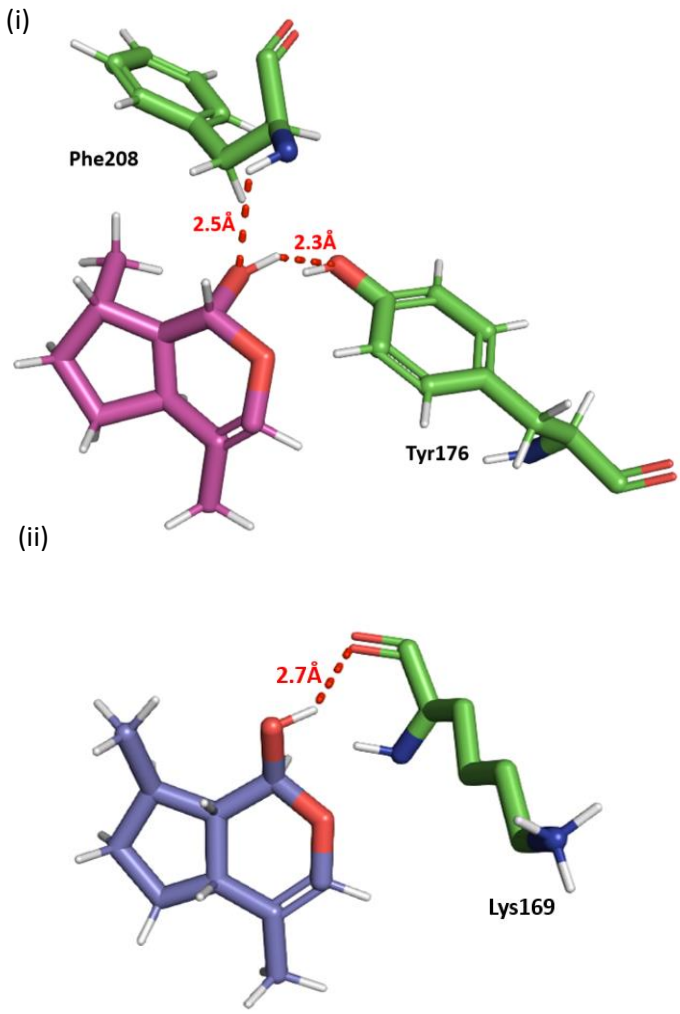


Figure 4.10: (i) OBP6 (green) interactions with the enantiomers of the sex pheromone. Natrually occurring (1*R*,4a*S*,7*S*,7a*R*)-nepetalactol **5** (pink) interacting with Tyr176 and Phe208 (red, dashed); (ii) Non-naturally occurring (1*S*,4a*R*,7*R*,7a*S*)-nepetalactol **32** (blue) with a potential interaction with Lys169 (red, dashed). Distance of interactions are in Å.

These results suggest that Tyr176 or Phe208 may be involved in OBP6–(1*R*,4a*S*,7*S*,7a*R*)-nepetalactol **5** interactions, and that although both enantiomers bind with similar energies, they may bind in differing conformations. To confirm the predicted interacting residues, the predicted binding site of OBP6 was mutated. One mutation of Tyr176 and two mutations of Phe208 were performed, generating three mutant homology models – OBP6-Phe176, OBP6-Ile208 and OBP6-Tyr208 (Table 4.3; Figure 4.11). Initially, tyrosine was mutated to phenylalanine to retain a similar structure within the binding pocket but remove the hydrogen bonding phenol. Phenylalanine was mutated to the vastly different, small, non-aromatic residue isoleucine. Small, but not significant, differences in the 3-dimensional structure were observed between the original protein and each of the mutant forms. The

second mutation involved changing the phenylalanine to a tyrosine. Tyrosine has more comparable properties to phenylalanine, and again only minor differences in the overall structure were observed. A mutation of Lys to Leu169 was also performed to investigate the potential interaction between the non-naturally occurring nepetalactol enantiomer and OBP6. Leucine is a non-polar residue of a similar size to lysine.

Table 4.3: Ramachandran data, discrete optimised protein energy (DOPE) scores and Z-scores from generated homology models of mutated *A. pisum* odorant binding protein 6.

A. Pisum **Protein Model**	**DOPE score**	**Ramachandran Data**			**Z-Score**
		% Residues Favoured	**% Residues Allowed**	**% Residues Disallowed**	
OBP6 *wild type*	-18156.41	84.3	13.5	2.1	-7.35
OBP6-Phe176	-17947.35	85.0	12.8	2.1	-7.33
OBP6-Ile208	-17833.52	86.4	10.0	3.6	-7.01
OBP6-Tyr208	-18299.81	87.1	11.4	1.4	-7.18
OBP6-Leu169	-17908.81	81.4	16.4	2.1	-6.80

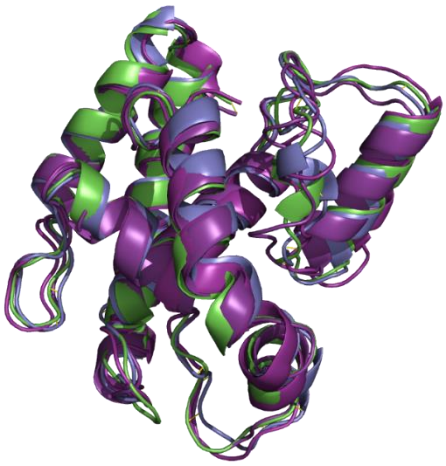


Figure 4.11: Homology model predicted structure of the mutants (varying shades of purple) aligned with wild type OBP6 (green)

Analogues of the sex pheromone components were also evaluated. Small modifications to the overall structure of nepetalactone and lactol were made to determine which structural properties were important for the activity of the ligand (Figure 4.12).

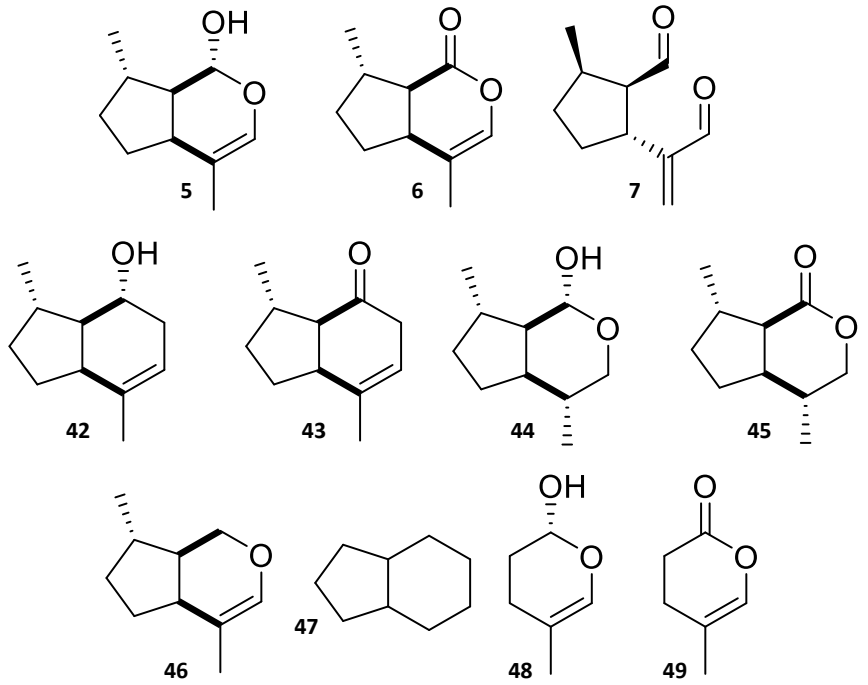


Figure 4.12: (1*R*,4a*S*,7*S*,7a*R*)-Nepetalactol **5** and (4a*S*,7*S*,7a*R*)-nepetalactone **6**, in addition to minor sex pheromone component (1*S*,2*R*,3*S*)-dolichodial **7**, with analogues **42**-**49**.

The mutated models, in addition to *wild type* OBP6, were then screened against sex pheromone components and the generated library of analogues for binding activity (Table 4.4). For mutant OBP6-Phe176, no change was seen in docking activity.

Table 4.4: Results from Autodock ligand-screening. Analogues of the aphid sex pheromone components and analogues were screened with OBP6 wild type and mutants. The values highlighted in **bold** represent the lowest energy interaction for that ligand. Data is shown to two significant figures.

Ligand	OBP6		OBP6-Phe176		OBP6-Ile208		OBP6-Tyr208	
	Binding energy (kcal mol^{-1})	**K_i (μM)**	**Binding energy (kcal mol^{-1})**	**K_i (μM)**	**Binding energy (kcal mol^{-1})**	**K_i (μM)**	**Binding energy (kcal mol^{-1})**	**K_i (μM)**
(1*R*,4a*S*,7*S*,7a*R*)-nepetalactol **5**	**-7.7**	**2.4**	-7.5	3.3	-5.4	110	-5.1	190
(4a*S*,7*S*,7a*R*)-nepetalactone **6**	**-7.5**	**3.0**	-7.3	4.5	-6.3	25	-5.6	78
(1*S*,4a*R*,7*R*,7a*S*)-nepetalactol **32**	**-7.7**	**2.3**	-7.5	3.4	-6.3	40	-4.7	360
(4a*R*,7*R*,7a*S*)-nepetalactone **33**	**-7.6**	**2.7**	-7.5	2.7	6.0	37	-5.3	120
(1*S*,2*R*,3*S*)-dolichodial **7**	**-7.0**	**7.4**	-7.1	6.8	-5.3	65	-4.8	310
Analogue **42**	**-8.8**	**0.38**	-7.8	2.0	-6.1	33	-5.5	92
Analogue **43**	**-7.8**	**2.1**	-7.4	3.5	-6.1	35	-5.6	86
Analogue **44**	**-7.4**	**3.6**	-7.4	3.7	-6.0	39	-4.9	240
Analogue **45**	**-8.0**	**1.4**	-7.9	1.6	-5.7	65	-5.0	220
Analogue **46**	-6.9	8.4	**-7.2**	**5.7**	-5.6	76	-6.1	34
Analogue **47**	**-7.9**	**1.5**	-7.9	2.6	-6.2	28	-6.4	32
Analogue **48**	**-5.7**	**69**	-5.4	95	-4.2	780	-4.2	840
Analogue **49**	**-5.6**	**76**	-5.3	100	-4.4	610	-4.2	850

By mutating Tyr176 to Phe176, thereby removing the phenol which nepetalactol was hypothesised to interact with, no change was seen in the binding activity of the protein, aside from a slightly decreased activity, with the exception of Analogue **46**. This suggests Tyr176 is only partially involved in the binding of the sex pheromone components. The two mutants generated of Phe208, however, show a significant decrease in activity compared to the wild type. Both OBP6-Ile208 and OBP6-Tyr208 had comparably poor binding activity with the sex pheromone components and analogues.

Both mutants generated showed a significant decrease in binding activity the sex pheromone components (Table 4.4.; Figure 4.12). This suggests the predicted hydrogen bond with Phe208 is important in OBP6-sex pheromone interactions. The models are an estimate of activity, and *in vitro* assays, including a site-directed mutagenesis of Phe208, would be required to confirm the presence and importance of this particular amino acid interaction.

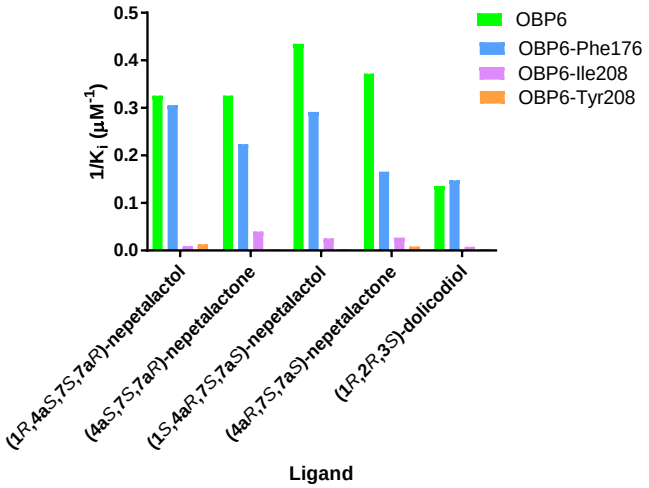


Figure 4.12: Sex pheromone binding with mutated OBP6. A significant decrease in activity (increase in K_i) is observed when the Phe208 residue is mutated.

Analogues of the sex pheromone gave varying results (Table 4.4; Figure 4.13). Analogues that removed the bicyclic structure (Analogues **48** and **49**) were seen to bind significantly worse to the OBPs than the bicyclic analogues. One analogue showed a significant increase in activity when interacting with OBP6 – a decrease of the K_i from approximately 2-3 µM to 380 nM was observed. This low binding constant suggests a high affinity interaction between Analogue **42** and OBP6. Analogue **42** lacks the heterocyclic oxygen – this modification could result in changes in the anomeric effect seen in this bicyclic lactol structure, resulting in a more favourable hydrogen bond interaction.

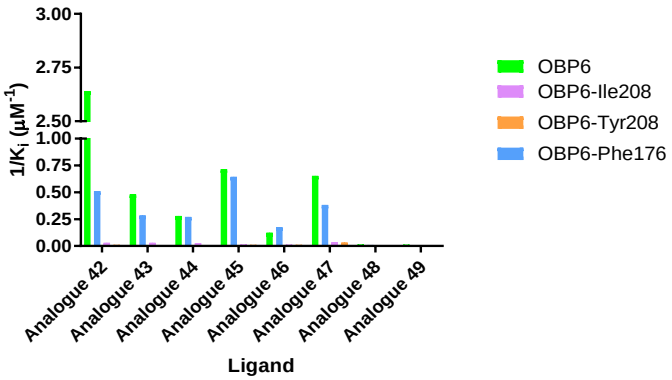


Figure 4.13: Sex pheromone component analogues binding with mutated OBP6 and mutants. Analogue **42** shows a significantly higher affinity for OBP6 than the sex pheromone components and other analogues

In future work, Analogue **42** should be synthesised, and its' predicted potent activity evaluated in *in vitro* assays and behavioural studies with aphids. If these experiments are successful in demonstrating the strong binding activity of the analogue, Analogue **42** could be a candidate for an olfactory inhibitor or blocker in aphids and may be useful in future pest control.

4.2.5 Sequence Alignment and Homology Modelling of *A. pisum* Odorant Receptors

Three odorant receptors from *A. pisum* were initially chosen for modelling for their known activity, with an additional six chosen for their high conservation within aphids. OR4 and OR5 have known functional activity and the olfactory co-receptor (ORCO) possesses the highest similarity to the template protein, ORCO from *Apocrypta bakeri,* in addition to being critical for the function of all other ORs.[99,100] Pair-wise sequence alignments and transmembrane domain assignment of all receptors showed a generally low sequence identity (OR2 20%, OR4 19%, OR5 18%, OR10 17%, OR17 17%, OR20 19%, OR22c 20% and OR39 17%; Clustal Omega[174]), comparable to other insect OR modelling studies[166], with the exception of *A. pisum* ORCO which shares 57% identity with *A. bakeri* ORCO. Each protein had 7 predicted transmembrane domains, as expected (Table 4.5.; Figure 4.14). ORCO has a relatively long predicted intracellular loop 3 of 41 residues (residue 211-336) compared with *A. bakeri* ORCO's 5 residue intracellular loop 3. All other predictions were similar to *A. bakeri's* ORCO transmembrane domains, with the exception of a long C-terminus segment in OR4.

Table 4.5: Transmembrane domain predictions of ORCO, OR4 and OR5 from *A. pisum*, compared with ORCO from *A. bakeri.* OR2, OR10, OR17, OR20, OR22c and OR39 are not displayed for clarity.

Protein	TM1	TM2	TM3	TM4	TM5	TM6	TM7
A. bakeri ORCO	31-61	65-100	112-150	170-223	230-284	289-330	338-352 367-387
A. pisum ORCO	43-65	75-93	134-156	192-210	337-356	366-385	437-460
A. pisum OR2	33-62	68-103	134-152	190-235	245-298	305-342	377-399
A. pisum OR4	37-60	69-105	127-148	181-207	256-280	288-317	327-385
A. pisum OR5	21-46	52-77	104-127	157-200	227-251	262-285	296-355

```
                  10        20        30        40        50        60        70        80        90       100
AbakORCO : --FKHQGLVADLLPNIRVMQGVGHFMFNYYSEGKKF---FHRIYCIVTLLLLLLQYGMMAVNLMMESDDVDDLTANTITMLFFLHPIVKMIYFPVRSKIF
ApisORCO : MGYKKDGLIKDLWPNIRLIQLSGLFISEYYDDYSGLAVLFRKIYSWITAIIIYSQFIFIVIFMVTKSNDSDQLAAGVVTTLFFTHSMIKFVYESTGTKSF

                 110       120       130       140       150       160       170       180       190       200
AbakORCO : YKTLAIWNNPNSHPLFAESNARFHALAITTKMRRLLFCVAGATIFSVISWTGITFIEDSP----------------IPRLMIRTFYPFNAMSGAGHVFALI
ApisORCO : YRTLSCWNNTSPHPLFAESHSRFHAKSLSRMRQLLIIVSIVTIFTTISWTTITFFGESVWKVPDPETFNQTMYVPVPRLMLHSWYPWDSGHGLGYIVAFV

                 210       220       230       240       250       260       270       280       290       300
AbakORCO : YQFYYLVISMAVSNSLDVLFCSWLLFACEQLQHLKAIMKPLM-------------------------------------------------ELSATGLTK-
ApisORCO : LQFYWVFIITSHSNLMELLFSSFLVHACEQLQHLKEILNPLIELSATLDSSVHNPAEIFRANSAKNQSINGIDHDYNGSYVNEITEYGTKGENEPNRKGP

                 310       320       330       340       350       360       370       380       390       400
AbakORCO : -----KQEMLVRSAIKYWVERHKHVVRLVTAVGDAYGVALLLHMLTTTITLTLLAYQATKVNGVNVYAATVIGYLLYTLGQVFLFCIFGNRLIEESSSVM
ApisORCO : NNLTSNQEVLVRSAIKYWVERHKHVVKYVSLITECYSSALLFHMLVSTVILTILAYQATKINGVNVFAFSTIGYLMYSFAQIFMFCIHGNELIEESSSVM

                 410       420       430       440       450       460
AbakORCO : EAAYSCHWYDGSEEAKTFVQIVCQQCQKAMSISGAKFFTVSLDLFASVLGAVVTYFMVLVQLK
ApisORCO : EAAYGCHWYDGSEEAKTFVQIVCQQCQKPLIVSGAKFFNVSLDLFASVLGAVVTYFMVLVQLK
```

```
                  10        20        30        40        50        60        70        80        90       100
AbakORCO : FKHQGL--VADLLPNIRVM-QGVGHFMFNYYSEGKKFPHRI-YCIVTLLLLLLQYGMMAVNLMMESDDVDDLTANTITMLFFLHPIVKMIYFPVRSKIFY
ApisOR4  : MTSIGKKNQRKEYQTLMTLAFFLDTSQYRYISRFVKQFYIFDWMVLVSVAAAFTIIEGNYRMPFVMEHIQYMIVGFYFTSIFVVFIIKKEAIMSNYNCIQ

                 110       120       130       140       150       160       170       180       190       200
AbakORCO : KTLAIWNNPNSHPLFAESNARFHALAITKMRRLLFCVAGATIFSVISWTGITFIEDSPIPRLMIRTFYPF--NAMSGAGHVFALIYQFYYLVISMAVSNS
ApisOR4  : TKFIQWSNK-RALHSNAAYKRNIKTVKSLSIPLAILSLSIALGPLISTINDIGKL-PLDNRAHFVLFWPTIVDTNKLSMYGIIYTLQVIFTIILYISVLS

                 210       220       230       240       250       260       270       280       290       300
AbakORCO : LDVLFCSWLLFACEQLQHLKAIMKPLMELSATGLTKKQEMLVRSAIKYWVERHKHVVRLVTAVGDAYGVALLLHMLTTTITLTLLAYQATKVNGVNVYAA
ApisOR4  : FNLGYMVPLNELITQFEMLLNGINDAFK-------YKMDKQFQTLFIDCIRHHQIIIKFLDDLKSYFKWMILIEIIVVQVILAILIYNLTKVNASLGYKV

                 310       320       330       340       350       360       370       380       390
AbakORCO : TVIGYLLYTLGQVFLFCIFGNRLIEESSSVMEAAYSCHWYDGSEEAKTFVQIVCQQCQKAMSISGAKF--FTVSLDLFASVLGAVVTYFMVLVQLK
ApisOR4  : KIAGSILFNLLPICFHCHVGEVVLSLHTRLSNHIYNMPWYDMPNKNKQLIVIMLQRTQRDLTLSSALFSSERASRSLISKVIKQVYTILNVLLKT-
```

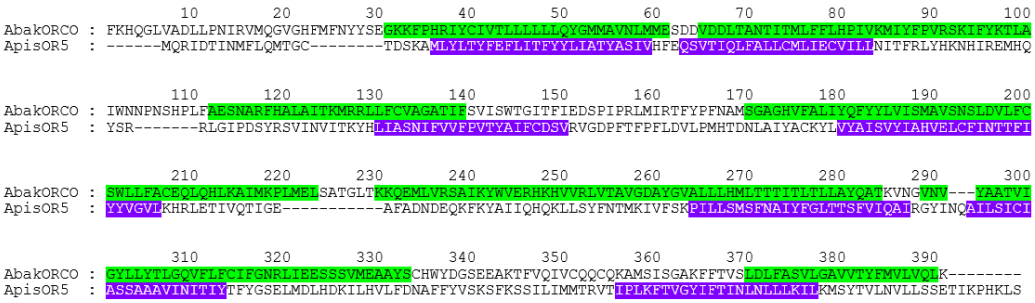


Figure 4.14: Sequence alignment and transmembrane domain predictions of *A. pisum* (Apis) ORCO, OR4 and OR5 (purple) against *A. bakeri* (Abak) ORCO (green). OR2, OR10, OR17, OR20, OR22c and OR39 are not displayed for clarity.

Homology models of insect ORs were generated using MODELLER 9.3 with a template of *A. bakeri* ORCO (Figure 4.15). Ligand-binding should take place in the extracellular domain of the protein, which includes multiple extracellular loops. For ORCO, the removal of 42-residue section of intracellular loop (IL3) improved the protein model. For other proteins, each loop was modelled individually with iTASSER, though, for OR2, OR4 and OR5, modelling of the longest loop, extracellular-loop 2, was the most critical in reducing the overall protein's energy. This loop of the OR protein, in addition to extracellular loop 3, has been shown to possesses critical binding domains in other insect ORs in previous work, and is particularly long in comparison to other loop portions of the protein.[88,175] The lowest energy predicted model, based on DOPE score (Table 4.6), was embedded into a lipid bilayer of 128 dipalmitoylphosphatidylcholine (DPPC) molecules using GROMACS before screening ligands.

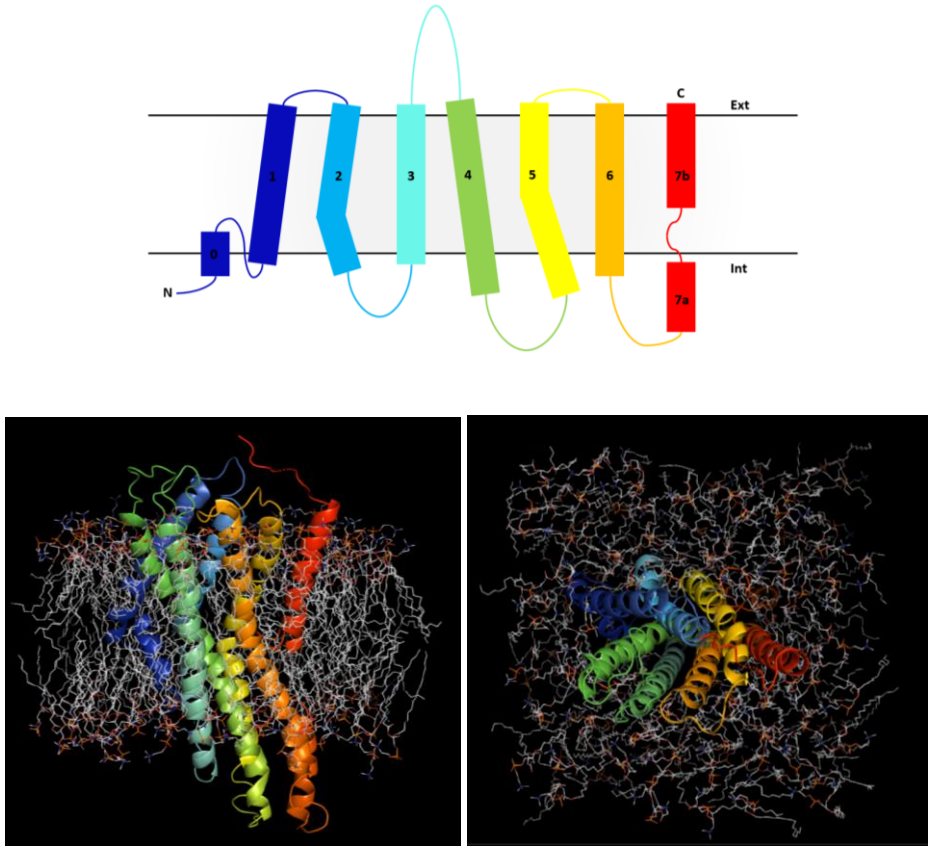


Figure 4.15: (i) The general transmembrane domains of insect odorant receptors; (ii) Homology models of *Acyrthosiphon pisum* odorant receptor 5 embedded in a lipid bilayer.

Table 4.6: DOPE scores (MODELLER) and Ramachandran analysis (PROCHECK) for the most favourable OR models.

***A. pisum* Protein Model**	**DOPE score**	**Ramachandran Data**		
		% Residues Favoured	**% Residues Allowed**	**% Residues Disallowed**
ORCO	-56024.94	93.5	6.0	0.5
OR2	-51562.26	89.7	9.8	0.5
OR4	-51244.34	88.5	10.4	1.1
OR5	-49737.77	91.7	7.2	1.1
OR10	-47680.4	89.3	9.3	1.4
OR17	-52616.25	87.1	11.9	1.0
OR20	-52364.49	87.7	10.7	1.6
OR22c-like	-46382.4	85.4	13.4	1.1
OR39	-55316.5	83.7	13.8	2.5

For each protein, the most favourable models had seven predicted transmembrane domains with an intracellular N-terminus and extracellular C-terminus, as predicted. All selected

models had over 98% allowed residues, with the exception of OR39 at 97.5%, with DOPE scores comparable to other insect OR models and the cryo-EM structure of ORCO from *A. bakeri* (Table 4.6).[88,166]

For visualisation, modelled protein structures were comprised into tetramers using Sam protein-protein docking (Figure 4.16).[176] The stoichiometry of the heteromeric ion channel that the ORCO and OR subunits form is not fully understood; it is known to forms a tetramer with at least one ORCO subunit and one OR subunit.[88,89,177] For this assembly, three ORCO units were compiled with one OR unit.

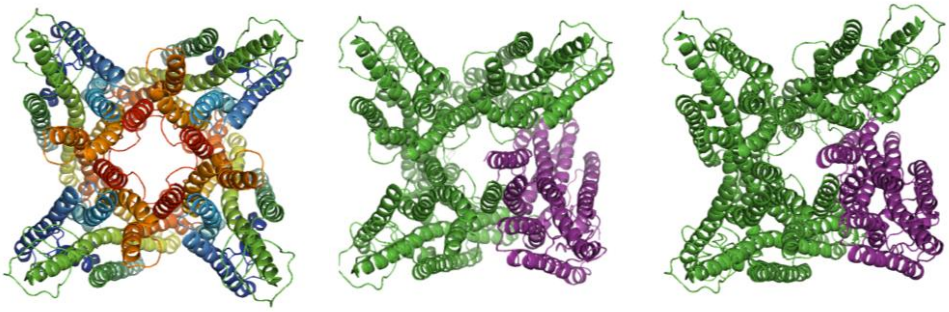


Figure 4.16: Homology models of *A. pisum* ORCO, OR4/ORCO and OR5/ORCO. ORCO (rainbow) is expressed as a tetramer, whereas the OR/ORCO complexes are expressed as three units of ORCO (green) and one unit of the specific OR (purple).

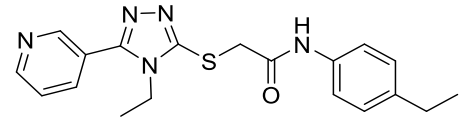
4.2.6 Ligand-Docking and Binding Site Assignment of *A. pisum* ORs

All ligand docking studies were performed using AutoDock 4.3 and the Racoon virtual screening tool. The predicted binding energy (kcal mol^{-1}) is reported for most, with inhibition constant (K_i) for the most critical docking conformations. Docking space was limited to the extracellular space of the odorant receptor, and all favourable conformations were found within the receptors predicted pocket.

Many olfactory ligands, such as aromatic compounds and unsaturated aldehydes, have a significant number of π-bonds, creating a challenge when predicting interactions *via* computational methods; Binding interactions are predicted by AutoDock 4.3 using a point-charge based force field. This results in several limitations when predicting binding conformations and kinetic values, in particular those involving π- π. Though these values are likely inaccurate to actual *in vitro* or *in silico* values, they can be used to score the binding conformations. Furthermore, π- π interactions that may appear in the true *in vivo* interactions can be predicted and estimated based on the positions of π bonds in the ligand and/or surrounding residues. Though efforts have been made to design more robust docking

software, for true analysis of non-covalent π- π interactions, a quantum mechanical analysis is required.[178] This would require a significant amount of computing power and molecular dynamics simulations.[159,168]

The unique ligand VUAA1 **50** (Figure 4.17) has been shown to act as an olfactory blocker in insects and has a high binding affinity with the olfactory receptor co-receptor (ORCO).[88,177] VUAA1 **50** is often used to test the functionality and viability of expressed ORCO.



50

Figure 4.17: The structure of olfactory ligand VUAA1 **50**.

VUAA1 **50** was screened with the modelled structure of ORCO from *A. pisum*. A predicted binding energy of -11 kcal mol^{-1}, with a surprisingly low K_i of 12nM, was observed. ORCO possess an electron rich binding site (Phe82, Phe83, Phe144, Trp149, Phe369, Tyr 332 (Tyr374), Tyr335 (Tyr377) consisting of aromatic residues (Figure 4.18). VUAA1 **50** is predicted to bind with a moderate hydrogen bond (3.1Å) between the oxygen of Tyr332's phenol and the nitrogen of the pyridine ring in VUAA1 **50**. Furthermore, this triazole ring is perfectly positioned (approximately 3.7Å and 4.0 Å respectively) to undergo sandwiched π-stacking with Trp149 and T-shaped π-stacking with Phe83. In addition to these interactions, varying strength π- π interactions are predicted between other unsaturated sites within the molecule and the aromatic binding site residues.

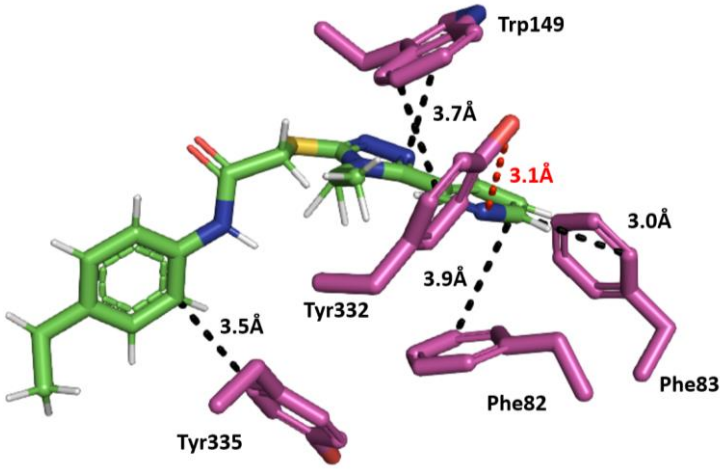


Figure 4.18: VUAA1 **50** in the highly conserved binding pocket of *A. pisum* ORCO. Interactions with key residues key residues can be seen – potential π-π interactions (black) with Phe82, Phe83, Trp149 and Tyr335, in addition to a moderate hydrogen bond (red) with Try332. For clarity, only side chains of residues are displayed.

This binding site is highly conserved; a screen of 66 ORCO genes, with an average sequence conservation of 65%, from different insect species showed high conservation of these key residues Phe82 (97%), Phe83 (98%), Trp149 (100%), Try332 (97%) and Tyr335 (100%). Previous work involving the specialist fly, *Mayetiola destructor,* which does not display endogenous activity with VUAA1 **50**, among other species elucidated some key binding residues for VUAA1 **50** in ORCO, including Phe83.[177]

ORCO is co-expressed as a predicted tetramer with all ORs in insects and is required for olfactory function.[88] Subsequently, ORCO is an ideal olfactory target for control, however its high conservation means any targeting will not be species specific. Odorant receptors are much more divergent and may prove a more promising target for species specific control.

Previous work screened *A. pisum* OR4 against a library of 57 molecules, and eight general plant volatiles were identified as being perceived by the aphid's antennal sensilla and OR4, demonstrating the broad-tuning of OR4.[100] These volatiles were 4-ethylacetophenone, salicylaldehyde, 4-ethylbenzaldehyde, 3-vinylbenzaldehyde, (1*S*)-(-)-verbenone, (*S*)-cis-verbenol, (-)-borneol and (1*R*)-(-)-myrtenal, with three specific aromatic volatiles (4-ethylacetophenone, salicylaldehyde, 3-vinylbenzaldehyde) producing the highest electrophysiological response from OR4.

OR4 was screened against the identified eight plant volatiles and the predicted binding sites analysed (Table 4.7). The aromatic volatiles bound in a highly aromatic binding site (Figure 4.19) involving key residues Phe53, Phe83, His161 and Trp166. Potential π -π interactions between the benzene ring of the three aromatic volatiles and the aromatic residues in the binding site will be relatively weak as most aromatic residues are over 4Å from the ligand. One interaction, with His161 at approximately 3.5Å is present in all the predicted binding sites of the active aromatic ligands. This well-defined binding site with multiple interactions explains the higher electrophysiological response seen when OR4 is activated by these ligands.[100]

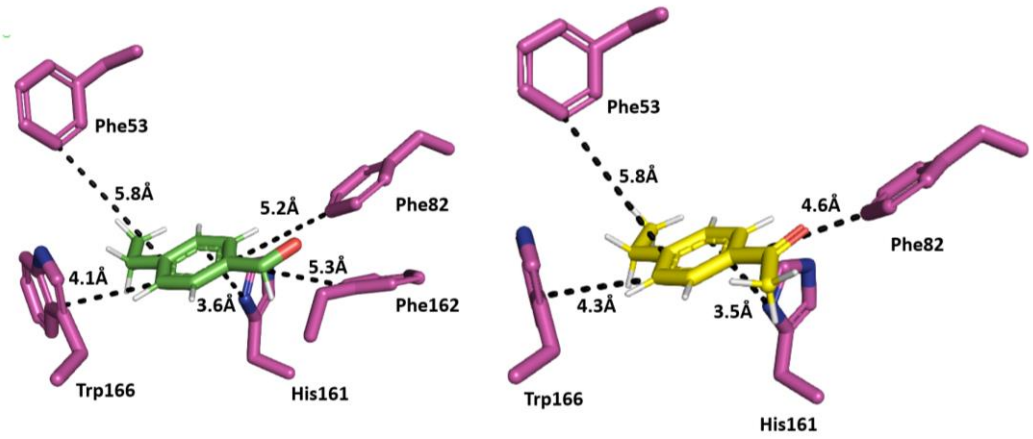


Figure 4.19: The aromatic binding site, Phe53, Phe82, His161, Phe162 and Trp166 of OR4, with the predicted π-π interactions (black) with 4-ethylbenzaldehyde and 4-ethylacetopheone, respectively, shown. For clarity, only side chains of residues are displayed.

Other electrophysiologically active ligands appear to bind in multiple other different sites. Alcohols (*S*)-cis-verbenol and (-)-borneol, for example, appear to form polar bonds between their hydroxyl group and Lys153 or Arg61 and Leu154, respectively (Figure 4.20). The other volatiles either bound in one of these sites or elsewhere *via* mainly hydrophobic interactions with residues found in extra cellular loop 2 (EL2 - Va163, Leu64, Phe165, Trp166, Pro167).

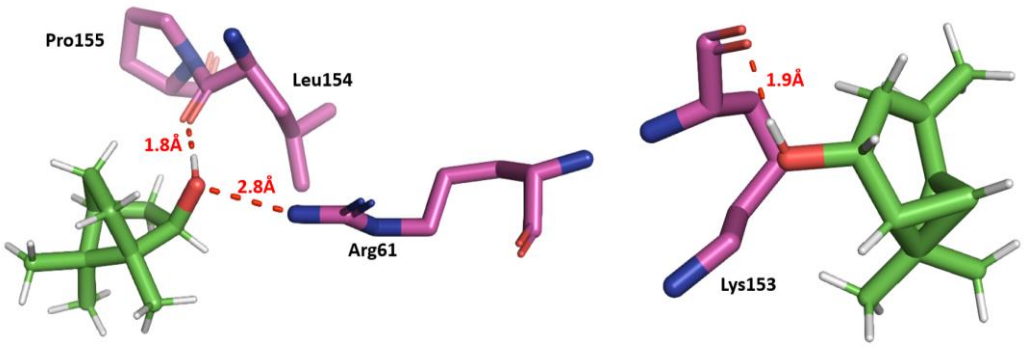


Figure 4.20: Two additional binding sites of OR4, with both (-)-borneol interacting *via* hydrogen bonds (red) with Leu154 and Arg61, and (*S*)-cis-verbenol interacting with Lys153.

The critical aromatic binding site (Phe53, Phe83, His161 and Trp166), in addition to the other key residues identified (Lys153, Leu154 and EL2) are all high conserved across OR4 and OR4-like proteins in aphid species. All except Phe53, at 50% conservation, and Phe83, at 87.5 %, are 100% conserved across the corresponding proteins in different eight aphid species (Figure 4.21). One hypothesis for this high conservation is that many aphid species feed on the same or similar host plants and so would be expected to detect the same volatile

compounds. The most divergent protein was for the species *Sipha flava* and *Melanaphis sacchari*, two species which utilise the same host plant, *Poaceae* or grasses.[179,180] The proline found in *A. pisum* OR4 at position 167 is highly conserved across many insect odorant receptors, and is predicted to be important for binding of ligands to ORs, often forming part of a much larger, well-defined binding site.[166] For OR4, Pro-167 forms part of one of the predicted binding sites and can be found on extra cellular loop 2.

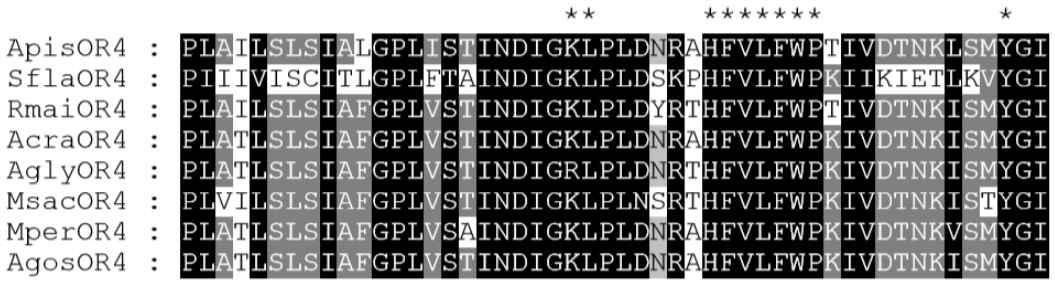


Figure 4.21: The conserved binding site of OR4, with key residues marked with asterisks (*).

In addition to the electrophysiologically active plant volatiles, a range of other compounds were screened against OR4. This included the aphid sex pheromone components, (4a*S*,7*S*,7aR)-nepetalactone **6** and (1*R*,4a*S*,7*S*,7a*R*)-nepetalactol **5** (Table 4.7). OR4 had predicted interactions with both the sex pheromone components, with (1*R*,4a*S*,7*S*,7a*R*)-nepetalactol **5** binding with a particularly low K_i of 1.8 µM. These predicted interactions suggest OR4 is a broadly tuned receptor with limited specificity. Studies observing these interactions *in vitro* have not yet been conducted, but these OR4 screening results could be used to direct mutagenesis experiments. Furthermore, the presence of multiple binding sites in OR4 suggests that binding site location may be more critical than the specific kinetics; the three compounds with the highest response levels were all predicted to bind in the same site.

Table 4.7: Binding energy and inhibition constants for the interactions of the functionalised odorant receptors from *A. pisum* (ORCO, OR4, OR5) with a variety of ligands.

Ligand	K_i	Binding Energy (kcal mol^{-1})
	ORCO	
VUAA1 **50**	12 nM	-10
	OR4	
4-ethylacetophenone	1.1 µM	-8.1
salicylaldehyde	4.4 µM	-7.3
4-ethylbenzaldehyde	3.7 µM	-7.4
3-vinylbenzaldehyde	3.3 µM	-7.5
(1S)-(-)-verbenone	29 µM	-6.2
(*S*)-cis-verbenol	2.9 µM	-7.6
(-)-borneol	14 µM	-6.6
(1*R*)-(-)-myrtenal.	79 µM	-5.6
VUAA1 **50**	1.0 µM	-8.2
(1*R*,4a*S*,7*S*,7a*R*)-nepetalactol **5**	6.2 µM	-7.1
(4aS,7S,7aR)-nepetalactone **6**	1.8 µM	-7.9
	OR5	
(*E*)-β-farnesene **17**	2.6 µM	-7.6
Geranyl acetate	19 µM	-6.4
(*S*)-Germacrene D **25**	2.1 µM	-7.8
(1*R*,4*E*,9*S*)-caryophyllene **26**	1.2 µM	-8.1
VUAA1 **50**	2.6 µM	-7.6

The aphid alarm pheromone, EBF **17**, has a repellent behavioural effect on aphids. OR5 has been shown to be critical in alarm pheromone perception in *A. pisum.*[99] In addition to **17**, OR5 has a high electrophysiological response to geranyl acetate, a behaviourally repellent compound.[99] It has been suggested that screening of OR5 could be used to identify and characterise potentially novel repellent compounds for pest control.[99]

An interaction between OR5 and EBF **17** was predicted with a low energy (K_i of 2.6 µM) and a strong interaction (-7.6 kcal mol^{-1}). Four key residues have been identified as being critical in this interaction His168, Tyr157, Tyr243 and Phe239 (Figure 4.22). EBF **17** appeared to bind in an electron rich site. This interaction comprises of mainly hydrophobic interactions, in addition to π-π interactions of the unsaturated bonds in EBF **17**, as seen between the various unsaturated bonds and three aromatic residues (Tyr157, Phe239, Tyr243) and His168. The

tightest of these interactions appears to be between Tyr157 and the two terminal unsaturated bonds of EBF **17**. Additionally, geranyl acetate (GA) was screened due to the observed electrophysiological activity of this compound.[99] GA was seen have a weaker interaction with OR5 (-6.44 kcal mol^{-1}) but bound in a similar site and also presented a potential π-π interaction with His168, in addition to some weaker interactions with residues Phe139 and His134.

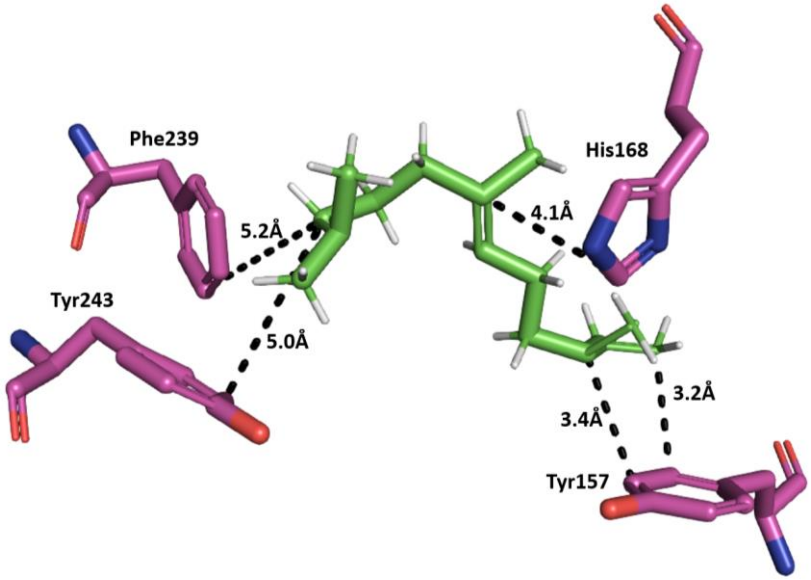


Figure 4.22: The aromatic binding site, Tyr157, His168, Phe239 and Tyr243 of OR5, with the predicted π-π interactions (black) with (*E*)-β-farnesene **17** shown.

Other aphid semiochemicals were screened, including (*S*)-germacrene D **25** (SGD), a known repellent compound, and (1*R*,4*E*,9*S*)-caryophyllene **26**, a compound that is biosynthetically similar to EBF **17** and known to be an inhibitor of alarm pheromone response.[181] Both these semiochemicals were predicted to bind with similar energy (SGD **25** -7.8 kcal mol^{-1} and (1*R*,4*E*,9*S*)-caryophyllene **26** at -8.1 kcal mol^{-1}) to EBF **17**, however, did not appear to fit within in the same predicted pocket. These two compounds were shown to have no electrophysiological activity with OR5.[99] This conflicting result may demonstrate that the specific binding site of ligands may be more critical for receptor function than specific binding energy. OBP3 and OBP7 have also been shown to be critical in the perception of both EBF **17** and GA[99], and this may play a role in why some compounds are seen to be electrophysiologically active over others.

The closest ORs from other aphid species to *A. pisum* OR5 are *M. persicae* OR43b-like and *A. gossypii* OR43b-like. The binding site of EBF **17**, including residues Tyr157, His168, Phe239 and Tyr243, is conserved across all three receptors (Figure 4.23). As both *M. persicae* and *A.*

gossypii both utilise the alarm pheromone EBF **17** [182,183], it is likely that the OR43b-like receptors are involved in its perception in these species.

Figure 4.23: The conserved binding site of OR5, with key residues marked with asterisks (*).

In addition to previously functionalised receptors, six highly conserved deorphanized receptors (OR2, OR10, OR17, OR20, OR22c, OR39) were modelled and screened for activity with the aphid sex pheromone components [(1*R*,4a*S*,7*S*,7a*R*)-nepetalactol **5** and (4a*S*,7*S*,7a*R*)-nepetalactone **6**] and their respective, non-naturally occurring enantiomers [(1*S*,4a*R*,7*R*,7a*S*)-nepetalactol **32** and (4a*R*,7*R*,7a*S*)-nepetalactone **33**] (Table 4.8). Little is understood about how sex pheromone components are perceived; in particular how enantiomeric discrimination occurs.

Table 4.8: Binding energy for the interactions of the deorphanized, highly conserved odorant receptors (OR2, OR10, OR17, OR20, OR22c, OR39) with the aphid sex pheromone components and respective enantiomers. The values highlighted in **bold** represent the lowest energy interaction for that ligand.

Ligand	Binding Energy (kcal mol^{-1})					
	OR2	OR10	OR17	OR20	OR22c	OR39
(4a*S*,7*S*,7a*R*)-nepetalactone **6**	-6.8	-7.0	-7.2	**-7.3**	-6.9	-6.7
(1*R*,4a*S*,7*S*,7a*R*)-nepetalactol **5**	-6.8	-6.3	-7.2	**-7.3**	-6.8	-6.3
(4a*R*,7*R*,7a*S*)-nepetalactone **33**	-6.8	-6.8	-7.3	**-7.4**	-6.8	-6.7
(1*S*,4a*R*,7*R*,7a*S*)-nepetalactol **32**	-6.8	-6.7	**-7.4**	-7.2	-6.5	-6.7

Of all the receptors, OR17 and OR20 showed the highest affinity interaction with the sex pheromone components. Unfortunately, no enantiomeric discrimination was seen based on binding affinity alone, and for five out of the six screened receptors both the natural and non-natural enantiomers bound in the same site. When binding with OR2, naturally occurring nepetalactol/lactone and their respective enantiomers bound in different sites (Figure 4.24). In particular, the natural lactol (1*R*,4a*S*,7*S*,7a*R*)-nepetalactol **5** formed a polar hydrogen bond with Met80, whereas no polar contacts were seen between OR2 and (1*S*,4a*R*,7*R*,7a*S*)-

nepetalactol **32** with Asp161. This may indicate a discriminatory ability of OR2, further demonstrating that the binding site is critical in determining activity.

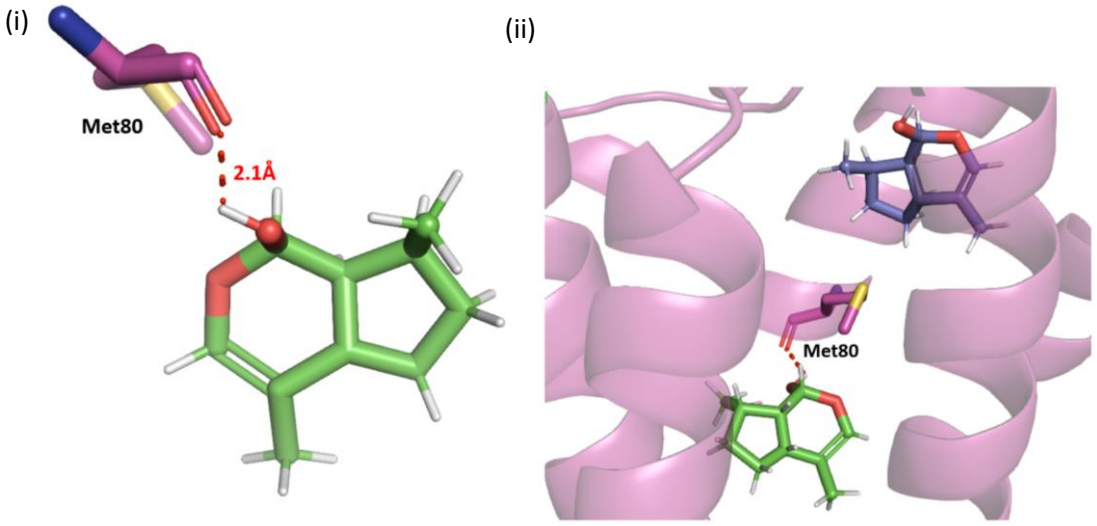


Figure 4.24: (i) The predicted hydrogen bonding (red) of (1*R*,4a*S*,7*S*,7a*R*)-nepetalactol **5** with Met80; (ii) Both (1*R*,4a*S*,7*S*,7a*R*)-nepetalactol **5** (green) (1*S*,4a*R*,7*R*,7a*S*)-nepetalactol **32** (blue) with OR2 (pink).

Future work is required to evaluate electrophysiological activity of these specific receptors, in addition to knockdown or knockout studies. Advances in molecular techniques, particularly the advent of CRISPR/Cas9, may allow for these genes to be entirely knocked out of aphids or placed into model species to determine activity.

4.3 CONCLUSIONS

In summary, OBPs and ORs from the pea aphid, *A. pisum*, were modelled and screened successfully with several ligands. OBP6 was predicted to have strong binding affinity to the aphid sex pheromone components, both the biologically active enantiomers and the inactive enantiomers. When this interaction was further investigated, a key residue, Phe208, was identified. Furthermore, screening of analogues of the sex pheromone component predicted a highly potent analogue that may be used in future pest management strategies.

Nine ORs from *A. pisum* have been modelled and ligand-docking studies conducted. The binding sites of VUAA1 **50** in ORCO, EBF **17** (the alarm pheromone) and geranyl acetate with OR5 and various plant volatiles with OR4 have all been predicted and described. All three receptors show high levels of conservation at the predicted binding site within aphid species. Conservation among receptors indicates closely related, if not identical, olfactory function.

It is clear from these results that binding affinity and energy is not the only factor that can be evaluated in a computational screen to determine subsequent biological roles and activities of specific receptors and receptor/ligand combinations. The mode-of-action of insect ORs is not yet fully understood, and how these receptors are activated may be understood by observing key binding residues. Computation screens can be confirmed by *in vitro* work, particularly by mutagenesis of predicted binding sites.

Predicted interactions of the OBPs, particularly the interactions of OBP6, will be investigated further in this thesis *via in vitro* ligand-binding assays. These will include fluorescence, mass spectrometry and nuclear magnetic resonance techniques. Future work in ORs should test the viability of screened receptor-ligand combinations, taking into consideration binding site specifics. Furthermore, significantly more work is required in determining the 3D-structures and stoichiometry of ORs and the OR/ORCO complex. The availability of one single structure, *A. bakeri* ORCO, has significantly increased modelling potential for these receptors, and subsequent models will be critical in understanding this complex and dynamic system in insects.

Chapter 5: Fluorescence Based Assays of *A. pisum* Odorant-Binding Proteins and Interactions with Ligands

5.1 INTRODUCTION

In the field of insect olfaction, ligand-binding studies between odorant-binding proteins and ligands are often conducted using fluorescence-based assays.[184] Fluorescence is a photochemical process, where an electron in an excited state loses energy by a form of radiative emission, producing a specific wavelength of light (Figure 5.1).

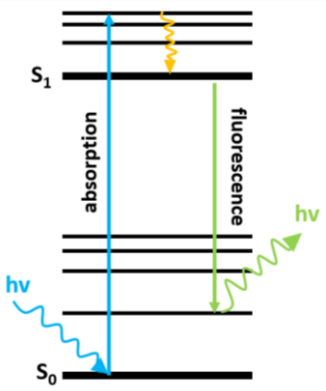


Figure 5.1: Jablonski diagram showing fluorescence. When electrons in fluorescence-active molecules (fluorophores) are excited by particular wavelengths of light, they are excited from the ground state to a higher energy level (blue). The electrons then relax by a combination of non-radiative emission (orange) and fluorescence (green), which emits at a specific wavelength photon differing from the excitation wavelength.

In fluorescent binding studies, the changing fluorescence of a specific fluorophore, such as the fluorescent amino acid tryptophan **54** or an introduced fluorescent probe, is monitored and used to calculate associated binding kinetics, such as the dissociation constant (K_D).[102,184] Fluorescence requires a fluorophore – a fluorescent chemical compound that can be excited and will subsequently re-emit light (Figure 5.1). Fluorophores (Figure 5.2) generally have an abundance of conjugated bonds, and can be synthetic, such as *N*-phenylnapthalen-1-amine **51** (1-NPN) and 8-anilino-1-napthalenesulfonic acid **52** (1,8-ANS), or biological, such as an intrinsic amino acid. In addition to tryptophan **54**, and another aromatic amino acid, tyrosine **53**, demonstrates fluorescent properties.[185] There are many forms of fluorescent protein-ligand binding assays that can be used for olfactory proteins, though generally a synthetic fluorophore is used in a titration and competitive inhibition-type assay.[102,114,125,186] The

fluorescence of intrinsic tryptophan **54** can be utilised, although tryptophan **54** is not a particularly common amino acid and is not present in all odorant-binding proteins (OBPs). Tryptophan's fluorescent properties exist as a result of the delocalised electrons in the connected indole ring system, and it is more sensitive to small environmental changes than tyrosine **53**.[185,187]

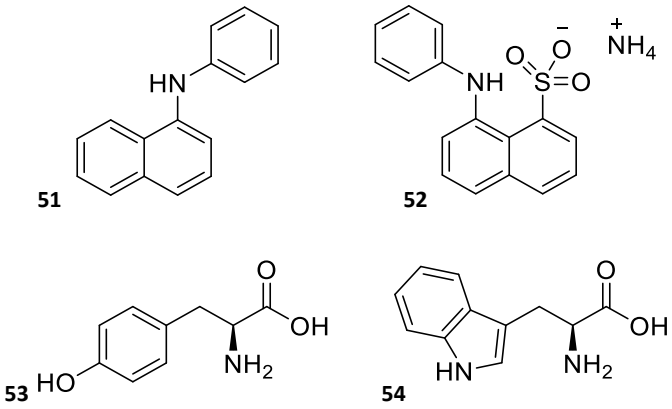


Figure 5.2: Synthetic fluorophores, *N*-phenylnapthalen-1-amine **51** (1-NPN) or 8-anilino-1-napthalenesulfonic acid **52** (1,8-ANS) and fluorescence amino acids, tyrosine **53** and tryptophan **54**.

Standard fluorescent probe assays to study the interaction between insect OBPs and their respective ligands are all generally performed using a standard method involving the fluorescent probe 1-NPN **51**.[105,112,118,125] However, this method does not work for all protein-ligand interactions and care should be taken when using it. Results can vary based on the choice of fluorescent ligand, and often fluorescence results do not correlate with structural and behavioural data.[186,188,189] The traditional approach to fluorescence assays relies on several assumptions. Firstly, an assumption is made that the fluorescent probe binds with the OBP. Fluorophores such as 1-NPN **51** are insoluble in aqueous conditions at 'high' concentrations (approximately 16 µM), often resulting in a false binding curve when titrated against the protein. Almost all reported use of 1-NPN **51** include a titration to a higher concentration than 16 µM. In addition, titration of the fluorescent ligand into the protein produces often unreliable and illogical kinetic data, if the assay is not well designed or the correct assay to use in the specific scenario.[190] Though some 1-NPN **51** will be solubilised by binding to the protein, this is highly dependent on both the protein concentration and binding affinity of 1-NPN **51** to the specific protein. There are many examples of this unreliable binding curve calculation in the literature relating to OBPs.[102,125,186,191,192] To overcome this issue, alternative fluorescent probes, such as 1,8-ANS **52** that have a greater aqueous solubility could be used.[186,188]

Secondly, it is assumed that the fluorescent probe and the ligand of interest bind in the same protein site (competitive binding). Quenching of fluorescence, resulting from non-specific binding or weak interactions between the fluorescent probe and the protein, may easily be mistaken for competitive binding.[188] As ligand is titrated into the protein-fluorescent probe complex, a change in fluorescence can be seen, usually a decrease in intensity and a shift in wavelength.[193] To ensure a true competitive binding, the assay can be inverted and 1-NPN **51** can be titrated into a protein-ligand complex. In Murphy *et al*., 2013[188], a shift is seen when 1-NPN **51** is titrated into AgamOBP22 with varying concentrations of citronellol. However, 1-NPN **51** fluorescence does not vary with any of the other ligands as their concentrations vary. This suggests that although it may initially appear that AgamOBP22 interacts with only citronellol, citronellol is actually the only ligand that can competitively displace 1-NPN **51**. When another fluorescent probe, 1,8-ANS **52**, is used, many other ligands show binding activity.[188] Although 1-NPN **51** is used prolifically through the literature, it has many drawbacks, many of which have been addressed in a small number of studies by the use of alternative probe or the use of intrinsic fluorescence.[115,186,190]

The use of intrinsic fluorescence removes the need for an introduced fluorescent probe completely.[185] By monitoring the change in the fluorescence of an amino acid, such as tryptophan **54**, direct interactions with the protein can be observed. There are still limitations to this method, namely the protein must contain a fluorescent amino acid residue, and this residue must also be in a position where binding will induce an environmental change. Practically, this often requires the fluorescent residue to be either within the binding site or relatively close.

A combination of intrinsic fluorescence and a fluorescent probe, such as 1-NPN **51**, has been used before to study OBP-ligand interactions.[115] However, the best approach to these fluorescence assays is to take into consideration each method, including the use of different fluorophores, and accounting for their limitations, and to use multiple methods to determine the most accurate result.

Overall, fluorescence assays are useful for initial screening of ligands with proteins for activity, being relatively quick and simple to perform, providing essential thermodynamic data. However, fluorescence and other techniques that focus on the kinetic and thermodynamic properties of the binding interaction, generally require high concentration and high purity proteins.[194] They also provide little information about the intricacies of the interaction that occurs when the ligand binds, including which amino acids are involved.

Alternative techniques, which provide more detailed information about interactions, include the use of mass spectrometry and nuclear magnetic resonance (NMR) spectroscopy. [194,195]

In this chapter, fluorescence assays are used to assess the interactions of odorant-binding proteins with pheromone components. By using *in silico* data (Chapter 4) to direct screening, predicted interactions will be tested. Fluorescence will be used to provide the first *in vitro* results of the binding and potential enantiomeric discrimination of the sex pheromone components.

5.2 RESULTS & DISCUSSION

5.2.1 Binding of Fluorescent Probe 1-NPN

Two *Acyrthosiphon pisum* odorant binding proteins were selected for fluorescence-based assays – OBP6 and OBP9. OBP6 was selected due to its predicted binding activity with the sex pheromone components from *in silico* studies (Chapter 4), whereas OBP9 was chosen as a protein which appeared to have no predicted binding activity with sex pheromone components, to act as a negative control. Furthermore, both proteins contain tryptophan **54** residues and therefore have intrinsic fluorescence, with one Trp residue found in OBP9 and three in OBP6.

Initially, binding assays were conducted to determine whether each protein bound with the fluorescent probe 1-NPN **51**. This probe is traditionally used in competitive binding studies with OBPs, where the fluorescence of 1-NPN **51** is monitored to determine whether it can be displaced by the ligands of interest. Not all OBPs bind to 1-NPN **51**, and therefore other fluorescent probes, such as 1,8-ANS **52** may also be used.[188]

1-NPN **51** is excited at 337 nm and fluoresces at approximately 460 nm. To generate an initial saturation curve, 1-NPN **51** was titrated into 25 mM Tris and 2 µM protein in 25 mM Tris to final concentrations of 0-16 µM. The fluorescence emission (across 390-600 nm) at excitation of 337 nm was recorded (Figure 5.3).

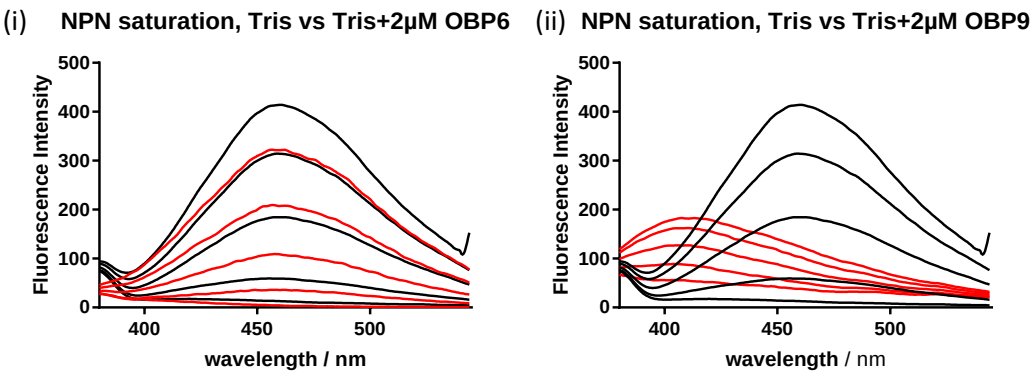


Figure 5.3: Titration of 1-NPN **51** into both 25 mM Tris (black) 2 μM protein in 25 mM Tris (red). The fluorescence intensity increases as 1-NPN **51** is titrated in to either sample; OBP6 (i), where fluorescence intensity increases but is slightly quenched, and OBP9 (ii), where fluorescence intensity increases but is significantly quenched and shifted to a lower wavelength.

The fluorescence of 1-NPN **51** (Figure 5.3; black) reduced in intensity, or was quenched, by the presence of the protein (Figure 5.3; red). For OBP6, a simple quench was observed; the emission wavelengths remained consistent and no shift was observed. This suggests that 1-NPN **51** does not bind with OBP6 directly, but a small or weak interaction occurs by which the fluorescence of the probe is quenched. When 1-NPN **51** is titrated into a solution of OBP9, the fluorescence is not only significantly quenched, but is shifted to approximately 405 nm (Figure 5.3 (ii)). This blue shift has been previously reported for interactions of OBPs and 1-NPN **51**.[118] The quench and blue shift are likely due to interactions with the fluorescent residues of OBP9. This shift occurs if 1-NPN **51** binds in close proximity to a fluorescent residue; a Förster resonance energy transfer (FRET; Figure 5.4) can occur between the two fluorophores, resulting in a new peak with shifted emission.

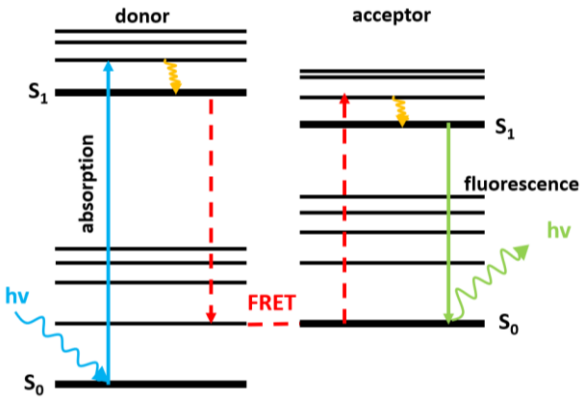


Figure 5.4: Förster resonance energy transfer (FRET; red) Jablonski diagram. Electrons are excited from their ground state (blue), however, the energy emitted from the 'donor' fluorophore is transferred to an 'acceptor' fluorophore, where it excites a different electron. This fluorophore then releases the energy by emitting a photo of a specific energy (green).

This titration demonstrates the interaction of OBP9 with 1-NPN **51**. A saturation curve can be generated and the K_D of 1-NPN **51** to OBP9 calculated by observing the relative increase in fluorescence at 405 nm (Figure 5.5).

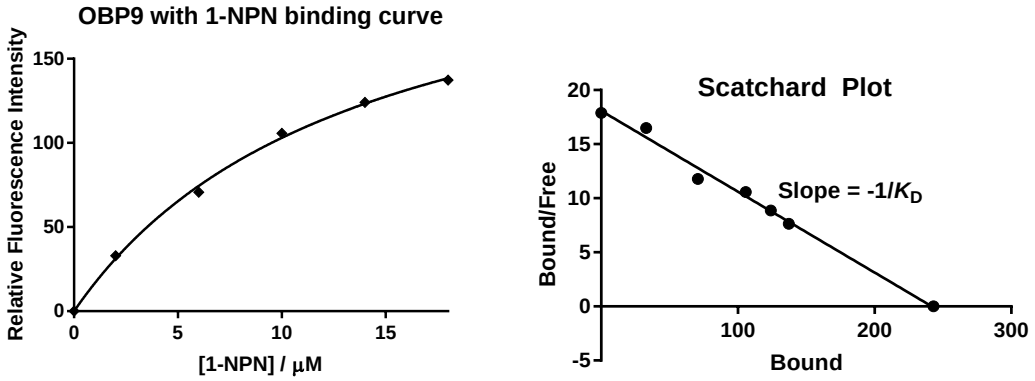


Figure 5.5: (i) The binding curve of 1-NPN **51** with OBP9. 1-NPN **51** is calculated to have a K_D of 13.59 ± 1.58 μM from this binding curve; (ii) Scatchard Plot, generated from the calculate B_{max} and K_D values.

From this assay, the binding affinity, or K_D, of 1-NPN **51** to OBP9 was calculated to be 13.59 ± 1.58 μM, a weak binding interaction. The K_D represents the concentration of ligand at which half of the protein is saturated and is calculated by the following equation, *via* a non-linear regression analysis:

$$y = \frac{B_{max} * x}{K_D + x}$$

The binding curve for OBP9 and 1-NPN **51** did not appear to reach saturation and gave a K_D which is close the maximum concentration that was titrated into the protein; the K_D from this specific assay is unlikely to be accurate. OBP6 did not appear to bind with 1-NPN **51**; although this method has been utilised in many previous binding studies, including with *A. pisum* OBP3 and OBP7[102], it is clear that it will not work for all OBPs. However, it should be possible to undertake competitive binding assays with OBP9.

Binding of insect OBPs and ligands can also be assessed by monitoring changes in the tryptophan **54** residue fluorescence. Intrinsic fluorescence can be observed either alone, with no fluorescent probes or ligands being introduced, or in a competitive binding assay similar to the 337 nm excitation assay. Tryptophan **54** can be excited at 280 nm and will emit with a maximum at approximately 330 nm. The tryptophan **54** residue fluorescence of OBP6 and OBP9 at 2 μM in 25 mM Tris was recorded (Figure 5.6).

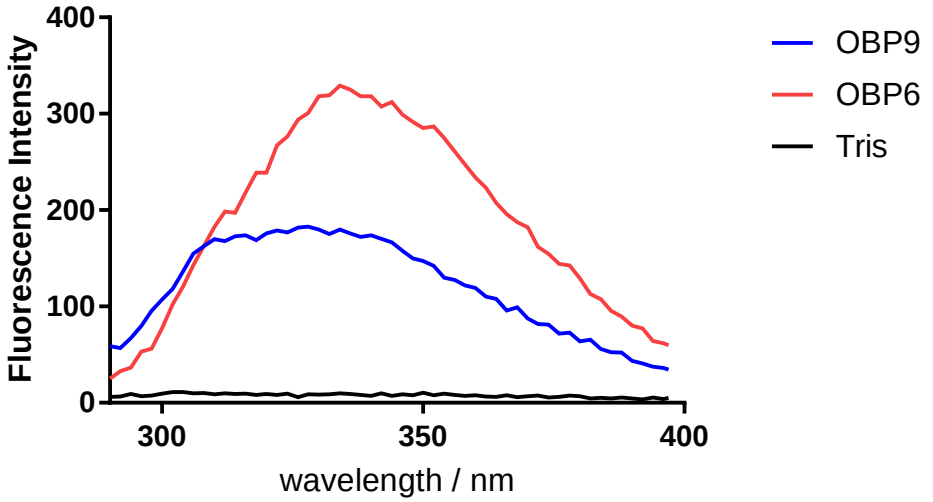


Figure 5.6: Tryptophan **54** fluorescence (excitation 280 nm) of 25 mM Tris (black) and 1 µM OBP6 (red) or OBP9 (blue) in 25 mM Tris.

The recorded intensity of OBP6 at an excitation of 280 nm was higher than OBP9 due to the number of fluorescent tryptophan **54** residues found in each protein. OBP9 also had a slightly shifted peak, demonstrating the effects a different positioning of the Trp residue within the protein can have. From *in silico* models (Chapter 4), the Trp residues of OBP6 were observed to be within the general binding pocket of the protein and buried deep within the 3-dimensional structure whereas in OBP9 the Trp residue is closer to the surface of the protein.

As with the assays at 337 nm, 1-NPN **51** can be titrated into the protein to determine whether binding is observed. Observing saturation of insect OBPs with 1-NPN **51** and tryptophan **54** has previously been reported to provide a new peak in the fluorescence at approximately 395 nm[115], and a subsequent decrease in fluorescence at 330 nm – this is seen for both OBP6 and 9 (Figure 5.7). The peak at 395 nm occurs due to a resonance energy transfer from the tryptophan **54** to the 1-NPN **51**.[115]

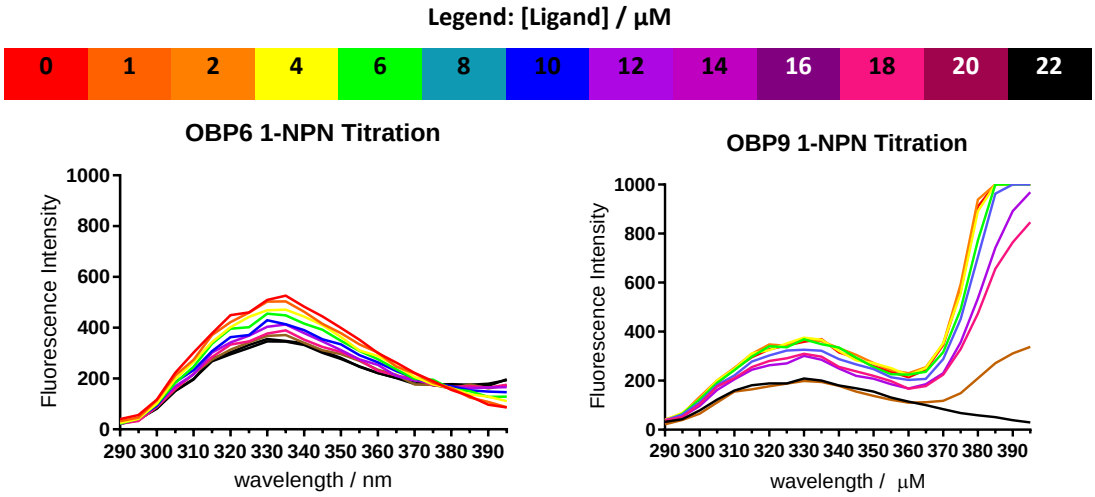


Figure 5.7: Tryptophan **54** fluorescence (280 nm excitation) of proteins with a titration of 1-NPN **51** to final concentrations of 0-16 µM with 0 µM 1-NPN in red and 16 µM 1-NPN in black.

Interestingly, the observation of this small peak at 395 nm and corresponding decrease at 330 nm in the tryptophan fluorescence of OBP6 suggests a weak binding of 1-NPN **51** to OBP6, not previously seen in the 337 nm fluorescence results. This may be due to weak binding of 1-NPN **51** with only one of the three tryptophan **54** residues, presumably the one located in the putative binding pocket highlighted in Chapter 4. In contrast, OBP9 has a strong binding interaction with 1-NPN **51** resulting in a dramatic increase in fluorescence at 395 nm. For both proteins, a small increase is observed at 340 nm. This relates to the quenching of 1-NPN **51** by the tryptophan **54** residue, as observed in the 337 nm assay (Figure 5.3). The comparison of these two proteins (Figure 5.8) demonstrates how weak and insignificant the interaction between 1-NPN **51** and OBP6 is in comparison to the interaction between 1-NPN **51** and OBP9.

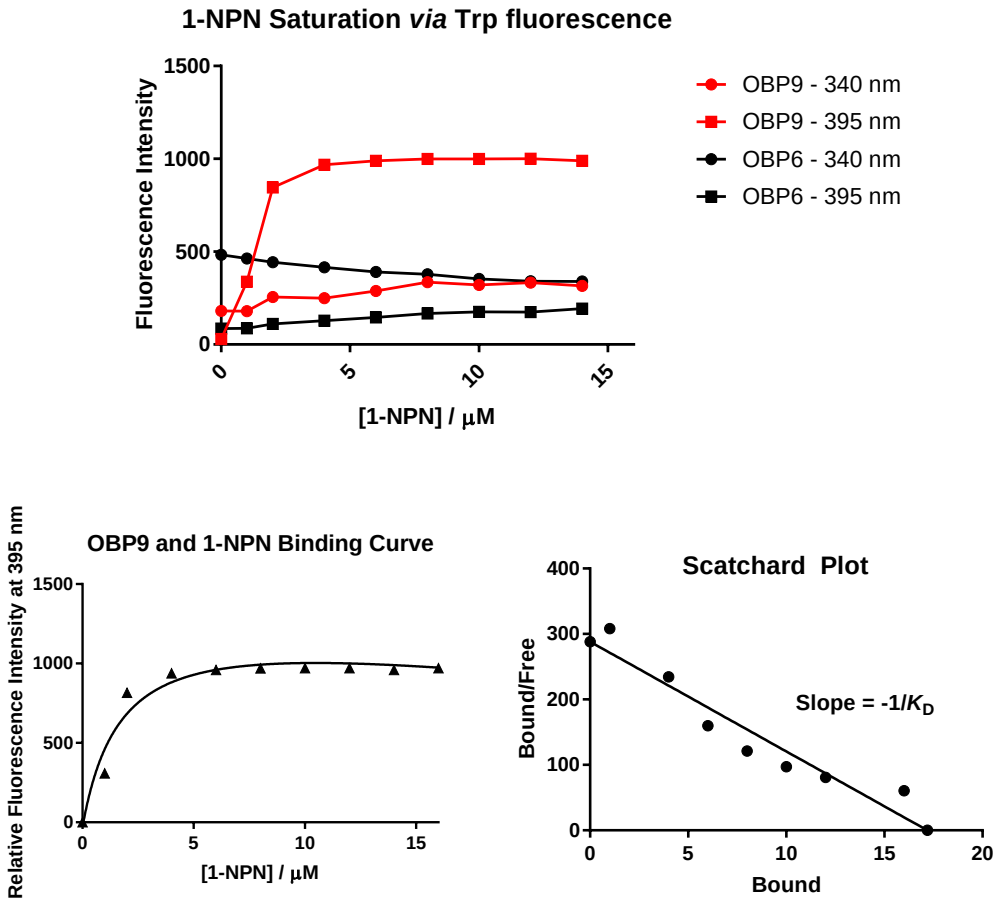


Figure 5.8: Comparison of tryptophan **54** interactions with 1-NPN **51** at 280 nm excitation. OBP9 (red) shows a significant increase in fluorescence at 395 nm – representing the interaction between 1-NPN **51** and Trp – compared to OBP6 (black). Both proteins display minor changes at 340 nm. The binding affinity (K_D) of 1-NPN **51** to OBP9 was calculated as 1.80 ± 0.72 µM.

The binding affinity of 1-NPN **51** to OBP9 was calculated, as in Figure 5.5, to be 1.80 ± 0.72 µM. This is a significant contrast from the results reported with the 337 nm fluorescence assay (K_D = 13.59 ± 1.58 µM). From this assay, it appears that 1-NPN **51** is a significantly stronger binder with OBP9 than from the 337 nm assays. As previously stated, fluorescence assays are notoriously unreliable with 1-NPN **51** as the fluorescence probe. It is clear that an assay based on the 337 nm excitation of 1-NPN **51** is not ideal for use with either of these proteins. For this preliminary calculation, only one repetition was completed (Figure 5.8).

The results of this assay complicate potential future experiment ideas. A competitive displacement assay using 1-NPN **51** could be used for OBP9, whereas a simple, competition free assay for OBP6 may be more appropriate. However, for comparable results, the same assay conditions should be used for each protein. As 1-NPN **51** appears to only bind weakly with OBP6, its presence in an assay may have negligible effect on the subsequent binding

results. To confirm this, preliminary binding assays were performed between OBP9 and (4a*S*,7*S*,7a*R*)-nepetalactone **6**, in addition to OBP6 with both (4a*S*,7*S*,7a*R*)-nepetalactone **6** and (*E*)-β-farnesene **17** (EBF). These assays observed the intrinsic fluorescence of tryptophan **54**, and two versions of the assay were performed – with and without 1-NPN **51** present in stoichiometric quantities (2 µM 1-NPN for 2 µM protein) (Figure 5.9; Table 5.1).

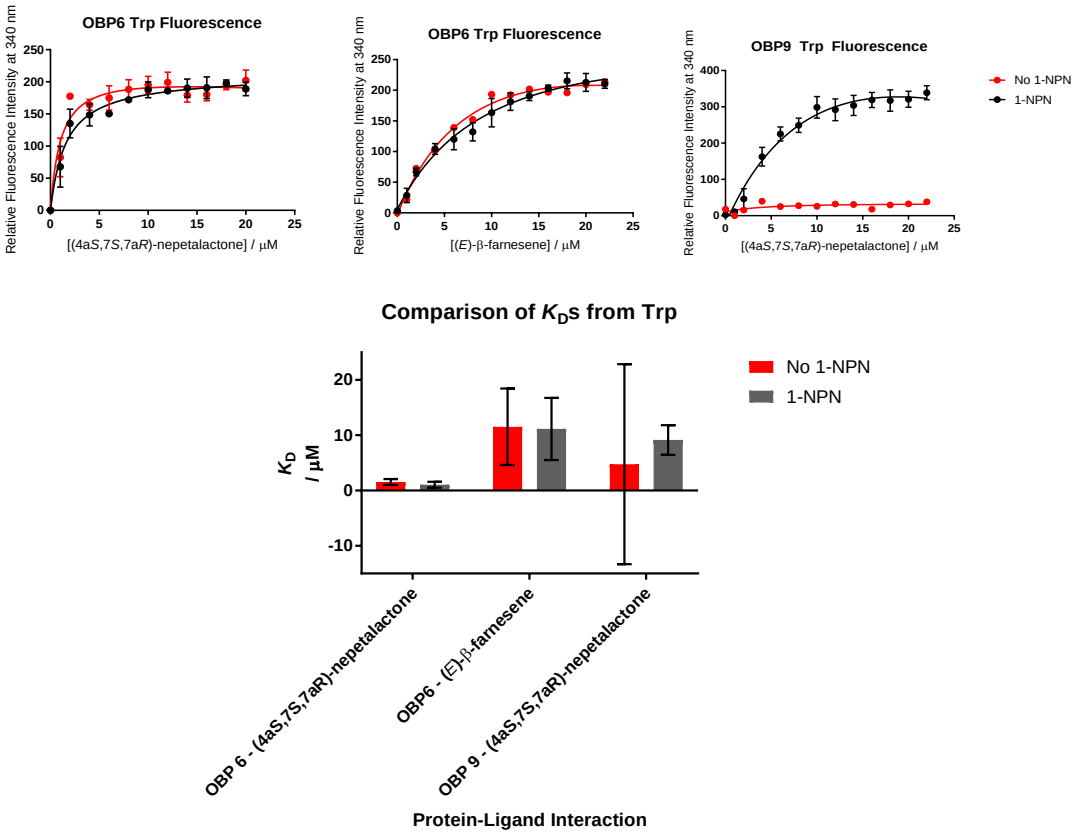


Figure 5.9: (i) Comparative preliminary binding assays of OBP6 in 25 mM Tris with (4a*S*,7*S*,7a*R*)-nepetalactone **6** and (*E*)-β-farnesene **17** and OBP9 in 25 mM Tris with (4a*S*,7*S*,7a*R*)-nepetalactone **6**, both with (black/grey) and without (red) an equivalent concentration of 1-NPN **51** present; (ii) a comparison of the preliminary calculated K_D's from the assay.

Table 5.1: Comparison of calculated K_D's from different preliminary binding assays with OBP6 and OBP9, both in 25 mM Tris.

Ligand	Protein	K_D / µM	
		No 1-NPN	With 2 µM 1-NPN
(4a*S*,7*S*,7a*R*)-nepetalactone **6**	OBP6	1.55 ± 0.53	1.06 ± 0.53
	OBP9	4.76 ± 18.08	9.15 ± 2.69
(*E*)-β-farnesene **17**	OBP6	11.53 ± 6.93	11.14 ± 5.61

When titrated with a ligand, the fluorescence intensity of the tryptophan **54** in OBP6 increased due to the ligand binding close to the Trp residue(s) and inducing changes in its environment. When OBP9 was titrated with a ligand with no 1-NPN **51** present, no change to the Trp fluorescence could be seen in comparison to an assay with 1-NPN **51**. No clear differences were observed between the two preliminary binding assays for OBP6; therefore, it was decided the presence of 1-NPN **51** would have negligible effect of the final binding assays to be conducted. For OBP9, the assay without the presence of 1-NPN **51** resulted in an unusable binding curve and a significant error on the calculated K_D, whereas an assay with 1-NPN **51** gave comparable results to OBP6. This is likely due to the much lower intensity of tryptophan **54** in OBP9 and its placement away from the binding pocket; the binding of a non-fluorescent ligand has limited effect on the electronic environment of Trp. 1-NPN **51** has a significantly higher effect on Trp fluorescence in contrast, due to its own fluorescence. The large error for the value of OBP9 occurs due to the calculation of K_D being made from an inaccurate binding curve, which was almost a straight line. The displacement of 1-NPN **51** will result, in theory, in the decrease of 395 nm fluorescence and potentially an increase in the fluorescence at 340 nm.

5.2.2 Ligand Binding Assays with 1-NPN

The observations of 1-NPN **51** binding with OBP6 and OBP9 complicate the options for binding studies, so to keep the experimental approach consistent an assay including 1-NPN **51** but observing tryptophan fluorescence was chosen. For OBP6, the presence of 1-NPN **51** had no effect on the subsequent calculated K_Ds, whereas 1-NPN **51** was required to conduct the assays with OBP9 (Figure 5.9).

To conduct the final assay, 2 µM of the proteins in 25 mM Tris was incubated with 2 µM 1-NPN **51**. Each ligand was then titrated into the mixture to final concentrations of 0-20 µM. Initially, OBP6 was titrated with the sex pheromone components, their respective enantiomers, alarm pheromone and plant volatile (*R/S*)-linalool. The fluorescence was measured from 290-400 nm after excitation at 280 nm (Figure 5.10).

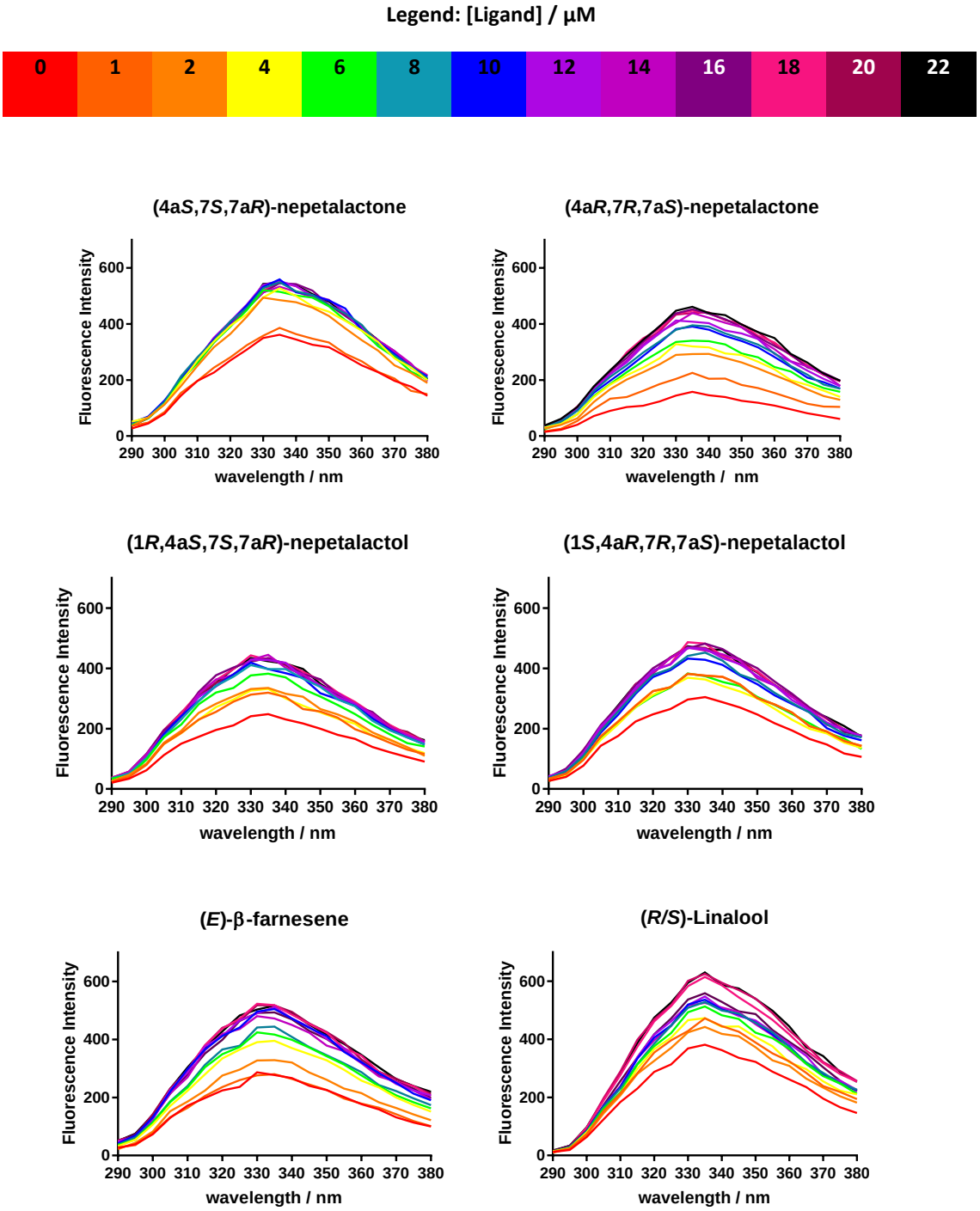
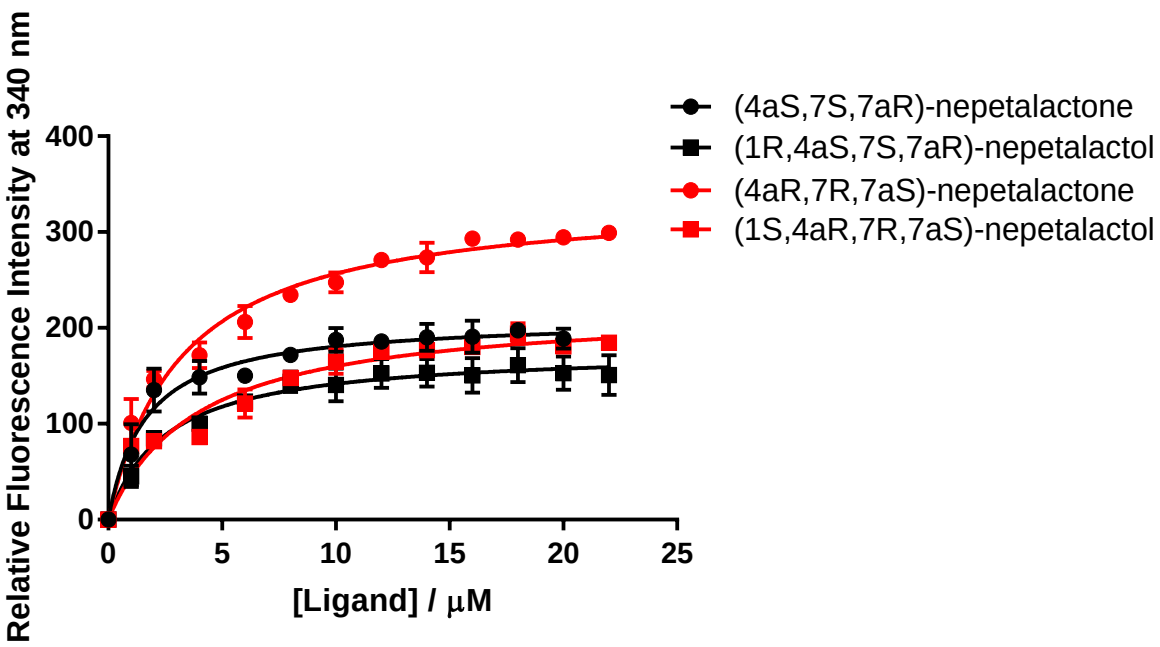


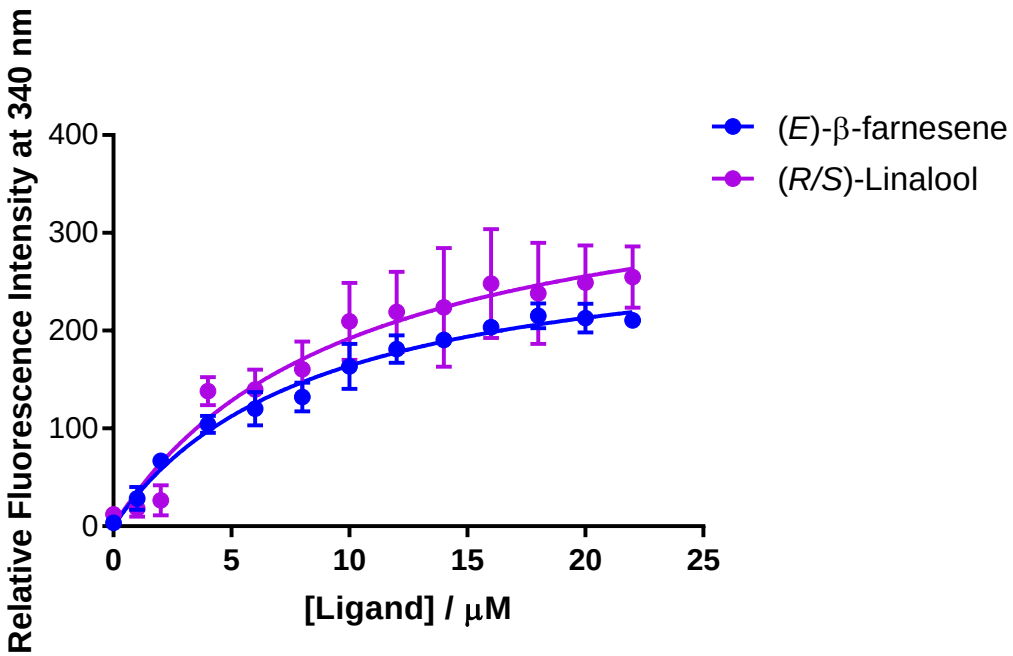
Figure 5.10: Tryptophan **54** fluorescence of OBP6 at 2 µM with 2 µM 1-NPN **51** titrated with various ligands to final concentrations of 0-22 µM. The lowest concentration of each ligand (0 µM) can be seen in red and the highest concentration (22 µM) in black. For each ligand, only one data set is presented for clarity.

An increase in fluorescence intensity at 340 nm was observed and the relative intensity plotted as a binding curve from which K_D values could be calculated (Figure 5.11).

(i)


Relative Fluorescence Intensity at 340 nm
400
300
200
100
0
0
5
10
15
20
25
[Ligand] / μM
(4aS,7S,7aR)-nepetalactone
(1R,4aS,7S,7aR)-nepetalactol
(4aR,7R,7aS)-nepetalactone
(1S,4aR,7R,7aS)-nepetalactol

(ii)


Relative Fluorescence Intensity at 340 nm
400
300
200
100
0
0
5
10
15
20
25
[Ligand] / μM
(E)-β-farnesene
(R/S)-Linalool

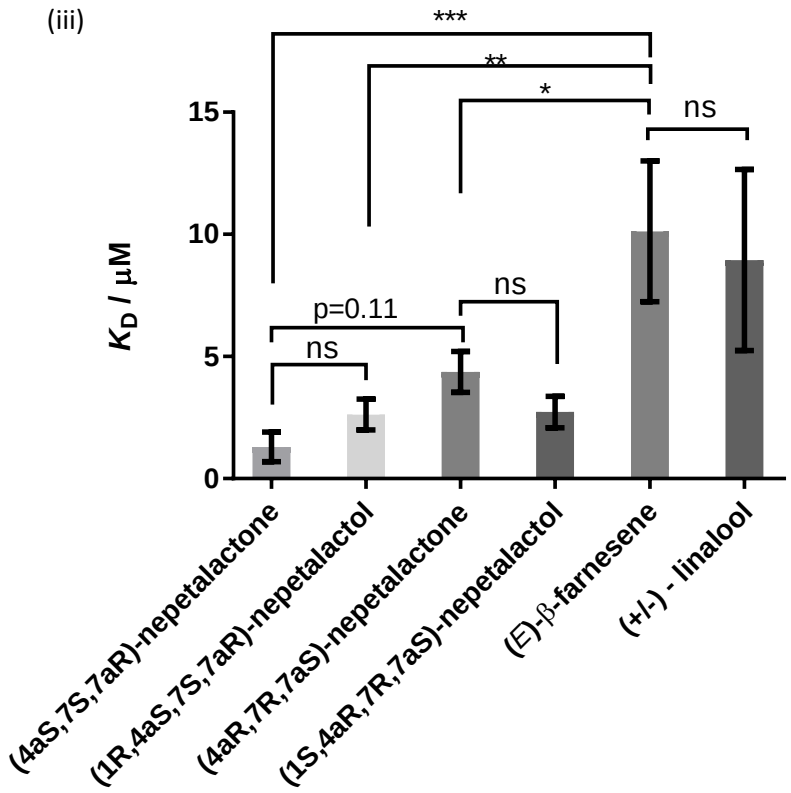


Figure 5.11: (i) Binding curves of OBP6 with the aphid sex pheromone components (black) and their respective enantiomers (red); (ii) Binding curves of OBP6 with the aphid alarm pheromone (blue) and linalool (purple); (iii) Calculated K_D values of OBP6 with various ligands.

Each ligand bound with OBP6, though a variety of binding affinities were observed (Figure 5.11, (iii)). The calculated K_D values from the fluorescence study were very close to the predicted values from *in silico* data (Table 5.1), suggesting that the model and predicted binding site were accurate. OBP6 was observed to bind the sex pheromone components with a much higher affinity versus the other tested ligands. The major difference between the two data sets was that in the *in silico* data, (4a*S*,7*S*,7a*R*)-nepetalactone **6** had the lowest affinity of all the sex pheromone components – the opposite is true from the *in vitro* tests.

Table 5.2: The calculated K_D values for OBP6 and various ligands from fluorescent binding studies versus the predicted K_Ds from *in silico* testing.
*For *in silico* (*R*)-linalool is reported, for fluorescence (*R/S*)-linalool is reported

Ligand	*In silico* K_D / µM	Fluorescence K_D / µM
(1*R*,4a*S*,7*S*,7a*R*)-nepetalactol **5**	2.4	2.62 ± 0.63
(4a*S*,7*S*,7a*R*)-nepetalactone **6**	3.1	1.30 ± 0.60
(1*S*,4a*R*,7*R*,7a*S*)-nepetalactol **32**	2.3	2.65 ± 0.80
(4a*R*,7*R*,7a*S*)-nepetalactone **33**	2.7	4.37 ± 0.81
(*E*)-β-farnesene **17**	12	10.12 ± 2.88
Linalool*	14	8.95 ± 3.71

Presently, no ligand-binding activity has been reported in the literature for OBP6 in aphids, despite OBP6, one of only two Plus-C OBPs found in aphids, being responsible for the second most abundant OBP mRNA in aphid antennae.[75] It is also a large OBP at 215 residues; it has been suggested that larger OBPs may have a longer C-terminal region, which can contribute to a conformational change by folding into the binding pocket when a ligand is bound.[77,102,132]

These studies suggest that OBP6 is responsible for binding the aphid sex pheromone components and respective enantiomers. The binding affinities calculated from these assays suggest a relatively strong binding affinity between the sex pheromone components and OBP6, compared to many OBP-ligand interactions; the initial binding data for the discrimination of the aphid alarm pheromone, EBF **17** with OBP3 described a weaker interaction (a K_D of approximately 6 µM) than observed between OBP6 and the sex pheromone components here.[102] Initial statistical analysis with a weighted analysis of variance (ANOVA), shows a significant difference between the K_D of each sex pheromone components, their corresponding enantiomers and the alternative ligands, EBF **17** and (*R/S*)-linalool (Figure 5.11; 8.4.3, Table 8.11). From initial analysis, no statistically significant difference was observed between the pairs of enantiomers, however, the K_D of naturally occurring (4a*S*,7*S*,7a*R*)-nepetalactone **6** appeared lower than it's respective enantiomer (4a*R*,7*R*,7a*S*)-nepetalactone **33** at initial glance (Figure 5.11; Table 5.1). An analysis, *via* a weighted t-test, of the differences between the K_D values for these enantiomers isolated from the data for all ligands gives a significant difference with a *p*-value < 0.001. Traditionally, statistical analysis of fluorescence data is rarely performed in the literature before a conclusion about the binding activity of a protein is reached.[102,125,186,191,192,196] This may be due to the difficulty in accounting for the associated errors from calculating K_D *via* the non-linear

regression analysis test – in this case a weighted approach was taken (8.4.3) – however, it is more likely due to the nature of the study and OBPs themselves. OBPs are notoriously promiscuous binders[112], and it is therefore unlikely a significantly weak binder will be observed, rather ligands that bind with a range of affinities. Additionally, fluorescence is a technique that relies on a high level of precision to avoid errors, including ensuring the sample is completely absent of light during measurements, the excitation wavelength is accurate and conducting the assay swiftly with consistent timings and temperatures.

Ligands with a weaker binding affinity to the protein appeared to have a much bigger range of observed fluorescence intensities and therefore larger errors on the final K_D value. Weak binding activity will cause the ligand to bind and release frequently over the course of the fluorescence measurements. Furthermore, EBF **17** has previously been shown to exhibit unusual binding activity due to its ability to form micelles around the fluorescent probe 1-NPN **51**.[102]

OBP9 was tested in an identical assay to OBP6 against the same ligands, with the exception of the ligand (*R/S*)-linalool; OBP9 was titrated with the sex pheromone components, their respective enantiomers and the alarm pheromone (Figure 5.12).

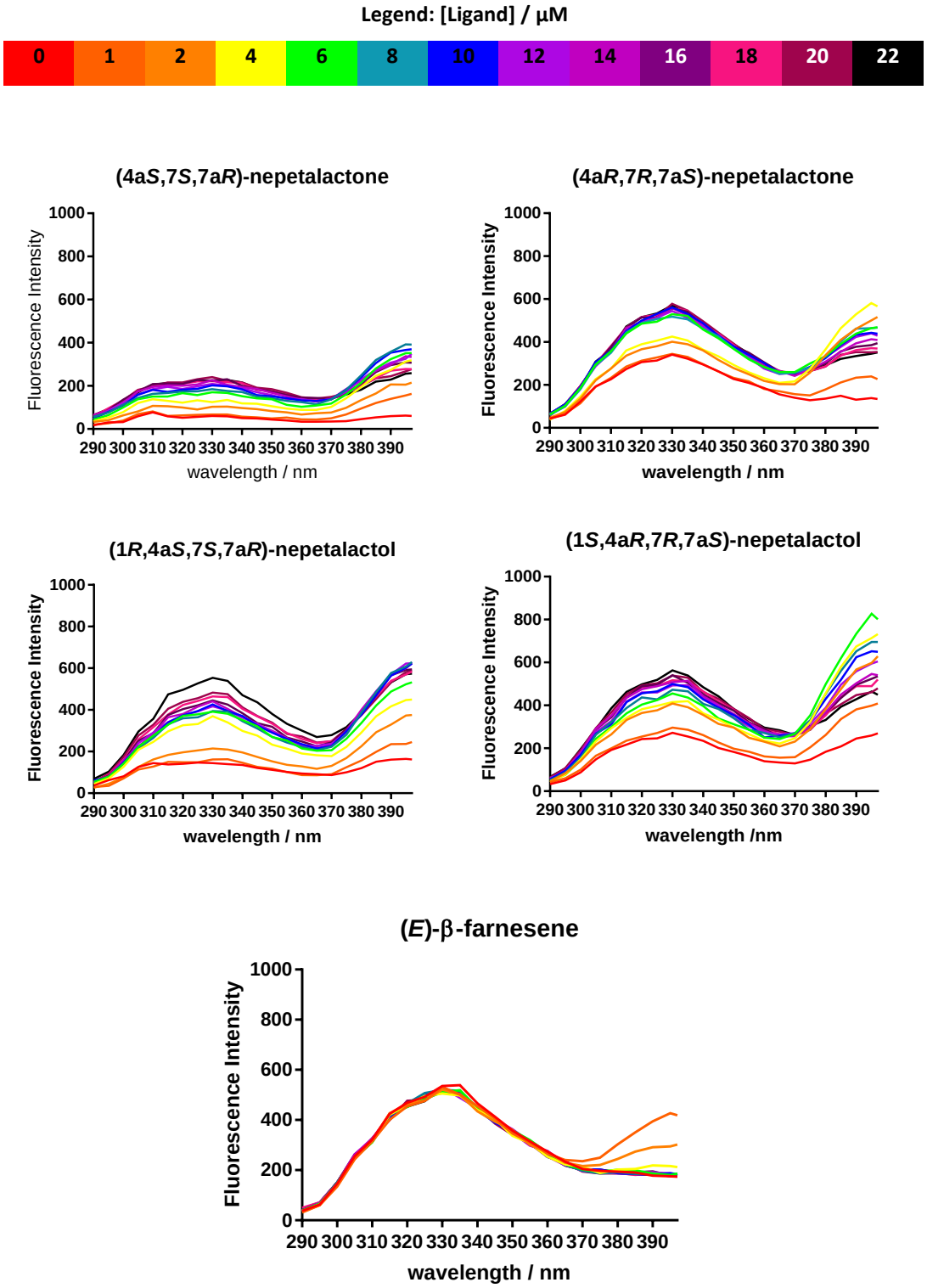


Figure 5.12: Tryptophan **54** fluorescence of OBP9 at 2 µM with 2 µM 1-NPN **51** titrated with various ligands to final concentrations of 0-22 µM. The lowest concentration of each ligand (0 µM) can be seen in red and the highest concentration (22 µM) in black. For each ligand, only one data set is presented for clarity.

OBP9 was expected to increase in fluorescence at 340 nm and decrease in fluorescence at 395 nm, as 1-NPN **51** is out-competed by the ligand. However, though an increase was seen

at 340 nm in most cases, no decrease in fluorescence at 395 nm was observed. OBP9 has a strong binding affinity to 1-NPN **51**, it is reasonable that these ligands may not have been able to competitively displace 1-NPN **51**. The fluorescence intensity at 395 nm appeared to increase as the ligand was titrated in – this could be linked to a quenching or the formation of micelles in the case of EBF **17**. This result is consistent with the observed *in silico* results (Table 5.2), where OBP9 had very weak predicted affinities to all the ligands of interest.

Table 5.3: The calculated K_D values for OBP9 and various ligands from fluorescent binding studies versus the predicted K_Ds from *in silico* testing.

Ligand	*In silico* K_D / µM	Fluorescence K_D / µM
(1*R*,4a*S*,7*S*,7a*R*)-nepetalactol **5**	130	5.74 ± 1.71
(4a*S*,7*S*,7a*R*)-nepetalactone **6**	57	6.49 ± 1.58
(1*S*,4a*R*,7*R*,7a*S*)-nepetalactol **32**	97	6.29 ± 1.99
(4a*R*,7*R*,7a*S*)-nepetalactone **33**	66	9.68 ± 4.57
(*E*)-β-farnesene **17**	190	8.32 ± 2.67

The *in silico* binding interactions between OBP9 and the various tested ligands have a very low affinity, and recorded fluorescence data contains large errors and unexpected patterns of fluorescence. It is unlikely that any of these calculated K_D values are accurate and can be used for comparison. 1-NPN **51** binds with such a high affinity to OBP9, making it difficult to displace, and intrinsic fluorescence of OBP9 is too weak to show observable change when a ligand is introduced in the absence of a fluorescent probe, it may be difficult to develop an assay that effectively demonstrates binding with OBP9. I*n silico* data predicted a strong interaction between (*Z*)-jasmone **30** and OBP9, with a predicted binding affinity of 2.31 µM. Given the success of the OBP6 binding assays, this ligand could be tested in future work.

5.2.3 Ligand Binding Assay with 1,8-ANS

As OBP6 did not appear to bind 1-NPN **51**, a different fluorescent probe, 1,8-ANS **52**, was used to undergo binding assays. 1,8-ANS **52** has a higher solubility and lower fluorescence than 1-NPN **51** and can therefore be titrated to much higher concentrations. 1,8-ANS **52** can be excited at 380 nm and emits with a maximum fluorescence at approximately 510 nm (Figure 5.13).

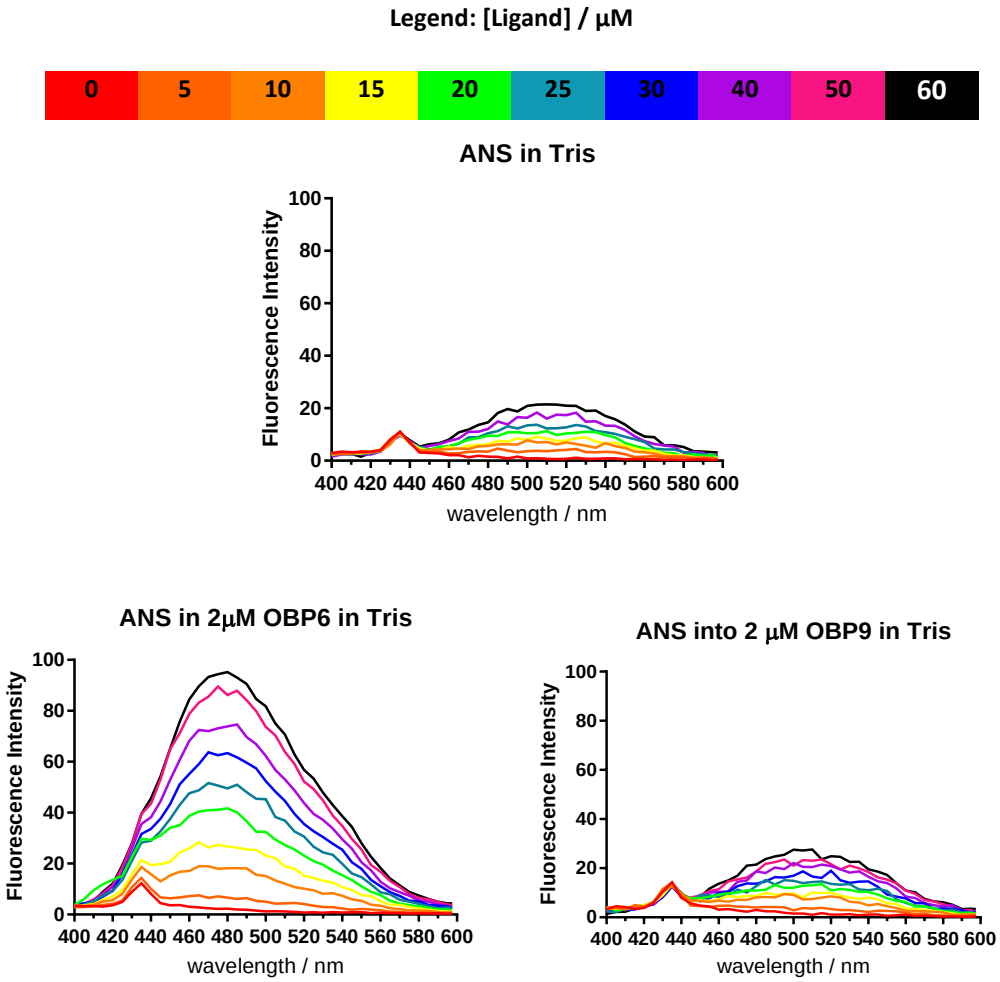


Figure 5.13: Fluorescence of 1,8-ANS **52** (excitation at 380 nm), titrated to concentrations of 0-60 µM. 1,8-ANS **52** was titrated into 25 mM Tris, 2 µM OBP6 in 25 mM Tris and 2 µM OBP9 in 25 mM Tris.

When titrated with OBP6, a blue shift to a maximum at 475 nm is seen, with a significant increase in fluorescence intensity. This dramatic change in fluorescence suggests that OBP6 binds 1,8-ANS **52**. Conversely, the fluorescence spectra for the titration of 1,8-ANS **52** into OBP9 appears to show no interaction. To confirm both these results, the fluorescence intensity at both 510 nm (free 1,8-ANS) and 475 nm (bound 1,8-ANS) were plotted (Figure 5.14). Due to the high concentration of 1,8-ANS **52** used, a logarithmic curve is plotted.

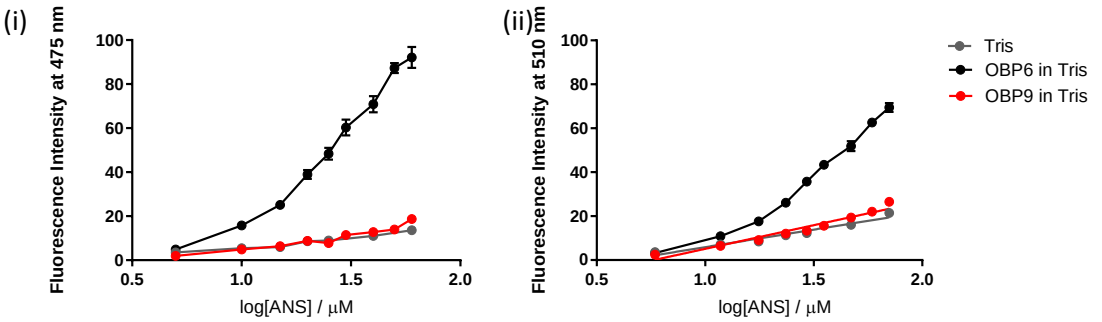


Figure 5.14: Fluorescence intensity of 1,8-ANS **52** titration (excitation of 380 nm) into 25 mM Tris (grey), 2µM OBP6 in 25 mM Tris (black) and 2µM OBP9 in 25 mM Tris (red), (i) at 475 nm (ii) at 510 nm.

When 1,8-ANS **52** is titrated in to OBP6, an increase intensity is seen, particularly at 475 nm. From these results, OBP9 does not interact with 1,8-ANS **52**. OBP6 appears to bind 1,8-ANS **52** with a very low affinity; titrations were performed up to 60 µM 1,8-ANS **52**, and full saturation was not achieved. Due to this, an accurate binding affinity could not be calculated, and a titration to a higher concentration of 1,8-ANS **52** is required. 1,8-ANS **52** is soluble in water up to 0.015 mgml^{-1}, or approximately 50 µM.[197] Although some ligand will be solubilised by the protein, the insolubility makes it difficult to generate binding curves at these levels of concentration.

In order to test the binding activity of OBP6, a 'reverse' binding assay was performed.[117] OBP6 was incubated with 2 µM (4a*S*,7*S*,7a*R*)-nepetalactone **6**, before titrating 1,8-ANS **52** into the mixture and measuring the change in fluorescence intensity at 475 nm (Figure 5.15).

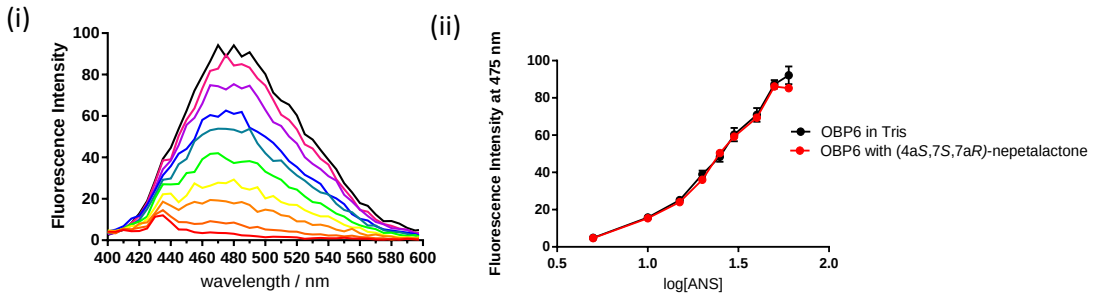


Figure 5.15: (i) Fluorescence of 1,8-ANS **52** (excitation at 380 nm), titrated to concentrations of 0-60 µM. 1,8-ANS **52** was titrated into 2 µM OBP6 with 2 µM (4a*S*,7*S*,7a*R*)-nepetalactone **6** in 25 mM Tris; (ii) Fluorescence intensity of 1,8-ANS **52** titrated into OBP6 and OBP6 with (4a*S*,7*S*,7a*R*)-nepetalactone **6**.

No difference was observed between OBP6 and OBP6 incubated with (4a*S*,7*S*,7a*R*)-nepetalactone **6**. Though fluorescence assays explored in this chapter show that OBP6 binds both (4a*S*,7*S*,7a*R*)-nepetalactone **6** and 1,8-ANS **52**, it appears that (4a*S*,7*S*,7a*R*)-

nepetalactone **6** was unable to displace 1,8-ANS **52** in this assay. To explore this result further, 1,8-ANS **52** was screened against a homology model of OBP6 using AutoDock (as Chapter 4). The docking results show that 1,8-ANS **52** binds to OBP6 with a low affinity (K_i of 24.70 µM), but in a completely different site to (4a*S*,7*S*,7a*R*)-nepetalactone **6** (Figure 5.16). This suggests (4a*S*,7*S*,7a*R*)-nepetalactone **6** was unable to displace 1,8-ANS **52** from this specific site, and instead bound in the active binding pocket, resulting in both ligands being bound to OBP6 simultaneously.

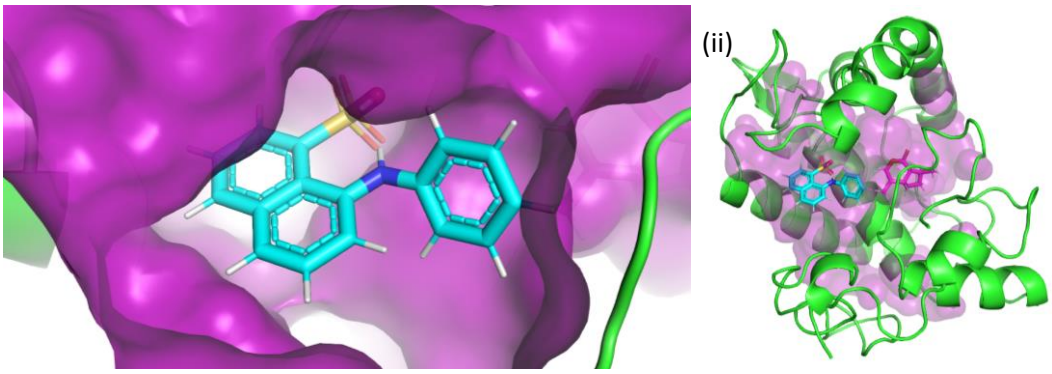


Figure 5.16: (i) 1,8-ANS **52** (blue) bound to OBP6 (green, pocket purple); (ii) Both 1,8-ANS **52** (blue) and (4a*S*,7*S*,7a*R*)-nepetalactone **6** (pink) bound to OBP6 (green, pocket purple) simultaneously.

5.2.4 Ligand Binding Assay without a Fluorescent Probe

Due to the success of the intrinsic assay with the fluorescent probe 1-NPN **51** and OBP6, a probe-free fluorescent assay was used to explore OBP6-ligand binding. To conduct the intrinsic assay, each ligand was then titrated into a solution of OBP6 (2 µM in 25 mM Tris) to final concentrations of 0-20 µM. OBP6 was titrated with the sex pheromone components, their respective enantiomers and the alarm pheromone. The fluorescence was measured from 290-400 nm after excitation at 280 nm (Figure 5.17).

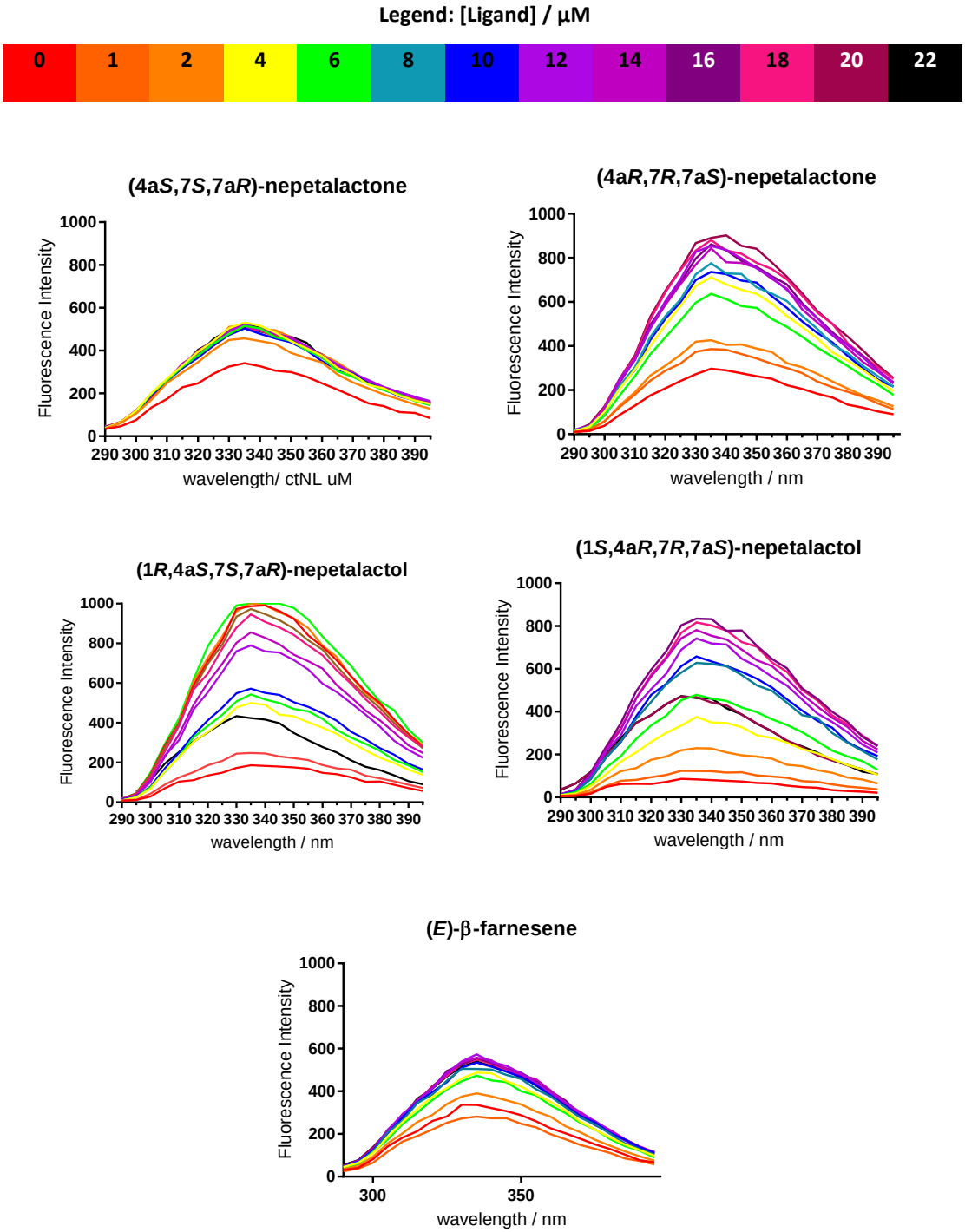


Figure 5.17: Tryptophan **54** fluorescence of OBP6 at 2 µM titrated with various ligands to final concentrations of 0-22 µM. The lowest concentration of each ligand (0 µM) can be seen in red and the highest concentration (22 µM) in black. For each ligand, only one data set is presented for clarity.

An increase in fluorescence intensity at 340 nm was observed and the relative intensity plotted as a binding curve from which K_D values could be calculated (Figure 5.18).

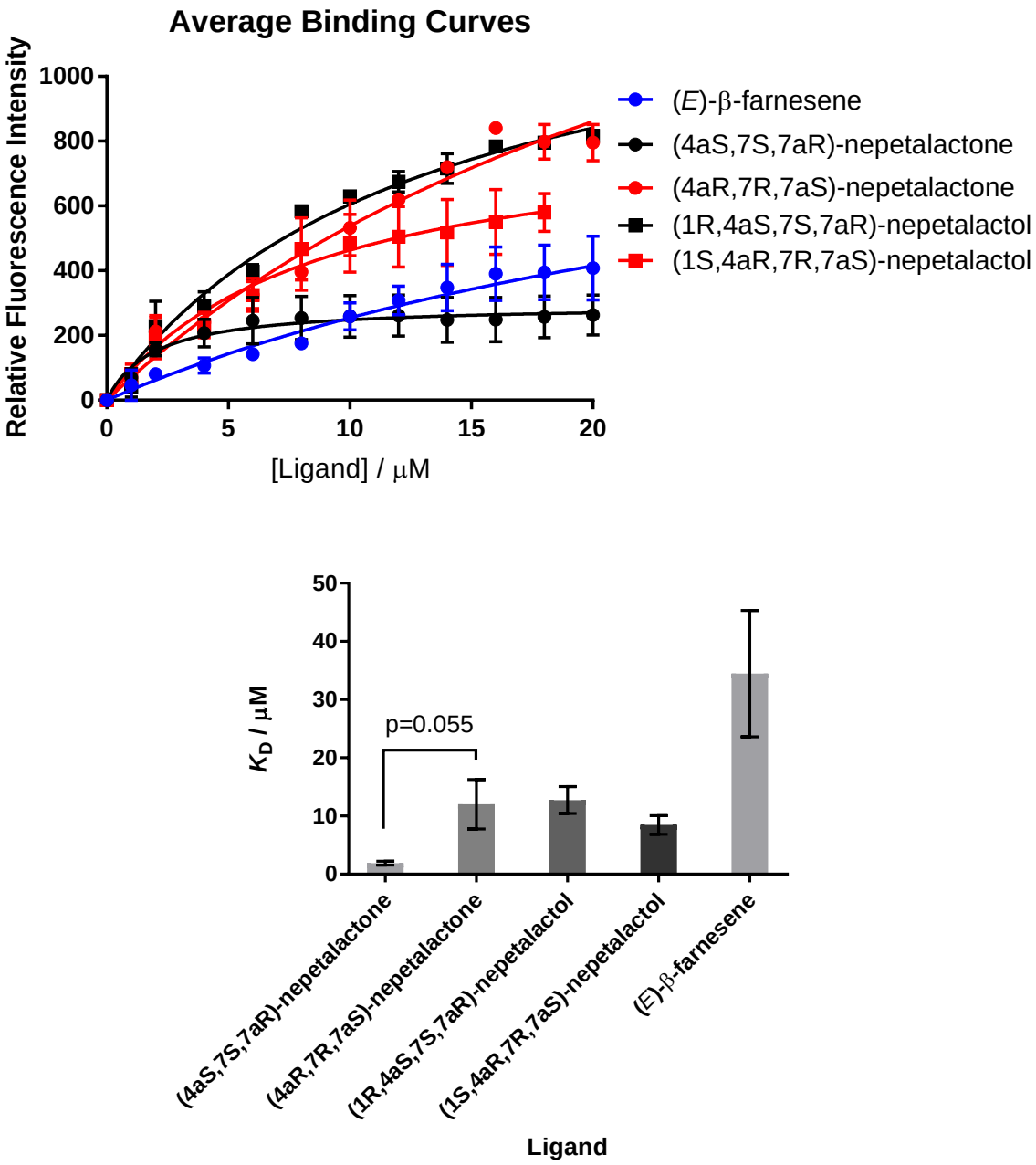


Figure 5.18: (i) Binding curves of OBP6 with the aphid sex pheromone components (black) and their respective enantiomers (red); (ii) Binding curves of OBP6 with the aphid alarm pheromone (blue) and linalool (purple); (iii) Calculated K_D values of OBP6 with various ligands.

As with the 1-NPN **51** fluorescent assay, (4a*S*,7*S*,7a*R*)-nepetalactone **6** was seen to have the lowest K_D. However, there were no statistically significant differences observed between the K_Ds of each ligand, with the lowest *p*-value, 0.055, calculated for the interaction between (4a*S*,7*S*,7a*R*)-nepetalactone **6** and the other ligands. This p-value is close to a significant level of 95% (94.5% significance) and highlights the much lower K_D observed for (4a*S*,7*S*,7a*R*)-

nepetalactone **6**. The assay with no fluorescent probe produced significantly higher errors than with 1-NPN **51**.

Overall, comparison of all the successful binding assay results (Table 5.4) show clear consistent trends. (4a*S*,7*S*,7a*R*)-Nepetalactone **6** consistently binds to OBP6 with a relatively high affinity (average K_D of 1.56 µM), and EBF **17** with a low affinity (average K_D of 12.59 µM). OBP9 does not appear to show any discrimination ability, and ligands which bind with a low affinity in general have a significantly higher associated error. Further investigation into the interaction of OBP6 with the sex pheromone components is required to fully explore any possibilities of discrimination capability of the protein between two different components or enantiomers.

Table 5.4: The calculated K_D values for OBP6 and OBP9 and various ligands from different binding assays.

NA = not available or not measured

Ligand	OBP6		OBP9
	1-NPN Assay K_D / µM	**Fluorescent Probe-Free Assay K_D / µM**	**1-NPN Assay K_D / µM**
(1*R*,4a*S*,7*S*,7a*R*)-nepetalactol **5**	2.62 ± 0.63	12.74 ± 2.31	5.74 ± 1.71
(4a*S*,7*S*,7a*R*)-nepetalactone **6**	1.30 ± 0.60	1.90 ± 0.35	6.49 ± 1.58
(1*S*,4a*R*,7*R*,7a*S*)-nepetalactol **32**	2.65 ± 0.80	8.46 ± 1.62	6.29 ± 1.99
(4a*R*,7*R*,7a*S*)-nepetalactone **33**	4.37 ± 0.81	12.01 ± 4.24	9.68 ± 4.57
(*E*)-β-farnesene **17**	10.12 ± 2.88	34.47 ± 10.85	8.32 ± 2.67
(*R*/S)-Linalool	8.95 ± 3.71	NA	NA

To further confirm the results of this chapter, knockdown or knockout in studies could be undertaken and behavioural studies subsequently performed. The removal of OBP6 may affect the aphid's ability to perceive the attractive sex pheromone. Previously, knockdown studies were successfully performed in the pea aphid, *Acyrthosiphon pisum*, based on *in vitro* data on the interactions of OBP3 and OBP7.[99,102,104,125] These studies demonstrated that both OBP3 and OBP7 were required for successful perception of the alarm pheromone, in addition to OR5. As with OBP3 and OBP7, OBP6 may also have a sister protein which is equally

responsible for the binding of the sex pheromone components, potentially adding and extra layer of discrimination.[102,104] *In silico* data do not identify any other potential candidates for sex pheromone binding, but highly expressed proteins such as OBP8 should also be investigated.[75]

5.2.5 Thermostability Assay

Intrinsically fluorescent amino acids, such as tryptophan **54** and tyrosine **53**, can be used to calculate thermostability of a protein. Tryptophan **54** (Figure 5.2) can be excited at 280 nm and will generally emit at 348 nm, whereas tyrosine **53** (Figure 5.2) can be excited at 275nm and emits at approximately 300 nm. When a fluorescent residue is observed within a protein, a shift in fluorescence can indicate the position within the protein has changed. Changes in the local electronic field, due to different surrounding amino acids and ligands, can alter the spread of electron density within the system, and therefore result in shifts in fluorescence wavelength.[187] For example, in folded proteins, such as OBP6 and OBP9, a blue shift to approximately 330 nm is observed due to the tryptophan's position within the tertiary structure of the protein, consistent with other insect OBPs and GOBPs.[115]

Protein unfolding can be determined by observing the shift from emission at 330 nm for the folded protein to approximately 350 nm for the unfolded protein. By monitoring the ratio of 330/350 nm fluorescence, a protein unfolding curve can be plotted and the midpoint, representing inflection temperature (Ti), can be calculated (Figure 5.19). The Ti can be seen more clearly by plotting the first derivative of the plot (Figure 5.19).

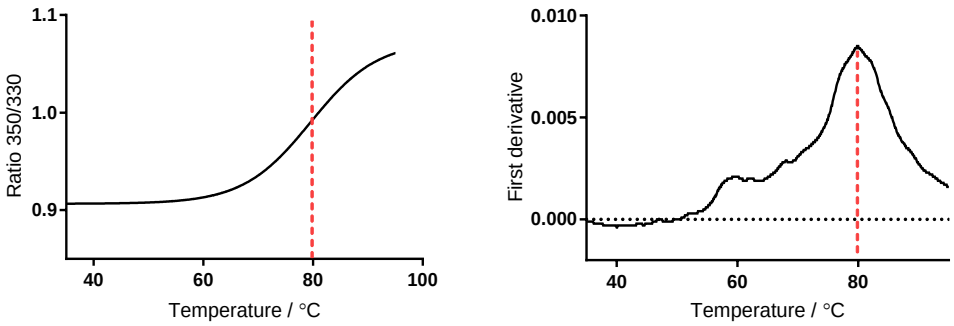


Figure 5.19: Thermostability assay of OBP6 monitored using fluorescence. The unfolding curve and first derivative of the unfolding curve can both be used to calculate the inflection temperature (Ti) to be 79.9 °C.

An inflection temperature of 79.9 °C was calculated for OBP6, indicating that OBP6 unfolds due to temperature impacts at approximately this temperature. This is a reasonably high Ti,

but not unexpected for an OBP. OBPs are notoriously thermally stable due to their multiple disulfide bonds formed between highly conserved cysteines.[101]

This form of assay could be used in future to observe ligand-binding properties of the protein by determining if the presence of a particular ligand thermally stabilises the thermal unfolding of the protein.

5.3 CONCLUSIONS

The results in this chapter suggest that OBP6 binds the sex pheromone. There was some evidence to suggest that OBP6 bound naturally-occurring (4a*S*,7*S*,7a*R*)-nepetalactone **6** with a higher affinity than the other sex pheromone components, including the non-naturally occurring enantiomer (4a*R*,7*R*,7a*S*)-nepetalactone **33**. The results suggest that odorant-binding proteins may play a role in enantiomeric discrimination that has previously been observed in behavioural bioassays. If odorant-binding proteins confer a conformational change when bound to an active ligand, the stereochemical difference between the ligands may result in more global changes in the protein 3-dimensional structure. If an OBP-ligand complex activates odorant receptors, enantiomeric discrimination by the OBPs may still be possible.

Though successful ligand-binding assessment for OBP6 has been carried out, other binding-methods are required to confirm the activity of this proteins. OBP9 has no confirmed binding activity from these assays – other assay types may be critical in observing any activity for this protein. Other ligands could also be tested in future, including other aphid semiochemicals and those which have shown low or no activity in previous aphid OBP binding studies, such as ethyl benzoate and octyl benzoate, shown to have no activity with OBP1,3 and 8 in previous work.[102] A thermostability assay has also been performed using fluorescent techniques, confirming OBP6 as a thermally stable protein.

The results of these experiments should be further explored using different techniques, including more intricate ligand-binding analyses in the form of mass spectrometry and nuclear magnetic resonance (NMR) spectroscopy.

Chapter 6: Alternative Techniques for the Study of *A. pisum* Odorant-Binding Protein Binding Activity

6.1 INTRODUCTION

Insect odorant-binding proteins (OBPs) and the binding of their respective ligands are usually studied using fluorescence-based assays (Chapter 5).[102,112,117,118,186] However, other techniques, such as mass spectrometry and nuclear magnetic resonance (NMR) spectroscopy, provide more comprehensive information about binding specifics, such as key residue interactions, and protein structure, in addition to the kinetic data that can also be obtained with fluorescence.[111,117,122,195,198]

Mass spectrometry, as previously introduced in Chapter 3, can be a useful tool for understanding not only the mass and subsequent make-up of proteins, but also their structure, kinetics and binding activity.[122,155] Whereas mass spectra of denatured proteins can provide some information about a protein, native electrospray ionisation mass spectrometry (native-ESI) spectrometry allows for direct observation of the formation of protein-ligand complexes.[122,123] This involves a soft ionisation mass spectrometry technique, where the protein remains in its native 3-dimensional structure (non-denatured).[122,123] If a ligand is bound to the protein, data representing the protein-ligand complex will be observed.[122,123]

Another approach for investigating protein-ligand interactions using mass spectrometry is to observe how the presence of a ligand stabilises the unfolding of a protein.[194] There are multiple methods which provide the means to do this, including the Stability of Proteins by Rates of Oxidation (SPROX) method, developed by West *et al.*[194] SPROX involves oxidising the sulfur atom of methionine residues in a protein at different stages of unfolding (Figure 6.1).[194] By observing the extent of methionine oxidations, and subsequent mass-shift in the spectra, the specific concentration of a denaturant at which a protein becomes unfolded can be measured.[194] When a ligand is bound to a protein, the overall structure is stabilised resulting in an increase in the denaturant concentration required to fully unfold the protein.

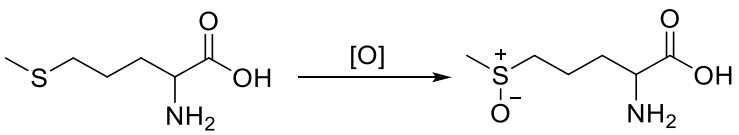


Figure 6.1: Oxidation of the sulfur atom in the amino acid methionine

To perform SPROX, a protein is exposed to a denaturant, such as guanidinium chloride, and the sulfur atoms of the methionine oxidised with hydrogen peroxide (Figure 6.1). Successful oxidation can be observed by the appearance of new species in the mass spectra, with the number of new species dependent on the number of methionines present in the protein and therefore number of oxidations that have occurred. For example, a protein with three methionines present may result in up to four species – the native protein, in addition to the singly, doubly, and triply oxidised protein species. From this mass spectra, an average number of oxidations for the protein can be calculated across different denaturant concentrations. The average number of oxidations will not only vary with the number of methionines present, but the positioning of each of these methionines in the protein; residues found deeper within the protein – in the binding pocket, for example – are less likely to be oxidised than the more exposed residues found on the protein's surface. If methionine residues are present both on the surface and deep within the pocket of the protein, the folding or unfolding of a protein can easily be monitored by observing the average oxidation of the protein over a range of denaturant concentrations.

NMR spectroscopy was initially developed to study small molecular structures. However, advancements in the technique have led to new applications and an ability to analyse more complex, macromolecular structures, including proteins. NMR spectroscopy is limited by the availability of NMR-active isotopes, which include ^{1}H, ^{13}C and ^{15}N.[199] Though ^{1}H is highly abundant, both ^{13}C and ^{15}N have a low natural abundance – 1.1% and 0.4% respectively. Both small molecules and proteins can be synthesised with a higher abundance of NMR-active isotopes (isotopically labelled). Isotopically labelled proteins are produced by generating recombinant bacteria that synthesise the specific protein of interest and feeding them with ^{15}N or ^{13}C-labelled media, allowing the bacteria to incorporate these isotopes into the final protein.[199] 2-Dimensional and 3-dimensional NMR spectroscopy can be performed to elucidate the structure of a proteins or structural changes that occur when a ligand is bound.[111,117,193,195,198,200] . For highly soluble proteins, such as OBPs, the use of NMR to study structure and function appears to be a suitable method. Evidence for a conformational shift in insect OBPs when binding, particularly the movement of the C-terminus tail, has been previously observed using NMR spectroscopy.[108,117,195] Though there are many different NMR experiments that can be performed to observe the structure and ligand binding of the proteins, however, most of these studies require double-labelling of the protein with NMR-active isotopes (^{15}N and ^{13}C), although limited binding studies can be performed with a single labelled (^{15}N) protein *via* ^{1}H-^{15}N HSQC NMR.

NMR techniques can be also be utilised to observe specific binding interactions between a ligand and protein without isotopic labelling. Saturation transfer difference (STD) is a technique that utilises the Nuclear Overhauser Effect (NOE) to study protein-ligand binding interactions (Figure 6.2).[201,202] NOE occurs when a nucleus is magnetically excited, with a neighbouring atom at equilibrium, resulting in relaxation between the two. The coupling effect occurs through physical space, rather than traditionally thought of through-bound coupling. Therefore, the smaller the physical distance between the nuclei, the bigger the expected change. With regards to protein-ligand binding experiments, by selectively irradiating the protein, any protons from ligands close (bound) to the protein can be observed (Figure 6.2; (i)). By calculating the difference between an initial NMR (off-resonance) spectrum, and a spectrum in which the sample has undergone selective irradiation (on-resonance), saturation transfer difference (STD) can be observed. Ligands that interact with the protein being affected by a saturation transfer, and appear in the final STD spectrum (Figure 6.2; (ii)).[201]

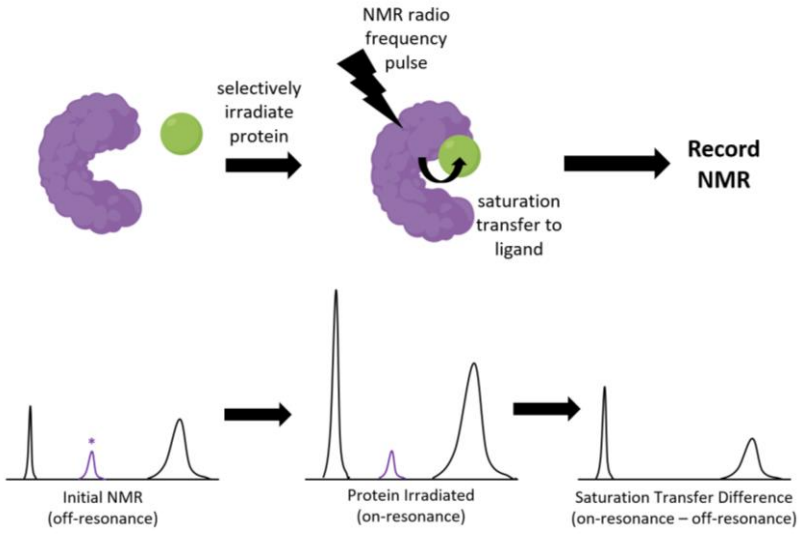


Figure 6.2: Saturation transfer difference (STD) nuclear magnetic resonance (NMR) spectroscopy. (i) The protein is selectively irradiated, and the saturation is subsequently transferred to any bound ligands; (ii) Predicted NMR spectra. Two peaks appear in the STD spectra with changed intensity, whilst one peak (*, purple) disappears, demonstrating its lack of interaction with the protein.

In many ligand-binding assays, the ligands bind and release freely from the protein over the course of the experiment. This is generally a benefit as it allows for kinetic data to be observed, however, for weak binders like OBPs, it can make obtaining accurate binding data more difficult. Furthermore, OBPs are highly soluble proteins that exist in the aqueous sensillum lymph of an insect antenna.[101,112] If they play a role in olfaction, it may involve

retrieving poorly-soluble ligands from the air surrounding the insect antennae *via* pores and solubilising them into the sensillum lymph.[112] By employing a biphasic assay system comprising of two distinct phases, i.e. aqueous and non-aqueous, the ability of OBPs to bind ligands across phases could be further explored. There are some limitations to this type of method, specifically accounting for the rate of diffusion of ligands across the layers, accurately monitoring changes in the amount of ligand present in each layer and choosing an appropriate non-aqueous layer. Gas chromatography analysis allows for the measurement of very low concentrations of compounds in a volatile solvent, and a method using gas chromatography to monitor concentration changes has been previous described. [115]

This chapter aims to explore alternative methodology for studying the interactions of insect OBPs and their respective ligands. Binding data from previous chapters, with specific focus on the elucidated interaction of OBP6 and the sex pheromone components (1*R*,4a*S*,7*S*,7a*R*)-nepetalactol **5** and (4a*S*,7*S*,7a*R*)-nepetalactone **6**, will be tested using a variety of different experiments. A range of techniques will be used including mass spectrometry, NMR spectroscopy and a two-phase/biphasic gas-chromatography based assay.

6.2 RESULTS & DISCUSSION

6.2.1 Stability of Proteins by Rates of Oxidation (SPROX)

A SPROX assay was preliminarily performed with OBP9. OBP9 has six methionines, and oxidation of the folded and unfolded state revealed three methionines that were more easily oxidised, with three that required protein denaturation to be oxidised (Figure 6.3)

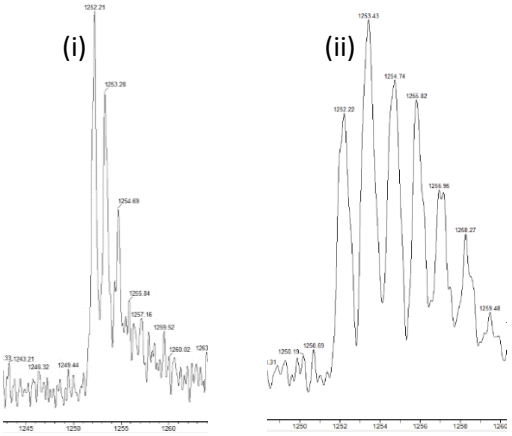


Figure 6.3: Mass-spectra displaying one charge state (+13) of oxidised OBP9. Spectra (i) shows two methionine oxidations that have occurred at a denaturant concentration of 0 M, whereas (ii) shows that 6 oxidations have occurred at a denaturant concentration of 5 M.

The average number of oxidations of a protein was calculated using the mass spectra by finding the average of the oxidations observed based on their abundance (quantified by peak area). The average number of oxidations was then plotted against the concentration of denaturant to generate an 'unfolding' curve, or 'SPROX' curve (Figure 6.4). The curve is fitted with a sigmoidal trend line *via* a non-linear regression, and the midpoint or $C_{1/2}$ calculated, with the average number of oxidations being inversely proportional to '$1 + e^{-[denaturant]}$'.[194] $C_{1/2}$ represents the concentration of denaturant where approximately half of the protein is unfolded.

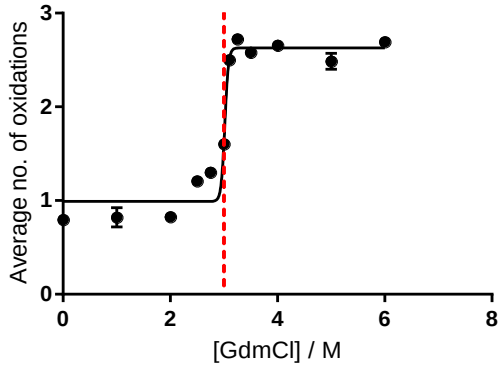


Figure 6.4: SPROX graph of OBP9 (concentration of denaturant, guanidinium chloride, vs average number of oxidations) and sigmoidal curve fitted. The midpoint ($C_{1/2}$) is 3.0M guanidinium chloride, represented with a dashed red line.

The $C_{1/2}$ for OBP9 was calculated to be 3.0 M of guanidinium chloride. After the initial curve generation, changes due to the introduction of ligand were observed. The presence of ligand should stabilise the protein's 3D structure, increasing the concentration of denaturant required to unfold to protein and subsequently shifting the curve to the right with an increase in $C_{1/2}$. The tighter the ligand binds, the more the complex should be stabilised. SPROX curves can therefore be used as a measure of ligand-binding affinity, and relevant kinetic data, such as a ΔG for the interaction, can be calculated.

OBP9 was tested by incubating with two different ligands, (4a*S*,7*S*,7a*R*)-nepetalactone **6** and (1*R*,4a*S*,7*S*,7a*R*)-nepetalactol **5**. The mass spectra for both ligands gave unclear results (Figure 6.5); on the addition of the ligand the noise on the spectra significantly increased, and the peaks for different charge states were unclear. This observation could not be explained, and was also seen with a variety of other protein and ligand combinations within this method.[203]

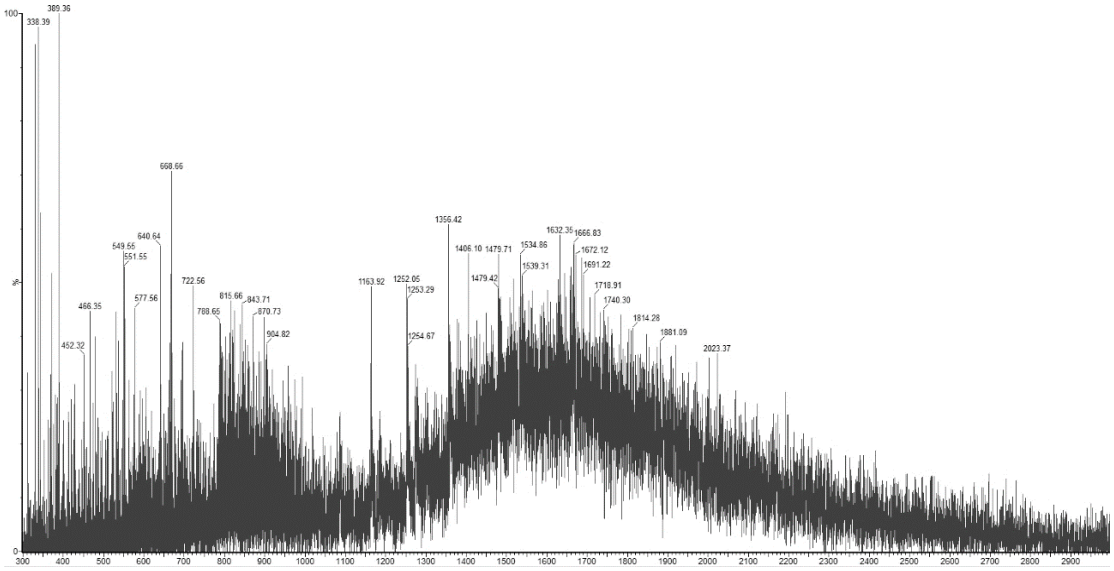


Figure 6.5: Mass spectrum of denatured OBP9 after being exposed to (4a*S*,7*S*,7a*R*)-nepetalactone **6** after undergoing the SPROX experiment. Data could not be reliably interpreted due to a high number of peaks and high signal to noise ratio.

The unusual observations could be due to several different factors. Firstly, the ligand may also be oxidised in the process. This could be avoided by using a more sulfur-directed or weaker oxidant than hydrogen peroxide, such as oxone or mCPBA. The messy spectra may also be due to interactions between the protein and ligand, and a negative control – a ligand which has been shown to not interact with the protein – could be introduced to determine whether this phenomenon occurs without binding. So far, no successful SPROX assay has been performed for an OBP[203], but further refinement of this method is required. The SPROX methodology could also be useful for studying other factors that affect a protein's stability, such as the mutation of specific residues, where no new species is introduced to the mixture.

6.2.2 Native ESI Mass Spectrometry

Following on from the unsuccessful SPROX mass spectrometry binding assay above, OBP9 was tested for binding with the sex pheromone components using native ESI-MS. Here, softer ionisation conditions are used, with no denaturant is introduced, in order to produce a mass spectrum of the native protein. The softer ionisation techniques result in the production of significantly less charge states. For OBP9, two charge states (+9 and +8) were observed (Figure 6.6).

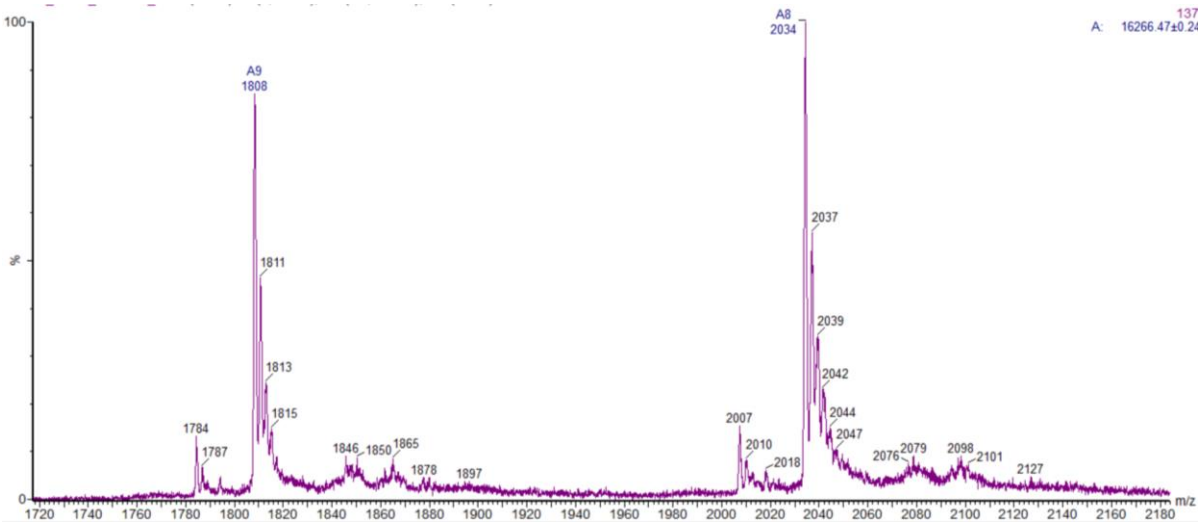


Figure 6.6: The native electrospray ionisation mass spectra (ESI-MS) of OBP9. Charge states +9 and +8 can be seen.

To observe binding of the ligand, native OBP9 was incubated with (4a*S*,7*S*,7a*R*)-nepetalactone **6** before injecting the complex into the mass spectrometer. This method has previously been utilised with OBPs from *Bombyx mori* to study their interaction with the sex pheromone bombykol, among other components.[122] If successful, charge states representative of the mass of the complex (mass of the protein plus the mass of the ligand) would be observed. The mass spectrum for OBP9 and (4a*S*,7*S*,7a*R*)-nepetalactone **6** appeared identical to the native protein without ligand (Figure 6.6) and no mass representative of a protein-ligand complex was observed.

Other experimental results reported in Chapters 4 and 5 demonstrated that OBP9 had little or no interaction with the sex pheromone components of *A. pisum*, including (4a*S*,7*S*,7a*R*)-nepetalactone **6**. This may explain why no protein-ligand complex was observed in this experiment. Due to time constraints, OBP6 was not screened using ESI-MS, and this experiment should be conducted in future.

6.2.3 Structural and Binding Studies with 2-Dimensional ^{1}H-^{15}N NMR

Due to the high expense of doubly labelled proteins, OBP6 was singly labelled with ^{15}N by expressing the protein in ^{15}N-labelled minimal media for preliminary experiments. A large-scale protein expression was performed, where the *E. coli* were grown to an OD_{600} of 0.5-0.6 in standard LB media, after which the cells were extracted and switched to ^{15}N-labelled minimal media before induction of protein expression.[199] This method ensured enough bacterial growth was obtained quickly, while still producing labelled proteins, and has been shown previously the give the best yields for isotopically labelling proteins produced from *E.*

coli.[199] Initially, to ensure the highest possible yield, experiments were performed with the His_6-tags intact. The proteins were purified *via* a nickel affinity column (as in 3.2.3) and fast-protein liquid-chromatography (FPLC, as in 3.2.5) to produce fractions of pure and highly concentrated ^{15}N-labelled OBP6-His_6 (Figure 6.7).

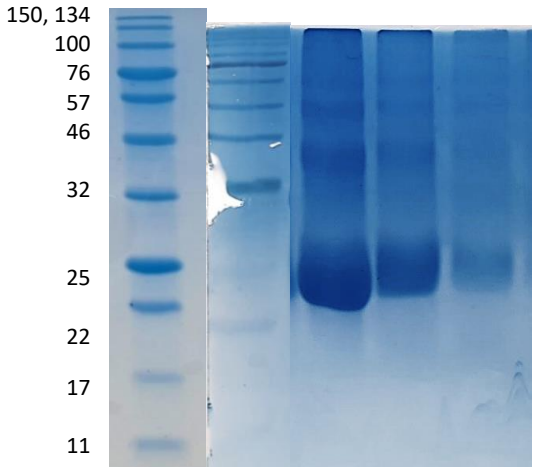


Figure 6.7: SDS-PAGE gel displaying ^{15}N labelled OBP6-His_6 fractions after one round of nickel affinity purification and one round of fast-protein liquid chromatography (FPLC). From L-R protein marker A, protein marker B, ^{15}N-OBP6-His_6 fractions 1-5. Molecular weight (kDa) of markers from the top is 150, 134, 100, 76, 57, 46, 32, 25, 22, 17, 11. A marker from another gel is included for clarity due to the poor photo quality of the SDS-PAGE of the labelled protein.

After successful purification, samples were transferred into 9:1 H_2O:D_2O to a final concentration of 8.5 mg ml^{-1} (0.34 mM). This sample was then used for initial HSQC NMR experiments which were performed on a 600 MHz Bruker NMR spectrometer (Figure 6.8).

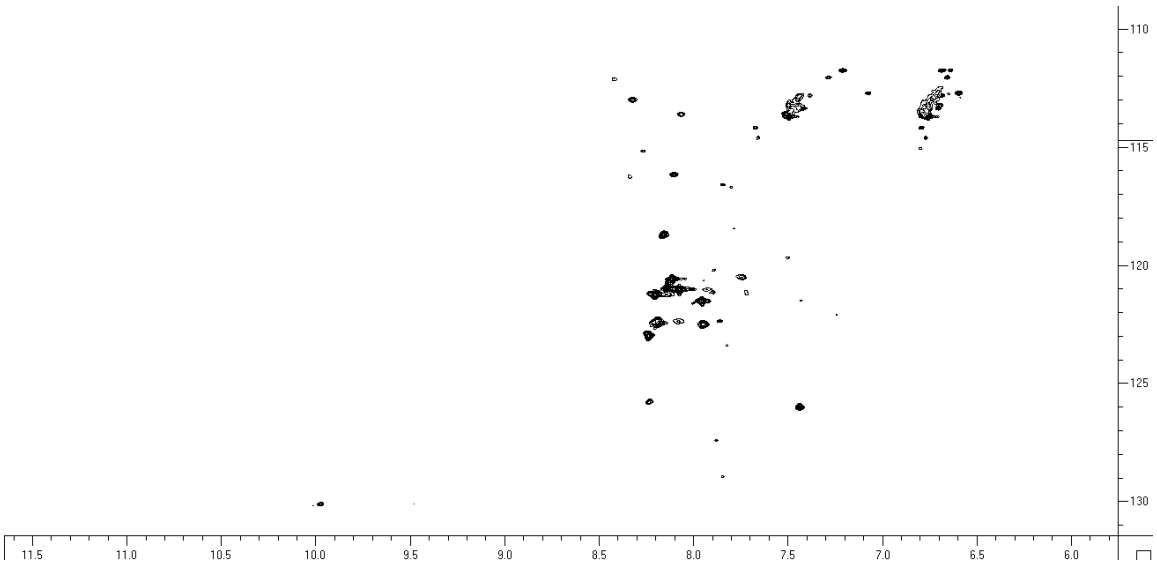


Figure 6.8: Initial HSQC NMR of ^{15}N labelled OBP6-His_6 (0.34 mM) in 9:1 H_2O:D_2O recorded on a 600 MHz Bruker NMR.

Initial HSQC-NMR analysis was not particularly clean and peak resolution was difficult to achieve. These issues may have arisen from the inclusion of a His_6 tag that protrudes from the globular structure of the overall protein (Figure 6.9). This causes the protein to tumble in the NMR experiment differently than how it would without the tag, with the main globular protein tumbling at a much slower rate than the His_6 tag tail. This His-tag and a few other residues are over-represented in the final spectra. This is due to the His-tag being able to relax and reorientate faster than other aspects of the proteins structure due to its small size protruding from the main structure (Figure 6.9). The fast relaxation results in a sharper peak and dominates signals. Furthermore, the presence of six histidines intensifies this effect, causing an overall dynamic range issue. This property can be useful in some NMR ligand-binding experiments and used to study changes in the 3-dimensional structure, however, in observing the standard HSQC NMR of a protein, it is a severe limitation, and the histidine residue peaks severely overshadow the rest of the protein. Additionally, the lack of dispersal of peaks is a good indicator that the protein is folded into α-helices.

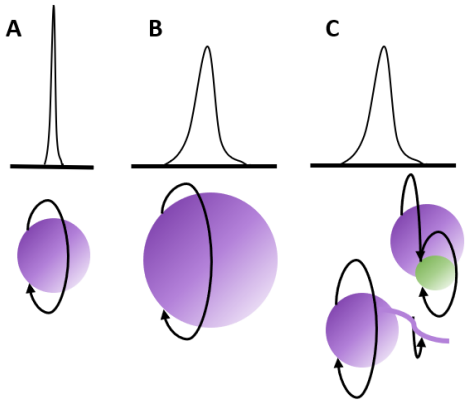


Figure 6.9: Differences in the peak resolution for NMR spectra of various different molecule due to the effects of tumbling. Small molecules (**A**) tumble faster, whereas larger molecules (**B**) and complexes or molecules with protruding tails (**C**) tumble slower, with different parts of the complex tumbling at different rates to others, resulting in broader peaks for some residues and sharper for others. Figure adapted from [204].

To overcome the issue with tumbling, OBP6-His_6 can be treated as a larger molecule and a higher magnetic field can be applied. In this case, OBP6 was tested in NMR experiments on a 900 MHz Varian spectrometer, using transverse relaxation optimised spectroscopy (TROSY) experiments with non-uniform sampling (NUS) (Figure 6.10). TROSY data collection uses HSQC NMRs with inbuilt properties to remove broad peaks, such as ^{1}H-^{15}N decoupling elements. NUS is an inbuilt method of only collecting a fraction of the NMR data and creating a prediction or projection of the remainder.

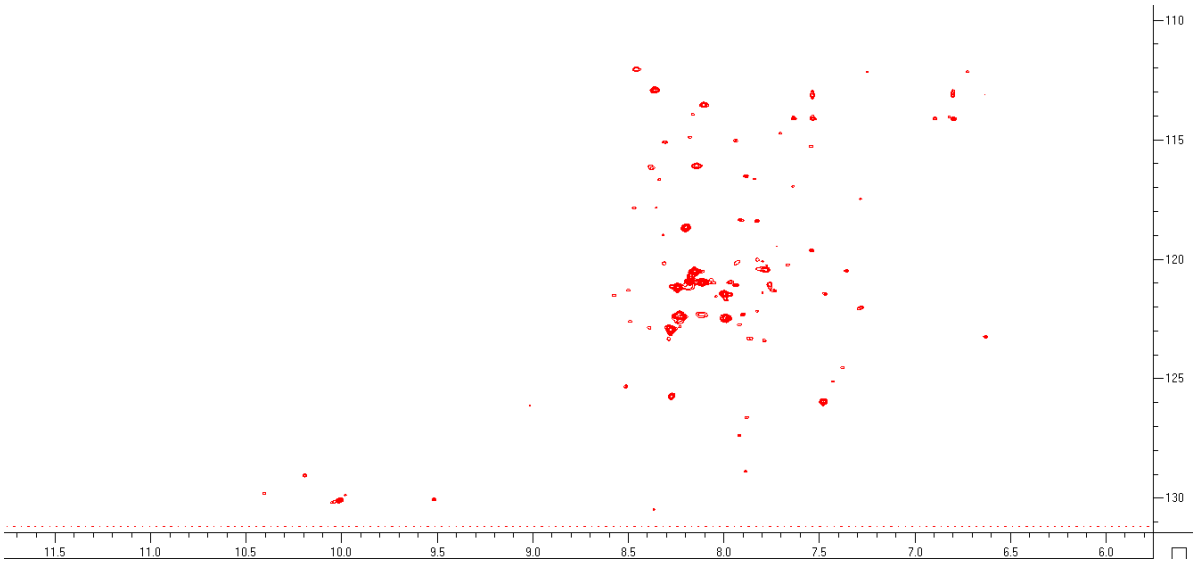
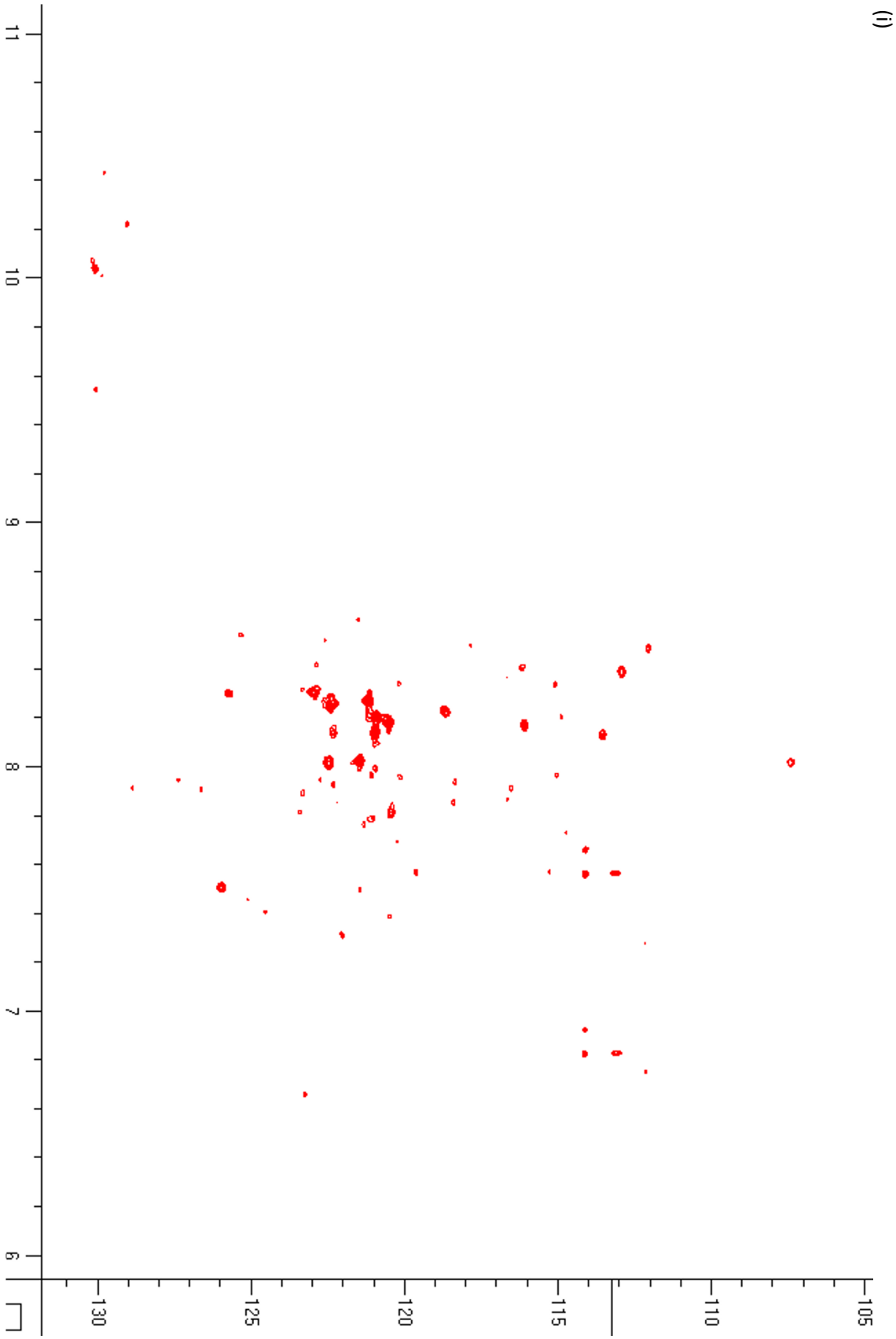
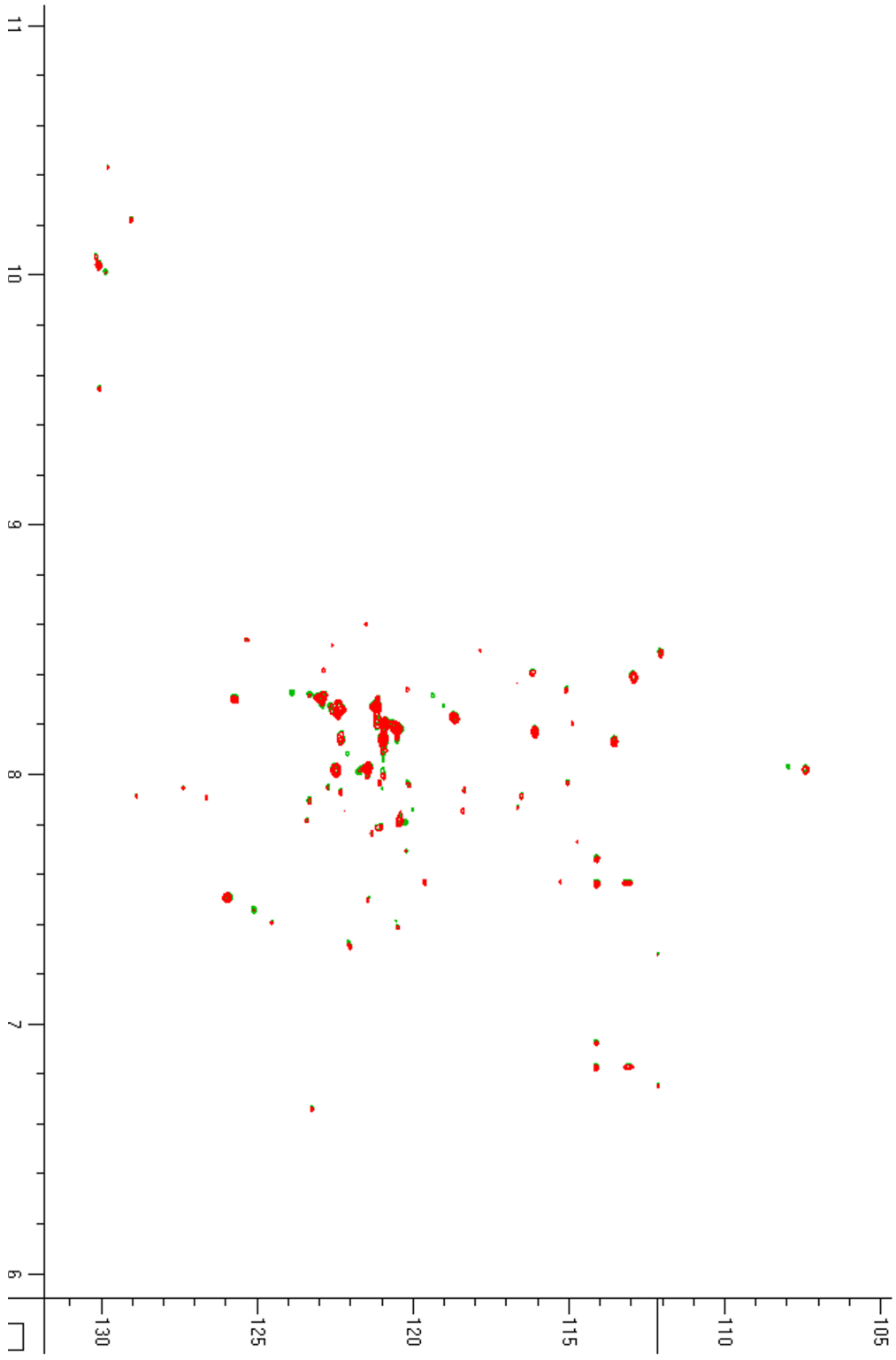


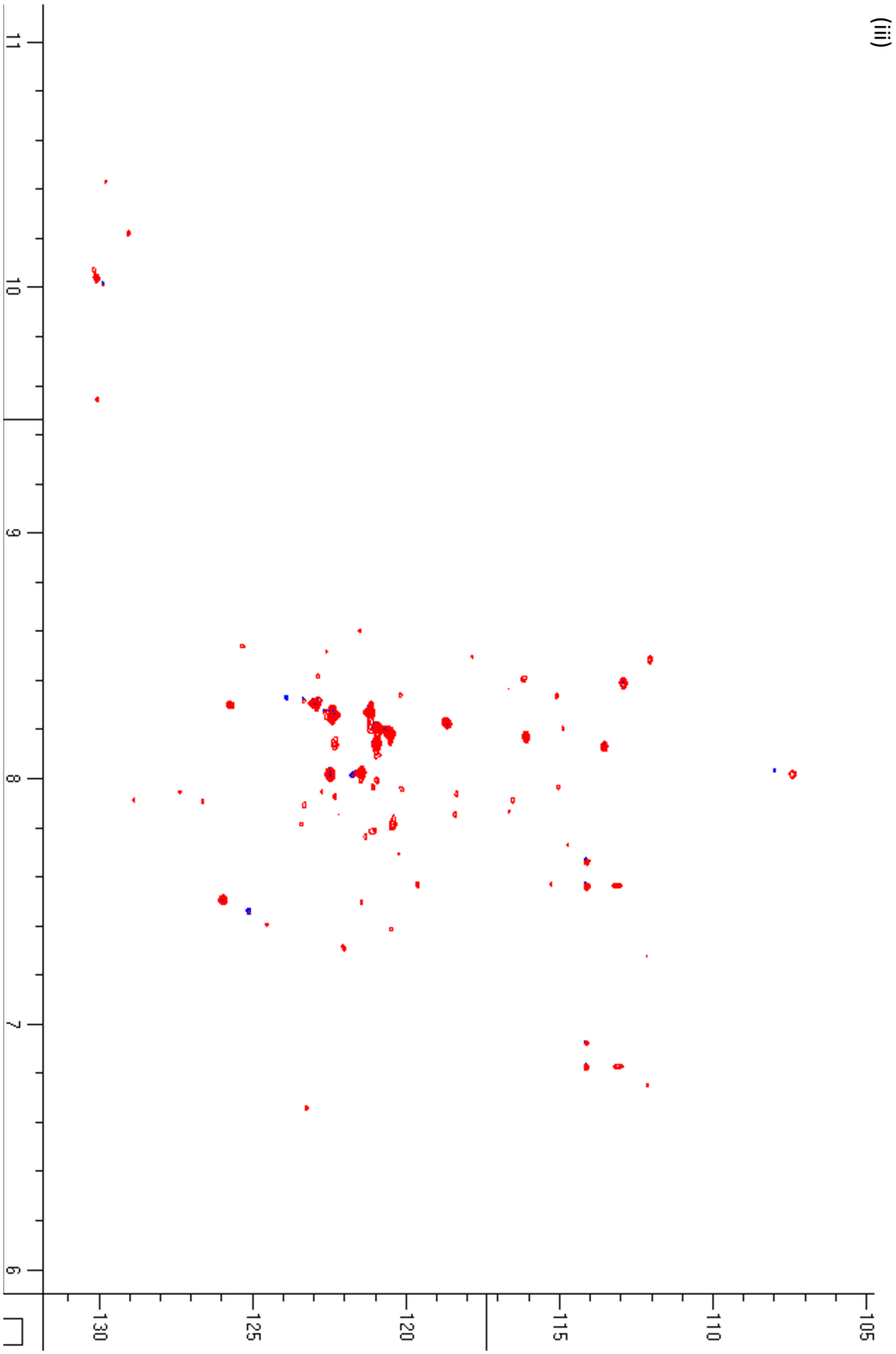
Figure 6.10: Initial transverse relaxation optimised spectroscopy (TROSY) experiments with non-uniform sampling (NUS) NMR of ^{15}N labelled OBP6-His$_6$ (0.34 mM) in 9:1 H_2O:D_2O recorded on a 900 MHz Varian NMR.

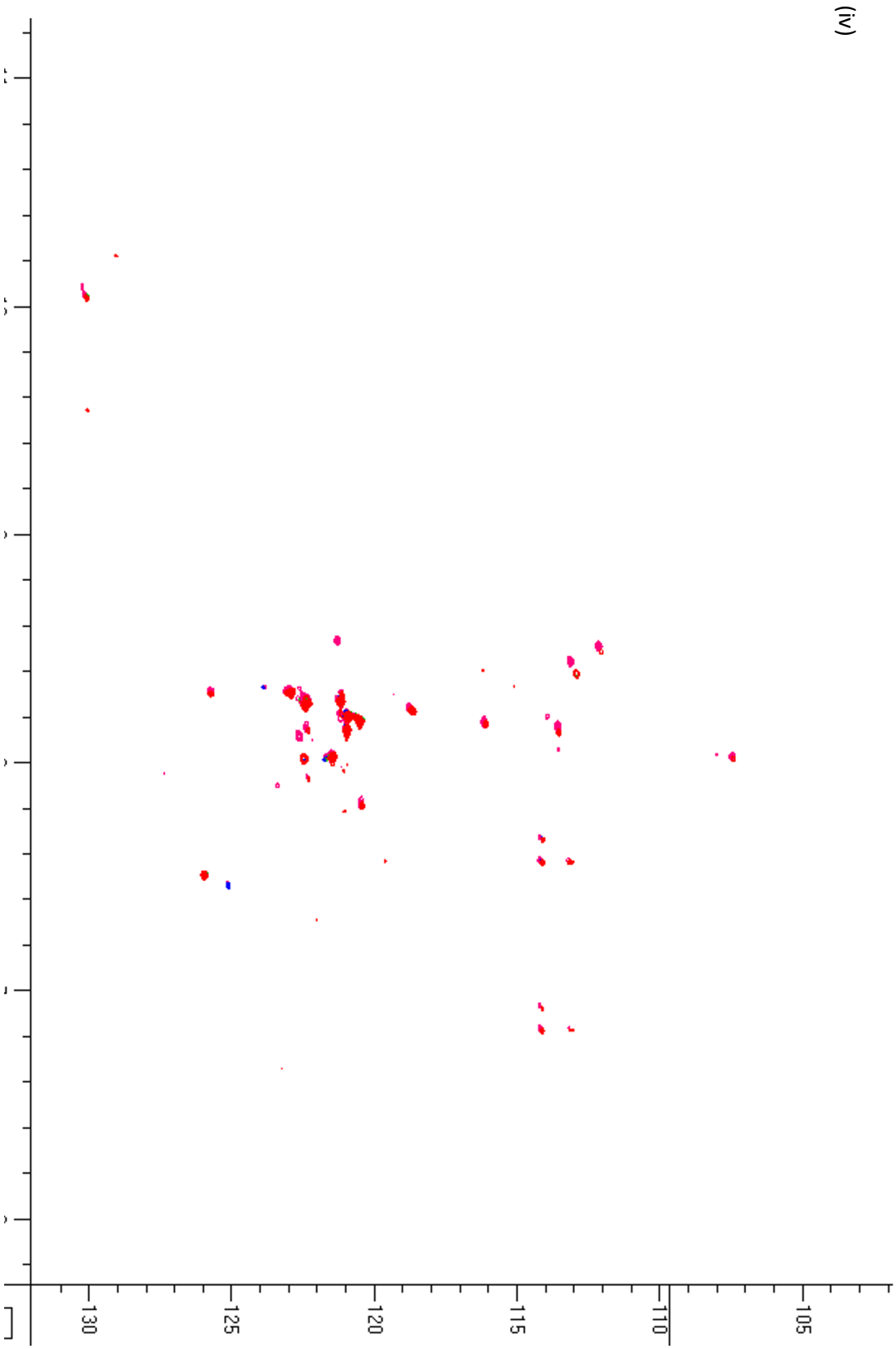
The TROSY experiment gave clearer NMR data with better visualisation of more peaks. Due to the single labelling nature of this experiment, it was not possible to fully assign the protein. However, changes due to the binding of a ligand could potentially be observed. To explore binding between OBP6 and the sex pheromone components, a TROSY NMR of OBP6-His$_6$ was recorded, and the ligand (4a*S*,7*S*,7a*R*)-nepetalactone **6** titrated in to record any observable changes (Figure 6.11).


(i)
11
10
9
8
7
6
130
125
120
115
110
105



(ii)


(iiii)
11
10
9
8
7
6
130
125
120
115
110
105


(iv)
130
125
120
115
110
105

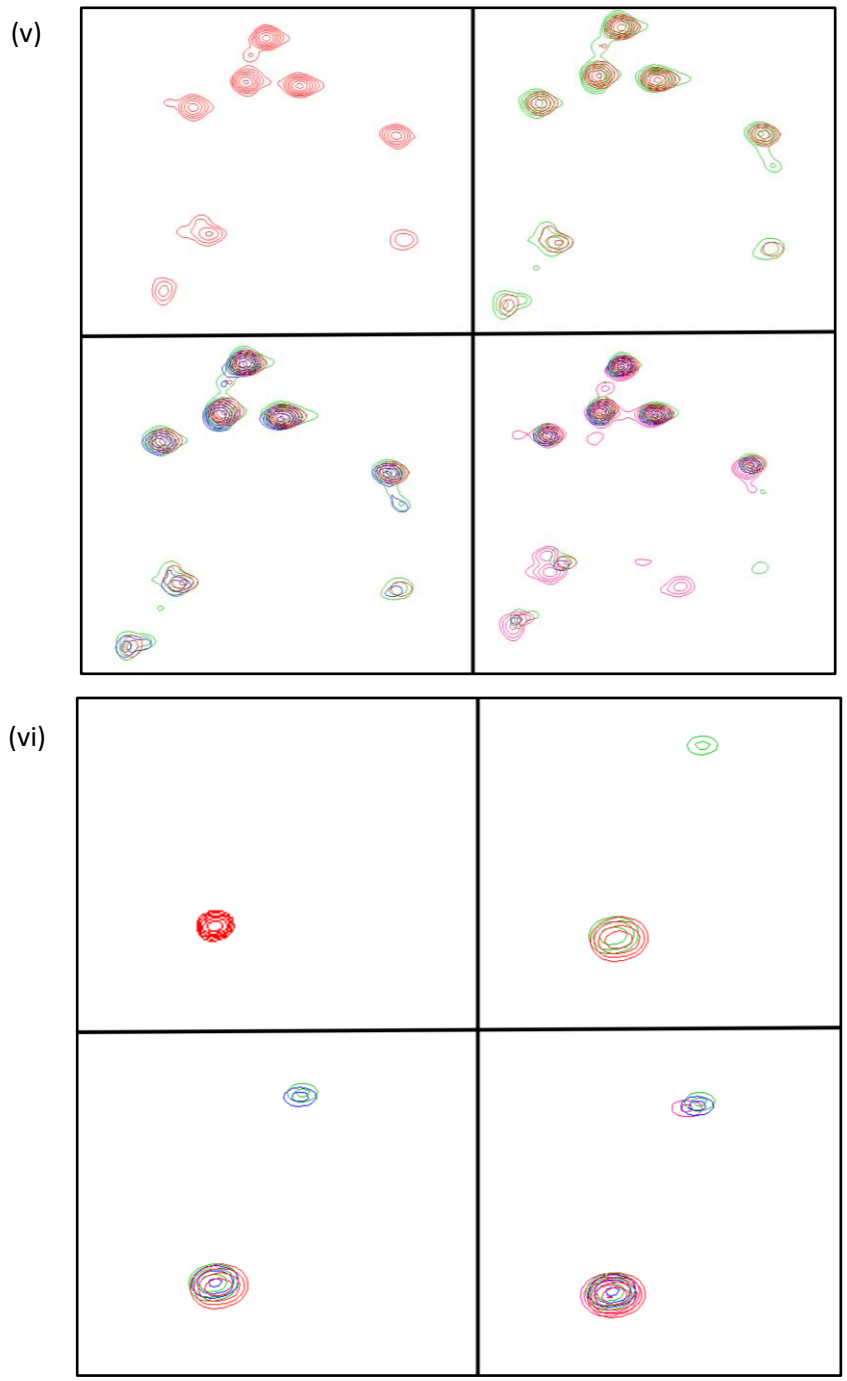


Figure 6.11: Initial transverse relaxation optimised spectroscopy (TROSY) experiments with non-uniform sampling (NUS) NMR of ^{15}N labelled OBP6-His$_6$ (0.34 mM) in 9:1 H_2O:D_2O recorded on a 900 MHz Varian NMR. (4a*S*,7*S*,7a*R*)-Nepetalactone **6** has been titrated in by adding to final concentrations (i) 0 mM (red); (ii) 0.15 mM (green); (iii) 0.30 mM (blue); (iv) 0.60 mM (pink). Zoom ins of particular segments can also be seen; (v) 8.0-8.4 ppm/120-124 ppm and (vi) 7.3-7.6 ppm/125-127 ppm

Spectra from the TROSY NMR experiment with OBP6-His$_6$ showed no significant changes when the ligand (4a*S*,7*S*,7a*R*)-nepetalactone **6** was titrated in. There are many reasons why

there may have been no observable change in these NMR experiments, one of which is that His_6 tag prevents or blocks the binding of the ligand, though this is unlikely, due to its position protruding out from the protein. The His_6-tag still causes issues with tumbling and the TROSY NMRs are still not as clear as is ideal for these experiments, with representing histidine residues still overshadowing the other protein peaks. In addition, OBP6 is not in an ideal buffer – the presence of many histidine residues may significantly affect the pH, and the pH may need to be altered to observe the protein and its binding activity.

Overall, these experiments were not successful in observing protein-ligand interactions, though there are many factors that can be altered in future attempts. TROSY NMR experiments with ligand titration should be repeated using OBP6 with the His_6 tag removed and (4a*S*,7*S*,7a*R*)-nepetalactone **6** in an appropriate buffer. Unfortunately, due to time constraints, this was not achieved during this thesis. In addition to optimisation of the singly labelled experiments, future work may involve assignment of the protein *via* double labelling and 3D NMR and validation of the data with other structural techniques.

6.2.4 Saturation Transfer Difference (STD) NMR spectroscopy

Saturation transfer difference (STD)-NMR spectroscopy is used to screen protein-ligand interactions. This method had not previously been used for protein-ligand binding studies within our group, therefore initial optimisation of the 500 MHz Bruker spectrometer for STD NMR experiments was required. A standard experiment to test STD-NMR capability is the study of the interaction between bovine serum albumin (BSA) and tryptophan **54**, with sucrose acting as a control non-binder. For the experiment, BSA was incubated with tryptophan **54** and a standard 1D proton (^{1}H) 'off-resonance' NMR recorded (Figure 6.10, red spectrum). The concentration of protein (30 μM) in comparison to ligand (3 mM) was relatively low (100-fold excess of ligand), and so the ligand peaks could be easily resolved in the NMR spectrum. A second spectrum was then recorded, with an on-resonance, carefully selecting a frequency that BSA resonates at, whilst not being close to the frequencies of the ligands of interest. By calculating the difference between these two spectra, an STD spectrum could be produced (Figure 6.12, blue spectrum).

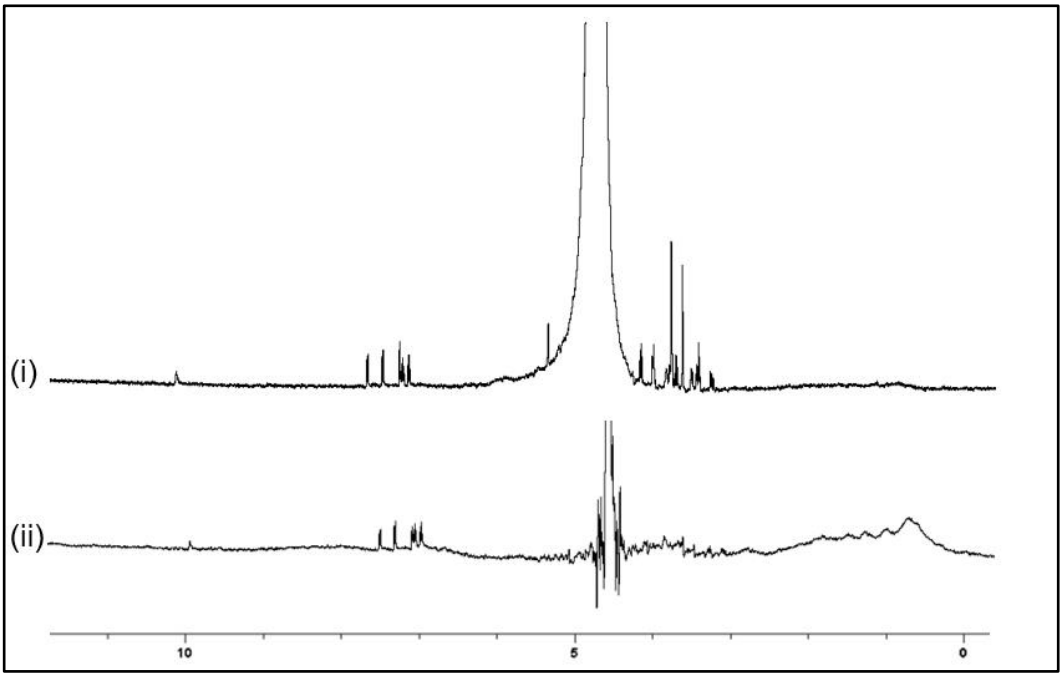


Figure 6.12: The saturation transfer difference (STD) nuclear magnetic resonance (NMR) spectra of bovine serum albumin (BSA) with tryptophan **54** (7.00-8.00 ppm) and sucrose (3.00-4.40 ppm). The initial ^{1}H can be seen (i), in addition to the STD NMR (ii), where only tryptophan peaks remain.

BSA is known to bind tryptophan **54** and have no interaction with sucrose; this is observed in the STD-NMR standard experiment (Figure 6.12). When an on-resonance of the protein is applied to the sample, resonance is transferred to the aromatic tryptophan protons, whereas the peaks from sucrose receive no resonance transfer (Figure 6.12). In the STD spectrum (Figure 6,12, blue spectra) only the tryptophan peaks remain. Furthermore, the relative intensities of the peaks for the tryptophan are subtly different – proteins interacting more closely with the protein will receive a higher resonance transfer than those further away, and therefore have a resulting higher intensity; this epitope mapping experiment allows the specific orientation of the ligand to be determined when bound to the protein.

After optimisation of the STD-NMR experimental parameters using BSA, a preliminary STD-NMR study with OBP6 and the natural enantiomer (4a*S*,7*S*,7a*R*)-nepetalactone **6** was undertaken, an interaction which was confirmed *via* other *in vitro* ligand-binding assays (Chapter 5). The nepetalactone, which was dissolved in *d*-dimethyl sulfoxide due to its insolubility in water, was incubated with OBP6 and a proton NMR spectrum with water suppression was reordered (Figure 6.13, red spectra). From this NMR analysis, a range of frequencies suitable for STD-NMR experiments, approximately 160-190 Hz, were identified. A control experiment using a frequency of -12000 Hz was also undertaken (Appendix Figure A.2).

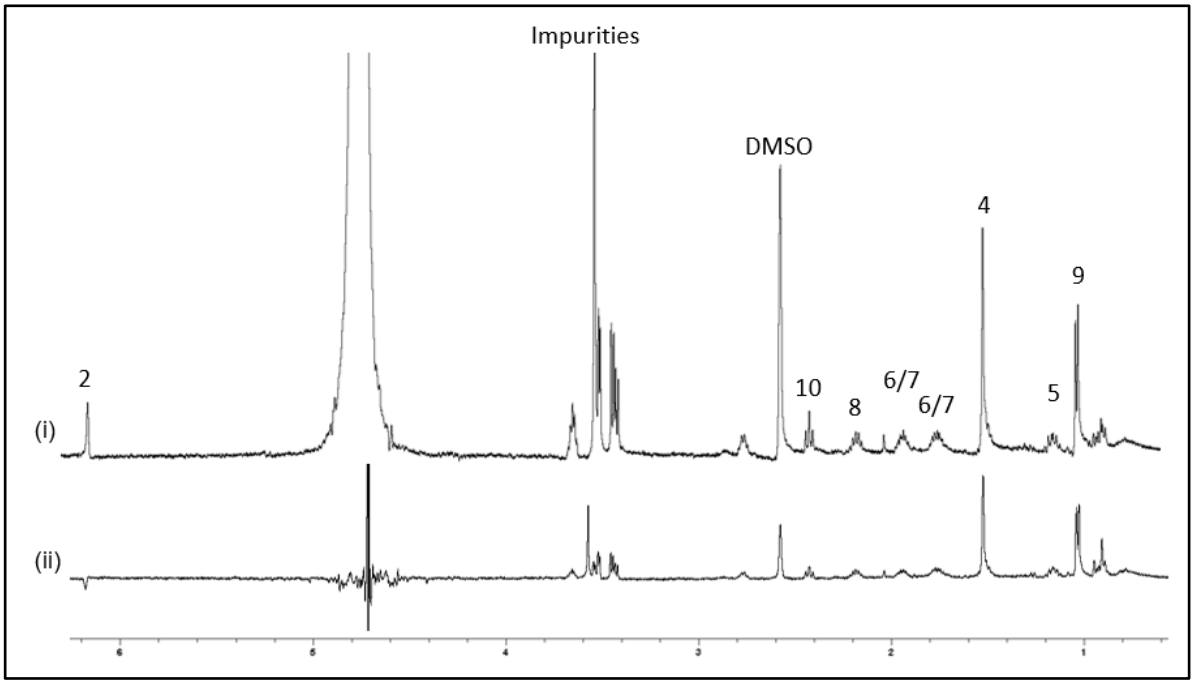


Figure 6.13: The STD NMR spectra of *A. pisum* OBP6 with (4a*S*,7*S*,7a*R*)-nepetalactone **6**. The initial ^{1}H spectra (i), in addition to the STD NMR (ii).

The STD-NMR experiment of OBP6 and (4a*S*,7*S*,7a*R*)-nepetalactone **6** yielded promising results (Table 6.1; Figure 6.14). To analyse the data, the relative intensity of each peak in the spectra was calculated and the difference observed. For the protons associated with (4a*S*,7*S*,7a*R*)-nepetalactone **6**, all peaks remained in the final STD-NMR spectra, indicating that they were in close association with the protein. The relative intensities of these proton peaks differed from the initial spectra; protons which are in the closest proximity to the protein will receive a higher amount of resonance transfer from the protein during the STD experiment.[201] This change can be used to accurately perform epitope mapping of individual protons, and therefore determine the specific interactions of the ligand with the protein.[205] Additionally, STD-NMR can be used to generate binding curves of the ligand by performing a titration.[205] To calculate epitope mapping within this protein-ligand interaction, intensities were normalised to the highest intensity peak and the change monitored (Table 6.1; Figure 6.14).

Table 6.1: Assignments and changes in relative intensity of different peaks in the STD NMR spectrum of OBP6 and (4a*S*,7*S*,7a*R*)-nepetalactone **6**.

* The peak for 2-H became negative after the STD experiment

δ / ppm (multiplicity)	Assignment	Epitope Mapping
1.21 (d)	9-H	110 %
1.50-1.59 (m)	5-H	75 %
1.64 (s)	4-H	100 %
1.89-1.98 (m)	6-H or 7-H	90 %
2.02-2.11 (m)	6-H or 7-H	60 %
2.31-2.39 (m)	8-H	70 %
2.05 (q)	10-H	50 %
2.71 (s)	*d*-DMSO	N/A
6.18-6.20 (m)	2-H	(-) 25 %

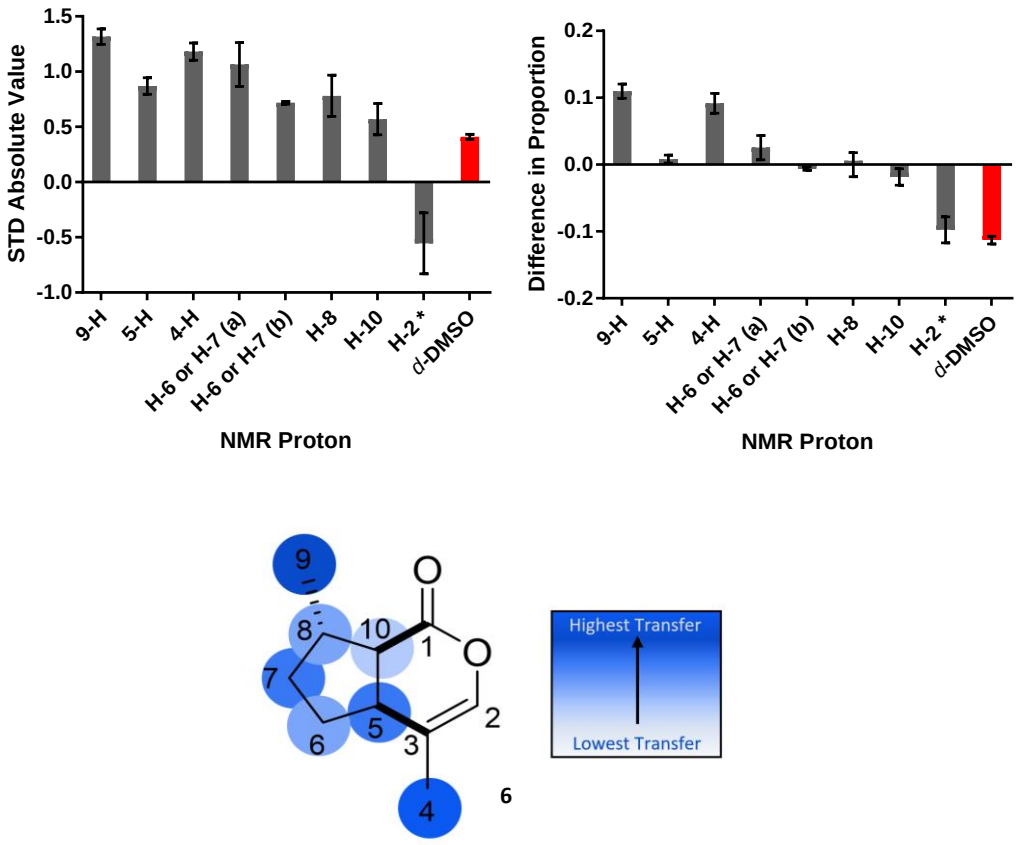


Figure 6.14: (i) STD absolute values[206] and changes in relative intensity of different protons in the STD NMR spectrum of OBP6 and (4a*S*,7*S*,7a*R*)-nepetalactone **6**; (ii) The structure of (4a*S*,7*S*,7a*R*)-nepetalactone **6** annotated with the epitope mapping results.

* H-6 and H-7 are indistinguishable by NMR

These results suggest that (4a*S*,7*S*,7a*R*)-nepetalactone **6** binds to OBP6, with an affinity high enough for saturation to be observed. The STD response varied across the molecule, with the most significant saturation occurring for the two methyl protons (9-H and 4-H; Figure 6.14). Modelling of the interaction between the sex pheromone components and OBP6 in Chapter 4 suggest the sex pheromone components fit into a small pocket, with these protons closest to the protein due to their peripheral position (Figure 6.15). There is little difference between the protons surrounding the ring, with 10-H giving the lowest level of interaction. In predicted models, (4a*S*,7*S*,7a*R*)-nepetalactone **6** fits in the binding pocket tightly (Figure 6.15). The proton closest to the pocket entrance, is 10-H, with the carbonyl group pointing out of the pocket (Figure 6.15).

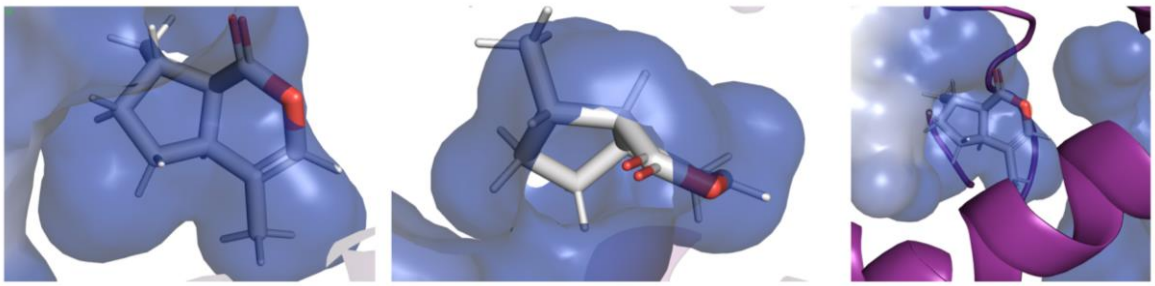


Figure 6.15: (4a*S*,7*S*,7a*R*)-nepetalactone **6** (white, with oxygens in red) in the predicted binding pocket of OBP6 (blue/purple).

Though the epitope mapping of (4aS,7S,7aR)-nepetalactone **6** bound to OBP6 is consistent with generated models, it was difficult to undertake quantitative analysis. For this experiment, only one saturation time was used in one experiment. As different protons have different relaxation times, the observed epitope mapping result could arise from saturation transfer within the nepetalactone protons.

An unusual result is observed for the alkene/lactone proton, 2-H. In the final STD spectra, 2-H (Figure 6.13; multiplet at 6.18-6.20 ppm) suggest a negative STD response. As the STD spectrum represents a difference between two NMRs, it is not impossible for a negative peak to appear, though it is unexpected; as the rest of the molecule appears to interact with OBP6 it is reasonable to expect that this proton would also demonstrate evidence of an interaction in the STD spectrum. This negative difference peak has been observed in other STD NMRs and was previously explained as due to a D_2O molecule interfering with the saturation of the ligand during spin and lock time.[206,207] The proton 2-H is sticking out of the predicted pocket in models (Figure 6.15). Despite being within the pocket, 2-H is also close to a second binding pocket within OBP6. This proton's position makes it solvent accessible and could result in a D_2O molecule associating or blocking the saturation transfer from the protein. In previous

literature, this effect has been observed with lactose ring structure, similar to the lactone structure seen here.[207] This effect may also be seen with nepetalactol and the proton adjacent to the hydroxyl group.

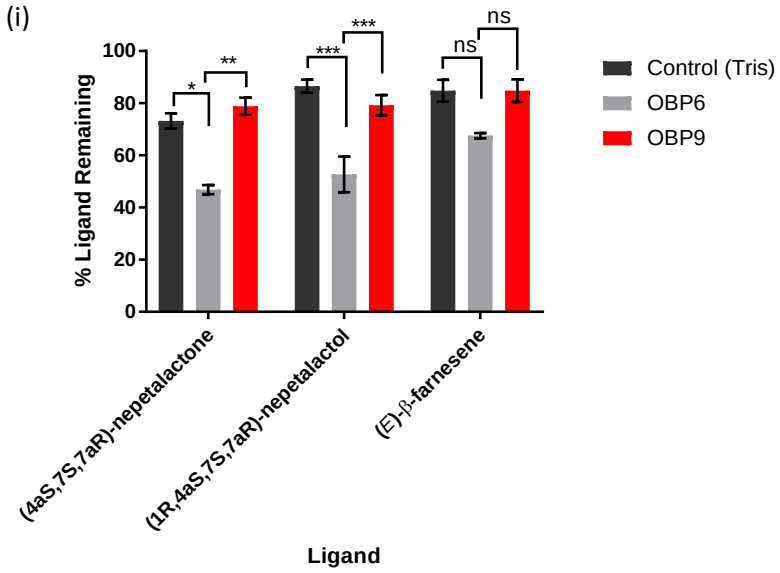
There are several advantages of the STD-NMR methodology. The method uses low concentrations of protein - the ligand to protein ratio need only be approximately 60:1 to 100:1. Furthermore, there are few restrictions to the proteins and ligands that can be used, except that they must be soluble in an NMR appropriate solvent. Protein-ligand interactions must be of a specific level of affinity for the resonance transfer to occur, and this may be the main restriction for observing the interactions of insect OBPs.[201] Ligands must bind with a high enough affinity for saturation transfer to occur, but an affinity of no higher than around 100 nM. This is as ligands must bind and unbind for the saturation transferred ligands to be observed in the NMR. However, a significant advantage of this method is that mixtures of NMR-distinguishable ligands can be tested simultaneously. This does reduce the possibility of testing two respective enantiomers simultaneously, due to their identical NMR spectra, however this is not possible with most ligand-binding assays. If the two enantiomers presented very clear differences in binding activity in individual assays, for example, one ligand that binds with a reasonable affinity and one that does not bind at all, a racemic mixture may give results that are some 'middle ground' between the two results. If the two enantiomers have similar binding activity, or bind in potentially different conformations, it may be possible to utilise a chiral shift reagent to achieve enantiomeric discrimination within the STD-NMR spectra. Whereas use of a covalently bound chiral shift reagent or chiral derivatizing agent such as Mosher's acid will change the binding activity of the ligand, a non-covalently bound reagent, such as $Eu(fod)_3$, could be used to distinguish between enantiomers in their unbound state. For the assay to give accurate binding activity results, the non-covalently bound chiral shift reagent would need to have a lower affinity for the ligand than the protein.

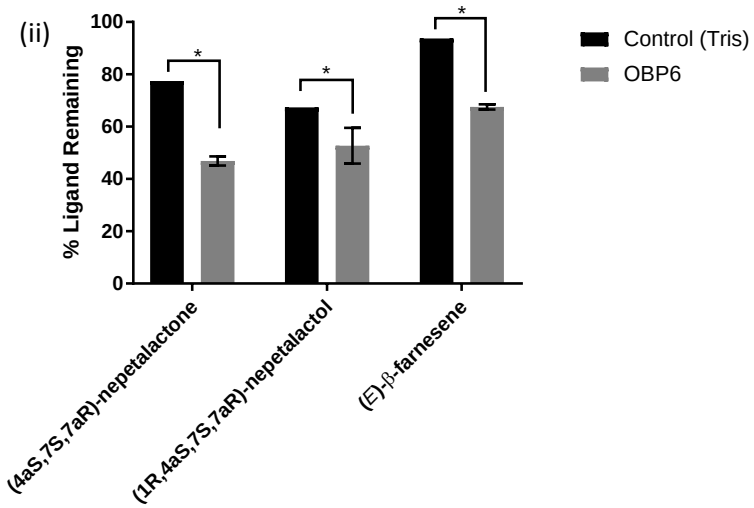
The results of this experiment provide further evidence of an interaction between OBP6 and naturally occurring (4a*S*,7*S*,7a*R*)-nepetalactone **6**, and supports the results of previous studies in Chapters 4 and 5). STD-NMR methodology has been the most successful of the alternative experiments used thus far, and although it has not been used previously to study insect OBP-ligand interactions, it may provide a viable and simple method for determining successful interactions in future.

6.2.5 Biphasic Gas Chromatography Assay

To investigate the interactions of OBPs and ligands across a biphasic system, more comparable to an *in vivo* representation, a biphasic assay was designed where ligand quantities would be monitored by gas-chromatography. Initially, OBP9 was investigated as in the mass spectrometry assays (6.2.1 and 6.2.2). The assay comprised of two immiscible phases – an aqueous phase containing protein, and a non-aqueous hexane phase containing the ligands of interest. Ligand concentrations were standardised, and a calibration curve generated for each ligand in order to accurately determine concentration and amount of ligand present using FID GC. The two layers were combined, gently mixed and allowed to settle before being left for two hours at room temperature. The amount of ligand present in the hexane layer was monitored in both a control and a sample containing OBP9. After initial assays, an optimised assay was developed to study the interaction of OBP6. In this assay, samples were centrifuged after mixing.[115] Three ligands were investigated, namely the naturally-occurring sex pheromone components (4a*S,*7*S,*7a*R*)-nepetalactone **6** and (1R,4a*S,*7*S,*7a*R*)-nepetalactol **5**, and the alarm pheromone (*E*)-β-farnesene **17**.

To compare the two assays, two values for each experiment were calculated – an overall percentage decrease in ligand amount in the hexane layer (Figure 6.16, (i) and (ii)) and a ratio of the concentration of protein present to the decrease in concentration of ligand present after 2 hours (Figure 6.16 (iii) and (iv)). For control samples, where no protein was present, the ratio of decrease could not be calculated. Statistical analysis was performed using an analysis of variance (ANOVA) in R 3.4.4.[208]


(i)
% Ligand Remaining
100
80
60
40
20
0
*
**
***
***
ns
ns
Control (Tris)
OBP6
OBP9
(4aS,7S,7aR)-nepetalactone
(1R,4aS,7S,7aR)-nepetalactol
(E)-β-farnesene
Ligand


(ii)
% Ligand Remaining
100
80
60
40
20
0
*
*
*
Control (Tris)
OBP6
(4aS,7S,7aR)-nepetalactone
(1R,4aS,7S,7aR)-nepetalactol
(E)-β-farnesene

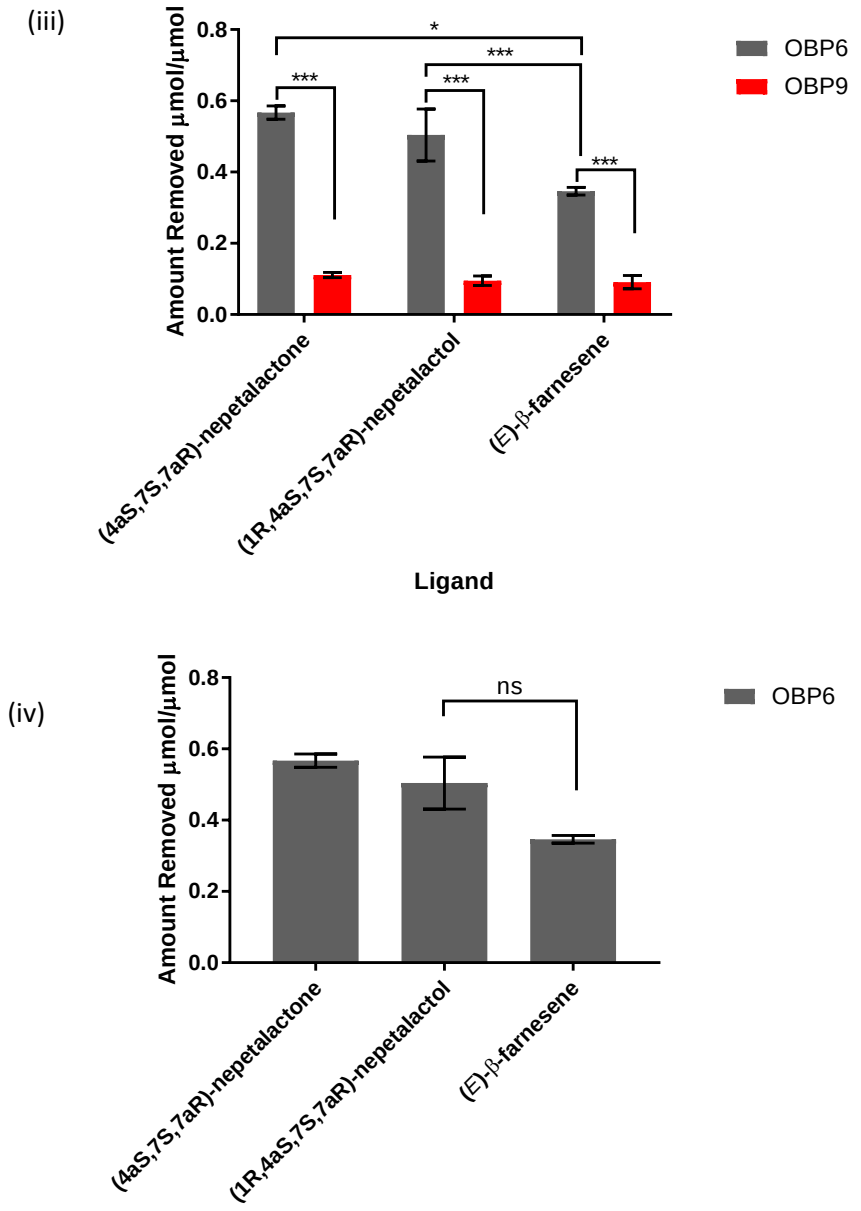


Figure 6.16: The percentage change in amount of ligand in the biphasic assay as monitored by gas-chromatography; (i) OBP6 and OBP9 compared to a control and (ii) OBP6 only, compared with a control. The ratio of µmol of protein to µmol of ligand removed from the layer, (iii) for OBP6 and OBP9 compared and (iv) for OBP6 alone. For statistical analysis, * = $p < 0.05$; ** = $p < 0.01$; *** = $p < 0.001$; ns = no significance

From these experiments, significant differences in both the percentage of ligand removed from the hexane layer, and the amount removed relative to the amount of protein present, can be clearly observed between OBP6 and OBP9, in addition to between OBP6 and the control. Biphasic systems with OBP6 present in the aqueous layer lost a significantly higher amount of ligand from the hexane layer than those with OBP9 or no protein present. Furthermore, the ratio of ligand (µmol per µmol protein) lost to the hexane layer was

significantly higher in samples with OBP6 than the control or those with OBP9. These results suggest that OBP6 is better able to solubilise or bind the investigated ligands from the hexane layer into the aqueous layer than OBP9, in addition to increasing the amount of ligand that is up taken into the aqueous layer than *via* natural diffusion.

There are some limitations to this methodology. Firstly, solvent may evaporate slightly over the course of the experiment, increasing the perceived concentration of ligands. This can be counteracted by careful sealing of the sample tubes, in addition to leaving samples in a location with a moderate and consistent temperature and low/no ventilation. Furthermore, this error would likely lead to a false negative result, where it appears less ligand is lost and will be consistent among all data sets. Data showing a significant reduction in the concentration of ligand in the hexane layer was still recorded regarding the activity of OBP6 in these experiments. In addition, the small change to the method between OBP9 and OBP6 may result in significant data for OBP6 and non-significant data for OBP9. However, isolation of the data for OBP6 (Figure 6.16 (iii) and (iv)) still demonstrates a significant difference from the control.

These experiments corroborate the data observed in the fluorescence binding experiments (Chapter 5) and the STD-NMR experiment (6.2.4). Overall, all ligands were transferred into the aqueous layer in high quantities with OBP6. Despite appearing that less of the alarm pheromone was transferred, there was no statistically significant difference between the ligands. Data from fluorescence assays suggest that OBP6 still binds (*E*)-β-farnesene, but with an approximately five times weaker affinity than the sex pheromone components.

6.2.6 Other Experiments

After successful binding interactions were observed between OBP6 and the sex pheromone components in various studies, the next step was to explore the structure of OBP6, ideally with a sex pheromone component in the binding pocket. The structure of many OBPs have been previously published, including *Nasonovia ribisnigri* and *Myzus persicae* OBP3 and OBP7 bound to the alarm pheromone, (*E*)-β-farnesene **17**.[104]

Screening trays were set up to attempt to crystallise *A. pisum* OBP6. After a few weeks, some precipitates i.e. potential crystal candidates started to appear. Unfortunately, experiments had to be halted due to the COVID-19 global pandemic and subsequent time constraints.

6.3 CONCLUSIONS

In this chapter, various methodologies were utilised to explore the interactions of *A. pisum* OBPs, specifically OBP9 and OBP6, with the natural enantiomers of the sex pheromone components. Mass spectrometry experiments were only performed using OBP9 and (4a*S,*7*S,*7a*R*)-nepetalactone **6**. The stability of proteins by rates of oxidation (SPROX) methodology was unsuccessful and requires further optimisation. Although ESI-MS analysis showed no binding between OBP9 and (4a*S,*7*S,*7a*R*)-nepetalactone **6**, this was consistent with the results of other experiments and clear spectra were obtained.

NMR correlation spectroscopy experiments with OBP6 were unsuccessful. This is mainly due to the presence of the His_6 tag, which significantly overshadowed other residues. Future experiments should be done with the tag removed, along with further optimisation and refinement of the method to determine the best buffer and conditions for this interaction.

STD-NMR spectroscopy gave promising results indicative of an interaction between OBP6 and (4a*S,*7*S,*7a*R*)-nepetalactone **6**. Estimated epitope mapping was performed, revealing results consistent with the *in silico* models and predicted interactions in Chapter 4. This method should be used to generate a binding curve by adding the ligand in increasing concentrations, and additionally, could be used to observe binding in the presence of multiple ligands. Experiments should be possible with individual enantiomers to observe enantiomeric differences but use of enantiomer mixtures would require the addition of a chiral shift reagent, due to the identical NMR spectra of enantiomers.

The most successful of the alternative experimental methods was the biphasic assay with gas chromatography. From this methodology, clear statistical differences between OBP6 and OBP9 could be observed. As with other methods, enantiomeric differences are impossible to observe via this methodology simultaneously due to their identical gas chromatography profiles (Kovats' Index/retention times) arising from their identical physical properties. One solution may be to use a chiral chromatography column in the GC. Gas chromatography relies on different physically properties, mainly polarity, of compounds to achieve separation. Chiral gas chromatography columns have an additional layer of separation by utilising a chiral stationary phase; the chiral stationary phase allows for the separation of enantiomers. The assay would need to be optimised, and chiral separation of the two enantiomers of nepetalactone and nepetalactol successfully completed.

It is clear from these multiple studies that OBP6 binds the sex pheromone components. As previously stated, future steps should involve the testing of this interaction *via in vivo*

behavioural and electrophysiological assays. These assays could also be repeated to include the enantiomers of the sex pheromone components, in addition to gathering a higher number of samples and hopefully mitigating some of the error.

Overall, exploration of alternative methods in this chapter has shown that there is some promise for their application in studying OBP-ligand interactions, with STD-NMR spectroscopy and biphasic gas chromatography being the most promising. In future, a wider range of ligands and proteins will be explored, and the techniques refined. The results of this chapter further confirm the interaction between OBP6 and the sex pheromone components, and this interaction should be explored in more detail through structural studies, such as NMR spectroscopy and X-ray crystallography.

Chapter 7: General Discussion

7.1 DISCUSSION

An increased understanding of insect olfaction, particularly in economically important pest species, can lead to the development of novel pest management tools. The pea aphid, *Acyrthosiphon pisum,* is a prevalent agricultural pest species and can be used as a model organism for all aphid species and insect olfaction in general. The aim of this thesis was to investigate the molecular mechanism of insect olfaction by evaluating the structure and function of odorant-binding proteins (OBPs), in addition to other olfactory proteins, using *A. pisum* as a model organism. OBPs are a highly conserved family of proteins that exist in high concentrations in the antennal lymph of insects.[101,112] Their role in insect olfaction is not fully understood.[113] The pea aphid sex pheromone comprise two components, (4a*S*,7*S*,7a*R*)-nepetalactone **6** and (1*R*,4a*S*,7*S*,7a*R*)-nepetalactol **5**. Aphids can discriminate between these and their respective, non-naturally occurring enantiomers; this discrimination ability has been demonstrated by observing differing electrophysiological responses to the different stereoisomers.[145] In this work, specific focus was put on the potential olfactory discrimination between sex pheromone components and their enantiomers, whilst trying to unravel the role of OBPs in insect olfaction.

Overall, the work of this thesis successfully described the synthesis and purification of both the *A. pisum* sex pheromone components and OBPs, *in silico* predictions of the structures and interactions between olfactory proteins and respective ligands, and *in vitro* investigations through a variety of methodologies including fluorescence, mass spectrometry, nuclear magnetic resonance spectroscopy and a gas chromatography biphasic assay.

7.1.1 Production of Aphid Sex Pheromone Components and Odorant-Binding Proteins

One of the sex pheromone components for *A. pisum*, (4a*S*,7*S*,7a*R*)-nepetalactone **6**, was obtained *via* purification of the essential oil isolated from the catmint plant *Nepeta cataria*. This was followed by sodium borohydride reduction to give the other sex pheromone component, (1*R*,4a*S*,7*S*,7a*R*)-nepetalactol **5**.[6] The non-natural enantiomers, (4a*R*,7*R*,7a*S*)-nepetalactone **33** and (1*S*,4a*R*,7*R*,7a*S*)-nepetalactol **32**, were successfully synthesised *via* a previously described synthetic route.[51,52] The synthesis starts with (*S*)-citronellol which was

oxidised *via* a selenium dioxide allylic oxidation. The subsequent diol is further oxidised to a dialdehyde, after which an enamine-mediated [4+2] cycloaddition was performed to give the bicyclic product. Finally, (1*S*,4a*R*,7*R*,7a*S*)-nepetalactol **32** was produced by hydrolysis of the cyclised product with toluensulfonic acid, and (4a*R*,7*R*,7a*S*)-nepetalactone **33** by a final oxidation. An improvement in the overall yield was achieved, and investigations into the mechanism of the enamine-mediated [4+2] cycloaddition performed.

Optimisation of the synthesis focused on two main steps. Firstly, the oxidation of (*R*)-8-hydroxycitronellol to (*R*)-8-oxocitronellal, for which four different oxidation methods were tested – TPAP/NMO, IBX, Swern and Dess-Martin. Both Swern and Dess-Martin oxidations were the most successful, providing the easiest methodology and highest yields. The second optimised step was the enamine mediated [4+2] cycloaddition of the dialdehyde. This step, which involved a reaction between an aniline and the dialdehyde to produce the bicyclic product, was investigated by changing substituents on the aniline. By altering the substituents, preliminary effects on the yield and diastereomeric excess were observed. Traditionally, this step is referred to as a Diels-Alder like mechanism, however, the altered substituent effects were more consistent with an inverse electron demand Diels-Alder (IEDDA) mechanism. Further optimisation and investigation of this step is required, however, in future this synthesis should be performed with a chloro-substituted aniline, as this increased both yield and diastereomeric excess – an overall improvement on the initial synthesis. In addition to the sex pheromone component synthesis, a racemic citronellol resolution methodology was unsuccessfully tested in an attempt to reduce overall cost of synthesis.

The synthesis and purification of *A. pisum* OBPs for further *in vitro* testing was generally successful. Cell lines (*E. coli* BL21(DE3)) expressing OBPs 1-10, except for OBP4, were transformed and small-scale (10 mL) expression tests performed. OBP6, 7 and 9 were expressed and purified at large scales (1-5 L). There were complications in expressing OBP6 – the cleavage enzyme to remove the His_6 tag over-cleaved the protein. This was rectified by introduction of an alternative cleavage site for an alternative enzyme *via* site-directed mutagenesis.

In summary, both *A. pisum* sex pheromone components and odorant-binding proteins were produced for future binding studies successfully, and the synthesis was optimised to provide greater yields and diastereomeric excess.

7.1.2 Predicted Interactions of Aphid Olfactory Proteins

Before *in vitro* studies were performed, *in silico* predictions were made to explore the structures and theoretical binding activities of the aphid olfactory proteins and respective ligands, with focus on both OBPs and odorant receptors (ORs).

Each of the OBPs were modelled using homology modelling techniques. The structures of both OBP3 and OBP7 have already been published for a closely related aphid species[104], but the other OBPs were modelled based on comparative modelling using a range of OBP structures from other insect species. Most of the models had an obvious binding pocket. Each model was then screened against a range of ligands using docking methods, and the specific interactions further investigated. Predicted interactions between ligands and OBP3/OBP7 could be seen, as expected from previous literature[102,104], however, the most interesting result was the predicted interactions of OBP6 from *A. pisum* with the sex pheromone components, (4a*S*,7*S*,7a*R*)-nepetalactone **6** and (1*R*,4a*S*,7*S*,7a*R*)-nepetalactol **5**, in addition to their enantiomers. These interactions had low predicted interaction energies and K_is (around 2 μM). As OBP6 is highly expressed in the aphid antenna, particularly in winged (sexual) aphids, this was a promising result and further information about the interaction was elucidated. Key residues were identified in these interactions and highlighted key differences between the sex component stereoisomers – Phe208 was identified as the critical residue for binding of the natural enantiomer of nepetalactol, (1*R*,4a*S*,7*S*,7a*R*)-nepetalactol **5**. Mutagenesis of Phe208 debilitated the binding activity of OBP6 in predicted interactions. This specific residue interaction could be tested by site-directed mutagenesis experiments in future, whereby Phe208 is mutated as in the *in silico* studies and binding tested *in vitro*.

The predicted sex pheromone binding activity from *in silico* screens aligns with expression analysis. OBP6 is a Plus-C OBP that is the second most abundant OBP mRNA found in the antennae, after alarm pheromone-binding OBP7.[75] Additionally, OBP6 is particularly highly expressed in the type II trichoid sensilla, where sex pheromone perception is thought to occur, and in winged (sexual) morphs.[75]

After identification of OBP6 as a candidate for sex pheromone binding, a range of sex pheromone component analogues were tested for binding activity. One analogue in particular (Analogue **42**) showed a significant increase in binding activity with OBP6 – a decrease of the K_i from approximately 2-3 μM to 380 nM was predicted. This analogue may be a good candidate for olfactory inhibition in future and work should be undertaken to

synthesis this analogue and test its behavioural effects on sexual aphids, from both the *A. pisum* species and other species that utilise the same sex pheromone components.

In addition to *in silico* studies with OBPs, ORs were also modelled and tested. OR models were built using comparative modelling techniques and the singly published structure of an OR, the cryo-EM structure of ORCO from *Apocrypta bakeri* (PDD 6C70).[88] The complete models were compiled into tetramers and were embedded in phospholipid bilayers for realistic ligand screening. Two ORs from *A. pisum* had been previously characterised *in vivo* – OR5 with the alarm pheromone (*E*)-β-farnesene, and OR4 with a variety of plant volatiles.[99,100] These ligand interactions were structurally characterised and key binding sites identified. Furthermore, the interactions of olfactory disrupter molecule VUAA1 **50** with ORs, including ORCO, were predicted, revealing the molecule fits perfectly within the receptor's structure, likely blocking the activity of any other ligand or the required movement of the receptors for olfactory function. Finally, deorphanized receptors of a high conservation were screened against the sex pheromone components to look for potential sex pheromone binding OR candidates. OR2 had potential for sex pheromone binding and enantiomeric discrimination ability. Though many interactions between receptor proteins and ligands were observed, the cross examination of previously published electrophysiological data with predicted interactions demonstrated the clear importance of binding site.[100,209] The critical binding site, and subsequent mode of action, for ORs/ORCO is yet to be fully understood, though evidence from previous studies and the work in this thesis predicted extracellular loop 2 (EL2) included critical residues for interactions.[175,210,211] Understanding the binding site in ORs could be important for understanding enantiomeric discrimination by insect olfactory systems.

Overall, *in silico* studies were successful, providing groundwork for future *in vitro* experiments in the thesis. Many potential interactions were identified, and these should be tested in future experiments. Furthermore, *in vivo* studies to determine the sex pheromone receptor in the pea aphid, directed by these predictions, should be undertaken, and development of CRISPR/Cas9 molecular techniques in this aphid species may provide a reliable method by which to do this.

7.1.3 The Role of OBP6 in Sex Pheromone Perception

As previously described, *in silico* studies indicated OBP6 as a potential candidate for sex pheromone binding. OBP6 has been a suggested as a potential candidate for sex pheromone perception previously[75], and the initial results were promising. Based on these results, *in vitro*

binding studies using the synthesised OBPs and sex pheromone components were performed.

In the field of insect olfaction, binding studies between OBPs and sex pheromone components are typically conducted *via* fluorescence-based competition assays with a fluorescent probe.[188] There are many limitations to these studies, and an alternative method using the intrinsic fluorescence of an aromatic residue, tryptophan **54**, was used. Both OBP6 and OBP9, used as a predicted non-interacting control protein that also contained tryptophan **54**, were tested with the aphid sex pheromone components and the alarm pheromone, (*E*)-β-farnesene **17**. A method whereby a fluorescent probe, *N*-phenylnapthalen-1-amine **51** (1-NPN), was introduced and the intrinsic fluorescence monitored, was used. 1-NPN **51** was seen to bind to OBP9 but not OBP6, however, its presence in OBP6 binding assays made no difference to the overall data. To keep the assay consistent between the two proteins, 1-NPN **51** was kept in the final assay.

Fluorescence binding studies between OBP6 and the sex pheromone components yielded promising data. The results of the experiment gave kinetic data consistent with the predicted values from *in silico* modelling, indicating that the models could be accurate. Furthermore, a significant difference in binding was observed between the sex pheromone components and the other ligands tested (alarm pheromone and the generic plant volatile, linalool). In addition to this general difference, subtle differences were observed between the individual sex pheromone components with the interaction between OBP6 and (4a*S*,7*S,*7a*R*)-nepetalactone **6** providing the lowest K_D value with 1.30 ± 0.60 µM. The fluorescence experiments with OBP9 were unsuccessful – this was likely due to its high affinity for the competitive fluorescent probe, 1-NPN **51**, used in this study, and the ligands' inability to displace the probe.

A second fluorescent probe, 8-anilino-1-napthalenesulfonic acid **52** (1,8-ANS), was used in binding experiments with OBP6 and OBP9. From initial experiments, it appeared OBP6 bound with 1,8-ANS **52**, however (4a*S*,7*S,*7a*R*)-nepetalactone **6**, a ligand shown to bind strongly with OBP6 in other fluorescence experiments, did not displace the probe. Modelling studies showed 1,8-ANS **52** binds in a completely different site to (4a*S*,7*S,*7a*R*)-nepetalactone **6**. The nepetalactone is subsequently unable to displace 1,8-ANS **52**, and instead both ligands are predicted to bind simultaneously. Unfortunately, no binding interaction between OBP9 and 1,8-ANS **52** was observed.

In addition to experiments involving a fluorescent probe, a probe-free study was performed. The experiment focused on monitoring changes in tryptophan **54** fluorescence on the addition of a ligand, in the absence of any other fluorescent species (aside from other fluorescent residues within the protein). OBP9 only possesses one tryptophan **54**, in contrast to OBP6's three. Despite tryptophan fluorescence being easily observed for OBP9, little difference could be seen when a ligand was introduced. This could be due to low or no interaction between the two species, or the positioning of tryptophan **54** within the protein. Successful studies were performed with OBP6. The results of the probe-free intrinsic studies reflected those of the 1-NPN assay. (4a*S*,7*S,*7a*R*)-nepetalactone **6** was shown to bind with the lowest K_D (1.90 ± 0.35 µM), with the other ligands binding with both K_Ds and error higher than the assay with 1-NPN **51**.

After the success of the fluorescence binding studies between OBP6 and the sex pheromone components, alternative studies were explored to delve deeper into the specifics of the interaction. First, trials of different mass spectrometry-based assays were attempted using OBP9. Both methods were partially successful, but no binding of a ligand could be observed. Firstly, for electrospray ionisation mass spectrometry (ESI-MS), OBP9 could be observed in two distinct low charge states, but no charge states representative of a complex between the protein and the ligand was observed when the two were incubated together and analysed through the mass-spectrometer as a native complex under soft ionisation techniques. This may be due to the fragile nature of the complex, or the lack of binding activity from OBP9. Secondly, the stability of proteins by rates of oxidation (SPROX) experiment proved successful in the generation of an unfolding curve for OBP9 alone across a range of denaturant concentrations, but no clear data could be obtained when a ligand was introduced. Both methodologies could be trialled with OBP6 in the future, however, ESI-MS is more promising than the SPROX experiment; the SPROX method was trialled with multiple protein-ligand combinations and was similarly unsuccessful in every case.[203]

Nuclear magnetic resonance (NMR) spectroscopy methods were more successful than mass spectrometry. Initially, 2-dimensional heteronuclear protein-observed NMRs of OBP6-His_6 were taken after single isotopic labelling with ^{15}N. Initial HSQC experiments were inconclusive – the His_6-tag caused tumbling issues resulting in the peaks representative of the histidine residues overshadowing the required residue peaks. Additionally, double labelling of the protein with ^{13}C would be required to determine structural data from 3-dimensional studies. When a ligand, (4a*S,*7S,7a*R*)-nepetalactone **6**, was titrated in to the OBP6-His_6 NMR sample, no visible change in the spectrum was observed. These specific studies could be repeated in

future after the removal of the His_6-tag to elucidate further information about the structure and binding of OBP6, and potentially to observe conformational changes when binding occurs. These conformational changes, usually occurring at the C-terminus, have been previously described in other insect OBPs.[108,117,212]

The successful ligand-observed NMR assay was the saturation transfer difference (STD) NMR. In this assay, ligand binding between OBP6 and (4a*S,*7S,7a*R*)-nepetalactone **6** was tested using a resonance transfer from the protein to any respective bound ligand. This method was initially optimised for use on the Rothamsted Bruker 500 MHz NMR using a standard BSA and tryptophan assay, followed by an analysis of OBP6 with the sex pheromone component. The final STD spectrum demonstrated a clear interaction between (4a*S,*7S,7a*R*)-nepetalactone **6** and OBP6, suggesting OBP6 binds the sex pheromone component. This result was consistent with the results of the fluorescence assay and the predicted results from the *in silico* experiments. Epitope mapping of the protein-ligand interaction revealed a tight interaction, with the most protruding methyl groups showing this highest level of resonance transfer with the protein. An unusual result was observed with the proton attached to the alkene/lactone side of the ring, in which it became negative in the STD-NMR spectrum. This has previously been remarked on with other protein-ligand interactions and is hypothesised to be due to a D_2O molecule blocking the interaction with the protein. This is consistent with the models and predicted positioning of OBP6, in which the side containing this proton is the most solvent accessible.

The final alternative assay used in this thesis was a biphasic gas-chromatography based assay. The biphasic assay was uniquely designed to provide a more realistic method of observing an OBP's binding activities, specifically determining whether it can abstract and bind ligands from a different phase. Solubilising ligands, typically hydrophobic in nature, from the air into an aqueous solution (the sensillum lymph) within the antennae is one of the main hypothesises for the role of OBPs.[112] Ligands were provided in a volatile non-aqueous solvent (hexane) and the amount of ligands that remained in the layer monitored *via* gas-chromatography, a particularly sensitive analytical technique. The assay was initially undertaken using OBP9, after which it was optimised, and the experiment repeated with OBP6. Three ligands were tested consistently – (4a*S,*7S,7a*R*)-nepetalactone **6**, (1*R*,4a*S,*7S,7a*R*)-nepetalactol **5** and (*E*)-β-farnesene. In the OBP9 experiment, no clear differences between the control samples and the sample containing OBP9 were observed. However, in the OBP6 assay, the amount of ligand in the hexane layer reduced to a significantly lower level than in the control or OBP9. Furthermore, the ratio of moles of ligand

up taken per mole of OBP was significantly higher in OBP6 than with OBP9. Overall, it appears OBP6 increases the amount of ligand than can be solubilised into the aqueous layer than in a control or OBP9. Unfortunately, no significant differences were observed between different ligands, but appeared that the effect for (*E*)-β-farnesene was lower than for the sex pheromone components. This result is consistent with the other ligand binding assays with OBP6, and further supports the role of OBP6 in sex pheromone perception.

Overall, the results of this thesis suggest that OBP6 binds the sex pheromone components, (4a*S,*7S,7a*R*)-nepetalactone **6** and (1*R*,4a*S,*7S,7a*R*)-nepetalactol **5**, and may be involved in the perception of the sex pheromone. For the aphid alarm pheromone, perception occurs using three proteins, a receptor OR5, and two OBPs, OBP3 and OBP7.[99] Though OBP6 binds the sex pheromone components, it is likely that is it is not the only protein involved in the process of identifying the sex pheromone components. The role of OBPs is not yet fully understood, and although OBP6 has been shown to bind sex pheromone components, these observations do not eliminate the potential role of OBPs as transporter proteins.[101,112,113] Furthermore, results from the biphasic GC assay suggest that OBP6 solubilises the sex pheromone components. Although it is difficult to fully elucidate the role of OBPs from these initial results, some levels of enantiomeric discrimination were observed, and these interactions should be investigated further. If true enantiomeric differences are seen, his would be the first observation of OBPs playing a discrimination role of this level.

In addition to understanding the role of OBPs, this work demonstrates successful prediction of interactions. *In silico* screening of ligands saves significant amounts of time, work and money, and should be considered as a valuable method for screening potential olfactory ligands for insects in future. The fluorescence, STD-NMR and GC assays were the most successful methods employed and should be optimised and repeated with the different enantiomers. Structural studies of OBP6 should be undertaken to generate a 3D structure and potentially observe the specific binding site of the sex pheromone components, as previously conducted for the alarm pheromone and OBP3 / OBP7.[104] Furthermore, this result could be tested in *in vivo* assays by knocking down or knocking out the gene for OBP6 in sexual aphids and determining whether it affects their ability to perceive and respond to the sex pheromone relative to wild-type aphids.

7.1.4 Enantiomeric Discrimination by OBPs

One of the main objectives of this thesis was to investigate whether OBPs played a role in enantiomeric discrimination in insect olfaction. The pea aphid's perception of the sex

pheromone appeared to be an excellent model for observing this possibility due to its behavioural enantiomeric discrimination between enantiomers of sex pheromone components reported previously.[145]

The enantiomers of the sex pheromone components were tested *in silico* and *in vitro* with fluorescence binding assays. As previously stated, there were no significant difference between either the sex pheromone components or their respective enantiomers. However, OBP6 did appear to bind the natural occurring lactone component, (4a*S*,7S,7a*R*)-nepetalactone **6**, with the lowest affinity, in both 1-NPN and probe-free assays. This difference was not significant in the analysis of variance (ANOVA), with a p values of 0.110 and 0.056 respectively, however, a t-test between the affinities for the two K_Ds gave a significant difference with the 1-NPN assay. There is clearly still more to elucidate about this interaction, which can hopefully be studied by alternative structural and binding methods.

Unfortunately, there are severe limitations in the screening of binding between different enantiomers. For most analytical methodologies, it is impossible to determine the difference between two enantiomers due to their identical physical properties in achiral environments. This is true of mass spectrometry and NMR spectroscopy in particular. The enantiomers could be tested individually, but a true assay comparing the two is not possible with either of these techniques. In STD NMR spectroscopy, mixtures of ligands can be introduced if they have distinct chemical shifts. This, unfortunately, rules out the screening of enantiomers by this method, unless methodology using a non-covalent chiral shift reagent can be developed. If one enantiomer was shown to be non-binding, with no peaks present in the final STD, a racemic mixture would show an intermediate amount of saturation. For the gas-chromatography assay, chiral GC is available. In future, the gas-chromatography method could be optimised to discriminate between the enantiomers.

In addition to binding studies, structural studies could give a greater understanding of the interactions of OBP6 and the sex pheromone component enantiomers. Solving the crystal structure of OBP6 with the components bound would allow for differences in the binding to be detected. Additionally, within the 2D heteronuclear NMRs, different residues may change when different enantiomers are introduced, indicating subtle differences in binding conformations.[108,117,200] Although both enantiomers may bind to OBP6, they bind in different active sites or induce different conformational changes. Observing this difference would suggest that OBPs play a larger role than in solubilisation and transport of odorants, and it is

potentially the OBP-ligand complex that activates the receptor or is critical for receptor activation in the olfactory process.

Overall, the work in this thesis preliminarily suggests that OBP6 may discriminate between the enantiomers of the sex pheromone components. Further work is required in future to explore the interactions of OBP6 in greater detail, in addition to potentially screening other OBPs and ORs to explore the ability of olfactory systems to discriminate between such structurally similar compounds in greater detail.

7.1.5 Applications of this work

There are many applications of this work. Primarily, this thesis explores the role of OBPs, specifically in sex pheromone perception in the pea aphid, *Acyrthosiphon pisum*. OBP6 was shown to be prolific at solubilising ligands from a non-aqueous to an aqueous layer, enforcing the role of OBPs as specific transport proteins. The work reported here will provide the foundation for further exploration of the interactions and specific role of OBP6, through both structural studies and *in vivo* assays. Knockdown studies have been performed with OBP3 and 7 in *A. pisum* previously *via* interference RNA, this could be repeated with OBP6.[99] Recent developments in knockout assays with CRISPR/Cas9 technology may also provide the means to remove the gene entirely from aphids. In both cases, behavioural changes in the absence of OBP6 could be observed.

In addition to contributing to a greater understanding of insect olfaction, knowledge of the interactions of OBPs could allow for screening of a wider range of ligands for their activity. Screening based on target proteins is a widely used technique in drug discovery, and this thesis has shown that this is possible with both i*n silico* methods and a range of *in vitro* methods. Firstly, many of the *in vitro* methods employed in this thesis are never (STD- NMR spectroscopy, SPROX mass spectrometry) or rarely (2D-NMR spectroscopy, ESI-MS)[122,195] utilised for studying insect OBPs. Though not all these methods were successful for the study of OBP6/OBP9, most had potential for optimisation for future assays. In addition to these *in vitro* assays, *in silico* screening of the OBPs in this thesis gave predicted results that were successfully shown *in vitro*. There were many unexplored results from the *in silico* screening, some of which may be novel and interesting interactions. In future, larger scale screening of OBPs with *in silico* methods should be considered as a viable and useful method for exploring the interactions of these proteins. From these screens, potential potent analogues were also designed, which could be used in pest control in future, and OBP6 may have potential as a biosensor to detect aphid sex pheromone and therefore aphids.

Finally, in addition to exploring the interactions of OBPs, odorant receptors (ORs) were also modelled and screened *in silico.* ORs remain an elusive group with only one available structure.[88] The screening provided a few leads for the potential OR responsible for sex pheromone perception in *A. pisum*. If these leads can be confirmed *in vivo*, these methods could be utilised to screen thousands of potential receptor/ligand combinations and reduce time spent on complex *in vivo* experiments.

7.2 GENERAL CONCLUSIONS AND FUTURE WORK

In conclusion, aphid sex pheromone components (4a*S,*7S,7a*R*)-nepetalactone **6** and (1*R*,4a*S,*7S,7a*R*)-nepetalactol **5** have been synthesised with significant optimisation of yields, and odorant-binding proteins from *A. pisum* have also been successfully produced. Screening of aphid olfactory proteins was also successful, and previously identified interactions were explored, in addition to new protein-ligand interaction leads being discovered, specifically the interaction between OBP6 and the aphid sex pheromone components.

The interactions between OBP6, the two sex pheromone components, (4a*S,*7S,7a*R*)-nepetalactone **6** and (1*R*,4a*S,*7S,7a*R*)-nepetalactol **5**, and their respective enantiomers, (4a*R,*7*R*,7a*S*)-nepetalactone **33** and (1*S*,4a*R,*7R,7a*S*)-nepetalactol **32**, were initially explored using fluorescence-based binding assays. The results of these assays confirmed the predicted results, though very little enantiomeric discrimination was observed. The binding of OBP6 to (4a*S,*7S,7a*R*)-nepetalactone **6** consistently had the highest affinity (and lowest K_Ds), and this interaction was explored further with NMR assays, including a successful assay using STD-NMR spectroscopy. Additionally, the interactions of OBP9 were explored using mass spectrometry methods, which were inevitably unsuccessful. Finally, both OBP6 and OBP9 were explored using a gas chromatography assay, where OBP6 appeared to bind or solubilise the sex pheromone components and the alarm pheromone, (*E*)-β-farnesene **17** better than the control.

Overall, the results of this thesis indicate that OBP6 may be involved in sex pheromone perception in the pea aphid, *A. pisum*. Though the focus of this work was the aphid species *A. pisum*, both (4a*S,*7S,7a*R*)-nepetalactone **6** and (1*R*,4a*S,*7S,7a*R*)-nepetalactol **5** act as sex pheromone components for multiple species of aphid. As with perception of the alarm pheromone in aphids, the elucidation of the role of an olfactory protein in one aphid species may be extended to others.[102,104,125] Furthermore, natural enemies of aphids often use the sex pheromone components to locate their aphid prey, and similar olfactory proteins may be

present in these species. In addition to this key finding, many other interactions and different or rarely used multi-disciplinary methodologies has been explored to understand the interactions of the aphid olfactory proteins.

Future work should initially explore the interaction of OBP6, the sex pheromone components and their respective enantiomers in greater detail. This may involve further optimisation of aforementioned techniques, site-directed mutagenesis of predicted key residues and structural studies, particularly x-ray crystallography. More work should be undertaken in understanding the enantiomeric discrimination ability of the pea aphid, initially *in vitro*. As this ability may arise from the odorant receptors and not the OBPs, these should be explored further. The results of the *in silico* screening could provide a basis for this. All determined results should also be tested *in vivo* via knockout or knockdown studies.

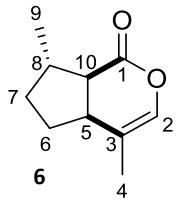
Finally, this thesis prompts a few potential pest control options, *via* the design of predicted potent analogues or screening of olfactory proteins for novel ligands. The pea aphid is a prevalent pest species worldwide, and the understanding gained of its olfactory system in this thesis should be utilised in managing aphid populations and mitigating crop losses worldwide.

8. Experimental

8.1 SYNTHETIC CHEMISTRY

8.1.1 Synthesis of (1R,4aS,7S,7aR)-4,7-dimethyl-1,4a,5,6,7,7a-hexahydrocyclopenta[c]pyran-1-ol and (1*R*,4a*S*,7*S*,7aR)-4,7-dimethyl-1,4a,5,6,7,7a-hexahydrocyclopenta[c]pyran-1-ol

8.1.1.1 Purification of (4aS,7S,7aR)-4,7-dimethyl-5,6,7,7a-tetrahydrocyclopenta[c]pyran-1(4aH)-one from steam-distilled Nepeta cataria oil



6

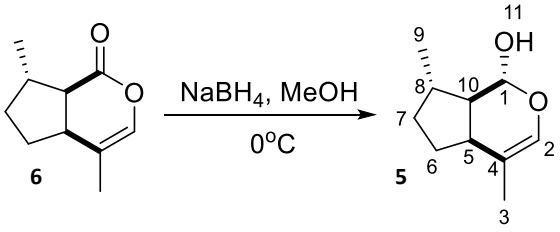
Steam-distilled oil from *Nepeta cataria*[6] (1.00 g) was purified on silica gel (3:2 EtOAc in Pet. Ether) to give a colourless oil containing (4a*S*,7*S*,7a*R*)-4,7-dimethyl-5,6,7,7a-tetrahydrocyclopenta[c]pyran-1(4aH)-one **6** (222 mg; 1.32 mmol). $[\alpha]_D$ +6.70 (c 3.3, CH_3OH); δ_H (500 MHz; $CDCl_3$) 1.21 (3H, d, *J* = 6.62 Hz, 9-H), 1.50-1.59 (1H, m, 5-H), 1.64 (3H, s, 4-H), 1.89-1.98 (2H, m, 6-H or 7-H), 2.02-2.11 (2H, m, 6-H or 7-H), 2.31-2.39 (1H, m, 8-H), 2.05 (1H, q, *J* = 7.79, 10-H), 6.18-6.20 (1H, m, 2-H); δ_C (125 MHz; $CDCl_3$) 15.6 (C4), 20.4 (C9), 31 (C6 or C7), 33.2 (C6 or C7), 39.8 (C8), 40.8 (C10), 115.2 (C3), 133.4 (C2), 171.0 (C1); *m/z* HRMS calculated for $[C_{10}H_{15}O_2]^+$ ($M+H^+$) 167.1066, found 167.1079. Data consistent with literature[213,214].

The compound was purified further using high-pressure liquid-chromatography (HPLC). A water and acetonitrile (ACN) solvent system was used (Table 3.1) with a semi-prep HPLC column (ACE 5 AQ V11-5053; 250 x 10 mm).

Table 8.1: High-pressure liquid chromatography (HPLC) method settings for purification of nepetalactol and nepetalactone with acetonitrile (ACN) and water. A flow rate of 5.0 mL min^{-1} was used.

HPLC Programming	
0 min	0% ACN
10 min	5% ACN → 100% ACN
40 min	100% ACN
55 min	5% ACN
Total Time	60 min

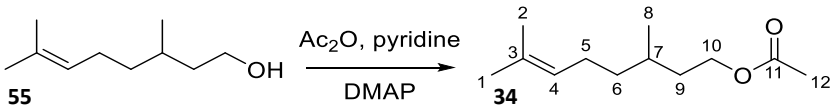
8.1.1.2 Reduction of (4aS,7S,7aR)-4,7-dimethyl-5,6,7,7a-tetrahydrocyclopenta[c]pyran-1(4aH)-one to (1R,4aS,7S,7aR)-4,7-dimethyl-1,4a,5,6,7,7a-hexahydrocyclopenta[c]pyran-1-ol



T2 oil, containing (4aS,7S,7aR)-4,7-dimethyl-5,6,7,7a-tetrahydrocyclopenta[c]pyran-1(4aH)-one **6**, (1.00g) was stirred in dry MeOH (20 mL) at 0°C. Sodium borohydride ($NaBH_4$; 497.8 mg; 13.2 mmol) in dry MeOH (10 mL) was added carefully to the mixture and stirred for a further 16 hours. Upon completion, the reaction was quenched with H_2O (10 mL), extracted with diethyl ether (10 mL), washed with H_2O (10 mL), dried over $MgSO_4$ and concentrated *in vacuo*, giving a crude oil containing **5** (167 mg; 0.99 mmol; 17%). $[\alpha]_D$ -14.4 (c 2.2, CH_3OH); δ_H (500 MHz; $CDCl_3$) 1.10 (1H, d, *J* = 7.06 Hz, 9-H), 1.31-1.40 (1H, m, 8-H), 1.57 (3H, s, 3-H), 1.61-1.73 (1H, m, 5-H), 1.78-2.03 (4H, m, 6-H and 7-H), 2.47 (1H, q, *J* = 7.79 Hz, 10-H), 4.85 (1H, m, 1-H), 6.03 (1H, s, 2-H). δ_C (125 MHz; $CDCl_3$) 16.2 (C3), 20.6 (C9), 30.4 (C6 or C7), 30.9 (C8), 35.4 (C6 or C7), 39.6 (C10), 50.7 (C5), 94.1 (C1), 113.8 (C4), 134.3 (C2); *m/z* HRMS calculated for $[C_{10}H_{17}O_2]^+$ ($M+H^+$) 168.1145, found 169.1265. Data consistent with literature[213,214].

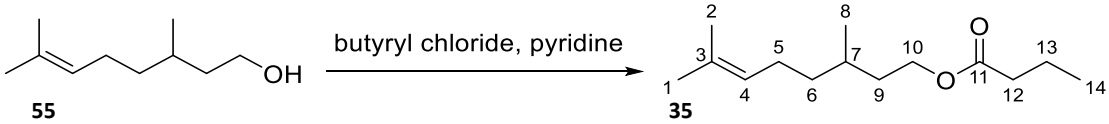
8.1.2 Racemic resolution of (*R/S)*-citronellol

8.1.2.1 Acetylation of (R/S)-citronellol to 3,7-dimethyloct-6-en-1-yl acetate



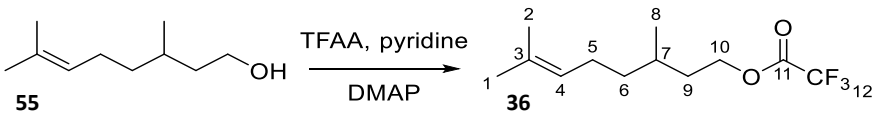
Acetic anhydride (Ac_2O; 0.12 mL; 1.23 mmol) and (*R/S*)-citronellol **55** (100 mg, 0.64 mmol) were added to a solution of 4-dimethylaminopyridine (DMAP; 2 mg; 0.016 mmol) in pyridine (10 mL) and the mixture stirred under N_2 for 12 hours. dH_2O (0°C, 10 mL) was added and the mixture extracted with EtOAc (30 mL) and washed with dH_2O (30 mL). The organic layers were concentrated *in vacuo* and the crude product purified on silica gel (1:5 EtOAc in Pet. Ether) to give a colourless oil, *3,7-dimethyloct-6-en-1-yl acetate* **34** (71 mg; 0.36 mmol; 56%). δ_H ($CDCl_3$, 500 MHz) 0.87 (3H, d, *J* 7.34 Hz, 8-H), 1.10-1.24 (2H, m, 6-H), 1.35-1.54 (1H, m, 9-H), 1.46-1.53 (1H, m, 7-H), 1.56 (3H, s, 1-H or 2-H), 1.65 (3H, s, 1-H or 2-H), 1.86-1.94 (2H, m, 5-H), 2.00 (3H, s, 12-H), 4.00-4.11 (2H, m, 10-H), 5.04 (1H, t, *J* = 7.15 Hz, 4-H); δc ($CDCl_3$, 125 MHz) 17.71 (C1 or C2), 19.49 (C8), 25.55 (C1 or C2), 25.71 (C5), 29.53 (C7), 35.5 (C9), 37.8 (C6), 63.02 (C10), 123.82 (C4), 131.39 (C3), 171.21 (C11), *m/z* HRMS calculated for $[C_{12}H_{21}O_2Na]^+$ ($M+Na^+$) 220.1434, found 220.1368.

8.1.2.2 Acetylation of (R/S)-citronellol to 3,7-dimethyloct-6-en-1-yl butyrate



Butyryl chloride (0.17 mL; 1.23 mmol) and (*R/S*)-citronellol **55** (100 mg; 0.64 mmol) were stirred in pyridine (5 mL) under N_2 for 12 hours. dH_2O (0°C; 10 mL) was added and the mixture extracted with EtOAc (30 mL) and washed with dH_2O (30 mL). The organic layers were concentrated *in vacuo* and the crude product purified on silica gel (1:5 EtOAc in Pet. Ether) to give a colourless oil, *3,7-dimethyloct-6-en-1-yl butyrate* **35** (95 mg; 0.42 mmol; 65%). δ_H ($CDCl_3$, 500 MHz) 0.89 (3H, d, *J* = 7.57, 8-H), 0.93 (3H, t, *J* = 7.48 Hz, 14-H), 1.12-1.19 (1H, m, 7-H), 1.22-1.35 (2H, m, 6-H), 1.37-1.47(2H, m, 13-H), 1.48-1.55 (2H, m, 9-H), 1.56 (3H, s, 1-H or 2-H), 1.58 (3H, s, 1-H or 2-H), 1.88-2.01 (2H, m, 5-H), 2.26 (2H, t, *J* = 7.58 Hz, 12-H), 4.04-4.14 (2H, m, 10-H), 5.06 (1H, t, *J* = 7.20, 4-H); δc ($CDCl_3$, 125 MHz) 13.4 (C14), 17.64 (C1 or C2), 18.59 (C1 or C2), 19.54 (C8), 25.44 (C5), 25.8 (C7), 29.43 (C9), 35.51 (C13), 36.54 (C12), 37.00 (C6), 62.86 (C10), 124.70 (C4), 131.44 (C3), 173.92 (C11). *m/z* HRMS calculated for $[C_{14}H_{26}O_2Na]^+$ ($M+Na^+$) 249.1825, found 249.1852.

8.1.2.3 Acetylation of (R/S)-citronellol to 3,7-dimethyloct-6-en-1-yl a2,2,2-trifluoroacetate



55 **36**

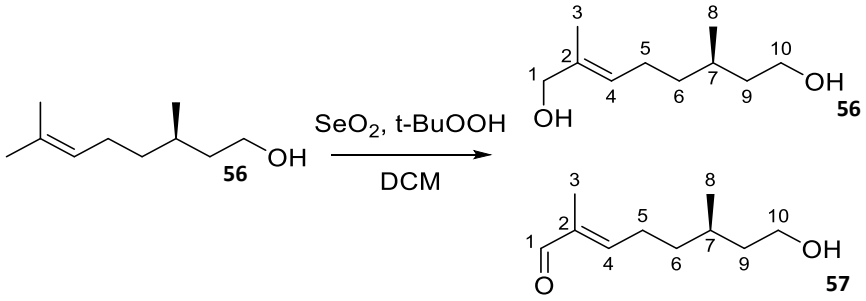
Trifluoracetic anhydride (TFAA; 0.17 mL; 1.23 mmol) was added dropwise to a solution of (*R/S*)-citronellol **55** (100 mg; 0.64 mmol) and 4-dimethylaminopyridine (DMAP; 1 mg; 0.008 mmol) in pyridine (5 mL) under N_2 at 0°C. The mixture was stirred for 10 min at 0°C, then warmed to RT and stirred for a further 40 min. ddH_2O (10 mL, 0°C) was added and the reaction stirred for a further 10 min. The mixture was then extracted with EtOAc (30 mL) and washed with ddH_2O (30 mL). The organic layers were concentrated *in vacuo* and purified over silica gel (1:4 EtOAc in pet ether) to give a colourless oil, *3,7-dimethyloct-6-en-1-yl a2,2,2-trifluoroacetate* **36** (121 mg; 0.50 mmol; 79%). δ_H ($CDCl_3$, 500 MHz) 0.93 (3H, d, *J* = 7.57 Hz, 8-H), 1.16-1.26 (2H, m, 6-H), 1.51-1.58 (2H, m, 9-H), 1.59 (3H, s, 1-H or 2-H), 1.67 (3H, s, 2-H or 2-H), 1.73-1.82 (1H, m, 7-H), 1.90-2.05 (2H, m, 5-H), 4.34-4.42 (2H, m, 10-H), 5.07 (1H, t, *J* = 7.15 Hz, 4-H); δc ($CDCl_3$, 125 MHz) 17.65 (C1 or C2), 19.28 (C8), 22.24 (C5), 25.27 (C9), 25.70 (C1 or C2), 34.84 (C7), 37.77 (C6), 66.78 (C10), 124.18 (C4), 131.71 (C3), 157.77 (q, J_{CF} = 42.44 Hz, C12), 172.46 (C11). *m/z* HRMS calculated for $[C_{12}H_{20}O_{23}]^+$ ($M+H^+$) 253.1410, found 253.1413.

8.1.2.4 Chiral gas-chromatography for separation of enantiomers

The three acetylated citronellol's, 3,7-dimethyloct-6-en-1-yl acetate **34**, 3,7-dimethyloct-6-en-1-yl butyrate **35** and 3,7-dimethyloct-6-en-1-yl a2,2,2-trifluoroacetate **36**, were analysed using a chiral gas chromatography (chiral-GC). Chiral-GC was performed using a chiral cool-on-column (β-cyclodextrin – 250.00 µm diameter x 50.0 m length) on an Agilent 6890N GC. A hydrogen carrier gas was utilised (rate of 1.9 mL min^{-1}), with a fixed temperature programme. The oven temperature programme of the GC, with initial temperature 30°C, was heated to 150°C at a rate of 5°C min^{-1}, followed by 230°C at a rate of 10°C min^{-1}, after which it was maintained at 230°C for 25.00 min. The final run time was 58.10 min.

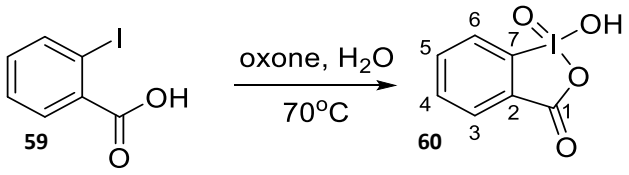
8.1.3 Synthesis of (1S,4aR,7R,7aS)-4,7-dimethyl-1,4a,5,6,7,7a-hexahydrocyclopenta[c]pyran-1-ol from (*R)*-citronellol

8.1.3.1 Oxidation of (R)-citronellol to (E)-2,(6R)-dimethylocta-2-ene-1,8-diol



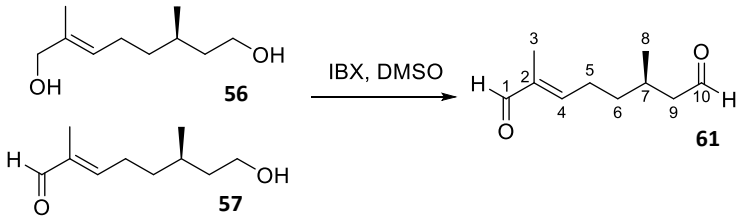
To a solution of (*R*)-citronellol **56** (1.00 g; 6.40 mmol) in DCM (20 mL) and under N_2, selenium dioxide (SeO_2; 71 mg; 0.64 mmol) was added, followed by *tert*-butyl hydroperoxide (*t*-BuOOH; 5M nonane; 1.6 mL; 8.00 mmol). The reaction was stirred at room temperature under N_2 for 72 hours, after which an aqueous solution of $Na_2S_2O_3$ (15 mL) was added and stirred vigorously for 15 min. The organic layers were separated, washed with $NaHCO_3$ (15 mL), dried over $MgSO_4$ and concentrated *in vacuo*. The crude product was then purified on silica gel (2:3 to 3:2 EtOAc in pet. ether) to give two colourless oils *(E)-2,(6,R)-dimethylocta-2-ene-1,8-diol* **56** (194 mg; 1.13 mmol; 18%). $[\alpha]_D$ +4.72 (c 1.8, CH_3OH); δ_H ($CDCl_3$, 500 MHz) 0.98 (3H, d, *J* = 6.45 Hz, 8-H), 1.33-1.45 (4H, m, 5-H and 6-H), 1.78 (3H, s, 3-H), 1.99-2.17 (1H, m, 7-H), 3.66-3.78 (4H, m, 9-H and 10-H), 4.02 (2H, s, 1-H), 5.42 (1H, t, *J* = 7.15 Hz, 4-H); δc ($CDCl_3$, 125 MHz) 9.28 (C3), 19.57 (C8), 25.05 (C7), 36.70 (C5 or C6), 39.82 (C5 or C6), 60.90 (C9), 61.13 (C10), 69.0 (C1), 126.5 (C4), 134.81 (C2); *m/z* HRMS calculated for $[C_{10}H_{20}O_2Na]^+$ ($M+Na^+$) 195.1356, found 195.1360. Data consistent with literature[214]. And *(E)-8-hydroxy,2,(6R/S)-dimethylocta-2-enal* **57** (197 mg; 1.15 mmol; 18%; combined yield 36%). δ_H ($CDCl_3$, 500 MHz) 0.94 (3H, d, *J* = 6.45 Hz, 8-H), 1.56-1.69 (4H, m, 5-H and 6-H), 1.70 (3H, s, 3-H), 2.32-2.47 (1H, m, 7-H), 3.66-3.78 (4H, m, 9-H and 10-H), 6.50 (1H, t, *J* = 7.15 Hz, 4-H), 9.41 (1H, s, 1-H); δc ($CDCl_3$, 125 MHz) 13.72 (C3), 19.66 (C8), 26.56 (C7), 35.62 (C5 or C6), 39.62 (C5 or C6), 60.90 (C9), 61.13 (C10), 139.37 (C2) 155.1 (C4), 195.5 (C1). The alol could not be fully isolated to provide full analytical data.

8.1.3.2 Synthesis of 2-iodoxybenzoic acid (IBX) from 2-iodobenzoic acid



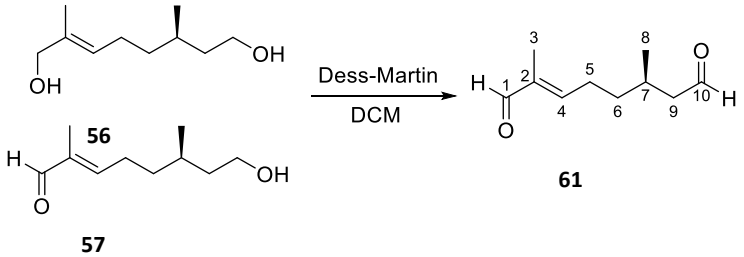
2-Iodobenzoic acid **59** (5 g; 20.00 mmol) was added to a mixture of Oxone (50% active ingredient; 35 g; 110 mmol) in H_2O (200 mL), and the mixture was heated to 70°C for 3 hours. The reaction was cooled to 0°C for a further 2 hours, before the precipitate was collected *via* filtration. The solid was rinsed with water (0°C; 50 mL) and acetone (0°C; 50 mL) before being dried for 48 hours in a vacuum desiccator, giving a white solid **60** (4.35g; 15.88 mmol; 79%). δ_H (*d*-DMSO, 500 MHz) 7.81 (1H, t, *J* = 7.54 Hz, 4-H), 7.92-8.02 (2H, m, 3-H and 5-H), 8.11 (1H, d, *J* = 8.36 Hz, 6-H); δ_C (*d*-DMSO, 125 MHz) 125.32 (C6), 130.47 (C3 or C5), 132.11 (C2 or C7), 133.32(C4), 133.68 (C3 or C5), 147.10 (C2 or C7), 167.91(C1). HRMS calculated for $[C_7H_6O_4I]^+$ ($M+H^+$) 280.9305, found 280.9249.

8.1.3.3 Oxidation of (E)-2,(6R)-dimethylocta-2-ene-1,8-diol and (E)-8-hydroxyl,2,(6R)-dimethylocta-2-enediol to (E)-2,(6R)-dimethyllocta-2-enediol with IBX



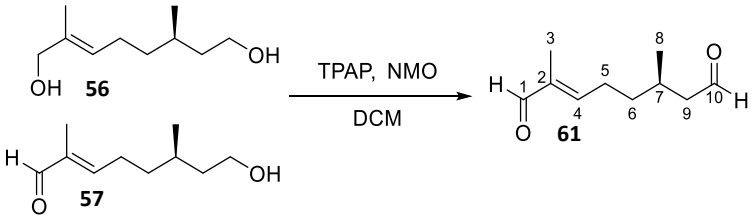
IBX **60** (1.60 g; 5.7 mmol) was added to a solution of *(E)*-2,(6*R*)-dimethylocta-2-ene-1,8-diol **56** (197 mg; 1.14 mmol) and *(E)*-8-hydroxy,2,(6*R*)-dimethylocta-2-enal **57** (194 mg; 1.14 mmol) in DMSO (30 mL). The mixture was stirred for 3 hours at room temperature, then diluted with diethyl ether (20 mL) and filtered over celite. The filtrate was diluted with H_2O (10 mL) and diethyl ether (10 mL), before being separated and washed with $NaHCO_3$ and brine. The extract was dried over $MgSO_4$, concentrated *in vacuo* the crude product purified by column chromatography on silica gel (3:2 EtOAc/Pet. ether) to give final product a colourless oil; *(E)-2,(6R)-dimethylocta-2-enediol* **61** (60 mg; 0.36 mmol; 16%). δ_H (500 MHz; $CDCl_3$) 1.05 (3H, d, *J* = 6.80 Hz, 8-H), 1.72 (3H, s, 3-H), 2.10-2.26 (1H, m, 7-H), 2.40-2.59 (4H, m, 5-H and 6-H), 2.36-2.51 (2H, m, 9-H), 5.10 (1H, t, *J* = 7.09 Hz, 4-H), 9.47 (1H, s, 1-H or 10-H), 9.75 (1H, s, 1-H or 10-H); Full analysis including ^{13}C NMR, HRMS and $[\alpha]_D$ could not be performed due to the instability of the dialdehyde.

8.1.3.4 Dess-Martin oxidation of (E)-2,(6R)-dimethylocta-2-ene-1,8-diol and (E)-8-hydroxyl,2,(6R)-dimethylocta-2-enediol to (E)-2,(6R)-dimethyllocta-2-enediol



Dess-Martin periodinane (3.44 g; 8.81 mmol) in DCM (10 mL) was added to a solution of *(E)*-2,(6*R*)-dimethylocta-2-ene-1,8-diol **56** and *(E)*-8-hydroxy,2,(6*R*)-dimethylocta-2-enal **57** (1.00 g combined; 5.87 mmol) in DCM (10 mL) under N_2. The mixture was stirred for 1 hour at room temperature, then diluted with $NaHCO_3$ (15 mL) and $Na_2S_2O_3$ (15 mL), before being separated and washed with $NaHCO_3$, brine and DCM. The extract was dried over $MgSO_4$, concentrated *in vacuo* the crude product purified by column chromatography on silica gel (1:1 EtOAc/pet. ether) to give final product a colourless oil; *(E)-2,(6R)-dimethylocta-2-enediol* **61** (535 mg; 3.17 mmol; 54%). δ_H (500 MHz; $CDCl_3$) 1.05 (3H, d, *J* = 6.80 Hz, 8-H), 1.72 (3H, s, 3-H), 2.10-2.26 (1H, m, 7-H), 2.40-2.59 (4H, m, 5-H and 6-H), 2.36-2.51 (2H, m, 9-H), 5.10 (1H, t, *J* = 7.09 Hz, 4-H), 9.47 (1H, s, 1-H or 10-H), 9.75 (1H, s, 1-H or 10-H); Full analysis including ^{13}C NMR, HRMS and $[\alpha]_D$ could not be performed due to the instability of the dialdehyde.

8.1.3.5 TPAP oxidation of (E)-2,(6R)-dimethylocta-2-ene-1,8-diol and (E)-8-hydroxy,2,(6R)-dimethylocta-2-enal to (E)-2,(6R)-dimethyllocta-2-enediol



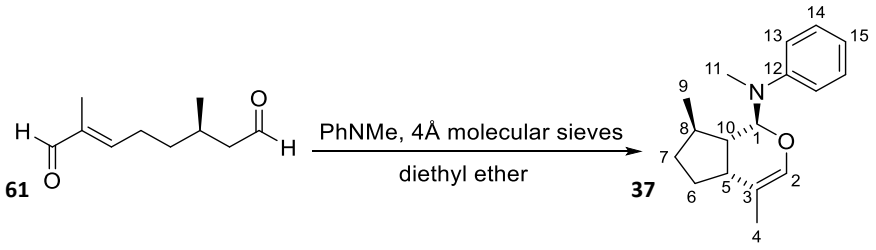
N-Methyl morpholine *N*-oxide (NMO; 1.07 g; 9.12 mmol), 4Å molecular sieves (1.00 g) and a mixture of *(E)*-2,(6*R*)-dimethylocta-2-ene-1,8-diol **56** (223 mg; 1.3 mmol) and *(E)*-8-hydroxy,2,(6*R*)-dimethylocta-2-enal **57** (167 mg; 0.98 mmol) were stirred in DCM (15 mL) for 10 min. Tetrapropylammonium perrutherate (TPAP; 38mg; 0.11 mmol) was then added and the mixture stirred for a further hour. The mixture was evaporated *in vacuo* to give the crude product. The crude product was then purified on silica gel (3:7 EtOAc in pet. ether) to give a colourless oil, *(E)-2,(6R)-dimethyllocta-2-enediol* **61** (111 mg; 0.66 mmol; 29%). δ_H (500 MHz; $CDCl_3$) 1.05 (3H, d, *J* = 6.80 Hz, 8-H), 1.72 (3H, s, 3-H), 2.10-2.26 (1H, m, 7-H), 2.40-2.59 (4H, m, 5-H and 6-H), 2.36-2.51 (2H, m, 9-H), 5.10 (1H, t, *J* = 7.09 Hz, 4-H), 9.47 (1H, s, 1-H or 10-

H), 9.75 (1H, s, 1-H or 10-H); Full analysis including ^{13}C NMR, HRMS and $[\alpha]_D$ could not be performed due to the instability of the dialdehyde.

8.1.3.6 Swern oxidation of (E)-2,(6R)-dimethylocta-2-ene-1,8-diol and (E)-8-hydroxy,2,(6R)-dimethylocta-2-enal to (E)-2,(6R)-dimethyllocta-2-enediol

Oxalyl chloride ($(COCl)_2$; 0.26 mL; 3.20 mmol) was stirred in DCM (10 mL) under N_2 and cooled to -78°C. DMSO (0.31 mL; 4.45 mmol) in DCM (5 mL) was added dropwise and the reaction stirred for a further 10 min. *(E)*-2,(6*R*)-dimethylocta-2-ene-1,8-diol **56** (333 mg; 1.78 mmol) was dissolved in DCM (5 mL) and added to the mixture, which was stirred for a further 45 min. Triethylamine (Et_3N; 1.24 mL; 8.90 mmol) was then added and the reaction allowed to warm to RT over 20 min. The white precipitate was dissolved in H_2O, the mixture extracted with DCM and washed with 2M HCl and brine, before being dried over $MgSO_4$ and concentrated *in vacuo*. The crude product was a yellow oil; *(E)*-2,(6*R*)-dimethylocta-2-enediol **61** (230 mg; 1.37 mmol; crude 77%)**.** δ_H (500 MHz; $CDCl_3$) 1.05 (3H, d, *J* = 6.80 Hz, 8-H), 1.72 (3H, s, 3-H), 2.10-2.26 (1H, m, 7-H), 2.40-2.59 (4H, m, 5-H and 6-H), 2.36-2.51 (2H, m, 9-H), 5.10 (1H, t, *J* = 7.09 Hz, 4-H), 9.47 (1H, s, 1-H or 10-H), 9.75 (1H, s, 1-H or 10-H); Full analysis including ^{13}C NMR, HRMS and $[\alpha]_D$ could not be performed due to the instability of the dialdehyde.

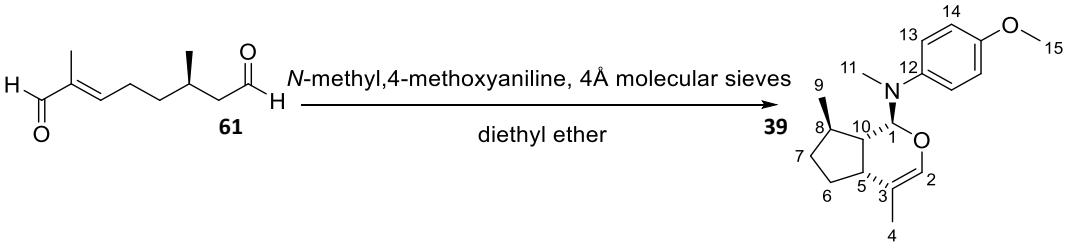
8.1.3.7 Cyclisation of (E)-2,(6R)-dimethylocta-2-enediol to (1S,4aR,7R,7aS)-N,4,7-trimethyl-N-phenyl-1,4a,5,6,7,7a-hexahydrocyclopenta[c]pyran-1-amine



(E)-2,(6*R*)-dimethylocta-2-enediol **61** (930 mg; 5.53 mmol) and 4Å molecular sieves (600 mg) were stirred in diethyl ether (10 mL) under N_2 for 30 minutes, after which *N*-methyl aniline (0.6 mL, 5.53 mmol) was added. The mixture was stirred for a further 16 hours at RT, then filtered over celite and concentrated *in vacuo*. The crude product was purified on silica gel (1:19 EtOAc:Pet ether) giving a product of a yellow oil (0.68 g; 2.60 mmol; 47%; de 72%) containing *(1S,4aR,7R,7aS)-N,4,7-trimethyl-N-phenyl-1,4a,5,6,7,7a-hexahydrocyclopenta[c]pyran-1-amine* **37**. $[\alpha]_D$ -13.6 (c 0.96, CH_3OH); δ_H ($CDCl_3$, 500 MHz)

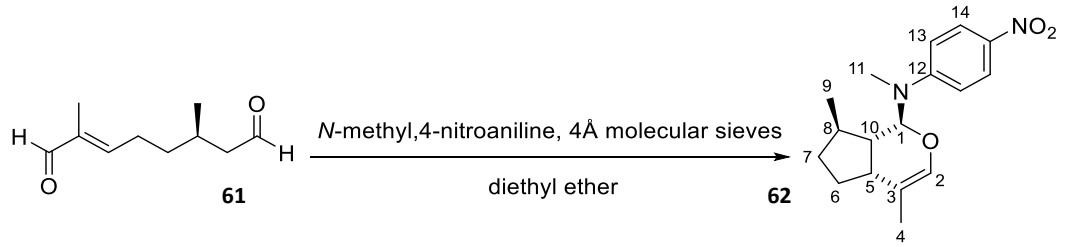
1.09 (1H, d, *J* = 6.62 Hz, 9-H), 1.64 (3H, s, 4-H), 1.82-1.94 (2H, m, 6-H and 7-H), 2.07-2.16 (2H, m, 5-H and 10-H), 2.36-2.50 (1H, m, 8-H), 2.99 (3H, s, 11-H), 4.72 (1H, d, *J* = 10.23 Hz, 1-H), 6.24 (1H, s, 2-H), 6.85 (1H, t, *J* = 7.41 Hz, 14-H or 15-H), 6.97 (2H, d, *J* = 8.15 Hz, 13-H), 7.24-7.30 (2H, m, 14-H or 15-H); δ_C ($CDCl_3$, 125 MHz) 17.06 (C4), 21.56 (C9), 32.68 (C5 or C10), 33.25 (C6 or C7), 36.46 (C6 or C7), 42.2 (C8), 45.54 (C5 or C10), 88.36 (C1), 133.08 (C3), 115.78 (C13), 119.22 (C14 or C15), 129.02 (C14 or C15), 137.68 (C2), 150.77 (C12); *m/z* HRMS calculated for $[C_{17}H_{24}ON]^+$ ($M+H^+$) 258.1852, found 258.1831.

8.1.3.8 Cyclisation of (E)-2,(6R)-dimethylocta-2-enediol to (1S,4aR,7R,7aS)-N,4,7-trimethyl-N-(4-methoxyphenyl)-1,4a,5,6,7,7a-hexahydrocyclopenta[c]pyran-1-amine



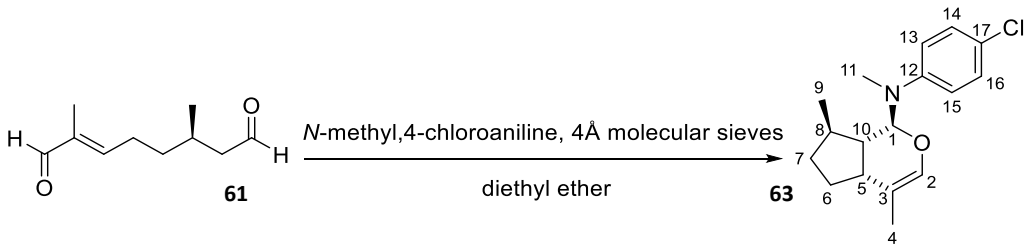
(E)-2,(6*R*)-dimethylocta-2-enediol **61** (150 mg; 0.89 mmol) and 4Å molecular sieves (200 mg) were stirred in diethyl ether (10 mL) under N_2 for 30 minutes, after which *N*-methyl, 4-methoxyaniline (122 mg; 0.89 mmol) was added. The mixture was stirred for a further 32 hours at RT, then filtered over celite and concentrated *in vacuo*. The crude product was a yellow oil containing *1S,4aR,7R,7aS)-N-(4-methoxyphenyl)-N,4,7-trimethyl-1,4a,5,6,7,7a-hexahydrocyclopenta[c]pyran-1-amine* **39** (156 mg; crude 61%; de 20%). δ_H ($CDCl_3$, 500 MHz) 1.07 (3H, d, *J* = 6.73 Hz, 9-H), 1.59 (3H, s, 4-H), 1.81-1.88 (1H, m, 6-H or 7-H), 1.89-1.97 (1H, m, 6-H or 7-H), 2.00-.12 (2H, m, 10-H and 5-H), 2.33-2.42 (1H, m, 8-H), 2.90 (3H, s, 11-H), 3.76 (3H, s, 15-H), 4.51 (1H, d, *J* = 10.30 Hz, 1-H), 6.20 (1H, s, 2-H), 6.81 (2H, d, *J* = 9.70 Hz, 13-H or 14-H), 6.96 (2H, d, *J* = 9.70 Hz, 13-H or 14-H); Full analysis including ^{13}C NMR, HRMS and $[\alpha]_D$ could not be performed due to the degradation of the product when purification was attempted.

8.1.3.9 Cyclisation of (E)-2,(6R)-dimethylocta-2-enediol to (1S,4aR,7R,7aS)-N,4,7-trimethyl-N-(4-nitrophenyl)-1,4a,5,6,7,7a-hexahydrocyclopenta[c]pyran-1-amine



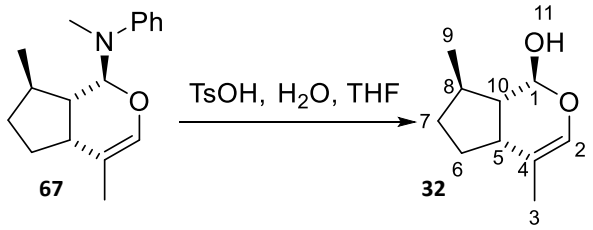
(E)-2,(6*R*)-dimethylocta-2-enediol **61** (200 mg; 1.78 mmol) and 4Å molecular sieves (300 mg) were stirred in diethyl ether (10 mL) under N_2 for 30 minutes, after which *N*-methyl, 4-nitroaniline (180 mg; 1.78 mmol) was added. The mixture was stirred for a further 72 hours at RT, then filtered over celite and concentrated *in vacuo* to give a crude product of a yellow oil containing **62** (290 mg). The crude product was analysed, revealing the reaction had not gone to completion. δ_H ($CDCl_3$, 500 MHz) 1.01 (3H, d, *J* = 6.75, 9-H), 1.75 (3H, s, 4-H), 2.20-2.47 (7H, m, 5-H, 6-H, 7-H, 8-H, 10-H), 3.05 (3H, s, 11-H), 4.81 (1H, d, *J* = 10.70, 1-H), 6.20 (1H, s, 2-H), 6.52 (2H, d, *J* = 9.43, 13-H or 14-H), 6.88 (4H, d, *J* = 9.83, 13-H or 14-H). Full analyses could not be performed due to the inability to isolate the product.

8.1.3.10 Cyclisation of (E)-2,(6R)-dimethylocta-2-enediol to (1S,4aR,7R,7aS)-N,4,7-trimethyl-N-(4-chlorophenyl)-1,4a,5,6,7,7a-hexahydrocyclopenta[c]pyran-1-amine



(E)-2,(6*R*)-dimethylocta-2-enediol **61** (950 mg; 5.64 mmol) and 4Å molecular sieves (200 mg) were stirred in diethyl ether (10 mL) under N_2 for 30 minutes, after which *N*-methyl, 4-chloroaniline (0.68 mL; 5.64 mmol) was added. The mixture was stirred for a further 16 hours at RT, then filtered over celite and concentrated *in vacuo*. The crude product was purified on silica gel (1:20 EtOAc:Pet ether) giving a product of a yellow oil (0.74 g; 45 %; 2.54 mmol; de 81%) containing *(1S,4aR,7R,7aS)-N,4,7-trimethyl-N-(4-chlorophenyl-1,4a,5,6,7,7a-hexahydrocyclopenta[c]pyran-1-amine* **63**. $[\alpha]_D$ – 17.7 (c 1.00, CH_3OH); δ_H ($CDCl_3$, 500 MHz) 1.06 (3H, d, *J* = 6.20, 9-H), 1.61 (3H, s, 4-H), 1.81-1.88 (2H, m, 6-H and 7-H), 2.04-2.15 (2H, m, 5-H and 10-H), 2.36-2.45 (1H, m, 8-H), 2.94 (3H, s, 11-H), 4.60 (1H, d, *J* = 10.34, 1-H), 6.19 (1H, s, 2-H), 6.53 (1H, d, *J* = 8.38, 13-H, 14-H, 15-H or 16-H), 6.87 (1H, d, *J* = 8.79, 13-H, 14-H, 15-H or 16-H), 7.13 (1H, d, *J* = 8.79, 13-H, 14-H, 15-H or 16-H), 7.18 (1H, d, *J* = 9.02, 13-H, 14-H, 15-H or 16-H). δ_C ($CDCl_3$, 125 MHz) 17.01 (C4), 21.53 (C9), 30.86 (C5 or C10), 33.03 (C11), 33.17 (C6 or C7), 36.52 (C6 or C7), 42.21 (C8), 45.48 (C5 or C10), 88.48 (C1), 113.33 (C3), 113.46 (C13, C14, C15 or C16), 117.07 (C13, C14, C15 or C16), 121.86 (C12 or C17), 124.08 (C12 or C17), 128.84 (C13, C14, C15 or C16), 129.03 (C13, C14, C15 or C16), 137.39 (C2); *m/z* HRMS calculated for $[C_{17}H_{23}ONCl]^+$ ($M+H^+$) 292.1463, found 292.1450.

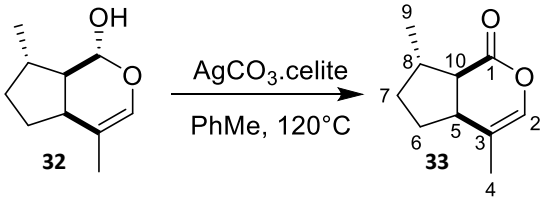
8.1.3.11 Hydrolysis of N,4,7-trimethyl-N-phenyl-1,4a,5,6,7,7a-hexahydrocyclopenta[c]pyran-1-amine to (1S,4aR,7R,7aS)-4,7-dimethyl-1,4a,5,6,7,7a-hexahydrocyclopenta[c]pyran-1-ol



para-Toluensulfonic acid (25 mg; 0.13 mmol) was added to a solution of crude *N*,4,7-trimethyl-*N*-phenyl-1,4a,5,6,7,7a-hexahydrocyclopenta[c]pyran-1-amine (30 mg; 0.12 mmol) in THF (2.5 mL) and H_2O (2.5 mL) and the reaction mixture stirred for 90 min at RT. $NaHCO_3$ was then added (5 mL) and the organic layers extracted with EtOAc and 2M HCl, dried over $MgSO_4$ and concentrated *in vacuo*. The crude product was purified on silica gel (9:1 pet. ether to EtOAc), giving a colourless oil*, (1S,4aR,7R,7aS)-4,7-dimethyl-1,4a,5,6,7,7a-hexahydrocyclopenta[c]pyran-1-ol* (10 mg; 0.06 mmol; 50%). $[\alpha]_D$ + 3.67 (c 0.31, CH_3OH); δ_H (500 MHz; $CDCl_3$) 1.10 (1H, d, *J* = 7.06 Hz, 9-H), 1.31-1.40 (1H, m, 8-H), 1.57 (3H, s, 3-H), 1.61-1.73 (1H, m, 5-H), 1.78-2.03 (4H, m, 6-H and 7-H), 2.47 (1H, q, *J* = 7.79 Hz, 10-H), 4.85 (1H, m, 1-H), 6.03 (1H, s, 2-H). δ_C (125 MHz; $CDCl_3$) 16.2 (C3), 20.6 (C9), 30.4 (C6 or C7), 30.9 (C8), 35.4 (C6 or C7), 39.6 (C10), 50.7 (C5), 94.1 (C1), 113.8 (C4), 134.3 (C2); *m/z* HRMS calculated for $[C_{10}H_{17}O_2]^+$ ($M+H^+$) 168.1145, found 169.1265. Data consistent with literature[213,214].

The final product was further purified using high-pressure liquid-chromatography (HPLC) using water/acetonitrile as detailed in Table 3.1.

8.1.3.12 Oxidation of (1S,4aR,7R,7aS)-4,7-dimethyl-1,4a,5,6,7,7a-hexahydrocyclopenta[c]pyran-1-ol to (4aS,7S,7aR)-4,7-dimethyl-5,6,7,7a-tetrahydrocyclopenta[c]pyran-1(4aH)-one



(1*S*,4a*R*,7*R*,7a*S*)-*N*,4,7-trimethyl-*N*-(4-chlorophenyl-1,4a,5,6,7,7a-hexahydrocyclopenta[c]pyran-1-amine **32** (43 mg, 0.26 mmol) was added to silver carbonate (347 mg, 1.26 mmol) and celite (240 mg) in toluene (5 mL) and the reaction refluxed at 120°C under N_2 for 1 hour. The final mixture was cooled to RT, filtered over celite and concentrated *in vacuo.* The crude product as purified on silica gel (1:19 EtOAc in Pet. Ether), giving final

product of a colourless oil containing *(4aS,7S,7aR)-4,7-dimethyl-5,6,7,7a-tetrahydrocyclopenta[c]pyran-1(4aH)-one* **33** (23 mg, 0.14 mmol, 52%). $[\alpha]_D$ -17.10 (c 2.7, CH_3OH); δ_H (500 MHz; $CDCl_3$) 1.21 (3H, d, *J* = 6.62 Hz, 9-H), 1.50-1.59 (1H, m, 5-H), 1.64 (3H, s, 4-H), 1.89-1.98 (2H, m, 6-H or 7-H), 2.02-2.11 (2H, m, 6-H or 7-H), 2.31-2.39 (1H, m, 8-H), 2.05(1H, q, *J* = 7.79, 10-H), 6.18-6.20 (1H, m, 2-H); δ_C (125 MHz; $CDCl_3$) 15.6 (C4), 20.4 (C9), 31 (C6 or C7), 33.2 (C6 or C7), 39.8 (C8), 40.8 (C10), 115.2 (C3), 133.4 (C2), 171.0 (C1); *m/z* HRMS calculated for $[C_{10}H_{15}O_2]^+$ ($M+H^+$) 167.1066, found 167.1079.

The final product was further purified using high-pressure liquid-chromatography (HPLC) using water/acetonitrile as detailed in Table 3.1.

8.2 EXPRESSION AND PURIFICATION OF OBPS

8.2.1 Media Recipes

Media for use in molecular biology experiments were prepared as in Table 8.2.

Table 8.2: Media recipes for LB, 2xYT and LB agar.

Media	Components (+ 1L H_2O)
LB liquid media	8g Tryptone, 5g Yeast Extract, 5g NaCl
2xYT liquid media	16g Tryptone, 10g Yeast Extract, 5g NaCl
LB agar	30g LB agar mix (without NaCL), 5g NaCl

Media was sterilised by autoclaving at 120°C, then cooled before adding an appropriate antibiotic to a working concentration (Table 8.3). Agar was kept at 60°C until poured into plates and allowed to set.

Table 8.3: Working concentrations of antibiotics added to media.

Antibiotic	Working Concentration / μgml^{-1}
Ampicillin	100
Kanamycin	50

8.2.2 Transformation of *E. coli* BL21(DE3) Competent Cells with *A. pisum* OBP Plasmids

Plasmids were supplied by Dr Jing-Jiang Zhou and colleagues (Rothamsted Research). All genes were previously cloned into a vector plasmid, pET45b or pET15b, containing an ampicillin resistance cassette, or pNIC28-Bsa4, containing a kanamycin resistance cassette. Each plasmid contained a hexa-histidine (His_6) tag encoded at the *N*-terminus of the protein, to allow for nickel affinity purification.

To transform cells, 5 µL of plasmid was added to BL21(DE3) competent *E. coli* cells and cooled to 0°C. The mixture was allowed to sit for 30 minutes, after which it was heat-shocked at 42°C for 70 sec and cooled to 0°C for a further 5 minutes. LB liquid media (800 µL) was added to the sample, followed by incubation (250 rpm; 37°C) for 45 minutes. The sample was centrifuged (4000 rpm) for 2 minutes and the pellet resuspended in LB media (100 µL). The resuspension was spread onto an LB-agar plate containing the appropriate antibiotic and incubated (37°C) for 12 hours. A clonal cell template for PCR was made from the overnight culture by diluting a scraping from a small, single colony in H_2O (10 µL).

8.2.3 Polymerase Chain Reaction (PCR)

A mixture for PCR was made by combining the cell template (1 µL), appropriate forward and reverse primers (1 µL of each), H_2O (7 µL) and a Taq polymerase mix (10 µL) containing dNTP, Taq polymerase and PCR buffer (100 mM Tris-HCl, 500 mM KCl, 15 mM $MgCl_2$, pH 8.3). The mixture was heated at 94°C for 5 minutes, followed by 25 cycles of heating at 94°C for 30 seconds, to 50°C for 30 seconds, then 72°C. Finally, the mixture was taken to 72°C for 30 minutes and stored at 4°C. Gel electrophoresis with a 1% agarose gel was used to check the PCR product and determine whether the transformation was successful.

8.2.4 Recombinant *E. coli* BL21(DE3) Starter Culture

A scraping from clonal culture or glycerol stock sample (500 µL) was added to 2xYT liquid media (10 mL) and incubated overnight (37°C; 250 rpm) to generate a starter culture of *E. coli* BL21(DE3).

8.2.5 Protein Expression Test from Recombinant BL21(DE3) *E. coli*

Starter culture of recombinant BL21(DE3) *E. coli* cells (500 µL) was added to 2xYT liquid media containing the appropriate antibiotic (10 mL) and incubated (37°C; 250 rpm) until the OD_{600} values reached 0.5-0.6. Isopropyl β-D-1-thiogalactopyranoside (IPTG; 10 µL; 1 M) was added to induce expression, and the mixture incubated for a further 3-4 hours (37°C; 250 rpm).

A sample of the culture (2 mL) was removed and centrifuged (12000rpm; RT) for 2 min, and the cell pellet resuspended in SDS-PAGE loading buffer (600µL). Samples were then denatured at 95°C for 10 minutes, before loading onto an SDS-PAGE gel.

8.2.6 Large Scale Expression, Refolding and Purification of OBPs

2xYT liquid media (600 mL) was inoculated with starter culture (3 mL) and incubated (37°C, 250 rpm) until an OD_{600} of 0.7-0.8 was reached. IPTG (300 µL; 1 M) was added to induce expression and the mixture incubated for a further 3-4 hours (37°C, 250 rpm). Cells were harvested by centrifuging for 15 minutes (3500 rpm).

The cell pellet was resuspended in ice-cold TBS buffer (10 mL; 25 mM Tris, 500mM NaCl, pH 8.0) and incubated on ice for 10 minutes. The mixture was sonicated for 4 minutes on ice, then centrifuged for 30 minutes (35000 rpm, 4°C). The pellet was resuspended in TBS+0.2% Triton X-100 (10 mL). The sonication and centrifugation steps were repeated, and the pellet resuspended in 8M urea (1.5 mL) and 10 mM 1,4-dithiothreitol in 100 mM Tris (1.5 mL; pH 8). The mixture was incubated for 1 hour at room temperature, then rapidly diluted with 27 mL TBS+5:0.5 mM GSH:GSSG and incubated overnight at RT.

The overnight mixture was centrifuged for 10 minutes (3500 rpm) and the supernatant filtered through a 0.22 µm filter. OBPs were purified using Histrap® columns, preconditioned with 25 mM imidazole buffer. The filtrate from the refolding step (Chapter 3; 3.2.5) was passed through the column, leaving the histidine-tagged protein bound to the nickel. An imidazole buffer (500 mM) was used to displace the protein and fractions were collected and analysed using SDS-PAGE.

8.2.7 Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE)

The resolving gel (1.0 mm width; 5 mL) was prepared as in Table 8.4. Approximately 1.5 mL of 2-isopronaol was added to remove bubbles and the gel left to set. After 20 minutes, the isopropanol was removed, and the stacking gel added. A comb was used to form wells before leaving the gel to set for a further 20 minutes.

Table 8.4: Composition of resolving and stacking gels for SDS-PAGE.
*added last to catalyse polymerisation

	Volume / µL	
	Resolving gel (15%; 5 mL)	**Stacking Gel (3%; 2mL)**
H_2O	1770	1185
40% acrylamide	1875	150
1.5M Tris buffer (pH 8.8)	1250	0
0.5M Tris buffer (pH 6.8)	0	250
10% SDS buffer	50	25
10% ammonium persulfate*	50	25
TEMED*	5	3

Samples for the gel (20 µL) were prepared by combining protein samples (10 µL) with sample buffer (10 µL; prepared as in Table 8.5). Samples were run alongside an appropriate marker (11-245 kDa). The gel was run at 200 V for 90-120 minutes and GelCode™ Blue Stain (Thermo-Fisher Scientific) was added for approximately 1 hour to resolve the protein bands.

Table 8.5: Composition of SDS-PAGE Protein Sample Buffer

Reagent	**Concentration**
Tris	80 mM
SDS	2.0%
Glycerol	10%
Bromophenol blue	0.0006%
DTT	0.1 M

8.2.8 Histidine-Tag Cleavage from Purified Protein

To remove the hexa-histidine tag from purified proteins, an appropriate cleavage enzyme was used (Table 8.6). The selection of the enzyme depended on the cleavage site encoded in the original vector plasmid. The protein was buffer-exchanged in $CaCl_2$ in TBS buffer (2 mM $CaCl_2$, 25 mM Tris; 500 mM NaCl), and enzyme added. Cleavage was monitored by observing a change in mass using mass spectrometry. After completion, excess tagged protein was removed by passing the mixture through a Histrap® column and collecting the eluent.

Table 8.6: Vector plasmids used for transformation and their his-tag cleavage sites and enzymes.

Plasmid	His-Tag FASTA sequence	Cleavage Site	Cleavage Enzyme
pET45b	MAHHHHHHVG TGSNDDDDKS PDP	DDDDK/S	Enterokinase
pNIC28-Bsa4	MHHHHHHSSGVDLGTENLYFQS	ENLYFQ/S	TEV (Tomato Etch Virus)
pET15b	MGSSHHHHHHSSGLVPRGSH	GLVPR/GS	Thrombin

8.2.9 Plasmid Extraction, Purification and Sequencing

Plasmids were extracted using a GeneJET Plasmid Miniprep Kit (Thermo-Fisher Scientific) and sequenced by Queen's Medical Centre laboratories, University of Nottingham or Eurofins (UK).

8.2.10 Protein Buffer Exchange and Concentration

All proteins were buffer exchanged and desalted using a PD-10 desalting column and concentrated using a Vivaspin 20 (10 kDa MWCO).

8.2.11 Site-Directed Mutagenesis

Site-directed mutagenesis (SDM) of OBP6-His_6 was performed to insert a thrombin tag. A Q5 Site-Directed Mutagenesis kit (New England Biolabs) was used to perform the mutagenesis. Primers (Table 8.7) were obtained from Sigma-Aldrich.

Table 8.7: Forward and reverse primers for site-directed mutagenesis (thrombin cleavage site insertion) of OBP6-His_6

Primers for thrombin cleavage site insertion	
Forward	cgcggcagcGCTGGGTACGATAGAACATG
Reverse	cggcaccagCGGATCCGGACTCTTGTC

Samples for PCR were prepared by combining plasmid containing the gene of interest (1 µL), forward and reverse primers (1 µL of each), H_2O (9 µL) and Q5 hot start master mix (12.5 µL). The mixture was heated at 98°C for 30 seconds, followed by 25 cycles of heating at 98°C for 10 seconds, to 64°C for 20 seconds, then 72°C for 2 minutes. Finally, the mixture was taken to 72°C for 2 minutes and stored at 4°C.

The PCR product (1 µL) was mixed with KLD reaction buffer (5 µL), KLD enzyme mix (1 µL) and H_2O (3 µL) and left at RT for 5 minutes. Finally, the mixture (5 µL) was used to transform NEB 5-alpha Competent *E. coli* cells as in 8.2.2. Results were validated by sequences as in 8.2.9.

8.2.12 Fast-Protein Liquid Chromatography

To purify proteins, gel filtration *via* fast-protein liquid chromatography (FPLC) was performed using an Akta FPLC. The FPLC was fitted with a Superdex size-exclusion column, and samples were exchanged into a TBS buffer (Tris 15 mM, NaCl 250 mM; Table 8.8). Protein samples were collected using an autosampler and subsequently concentrated as 8.2.10.

Table 8.8: Run conditions for size exclusion fast-protein liquid chromatography (FPLC).

Step	Time	Flow Rate	Buffers
Wash	20 min	3.0 mL min^{-1}	1:1 20% EtOH:H_2O
Wash	20 min	3.0 mL min^{-1}	1:1 TBS:H_2O
Equilibrate	40 min	1.5 mL min^{-1}	TBS
Inject Sample			
Run	60-180 min	1.5 mL min^{-1}	TBS
Wash & Store	40 min	3.0 mL min^{-1}	20% EtOH

8.3 *IN SILICO* MODELLING AND DOCKING

8.3.1 Sequence Data and Protein Models

A. pisum OBP and OR sequences were previously reported by a variety of literature (Table 8.9).[103]

Table 8.9: NCBI accession codes for *Acyrthosiphon pisum* olfactory proteins

***A. Pisum* Protein**	**Residues**	**NCBI Ascension Code**
OBP1	159 aa	NP_001153526.1
OBP2	243 aa	NP_001153528.1
OBP3	141 aa	NP_001153529.1
OBP4	193 aa	NP_001153530.1
OBP5	221 aa	NP_001153531.1
OBP6	160 aa	NP_001153532.1
OBP7	155 aa	NP_001153533.1
OBP8	162 aa	NP_001153534.1
OBP9	165 aa	NP_001153535.1
OBP10	143 aa	NP_001153525.1
OBP11	141 aa	XP_008178459.1
OR1 (ORCO)	463 aa	AQS60741.1
OR2	403 aa	AQS60742.1
OR4	368 aa	ARJ54248.1
OR5	367 aa	KX890157.1
OR10	369 aa	AQS60745.1
OR17	430 aa	AQS60746.1
OR20	420 aa	AQS60747.1
OR22c	403 aa	XP_003245950.2
OR39	426 aa	AQS60753.1

All previously published protein structures were accessed from the Protein Data Bank (PDB).[215]

8.3.2 Sequence Alignment and Transmembrane Domain Prediction

Sequence alignments were performed using Cluster Omega and the PRALINE server.[216,217] Conservation mapping was performed using ConSurf. Transmembrane domain predications of the receptor proteins were performed using a consensus approach and four different serves – HMMTOP, TMpred, PHOBIUS and TMHMM.[218–221] Alignments and conservation

maps were analysed in GeneDoc[222], and phylogenetic trees were generated in FigTree v1.4.3.[223]

8.3.3 Homology Modelling

For odorant-binding proteins, protein structures were initially predicted using the iTASSER database[141], which takes a hierarchical approach by identifying structural templates from the Protein Data Bank. All predicted protein structures were minimised using the Yasara minimisation server.[170]

For odorant receptors, homology based modelling was performed using the olfactory receptor co-receptor, ORCO, from *Apocrypta bakeri* as a template. The structure of AbakORCO was obtained from the Protein Data Bank (PDB ID 6C70; resolution 3.5 Å).[88] Pairwise sequence alignment was performed using PRALINE[217] and the predicted transmembrane domains were manually aligned and annotated. The available ORCO structure shares a generally low sequence identity with the *A. pisum* odorant receptors, however, the general seven transmembrane structure should be conserved. The pairwise alignment served as a template for homology modelling using MODELLER 9.3 with loop refinement.[172] The secondary structure of long extracellular loops 2 and 3 (EL2 and EL3) were individually predicted using the iTASSER server.[141] Approximately 25 models were generated for each protein, and these were subsequently assessed using discrete optimised protein energy (DOPE) from MODELLER, in addition to Ramachandran and Z-score analysis, performed in PROCHECK and ProSA respectively.[171,173]

All proteins structures were visualised in in PyMol 2.3.4.[91]

8.3.4 Ligand Screening

Ligands were prepared in Chem3D 16.0 and AutoDock 4.2 (Python Molecule Viewer), then screened against computer generated models using AutoDock4.2 and the Racoon virtual screening tool.[161] A Lamarckian Genetic Algorithm was selected for the simulation.

OBPs and ORs were all screened against a wide range of ligands. The binding energy of the complex and subsequent K_i were calculated. In some cases, mutants were also generated to determine critical amino acid residues by manually mutating single residues.

8.3.4 Molecular Dynamics

Molecular dynamics simulations were performed using GROMACS 2019. For odorant-binding proteins, the OLPS force field was used, and for odorant receptors, a modified Gromas 53a6 force field was utilised.[168,169]

For odorant-binding proteins, models were solvated and neutralised with ions (for negative Cl-, for positive Na+) before energy minimisation and equilibration (temperature and pressure) calculations were performed. For each protein, a 1 ns simulation was performed. The stability of the protein was then assessed by observing temperature and pressure stabilisation of the simulated models.

Molecular dynamics simulations were performed using the GROMACS 9.3 package. For OBPs, and OLPS force field was used, whereas for odorant receptors, the Gromos53a6 forcefield was utilised, with some modifications made to parameters to include lipids. Lipid bilayer topology and structure were obtained from Dr. Peter Tielman.[224] Receptor models were embedded into a lipid bilayer of 128 dipalmitoylphosphatidylcholine (DPPC) molecules. All models were solvated with explicit solvent (SPC) water and neutralised with either Cl- or Na+ ions. Steepest-descent methods were used to minimise the energy of each system, at a reference temperature of 300 K. All bonds were constrained with LINCS and long-range electrostatics were handled using PME.

8.4 FLUORESCENT BINDING STUDIES

8.4.1 Fluorescence Measurements

All fluorescent measurements were undertaken using a Perkin-Elmer LS50B fluorescence spectrophotometer, using a 2mL quartz cuvette, and the settings described below in Table 8.10, unless otherwise stated. Spectra were recorded using FL WinLab software.

Table 8.10: Settings for fluorescence measurements using the Perkin-Elmer LS50B fluorescence spectrophotometer

Setting	Intrinsic Fluorescence (Tryptophan 54)	Probe Fluorescence (1-NPN 51)	Probe Fluorescence (1,8-ANS 52)
Excitation Wavelength	280 nm	337 nm	380 nm
Emission Wavelengths	290 – 400 nm	350 – 600 nm	400 – 600 nm
Excitation Slit	5.0 nm	5.0 nm	5.0 nm
Emission Slit	5.0 nm	5.0 nm	5.0 nm

8.4.2 Ligand Binding and Saturation Curves

Saturation of OBPs with fluorescent probe, 1-NPN **51** or 1,8-ANS **52** (both Sigma-Aldrich) was initially measured by titrating a 2 µM protein sample (2 mL in 25 mM Tris-HCl) with aliquots of 1 mM ligand in methanol to final concentrations of 1-16 µM. The fluorescence intensity was recorded.

The competitive binding of ligands was measured by observing the intrinsic fluorescence of tryptophan **54**. Titrations were performed with aliquots of 1 mM ligand in methanol to final concentrations of 1-20 µM, either after the addition of fluorescent probe to a final concentration of 1 µM or in the absence of fluorescent probe.

8.4.3 Analysis of Fluorescence data

To generate K_D values, relative fluorescence intensity was plotted against the concentration of ligand as a binding curve. K_D values were generated in GraphPad Prism 7 using a non-linear regression and the inbuilt equation:

$$y = \frac{B_{max} * x}{K_D + x}$$

Each calculated K_D value had an associated error from the non-linear regression. To account for these errors in statistical analysis, values were weighted in direct proportion to their error – the higher the error, the lower the weighting of the values. The 'weight' factor was calculated using the following equation:

$$weight = \frac{1}{(SE)^2}$$

Statistical analysis of fluorescence data was performed using R. A one-way weighted ANOVA was performed between ligands for each protein, and a two-way weighted ANOVA was performed to investigate the interactions between proteins and ligands. A Tukey Test was used for post-hoc analysis (Table 8.11).

Table 8.11: Statistical analysis results for a weighted analysis of variance (ANOVA) of the binding affinities (K_D) difference between different ligands and OBP6.

Assay	Interaction		*p*-value
	Ligand A	Ligand B	
1-NPN	(1*R*,4a*S*,7*S*,7a*R*)-Nepetalactol **5**	(*E*)-β-Farnesene **17**	0.0021
	(4a*S*,7*S*,7a*R*)-Nepetalactone **6**		0.00017
	(1*S*,4a*R*,7*R*,7a*S*)-Nepetalactol **32**		0.010
	(4a*R*,7*R*,7a*S*)-Nepetalactone **33**		0.015
	(1*R*,4a*S*,7*S*,7a*R*)-Nepetalactol **5**	(±)-Linalool	0.023
	(4a*S*,7*S*,7a*R*)-Nepetalactone **6**		0.0014

The interaction between (4a*S*,7*S*,7a*R*)-nepetalactone **6** and (4a*R*,7*R*,7a*S*)-nepetalactone **33** did not give a significant difference, with a *p*-value of 0.11 in the 1-NPN assay and a *p*-value of 0.055 in the fluorescent probe free assay. A weighted t-test was subsequently performed, giving a significant difference with a *p*-value of 2.27 x 10^{-4}.

8.4.3 Thermostability Assay

Thermostability measurements were performed using a NanoTemper NanoDSF fluorescence-based protein stability instrument using a 0.8 $mgml^{-1}$ sample of *A. pisum* OBP6.

8.5 MASS SPECTROMETRY

8.6.1 General Mass Spectrometry

All mass spectrometry was undertaken using electrospray ionisation and conducted on the QTof2 or QTof3 under denaturing conditions (80% CH_3CN, 0.1% formic acid) and the settings detailed in Table 8.11, unless otherwise stated.

Mass spectra were analysed using MassLynx 3.0 software.

Table 8.11: Denatured mass spectrometry settings for the QTof 3.0

Setting	Value
Capillary Voltage	2.80 kV
Cone Voltage	35 V
Desolvation Temperature	80°C
Source Temperature	50°C
Rf Lens	1.0

8.6.2 Sample Preparation using the ZipTip®

Samples for mass spectrometry were prepared by ZipTip®. The ZipTip® was conditioned by washing with elution buffer (80% CH_3CN, 0.1% formic acid; 5x10 µL), followed by equilibrating with wash buffer (5% MeOH, 0.1% trifluoroacetic acid; 7x10 µL). Trifluoroacetic acid (TFA) was added to the sample to a concentration of 0.1%, which was adsorbed onto the column (20 aspirations). The sample was then washed with wash buffer (20x10 µL) and eluted into elution buffer (10 µL, 20 aspirations).

8.6.3 Stability of Proteins from Rates of Oxidations (SPROX)

Samples were prepared using a constant concentration of protein against a gradient of concentrations of a denaturant (guanidinium chloride; 0-6M) as seen in Table 3.8.

Table 8.12: Sample composition for SPROX reactions with guanidinium chloride concentrations 0-6M

[Guanidinium Chloride] / M	Volume Added / µL		
	Sample	Water	Guanidinium Chloride (8M)
0	7.5	22.5	0
1	7.5	18.75	3.75
2	7.5	15	7.5
3	7.5	11.25	11.25
4	7.5	7.5	15
5	7.5	3.75	18.75
6	7.5	0	22.5

To each of these, hydrogen peroxide (1 µL, 9.8 M) was added. After 90 seconds, the reaction was quenched by ZipTip®, and an ESI-MS recorded. Each SPROX reaction was conducted for 3 repetitions to obtain an average.

Mass spectrometry data was analysed by calculating the average number of oxidations for the protein. Initially, weighted average mass, $m/z_{wt,av}$, was calculated using the mass (m_i) and intensity (I_i) of peaks for different oxidation states, including the native state, as follows:

$$m/z_{wt,av} = \sum m_i I_i / \sum I_i$$

This was then converted into the average weighted change, $\Delta\text{Mass}_{wt,av}$ in mass by subtracting the value from the mass of the native protein, Mass_{tot}, as follows:

$$\Delta\text{Mass}_{wt,av} = \text{Mass}_{tot} - m/z_{wt,av}$$

Finally, average number of oxidations, OXI_{av}, was calculated by dividing the average weighted change in mass by the molecular weight of oxygen, 16.

$$\mathrm{OXI}_{av} = \Delta\mathrm{Mass}_{wt,av} / 16$$

Data was plotted using Python2.7. A four-parameter logistic curve was used to fit a sigmoidal trend line and calculate the midpoint, $C_{1/2}$:

$$f(x) = \frac{1}{1 + e^{-x}}$$

After the midpoint concentration was found, further SPROX reactions were undertaken using a wider variety of concentrations close to this value. For example, if the midpoint value was found to be 3.0 M, a range from 2.5-3.5 M concentrations were conducted.

[GdmCl] / M	Average no. of oxidations
0	0.80 ± 0.03
1	0.82 ± 0.10
2	0.83 ± 0.05
2.5*	1.21
2.75*	1.30
3	1.60 ± 0.04
3.1*	2.50
3.25*	2.72
3.5*	2.58
4	2.66 ± 0.05
5	2.49 ± 0.09
6	2.69 ± 0.05

8.6 NUCLEAR MAGNETIC RESONANCE (NMR) SPECTROSCOPY

8.7.1 General NMR spectroscopy

Samples were run using an AVANCE Bruker DRX-500 MHz Nuclear Magnetic Resonance spectrometer with a 5 mm BBO BB-1H probe and set at 500 MHz for ^{1}H spectra and 125 MHz for ^{13}C spectra. Analysis of Bruker data was performed using Topsin 4.0.7, and analysis of Varien data performed with CCPNMR V2 and NMRPipe.[225,226] Standard NMR tubes were used with a sample volume of 600 µL unless otherwise stated. Deuterated chloroform ($CDCl_3$), methanol (CD_3OD) and dimethyl sulfoxide (*d*-DMSO) were stored over 4 Å molecular sieves

and used as both sample solvents and internal standards. For assignment of small molecules, additional 2D-NMR spectroscopy experiments were performed.

8.7.2 Standard sample preparation

Protein samples for NMR were desalted and buffer exchanged into 9:1 H_2O:D_2O unless stated otherwise (as 8.2.10). Ligand samples were not soluble in D_2O and samples were prepared in *d*-DMSO. This resulted in a final NMR solvent including H_2O, D_2O and *d*-DMSO.

8.7.3 ^{15}N Labelled Media Recipes

Stock solutions to generate ^{15}N labelled minimal media for NMR experiments were prepared as in Table 8.13.

Table 8.13: Stock solutions for ^{15}N labelled minimal media

Media/Stock	Component	Amount (+ 1L H_2O)
M9 Media Mix X10	$N^{15}H_4Cl$	5 g
	NaCl	5 g
	Na_2HPO_4	60 g
	KH_2PO_4	30 g
Trace Elements X100	EDTA	5 g
	$FeCl_3.6H_2O$	0.83 g
	$ZnCl_2$	84 mg
	$CuCl_2.2H_2O$	13 mg
	$CoCl_2.6H_2O$	10 mg
	H_3BO_3	10 mg
	$MnCl_2.6H_2O$	1.6 mg
1M $CaCl_2$		110.98 g
1M $MgSO_4$		120.37 g
1 mg ml^{-1} Biotin		1 g
1 mg ml^{-1} Thiamin		1 g
20% *w/v* Glucose		200 g

All stock solutions were prepared with autoclave sterilised, distilled H_2O and passed over a 0.22 µm filter for further sterilisation. The stock solutions were combined fresh for large scale protein expression as in Table 8.14.

Table 8.14: Final ^{15}N labelled minimal media recipe

Stock Solution	Amount
M9 Media Mix X10	100 mL
Trace Elements X100	10 mL
1M $CaCl_2$	0.3 mL
1M $MgSO_4$	1 mL
1 mg ml^{-1} Biotin	1 mL
1 mg ml^{-1} Thiamin	1 mL
20% w/*v* glucose	20 mL
ddH_2O	866.7 mL

8.7.4 Production and purification of 15-N labelled protein

Production of 15-N labelled protein was undertaken with the Marley method.[199] Proteins were grown from a starter culture as in 8.2.4 and 8.2.5. When an OD_{600} of 0.5-0.6 was reached, the culture was centrifuged (5000 rpm; 30 min; RT) to produce a pellet. This pellet was subsequently washed with M9 salts (M9 media mix with no $N^{15}H_4Cl$), pelleted again (5000 rpm; 15 min; RT) and resuspended in 1 L of labelled M9 minimal media. The mixture was incubated for 1 hour (37 °C; 180 rpm), then induced with ITPG (1 mL; 1 M). The culture was left to incubate for 4 hours (37 °C; 180 rpm) to allow for protein production. Once incubated, proteins were purified and analysed as standard (8.2.6-8.2.9).

8.6.5 ^{1}H-^{15}N HSQC NMR of Singly Labelled Protein

For ^{1}H-^{15}N HSQC NMR, samples were run on a Bruker Avance Neo 600 MHz and a Varian 900 MHz NMR with a 5 mm triple resonance 1H/13C/15N cold probe with Z-axis gradient. Samples were run in a 5 mm Shigemi tube (sample size approximately 300 µL). Each sample was gently mixed and centrifuged (14000 rpm, 10 min) before adding to the sample tube. For ligand titrations, the sample was removed from the Shigemi tube, ligand added, gently mixed and centrifuged (14000 rpm, 10 min). For each sample, the NMR was manually tuned, check with a proton (^{1}H) 1-dimensional NMR, before a TROSY NMR was run.

8.7.6 Saturation Transfer Difference (STD) NMR

An initial test assay containing Bovine Serum Albumin (BSA) protein with tryptophan **54** and sucrose was prepared for STD NMR. Samples were prepared as in Table 8.15. After running an initial ^{1}H NMR spectra, a saturation transfer difference (STD) programme was used. Each

run used the 'stdiff' pulse programme provided by Bruker, ran for 192 scans with a saturation time of 3 s and a total relaxation delay of 5.17 s - 3.17 s acquisition time and 2 s delay (D1). An on-resonance frequency of 192 Hz, with a control frequency of -12000 Hz, were used. Finally, gradient pulsed water suppression was used.

Table 8.15: Composition of the final STD assays, both the BSA test assay and OBP6 assay.

Component	Test Assay	OBP6 Assay
Protein (Unlabelled)	BSA; 100 µM	OBP6; 30 uM
Suspected Binder	Tryptophan **54**; 10 mM	(4aS,7S,7aR)-nepetalactone **6**; 3 mM
Suspected Non-Binder	Sucrose; 10 mM	Not Included

Spectra were processed and analysed in TopSpin 4.0.7. STD absolute values were calculated by observing the change in proportions between the off-resonance spectrum and the final STD spectrum using the following equation:

$$STD\ absolute\ value = \frac{I_0 - I_{STD}}{I_0}$$

in which the term (I_0 - I_{STD}) represents the ratio of peak intensity in the STD spectrum and I_0 the ratio of intensity in the off resonance spectrum.[206] A second value representing the proportionate change was calculated using the following equation:

$$Difference\ in\ proportion = I_0 - (I_0 - I_{STD})$$

Finally, epitope mapping was performed by calculating the relative peak integration in the STD spectrum versus the off-resonance spectrum of a peak as a percentage.

8.7 BIPHASIC ASSAY WITH GAS CHROMATOGRAPHY

8.7.1 General Gas Chromatography

An Agilent 6890A GC fitted with a non-polar HP1 equipped with a cool-on-column injector and a flame ionisation detector was used for gas chromatography. Unless otherwise stated, 2 µL samples were injected.

Biphasic GC was performed using a HP-1 column (320.00 µm diameter x 46.2 m length). A hydrogen carrier gas was utilised (rate of 3.1 mL/min), with a fixed temperature programme. The oven temperature programme of the GC, with initial temperature 30°C, was heated to 100°C at a rate of 5°C min^{-1}, maintained at 100°C for 10 min, then heated to 250°C at a rate

of 10°C $min^{-1,}$ after which it was maintained at 250°C for 45 min. The final run time was 84.10 min.

8.7.2 Preliminary Biphasic Assay with OBP9

For the preliminary biphasic assay with OBP9, a solution of OBP9 (0.5 mL of 1.22 µM in 25 mM Tris) was added to a small vial (2 mL size). A ligand solution (0.5 mL of 0.02 $mgml^{-1}$ solution of (4a*S,*7*S,*7a*R*)-nepetalactone **6**, (1R,4a*S,*7*S,*7a*R*)-nepetalactol **5** and (*E*)-β-farnesene **17**) was carefully added on top, to create a two-layer or biphasic system. The vial was gently mixed, allowed to settle for 1 h and finally, the samples were incubated (RT, 2h) before a GC sample (2 µL) was taken and analysed.

8.7.3 Preliminary Biphasic Assay with OBP6

For the final biphasic assay with OBP6, a solution of OBP6 (100µL of 5µM in 25 mM Tris) was added to a small vial (2 mL size). A ligand solution (80 µL of 12 µM solution in hexane) was carefully added on top, to create a two-layer or biphasic system. The vial was gently mixed before being centrifuged (5000 rpm, 15 minutes). Finally, samples were incubated (RT, 2h) and a GC sample (2 µL) taken and analysed *via* the same GC method as in 8.7.2.

8.7.4 Analysis of Gas Chromatography Data

Quantification of the amount of ligand per sample was undertaken by generating a calibration curve for each ligand in hexane in the GC at various concentrations. The amount of ligand present was reported as a mg quantity and µM quantity.

Statistical analysis was performed in R 3.4.4.[208]

8.8 OTHER STUDIES

8.9.1 X-Ray Crystallography

An initial sample of OBP6 was prepared as in section 8.2. A large-scale expression and purification of *E. coli* containing. The His_6-tag was cleaved, and the protein purified further with fast-protein liquid chromatography (FPLC). Sample quantities are documented in Table 8.17.

Table 8.17: Final amounts of protein for crystallography

Experimental Step	Sample Quantity / mg
Purification from *E. coli*	30 mg
Cleavage of His_6-tag	27 mg
Fast protein liquid chromatography	15 mg

Final buffer exchanged and concentration	13.64 mg

Two final protein concentrations were generated for crystallography: 12 $mgml^{-1}$ (15 mM Tris, 250 mM NaCl) and 6 $mgml^{-1}$ (32.5 mM Tris, 175 mM NaCl, 0.5% glycerol). The samples were then centrifuged (13000 g; 10 min; 7°C) to remove any dust, debris or precipitate.

A high-throughput screening was performed with 96-well plates and the 'mosquito' robot. Four sets of conditions were used, PACT *Premier*™, SG1™ Screen, JCSG-*Plus*™ and Morpheus®, resulting in a total of 384 conditions.

References

(1) Taylor, L. Migration and the Spatial Dynamics of an Aphid, *Myzus Persicae*. *J. Anim. Ecol.* **1977**, *46* (2), 411–423.

(2) S. Richards et al. Genome Sequence of the Pea Aphid *Acyrthosiphon Pisum*. *PLoS Biol.* **2010**, *8* (2). https://doi.org/10.1371/journal.pbio.1000313.

(3) Brisson, J. A.; Stern, D. L. The Pea Aphid, *Acyrthosiphon Pisum*: An Emerging Genomic Model System for Ecological, Developmental and Evolutionary Studies. *BioEssays* **2006**, *28*, 747–755. https://doi.org/10.1002/bies.20436.

(4) Ishikawa, A.; Ogawa, K.; Gotoh, H.; Walsh, T. K.; Tagu, D.; Brisson, J. A.; Rispe, C.; Jaubert-Possamai, S.; Kanbe, T.; Tsubota, T.; et al. Juvenile Hormone Titre and Related Gene Expression during the Change of Reproductive Modes in the Pea Aphid. *Insect Mol. Biol.* **2012**, *21* (1), 49–60. https://doi.org/10.1111/j.1365-2583.2011.01111.x.

(5) Cannon, C.; Bunn, B.; Petrizzo, E.; Alston, D. G.; Murray, M. Aphid Pests on Vegetables. *Utah Plant Pest Diagnostic Laboratory*. 2017, pp 1–9.

(6) Birkett, M. A.; Pickett, J. A. Aphid Sex Pheromones: From Discovery to Commercial Production. *Phytochemistry* **2003**, *62* (5), 651–656. https://doi.org/10.1016/S0031-9422(02)00568-X.

(7) Dixon, A. F. G. The Life-cycle and Host Preferences of the Bird Cherry-oat Aphid, *Rhopalosiphum Padi* L., and Their Bearing on the Theories of Host Alternation in Aphids. *Ann. Appl. Biol.* **1971**, *68* (2), 135–147. https://doi.org/10.1111/j.1744-7348.1971.tb06450.x.

(8) McLean, D. L.; Kinsey, M. G. Probing Behavior of the Pea Aphid, *Acyrthosiphon Pisum*. II. Comparisons of Salivation and Ingestion in Host and Non-Host Plant Leaves1. *Ann. Entomol. Soc. Am.* **1968**, *61* (3), 730–739. https://doi.org/10.1093/aesa/61.3.730.

(9) Harrington, R.; Clark, S. J.; Welham, S. J.; Verrier, P. J.; Denholm, C. H.; Hullé, M.; Maurice, D.; Rounsevell, M. D.; Cocu, N. Environmental Change and the Phenology of European Aphids. *Glob. Chang. Biol.* **2007**, *13* (8), 1550–1564. https://doi.org/10.1111/j.1365-2486.2007.01394.x.

(10) Sack, C.; Stern, D. L. Sex and Death in the Male Pea Aphid, *Acyrthosiphon Pisum*: The Life-History Effects of a Wing Dimorphism. *J. Insect Sci.* **2007**, *7* (45), 1–9. https://doi.org/10.1673/031.007.4501.

(11) Agriculture and Horticulture Development Board. Encyclopaedia of Pests and Natural Enemies in Field Crops. 2015, pp 43–76.

(12) Dixon, A. F. G.; Kindleman, P.; Leps, J.; Holman, J. Why There Are So Few Species of Aphids , Especially in the Tropics. *Am. Nat.* **1987**, *129* (4), 580–592.

(13) Dedryver, C. A.; Le Ralec, A.; Fabre, F. The Conflicting Relationships between Aphids and Men: A Review of Aphid Damage and Control Strategies. *Comptes Rendus - Biol.* **2010**, *333* (6–7), 539–553. https://doi.org/10.1016/j.crvi.2010.03.009.

(14) Douglas, A. E. Phloem-Sap Feeding by Animals: Problems and Solutions. *J. Exp. Bot.* **2006**, *57* (4), 747–754. https://doi.org/10.1093/jxb/erj067.

(15) Ali, M.; Anwar, S.; Shuja, M. N.; Tripathi, R. K.; Singh, J. The Genus Luteovirus from Infection to Disease. *Eur. J. Plant Pathol.* **2018**, 1–20. https://doi.org/10.1007/s10658-

018-1425-8.

(16) Fokkema, N. J.; Riphagen, I.; Poot, R. J.; de Jong, C. Aphid Honeydew, a Potential Stimulant of Cochliobolus Sativus and Septoria Nodorum and the Competitive Role of Saprophytic Mycoflora. *Trans. Br. Mycol. Soc.* **1983**, *81* (2), 355–363. https://doi.org/10.1016/S0007-1536(83)80087-4.

(17) Pimentel, D.; McLaughlin, L.; Zepp, A.; Lakitan, B.; Kraus, T.; Kleinman, P.; Vancini, F.; Roach, W. J.; Graap, E.; Keeton, W. S.; et al. Environmental and Economic Effects of Reducing Pesticide Use. *Bioscience* **1991**, *41* (6), 402–409. https://doi.org/10.2307/1311747.

(18) Neubauer, I.; Raccah, B.; Aharonson, N.; Swirski, E.; Ishaaya, I. Systemic Effect of Aldicarb, Dimethoate and Ethiofencarb on Mortality and Population Dynamics of the Spirea Aphid, *Aphis Citricola* Van Der Goot. *Crop Prot.* **1983**, *2* (2), 211–218. https://doi.org/10.1016/0261-2194(83)90046-7.

(19) Musau, D. M.; Parry, W. H. Comparison of the Potential of Organophosphorous Insecticides and Soaps in Conifer Aphid Control. *Crop Prot.* **1988**, *7* (4), 267–272. https://doi.org/10.1016/0261-2194(88)90048-8.

(20) Bass, C.; Puinean, A. M.; Zimmer, C. T.; Denholm, I.; Field, L. M.; Foster, S. P.; Gutbrod, O.; Nauen, R.; Slater, R.; Williamson, M. S. The Evolution of Insecticide Resistance in the Peach Potato Aphid, *Myzus Persicae*. *Insect Biochem. Mol. Biol.* **2014**, *51* (1), 41–51. https://doi.org/10.1016/j.ibmb.2014.05.003.

(21) Puinean, A. M.; Foster, S. P.; Oliphant, L.; Denholm, I.; Field, L. M.; Millar, N. S.; Williamson, M. S.; Bass, C. Amplification of a Cytochrome P450 Gene Is Associated with Resistance to Neonicotinoid Insecticides in the Aphid *Myzus Persicae*. *PLoS Genet.* **2010**, *6* (6), 1–11. https://doi.org/10.1371/journal.pgen.1000999.

(22) Devonshire, A. L.; Moores, G. D. A Carboxylesterase with Broad Substrate Specificity Causes Organophosphorus, Carbamate and Pyrethroid Resistance in Peach-Potato Aphids (Myzus Persicae). *Pestic. Biochem. Physiol.* **1982**, *18* (2), 235–246. https://doi.org/10.1016/0048-3575(82)90110-9.

(23) Porretta, D.; Gargani, M.; Bellini, R.; Medici, A.; Punelli, F.; Urbanelli, S. Defence Mechanisms against Insecticides Temephos and Diflubenzuron in the Mosquito Aedes Caspius: The P-Glycoprotein Efflux Pumps. *Med. Vet. Entomol.* **2008**, *22* (1), 48–54. https://doi.org/10.1111/j.1365-2915.2008.00712.x.

(24) Scott, J. G. Investigating Mechanisms of Insecticide Resistance: Methods, Strategies, and Pitfalls. In *Pesticide Resistance in Arthropods*; 1990; pp 39–57.

(25) Snyder, W. E.; Ives, A. R. Interactions between Specialist and Generalist Natural Enemies: Parasitoids, Predators, and Pea Aphid Biocontrol. *Ecology* **2003**, *84* (1), 91–107. https://doi.org/10.1890/0012-9658(2003)084[0091:IBSAGN]2.0.CO;2.

(26) Bianchi, F. J. J. A.; Booij, C. J. H.; Tscharntke, T. Sustainable Pest Regulation in Agricultural Landscapes: A Review on Landscape Composition, Biodiversity and Natural Pest Control. *Proc. R. Soc. B Biol. Sci.* **2006**, *273* (1595), 1715–1727. https://doi.org/10.1098/rspb.2006.3530.

(27) Nakashima, Y.; Ida, T. Y.; Powell, W.; Pickett, J. A.; Birkett, M. A.; Taki, H.; Takabayashi, J. Field Evaluation of Synthetic Aphid Sex Pheromone in Enhancing Suppression of Aphid Abundance by Their Natural Enemies. *BioControl* **2016**, *61* (5), 485–496.

https://doi.org/10.1007/s10526-016-9734-3.

(28) Powell, W.; Pickett, J. A. Manipulation of Parasitoids for Aphid Pest Management: Progress and Prospects. *Pest Manag. Sci.* **2003**, *59* (2), 149–155. https://doi.org/10.1002/ps.550.

(29) B C Alavo, T. The Insect Pathogenic Fungus *Verticillium Lecanii* (Zimm.) Viegas and Its Use for Pests Control: A Review. *J. Exp. Biol. Agric. Sci.* **2015**, *3* (4), 337–345. https://doi.org/10.18006/2015.3(4).337.345.

(30) Pickett, J. A.; Aradottir, G. I.; Birkett, M. A.; Bruce, T. J. A.; Chamberlain, K.; Khan, Z. R.; Midega, C. A. O.; Smart, L. E.; Woodcock, C. M. Aspects of Insect Chemical Ecology: Exploitation of Reception and Detection as Tools for Deception of Pests and Beneficial Insects. *Physiol. Entomol.* **2012**, *37* (1), 2–9. https://doi.org/10.1111/j.1365-3032.2011.00828.x.

(31) Howard, R. W.; Blomquist, G. J. Ecological, Behavoiral, and Biochemical Aspects of Insect Hydrocarbons. *Annu. Rev. Entomol.* **2005**, *50*, 371–393. https://doi.org/10.1146/annurev.ento.50.071803.130359.

(32) Nordlund, D. A.; Lewis, W. J. Terminology of Chemical Releasing Stimuli in Intraspecific and Interspecific Interactions. *J. Chem. Ecol.* **1976**, *2* (2), 211–220. https://doi.org/10.1007/BF00987744.

(33) Burger, B. V. Mammalian Semiochemicals. In *The Chemistry of Pheromones and Other Semiochemicals II*; Springer, Berlin, Heidelberg, 2004; pp 231–278.

(34) Campagna, S.; Mardon, J.; Celerier, A.; Bonadonna, F. Potential Semiochemical Molecules from Birds: A Practical and Comprehensive Compilation of the Last 20 Years Studies. *Chem. Senses* **2012**, *37* (1), 3–25. https://doi.org/10.1093/chemse/bjr067.

(35) Magurran, A. E.; Irving, P. W.; Henderson, P. A. Is There a Fish Alarm Pheromone? A Wild Study and Critique. *Proc. R. Soc. B Biol. Sci.* **1996**, *263* (1376), 1551–1556. https://doi.org/10.1098/rspb.1996.0227.

(36) Pickett, J. A.; Allemann, R. K.; Birkett, M. A. The Semiochemistry of Aphids. *Nat. Prod. Rep.* **2013**, *30* (10), 1277. https://doi.org/10.1039/c3np70036d.

(37) Dawson, G.; Griffiths, D.; Janes, N. Identification of an Aphid Sex Pheromone. *Nature*. 1987, pp 614–616. https://doi.org/10.1038/325614a0.

(38) Campbell, C. A. M.; Dawson, G. W.; Griffiths, D. C.; Pettersson, J.; Pickett, J. A.; Wadhams, L. J.; Woodcock, C. M. Sex Attractant Pheromone of Damson-Hop Aphid *Phorodon Humuli* (Homoptera, Aphididae). *J. Chem. Ecol.* **1990**, *16* (12), 3455–3465. https://doi.org/10.1007/BF00982110.

(39) Dawson, G. W.; Griffiths, D. C.; Merritt, L. A.; Mudd, A.; Pickett, J. A.; Woodcock, C. M. Aphid Semiochemicals - A Review and Recent Advances on the Sex Pheromone. *J. Chem. Ecol.* **1990**, *16* (11), 3019–3030.

(40) Jeon, H.; Han, K. S.; Boo, K. S. Sex Pheromone of *Aphis Spiraecola* (Homoptera: Aphididae): Composition and Circadian Rhythm in Release. *J. Asia. Pac. Entomol.* **2003**, *6* (2), 159–165. https://doi.org/10.1016/S1226-8615(08)60181-8.

(41) Symmes, E. J.; Dewhirst, S. Y.; Birkett, M. A.; Campbell, C. A. M.; Chamberlain, K.; Pickett, J. A.; Zalom, F. G. The Sex Pheromones of Mealy Plum (*Hyalopterus Pruni*) and Leaf-Curl Plum (*Brachycaudus Helichrysi*) Aphids: Identification and Field Trapping of

Male and Gynoparous Aphids in Prune Orchards. *J. Chem. Ecol.* **2012**, *38* (5), 576–583. https://doi.org/10.1007/s10886-012-0121-y.

(42) Gabryś, B. J.; Gadomski, H. J.; Klukowski, Z.; Pickett, J. A.; Sobota, G. T.; Wadhams, L. J.; Woodcock, C. M. Sex Pheromone of Cabbage Aphid *Brevicoryne Brassicae*: Identification and Field Trapping of Male Aphids and Parasitoids. *J. Chem. Ecol.* **1997**, *23* (7), 1881–1890. https://doi.org/10.1023/B:JOEC.0000006457.28372.48.

(43) Stewart-Jones, A.; Dewhirst, S. Y.; Durrant, L.; Fitzgerald, J. D.; Hardie, J.; Hooper, A. M.; Pickett, J. A.; Poppy, G. M. Structure, Ratios and Patterns of Release in the Sex Pheromone of an Aphid, *Dysaphis Plantaginea*. *J. Exp. Biol.* **2007**, *210* (24), 4335–4344. https://doi.org/10.1242/jeb.009944.

(44) Goldansaz, S. H.; Dewhirst, S.; Birkett, M. A.; Hooper, A. M.; Smiley, D. W. M.; Pickett, J. A.; Wadhams, L.; McNeil, J. N. Identification of Two Sex Pheromone Components of the Potato Aphid, *Macrosiphum Euphorbiae* (Thomas). *J. Chem. Ecol.* **2004**, *30* (4), 819–834. https://doi.org/10.1023/B:JOEC.0000028434.19319.b4.

(45) Hardie, J.; Storer, J.; Nottingham, S. F.; Peace, L.; Harrington, R.; Merritt, L. A.; Wadhams, L. J.; Wood, D. The Interaction of Sex Pheromone and Plant Volatiles for Field Attraction of Male Bird-Cherry Aphid, Rhopalosiphum Padi. In *Brighton Crop Protection Conference, Pests and Diseases vol 3*; 1994; pp 1223–1230.

(46) Lilley, R.; Hardie, J. Cereal Aphid Responses to Sex Pheromones and Host-Plant Odours in the Laboratory. *Physiol. Entomol.* **1996**, *21* (4), 304–308. https://doi.org/10.1111/j.1365-3032.1996.tb00869.x.

(47) Hardie, J.; Nottingham, S. F.; Dawson, G. W.; Harrington, R.; Pickett, J. A.; Wadhams, L. J. Attraction of Field-Flying Aphid Males to Synthetic Sex Pheromone. *Chemoecology* **1992**, *3* (3–4), 113–117. https://doi.org/10.1007/BF01370138.

(48) Boo, K. S.; Choi, M. Y.; Chung, I. B.; Eastop, V. F.; Pickett, J. A.; Wadhams, L. J.; Woodcock, C. M. Sex Pheromone of the Peach Aphid, *Tuberocephalus Momonis*, and Optimal Blends for Trapping Males and Females in the Field. *J. Chem. Ecol.* **2000**, *26* (3), 601–609. https://doi.org/10.1023/A:1005415919226.

(49) Marsh, D. Sex Pheromone in the Aphid *Megoura Viciae*. *Nature* **1972**, *238*, 31–32. https://doi.org/10.1038/newbio238031a0.

(50) Pettersson, J. An Aphid Sex Attractant. *Biol. Stud. Entomol. Scand.* **1970**, *1*, 63–73.

(51) Dawson, G. W.; Pickett, J. a.; Smiley, D. W. M. The Aphid Sex Pheromone Cyclopentanoids: Synthesis in the Elucidation of Structure and Biosynthetic Pathways. *Bioorganic Med. Chem.* **1996**, *4* (3), 351–361. https://doi.org/10.1016/0968-0896(96)00012-0.

(52) Schreiber, S. L.; Meyers, H. V.; Wibery, K. B. Stereochemistry of the Intramolecular Enamine/Enal (Enone) Cycloaddition Reaction and Subsequent Transformations. *J. Am. Chem. Soc.* **1986**, *108*, 8274–8277.

(53) Pickett, J. A.; Griffiths, D. C. Composition of Aphid Alarm Pheromones. *J. Chem. Ecol.* **1980**, *6* (2), 349–360. https://doi.org/10.1007/BF01402913.

(54) Beale, M. H.; Birkett, M. A.; Bruce, T. J. A.; Chamberlain, K.; Field, L. M.; Huttly, A. K.; Martin, J. L.; Parker, R.; Phillips, A. L.; Pickett, J. a; et al. Aphid Alarm Pheromone Produced by Transgenic Plants Affects Aphid and Parasitoid Behavior. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103* (27), 10509–10513.

https://doi.org/10.1073/pnas.0603998103.

(55) Francis, F.; Vandermoten, S.; Verheggen, F.; Lognay, G.; Haubruge, E. Is the (E)-β-Farnesene Only Volatile Terpenoid in Aphids? *Jen* **2005**, *129* (1), 6–11. https://doi.org/10.1111/j.1439-0418.2005.00925.6.

(56) Pickett, J. The Chemical Ecology Of Aphids. *Annu. Rev. Entomol.* **1992**, *37* (1), 67–90. https://doi.org/10.1146/annurev.ento.37.1.67.

(57) Vandermoten, S.; Mescher, M. C.; Francis, F.; Haubruge, E.; Verheggen, F. J. Aphid Alarm Pheromone: An Overview of Current Knowledge on Biosynthesis and Functions. *Insect Biochem. Mol. Biol.* **2012**, *42* (3), 155–163. https://doi.org/10.1016/j.ibmb.2011.11.008.

(58) Kang, S.-K.; Chung, G.-Y.; Lee, D.-H. A Convenient Synthesis of (E)-β-Farnesene. *Bull. Korean Chem. Soceity* **1987**, *8* (4), 351–353.

(59) Bruce, T. J. A.; Pickett, J. A. Perception of Plant Volatile Blends by Herbivorous Insects - Finding the Right Mix. *Phytochemistry* **2011**, *72* (13), 1605–1611. https://doi.org/10.1016/j.phytochem.2011.04.011.

(60) Webster, B.; Bruce, T.; Dufour, S.; Birkemeyer, C.; Birkett, M.; Hardie, J.; Pickett, J. Identification of Volatile Compounds Used in Host Location by the Black Bean Aphid, *Aphis Fabae*. *J. Chem. Ecol.* **2008**, *34* (9), 1153–1161. https://doi.org/10.1007/s10886-008-9510-7.

(61) Bruce, T. J.; Pickett, J. A. Plant Defence Signalling Induced by Biotic Attacks. *Curr. Opin. Plant Biol.* **2007**, *10* (4), 387–392. https://doi.org/10.1016/j.pbi.2007.05.002.

(62) Du, Y. J.; Poppy, G. M.; Powell, W.; Pickett, J. a; Wadhams, L. J.; Woodcock, C. M. Identification of Semiochemicals Released during Aphid Feeding That Attract Parasitoid *Aphidius Ervi*. *J. Chem. Ecol.* **1998**, *24* (8), 1355–1368. https://doi.org/10.1023/A:1021278816970.

(63) Hardie, J.; Peace, L.; Pickett, J. A.; Smiley, D. W. M.; Storer, J. R.; Wadhams, L. J. Sex Pheromone Stereochemistry and Purity Affect Field Catches of Male Aphids. *J. Chem. Ecol.* **1997**, *23* (11), 2547–2554. https://doi.org/10.1023/B:JOEC.0000006665.05409.97.

(64) Boo, K. S.; Kang, S. S.; Park, J. H.; Pickett, J. A.; Wadhams, L. J. Field Trapping of *Chrysopa Cognata* (Neuroptera: Chrysopidae) with Aphid Sex Pheromone Components in Korea. *J. Asia. Pac. Entomol.* **2003**, *6* (1), 29–36. https://doi.org/10.1016/S1226-8615(08)60164-8.

(65) Koczor, S.; Szentkirayi, F.; Birkett, M. A.; Pickett, J. A.; Voigt, E.; Toth, M. Attraction of *Chrysoperla Carnea* Complex and *Chrysopa Spp.* Lacewings (Neuroptera: Chrysopidae) to Aphid Sex Pheromone Components and a Synthetic Blend of Floral Compounds in Hungary. *Pest Manag. Sci.* **2010**, *66* (12), 1374–1379. https://doi.org/10.1002/ps.2030.

(66) Bruce, T. J.; Aradottir, G. I.; Smart, L. E.; Martin, J. L.; Caulfield, J. C.; Doherty, A.; Sparks, C. A.; Woodcock, C. M.; Birkett, M. A.; Napier, J. A.; et al. The First Crop Plant Genetically Engineered to Release an Insect Pheromone for Defence. *Sci. Rep.* **2015**, *5*, 11183. https://doi.org/10.1038/srep11183.

(67) Bruce, T. J. A.; Birkett, M. A.; Blande, J.; Hooper, A. M.; Martin, J. L.; Khambay, B.; Prosser, I.; Smart, L. E.; Wadhams, L. J. Response of Economically Important Aphids to

Components of *Hemizygia Petiolata* Essential Oil. *Pest Manag. Sci.* **2005**, *61* (11), 1115–1121. https://doi.org/10.1002/ps.1102.

(68) Touchet, S.; Chamberlain, K.; Woodcock, C. M.; Miller, D. J.; Birkett, M. A.; Pickett, J. A.; Allemann, R. K. Novel Olfactory Ligands via Terpene Synthases. *Chem. Commun.* **2015**, *51* (35), 7550–7553. https://doi.org/10.1039/C5CC01814E.

(69) Leal, W. S. Odorant Reception in Insects : Roles of Receptors , Binding Proteins , and Degrading Enzymes. *Annu. Rev. Entomol. is* **2013**, *58*, 373–391. https://doi.org/10.1146/annurev-ento-120811-153635.

(70) Schneider, D. 100 Years of Pheromone Research - An Essay on Lepidoptera. *Naturwissenschaften* **1992**, *79* (6), 241–250. https://doi.org/10.1007/BF01175388.

(71) Sandler, B. H.; Nikonova, L.; Leal, W. S.; Clardy, J. Sexual Attraction in the Silkworm Moth: Structure of the Pheromone-Binding-Protein-Bombykol Complex. *Chem. Biol.* **2000**, *7* (2), 143–151. https://doi.org/10.1016/S1074-5521(00)00078-8.

(72) Hansson, B. S. Olfaction in Lepidoptera. *Experientia* **1995**, *51* (11), 1003–1027. https://doi.org/10.1007/BF01946910.

(73) Butenandt, A. von; Beckmann, R.; Stamm, D.; Hecker, E. Über Den Sexual-Lockstoff Des Seidenspinners *Bombyx Mori*. Reindarstellung Und Konstitution. *Zeitschrift für Naturforschung. B, J. Chem. Sci.* **1959**, *14b*, 283–284.

(74) Adams, M.; Celniker, S.; Holt, R.; Evans, C.; Gocyane, J.; Amanatides, P.; Scherer, S.; Li, P.; Hoskins, R. The Genome Sequence of *Drosophila Melanogaster*. *Science (80-.).* **2000**, *287* (March), 2185–2195. https://doi.org/10.1126/science.287.5461.2185.

(75) De Biasio, F.; Riviello, L.; Bruno, D.; Grimaldi, A.; Congiu, T.; Sun, Y. F.; Falabella, P. Expression Pattern Analysis of Odorant-Binding Proteins in the Pea Aphid *Acyrthosiphon Pisum. Insect Sci.* **2014**, *00*, 1–15. https://doi.org/10.1111/1744-7917.12118.

(76) Larter, N. K.; Sun, J. S.; Carlson, J. R. Organization and Function of *Drosophila* Odorant Binding Proteins. *Elife* **2016**, *5*, e20242. https://doi.org/10.7554/eLife.20242.

(77) Gomez-Diaz, C.; Reina, J. H.; Cambillau, C.; Benton, R. Ligands for Pheromone-Sensing Neurons Are Not Conformationally Activated Odorant Binding Proteins. *PLoS Biol.* **2013**, *11* (4). https://doi.org/10.1371/journal.pbio.1001546.

(78) Clyne, P. J.; Warr, C. G.; Freeman, M. R.; Lessing, D.; Kim, J.; Carlson, J. R. A Novel Family of Divergent Seven-Transmembrane Proteins: Candidate Odorant Receptors in Drosophila. *Neuron* **1999**, *22* (2), 327–338. https://doi.org/10.1016/S0896-6273(00)81093-4.

(79) Dobritsa, A. A.; Van Der Goes Van Naters, W.; Warr, C. G.; Steinbrecht, R. A.; Carlson, J. R. Integrating the Molecular and Cellular Basis of Odor Coding in the Drosophila Antenna. *Neuron* **2003**, *37* (5), 827–841. https://doi.org/10.1016/S0896-6273(03)00094-1.

(80) Buck, L.; Axel, R. A Novel Multigene Family May Encode Odorant Receptors: A Molecular Basis for Odor Recognition. *Cell* **1991**, *65* (1), 175–187. https://doi.org/10.1016/0092-8674(91)90418-X.

(81) Kristoffersen, L.; Hansson, B. S.; Anderbrant, O.; Larsson, M. C. Aglomerular Hemipteran Antennal Lobes - Basic Neuroanatomy of a Small Nose. *Chem. Senses*

2008, *33* (9), 771–778. https://doi.org/10.1093/chemse/bjn044.

(82) Benton, R. On the ORigin of Smell: Odorant Receptors in Insects. *Cell. Mol. Life Sci.* **2006**, *63* (14), 1579–1585. https://doi.org/10.1007/s00018-006-6130-7.

(83) Kollmann, M.; Minoli, S.; Bonhomme, J.; Homberg, U.; Schachtner, J.; Tagu, D.; Anton, S. Revisiting the Anatomy of the Central Nervous System of a Hemimetabolous Model Insect Species: The Pea Aphid Acyrthosiphon Pisum. *Cell Tissue Res.* **2011**, *343* (2), 343–355. https://doi.org/10.1007/s00441-010-1099-9.

(84) Pflugfelder, O. Vergleichend-Anatomische Experimentelle Und Embryologische Untersuchungen Über Das Nervensystem Und Die Sinnesorgane Der Rhynchoten. *Zoologica* **1937**, *34*, 1–102.

(85) Strausfeld, N. J.; Sinakevitch, I.; Brown, S. M.; Farris, S. M. Ground Plan of the Insect Mushroom Body: Functional and Evolutionary Implications. *J. Comp. Neurol.* **2009**, *513* (3), 265–291. https://doi.org/10.1002/cne.21948.

(86) Seki, Y.; Dweck, H. K. M.; Rybak, J.; Wicher, D.; Sachse, S.; Hansson, B. S. Olfactory Coding from the Periphery to Higher Brain Centers in the Drosophila Brain. *BMC Biol.* **2017**, *15* (1), 18–22. https://doi.org/10.1186/s12915-017-0389-z.

(87) Rosenbaum, D. M.; Rasmussen, S. G.; Kobilka, B. K. The Structure and Function of G-Protein-Coupled Receptors. *Nature* **2009**, *459* (7245), 356–363. https://doi.org/doi:10.1038/nature08144.

(88) Butterwick, J. A.; del Mármol, J.; Kim, K. H.; Kahlson, M. A.; Rogow, J. A.; Walz, T.; Ruta, V. Cryo-EM Structure of the Insect Olfactory Receptor Orco. *Nature* **2018**, *560* (7719), 447–452. https://doi.org/10.1038/s41586-018-0420-8.

(89) Sato, K.; Pellegrino, M.; Nakagawa, T.; Nakagawa, T.; Vosshall, L. B.; Touhara, K. Insect Olfactory Receptors Are Heteromeric Ligand-Gated Ion Channels. *Nature* **2008**, *452* (7190), 1002–1006. https://doi.org/10.1038/nature06850.

(90) Wicher, D.; Schafer, R.; Bauernfeind, R.; Stensmyr, M. C.; Heller, R.; Heinemann, S. H.; Hansson, B. S. Drosophila Odorant Receptors Are Both Ligand-Gated and Cyclic-Nucleotide-Activated Cation Channels. *Nature* **2008**, *452* (7190), 1007-U10. https://doi.org/10.1038/nature06861.

(91) Schrödinger, L. The PyMOL Molecular Graphics System. 2015.

(92) Li, Y.; Zhang, J.; Chen, D.; Yang, P.; Jiang, F.; Wang, X.; Kang, L. CRISPR/Cas9 in Locusts: Successful Establishment of an Olfactory Deficiency Line by Targeting the Mutagenesis of an Odorant Receptor Co-Receptor (Orco). *Insect Biochem. Mol. Biol.* **2016**, *79*, 27–35. https://doi.org/10.1016/j.ibmb.2016.10.003.

(93) Fandino, R. A.; Haverkamp, A.; Bisch-Knaden, S.; Zhang, J.; Bucks, S.; Nguyen, T. A. T.; Schröder, K.; Werckenthin, A.; Rybak, J.; Stengl, M.; et al. Mutagenesis of Odorant Coreceptor Orco Fully Disrupts Foraging but Not Oviposition Behaviors in the Hawkmoth *Manduca Sexta* . *Proc. Natl. Acad. Sci.* **2019**, *116* (31), 15677–15685. https://doi.org/10.1073/pnas.1902089116.

(94) Trible, W.; Olivos-Cisneros, L.; McKenzie, S. K.; Saragosti, J.; Chang, N. C.; Matthews, B. J.; Oxley, P. R.; Kronauer, D. J. C. Orco Mutagenesis Causes Loss of Antennal Lobe Glomeruli and Impaired Social Behavior in Ants. *Cell* **2017**, *170* (4), 727-735.e10. https://doi.org/10.1016/j.cell.2017.07.001.

(95) Malnic, B.; Hirono, J.; Sato, T.; Buck, L. B. Combinatorial Receptor Codes for Odors. *Cell* **1999**, *96* (5), 713–723. https://doi.org/10.1016/S0092-8674(00)80581-4.

(96) Behrens, M.; Briand, L.; de March, C. A.; Matsunami, H.; Yamashita, A.; Meyerhof, W.; Weyand, S. Structure-Function Relationships of Olfactory and Taste Receptors. *Chem. Senses* **2018**, *43* (2), 81–87. https://doi.org/10.1093/chemse/bjx083.

(97) Smadja, C.; Shi, P.; Butlin, R. K.; Robertson, H. M. Large Gene Family Expansions and Adaptive Evolution for Odorant and Gustatory Receptors in the Pea Aphid, *Acyrthosiphon Pisum*. *Mol. Biol. Evol.* **2009**, *26* (9), 2073–2086. https://doi.org/10.1093/molbev/msp116.

(98) Cao, D.; Liu, Y.; Walker, W. B.; Li, J.; Wang, G. Molecular Characterization of the Aphis Gossypii Olfactory Receptor Gene Families. *PLoS One* **2014**, *9* (6). https://doi.org/10.1371/journal.pone.0101187.

(99) Zhang, R.; Wang, B.; Grossi, G.; Falabella, P.; Liu, Y.; Yan, S.; Lu, J.; Xi, J.; Wang, G. Molecular Basis of Alarm Pheromone Detection in Aphids. *Curr. Biol.* **2017**, *27* (1), 55–61. https://doi.org/10.1016/j.cub.2016.10.013.

(100) Zhang, R. Bin; Liu, Y.; Yan, S. C.; Wang, G. R. Identification and Functional Characterization of an Odorant Receptor in Pea Aphid, *Acyrthosiphon Pisum*. *Insect Sci.* **2019**, *26* (1), 58–67. https://doi.org/10.1111/1744-7917.12510.

(101) Pelosi, P.; Maida, R. Odorant-Binding Proteins in Insects. *Comp. Biochem. Physiol.* **1995**, *111* (B), 503–514. https://doi.org/10.1016/S0083-6729(10)83010-9.

(102) Qiao, H.; Tuccori, E.; He, X.; Gazzano, A.; Field, L.; Zhou, J. J.; Pelosi, P. Discrimination of Alarm Pheromone (E)-β-Farnesene by Aphid Odorant-Binding Proteins. *Insect Biochem. Mol. Biol.* **2009**, *39* (5–6), 414–419. https://doi.org/10.1016/j.ibmb.2009.03.004.

(103) Zhou, J. J.; Vieira, F. G.; He, X. L.; Smadja, C.; Liu, R.; Rozas, J.; Field, L. M. Genome Annotation and Comparative Analyses of the Odorant-Binding Proteins and Chemosensory Proteins in the Pea Aphid Acyrthosiphon Pisum. *Insect Mol. Biol.* **2010**, *19*, 113–122. https://doi.org/10.1111/j.1365-2583.2009.00919.x.

(104) Northey, T.; Venthur, H.; De Biasio, F.; Chauviac, F. X.; Cole, A.; Ribeiro, K. A. L.; Grossi, G.; Falabella, P.; Field, L. M.; Keep, N. H.; et al. Crystal Structures and Binding Dynamics of Odorant-Binding Protein 3 from Two Aphid Species *Megoura Viciae* and *Nasonovia Ribisnigri*. *Sci. Rep.* **2016**, *6* (April), 1–13. https://doi.org/10.1038/srep24739.

(105) Leal, W. S.; Chen, A. M.; Ishida, Y.; Chiang, V. P.; Erickson, M. L.; Morgan, T. I.; Tsuruda, J. M. Kinetics and Molecular Properties of Pheromone Binding and Release. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (15), 5386–5391. https://doi.org/10.1073/pnas.0501447102.

(106) Kowcun, A.; Honson, N.; Plettner, E. Olfaction in the Gypsy Moth, *Lymantria Dispar*: Effect of PH, Ionic Strength, and Reductants on Pheromone Transport by Pheromone-Binding Proteins. *J. Biol. Chem.* **2001**, *276* (48), 44770–44776. https://doi.org/10.1074/jbc.M104688200.

(107) Leite, N. R.; Krogh, R.; Xu, W.; Ishida, Y.; Iulek, J.; Leal, W. S.; Oliva, G. Structure of an Odorant-Binding Protein from the Mosquito *Aedes Aegypti* Suggests a Binding Pocket Covered by a PH-Sensitive "Lid." *PLoS One* **2009**, *4* (11), 1–7. https://doi.org/10.1371/journal.pone.0008006.

(108) Zubkov, S.; Gronenborn, A. M.; Byeon, I. J. L.; Mohanty, S. Structural Consequences of the PH-Induced Conformational Switch in A. Polyphemus Pheromone-Binding Protein: Mechanisms of Ligand Release. *J. Mol. Biol.* **2005**, *354* (5), 1081–1090. https://doi.org/10.1016/j.jmb.2005.10.015.

(109) Wogulis, M.; Morgan, T.; Ishida, Y.; Leal, W. S.; Wilson, D. K. The Crystal Structure of an Odorant Binding Protein from *Anopheles Gambiae*: Evidence for a Common Ligand Release Mechanism. *Biochem. Biophys. Res. Commun.* **2006**, *339* (1), 157–164. https://doi.org/10.1016/j.bbrc.2005.10.191.

(110) Han, L.; Zhang, Y. J.; Zhang, L.; Cui, X.; Yu, J.; Zhang, Z.; Liu, M. S. Operating Mechanism and Molecular Dynamics of Pheromone-Binding Protein ASP1 as Influenced by PH. *PLoS One* **2014**, *9* (10), 1–9. https://doi.org/10.1371/journal.pone.0110565.

(111) Mosbah, A.; Campanacci, V.; Lartigue, A.; Tegoni, M.; Cambillau, C.; Darbon, H. Solution Structure of a Chemosensory Protein from the Moth *Mamestra Brassicae*. *Biochem. J.* **2003**, *369*, 39–44. https://doi.org/10.1007/s13355-010-0007-9.

(112) Pelosi, P.; Zhou, J. J.; Ban, L. P.; Calvello, M. Soluble Proteins in Insect Chemical Communication. *Cell. Mol. Life Sci.* **2006**, *63* (14), 1658–1676. https://doi.org/10.1007/s00018-005-5607-0.

(113) Pelosi, P.; Iovinella, I.; Zhu, J.; Wang, G.; Francesca, R. Beyond Chemoreception : Diverse Tasks of Soluble Olfactory Proteins in Insects. **2018**, *93*, 184–200. https://doi.org/10.1111/brv.12339.

(114) Siciliano, P.; He, X. L.; Woodcock, C.; Pickett, J. A.; Field, L. M.; Birkett, M. A.; Kalinova, B.; Gomulski, L. M.; Scolari, F.; Gasperi, G.; et al. Identification of Pheromone Components and Their Binding Affinity to the Odorant Binding Protein CcapOBP83a-2 of the Mediterranean Fruit Fly, *Ceratitis Capitata*. *Insect Biochem. Mol. Biol.* **2014**, *48* (1), 51–62. https://doi.org/10.1016/j.ibmb.2014.02.005.

(115) Zhou, J. J.; Robertson, G.; He, X.; Dufour, S.; Hooper, A. M.; Pickett, J. A.; Keep, N. H.; Field, L. M. Characterisation of Bombyx Mori Odorant-Binding Proteins Reveals That a General Odorant-Binding Protein Discriminates Between Sex Pheromone Components. *J. Mol. Biol.* **2009**, *389* (3), 529–545. https://doi.org/10.1016/j.jmb.2009.04.015.

(116) Zhu, J.; Pelosi, P.; Liu, Y.; Lin, K. J.; Yuan, H. Bin; Wang, G. R. Ligand-Binding Properties of Three Odorant-Binding Proteins of the Diamondback Moth *Plutella Xylostella*. *J. Integr. Agric.* **2016**, *15* (3), 580–590. https://doi.org/10.1016/S2095-3119(15)61067-X.

(117) Ziemba, B. P.; Murphy, E. J.; Edlin, H. T.; Jones, D. N. M. A Novel Mechanism of Ligand Binding and Release in the Odorant Binding Protein 20 from the Malaria Mosquito *Anopheles Gambiae*. *Protein Sci.* **2013**, *22* (1), 11–21. https://doi.org/10.1002/pro.2179.

(118) Leal, G. M.; Leal, W. S. Binding of a Fluorescence Reporter and a Ligand to an Odorant-Binding Protein of the Yellow Fever Mosquito, *Aedes Aegypti*. *F1000Research* **2015**, *3* (0), 305. https://doi.org/10.12688/f1000research.5879.2.

(119) Zhang, T. T.; Mei, X. D.; Feng, J. N.; Berg, B. G.; Zhang, Y. J.; Guo, Y. Y. Characterization of Three Pheromone-Binding Proteins (PBPs) of Helicoverpa Armigera (Hübner) and Their Binding Properties. *J. Insect Physiol.* **2012**, *58* (7), 941–948. https://doi.org/10.1016/j.jinsphys.2012.04.010.

(120) Laughlin, J. D.; Ha, T. S.; Jones, D. N. M.; Smith, D. P. Activation of Pheromone-Sensitive Neurons Is Mediated by Conformational Activation of Pheromone-Binding Protein. *Cell* **2008**, *133* (7), 1255–1265. https://doi.org/10.1016/j.cell.2008.04.046.

(121) Kim, M. S.; Repp, A.; Smith, D. P. LUSH Odorant-Binding Protein Mediates Chemosensory Responses to Alcohols in Drosophila Melanogaster. *Genetics* **1998**, *150* (2), 711–721.

(122) Hooper, A. M.; Dufour, S.; He, X.; Muck, A.; Zhou, J.-J.; Almeida, R.; Field, L. M.; Svatos, A.; Pickett, J. A. High-Throughput ESI-MS Analysis of Binding between the Bombyx Mori Pheromone-Binding Protein BmorPBP1, Its Pheromone Components and Some Analogues. *Chem. Commun. (Camb).* **2009**, No. 38, 5725–5727. https://doi.org/10.1039/b914294k.

(123) Oldham, N. J.; Krieger, J.; Breer, H.; Fischedick, A.; Hoskovec, M.; Svatos, A. Analysis of the Silkworm Moth Pheromone. *Angew. Chemie Int. Ed.* **2000**, *39* (23), 4341–4343. https://doi.org/10.1002/1521-3773(20001201)39:23<4341::AID-ANIE4341>3.0.CO;2-1.

(124) Wang, S.; Duan, H.; Yang, X.; Sun, Y. F.; Du, S.; Qin, Y.; Duan, H.; Yang, X. Computer-Aided Rational Design of Novel EBF Analogues with an Aromatic Ring. *J. Mol. Model.* **2016**, *22* (144). https://doi.org/10.1007/s00894-016-3011-3.

(125) Sun, Y. F.; De Biasio, F.; Qiao, H. L.; Iovinella, I.; Yang, S. X.; Ling, Y.; Riviello, L.; Battaglia, D.; Falabella, P.; Yang, X. L.; et al. Two Odorant-Binding Proteins Mediate the Behavioural Response of Aphids to the Alarm Pheromone (E)-ß-Farnesene and Structural Analogues. *PLoS One* **2012**, *7* (3), 1–10. https://doi.org/10.1371/journal.pone.0032759.

(126) Sun, D.; Guo, Z.; Liu, Y.; Zhang, Y. Progress and Prospects of CRISPR/Cas Systems in Insects and Other Arthropods. *Front. Physiol.* **2017**, *8* (SEP), 1–22. https://doi.org/10.3389/fphys.2017.00608.

(127) Dong, X.-T.; Liao, H.; Zhu, G.-H.; Khuhro, S. A.; Ye, Z.-F.; Yan, Q.; Dong, S.-L. CRISPR/Cas9 Mediated PBP1 and PBP3 Mutagenesis Induced Significant Reduction in Electrophysiological Response to Sex Pheromones in Male Chilo Suppressalis. *Insect Sci.* **2017**, *276* (November), 1–10. https://doi.org/10.3389/fevo.2017.00142.

(128) Chang, H.; Liu, Y.; Yang, T.; Pelosi, P.; Dong, S.; Wang, G. Pheromone Binding Proteins Enhance the Sensitivity of Olfactory Receptors to Sex Pheromones in *Chilo Suppressalis*. *Sci. Rep.* **2015**, *5* (August), 13093. https://doi.org/10.1038/srep13093.

(129) Zhu, G. H.; Xu, J.; Cui, Z.; Dong, X. T.; Ye, Z. F.; Niu, D. J.; Huang, Y. P.; Dong, S. L. Functional Characterization of SlitPBP3 in Spodoptera Litura by CRISPR/Cas9 Mediated Genome Editing. *Insect Biochem. Mol. Biol.* **2016**, *75*, 1–9. https://doi.org/10.1016/j.ibmb.2016.05.006.

(130) Matsuo, T.; Sugaya, S.; Yasukawa, J.; Aigaki, T.; Fuyama, Y. Odorant-Binding Proteins OBP57d and OBP57e Affect Taste Perception and Host-Plant Preference in *Drosophila Sechellia*. *PLoS Biol.* **2007**, *5* (5), 0985–0996. https://doi.org/10.1371/journal.pbio.0050118.

(131) Damberger, F.; Horst, R.; Wüthrich, K.; Peng, G.; Nikonova, L.; Leal, W. S. NMR Characterization of a PH-Dependent Equilibrium between Two Folded Solution Conformations of the Pheromone-Binding Protein from Bombyx Mori . *Protein Sci.* **2000**, *9* (5), 1038–1041. https://doi.org/10.1110/ps.9.5.1038.

(132) Pesenti, M. E.; Spinelli, S.; Bezirard, V.; Briand, L.; Pernollet, J. C.; Tegoni, M.; Cambillau, C. Structural Basis of the Honey Bee PBP Pheromone and PH-Induced Conformational Change. *J. Mol. Biol.* **2008**, *380* (1), 158–169. https://doi.org/10.1016/j.jmb.2008.04.048.

(133) Dani, F. R.; Iovinella, I.; Felicioli, A.; Niccolini, A.; Calvello, M. A.; Carucci, M. G.; Qiao, H.; Pieraccini, G.; Turillazzi, S.; Moneti, G.; et al. Mapping the Expression of Soluble Olfactory Proteins in the Honeybee. *J. Proteome Res.* **2010**, *9* (4), 1822–1833. https://doi.org/10.1021/pr900969k.

(134) Wang, S.; Minter, M.; Homem, R. A.; Michaelson, L. V.; Venthur, H.; Lim, K. S.; Withers, A.; Xi, J.; Jones, C. M.; Zhou, J.-J. Odorant Binding Proteins Promote Flight Activity in the Migratory Insect, Helicoverpa Armigera. *Mol. Ecol.* **2020**. https://doi.org/doi: 10.1111/MEC.15556.

(135) Zhang, Y. N.; Zhu, X. Y.; Fang, L. P.; He, P.; Wang, Z. Q.; Chen, G.; Sun, L.; Ye, Z. F.; Deng, D. G.; Li, J. B. Identification and Expression Profiles of Sex Pheromone Biosynthesis and Transport Related Genes in *Spodoptera Litura*. *PLoS One* **2015**, *10* (10), 1–22. https://doi.org/10.1371/journal.pone.0140019.

(136) Benton, R. Sensitivity and Specificity in Drosophila Pheromone Perception. *Trends Neurosci.* **2007**, *30* (10), 512–519. https://doi.org/10.1016/j.tins.2007.07.004.

(137) Ban, L.; Napolitano, E.; Serra, A.; Zhou, X.; Iovinella, I.; Pelosi, P. Identification of Pheromone-like Compounds in Male Reproductive Organs of the Oriental Locust *Locusta Migratoria*. *Biochem. Biophys. Res. Commun.* **2013**, *437* (4), 620–624. https://doi.org/10.1016/j.bbrc.2013.07.015.

(138) Forêt, S.; Wanner, K. W.; Maleszka, R. Chemosensory Proteins in the Honey Bee: Insights from the Annotated Genome, Comparative Analyses and Expressional Profiling. *Insect Biochem. Mol. Biol.* **2007**, *37* (1), 19–28. https://doi.org/10.1016/j.ibmb.2006.09.009.

(139) Sun, J. S.; Larter, N. K.; Chahda, J. S.; Rioux, D.; Gumaste, A.; Carlson, J. R. Humidity Response Depends on the Small Soluble Protein Obp59a in Drosophila. *Elife* **2018**, *7*, 1–19. https://doi.org/10.7554/eLife.39249.

(140) Sun, Y.; Qiao, H.; Ling, Y.; Yang, S.; Rui, C.; Pelosi, P.; Yang, X. New Analogues of (E)-β-Farnesene with Insecticidal Activity and Binding Affinity to Aphid Odorant-Binding Proteins. *J. Agric. Food Chem.* **2011**, *59* (6), 2456–2461. https://doi.org/10.1021/jf104712c.

(141) Pandit, S. B.; Zhang, Y.; Skolnick, J. TASSER-Lite: An Automated Tool for Protein Comparative Modeling. *Biophys. J.* **2006**, *91* (11), 4180–4190. https://doi.org/10.1529/biophysj.106.084293.

(142) Lin, W.; Yu, Y.; Zhou, P.; Zhang, J.; Dou, L.; Hao, Q.; Chen, H.; Zhu, S. Identification and Knockdown of the Olfactory Receptor (OrCo) in Gypsy Moth, Lymantria Dispar. *Int. J. Biol. Sci.* **2015**, *11* (7), 772–780. https://doi.org/10.7150/ijbs.11898.

(143) Brezolin, A. N.; Martinazzo, J.; Muenchen, D. K.; de Cezaro, A. M.; Rigo, A. A.; Steffens, C.; Steffens, J.; Blassioli-Moraes, M. C.; Borges, M. Tools for Detecting Insect Semiochemicals: A Review. *Anal. Bioanal. Chem.* **2018**, *410* (17), 4091–4108. https://doi.org/10.1007/s00216-018-1118-3.

(144) Wasilewski, T.; Szulczyński, B.; Kamysz, W.; Gębicki, J.; Namieśnik, J. Evaluation of

Three Peptide Immobilization Techniques on a Qcm Surface Related to Acetaldehyde Responses in the Gas Phase. *Sensors (Switzerland)* **2018**, *18* (11), 1–15. https://doi.org/10.3390/s18113942.

(145) Woodcock, C.; Birkett, M. A. *Personal Communication: Aphid Electrophysiological Data*.

(146) Zada, A.; Harel, M. Enzymatic Transesterification of Racemic Lavandulol: Preparation of the Two Enantiomeric Alcohols and of the Two Enantiomers of Lavandulyl Senecioate. *Tetrahedron Asymmetry* **2004**, *15* (15), 2339–2343. https://doi.org/10.1016/j.tetasy.2004.06.015.

(147) Hooper, A. M.; Donato, B.; Woodcock, C. M.; Park, J. H.; Paul, R. L.; Boo, K. S.; Hardie, J.; Pickett, J. A. Characterization of (1R,4S,4aR,7S,7aR)-Dihydronepetalactol as a Semiochemical for Lacewings, Including Chrysopa Spp. and *Peyerimhoffina Gracilis*. *J. Chem. Ecol.* **2002**, *28* (4), 849–864. https://doi.org/10.1023/A:1015201129331.

(148) Guillemonat, A. Oxidation of Ethylenic Hydrocarbons with Selenium Dioxide. *Ann. Chim. Fr.* **1939**, *11*, 143–211.

(149) Stephenson, L. M.; Speth, D. R. Mechanism of Allylic Hydroxylation by Selenium Dioxide. *J. Org. Chem.* **1979**, *44* (25), 4683–4689. https://doi.org/10.1021/jo00393a045.

(150) Sauer, J. Diels-Alder Reactions II: The Reaction Mechanism. *Angew. Chemie Int. Ed. English* **1967**, *6* (1), 16–33. https://doi.org/10.1002/anie.196700161.

(151) García, J. I.; Mayoral, J. A.; Salvatella, L. The Source of the Endo Rule in the Diels-Alder Reaction: Are Secondary Orbital Interactions Really Necessary? *European J. Org. Chem.* **2004**, No. 1, 85–90. https://doi.org/10.1002/ejoc.200400424.

(152) Bodwell, G. J.; Pi, Z.; Pottie, I. R. Electron Deficient Dienes. 2. One Step Synthesis of a Coumarin-Fused Electron Deficient Diene and Its Inverse Electron Demand Diels-Alder Reactions with Enamines. *Synlett* **1999**, *10* (4), 477–479. https://doi.org/10.1055/s-1999-2645.

(153) Prestwich, G. D. Bacterial Expression and Photoaffinity Labeling Of. *Protein Sci.* **1993**, *2*, 420–428.

(154) Chung, C. T.; Niemela, S. L.; Miller, R. H. One-Step Preparation of Competent Escherichia Coli: Transformation and Storage of Bacterial Cells in the Same Solution. *Biochemistry* **1989**, *86* (April), 2172–2175.

(155) Przybylski, M.; Glocker, M. O. Electrospray Mass Spectrometry of Biomacromolecular Complexes with Noncovalent Interactions - New Analytical Perspectives for Supramolecular Chemistry and Molecular Recognition Processes. *ChemInform* **2010**, *27* (31), 806–826. https://doi.org/10.1002/chin.199631320.

(156) Winkler, R. ESIprot: A Universal Tool for Charge State Determination and Molecular Weight Calculation of Proteins from Electrospray Ionization Mass Spectrometry Data. *Rapid Commun. Mass Spectrom.* **2010**, *24*, 1457–1466. https://doi.org/10.1002/rcm.

(157) ExPASy. ProtParam https://web.expasy.org/protparam/.

(158) Gasteiger, E.; Hoogland, C.; Gattiker, A.; Duvaud, S.; Wilkins, M. R.; Appel, R. D.; Bairoch, A. The Proteomics Protocols Handbook. *Proteomics Protoc. Handb.* **2005**, 571–608. https://doi.org/10.1385/1592598900.

(159) Salmaso, V.; Moro, S. Bridging Molecular Docking to Molecular Dynamics in Exploring Ligand-Protein Recognition Process: An Overview. *Front. Pharmacol.* **2018**, *9* (AUG), 1–16. https://doi.org/10.3389/fphar.2018.00923.

(160) Venthur, H.; Mutis, A.; Zhou, J. J.; Quiroz, A. Ligand Binding and Homology Modelling of Insect Odorant-Binding Proteins. *Physiol. Entomol.* **2014**, *39* (3), 183–198. https://doi.org/10.1111/phen.12066.

(161) Forli, S.; Huey, R.; Pique, M. E.; Sanner, M.; Goodsell, D. S.; Arthur, J. Computational Protein-Ligand Docking and Virtual Drug Screening with the AutoDock Suite. *Nat. Protoc.* **2016**, *11* (5), 905–919. https://doi.org/10.1038/nprot.2016.051.Computational.

(162) Carugo, O.; Djinovic-Carugo, K. Half a Century of Ramachandran Plots. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2013**, *69* (8), 1333–1341. https://doi.org/https://doi.org/10.1107/S090744491301158X.

(163) Chen, V. B.; Arendall, W. B.; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C. MolProbity: All-Atom Structure Validation for Macromolecular Crystallography. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66* (1), 12–21. https://doi.org/10.1107/S0907444909042073.

(164) Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. Automated Docking Using a Lamarckian Genetic Algorithm and an Empirical Binding Free Energy Function. *J. Comput. Chem.* **1998**, *19* (14), 1639–1662.

(165) Hopf, T. A.; Morinaga, S.; Ihara, S.; Touhara, K.; Marks, D. S.; Benton, R. Amino Acid Coevolution Reveals Three-Dimensional Structure and Functional Domains of Insect Odorant Receptors. *Nat. Commun.* **2015**, *6*, 1–7. https://doi.org/10.1038/ncomms7077.

(166) Batra, S.; Corcoran, J.; Zhang, D. D.; Pal, P.; Umesh, K. P.; Kulkarni, R.; Löfstedt, C.; Sowdhamini, R.; Olsson, S. B. A Functional Agonist of Insect Olfactory Receptors: Behavior, Physiology and Structure. *Front. Cell. Neurosci.* **2019**, *13* (April), 1–13. https://doi.org/10.3389/fncel.2019.00134.

(167) Van Gunsteren, W. F.; Berendsen, H. J. C. Molecular Dynamics Computer Simulation. Method, Application and Perspectives in Chemistry. *Sect. Title Phys. Org. Chem.* **1990**, *102* (9), 1020–1055.

(168) Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. GROMACS: Fast, Flexible, and Free. *J. Comput. Chem.* **2005**, *26* (16), 1701–1718. https://doi.org/10.1002/jcc.20291.

(169) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindah, E. Gromacs: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015**, *1–2*, 19–25. https://doi.org/10.1016/j.softx.2015.06.001.

(170) Krieger, E.; Joo, K.; Lee, J.; Lee, J.; Raman, S.; Thompson, J.; Tyka, M.; Baker, D.; Karplus, K. Improving Physical Realism, Stereochemistry, and Side-Chain Accuracy in Homology Modeling: Four Approaches That Performed Well in CASP8. *Proteins Struct. Funct. Bioinforma.* **2009**, *77* (9), 114–122. https://doi.org/10.1002/prot.22570.

(171) Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M. PROCHECK: A Program to Check the Stereochemical Quality of Protein Structures. *J. Appl.*

Crystallogr. **1993**, *26* (2), 283–291. https://doi.org/10.1107/s0021889892009944.

(172) Webb, B.; Sali, A. Comparative Protein Structure Modeling Using MODELLER. *Curr. Protoc. Bioinforma.* **2017**, *54* (5.6), 1–37. https://doi.org/doi:10.1002/cpbi.3.

(173) Wiederstein, M.; Sippl, M. J. ProSA-Web: Interactive Web Service for the Recognition of Errors in Three-Dimensional Structures of Proteins. *Nucleic Acids Res.* **2007**, *35* (2), 407–410. https://doi.org/10.1093/nar/gkm290.

(174) Madeira, F.; Park, Y. M.; Lee, J.; Buso, N.; Gur, T.; Madhusoodanan, N.; Basutkar, P.; Tivey, A. R. N.; Potter, S. C.; Finn, R. D.; et al. The EMBL-EBI Search and Sequence Analysis Tools APIs in 2019. *Nucleic Acids Res.* **2019**, *47* (W1), W636–W641. https://doi.org/10.1093/nar/gkz268.

(175) Nichols, A. S.; Luetje, C. W. Transmembrane Segment 3 of Drosophila Melanogaster Odorant Receptor Subunit 85b Contributes to Ligand-Receptor Interactions. *J. Biol. Chem.* **2010**, *285* (16), 11854–11862. https://doi.org/10.1074/jbc.M109.058321.

(176) Ritchie, D. W.; Grudinin, S. Spherical Polar Fourier Assembly of Protein Complexes with Arbitrary Point Group Symmetry. *J. Appl. Crystallogr.* **2016**, *49*, 158–167. https://doi.org/10.1107/S1600576715022931.

(177) Corcoran, J. A.; Sonntag, Y.; Andersson, M. N.; Johanson, U.; Löfstedt, C. Endogenous Insensitivity to the Orco Agonist VUAA1 Reveals Novel Olfactory Receptor Complex Properties in the Specialist Fly *Mayetiola Destructor*. *Sci. Rep.* **2018**, *8* (1), 1–13. https://doi.org/10.1038/s41598-018-21631-3.

(178) Brylinski, M. Aromatic Interactions at the Ligand-Protein Interface: Implications for the Development of Docking Scoring Functions. *Chem Biol Drug Des.* **2018**, *91* (2), 380–390. https://doi.org/doi:10.1111/cbdd.13084.

(179) Gonzales, W. L.; Ramirez, C. C.; Olea, N.; Niemeyer, H. M. Host Plant Changes Produced by the Aphid Sipha Flava: Consequences for Aphid Feeding Behaviour and Growth. *Entomol. Exp. Appl.* **2002**, *103* (2), 107–113. https://doi.org/10.1046/j.1570-7458.2002.00964.x.

(180) Nibouche, S.; Mississipi, S.; Fartek, B.; Delatte, H.; Reynaud, B.; Costet, L. Host Plant Specialization in the Sugarcane Aphid *Melanaphis Sacchari*. *PLoS One* **2015**, *10* (11), 1–13. https://doi.org/10.1371/journal.pone.0143704.

(181) Abassi, S. A. L.; Birkett, M. A.; Pettersson, J.; Pickett, J. A.; Wadhams, L. J.; Woodcock, C. M. Response of the Seven-Spot Ladybird to an Aphid Alarm Pheromone and an Alarm Pheromone Inhibitor Is Mediated by Paired Olfactory Cells. *J. Chem. Ecol.* **2000**, *26* (7), 1765–1771. https://doi.org/10.1023/A:1005555300476.

(182) Bowers, W.; Nault, L.; Webb, R.; Dutky, S. Aphid Alarm Pheromone: Isolation, Identification, Synthesis. *Science (80-.).* **1972**, *177* (4054), 1121–1122. https://doi.org/DOI: 10.1126/science.177.4054.1121.

(183) Edwards, L. J.; Siddall, J. B.; Dunham, L. L.; Uden, P.; Kislow, C. Trans-β-Farnesene, Alarm Pheromone of the Green Peach Aphid, Myzus Persicae (Sulzer). *Nature* **1973**, *241* (5385), 126–127. https://doi.org/10.1038/246421a0.

(184) Mahendrarajah, K.; Dalby, P. A.; Wilkinson, B.; Jackson, S. E.; Main, E. R. G. A High-Throughput Fluorescence Chemical Denaturation Assay as a General Screen for Protein-Ligand Binding. *Anal. Biochem.* **2011**, *411* (1), 155–157. https://doi.org/10.1016/j.ab.2010.12.001.

(185) Eftink, M. R. Intrinsic Fluorescence of Proteins. In *Topics in Fluorescence Spectroscopy: Volume 6 Protein Fluorescence*; 2000; pp 1–15.

(186) Tan, J.; Zaremska, V.; Lim, S.; Knoll, W.; Pelosi, P. Probe-Dependence of Competitive Fluorescent Ligand Binding Assays to Odorant-Binding Proteins. *Anal. Bioanal. Chem.* **2020**, *412* (3), 547–554. https://doi.org/10.1007/s00216-019-02309-9.

(187) Vivian, J. T.; Callis, P. R. Mechanisms of Tryptophan Fluorescence Shifts in Proteins. *Biophys. J.* **2001**, *80* (5), 2093–2109. https://doi.org/10.1016/S0006-3495(01)76183-8.

(188) Murphy, E. J.; Booth, J. C.; Davrazou, F.; Port, A. M.; Jones, D. N. M. Interactions of *Anopheles Gambiae* Odorant-Binding Proteins with a Human-Derived Repellent: Implications for the Mode of Action of N,N-Diethyl-3-Methylbenzamide (DEET). *J. Biol. Chem.* **2013**, *288* (6), 4475–4485. https://doi.org/10.1074/jbc.M112.436386.

(189) Gong, Y.; Plettner, E. Effects of Aromatic Compounds on Antennal Responses and on the Pheromone-Binding Proteins of the Gypsy Moth (*Lymantria Dispar*). *Chem. Senses* **2011**, *36* (3), 291–300. https://doi.org/10.1093/chemse/bjq130.

(190) Murphy, E. J.; Booth, J. C.; Davrazou, F.; Port, A. M.; Jones, D. N. M. Interactions of Anopheles Gambiae Odorant-Binding Proteins with a Human-Derived Repellent: Implications for the Mode of Action of N,N-Diethyl-3-Methylbenzamide (DEET). *J. Biol. Chem.* **2013**, *288* (6), 4475–4485. https://doi.org/10.1074/jbc.M112.436386.

(191) Du, S.; Yang, Z.; Qin, Y.; Wang, S.; Duan, H.; Yang, X. Computational Investigation of the Molecular Conformation-Dependent Binding Mode of (E)-β-Farnesene Analogs with a Heterocycle to Aphid Odorant-Binding Proteins. **2018**, *7*.

(192) Lizana, P.; Machuca, J.; Larama, G.; Quiroz, A.; Mutis, A.; Venthur, H. Mating-Based Regulation and Ligand Binding of an Odorant-Binding Protein Support the Inverse Sexual Communication of the Greater Wax Moth, Galleria Mellonella (Lepidoptera: Pyralidae). *Insect Mol. Biol.* **2020**, No. 56. https://doi.org/10.1111/imb.12638.

(193) Bucci, B. K.; Kruse, S. W.; Thode, A. B.; Alvarado, S. M.; Jones, D. N. M. Effect of N-Alcohols on the Structure and Stability of the Drosophila Odorant Binding Protein LUSH. *Biochemistry* **2006**, *45* (6), 1693–1701. https://doi.org/10.1021/bi0516576.

(194) West, G. M.; Tang, L.; Fitzgerald, M. C. Thermodynamic Analysis of Protein Stability and Ligand Binding Using a Chemical Modification- and Mass Spectrometry-Based Strategy. *Anal. Chem.* **2008**, *80* (11), 4175–4185.

(195) Jones, D. N. M.; Wang, J.; Murphy, E. J. Complete NMR Chemical Shift Assignments of Odorant Binding Protein 22 from the Yellow Fever Mosquito, *Aedes Aegypti*, Bound to Arachidonic Acid. *Biomol. NMR Assign.* **2019**, *0* (0), 0. https://doi.org/10.1007/s12104-019-09875-0.

(196) Tan, J.; Zaremska, V.; Lim, S.; Knoll, W.; Pelosi, P. Probe-Dependence of Competitive Fluorescent Ligand Binding Assays to Odorant-Binding Proteins. *Anal. Bioanal. Chem.* **2019**. https://doi.org/10.1007/s00216-019-02309-9.

(197) Tetko, I. V.; Gasteiger, J.; Todeschini, R.; Mauri, A.; Livingstone, D.; Ertl, P.; Palyulin, V. A.; Radchenko, E. V.; Zefirov, N. S.; Makarenko, A. S.; et al. Virtual Computational Chemistry Laboratory - Design and Description. *J. Comput. Aided. Mol. Des.* **2005**, *19* (6), 453–463. https://doi.org/10.1007/s10822-005-8694-y.

(198) Lee, D.; Damberger, F. F.; Peng, G.; Horst, R.; Güntert, P.; Nikonova, L.; Leal, W. S.;

Wüthrich, K. NMR Structure of the Unliganded *Bombyx Mori* Pheromone-Binding Protein at Physiological PH. *FEBS Lett.* **2002**, *531* (2), 314–318. https://doi.org/10.1016/S0014-5793(02)03548-2.

(199) Marley, J.; Lu, M.; Bracken, C. A Method for Efficient Isotopic Labeling of Recombinant Proteins. *J. Biomol. NMR* **2001**, *20* (1), 71–75. https://doi.org/10.1023/A:1011254402785.

(200) Davrazou, F.; Dong, E.; Murphy, E. J.; Johnson, H. T.; Jones, D. N. M. New Insights into the Mechanism of Odorant Detection by the Malaria-Transmitting Mosquito *Anopheles Gambiae*. *J. Biol. Chem.* **2011**, *286* (39), 34175–34183. https://doi.org/10.1074/jbc.M111.274712.

(201) Mayer, M.; Meyer, B. Characterization of Ligand Binding by Saturation Transfer Difference NMR Spectroscopy. *Angew. Chemie - Int. Ed.* **1999**, *38* (12), 1784–1788. https://doi.org/10.1002/(SICI)1521-3773(19990614)38:12<1784::AID-ANIE1784>3.0.CO;2-Q.

(202) Xia, Y.; Zhu, Q.; Jun, K. Y.; Wang, J.; Gao, X. Clean STD-NMR Spectrum for Improved Detection of Ligand-Protein Interactions at Low Concentration of Protein. *Magn. Reson. Chem.* **2010**, *48* (12), 918–924. https://doi.org/10.1002/mrc.2687.

(203) Bernardo, V.-B.; Oldham, N. J. Personal Communication. 2018.

(204) Selenko, P.; Wagner, G. Looking into Live Cells with In-Cell NMR Spectroscopy. *J. Struct. Biol.* **2007**, *158* (2), 244–253. https://doi.org/10.1016/j.jsb.2007.04.001.

(205) Mayer, M.; Meyer, B. Group Epitope Mapping by Saturation Transfer Difference NMR to Identify Segments of a Ligand in Direct Contact with a Protein Receptor. *J. Am. Chem. Soc.* **2001**, *123* (25), 6108–6117. https://doi.org/10.1021/ja0100120.

(206) Puchner, C.; Eixelsberger, T.; Nidetzky, B.; Brecker, L. Saturation Transfer Difference NMR to Study Substrate and Product Binding to Human UDP-Xylose Synthase (HUXS1A) during Catalytic Event. *RSC Adv.* **2015**, *5* (106), 86919–86926. https://doi.org/10.1039/c5ra18284k.

(207) Brecker, L.; Straganz, G. D.; Tyl, C. E.; Steiner, W.; Nidetzky, B. Saturation-Transfer-Difference NMR to Characterize Substrate Binding Recognition and Catalysis of Two Broadly Specific Glycoside Hydrolases. *J. Mol. Catal. B Enzym.* **2006**, *42* (3–4), 85–89. https://doi.org/10.1016/j.molcatb.2006.07.004.

(208) R. R Core Team, R foundation for Statistical Computing: Vienna, Austria 2018.

(209) Hou, X.; Zhang, D. D.; Yuvaraj, J. K.; Corcoran, J. A.; Andersson, M. N.; Löfstedt, C. Functional Characterization of Odorant Receptors from the Moth *Eriocrania Semipurpurella*: A Comparison of Results in the Xenopus Oocyte and HEK Cell Systems. *Insect Biochem. Mol. Biol.* **2020**, *117* (November 2019), 103289. https://doi.org/10.1016/j.ibmb.2019.103289.

(210) Xu, P.; Leal, W. S. Probing Insect Odorant Receptors with Their Cognate Ligands: Insights into Structural Features. *Biochem. Biophys. Res. Commun.* **2013**, *435* (3), 477–482. https://doi.org/10.1016/j.bbrc.2013.05.015.

(211) Sato, K.; Pellegrino, M.; Nakagawa, T.; Nakagawa, T.; Vosshall, L. B.; Touhara, K. Insect Olfactory Receptors Are Heteromeric Ligand-Gated Ion Channels. *Nature* **2008**, *452* (7190), 1002–1006. https://doi.org/10.1038/nature06850.

(212) Campanacci, V.; Lartigue, A.; Hällberg, B. M.; Jones, T. A.; Giudici-Orticoni, M. T.; Tegoni, M.; Cambillau, C. Moth Chemosensory Protein Exhibits Drastic Conformational Changes and Cooperativity on Ligand Binding. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100* (9), 5069–5074. https://doi.org/10.1073/pnas.0836654100.

(213) Willot, M.; Radtke, L.; Könning, D.; Fröhlich, R.; Gessner, V. H.; Strohmann, C.; Christmann, M. Total Synthesis and Absolute Configuration of the Guaiane Sesquiterpene Englerin A. *Angew. Chemie - Int. Ed.* **2009**, *48* (48), 9105–9108. https://doi.org/10.1002/anie.200905032.

(214) Uesato, S.; Kobayashi, K.; Inouye, H. Studies on Monoterpene Glucosides and Related Natural Products: Synthesis of 13C-Labeled Acyclic Monoterpenes for Studies on the Mechanism of the Iridane Skeleton Formation in the Biosynthesis of Iridoid Glucosides. *Chem. Pharm. Bull.* **1982**, *30* (3), 927–940.

(215) Berman, H. M.; Battistuz, T.; Bhat, T. N.; Bluhm, W. F.; Bourne, P. E.; Burkhardt, K.; Feng, Z.; Gilliland, G. L.; Iype, L.; Jain, S.; et al. The Protein Data Bank. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2002**, *58* (6 I), 899–907. https://doi.org/10.1107/S0907444902003451.

(216) Heringa, J. Two Strategies for Sequence Comparison: Profile-Preprocessed and Secondary Structure-Induced Multiple Alignment. *Comput. Chem.* **1999**, *23*, 341–364.

(217) Pirovano, W.; Feenstra, K. A.; Heringa, J. PRALINE^{TM}: A Strategy for Improved Multiple Alignment of Transmembrane Proteins. *Bioinformatics* **2008**, *24* (4), 492–497. https://doi.org/10.1093/bioinformatics/btm636.

(218) Hofmann, K.; Stoffel, W. TMbase - A Database of Membrane Spanning Proteins Segments. *Biol. Chem.* **1993**.

(219) Tusnady, G. E.; Simon, I. The HMMTOP Transmembrane Topology Prediction Server. *Bioinformatics* **2001**, *17*, 849–850.

(220) Käll, L.; Krogh, A.; Sonnhammer, E. L. L. A Combined Transmembrane Topology and Signal Peptide Prediction Method. *J. Mol. Biol.* **2004**, *338* (5), 1027–1036.

(221) Krogh, A.; Larsson, G.; von Heijne, G.; Sonnhammer, E. L. L. Predicting Transmembrane Protein Topology with a Hidden Markov Model: Application to Complete Genomes. *J. Mol. Biol.* **2001**, *305* (3), 567–580.

(222) GeneDoc https://genedoc.software.informer.com/2.7/.

(223) FigTree http://tree.bio.ed.ac.uk/software/figtree/.

(224) Peter Tieleman. Biocomputing Group http://wcm.ucalgary.ca/tieleman/downloads.

(225) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. NMRPipe: A Multidimensional Spectral Processing System Based on UNIX Pipes. *J. Biomol. NMR* **1995**, *6*, 277–293.

(226) Leicester, U. of. CCPNMR. 2013.

(227) Addgene http://www.addgene.org/.

Appendix

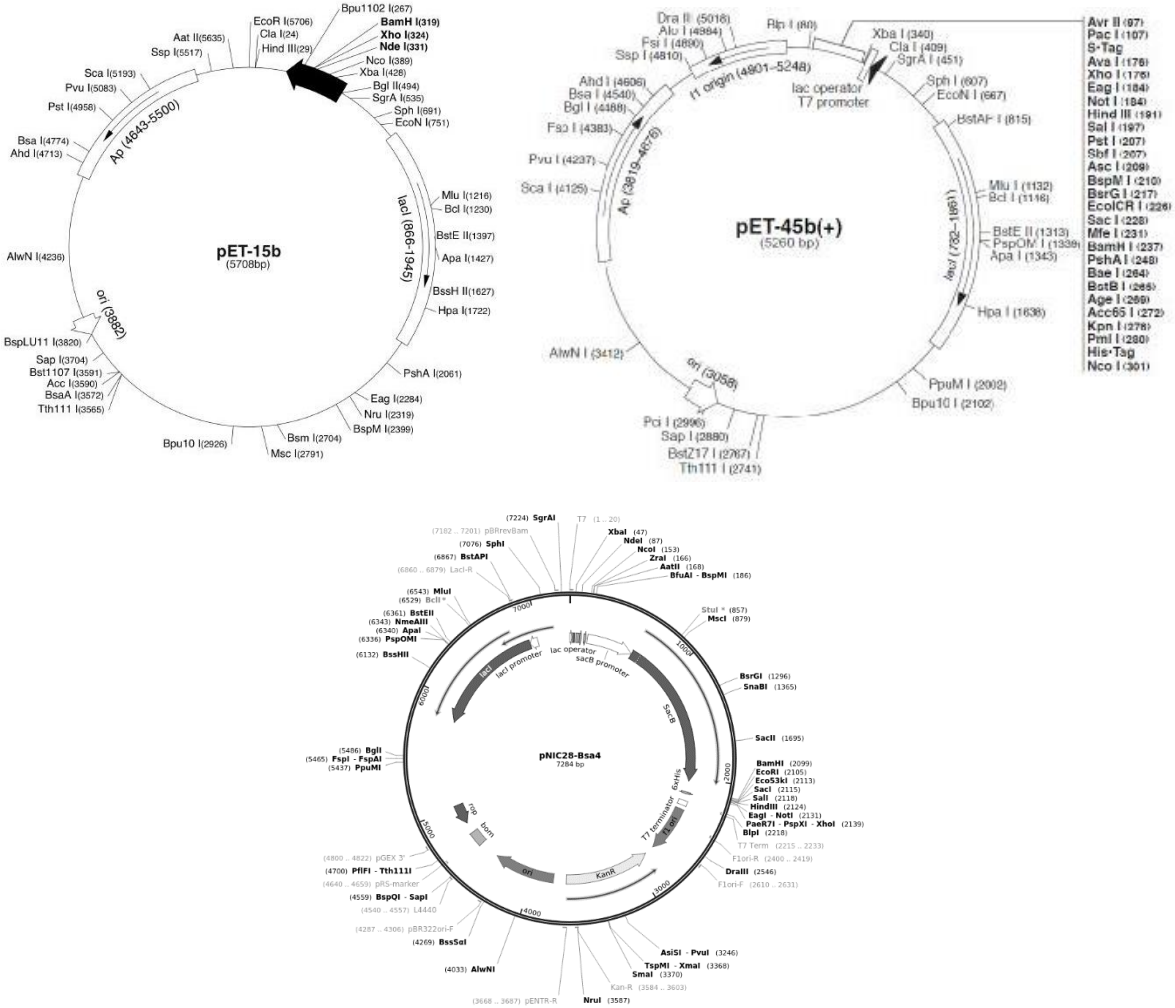


Figure A.1: Vector maps for the three utilised plasmids, pET-15b, pET45b and pNIC28-Bsa4. Images from [227]

Table A.1: Results from Autodock ligand-screening. The aphid sex pheromone components and respective enantiomers were tested, along with the aphid alarm pheromone.

NA = no favourable docking conformations were found in the screening

Ligand	ApisOBP1		ApisOBP2		ApisOBP3		ApisOBP4		ApisOBP5		ApisOBP10		ApisOBP11	
	Binding energy (kcal mol^{-1})	**K_i (μM)**	**Binding energy (kcal mol^{-1})**	**K_i (μM)**	**Binding energy (kcal mol^{-1})**	**K_i (μM)**	**Binding energy (kcal mol^{-1})**	**K_i (μM)**	**Binding energy (kcal mol^{-1})**	**K_i (μM)**	**Binding energy (kcal mol^{-1})**	**K_i (μM)**	**Binding energy (kcal mol^{-1})**	**K_i (μM)**
(1*R*,4a*S*,7*S*,7a*R*)-nepetalactol **5**	-5.85	51.09	8.79	NA	-6.14	31.33	-5.94	44.08	-4.28	725.14	1.26	NA	-5.92	46.07
(4a*S*,7*S*,7a*R*)-nepetalactone **6**	-6.60	14.47	2.33	NA	NA	NA	-5.86	50.36	-5.00	214.56	2.60	NA	-6.44	18.94
(1*S*,4a*R*,7*R*,7a*S*)-nepetalactol **32**	-6.05	37.0	10.78	NA	NA	NA	-6.45	18.55	-4.24	784.16	10.06	NA	-5.94	43.96
(4a*R*,7*R*,7a*S*)-nepetalactone **33**	-6.53	16.37	10.1	NA	NA	NA	-6.17	30.01	-5.07	192.32	NA	NA	-6.39	20.67

Table A.2: Results from Autodock ligand-screening. The aphid sex pheromone components and respective enantiomers were tested, along with the aphid alarm pheromone.

NA = no favourable docking conformations were found in the screening

Ligand	ApisOBP1		ApisOBP2		ApisOBP3		ApisOBP4		ApisOBP5		ApisOBP10		ApisOBP11	
	Binding energy (kcal mol^{-1})	**K_i (µM)**	**Binding energy (kcal mol^{-1})**	**K_i (µM)**	**Binding energy (kcal mol^{-1})**	**K_i (µM)**	**Binding energy (kcal mol^{-1})**	**K_i (µM)**	**Binding energy (kcal mol^{-1})**	**K_i (µM)**	**Binding energy (kcal mol^{-1})**	**K_i (µM)**	**Binding energy (kcal mol^{-1})**	**K_i (µM)**
(*E*)-β-farnesene **17**	-6.63	13.84	7.20	NA	-3.22	4390	-5.51	90.71	-4.00	1160.00	-0.86	2.34 x 10^5	-6.16	30.57
(*S*)-germacrene D **25**	-7.56	2.88	42.76	NA	NA	NA	-5.52	89.51	-4.69	362.97	-3.12	5.14	-7.25	4.88
(1*R*,4*E*,9*S*)-caryophyllene **26**	-8.10	1.15	119.86	NA	NA	NA	-6.62	14.06	-4.93	244.74	22.76	NA	-7.46	3.40
Myrcene	-5.05	197.69	1.50	NA	-5.81	54.81	-4.61	419.57	-4.31	697.40	-4.51	498.18	-4.75	329.06
Ocimene	-5.27	137.93	-0.66	6.91 x 10^5	-6.31	23.83	-4.81	297.90	-4.41	581.33	-4.29	720.97	-5.01	213.44
(*R*)-linalool	-5.26	139.24	1.6	NA	-6.05	36.94	-4.66	318.30	-4.14	923.66	-4.38	619.76	-4.95	233.60
(*S*)-linalool	-5.20	153.43	-1.7	1.39 x 10^5	-5.39	111.49	-4.57	448.12	-4.32	675.86	-3.33	3630.00	-4.82	292.68
(*Z*)-jasmone **30**	-6.29	24.51	-1.55	7.2 x 10^4	-6.31	23.74	-5.56	83.97	-5.14	171.84	-4.96	230.18	-6.30	24.25

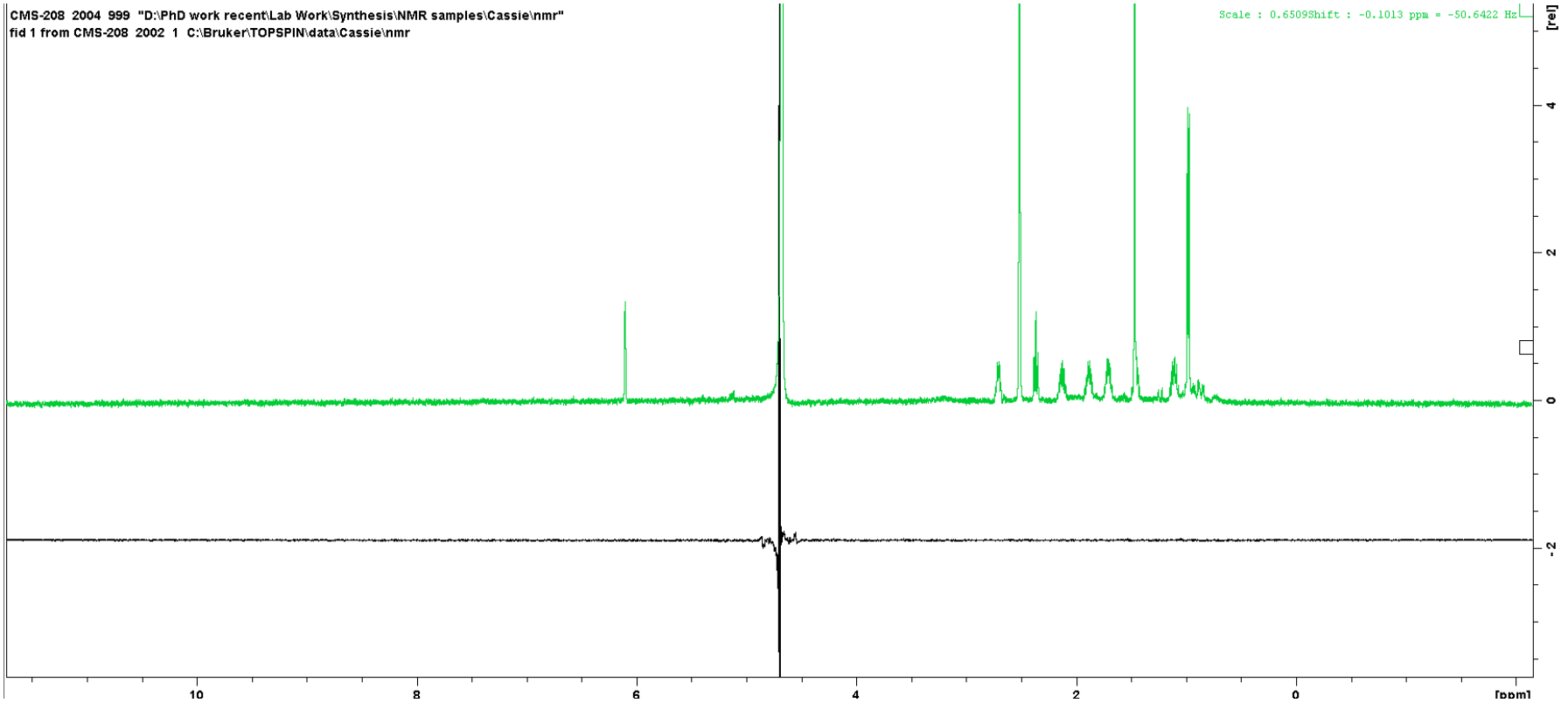


Figure A.2: The control STD NMR spectra of *A. pisum* OBP6 with (4a*S*,7*S*,7a*R*)-nepetalactone **6**. The initial ^{1}H spectra (green spectra), in addition to the STD NMR (black spectra).