Universidad de La Frontera
Facultad de Ingeniería y Ciencias
Programa de Doctorado en Ciencias de Recursos Naturales

Identification of chemical compounds with putative semiochemical activity towards adult *Hylamorpha elegans* Burmeister (Coleoptera: Scarabaeidae) through a computational reverse chemical ecology approach

Doctoral thesis
In fulfillment of the
Requirements for the Degree
Doctor in Sciences of Natural Resources
By

Herbert Max Venthur Peña

Temuco – Chile 2015

Identification of chemical compounds with putative semiochemical activity towards adult *Hylamorpha elegans* Burmeister (Coleoptera: Scarabaeidae) through a computational reverse chemical ecology approach

Esta tesis fue realizada bajo la supervisión del Dr. Andrés Quiroz y Dra. Ana Mutis, ambos pertenecientes al Departamento de Ciencias Químicas y Recursos Naturales de la Universidad de La Frontera y es presentada para su revisión por los miembros de la comisión examinadora.

Herbert Max Venthur Peña

Dr. Francisco Matus ***Director Programa de Doctorado en Ciencias de Recursos Naturales***	***Dr. Andrés Quiroz***
Dr. Juan Carlos Parra ***Director Académico de Postgrado Universidad de La Frontera***	***Dra. Ana Mutis***
	Dr. Bruce Cassels
	Dr. Gonzalo Gajardo
	Dr. León Bravo
	Dr. Milko Jorquera

“No quiero temer a fracasar en mis objetivos…

pero sí a no intentar cumplirlos.”

Agradecimientos

Hoy pareciera que cierro una gran y desafiante etapa en mi vida. Veo a mi familia a mi lado, mi hijo jugando, mi pareja acompañándome, tal como mi madre y hermano lo hacen, y mientras escucho mi banda favorita, quisiera plasmar en estas líneas mis agradecimientos más sinceros a quienes han hecho de esto una realidad.

Quiero agradecer a mis tutores, Dr. Andrés Quiroz y Dra. Ana Mutis por su sabiduría entregada al igual que su paciencia y tolerancia. La confianza que siento han depositado en mí posee un valor incalculable, y espero poder mantenerlo en el tiempo. Asimismo, quiero agradecer el apoyo financiero de las becas Conicyt de doctorado, pasantía y gastos operacionales. También, a todos los miembros del Laboratorio de Química Ecológica, Leo, Loreto, Marcelo, Pao, Emilio, y tantos otros que han contribuido de manera importante en mi formación, no sólo como investigador y profesional, si no también como persona. No puedo dejar atrás a quienes de manera sorpresiva, pero grata, contribuyeron a que esté escribiendo esto. Mis agradecimientos al Dr. Jing-Jiang Zhou por su hospitalidad y gran ayuda durante mi pasantía doctoral. También, a las grandes personas que conocí en aquel enriquecedor viaje, Luciana Palazzo, Leonela Caravajal, Diego Magalhães, Alex Barton, entre otros.

De manera más generalizada pero no menos importante, quisiera agradecer a todos los profesores del Departamento de Ciencias Químicas y Recursos Naturales que alguna vez me hicieron clases y que sin duda fueron parte esencial de mi formación profesional. A los profesores Fernando Perich, Fernando Pardo, Felipe Gallardo, Mario Pino, Marysol Alvear, Juan Granifo, Moisés Vargas, grandes nombres que vienen a la mente.

Mi búsqueda y generación de nuevo conocimiento ha sido influenciada y apoyada fuertemente por las personas que me han rodeado. Un gran “gracias” a mis amigos de la vida, Héctor Jiménez, Ricardo Gutierrez y Jorge Rojas quienes han estado ahí siempre. Por último, agradecimientos de corazón a mi familia, mi hijo Matthias, mi princesa y compañera de vida Jhoselyn, mamá Ketty, mi hermano Rolando, Bárbara, tía Paola, Bastián

y Fernanda por su constante apoyo y preocupación. Gracias mamá, hermano y amor por escucharme todas esas tardes de once en donde les comentaba con ceño fruncido mis preocupaciones entorno al Doctorado, las que a final de cuentas resultaron ser positivas.

Finalmente, agradecer a tres personas que sé me cuidan y guían desde algún lugar ahí arriba. A mi papá Carlos Venthur y abuelos Inés y Pedro, los llevo en mi mente y corazón hoy y siempre.

¡Gracias!

Summary and outline of this thesis

In Chile, *Hylamorpha elegans* (Coleoptera: Scarabaeidae, Rutelinae) is recognized as an underground pest during its larval stage feeding on roots of crops, such as cereals and grass. Thus, *H. elegans* represents an important pest in Central South and South of Chile. Considering its difficult control, it is thought that the application of recent molecular approaches could provide new opportunities to develop specific control strategies for *H. elegans*. For example, computational reverse chemical ecology, as a complement between experimental and *in silico* techniques, uses insect odorant-binding proteins (OBPs) as targets. As the main function of OBPs is to carry semiochemicals inside the olfactory organs of insects, their putative selectivity becomes to OBPs as an attractive field of research, where new semiochemicals for pest control can be discovered. Therefore, the main objective of this thesis considered the identification of chemical compounds with putative semiochemical activity towards adult *H. elegans*. Here, molecular biology techniques along with *in silico* methods have provided results enough to identify and characterize an OBP from antennae of male and female *H. elegans*. The protein, called HeleOBP1, was cloned, sub-cloned, recombinantly expressed and purified. Fluorescence binding assays were performed using several ligands with the recombinant protein as well as molecular docking with its modelled structure. Finally, electroantennography (EAG) and behavioural assays to determine their bioactivity were used. From 29 ligands, α-ionone, β-ionone and 2-phenyl acetaldehyde showed the strongest binding to HeleOBP1. On the other hand, HeleOBP1-ligand complexes obtained from molecular docking suggested the participation of His102, Tyr105, Tyr113 and Met114. It is thought that these residues could form π-π interactions with the unsaturations and aromatic rings of ligands. Moreover, our findings show high sensitivity of males and a moderately attractive effect for β-ionone at low concentration suggested by EAG and behavioural recordings. Finally, it is proposed that β-ionone might be detected specifically by HeleOBP1 and other OBPs could participate for the recognition of other semiochemicals. Further perspectives for the application of computational reverse chemical ecology methods are discussed.

Table of contents

Chapter 1. General introduction
1.1. Introduction 1
1.2. Hypothesis and objectives 4
1.2.1. Hypothesis 4
1.2.2. General objective 4
1.2.3. Specific objectives 5

Chapter 2. Review: Ligand binding and homology modelling of insect odorant-binding proteins
2.2. Introduction 8
2.3. Method 11
2.4. Odorant-binding proteins: classification, function and mechanisms 11
2.4.1. Classification 11
2.4.2. Function and specificity 12
2.4.3. Ligand binding and release by Lepidoptera OBPs 16
2.5. Structure prediction of OBPs by homology modelling 19
2.5.1. Steps, limitations and improvements within the homology modelling 20
2.5.2. Ligand binding affinity to homology models of OBPs: methods and interaction 29
2.6. Conclusions 32
2.7. Acknowledgements 33

Chapter 3. β-ionone as putative semiochemical suggested by ligand binding on the odorant-binding protein 1 of *Hylamorpha elegans* (Burmeister) and electroantennographic recordings
3.2. Introduction 36
3.3. Material and Methods 38
3.3.1. Volatile trapping and analysis 38
3.3.2. Insects and protein analysis 39
3.3.3. cDNA cloning and amplification by PCR 40
3.3.4. Sub-cloning in pNIC28-Bsa4 vector 40

3.3.5. Bacterial expression and purification 41
3.3.6. Fluorescence binding assays 42
3.3.7. Multiple template-based homology modelling 43
3.3.8. Molecular dynamics 43
3.3.9. Molecular docking 44
3.3.10. Electroantennographic recordings 44
3.3.11. Statistical analysis 45
3.4. Results 45
3.4.1. Sequence analysis, sub-cloning and expression of HeleOBP1 45
3.4.2. Fluorescence binding assays 47
3.4.3. Protein structure prediction and molecular docking 50
3.4.4. Electroantennographic responses 52
3.5. Discussion 54
3.6. Acknowledgements 57

Chapter 4. Behavioural responses of adult *Hylamorpha elegans* to three active molecules in both OBP-binding and electroantennography assays
4.1. Introduction 60
4.2. Materials and methods 61
4.2.1. Insect collection 61
4.2.2. Y-Tube olfactometric bioassays 61
4.2.3. Statistical analysis 62
4.3. Results 62
4.4. Discussion 64
4.5. Acknowledgements 66

Chapter 5. General discussion, concluding remarks and further perspectives
5.1. General discussion 68
5.2. Concluding remarks 75
5.3. Further perspectives 76
References 78
Appendices 103

Chapter 1

General introduction

General introduction

1.1. Introduction

Nowadays, agriculture has arisen as an important area of research where the objective is to enhance not only the performance of production, but also the defence against climate conditions and herbivores. In Chile, agriculture has been established as one of the most important economical sources for the country. Examples are wheat and beet crops, where their annual production of 1,358,129 and 1,732,032 tons, respectively (i.e. 2013/14 season) are the highest over oats, barley, maize, among others (ODEPA, 2015). However, the high production of wheat has been reduced over the years due to the attack of an endemic beetle, *Hylamorpha elegans* Burmeister (Coleoptera: Scarabaeidae). Commonly known as San Juan Verde, this beetle feeds on roots of crops during its larval stage (i.e. white grub) causing evident decomposition of plants and a subsequently occurrence of yellow spots (Aguilera *et al*., 1996). Artigas (1994) highlighted this white grub as one of the most harmful for wheat from Region del Maule to Los Lagos. It is worth mentioning that in spite of years, *H. elegans* is still considered a pest though other beetles such as *Phytoloema herrmanni*, *Sericoides* spp., *Schizochelus serratus* and *Brachysternus prasinus* also contribute to crop damages. However, from this white grub complex *H. elegans* emerge as the most important.

Insecticides and entomopathogenic fungi have been proposed as control methods for white grubs of *H. elegans* (Rodríguez *et al*., 2004; Aguilera, 2007). However, low effectiveness as been obtained. Thus, alternative strategies from chemical ecology were studied for the beetle by Quiroz *et al*. (2007), trying to identify some molecules that can mediate the sexual behaviour of *H. elegans* (i.e. sex pheromones). With the idea of finding specific attractants, the main purpose was the design of environmentally friendly control strategies. Results from the volatile profile emitted by virgin conspecific females and field behavioural assays, indicated 1,4-benzoquinone and 1,4-hydroquinone as putative sex pheromones. Moreover, 1,4-benzoquinone plus essential oils from *Nothofagus obliqua*, the most reported host plant, showed an enhanced attractive effect suggesting a role for volatiles released from the tree in the behaviour of adult *H. elegans*. Considering the above,

it is thought that males could have more skills for finding and recognizing females while they are feeding on leaves of *N. obliqua* since they have more chemosensory sensilla than females (Mutis *et al*., 2014). For instance, the participation of host plant volatiles has been suggested to play an important role for mate finding in scarab beetles (Ruther *et al*. 2000; Reinecke *et al*. 2002; Quiroz *et al*. 2007). Males are able to recognize females while they are eating through a sexual kairomone, which is proposed to be released by the attacked plant. Field observations are consistent with this behaviour in *H. elegans*.

On the other hand, since the discovery of olfactory proteins in the silk moth *Bombyx mori* (Vogt and Riddiford, 1981), an outstanding field of research has arisen due to the potential that these proteins offer for the elucidation of putative semiochemicals and the subsequent design of new control strategies. Proteins, such as olfactory receptors (ORs), olfactory co-receptors (Orco), sensory neuron membrane proteins (SNMPs), odorant-degrading enzymes (ODEs), chemosensory proteins (CSPs) and odorant-binding proteins (OBPs) comprise the army of macromolecules that govern the olfactory system at perireceptor level in insects. These proteins are located in antennae, specifically in small olfactory hairs called sensilla. There, OBPs (the most studied so far) have provided the target for semiochemical-discovery due to their main function, which is thought to be the transport of hydrophobic volatiles from olfactory pores to the vicinity of ORs (Kaissling, 2013). Due to that function, OBPs have been proposed as the target for the development of an approach called reverse chemical ecology, which means the identification of putative semiochemicals from molecular approaches, avoiding some conventional techniques (e.g. extensive field assays) (Leal, 2005). Lepidopteran OBPs, for example, are classified according to their presence in males or females and which semiochemicals can be bound in two main categories: (1) pheromone-binding proteins (PBPs) when the protein binds pheromones and (2) general odorant-binding proteins (GOBPs) when the protein binds general odorants. This previous classification has been supported since the beginning of 21st century, when the binding properties of insect OBPs started to be studied extensively. An example is how the moth *Mamestra brassicae* PBP1 (MbraPBP1) and giant moth *Anthearea polyphemus* PBP1 (ApolPBP1) were able to bind not only their pheromones, but also other compounds, such as fatty acids (Campanacci *et al*., 2001). This last, showed that OBPs are not as specific as was thought when they were discovered. From findings like the

mentioned before, reverse chemical ecology has been improving the way how putative semiochemical are identified including research at structural level.

X-ray crystallography has provided the necessary structural information to complement the study of insect OBPs. Although OBPs are small (~ 20 kDa) and usually monomers, their crystallization has been reported as difficult. However, several insect OBP structures have been solved with high resolution and deposited in database, such as Protein Data Bank (http://www.rcsb.org/pdb/home/home.do). The first crystallized OBP corresponds to the moth *B. mori* PBP1 (BmorPBP1), which showed how bombykol (the sex pheromone of *B. mori*) binds strongly to Ser56 through hydrogen bonds (Sandler *et al*., 2000). Later, Zhou *et al*. (2009) reported that other OBP from antennae of *B. mori*, BmorGOBP2, could bind bombykol through hydrogen bonds with Arg110 rather than Ser56. Probably, residues such as Arg110 and Ser56 could be involved in the specific binding of ligands explaining, for instance, how analogues (10*E*,12*Z*)-hexadecadienyl acetate and (10*E*,12*Z*)-octadecadien-1-ol bind to BmorPBP1 with higher affinity than bombykol (Hooper *et al*., 2009). These findings have shown the key role that some residues in the binding site of OBPs could play, not only for the discovery of putative semiochemical, but also for the design of potent analogues.

Despite crystallography has been sometimes successfully applied for insect OBPs, its limitations (i.e. time consuming and economical costs) result in a limited number of three-dimensional (3D) structures solved, with around 60 crystals available. Considering the enormous amount of OBP sequences that have been determined so far, the dozens of crystallized OBPs seem even lower. However, computing approaches have arisen to expand the information with OBPs that have not been crystallized yet (Venthur *et al*., 2014). Thus, homology modelling, as the most reliable method currently used, uses experimental 3D structures as templates to obtain a modelled structure of a target sequence (i.e. insect OBPs) (Bordoli and Schwede, 2012). Subsequently, to identify binding sites and amino acids that participate in ligand binding, the model can be used for binding simulations, called molecular docking. The study on the OBP1 of the scarab beetle *Holotrichia oblita* (HoblOBP1) is a clear example of integration of molecular modelling approaches with experimental ones. There, it was demonstrated the role of some amino acids (e.g. Met48,

Ile48 and Tyr111) for the binding of β-ionone, myrcene, hexyl benzoate and cinnamaldehyde using homology modelling and molecular docking along with experimental assays (Zhuang *et al*. 2014).

Considering the economical importance of *H. elegans*, its difficult control and the lack of upbringing methods for research, the scarab beetle represents a suitable methodological model for the identification of potential semiochemicals using computational reverse chemical ecology. With limited research performed on *H. elegans* reporting only two putative sex pheromones (1,4-hydroquinone and 1,4-benzoquinone), the application of computational techniques on reverse chemical ecology represents the opportunity to investigate how useful are *in silico* approaches for semiochemical discovery. Therefore, the main objective of this thesis was the identification of chemical compounds with putative semiochemical activity towards adult *H. elegans* through a computational reverse chemical ecology approach.

1.2. Hypothesis and objectives

1.2.1. Hypothesis

The presence of OBPs in several species of insects has provided alternative approaches (experimental and computational) for the design of new control strategies taking into account binding affinity of pheromones, attractants and repellents to these proteins. Considering the above, the hypothesis of this work is:

> Adults *Hylamorpha elegans* have OBPs with binding selectivity to molecules released from the main host plant, *Nothofagus obliqua*, and/or from other scarab beetles of the same subfamily.

1.2.2. General objective

- To identify chemical compounds with putative semiochemical activity towards adult *H. elegans* by computational reverse chemical ecology.

1.2.3. Specific objectives

- To determine the primary and three-dimensional structure of an OBP from antennae of *H. elegans*.
- To select high affinity ligands to an OBP of *H. elegans* by both molecular modelling and experimental techniques.
- To evaluate the biological activity of high affinity ligands by olfactometry and electroantennography.

Chapter 2

Review: Ligand binding and homology modelling of insect odorant-binding proteins

Published in Physiological Entomology (2014) 39, 183 - 198

Ligand binding and homology modelling of insect odorant-binding proteins

Herbert Venthur[1,2,3], Jing-Jiang Zhou[3], Ana Mutis[2] and Andrés Quiroz[2]

[1]Programa de Doctorado en Ciencias de Recursos Naturales, Universidad de La Frontera, Temuco, Chile

[2]Laboratorio de Química Ecológica, Departamento de Ciencias Químicas y Recursos Naturales, Universidad de La Frontera, Temuco, Chile

[3]Department of Biological Chemistry and Crop Protection, Rothamsted Research, Harpenden, Herts. AL5 2JQ, UK

Abstract

This review describes main characteristics of odorant-binding proteins (OBPs) for homology modelling and presents a summary of structure prediction studies on insect OBPs along with its steps, limitations and improvements. This technique involves a computing approach to model protein structures and is based on a comparison between a target (unknown structure) and one or more templates (experimentally determined structures). As targets for structure prediction, OBPs are thought to play a functional role for recognition, desorption, scavenging, protection and transportation of hydrophobic molecules (odorants) across an aqueous environment (lymph) to olfactory receptor neurons (ORNs) located in sensilla, main olfactory units of insect antennae. Lepidoptera pheromone binding proteins, a subgroup of OBPs are characterized by remarkable structural features, where high sequence identities (about 30%) among these OBPs and a large number of available templates could facilitate the prediction of precise homology models. About 30 studies have been performed on insect OBPs using homology modelling as a tool to predict their structures. Although some of them have assessed the ligand binding affinity using structural information and biochemistry measurements, a few have performed docking and molecular dynamic (MD) simulations as a virtual method to predict best ligands. Docking and MD simulations are discussed in the context of discovery of novel semiochemicals (super-ligands) using homology modelling to conceive further strategies in insect management.

Keywords. homology modelling, odorant-binding protein, molecular docking, ligand binding, molecular dynamic simulation, chemical ecology

Correspondence: Andrés Quiroz, Universidad de La Frontera, Temuco, Chile. Tel: +56 045 2732419; fax: +56 045 325440; e-mail: andres.quiroz@ufrontera.cl

2.2. Introduction

Host-seeking, ovipositing and mating of insects are governed mainly by odor perception through sensory organs such as antennae. These organs contain a well described olfactory system (Jacquin-Joly and Merlin, 2004; Leal, 2005), which perceives and triggers a behavioural response to chemical signals. Antennae are characterized by having specialized units called sensilla, which are comprised of one or more olfactory receptor neurons (ORNs) with their dendrites bathed in sensillum lymph. These ORNs modulate ion potentials across their plasma membrane thus participate in the transduction of chemical signals into electrical signals. Through key proteins in the ORNs, the olfactory system carries out a dynamic process of odor perception and discrimination. The main olfactory proteins include odorant-binding proteins (OBPs), chemosensory proteins (CSPs), odorant-degrading enzymes (ODEs), sensory neuron membrane proteins (SNMPs), olfactory co-receptors (Orco) (formerly OR83b receptors) and olfactory receptors (ORs). It is thought that volatile odorant molecules such as pheromones are transported by OBPs or CSPs across sensillum lymph to ORs located in dendritic membrane of ORNs in sensilla (Vogt and Riddiford, 1981; Klein, 1987; Maida *et al*., 1993; Jacquin-Joly *et al*., 2001; Zhou *et al*., 2006) and act as the first filter of olfactory information (Leal, 2003). Since the first OBP was identified by Vogt and Riddiford (1981), many more have been characterized. These proteins are soluble with a molecular weight of 13-16 kDa and abundant in sensillum lymph. Some evidences support that OBPs, particularly members of pheromone binding protein (PBP) subgroup, can selectively bind to odorants, such as the pheromone (*E*,*Z*)-10,12-hexadecadienol (bombykol) as a specific ligand to the *Bombyx mori* PBP (BmorPBP1) (Sandler *et al*., 2000) or the pheromone component (*E*,*Z*)-6,11-

hexadecadienal to *Antheraea polyphemus* PBP (ApolPBP1). However, it has been shown that they also bind to a wide range of odorant molecules as reported by Honson *et al*. (2005) and Zhou (2010). These papers summarized all the early work on OBP/ligand interactions determined experimentally and the various assays that are still used today to study the ligand binding of OBPs. More studies are necessary to fully understand the selectivity of OBPs during the odor recognition, although some important parameters of odorant molecules are known, such as length of carbon chain, functional groups, location of methyl ends and specific position of unsaturation as well as hydrophobic characteristics (Zhou, 2010, Yu and Plettner, 2013). These provide useful information for structural homology modelling.

Ligands (e.g. odorant molecules) bind to proteins through physicochemical forces, which are non-covalent interactions, such as hydrogen bonds, ionic bonds, van der Waals (vdW) interactions, among others. Different amino acids contribute to the interactions according to their structures and positions in the ligand binding site of proteins. The most common approach to understand these interactions in the potency and specificity of ligand binding is to obtain the three dimensional (3D) structure of the binding proteins bound with ligands (Ravna and Sylte, 2012) as has been demonstrated in the search of the binding sites and their specificities for the ligands of ABC transporters (Gajendrarao *et al*., 2010). X-ray crystallography and nuclear magnetic resonance (NMR) are optimal techniques so far for the analysis of ligand-protein interactions based on structural information due to their high precision (high resolution), but they are expensive and time consuming. Moreover, it is not always possible to obtain protein crystals in the presence of testing ligands. The computational approaches can contribute as complementary tools for 3D structure prediction (Paas *et al*., 2000). One of them is comparative or homology modelling, a reliable method currently available on the assumption that proteins that fold within a structural family are more functionally conserved than the primary sequences. This structural prediction approach is based on evolutionary relationships between a target protein and template proteins for which a 3D structure, determined by crystallographic or NMR experiments, exists.

OBPs are present at high concentrations (up to 10 mM) in sensillum lymph (Klein, 1987). There is an increased interest on OBPs and other olfactory proteins of invertebrates in the hope to prevent human diseases transmitted by insect vectors and to provide alternative pest control strategies from the use of pesticides by manipulating insect pest behaviours (Leite *et al*., 2009; Lagarde *et al*., 2011). An OBP-based screening of putative bioactive chemicals with homology modelling can serve as a good complement to study ligand-OBP interactions and research for 'super-ligands' together with robust biological assays for insect behavioural manipulation. Leal (2005) proposed a reverse chemical ecology concept, which utilizes a protein-based screening of attractants, pheromones and repellents through their binding affinity to OBPs as an interesting approach of using these chemicals in pest management. An advantage of this approach is that insect OBPs are structurally different from vertebrate OBPs and have no sequence homology to vertebrate OBPs. Insect OBPs are commonly located in sensilla of antennae. However, vertebrate OBPs are located in nose that is involved in both olfaction and respiration (Pelosi and Maida, 1990). Insect OBPs mainly have α-helices domains and vertebrate OBPs have only β-strands and a short α-helix. Nevertheless, both insect and vertebrate OBPs have conserved disulfide bridges and are small soluble proteins. Ligand binding assays indicate that both insect and vertebrate OBPs bind to a wide range of volatiles with dissociation constants of either micro or milli molar concentration (Tegoni *et al*., 2000; Pelosi, 2001; Briand *et al*., 2002; Löbel *et al*., 2002; Nespoulous *et al*., 2004; Grolli *et al*., 2006; Wei *et al*., 2008; Brimau *et al*., 2010). However, electrophysiological recording showed the evidence of selective binding of BmorPBPs and ApolPBPs to pheromones (Pophof, 2004). Moreover, selective binding of ApolPBP1 at pH 6.5 and pH 4.5 was reported by Katre *et al*. (2009) at nano molar range. Although there are many studies that have used homology modelling in structure-based drug discovery, only a few have been performed on insects OBPs. Due to the high similarity of OBPs across Lepidoptera species and a large number of experimentally determined 3D structures (Tegoni *et al*., 2004; Pelosi *et al*., 2006; Damberger *et al*., 2007; Zhou, 2010), these proteins could be used as good targets for homology modelling and molecular dynamic simulations to obtain best structural models in terms of energy. Thus, the objective of this review is to present current knowledge of

computer-assisted protein modelling by homology, focusing on Lepidoptera OBPs to predict the ligand binding affinity.

2.3. Method

We review the classification and the functions of insect OBPs, and the mechanisms of their binding and releasing air-born ligands to the olfactory receptors in the chemo-sensilla on insect antennae. We further describe the steps and limitations of OBP structure prediction by homology modelling and propose some improvements for the prediction. We finally summarize ligand binding affinity of homology models including methods and interactions.

2.4. Odorant-binding proteins: classification, function and mechanisms

2.4.1. Classification

Lepidoptera OBPs can be divided into three classes based on their amino acid sequences (Figure 1) and their structural characteristics. Thus, OBPs that bind to pheromones are referred as pheromone binding proteins (PBPs), e.g. BmorPBP1 (Krieger *et al*., 1996). General odorant-binding proteins (GOBPs) such as GOBP1 and GOBP2 are present in both female and male of tobacco hawk moth *M*and*uca sexta* (Vogt *et al*., 1991) and proposed to bind to general odorants. Less described are antennal-binding proteins X (ABPx), which were reported first in *B. mori* (Krieger *et al*., 1996) and their ligand binding have only recently reported (He *et al*., 2010). As shown in Figure 1, an alternative classification for OBPs consists of Classic OBPs; Plus-C OBPs; Minus-C OBPs and Atypical OBPs (Hekmat-Scafe *et al*., 2002 and Zhou, 2004). The main sequence differences are the number of cysteine (Cys) residues and their conservation. According to this, Classic OBPs have 6 conserved Cys residues at specific positions and include PBPs, GOBPs and ABPx, but Plus-C OBPs have 2 additional Cys and a conserved proline (Hekmat-Scafe *et al*., 2002; Zhou *et al*., 2004). Minus-C OBPs have less than 6 Cys and Atypical OBPs are characterized by have 6 Cys residues as Classical OBPs but with additional ones in the C-terminal region (Xu *et al*., 2003). For OBPs with the 6 conserved

Cys residues, there is also a classification according to its chain and C-terminus length (Tegoni *et al*., 2004): 1) long-chain OBPs with ~140 amino acids (e.g. moth's OBPs of *B. mori* and *A. polyphemus*); 2) medium-chain OBPs with ~120 amino acids (e.g. some OBPs of the mosquito *Anopheles gambiae* and the bee *A. mellifera*) and 3) short-chain OBPs with ~100 amino acids (e.g. OBPs from the cockroach *Leucophaea maderae*). More recently, a novel subclass of OBPs according to its C-termini and 3D structure was proposed by Lagarde *et al*. (2011), called C8 OBP class. Their findings for the *A. gambiae* OBP7 (AgamOBP7) suggested a more evolved protein due to an increase in structural complexity with 8 Cys residues, 4 disulfide bridges and a C-terminus slightly longer than Classic OBPs. Moreover, it has been proposed that AgamOBP7 had evolved from Classic OBPs, which served as a basal group of OBPs according to phylogenetic analyses (Vieira and Rozas, 2011). We focus on Classic OBPs, because they have well conserved characteristics such as 6 Cys residues, 3 disulfide bridges and an average molecular weight of ~14 kDa. Furthermore, most of well-studied OBPs such as PBPs and GOBPs of Lepidoptera species and the odorant binding protein LUSH for the pheromone 11-*Z*-vaccenyl acetate of *Drosophila melanogaster* are Classic OBPs.

2.4.2. Function and specificity

Insects' perception of their environment such as host plants, prey and potential mates are guided by chemical signals, which are called semiochemicals. If a semiochemical is a small, volatile and hydrophobic molecule they cannot easily cross a polar environment such as insect sensillum lymph. It is thought that OBPs are the key component to solubilize and transport these molecules across the lymph as far as reaching ORs. It is clear now at least for hydrophobic ligands such as Lepidoptera sex pheromones that OBPs play an important role in ligand capturing and transporting to achieve the ligand-OR interaction, and contribute to the subsequent signal transduction (Kaissling, 2013). However, experimental evidence suggests three outstanding functions for OBPs and may all play dynamic roles in ligand selectivity: (1) Ligand scavenger responsible of protection of ORs from saturation, which was proposed using a kinetic model for *Lymantria dispar* PBP2 (LdisPBP2) by Gong *et al*. (2009); (2) Ligand desorption from the cuticular wax layer of olfactory pores to

the lymph as it was suggested for LdisPBP1 and LdisPBP2 (Kowcun *et al*., 2001) and (3) Ligand recognition function proposed for *D. melanogaster* OBP (LUSH) acting as an activator of ORNs tuned to pheromones (Laughlin *et al*., 2008). In contrast, Gomez-Diaz *et al*. (2013) reported that a high concentration of the sex pheromone (*Z*)-11-octadecenyl acetate (*cis*-vaccenyl acetate) of *D. melanogaster* activates ORs in the absence of LUSH, indicating that LUSH could not have a direct role as the activator of ORNs. It has also been proposed that OBPs protect odorant molecules from odorant degrading enzymes (ODEs) for degradation during its journey to ORs (Ishida and Leal, 2002; Leal, 2005, Leal, 2013).

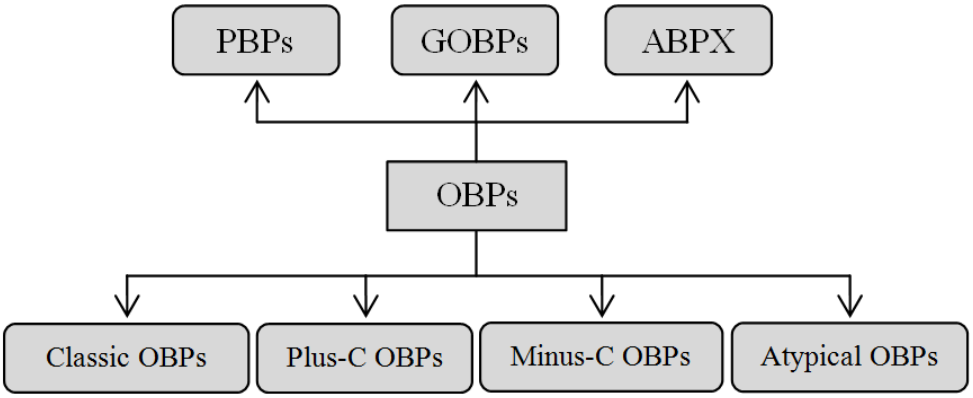


Figure 2.1. Classification scheme of insect odorant binding proteins (OBPs) according to their primary protein sequence. ABPX, antennal-binding protein X; GOBP, general odorant-binding protein; PBP, pheromone-binding protein.

When semiochemicals are pheromones, there is an intraspecific interaction between insects through associated PBPs. For example, male moth *A. polyphemus* has three PBPs, which interact with its sex pheromone components: (*E*6,*Z*11)-hexadecadienal; (*E*4,*Z*9)-tetradecadienyl-1-acetate and (*E*6,*Z*11)-hexadecadienyl-1-acetate (Bette *et al*., 2002; Maida *et al*., 2003). Kairomones (benefits to receivers) and allomones (benefits to sender) are involved in interspecific interactions, in which GOBPs may be associated. An example is the GOBP2 of the meadow moth *Loxostege sticticalis* (LstiGOBP2), which was shown to have high affinities to plant volatiles from essential oils, such as (*E*)-2-hexenal and (*Z*)-3-hexen-1-ol, and to the pheromone component (*E*)-11-tetradecen-1-yl acetate (Yin *et al*., 2012). In *B. mori*, the pheromone bombykol must reach and activate ORs, which is mediated by BmorPBP1. Große-Wilde *et al*. (2006) have studied the specificity of

BmorPBP1 to bombykol and its analogue bombykal by measuring the activation of ORs with the BmorPBP1-pheromone complex. They demonstrated that an OR response could be achieved with bombykol solubilized with BmorPBP1 as well as with the organic solvent dimethyl sulfoxide (DMSO). For bombykal BmorPBP1 could not substitute DMSO, indicating that BmorPBP1 has a higher specificity to bombykol. They postulated that like *A. pernyi* and *A. polyphemus* (Maida *et al*., 2000), a specialized PBP for bombykal could exist in *B. mori*. The interaction with ORs and other downstream events are not well understood yet. Apart from Lepidoptera PBPs, the ligand specificity of insect OBPs has so far not been exclusively demonstrated. Nevertheless, the sex pheromone specificity of Lepidoptera PBPs has provided an interesting opportunity to study protein-ligand interactions by using OBPs as attractive targets for homology modelling and ligand screening. The observed OBP specificity allows the interaction between a ligand and a protein, forming a ligand-OBP complex as an encapsulation, a term proposed by Leal (2005). The competitive binding assays using *N*-phenyl-1-naphthylamine (1-NPN) as fluorescence probe have been used to study the ligand binding affinity to OBPs (Lescop *et al*., 2009; Zhou, 2010). Only few studies have provided the binding affinities in connection with structural data. Hydrogen bonds seem to have an important role as specific interactions for the recognition of some ligands as it was reported for LUSH and its alcohol-binding site (Kruse *et al*., 2003) and other Lepidoptera OBPs (Zhou *et al*., 2009). The first attempt was made by the study of Sandler *et al*. (2000), in which BmorPBP1 was crystallized in complex with bombykol. The authors reported that Ser56 formed a hydrogen bond with the hydroxyl of bombykol, which plays a role in the specificity of the PBP. However, a more deep study that was performed on bombykol-BmorPBP complex by Klusak *et al*. (2003) using *ab initio* methods suggested that not only hydrogen bonds but also cation-π and π-π interactions have an important role in the bombykol binding. The a*b initio* calculations suggest that bombykol could adopt two conformations (A and B). A single hydrogen bond with Ser56 was formed with conformation A, while two hydrogen bonds were formed with conformation B involving Ser56 and Met61. BmorPBP1 has been shown to bind, when measured with Chip-assisted high-throughput ESI-MS analysis, to analogues of bombykol, where the compounds with different chain lengths such as (10*E*,12*Z*)-hexadecadienyl acetate and (10*E*,12*Z*)-octadecadien-1-ol bind to the protein with a higher affinity than

bombykol (Hooper *et al*., 2009; He *et al*., 2010). Likewise, Campanacci *et al*. (2001) reported that PBP1 of *Mamestra brassicae* (MbraPBP1) and ApolPBP1 cannot discriminate pheromones of other compounds. By competitive binding assays using 1-aminoanthracene (AMA), fatty acids showed significant affinity. Zhou *et al*. (2009) and He *et al*. (2010) demonstrated that BmorPBP1 binds well to both bombykol and bombykal, while a GOBP (BmorGOBP2) binds differently to these compounds. The authors suggested that bombykol binding to BmorGOBP2 involves hydrogen bonding to Arg110 rather than to Ser56 as found for BmorPBP1. In addition, the hydroxyl group of bombykol forms additional hydrogen bond with Glu98 via a water molecule as it has been predicted through molecular dynamic simulations by Gräter *et al*. (2006a; 2006b), whilst bombykal cannot form such additional hydrogen bond. This explains the ligand discrimination by BmorGOBP2 at structural level. Interestingly Zhou *et al*. (2009) also demonstrated by crystallography there is no conformational change among the BmorGOBP2 structures bound with sex pheromone components and their analogues, which is consistent with a recent study on *Drosophila* LUSH (Gomez-Diaz *et al*., 2013). A crystallized structure of an OBP of *A. gambiae* (AgamOBP1) showed numerous contacts with *N*,*N*-diethyl-3-methylbenzamide (DEET) at its binding site (e.g. van der Waals interactions). One hydrogen bond was formed via a water molecule with Trp114, Cys95 and Gly92 (Tsitsanou *et al*., 2012). Although DEET seems to have a specific interaction with the OBP, it has been recently reported that eugenyl acetate could be a better repellent than DEET due to a better affinity with the binding site of AgamOBP1 by molecular modelling (Affonso *et al*., 2013). However, binding assays and robust behaviour bioassays are necessary to corroborate those findings.

Although OBPs are divergent across insect species, even within same species (Pelosi *et al*., 2006), general structural characteristics of OBPs are conserved, such as 6 α-helical domains and 3 disulfide bridges (from 6 Cys residues). These features are well conserved in the GOBP of honeybee *A. mellifera* (ASP2). However, ASP2 showed a broad specificity for ligands compared to PBPs (Lescop *et al*., 2009). Structurally different compounds, such as 2-isobutyl-3-methoxypyrazine, isoamyl acetate, 1,8-cineol and 2-heptanone showed significant affinity for ASP2 by weak and non-specific interactions. Therefore, more evidences and robust bioassay techniques are necessary to establish the discriminatory capability of insect OBPs.

2.4.3. Ligand binding and release by Lepidoptera OBPs

It has been proposed that, once bombykol-BmorPBP1 complex reaches the proximity of dendritic membrane, a pH-induced conformational change occurs to release the ligand. Thus, studies have been performed with crystal structures of OBPs to determine the key features involved in the pH-dependent change. For example, for the moth *B. mori*, it has been suggested that its pheromone is released to ORs through a conformational change of BmorPBP1 by acid pH near the ORNs membrane. A main change is denoted by the observation of shifting the C-terminus, as a characteristic feature of OBPs with a long C-terminus in Lepidoptera PBPs, from an extended form to α-helix by this pH change (from 6.5 to 4.5) (Horst *et al*., 2001; Lautenschlager *et al*., 2005; Leal, 2005). Thus, at acid pH due to membrane negative charges near the surface of dendritic membrane, there is a release of the pheromone (bombykol) because the C-terminus replaces bombykol and occupies the binding site of BmorPBP1 as α-helix. However, the pH of antennal lymph has not been measured and the nature of any pH homogeneities in the lymph space is also not known. Therefore, pH-dependent conformational changes should be considered as an interesting phenomenon on these types of proteins due to no clear function has been attributed yet. It seems that mechanisms of ligand binding and release are not always dependent on the C-terminus of OBPs. Zhou *et al*. (2009) reported a significant difference in the C-terminal conformation between BmorPBP1 and BmorGOBP2. Whereas BmorPBP1 have an extended C-terminus occupying its binding site, BmorGOBP2 has a longer C-terminus, which forms a α-helix and does not occupy the binding site nor participate in the ligand binding. The authors also demonstrated no significant conformational changes among six different ligand-BmorGOBP2 complexes. Furthermore, Xu *et al*. (2011) studied the pheromone binding to the PBP1 of the moth *Amyelois transitella* (AtraPBP1) and how C-terminus controls this binding according to pH. The pheromone (*Z*11,*Z*13)-hexadecadienal binds more strongly to the pheromone at neutral pH than at acid pH. However, the deletion of the C-terminus of AtraPBP1 increased the pheromone binding affinity by 100-fold at pH 5.0 and 1.5-fold at pH 7.0. Probably, there is an occupation of the binding site of AtraPBP1 by C-terminus at neutral pH as well as acid

pH, which decreases the binding affinity of (*Z*11,*Z*13)-hexadecadienal. Furthermore, three wild-type PBPs (PBP1, PBP2 and PBP3) of two sibling species *Helicoverpa armigera* and *H. assulta* showed a decrease in the binding affinity at acid pH. The mutant PBPs without C-terminal segment had a similar binding affinity at the same pH (Guo *et al*., 2012). Zubkov *et al*. (2005) found that the C-terminus of ApolPBP1 does not have a major role for the displacement of ligands. Instead, there is a reorientation of helices α1, α3 and α4 at acid pH, which causes protonation of histidine residues (His69, His70 and His95) inside the binding site, leading to an opening of the binding site, and the pheromone is released near the dendritic membrane. Pesenti *et al*. (2009) reported that the *A. mellifera* PBP (ASP1) had a contradictory conformational change from that observed in BmorPBP1. Thus, ligands bind to this protein at low pH (4.0) and the release occurs at pH 7.0. However, their findings indicated that the binding is also possible at neutral pH, therefore, the authors proposed an uncommon conformational change. At pH 7.0, ASP1 can form dimers through the absence of the C-terminus from the core of the protein and the participation of the N-terminus to form a more stable dimer structure. This dimeric form binds ligands but with lower affinity than the acidic monomer. There is another more direct ligand release mechanism for *L. maderae* PBP (LmaPBP) suggested by Lartigue *et al*. (2003). This mechanism, in contrast to BmorPBP1, is based on the absence of a helix inside of the PBP and the presence of a significant amount of hydrophilic residues in the binding site.

Kinetics studies have been performed for LdisPBP2 by fluorescence binding assays and tryptophan anisotropy measurements, obtaining the association in timescale and multimerization, respectively (Gong *et al*., 2009). The results indicated that LdisPBP2 binds (+)- and (-)-disparlure in a rapid step and also in a slow step. The slow step could be more important to obtain an active ligand for ORs due to an internal binding. On the contrary, the rapid step seems to be related with a binding to an external site. Moreover, the authors reported that over hours, LdisPBP2 increases its presence as dimer, which has smaller binding capacity compared with monomer. Subsequently, the same authors proposed a more detailed binding mechanism. It seems that two steps are necessary to obtain a final complex. First, LdisPBPs and the ligand form an intermediate complex through diffusion-controlled collision. The second and final step involves the relocation of the ligand to a different binding site from external to an internal one (Gong *et al*., 2010). In

this review, we have proposed a schematic representation that summarizes the most outstanding research on Lepidoptera OBPs and its ligand binding and release mechanisms including highlighted mechanisms for ASP1 and LmaPBP (Figure 2).

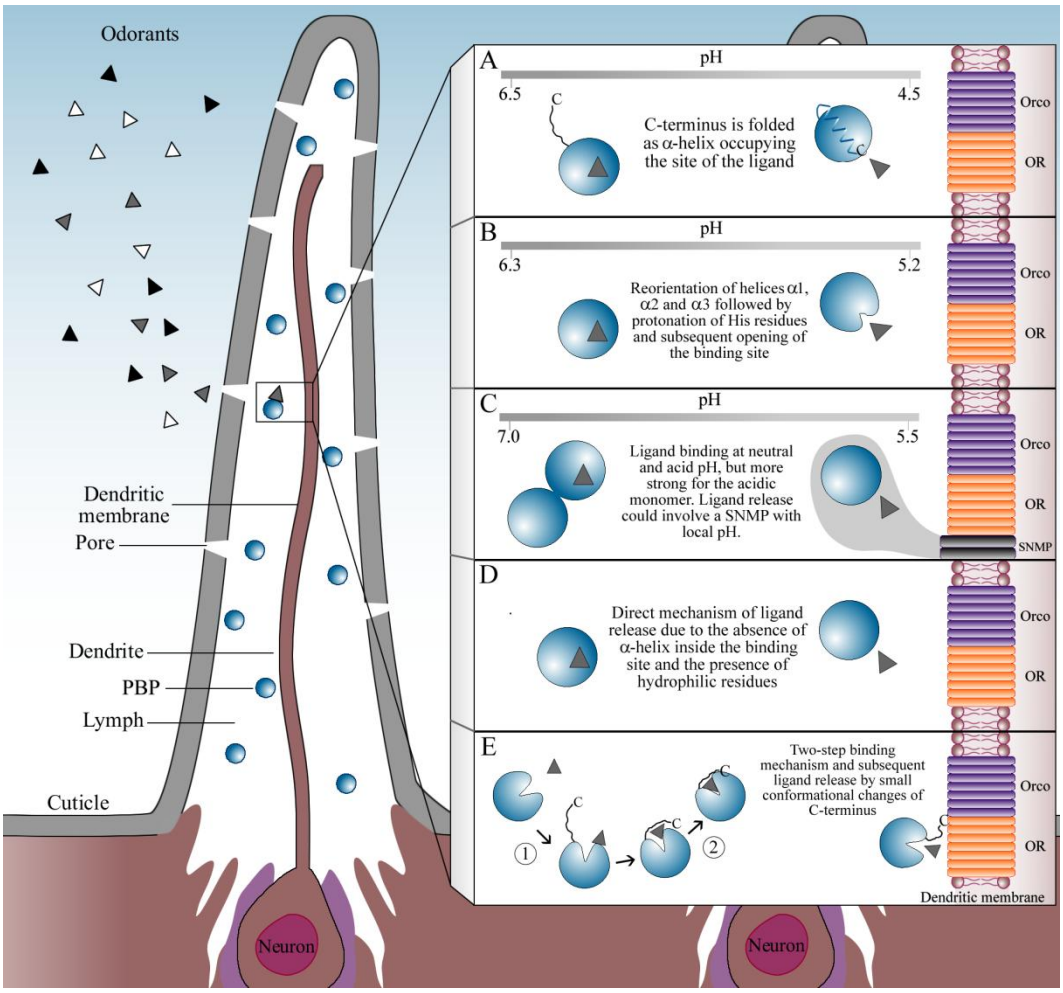


Figure 2.2. Schematic representation of the main mechanisms of ligand binding and release described for PBPs. Background section represents a commonly multipore sensillum presented in moths, such as *B. mori* and *A. polyphemus*. From the external environment, different odorants (white, grey and black triangles) can enter to sensilla by diffusion through cuticle pores. PBPs (blue spheres) are secreted by auxiliary cells and located in the sensillium lymph. These proteins are proposed as the first filter of selection and discrimination of odorants, which are represented by the ligand-PBP complex (grey triangle-sphere) in the lymph. (A) Ligand binding and release mechanism proposed for BmorPBP (Horst *et al*., 2001; Lautenschlager *et al*., 2005; Leal, 2005). (B) Ligand binding and release mechanism proposed for ApolPBP1 (Zubkov *et al*.2005). (C) Uncommon

ligand binding and release mechanism proposed for *A. mellifera* ASP1 (Pesenti *et al*.2009). (D) Ligand binding and release mechanism proposed for LmaPBP (Lartigue *et al*., 2003). (E) Ligand binding and release mechanism proposed for LdisPBPs (Gong *et al*., 2009; Gong *et al*., 2010). Inactivation or degradation (by ODEs) is not represented in this scheme because is not part of the main discussion of this revision.

2.5. Structure prediction of OBPs by homology modelling

About four years ago, Zhou (2010) mentioned that 47 crystal structures of insect OBPs were deposited in Protein Data Bank (PDB) (http://www.rcsb.org/pdb). A search with "odorant binding protein" as keywords returned the structures of OBPs in 8 insect species, despite of more than hundreds OBPs reported in the enormous number of insect species. To date, more than 60 crystal structures have been determined by X-ray crystallography. Only 13 of them have a high resolution (<1.5 Å), 45 have a resolution of 1.5 – 2.5 Å and 3 have a resolution of 2.5 – 3.0 Å. Thus, *A. mellifera*, *D. melanogaster*, *A. gambiae* and *B. mori* have the greatest number of crystal structures including OBPs and PBPs in some cases. These structures were solved under different conditions, such as ligand-OBP complex, apo-OBP form and at different pHs. The PBPs from moths frequently have 7 α-helices and 3 conserved disulfide bridges. The high sequence identity provides a similar global fold among these OBPs, which is crucial for homology modelling, and the structural studies of these OBPs have provided some outstanding information for computer-assisted modelling of these proteins. For instance, early studies on homology modelling with OBPs were carried out by Honson and Plettner (2006) to determine the three-dimensional arrangement of disulfide bridges on *L. dispar* PBPs (LmarPBP1 and LmarPBP2). Based on the crystal structure of BmorPBP1, the homology models revealed that the most exposed and accessible disulfide bridge was C2-C5 (i.e. Cys50-Cys108), which was the most easily reduced disulfide bridge by cyanylation using tris(2-carboxyethyl)phosphine (TCEP), 2-mercaptoethanol and dithiothreitol (DTT). The structural models of these proteins suggest that the easy reduction of C2-C5 is due to a steric property with Cys side chains more exposed. Moreover, the conformation of C2-C5 could be strongly influenced by electronic effects from a nearby aspartic acid (Asp106), which stabilized C5 through a proton transfer.

2.5.1. Steps, limitations and improvements within the homology modelling

Knowledge of 3D protein structures allows understanding molecular mechanisms, evolutionary relationships as well as detail information about binding sites and ligand-protein interactions. The prediction of 3D protein structure from their amino acid sequence can be carried out by homology or comparative modelling from the protein structures determined previously by X-ray crystallography or NMR spectroscopy. This has provided an important progress in the understanding of protein functions for more than four decades with studies started by Levinthal (1966). Homology modelling is based on an evolutionary relationship between target and template proteins, sharing a degree of structural similarity. The proteins are called "target" when their 3D structure is unknown and "template" for proteins with their 3D structures experimentally determined. Thus, this protein structural modelling performs a comparison between the amino acid sequence of target and template(s). Templates are stored and available with a code of four digits called PDB code in the Protein Data Bank (http://www.rcsb.org/pdb/home/home.do). An example is the alignment between *Acyrthosiphon pisum* OBP3 target and *L. maderae* PBP with code 1ORG chain A as template for aphid OBPs modelling (Qiao *et al*., 2009). Similarly, other authors have used the same template to predict the 3D structure of aphids OBPs of *A. pisum* and *Sitobion avenae* (Sun *et al*., 2011; Sun *et al*., 2012; Zhong *et al*., 2012) and an OBP of the plant bug *Adelphocoris lineolatus* (Wang *et al*., 2013). Moreover, a template can be used to model not only a full-length but also a part of target, which is possible according to Chang and Swaan (2006). For example, Paramasivan *et al*. (2007) used two templates for modelling *Culex quinquefasciatus* OBP2. The first 15 residues were modeled with 1DI1 template of *Penicillium roqueforti* synthase and the rest with 1R5R of *A. mellifera* ASP1. Such templates were selected based on the folding of the target by a fold recognition server called GenTHREADER (http://bioinf.cs.ucl.ac.uk/psipred//?program=psipred). Another modelling strategy is to model regions of proteins independently as loops or transmembrane domains with different templates for each region especially when proteins have multiple domains. Swiss-Model (Arnold *et al*., 2006) and Modeller (Eswar *et al*., 2006) are popular and freely available software for modelling process, which offer a

friendly platform for the protein homology modelling. Swiss-Model is suitable for beginners as an automated and online server and does not need to be downloaded and installed. Instead, this server offers the building of homology models with a workspace through web connection (Arnold *et al*., 2006). On the other hand, Modeller must be downloaded and installed. Although Modeller automatically calculates a model, it must be provided with an alignment between target and template as well as the user scripts as commands. Once the model has been built, there is software to visualize the model, such as PyMOL, RasMol, Chimera and Swiss Pdb-Viewer, among others. Homology modelling studies of insect OBPs so far are listed in Table 1 including templates and software used. Apart from homology modelling, there are other types of computing approaches in protein structure prediction according to several reviews on the subject (Sánchez *et al*., 2000; Schwede *et al*., 2007; Ravna and Sylte, 2012). These are threading methods, *ab initio* and integrative or hybrid methods. However, homology modelling is the most precise and reliable method currently used (Bordoli and Schwede, 2012; Ravna and Sylte, 2012). There are numerous reviews about homology modelling with a similar content structure, highlighting the steps, precision, limitations for this modelling and an endless list of modelling software's available. Therefore, in the next subsection we summarize important concepts in homology modelling, and focus on ligand affinities of insect OBPs.

Protein structure modelling based on homology between target and template consists of four main steps (Figure 3): 1) template identification; 2) target-template alignment; 3) model building and 4) model refinement and validation. Template identification is the first step to obtain a homology model. Such identification is usually carries out with basic local alignment search tool (BLAST). This tool provides regions of similarity and an identity percentage of nucleotide or protein sequences between target and templates. A template must have some characteristics according to Orry and Abagyan (2012) to determine a ligand-protein interaction, such as high resolution, sharing regions or substructures with target and be bound as holo-protein.

Table 2.1. Summary of homology modelling studies to determine odorant-binding protein structures of insects.

OBP	Species	Template(s)	PDB code	Software	Reference
OBP	*L. migratoria*	*A. mellifera* PBP	3BFH	Discovery Studio 2.0	Jiang *et al*. 2009
OBP2, OBP3 and OBP7	*S. avenae*	*L. maderae* PBP	1ORG	Swiss-Model	Zhong *et al*. 2012
PBP1	*Spodoptera litura*	*A. Polyphemus* and *B. mori* PBP	1IS8, 1QWV	Modeller	Liu *et al*. 2012
OBP	*A. pisum*	*L. maderae* PBP	1ORG	Swiss-Model	Qiao *et al*. 2009
OBP4	*A. gambiae*	*A. gambiae* OBP1	2ERB	Swiss-Model	Qiao *et al*. 2011
OBP1, OBP2 and OBP3	*Helicoverpa armigera*	*B. mori* PBP	1DQE	Discovery Studio 2.0	Zhang *et al*. 2012
OBP5	*H. armigera*	*A. aegypti* OBP1	3K1E	Discovery Studio 2.0	Zhang *et al*. 2012
OBP2 and OBP3	*L. migratoria*	*A. gambiae* OBP1	2ERB, 1DQE	Swiss-Model	Yu *et al*. 2009
OBP	*Phormia regina*	*B. mori* PBP		Insight II	Tsuchihara *et al*. 2005
OBP3	*A. pisum*	*L. maderae* PBP	1ORG	Swiss-Model	Sun *et al*. 2011
OBP10	*H. armigera*, *H. assulta*	*C. quinquefasciatus* OBP	3OGN	Swiss-Model	Sun *et al*. 2012
OBP3 and OBP7	*A. pisum*	*L. maderae* PBP and *D. melanogaster* LUSH	1ORG, 3B6X	Swiss-Model	Sun *et al*. 2012
OBP15 and OBP21	*A. mellifera*	*A. mellifera* OBP14	3RZS	Modeller	Spinelli *et al*. 2012
OBP	*D. melanogaster*	*A. gambiae* OBP1	2ERB	Swiss-Model	Sánchez-Gracia and Rozas, 2008
OBP1	*A. gambiae*	*A. gambiae* OBP1	2ERB, 3N7H	Swiss-Model	Rusconi *et al*., 2012
OBP2	*C. quinquefasciatus*	*A. mellifera* ASP1 and *P. roqueforti* Synthase	1R5R, 1DI1	Modeller	Paramasivan *et al*., 2007
OBP1 and OBP2	*Scleroderma guani*	*Pyrococcus horikoshii* PH1010 and *A. mellifera* ASP1	3D76, 3BJH	Swiss-Model	Li *et al*., 2011
PBP1	*Lymantria dispar*	*B. mori* PBP	1DQE	Swiss-Model	Honson *et al*., 2003
PBP1 and PBP2	*L. dispar*	*B. mori* PBP	1DQE	Swiss-Model	Honson and Plettner, 2006
OBP1	*Adelphocoris lineolatus*	*B. mori* PBP	1DQE	Discovery Studio 2.0	Gu *et al*., 2011

Table 2.1. Continued.

OBP	Species	Template(s)	PDB code	Software	Reference
PBP1	*M. brassicae*	*B. mori* PBP	1DQE	TURBO-FRODO	Campanacci *et al*. 2001
OBP1	*A. gambiae*	*A. gambiae* OBP1	2ERB	-	Biessmann *et al*., 2010
OBP1	*L. migratoria*	*B. mori* PBP	1DQE	Insight/Homology	Ban *et al*., 2003
PBP3	*Ostrinia* spp.	*B. mori* GOBP2, *A. polyphemus* PBP and *B. mori* PBP	2WCJ, 1QWV, 1DQE	Modeller	Allen and Wanner, 2011
PBP1, PBP2 and PBP3	*H. armigera* and *H. assulta*	*B. mori* PBPs	1DQE, 2FJY	Swiss-Model	Guo *et al*., 2012
OBP5	*A. lineolatus*	*D. melanogaster* LUSH	1OOI	Discovery Studio 2.0	Wang *et al*., 2013
PBP1 and PBP2	*L. dispar*	*A. polyphemus* PBP	1QWV	Discovery Studio 3.0	Yu *et al*., 2012
PBP1 and PBP2	*L. dispar*	*B. mori* GOBP2	2WCJ	Swiss-Model	Yu and Plettner, 2013
OBP2	*A. cerana*	*A. mellifera* OBP2	1TUJ	Swiss-Model	Li *et al*., 2013
OBP7	*H. armigera*	*A. gambiae* OBP1	3N7H	Swiss-Model	Sun *et al*., 2013
OBP37 and OBP39	*A. albopictus*	*C. quinquefasciatus OBP1* and *A. aegypti OBP1*	2L2C, 3K1E	Swiss-Model	Deng *et al*., 2013
OBP3 and OBP4	*Holotrichia oblita*	*A. gambiae* OBP20	3VB1	Swiss-Model	Wang *et al*., 2013
OBP1	*H. oblita*	*A. gambiae* OBP1	2ERB	Swiss-Model	Zhuang *et al*., 2014
ASP2	*A. cerana*	-	-	I-TASSER	Lu *et al*., 2014

The sequence identity percentages are important indicators for establishing the best template. Thus, 30% or more is considered as good as to obtain a suitable and precise model. If the sequence identity is less than 30%, the model quality decreases (Schwede *et al*., 2007). Based on the above, the sequence identity of the targets LdisPBP1 and LdisPBP2 with the template BmorPBP (1DQE) was 61% and 48%, respectively (Honson *et al*., 2003), which provided a good homology model for both PBPs. However, this study did not provide an evaluation on the quality of the predicted models. This is commonly performed with ProCheck (http://www.ebi.ac.uk/thornton-srv/software/PROCHECK/), which evaluates the stereochemical quality of the structure predicted. A special case was reported by Gu *et al*. (2011), with a reliable homology model for the lucerne plant bug *A. lineolatus* OBP (AlinOBP1) based on a low sequence identity of 16.8% with the template bombykol-BmorPBP1 complex (1DQE). The homology recognition program called FUGUE, which comprises a sequence-structure alignment to find common folds even with low sequence similarity between target and template (Shi *et al*., 2001), was used. Thus, a z-score was related with a 99% of confidence level, although the discrete optimized protein energy (DOPE) score is not included. The DOPE score is commonly used for homology models obtained from Modeller, which assesses the quality of the predicted structure based on statistical potentials. The lowest DOPE scores are related to the best models predicted. Homology models with more than 50% sequence identity have a resolution of ~1 Å Cα RMSD from templates, which is considered as a precise model (Ginalski, 2006). After the template is identified it is necessary to produce a good alignment between the target and the template. The alignment can be classified into pairwise sequence alignment (PSA) and multiple sequence alignment (MSA). The PSAs are carried out by BLAST and the MSAs are carried out by programs such as ClustalW (Thompson *et al*., 1994), MUSCLE (Edgar, 2004) and T-Coffee (Notredame *et al*., 2000). These programs are online bioinformatic tools, where ClustalW and MUSCLE have higher speed in running time compared to T-Coffee for a large number of sequences according to Hang (2008). The third step of homology modelling is the construction of main structures including core modelling as a structurally conserved region in proteins, loop modelling and side chains and backbone modelling (Leach, 2001; Ravna and Sylte, 2012). For structure model building, there are two main approaches highlighted by Leach (2001), Schwede *et al*.

(2007) and Ravna and Sylte (2012). The first approach is called rigid-body method or rigid fragment assembly, which consists of the construction of an initial model of a target from the structurally conserved core region of a template protein. Satisfaction of spatial restraints is the second approach for homology model building. This is an optimization strategy, in which restraints such as distance atoms, solvent and torsion angles are considered. Once an initial protein has been modeled, the refinement must be done. The initial models often have a high energy associated and a low resolution, thus an energy minimization is performed to refine the initial models. This step begins with the assumption that native proteins have the lowest energy conformation. Therefore, the aim of protein structure refinement with molecular dynamics (MD) using force fields is to obtain a structural model with minimum energy. Force fields estimate the energy associated to each possible conformation of the protein structures modeled to achieve the correct covalent geometry, avoid atomic overlaps, select the nearest conformation to the native structure and assess the model quality (Bordner, 2012). The MD consists of the simulation of protein motion according to Newton's law of motion to obtain the energy of a particular protein conformation through calculations by force fields, which is usually performed with AMBER (Weiner and Kollman, 1981), GROMACS and NAMD (Phillips *et al*., 2005) package (Berendsen *et al*., 1995). It is noteworthy that the MD simulations are not only used for refinement of homology models, but also to determine structural changes in proteins. For example, the MD simulation was used to assess the binding dynamics and specificity of insect PBPs. Thus, Gräter *et al*. (2006) reported a similar binding affinity for bombykol and bombykal to BmorPBP1 through MD simulations, considering entropic contributions to the free energy of the protein/ligand binding complexes. Although bombykol acts as hydrogen-bond donor and bombykal as hydrogen-bond acceptor, the MD simulation showed that the binding modes of both bombykol and bombykal are similar, forming and breaking hydrogen bonds in a reversible way. More recently, Chu *et al*. (2013) provided a putative pH-induced ligand-releasing mechanism, which was predicted using MD simulation. Their findings suggested that the OBP1 of *C. quinquefasciatus* (CquiOBP1) releases the mosquito oviposition pheromone (MOP) at low pH due to a cleavage of hydrogen bonds in the binding site. These simulations have been performed on crystal structures of OBPs. However, Yu *et al*. (2012) reported a constant pH molecular

dynamics simulation on the modeled structures of *L. dispar* PBPs (LdisPBPs) at both pH 7.3 and pH 5.5. Subsequently, the authors corroborated their findings by circular dichroism, where conformational changes were revealed.

Overall, MD simulation involves classical mechanics based on Newtonian physics in order to reduce the computational cost (Durrant and McCammon, 2011). In contrast to quantum mechanics, classical mechanics or molecular mechanics (MM) works to find the global minimum of free energy in a protein structure to describe a conformation (Bordner, 2012) considering atoms and bonds as single particles and springs, respectively. All refinement processes on homology models of insect OBPs have used MM. An example is the predicted structure of the salivary OBP2 of the mosquito *C. quinquefasciatus*, where GROMOS96 (Scott *et al*., 1999) was used as force field based on MM in the refinement step (Paramasivan *et al*., 2007). Likewise, a refined homology model of *Locusta migratoria* OBP1 (LmigOBP1) was obtained using CHARMM force field (Brooks *et al*., 1983) by Jiang *et al*. (2009). Despite these studies on the refinement of predicted OBPs structures, there is no information on how long these simulations were performed. Nowadays, MD at long-timescale (e.g. nanoseconds) seems to be useful for refinement of homology models from proteins of small or medium size (Fan and Mark, 2004; Raval *et al*., 2012), such as insect OBPs.

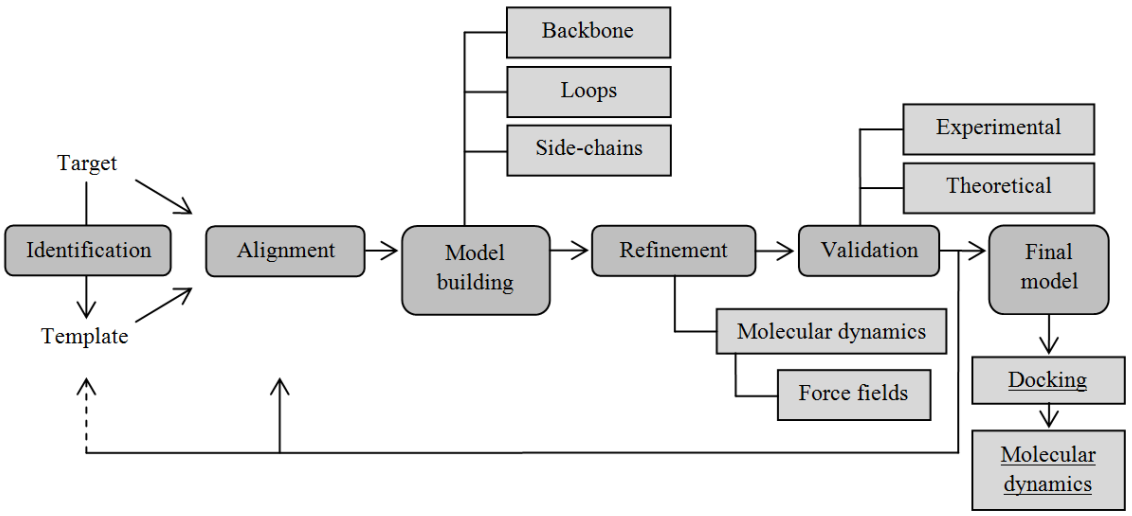


Figure 2.3. General building scheme of proteins by homology modelling applied so far to insect odorant-binding proteins. Underlined section indicates an optional step when ligand binding is studied.

Model validation is a crucial step to ensure its quality. There are two approaches for model validation: 1) experimental and 2) theoretical. Experimental approach compares the results from modelling studies against biological observations, and establishes the model quality. If the biological results are in accordance with the homology model, it is considered correct (Chang and Swaan, 2006; Ravna and Sylte, 2012). The theoretical validation analyzes the stereochemical quality of the model. Thus, when the model is not satisfactory, the steps are repeated from template identification or target-template alignment forward on (Martí-Renom *et al*., 2000). Therefore, all the steps mentioned above are usually repeated iteratively until the best homology model is obtained.

Insect OBPs are flexible and their motions are probably very important in their ligand recognition patterns and functions. To date, optional steps are being developed after the validation of multiple homology models, such as docking and MD simulations of ligand-protein complexes. Docking (also called molecular docking) is discussed in more detail in the following section. It is important to note that homology models as well as experimental protein structures are only a snapshot of a specific conformation. Therefore, all the results that are determined represent only one state of the protein without information of its dynamic. Thus, studies on drug discovery have directed efforts in the determination of how stable ligand-protein complexes are (Alonson *et al*., 2006). An example is the identification of an inhibitor binding site in Human Sirtuin 2 (SIRT2), which is a protein from histone deacetylase family (Sakkiah *et al*., 2013). In this work, five well known inhibitors were docked to SIRT2, such as suramin, mol-6, sirtinol, 67 and nf675. Plotted RMSD values revealed that the five complexes between the inhibitors and SIRT2 were stable after approximately 10 ns of simulations. The stability of certain protein structures or ligand-protein complexes is commonly observed when RMSD values become unchanged with time between the predicted structural conformations and surrounding environment (i.e. water molecules and ions) (Sokkar *et al*., 2011).

Despite homology modelling is highlighted as the most reliable and precise method for protein structure prediction, there are usually three limitations: 1) template selection; 2) precise target-template alignment and 3) imprecisions in refinement. These are considered the main sources of errors in homology modelling (Arnold *et al*., 2006; Larsson *et al*.,

2008). Selection of a proper template depends on the evolutionary relationship with target, hence, structural and functional divergence from the target. To overcome the imprecisions from template selection and sequence alignment, one should select the template with the highest sequence identity (≥30%) to target, use MSA with manual intervention if necessary and use programs or servers like FUGUE and GenTHREADER for low sequence identities. Each two years the critical assessment of methods of protein structure prediction (CASP) (http://www.predictioncenter.org/index.cgi) assesses and reports the progress of modelling techniques. If the template selection for a full-length target is difficult because the sequence identity (less than 30%) or the template resolution is too low, the modelling of some conserved regions is feasible. For instance, homology modelling of α4β2 nicotinic acetylcholine receptor (nAChR) with homopentameric acetylcholine binding protein (AChBP) from the sea slug *Aplysia californica* as template had a sequence identity of 18 to 20%. Despite of such low identity, the shared features between the target nAChR and template AChBP, such as a similar fold and highly conserved residues in binding site reported by Hansen *et al*. (2005), allowed to predict a nAChR model using the AChBP as a suitable template (Iturriaga-Vásquez *et al*., 2010).

It is important to get precise alignment between targets and templates. Distant evolutionary relationships, specific regions that do not aligned reliably and MSA programs themselves could lead to imprecisions in the alignment. Arnold *et al*. (2006) highlighted that when imprecisions of a target-template alignment cannot be determined, visual and manual intervention can significantly improve the model quality. If the precision of a sequence alignment decreases, the percentage of equivalent Cα atoms for the superimposed target and template structures increases, this is measured with RMSD. Raval *et al*. (2012) noted that the results of homology modelling have a low resolution by imprecise refinement during MD simulation with errors reflected in the values of 3Å for atomic coordinates. The timescale of MD application and the precision of force fields could also be the sources of errors. However, the main failure of MD within the refinement process seems to be caused by force fields errors (Raval *et al*., 2012). The same authors proposed a restricted conformational sampling for the initial homology model to improve the predicted structures. This conformational sampling emerges from the MD (referred to the protein motion possibly with surrounding solvent) and the use of energy functions (force fields)

(Bordner, 2012), comprising sampling steps to find the global optimal structures with lowest energy.

2.5.2. Ligand binding affinity to homology models of OBPs: methods and interaction

Many types of biological effects are produced in vertebrates and invertebrates due to the ligand-protein interactions. The transport function of OBPs is carried out through an interaction between ligands and the binding sites of these proteins. As previously mentioned, homology modelling as an inexpensive and quick method allows to predict the binding site in protein structures. Once a good quality homology model is obtained, it is possible to perform molecular docking of ligands, which predicts how and where small molecules could be bound to the protein model. Before docking, an important step is the identification of the binding pocket to find the cavities suitable as potential binding sites for ligands. For this a ligand structure preparation is necessary, this step can be achieved by searching ligands in databases (e.g. PubChem http://www.pubchem.ncbi.nlm.nih.gov/) or drawing them using software (e.g. Chem3D or Spartan). Then, molecular docking is applied a number of times enough to achieve the best ligand-protein conformation, which is based on the free energy associated to possible conformations. The docked conformations are commonly ranked according to an increase of energy, where the conformation with the lowest energy is ranked on the first place (best ligands). This method requires algorithms for the docking of ligands, such as Monte Carlo docking (Metropolis *et al*. 1953), molecular dynamics docking (McCammon *et al*., 1977; Mangoni *et al*., 1999), genetic algorithms (Morris *et al*., 1998) and ligand fragment-based method (Rarey *et al*., 1996). The use of these docking algorithms depends on the success of its application on some protein family. An example is the study of volatile compounds affinity to LmigOBP1 performed by Jiang *et al*. (2009). The specific binding between LmigOBP1 and pentadecanol (the ligand with the highest affinity) was carried out with CDOCK Software (Wu *et al*., 2003) through a molecular dynamic algorithm. Docking simulations in this work suggested that Asparagine (Asn74) is a key amino acid in the binding site of this OBP. Other example is to dock ligands with protein crystal structures, which is ideal because of the high resolution of these experimental structures compared with homology

models. Thus, He *et al*. (2010) carried out a docking simulation on BmorGOBP2, using a genetic algorithm with AutoDock software. Their findings showed that two hydrogen bonds were formed with Arg110 and Glu98 in bombykol-BmorGOPB2 complex, which is consistent with 3D structure of bombykol-BmorGOPB2 complex reported by Zhou *et al*. (2009). Furthermore, from 7 analogue ligands, docking predicted that (10*E*,12*Z*)-hexadecadienyl acetate had lower energy than bombykol and bombykal, thus it binds better to BmorGOBP2.

Molecular docking has been used mostly in structure-based drug discovery. Kolb *et al*. (2009) have highlighted that this method is suitable for ligand screening compared with the empirical method called high-throughput screening (HTS), allowing hit rates of 10-fold to 1000-fold higher than HTS. Besides predicting the presence of the binding cavities in proteins for ligand binding, molecular docking methods can also indicate amino acid residues that form the cavities. It can also provide a fine selection of ligands from extensive libraries, thus allows a deeper analysis of non-covalent interactions such as vdW interactions. For instance, Jiang *et al*. (2009) reported vdW interactions for hydrophobic residues in the binding pocket of LmigOBP1. The same authors suggested that these interactions can occur through benzene rings even with Tyr109 and Tyr117 as hydrophilic residues. To date, only 12 studies have performed molecular docking on the homology models of insect OBPs and the relevant information about key residues for ligand binding have been published (Honson *et al*., 2003; Jiang *et al*., 2009; Biessman *et al*., 2010; Rusconi *et al*., 2012; Yu *et al*. 2012; Zhang *et al*., 2012; Li *et al*., 2013; Wang *et al*., 2013; Yu and Plettner, 2013; Lu *et al*., 2014; P D *et al*., 2014; Zhuang *et al*., 2014) (Table 2). Although different software have been used for docking, the most recent studies seem to be coinciding in the development of detailed analysis with more parameters included, such as energy minimizations, theoretical evaluations and more sophisticated algorithms. An example is the comprehensive study performed by P D *et al*. (2014) for a rapid screening of active semiochemicals using a computational reverse chemical ecology approach on a GOBP of the oriental fruit fly *Bactrocera dorsalis*.

Table 2.2. Molecular docking on homology models of insect odorant-binding proteins.

OBP	Ligands	Software	Reference
A. gambiae OBP1	Indole	QUANTA-CHARMM	Biessman *et al*., 2010
L. migratoria OBP	Pentadecanol	CDOCK	Jiang *et al*., 2009
H. armigera OBP1, OBP2 and OBP3	(*Z*)-9-hexadecenal and (*Z*)-11-hexadecenal	-	Zhang *et al*., 2012
L. dispar PBP1	Aziridine		Honson *et al*., 2003
A. gambiae OBP1	*N*-phenyl-1-naphylamine (Generic ligand)	AutoDock 4.0	Rusconi *et al*., 2012
L. dispar PBP1 and PBP2	Palmitic acid n-butyl ester, bis(3,4-epoxycyclohexylmethyl) adipate and L-trans-epoxysuccinyl-isoleucyl-proline methyl ester propylamide	CDOCK	Yu *et al*., 2012
L. dispar PBP1 and PBP2	(+) and (-) disparlure	Molecular Operating Environment (MOE)	Yu and Plettner, 2013
A. cerana OBP2	*N*-phenyl-1-naphylamine (Generic ligand)	Molegro Virtual Docker 4.2	Li *et al*., 2013
A. lineolatus OBP5	*cis*-nerolidol	CDOCK	Wang *et al*., 2013
H. oblita OBP1	Hexyl benzoate, β-ionone, cinnamaldehyde and myrcene	AutoDock 4.0	Zhuang *et al*., 2014
A. cerana ASP2	Linalool, geraniol, β-ionone, 4-allylveratrole, phenylacetaldehyde, dibutyl phthalate, isoamylacetate, methyl-*p*-hydroxyl benzoate and butanedione	Molegro Virtual Docker 4.2	Lu *et al*., 2014

Although the docking method is a powerful tool to predict ligand-binding interactions, other experimental methods are also used in addition to with homology modelling, such as fluorescence competitive binding assays (FBA), radioactively labeled ligand (RLL) and site-directed mutagenesis (SDM). FBA and RLL comprise two technologies for receptor-ligand binding assays. Through radio-isotopic labels such as ^{3}H, RLL can be used for determining of receptor distribution and subtypes, screening of ligands and quantification. However, this binding assay has major disadvantages (e.g. radioactive wastes, high costs, health hazards and requirement of special licenses) compared with FBA, which is a widely used method to determine affinity of insect OBPs to a wide range of compounds (Zhou *et al*., 2004; de Jong *et al*., 2005). FBA provides a ligand screening with the use of fluorophores (commonly 1-NPN), and quantification through displacement of the fluorophore by ligands. An example is the FBAs used by Zhang *et al*. (2012) for 113 compounds, using homology modelling as a complement. Whereas these binding assays indicated that (*E*)-β-farnesene, ethyl butyrate, ethyl heptanoate and acetic acid 2-methylbutyl ester were the best ligands for *H. armigera* OBP5, the homology 3D model revealed key residues in the OBP5 binding site, which may have an important role in ligand specificity. Furthermore, the protein structure prediction by homology modelling along with mutagenesis allows to examine ligand-protein interactions, the role of non-conserved residues in the binding sites and ligand binding modes (Cavasotto and Phatak, 2009). The ligand binding affinity was altered through site-directed mutagenesis in *A. gambiae* OBP1 reported by Rusconi *et al*. (2012). Also physicochemical properties were changed with the mutations on the residues in the binding site, such as Phe59, Met84, His111, Trp114, Tyr122, Phe123 and Leu124. Thus, the mutated OBPs exhibited a lower binding affinity in the binding assays. More recently, Zhuang *et al*. (2014) reported that Tyr111 of the scarab beetle *Holotrichia oblita* OBP (HoblOBP1) is a key residue in the binding site of the protein. Through homology modelling, molecular docking and subsequent site-directed mutagenesis along with fluorescence binding assays, the authors demonstrated the role of Tyr111 in the binding of β-ionone, myrcene, hexyl benzoate and cinnamaldehyde.

2.6. Conclusions

Insect OBP research has evolved and is ready to offer alternative approaches for ligand screening aiming to identify semiochemicals in the hope for their eventual use in pest management. Due to intense studies on sensory system of insects, OBPs as semiochemical carriers have become attractive targets for ligand screening, which is the core issue of reverse chemical ecology. The absence of crystal structures of OBPs for a broad range of insect species has limited the progress. Homology modelling could provide information about 3D structure of these proteins and predict amino acid residues that could be involved in the formation of ligand-protein complexes. It is important to understand the factors that influence the quality of homology model building and the sources of errors in homology modelling such as the sequence alignment, selection of template(s), refinement and the type of interactions between the ligand and the binding sites. The appreciable sequence identities among Lepidoptera OBPs could provide precise homology models, and along with molecular docking (as virtual method to predict best ligands) could give a great leap in the search for super-ligands to manipulate insect behaviours. Noteworthy, OBP homology models as well as crystal and NMR structures provide only static representations of dynamic olfaction systems. Once molecular modelling is successfully performed, further robust bioassays need to be done and super-ligands need to be screened with insects to find potential pest control agents.

2.7. Acknowledgements

We would like to thank Dr. Andrés Avila for its valuable and helpful comments on an earlier version of this manuscript. This work was supported by CONICYT (21110933). We also thank two reviewers for their professional and valuable comments for the improvement of this work. Rothamsted Research receives grant-aided support from the Biotechnology and Biological Sciences Research Council (BBSRC) of the United Kingdom.

Chapter 3

β-ionone as putative semiochemical suggested by ligand binding on the odorant-binding protein 1 of *Hylamorpha elegans* (Burmeister) and electroantennographic recordings

Accepted in Entomological Science

β-ionone as putative semiochemical suggested by ligand binding on the odorant-binding protein 1 of *Hylamorpha elegans* (Burmeister) and electroantennographic recordings

Herbert Venthur[1,2], Jing-Jiang Zhou[4,*], Ana Mutis[2], Ricardo Ceballos[8], Rodrigo Mella-Herrera[5], Giovanni Larama[7], Andrés Avila[7], Patricio Iturriaga-Vásquez[6], Manuel Faundez-Parraguez[6], Marysol Alvear[3], and Andrés Quiroz[2,*]

[1]Programa de Doctorado en Ciencias de Recursos Naturales, Universidad de La Frontera, Temuco, Chile

[2]Laboratorio de Química Ecológica, Departamento de Ciencias Químicas y Recursos Naturales, Universidad de La Frontera, Temuco, Chile

[3]Departamento de Ciencias Químicas y Recursos Naturales, Universidad de La Frontera, Casilla 54-D, Temuco 4811230, Araucanía, Chile

[4]Department of Biological Chemistry and Crop Protection, Rothamsted Research, Harpenden, Herts. AL5 2JQ, UK

[5]Center of Waste Management and Bioenergy, Scientific and Technological Bioresource Nucleus, Universidad de La Frontera, Temuco, Chile

[6]Laboratorio de Síntesis Orgánica, Departamento de Química, Facultad de Ciencias, Universidad de Chile, Las Palmeras, Santiago, Chile

[7]Centro de Modelación y Computación Científica, Universidad de La Frontera, Temuco, Chile

[8]Laboratorio de Ecología Química, Centro Tecnológico de Control Biológico, Instituto de Investigaciones Agropecuarias (INIA)-Quilamapu, Chillán, Chile

Correspondence: Andrés Quiroz, Universidad de La Frontera, Temuco, Casilla 54-D, Chile. Tel.: +56 045 2732419; e-mail: andres.quiroz@ufrontera.cl ; Jing-Jiang Zhou, Rothamsted Research, Harpenden, Herts. UK. Tel: +44 1582 763133; email: jing-jiang.zhou@rothamsted.ac.uk

Abstract

Nowadays, odorant-binding proteins (OBPs) are considered the first filter of olfactory information for insects and constitute an interesting target for pest control. Thus, an OBP from the scarab beetle *Hylamorpha elegans* Burmeister (Coleoptera: Scarabaeidae, Rutelinae) (HeleOBP1) was identified, and ligand binding assays based on fluorescence and *in silico* approaches were performed, followed by a simulated binding assay. Fluorescence binding assays showed slight binding for most of the ligands tested, including host plant volatiles. A high binding affinity was obtained for β-ionone, a scarab beetle-related compound. However, the binding of its analogue α-ionone was weaker though is still considered good. On the other hand, through a three-dimensional model of HeleOBP1 constructed by homology, molecular docking was carried out with 29 related ligands to the beetle. Results expressed as free binding energy and fit quality (FQ) indicated strong interactions of sesquiterpenes and terpenoids (α- and β-ionone) with HeleOBP1 as well as some aromatic compounds. Residues such as His102, Tyr105 and Tyr113 seemed participate for the interactions previously mentioned. Both *in silico* scores supported the experimental affinity for the strongest ligands. Therefore, the activity of α-ionone, β-ionone and 2-phenyl acetaldehyde at antennal level was studied using electroantenography (EAG). Results showed that the three ligands are electrophysiologically active. However, an aliquot of β-ionone (represented by 3.0 ng) highlighted due to elicit stronger EAG responses in antennae of males than females. Finally, the role of these ligands as potential semiochemicals for *H. elegans* is discussed.

Keywords. Electroantennography, fluorescence binding assay, homology modelling, molecular docking, olfactory protein, Scarabaeidae

3.2. Introduction

More than thirty years have passed since the first insect odorant-binding protein (OBP) was discovered (Vogt and Riddiford, 1981). This OBP, identified in sensilla of the giant moth *Antheraea polyphemus* and named as pheromone-binding protein (PBP), is present mainly in males and it is proposed to bind pheromones. Although it is strongly suggested that

insect OBPs transport odorants from olfactory pores of sensilla to olfactory receptors (ORs) (Kaissling, 2013), other functions have been reported for insect OBPs: (1) ligand scavenger (Gong *et al*., 2009); (2) ligand desorption (Kowcun *et al*., 2001); (3) ligand recognition (Laughlin *et al*. 2008) and (4) ligand protection (Ishida and Leal, 2002). Nowadays, OBPs are divided in three subclasses, such as PBP, general odorant-binding protein (GOBP) and antennal-binding protein X (ABPx) for Lepidoptera. However, for Coleoptera, no difference has been made yet. It has been proposed that OBPs in scarab beetles could have two families: (1) Family of OBPs with 2 proteins, where OBP1 is more conserved than OBP2, and (2) family of conserved PBPs as the unique proteins present in antennae of beetles (Deng *et al*., 2012). However, very few studies on binding for scarab OBPs have been performed. Moreover, a limited number of these proteins have been identified in scarab beetles, considering that information from the only Coleopteran genome known to date, *Tribolium castaneum*, indicates the presence of up to 50 OBPs.

Considering the economic importance of scarab beetles during their larval stage (Leal, 1998; Vuts *et al*., 2014), the binding characteristics of scarab beetle's OBPs have been studied recently. For instance, fluorescence binding assays have been performed on the *Holotrichia parallela* OBP1 (HparOBP1), which showed binding affinities to a wider range of volatiles, such as β-ionone, hexyl benzoate and cinnamaldehyde (Ju *et al*., 2012). More recently, a putative cooperation among the *H. oblita* OBPs (HoblOBP1, HoblOBP2 and HoblOBP4) has been proposed. These OBPs may act as heterodimers, enhancing the binding affinity of ligands when the proteins are together (Wang *et al*., 2013). Later, Zhuang *et al*. (2014) reported the role of Tyr111 in the binding site of HoblOBP1 using molecular modelling approaches as well as experimental techniques to show how hexyl benzoate, β-ionone, cinnamaldehyde and myrcene cannot bind to the protein in the absence of Tyr111. In Chile, *Hylamorpha elegans* is a characteristic beetle belonging to Rutelinae subfamily, which is distributed from Region del Maule to Region de Los Lagos. This insect acquires a significant importance due to the damage caused to cereals and grass crops, where it feeds on roots during its larval stage. Likewise, the adult stage of *H. elegans* is characterized by feeding on leaves of trees such as *Nothofagus* species, in special *Nothofagus obliqua*. Less information is available around the adult stage of this scarab

beetle and the odorants involved in its life cycle. Only two putative sex pheromones, such as 1,4-hydroquinone and 1,4-benzoquinone have been reported (Quiroz *et al*., 2007). Likewise, a synergistic effect was proposed when 1,4-benzoquinone plus essential oil from *N. obliqua* significantly attracted males of the scarab beetle to traps baited with this blend. Hence, compounds from *N. obliqua* seem to be important in the behaviour of *H. elegans*. An outstanding mating behaviour driven by host plant volatiles has been suggested to play an important role for scarab beetles in several studies (Ruther *et al*., 2000; Reinecke *et al*., 2002; Quiroz *et al*., 2007). Males are able to recognize females while they are eating through a sexual kairomone, which is proposed to be released by the attacked plant. Field observations are consistent with this behaviour in *H. elegans*. Actually, a recent study on morphology and distribution of sensilla suggests that males of *H. elegans* have more chemosensory sensilla than females due to the ability of males to find and recognize females while they are feeding on leaves of *N. obliqua* (Mutis *et al*., 2014). Another special behaviour has been highlighted for this beetle, where after copulation females fly to crops, such as red clover *Trifolium pratense*, to deposit their fertilized eggs (Artigas, 1994), being this where the importance of the beetle as underground pest lies. It is thought that such behaviour could be driven by volatiles compounds emitted by crops.

In this study the binding characteristics of an OBP (HeleOBP1) presents in antennae of both males and females of *H. elegans* were examined. Thus, host plant volatiles, putative sex pheromones and semiochemicals reported for other scarab beetles were considered for binding. Here, we have performed fluorescence binding assays to evaluate the affinity of the ligands mentioned before. Likewise, homology modelling and subsequently molecular docking were applied for a binding simulation. Finally, three strong ligands were selected for electroantennography, which showed significant activity at certain concentrations.

3.3. Material and Methods

3.3.1. Volatile trapping and analysis

Volatiles were trapped by Porapak Q according to methodology reported by Quiroz *et al*. (1999) as well as by solid-phase micro extraction (SPME) according to Palma *et al*. (2012) with some modifications. Both methodologies for volatiles trapping were carried out in the Experimental Station INIA Carillanca, Vilcún, Chile. For two branches randomly chosen of *N. obliqua*, the apical 30 cm of each branch 1 m above the ground was enclosed in a glass bell. Volatiles were adsorbed on 100 mg Porapak Q columns during 24 h. The air was dried and purified by passage through activated 5-Å molecular sieves and charcoal. Trapped volatiles were desorbed from Porapak Q with hexane and concentrated to 250 µl under a flow of nitrogen, obtaining an extract of volatiles from *N. obliqua*. On the other hand, *N. obliqua* branches were enclosed to a glass bell at room temperature for 180 min. Thus, volatiles were trapped by SPME holder containing a 65-µm polydimethylsiloxane/divinylbenzene (PDMS/PVB) fiber, which was carefully arranged inside the glass bell. An empty glass bell was used as control. Likewise, six virgin adult *H. elegans* were put alive in two glass vials (i.e. females and males) with PTFE septum at room temperature for 24 h without food to avoid interference odors. Volatiles were trapped using an SPME holder with PDMS/PVB fiber. Volatiles collected by Porapak Q and SPME from *N. obliqua* and beetles were analyzed by gas chromatograph coupled with mass spectrometer (GC-MS) equipped with a BP-1 capillary column (30m, 0.22mm, 0.25µm) and with Helium as gas carrier (flow 0.5 ml/min). Ionization was by electron impact at 70 eV at 250°C. The GC oven was programmed to remain at 40°C for 2 min and increased at 5°C/min to 280°C. Kovats indices (KI) of volatiles were determined relative to the retention times of a series of *n*-alkanes with linear interpolation. Volatiles were identified through comparison between KI and mass spectra with available commercial standards (Appendix 3.1).

3.3.2. Insects and protein analysis

Both male and female of *H. elegans* adults were captured by trap light from Regional Research Center INIA-Carillanca in Vilcún, Araucanía, Chile. Beetles were captured during the seasonal flight period, from 11 November 2011 to 27 January 2012 according to Quiroz *et al*. (2007). Scarab beetles were recorded daily and taken to the laboratory for

determining sex and used for both protein analysis and RNA extraction. For protein analysis, the antennae and hindlegs from both males and females of *H. elegans* were anesthetized on ice and collected according to the methodology described by Ishida *et al*. (2002). After homogenizing in 10 mM Tris-HCl (pH 8.0), samples were centrifuged 2 x 10 min at 13 500 g and 4°C. The supernatant was concentrated by centrifugation under vacuum, analyzed by 15% native polyacrylamide gel (PAGE).

3.3.3. cDNA cloning and amplification by PCR

cDNA was synthesized by SMART RACE cDNA Amplification Kit (Clontech) from total RNA obtained from 50 antennae of male and female of *H. elegans* using RNA latter and the RNeasy Kit (Qiagen). For cDNA cloning, degenerated primers for *Phyllopertha diversa* OBPs (PdivOBP_1 and PdivOBP_2) were used (Appendix 3.2) (Wojtasek *et al*. 1999) in combination with poly T primer. Polymerase chain reaction (PCR) was carried out with PCR Master Mix 2X (Fermentas, Thermo). Thus, 40 cycles of stepwise amplification program were carried out with 95°C for 30 s, 42°C for 30 s and 72°C for 2 min. The PCR products were sent to sequence (Macrogen), analyzed by Sequence Scanner Software V1.0 and comparisons of sequences were carried out with nucleotide and protein BLAST (http://blast.ncbi.nlm.nih.gov/). Once the best sequences were confirmed as belonging to OBPs, EMBOSS Translation (http://www.ebi.ac.uk/Tools/st/) was used to obtain the amino acid sequence of the protein. Gene-specific primers (GSPs) were designed according to 3' sequence of the protein and used to obtain 5'-RACE (Appendix 3.2). The quality of first cDNA strand was tested by PCR using primers designed according to conserved regions of actine in insects: Actine-1, 5'-AA(C/T)TGGGA(C/T)GA(C/T)ATGGA(A/G)AA-3' and Actine-2, 5'-GCCAT(C/T)TC(C/T)TG(C/T)TC(A/G)AA(A/G)TC-'3. Actine DNA was amplified by 45 cycles with 1 min for denaturation at 94°C, 2 min for annealing at 45°C and 3 min for amplification at 72°C. The PCR products were analyzed in 0.8% agarose gel.

3.3.4. Sub-cloning in pNIC28-Bsa4 vector

In order to obtain construct for bacterial expression, previously cloned OBP gene was used for ligase-independent cloning (LIC). This method is characterized by an efficient cloning, which does not require restriction enzymes (Aslanidis and de Jong, 1990). LIC of PCR products has been improved in time to a high-throughput (HTP) cloning reaching efficiencies greater than 80% (Alzari *et al*., 2005). BsaI restriction enzyme was used to digest pNIC28-Bsa4 vector for 2 h at 50°C. Digested vector was loaded on 1% agarose gel and purified. Linearized vector was treated with 0.5 µL of T4 DNA polymerase and 10 mM deoxyguanosinetriphosphate (dGTP) for 30 min at 22°C. T4 DNA polymerase was inactivated by incubation for 20 min at 75°C. On the other hand, the identified OBP from *H. elegans* (HeleOBP1) was cloned into pGEM-T easy cloning vector. PCR for the construct using primers with BsaI adapters (Appendix 3.2) was performed as follow: 94°C for 3 min, 35 cycles of 94°C for 30 s, 60°C for 30 s, 72°C for 1 min and extension at 72°C for 10 min. PCR product was cleaned and treated with T4 DNA polymerase plus 10 mM deoxycytidinetriphosphate (dCTP) following the same conditions for pNIC28-Bsa4 vector. Both the treated vector and insert were mixed with 1:10 molar ratio in a total volume of 10 ul and incubated for 1 h at room temperature followed by transformation into BL21(DE3) competent cells. Correct gene insertions were tested by double digestion with both NdeI and SalI enzymes followed by sequencing.

3.3.5. Bacterial expression and purification

Positive colonies for the construct pNIC28-Bsa4-HeleOBP1 were used to inoculate 10 mL of fresh LB/kanamycin medium at 37°C. Protein expression was induced by adding isopropyl β-D-1-thiogalactopyranoside (IPTG) to 0.5 mM final concentration when the OD_{600} was 0.5-0.6. Cells were incubated for 3 h. To check the expression, aliquots of 200 uL of both induced and un-induced cells were used for SDS-PAGE analysis. Protein expression was scaled up with 500 mL of fresh LB/kanamycin medium in 2 L Erlenmeyer flask and induced with IPTG 0.5 mM final concentration at 37°C, overnight. Cells were collected and resuspended in 20 mM Tris-HCl pH 7.4. Resuspended cells were lysed by sonication and centrifuged. After SDS-PAGE analysis of supernatant and pellet, recombinant HeleOBP1 (rHeleOBP1) (presents in inclusion body) was solubilized in 3.5

mL of 5 M Urea, 2.5 mL of 10 mM DTT, 1 mL of 100 mM Cystine and 15 mL of 5 mM Cysteine. Sample was shaken at 4°C overnight. Dialysis of protein sample was carried out against 20 mM Tris-HCl pH 7.4. His-tagged HeleOBP1 was purified by two rounds of Ni ion affinity chromatography in an ÄKTA FPLC system (GE Healthcare, Hatfield, UK). Fractions were collected for each round and analysed by SDS-PAGE. The pure protein was delipidated at pH 4.5 with 100 µL of Lipidex-1000 (Perkin-Elmer) for 1 h on ice according to Siciliano *et al*. (2014). Refolding was carried out overnight by dialysis against 20 mM Tris-HCl pH 7.4 at 4°C.

3.3.6. Fluorescence binding assays

The fluorescence measurements were performed in a Luminescence Spectrometer LS50B (Perkin-Elmer) at 25°C with a 1 cm light path quartz cuvette and 5.0 nm slit for excitation and emission. Purified protein was diluted in 20 mM Tris-HCl pH 7.4 to a 2 µM of concentration. Likewise, the fluorescent probe *N*-phenyl-1-naphthylamine (1-NPN) was dissolved in high performance liquid chromatography purity grade methanol to 1 mM stock solution. To test the binding affinity of 1-NPN to HeleOBP1, 2 µM solution of the protein was titrated with 1 mM of the probe with concentration from 2 to 24 µM. 1-NPN was excited at 337 nm, and the emission spectra were recorded from 380 to 440 nm. The affinities of 29 ligands were measured in competitive binding assays using 1-NPN as the fluorescent reporter and 4-40 µM concentrations for each competitor.

Binding data were collected and maximum fluorescence intensity values were plotted against free ligand concentrations. Bound ligand was determined from the values of fluorescence intensity assuming that the protein was 100% active with stoichiometry of 1:1 (ligand:protein) at saturation. Scatchard plots were used to linearize curves. Thus, IC_{50} values were used to calculate the dissociation constants for each ligand with the equation: $K_i=[IC_{50}]/(1+[1\text{-NPN}]/K_{1\text{-NPN}})$, where [1-NPN] is the free concentration of 1-NPN and $K_{1\text{-NPN}}$ is the dissociation constant of the complex HeleOBP1/1-NPN.

3.3.7. Multiple template-based homology modelling

The amino acid sequence of HeleOBP1 was submitted to BLASTP program available on the NCBI website (http://www.ncbi.nlm.nih.gov/). Templates were selected based on the sequence identity (i.e. >30%) between crystal structures and HeleOBP1. Multiple structure alignments were generated by SALIGN command which is implemented in MODELLER. Multiple templates were used to increase the accuracy of multiple structure alignment as was reported by Sokkar *et al* (2011). Two hundred models of HeleOBP1 were obtained using MODELLER9.10 (http://salilab.org/modeller). Best models were selected according to discrete optimized protein energy (DOPE) score provided by the software. Likewise, best models were assessed using the theoretical validation package ProCheck (Laskowski *et al*., 1993). The modeled protein was visualized with PyMOL (http://www.pymol.org).

3.3.8. Molecular dynamics

Simulations were performed with NAMD v2.9 installed in the high-performance computer (HPC) Troquil Linux cluster at Centro de Modelación y Computación Científica (CMCC) from Universidad de La Frontera. CHARMM36 force field was used for all the simulations. The best modeled protein was solvated with water (TIP3P model) in a cubic box with a minimum distance of 5 Å between the protein and the edge of the box. The system net charge was neutralized by adding Na^+ or Cl^- randomly placed in the box. Likewise, the system was simulated under periodic boundary conditions with a cutoff radius of 12 Å for non-bonded interactions and a time step of 2 fs. Alpha-carbons (Cα) of secondary structures were fixed with a constant force of 1 kcal/mol/Å. A first energy minimization of 2 000 steps was performed followed by heating through short simulations of 1 ps at 50, 100, 150, 200, 250 and 300 K. Long simulations were kept at 300 K and 1 bar pressure in the NTP (referred to a constant number of particles, temperature and pressure) during 10 ns. Root-mean-square deviation (RMSD) trajectory tool was used to calculate the RMSD with reference to the starting structure. When the plotted RMSD did not showed any big changes, coordinates were analyzed every 50 frames to obtain the best structure (lowest

energy). Putative binding site and its volume were calculated by CASTp server (http://sts-fw.bioengr.uic.edu/castp/calculation.php) (Dundas *et al*., 2006).

3.3.9. Molecular docking

Molecular docking was carried out three times for each ligand by AutoDock4.2 (Morris *et al*., 2009) using the refined structure of HeleOBP1. Energy minimization for all chemical structures was performed by Chem3D software (http://www.cambridgesoft.com). Two hundred runs of Lamarckian genetic algorithm (GA) as the best method to find the lowest energy structures were used (Morris *et al*., 1998). Polar hydrogens were added by the interface AutoDock Tools as well as establish torsional bonds. Grid box with 40x40x40 points and default space of 0.375 Å was prepared by AutoGrid. The best binding modes were selected according to the lowest binding energy and the average of the three replicates was used to determine K_i according to the equation $K_i = e^{\Delta G/RT}$. For the best conformations, two parameters for comparisons were selected, (1) free binding energy and (2) fit quality (FQ), being this last an independent-size score for comparison of wider range of sizes in ligands. The ligand efficiency (LE) of a compound can be defined as the binding energy divided by its molecular size. To calculate FQ, the following equations $LE\_Scale=0.0715+(7.5328/HA)+(25.7079/HA^2)+(361.4722/HA^3)$ and $FQ=LE/LE\_Scale$ were used according to Bembenek *et al*. (2009). Where LE_Scale is the scaling of raw ligand efficiencies and HA is the number of heavy (non-hydrogens) atoms. Thus, compounds considered as good ligands are those with FQ near 1.0. On the contrary, lower FQs are related to poor ligand efficiencies (Reynolds *et al*., 2007; Bembenek *et al*., 2009). Therefore, to narrow the number of strong ligands according to the *in silico* approach we considered FQs ≥ 0.70 as cutoff.

3.3.10. Electroantennographic recordings

Antennal response of adult females and males of *H. elegans* was determined by EAG according to the methodology described by Reinicke *et al*. (2005) with some modifications. Five antennae per sex were excised from the head, where lamellae were carefully separated

from the rest of antennal segments. Subsequently, a lateral lamella was excised in order to expose one side of the middle lamella, where olfactory sensilla are located (Mutis *et al*., 2014). Signals from antennae were conducted and recorded by Syntech equipments (Kirchzarten, Germany). Data recorded were displayed and analyzed by the software GcEad 2012 v1.2.4. An aliquot of 30 µL of α-ionone, β-ionone and 2-phenyl acetaldehyde in hexane (0.1, 1.0, 10, 100 and 1 000 ppm, being 3.0, 30, 300, 3 000 and 30 000 ng of stimulus, respectively) was loaded in a piece of filter paper (2 cm^2), which was then inserted into a glass Pasteur pipette. Each odor stimulus was delivered as a continuous airstream (500 mL min^{-1}) from the Pasteur pipette for 2.0 sec. Intervals of 60 sec between puffs were used to ensure antennal recovery.

3.3.11. Statistical analysis

To minimize any variation among antennae, EAG responses were corrected according to the solvent amplitude before and after the stimulus. This was carried out by the formula $R_c=R_s-[(R_{sb}+R_{sa})/2]$ (Guo and Li, 2009), where R_c = response corrected, R_s = response to the stimulus, R_{sb} = response to the solvent before and R_{sa} = response to the solvent after. Thus, to determine differences of EAG responses to the concentrations used for each stimulus, data were submitted to analysis of variance ($P<0.05$) and the Tukey test ($P<0.05$) for group separation. Finally, t-student test ($P<0.05$) was used to evaluate differences in EAG response between *H. elegans* males and females.

3.4. Results

3.4.1. Sequence analysis, sub-cloning and expression of HeleOBP1

The cDNA obtained from antennae of *H. elegans* consisted in 348 bp with high sequence identity of 95-98% to other PBPs and OBPs reported for scarab beetles. The gene was clustered with a big group of scarab beetle PBPs and named as HeleOBP1 as less information is available in terms of binding (Appendix 3.3) Thus, the *HeleOBP1* gene was cloned from the antennae cDNA into pGEM vector (Promega) with the aim to obtain

recombinant protein to perform binding assays. Ligase independent cloning (LIC) was then performed to subclone *HeleOBP1* gene by PCR of pGEM-HeleOBP1 construct into an expression vector pNIC28 (a kind gift from Prof. Nick Keep, University of London) for the recombinant protein expression. The recombinant protein was expressed as inclusion bodies. Therefore, the protein was denatured with urea/DTT, re-natured using cysteine-cystine redox reaction and dialysed to obtain a soluble form as reported by Plettner *et al*. (2000). However, the expressed OBP may contain endogenous ligands from the bacterial cells (Lagarde *et al*., 2011). Hence, delipidation was performed on purified protein (Siciliano *et al*., 2014). SDS-PAGE analysis of bacterial pellets and purified HeleOBP1 are shown in Figure 3.1 as homogenous recombinant protein. It is worth noting that a short His-tag section (22 amino acids) from pNIC28 vector was added to the N-terminal of the OBP. Although the His-tag part could represent interference for binding, its expression in pNIC28-Bsa4 vector along with the TEV cleavage site associated, make the removal of His-tag experimentally costly and difficult at high scale. Furthermore, we considered the fact that in some insect OBPs the C-terminal section, instead the N-terminal, has an important role for binding. Therefore, it was thought that the His-tagged OBP would still represent a functional recombinant protein from which to obtain binding data.

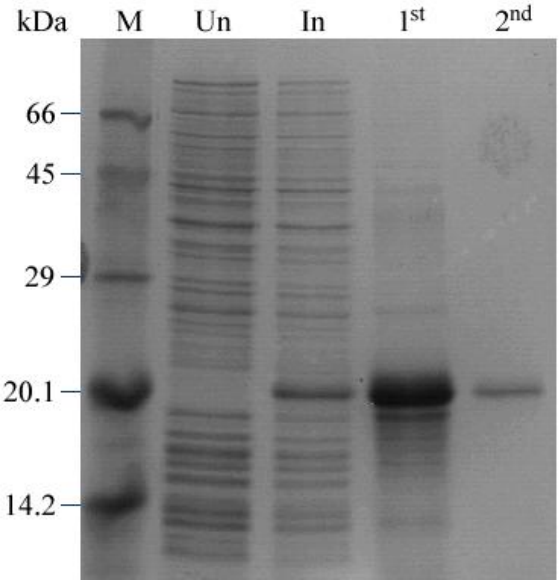


Figure 3.1. HeleOBP1 protein expression and purification. Protein from inclusion bodies was purified with two rounds of Ni ion affinity chromatography. M: protein weight marker (BSA: 66 kDa, Ovalbumin: 45 kDa, Carbonic anhydrase: 29 kDa, Trypsin inhibitor: 20.1

kDa, Lactalbumin: 14.2 kDa); Un: un-induced bacterial cells; In: induced bacterial cells with IPTG at 0.5 mM final concentration; 1st: First purification round and 2nd: Protein with second purification round, delipidation and 2 µM concentration for fluorescence binding assays.

3.4.2. Fluorescence binding assays

To determine the binding affinity of recombinant HeleOBP1 (rHeleOBP1) to several volatiles, we first measured the affinity of the fluorescent probe 1-NPN to rHeleOBP1. The results showed a good binding between 1-NPN and rHeleOBP1 (Figure 3.2) with a binding constant (K_d) of 9.52 µM. Once corroborated the binding of HeleOBP1/1-NPN, the competitive binding was performed using host plant *N. obliqua* volatiles identified by our laboratory plus volatiles reported before by Quiroz *et al*. 1999. Thus, no strong binding was found for most volatiles tested (Table 3.1). However, some of them (e.g. benzaldehyde, cinnamaldehyde and host plant volatiles) could displace 1-NPN from the binding site of HeleOBP1, but high concentrations (> 40 µM) were needed. The low binding affinity to the volatiles tested shows that HeleOBP1 is not specific or its true ligand(s) is not found and tested in this study.

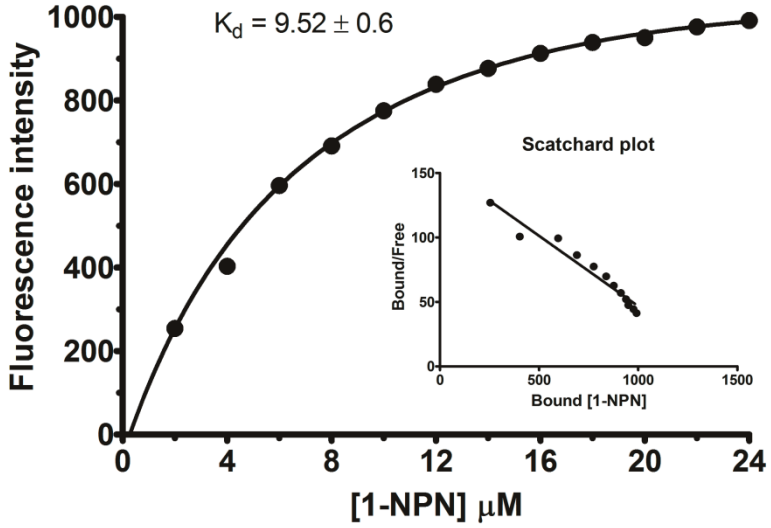


Figure 3.2. Binding curve of 1-NPN and Scatchard plot for the OBP of *H. elegans*. A 2 µM solution of HeleOBP in Tris buffer was titrated with 1 mM solution of 1-NPN in

methanol to final concentrations of 2-24 µM. The dissociation constant (K_d) (from the average of three replicates) was determined using Prism software.

Table 3.1. Binding affinities of HeleOBP1 to host plant volatiles and compounds with reported significance for scarab beetles.

Ligand	K_i (µM)	IC_{50} (µM)	Molecular Formula	HPV	PP	K	A
Terpenes							
β-caryophyllene	-	-	$C_{15}H_{24}$	◆			
α-gurjunene	-	-	$C_{15}H_{24}$	◆			
(+)-aromadendrene	-	-	$C_{15}H_{24}$	◆			
β-ocimene	-	-	$C_{10}H_{16}$	◆			
α-pinene	-	-	$C_{10}H_{16}$				●
Geraniol	46.2	34	$C_{10}H_{18}O$				●
α-ionone	21.4	16	$C_{13}H_{20}O$				●
β-ionone	6.9	5	$C_{13}H_{20}O$				●
β-myrcene	-	-	$C_{10}H_{16}$	◆			
(±)-linalool	50.3	37	$C_{10}H_{18}O$	◆			
Alkanes							
Dodecane	-	-	$C_{12}H_{26}$	◆			
Tetradecane	-	-	$C_{14}H_{26}$	◆			
Alcohols							
(*E*)-2-nonen-1-ol	47.6	35	$C_9H_{16}O$		●		
Heptan-2-ol	-	-	$C_7H_{16}O$				●
Ester							
(*Z*)-3-hexenyl acetate	-	-	$C_8H_{14}O_2$	◆			
Hexyl acetate	53.0	39	$C_8H_{16}O_2$	●			
Aldehydes							
Nonanal	54.4	40	$C_9H_{18}O$	●			
Decanal	49.0	36	$C_{10}H_{20}O$	●			
Ketone							
Acetoin	-	-	$C_4H_8O_2$		●		
Aromatics							
Phenol	-	-	C_6H_6O	●			
Benzaldehyde	49.0	36	C_7H_6O	●			
1,4-benzoquinone	-	-	$C_6H_4O_2$		●		
1,4-hydroquinone	53.0	39	$C_6H_6O_2$		●		
Eugenol	44.9	33	$C_{10}H_{12}O_2$				●
2-phenyl acetaldehyde	16.3	12	C_8H_8O			●	
2-phenyl ethanol	42.2	31	$C_8H_{10}O$			●	
2-phenylethyl propionate	44.9	33	$C_{11}H_{14}O_2$				●
2-phenyl acetonitrile	51.7	38	C_8H_7N			●	
Cinnamyl alcohol	39.4	29	$C_9H_{10}O$				●

Where applicable, we report the dissociation constant (K_i) and concentration of ligand halving the initial fluorescence intensity (IC_{50}). We also report the chemical formula of ligands and their role as

host plant volatiles (HPV), putative pheromone (PP), kairomone (K) or attractants (A). '-' represents data not available. '●' represents the role of the ligand for scarab beetles reported in literature and '◆' represents the ligands identified from the host plant of *H. elegans* in this study.

Putative sex pheromones for *H. elegans* such as 1,4-hydroquinone and 1,4-benzoquinone were reported by Quiroz *et al*. (2007) and also used as competitors for the binding study. (*E*)-2-nonen-1-ol, a sex pheromone reported for several *Anomala* beetles, was also included. Moreover, we have identified a compound released from females of *H. elegans*, acetoin. Results showed no strong binding between HeleOBP1 and the putative sex pheromones despite HeleOBP1 could bind with low affinity to 1,4-hydroquinone (K_i 53.0 µM) and (*E*)-2-nonen-1-ol (K_i 47.6 µM). Considering that host plant volatiles identified here were not strong competitors to displace 1-NPN, we selected other volatiles with a significant reported role for the Rutelinae subfamily, where *H. elegans* belongs. Thus, both kairomones and attractants from the pherobase database (http://www.pherobase.com/) were selected. From the scarab beetle-related compounds, α-ionone, β-ionone and 2-phenyl acetaldehyde showed good binding to HeleOBP1. However, β-ionone was the strongest in terms of binding. The binding affinities of these compounds and other compounds tested in this study are shown in Appendix 3.4. The results indicated that aromatics compounds such as geraniol, eugenol, 2-phenyl ethanol, 2-phenylethyl propionate, 2-phenyl acetonitrile and cinnamyl alcohol, had an appreciable binding affinity to HeleOBP1 at relatively high concentrations (K_i 46.2, 44.9, 42.2, 44.9, 51.7 and 39.4 µM, respectively). However, 2-phenyl acetaldehyde had the lowest K_i (16.3 µM) and the highest affinity to HeleOBP1 among the aromatic ligands. Surprisingly, both α-ionone and β-ionone were bound strongly to HeleOBP1 with K_i values of 21.4 µM for α-ionone and 6.9 µM for β-ionone (Figure 3.3). Moreover, the appreciable difference in the binding affinity to HeleOBP1 between these two terpenoid isomers suggests that HeleOBP1 can bind selectively to terpenoids and discriminate between two highly similar ligands.

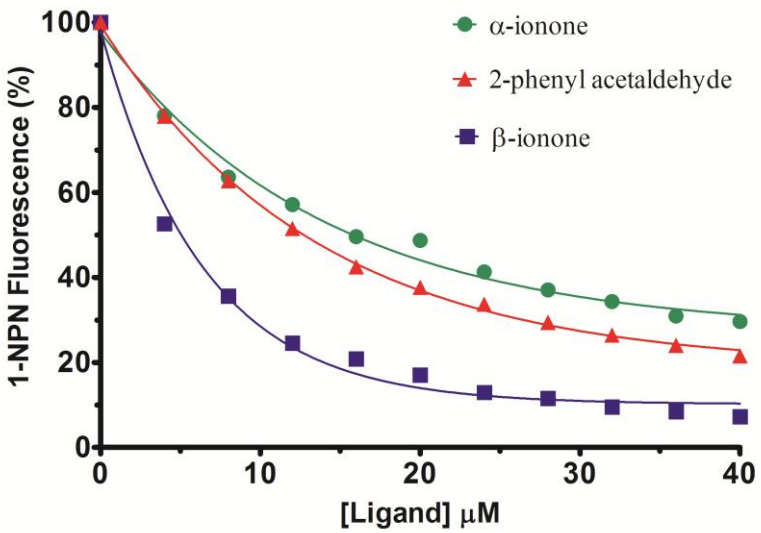


Figure 3.3. Competitive binding of HeleOBP1 to α-ionone, β-ionone and 2-phenyl acetaldehyde. A 2 μM solution of the protein plus 1-NPN was titrated with 1 mM solutions of each ligand in methanol to final concentrations of 4-40 μM.

3.4.3. Protein structure prediction and molecular docking

There are no OBP structures determined for any beetles yet. However, experimentally-determined three-dimensional (3D) structures of insect OBPs, such as moths and mosquitoes are available in Protein Data Bank (PDB) (http://www.rcsb.org/pdb/home/home.do). In an attempt to understand the interaction between HeleOBP1 and the chemicals tested in this study, we predicted the 3D structure of HeleOBP1 by homology modelling, which is characterized as the best method currently used for protein structure prediction (Bordoli and Schwede, 2012; Ravna and Sylte, 2012). Moreover, we used multiple templates to improve the quality of models (Larsson *et al*., 2008). The OBPs from mosquito *Anopheles gambiae* (AgamOBP1) (PDB: 2ERB) and *Culex quinquefasciatus* (CquiOBP1) (PDB: 2L2C and 3OGN) were used as templates to build the structure model of HeleOBP1. There are 33-35% of sequence identities between HeleOBP1 and the templates, which is considered good enough to obtain accurate models (Schwede *et al*., 2007). The 3D structure of HeleOBP1 suggests the presence of six α-helices as the main feature of insect OBPs (Pelosi *et al*., 2006) located between amino acids as follows: Glu3-Thr20 (α1), Glu24-Asp32 (α2), Glu40-Met53 (α3), Val64-Ile70 (α4),

Asp73-Arg84 (α5) and Pro94-Thr108 (α6). Likewise, a multiple sequence alignment between the templates and HeleOBP1 indicates the presence of six conserved Cys residues (Appendix 3.5). Subsequently, the 3D model of the protein showed three disulfide bridges connected as follows: Cys16-Cys48, Cys44-Cys95 and Cys86-Cys104. The presence of three disulfide bridges allow us to classify HeleOBP1 as a classic OBP as has been suggested by several authors (Zhou, 2010; Fan *et al*., 2011). The modeled structure of HeleOBP1 was then used to calculate *in silico* its binding affinity to a wide range of chemicals by molecular docking based on free binding energy and a size-independent score called fit quality (FQ) (Appendix 3.8). It is worth mentioning that there was a strong dependency between the free binding energy, commonly calculated by docking software such as Autodock, and ligand molecular size (Appendix 3.6) as reported in other studies (Reynolds *et al*., 2007). This allowed us to use FQ for the affinity calculation and compare the score with those obtained by the fluorescent binding assay. The molecular docking suggests that HeleOBP1 could provide a good accommodation for acetoin, phenol, benzaldehyde, sesquiterpenes, α-ionone and β-ionone in terms of FQ. However, free binding energy still suggests α- and β-ionone as highly stable binders to HeleOBP1. Thus, the *in silico* binding assays suggested the participation of four main residues, such as His102, Tyr105, Tyr113 and Met114 in the stabilization of the complex of HeleOBP1 with the strong binding ligands (α-ionone, β-ionone and 2-phenyl acetaldehyde) (Figure 3.4). Acetoin showed a high FQ (over 1.0) and the highest K_i (3.52 mM), suggesting a good fitting but bad protein-ligand stability. On the other hand, the binding of 2-phenyl acetaldehyde showed a FQ=0.73 and an experimental K_i of 16.3 μM suggesting a good fitting for the ligand.

Molecular docking was successful in render the difference between β-ionone (FQ = 0.82; -6.99 kcal mol^{-1}) and 2-phenyl acetaldehyde (FQ = 0.73; -4.78 kcal mol^{-1}) (Table 3.2). Similarly, but less noticeable, the *in silico* approach suggested α- and β-ionone as strong ligands in terms of both free binding energy and FQ. Here, four main residues are predicted to participate in HeleOBP1-ligand complexes supporting, particularly, the good binding of β-ionone.

Table 3.2. Binding affinities of HeleOBP1 to the strongest ligands suggested by i*n silico* and in vitro binding assays.

Ligand	$K_{i(exp)}$ (μM)	Binding energy (kcal mol^{-1})	FQ
α-ionone	21.4	-6.88	0.81
β-ionone	6.9	-6.99	0.82
2-phenyl acetaldehyde	16.3	-4.78	0.73

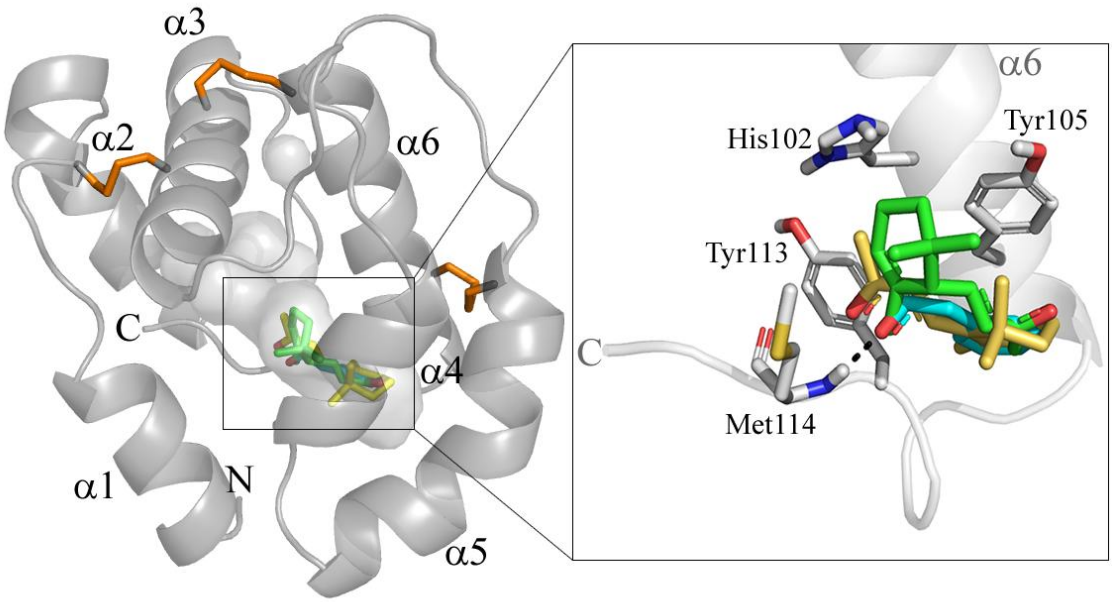


Figure 3.4. The homology model of HeleOBP with α-helices displayed as grey ribbons (α1-α6) in complex with the strongest ligands obtained from experimental binding assays. N- and C-terminals are indicated as N and C, respectively. Disulfide bonds are highlighted as orange sticks. Square indicates the section with zoom in as well as the docked conformations of α-ionone (green sticks), β-ionone (yellow sticks) and 2-phenyl acetaldehyde (light blue sticks). Hydrogen bond is indicated as dashed lines.

3.4.4. Electroantennographic responses

The three strongest ligands (α-ionone, β-ionone and 2-phenyl acetaldehyde) in terms of binding to HeleOBP1 were selected for EAG recordings. The three ligands elicited EAG responses in antennae of both *H. elegans* males and females (Figure 3.5). However, no

dose-dependent responses were obtained for most of the tested antennae. Despite this, from 10 to 1 000 ppm males responded with higher amplitudes to 2-phenyl acetaldehyde, showing a slight dose-dependent response (Figure 3.5C). On the other hand, a significant difference between 0.1 and 1.0 ppm of β-ionone was obtained for females (Figure 3.5A). The same result was obtained for β-ionone at 1 000 ppm compared with 1.0 ppm in females. On the other hand, EAG responses for male antennae indicate significant difference between doses 0.1 and 1.0 ppm, being this similar to the EAG responses of females. Similarly, statistical analyses suggest that α-ionone does not elicit any dose-dependent response in both males and females (Figure 3.5B). The two terpenoids elicited stronger responses in males, with over 0.75 and 1.00 mV from 0.1 ppm for α-ionone and from 1.0 ppm for β-ionone. Despite 2-phenyl acetaldehyde showed an increasing EAG response in males, less than 0.25 mV amplitudes were obtained at lower doses (i.e. 0.1 and 1.0 ppm). It is worth mentioning that the ranges of amplitudes obtained for α-ionone in females and males throughout the doses were similar, and no significant difference was found. However, a rough comparison of dose between α- and β-ionone in females and males, suggests that antennae were more sensitive for α-ionone at 0.1 ppm considering their high amplitudes.

Analyses between sexes showed no difference in EAG response for almost all the stimulus doses (Figure 5). However, females responded significantly with higher amplitude to 2-phenyl acetaldehyde at 1.0 ppm than males. Likewise, a significant difference was obtained at 1 000 ppm for the same chemical, where males were more sensitive. An outstanding response for β-ionone at 0.1 ppm was obtained from males, being more electrophysiologically sensitive than females. It is worth mentioning, that this terpenoid has been selected as a strong OBP-binder, with inhibition constant (K_i) of 6.9 μM according to our previous results.

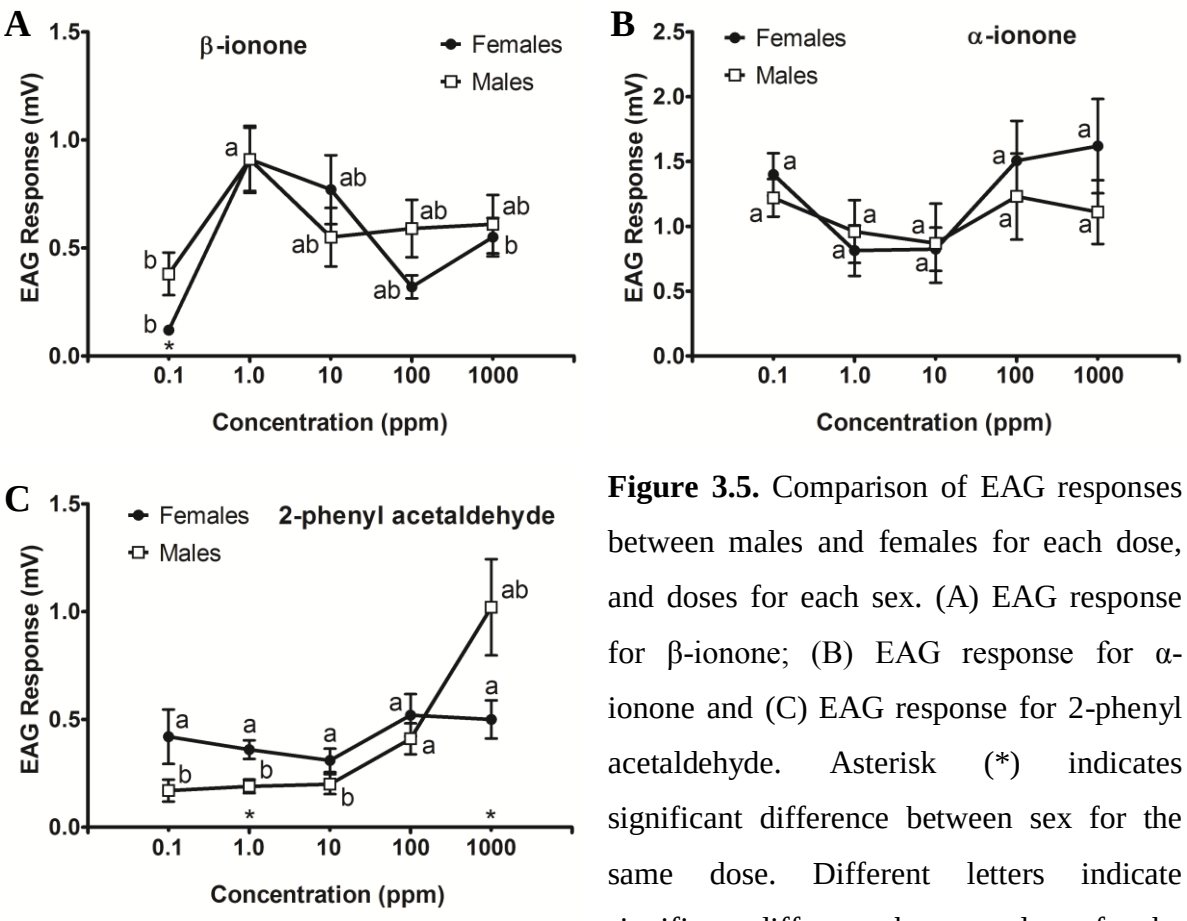


Figure 3.5. Comparison of EAG responses between males and females for each dose, and doses for each sex. (A) EAG response for β-ionone; (B) EAG response for α-ionone and (C) EAG response for 2-phenyl acetaldehyde. Asterisk (*) indicates significant difference between sex for the same dose. Different letters indicate significant difference between doses for the same sex.

3.5. Discussion

Here, we have identified a full-length OBP from antennae of both males and females of *H. elegans*. This protein has been temporarily considered as OBP due to the presence in both sexes and lack of binding information. Thus, in order to characterize the binding properties of HeleOBP1 and find putative semiochemicals, we have used several compounds from the main host plant, the tree *N. obliqua* (Giganti and Dapoto, 1990; Artigas, 1994; Quiroz *et al*., 1999; Klein and Waterhouse, 2000; Lanfranco *et al*., 2001), putative sex pheromones and semiochemicals reported in literature for scarab beetles, specifically for Rutelinae subfamily. All these compounds were tested using both fluorescence and *in silico* binding assays, whose results led the application of EAG.

Some putative sex pheromones and a potential source of sexual kairomones were tested here. However, the antennae-identified protein, HeleOBP1, showed a highly selective binding to compounds reported for other scarab beetles, such as α-ionone, β-ionone and 2-phenyl acetaldehyde. Despite of their strong binding to HeleOBP1, the physiological role of these compounds in *H. elegans* is unclear. Deng *et al*. (2012) reported a strong binding of β-ionone to HoblOBP1 and HoblOBP2 with a K_i of 6.35 and 5.36 μM, respectively. Moreover, β-ionone has been identified as highly attractive for both males and females of the scarab beetle *Anomala transvaalensis* (Donaldson *et al*. 1990). On the other hand, 2-phenyl acetaldehyde has been reported as kairomone to which the scarab beetle *A. octiescostata* is attracted. This chemical is highlighted as part of a chemically active mixture from dandelion *Taraxacum officinale* by a mixture of volatiles, such as (*Z*)-3-hexenyl acetate, benzaldehyde, 2-phenyl acetaldehyde and phenyl acetonitrile, among others (Leal *et al*., 1994d).

Although, α-ionone, β-ionone and 2-phenyl acetaldehyde do not show the strongest binding affinity found for insect OBPs, in this study they are considered as the strongest ligands supported not only by fluorescence binding assays, but also by molecular docking. Considering that a 3D model of HeleOBP1 was constructed, the identification of its binding site could represent a challenging task. However, the presence of ligands already bound to the templates (e.g. AgamOBP1-PEG complex and CquiOBP1-MOP complex) facilitated the identification of a potential binding site in HeleOBP1. Likewise, the consistency between the modeled binding site in HeleOBP1 and the predicted by CASTp server, provided us the area where molecular docking could be applied. Thus, this last *in silico* method suggested that α-ionone, β-ionone and 2-phenyl acetaldehyde seem to interact with His102, Tyr105 and Tyr113. Actually, His102 and Tyr113 are well conserved residues in part of the binding site of the templates AgamOBP1 and CquiOBP1 (Appendix 3.7). It is likely that π-π interaction, where both aromatic properties and unsaturations, could have an important role when these residues participate. For example, it has been proposed that these type of interactions could be established when Tyr111 in the binding site of HoblOBP1 interacts with cinnamaldehyde and β-ionone (Zhuang *et al*., 2014). Probably, Tyr113 in HeleOBP1 could play a similar role for β-ionone binding since the residue is well aligned

to the Tyr111. On the other hand, α- and β-ionone have a highly similar structure, though different binding modes were obtained (Figure 4). One big conformational cluster was obtained from the complexes HeleOBP1-ionones in Autodock, where limited movement of α- and β-ionone (i.e. two rotatable bonds) was showed. It is probable that these docking characteristics resulted in one stable binding mode for both ionones but opposite from each other. A closest view of the binding modes of these terpenoids would involve the dynamics of the complexes, where the movement of ligands into the binding site and a likely formation of hydrogen bonds could explain the above finding. For instance, a recent study on the moth *Loxostege sticticalis* has reported the key role of Thr15 and Trp43 in the binding site of *L. sticticalis* OBP1 (LstiOBP1). Multiple hydrogen bonds seemed to be interrupted when both Thr15 and Trp43 were mutated to Ala. A decrease of the binding affinity of ligands such as heptanol, (*E*)-11-tetradecenyl acetate, cinnamic aldehyde and (*E*)-2-hexenal, to over 40 μM of K_d was obtained (Yi *et al*., 2015). For 2-phenyl acetaldehyde case, it seems that the carbonyl group is able to form a hydrogen bond with the amino moiety of Met114. This residue could participate actively in forming the binding site of HeleOBP1 (Figure 4). The modeled structure of HeleOBP1 showed the C-terminal section forming a lid on the binding site, which makes Met114 free to establish interactions. Therefore, the C-terminal section might play a key role for ligand binding as it has been widely proposed for moths OBPs, such as *Bombyx mori* PBP1 (BmorPBP1) (Horst *et al*., 2001; Lautenschlager *et al*., 2005; Leal, 2005).

The binding of HeleOBP1 supported by fluorescence binding assays and molecular docking indicated β-ionone as a strong OBP-binder as it has been found for other scarab beetles, such as *H. oblita*. However, contrary to males, low EAG responses to the terpenoid have been reported for females (Deng *et al*., 2012), which is consistent with the results presented here. Thus, it is possible that a number of β-ionone-tuned ORs are part of male's antennae. Hence, explaining the sensitivity obtained at low concentrations by males. Despite the results obtained, it is still unclear where this compound is present in the context of *H. elegans*. Probably, this terpenoid might play a key role at some stage of the life cycle of *H. elegans*, considering that β-ionone has been reported in red clover *T. pratense* (Figueiredo *et al*., 2007). However, contrary to the sensitivity showed by males, gravid

females have been reported flying to crops, such as *T. pratense* (Artigas, 1994). This last behaviour could be explained because females are more sensitive to several factors of host plants, such as non-volatile secondary metabolites, surface structure, tissue toughness and water content, than only volatiles as semiochemicals. Thus, females could ensure a host plant with high nutritive value for themselves and their offspring. On the other hand, it has been proposed that due to the priority of mate finding, males are more sensitive to volatiles from either host plants or damaged plants, where they can find potential mates (Fernandez and Hilker, 2007). Although our results indicate β-ionone as a potential bioactive volatile, more evidence is necessary to conceive it as semiochemical for *H. elegans*.

Hitherto, only one significant attractive blend for *H. elegans* (i.e. 1,4-benzoquinone and essential oil from *N. obliqua*) has been reported. Likewise, neither pheromone nor kairomone has been identified for this scarab beetle. Although the binding characteristics of HeleOBP1 examined here suggest that the protein is not tuned to host plant volatiles, binding mechanisms could play a crucial role for active ligands at OR level, as has recently been proposed by Murphy *et al*. (2013). Likewise, it is probable that other OBPs could be actively participating in host plant volatile transport. Further experiments are necessary to determine other OBPs present in *H. elegans* as well as the role of the semiochemical-candidate compounds found in this work. Finally, our results represent the first step of a semiochemical screening in *H. elegans* starting from molecular approaches.

3.6. Acknowledgements

This research was partially supported by the supercomputing infrastructure of the NLHPC (ECM-02) at Centro de Modelación y Computación Científica, Universidad de La Frontera, CMCC-UFRO. This work was supported by CONICYT (21110933), DIUFRO DI 10-0018 and Rothamsted Research, which receives grant-aided support from the Biotechnology and Biological Sciences Research Council (BBSRC) of the United Kingdom. Finally, we would like to thank Dr. Ricardo Ceballos from Instituto de Investigaciones Agropecuarias (INIA), Quilamapu, Laboratorio de Ecología Química, Chillán for his valuable help during the development and analysis of electroantennographic recordings.

Chapter 4

Behavioural responses of adult *Hylamorpha elegans* to three active molecules in both OBP-binding and electroantennography assays

Behavioural responses of adult *Hylamorpha elegans* to three active molecules in both OBP-binding and electroantennography assays

Herbert Venthur[1,2], Ana Mutis[2], Ricardo Ceballos[3], Andrés Quiroz[2]

[1]Doctoral Program in Sciences of Natural Resources, Universidad de La Frontera, Temuco, Chile

[2]Laboratorio de Química Ecológica, Departamento de Ciencias Químicas y Recursos Naturales, Universidad de La Frontera, Temuco, Chile

[3]Laboratorio de Ecología Química, Centro Tecnológico de Control Biológico, Instituto de Investigaciones Agropecuarias (INIA)-Quilamapu, Chillán, Chile

Abstract

Hylamorpha elegans (Coleoptera: Scarabaeidae, Rutelinae) is recognized as a pest during its larval stage feeding on roots of crops, such as cereals and grass. Moreover, it acts as active defoliator of *Nothofagus obliqua* trees during its adult stage. An odorant-binding protein (OBP) called HeleOBP1 has been identified and studied, determining three strong ligands in terms of binding affinity, such as α-ionone, β-ionone and 2-phenyl acetaldehyde. Likewise, these three ligands have shown electrophysiological activity in antennae of adults *H. elegans* female and male. Considering the above, the role of these ligands (or odorants) has been evaluated through the response of *H. elegans* in vivo by Y-tube olfactometry. Thus, 20 beetles per ligand were used for females and males, considering 3.0, 30 and 300 ng of dose. Main results indicate β-ionone as a moderately attractive odorant for males from 3.0 to 30 ng of the stimulus. Despite no strong attractive effect was found, β-ionone could play a role at some point of the flight season of males *H. elegans*. Further experiments should be performed to establish the specific role of this compound.

Keywords. Olfactometry, coleoptera, scarabaiedae, *Hylamorpha elegans*

4.1. Introduction

Highlighted by attacking several plants and producing important economic damages, special efforts have been focused on beetles from Rutelinae subfamily (Leal, 1998). In Chile, *Hylamorpha elegans* Burmeister (Coleoptera: Scarabaeidae, Rutelinae) is an economically important pest for cereal and grass crops, considering that the production of cereal is over 1.100.000 tons and the sown area is over 280.000 ha (Danty and Olfos, 2012). This beetle, commonly known as San Juan Verde, feeds on roots of crops causing evident decomposition of plants and a subsequently occurrence of yellow spots during its larval stage (i.e. white grub) (Aguilera *et al*., 1996). Thus, some research has been focused on adults of *H. elegans* to determine semiochemicals that allow further development of control strategies. For instance, through conventional chemical ecology techniques, Quiroz *et al*. (2007) identified semiochemicals released from virgin conspecific females, which can be involved in the behaviour of the scarab beetle, suggesting 1,4-benzoquinone and 1,4-hydroquinone as putative sex pheromones. Moreover, the authors reported that the essential oils from *Nothofagus obliqua*, its secondary host plant, could have a role in the behaviour of adult *H. elegans*. Based on the above, it has been proposed that *N. obliqua* could serve as a source of putative semiochemicals with attractant or repellent activities on the adult stage of *H. elegans*. However, our previous results suggest that other volatiles could be tuned to part of the olfactory system of the beetle.

The above is based on the only olfactory protein identified for *H. elegans* so far, called odorant-binding protein 1 (OBP1 or HeleOBP1 specifically for *H. elegans*). Located in the antennae of insects and specifically inside of sensilla (olfactory hairs commonly on the surface of antennae), OBPs have provided important information to determine putative semiochemicals. The main function of OBPs includes the transport of odorants from olfactory pores to the dendritic membrane of olfactory receptor neurons (ORNs) (Kaissling, 2013). Likewise, it is believed that these proteins have a key role in the solubilization and protection of odorants (Leal, 2005). Thus, these properties are essential for a selective and sensitive olfactory system as the first filter of olfactory information (Leal, 2003).

Previous results for HeleOBP1 indicate three strong ligands (α-ionone, β-ionone and 2-phenyl acetaldehyde) in terms of binding affinity. Moreover, EAG recordings have shown that these three ligands are active at olfactory receptor level due to notorious electrophysiological activity. However, it is known that a final stage is necessary to conceive these odorants as potential semiochemicals, such as the ligand processing at nervous system level measured in vivo. Therefore, the objective of this study was to determine the olfactometric activity elicited by α-ionone, β-ionone and 2-phenyl acetaldehyde at different doses on adult *H. elegans.*

4.2. Materials and methods

4.2.1. Insect collection

Both male and female of *H. elegans* adults were captured either by trap light or by hand from Regional Research Center INIA-Carillanca in Vilcún, Araucanía, Chile. Beetles were captured during the seasonal flight period, from middle of November 2014 to the beginning of February 2015 according to flight curves. Scarab beetles were recorded daily and taken to the laboratory for further assays.

4.2.2. Y-Tube olfactometric bioassays

Behavioural responses of 20 *H. elegans* males and females to α-ionone, β-ionone and 2-phenyl acetaldehyde were tested by using a two-way olfactometer according to the methodology described by Mutis *et al*. (2010). Three doses, 3.0, 30 and 300 ng equivalents to 0.1, 1.0 and 10 ppm, were used. Each arm of Y-tubes along with the base of the tubes was 16 cm long and 3.0 cm of internal diameter. Both arms of the Y-tube were connected to a glass tube containing a stimulus or control (i.e. hexane) applied on a filter paper (2 cm^2). An aliquot of 50 µL of standard solutions prepared in hexane was loaded onto Whatman N°1 filter paper. The sequence of the odor treatment was randomized. The Y-tube base was connected to a vacuum pump, where the airflow through the arms and stem of the tube was at a rate of 600 mL min^{-1}. One male or female beetle per treatment was

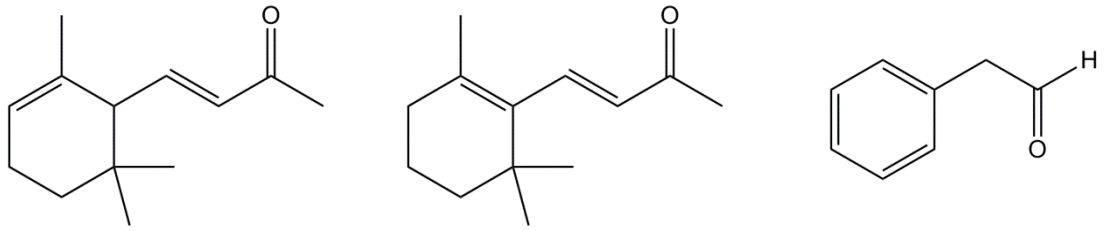
randomly selected and introduced into the base of the Y-tube. Assays were considered successful when beetles walked from the Y-tube base up to 3 cm past the Y-junction and remained there for at least 20 s. Thus, a maximum of 5 min were given to the beetles to respond. After each assay, the Y-tube was replaced with a clean one. The Y-tubes were cleaned using neutral detergent, distilled water and ethanol to remove any organic residue, and finally dried in the oven at 75°C.

4.2.3. Statistical analysis

For the olfactometry bioassay of male and female *H. elegans*, their preference was assessed with the percentage of individuals that selected the stimuli instead of the control (i.e. hexane). No significant difference between sexes was the first hypothesis tested here, where the ratio of males and females that select the stimuli is 1:1. The second hypothesis was to prove that there is no significant difference between doses in each stimulus, where the ratio of individuals, from the same sex, that select one dose is 1:1:1. To determine the above, data were analyzed by the G-test of goodness of fit.

4.3. Results

The statistical analysis carried out for the behavioural responses of *H. elegans* shows how attractive can be the three odorants selected (Figure 4.1) so far through percentage of preference. For instance, Figure 4.2 is indicating that at 3.0, 30 or 300 ng of α-ionone, no significant difference is found among these doses. However, a trend can be highlighted, where females seem to increase their preference as the amount of α-ionone increases. On the other hand, a significant difference was determined for females but not for males, with higher preference at 30 ng than 3.0 ng of β-ionone. For 2-phenyl acetaldehyde, no significant difference was obtained among the doses tested for both males and females. It is worth mentioning that males seem to show an invariable preference for α-ionone and β-ionone throughout the amount of doses used, with a 60% of preference. On the contrary, females seem to show an increase of preference for these terpenoids as the amount of stimulus increases.



α-ionone β-ionone 2-phenyl acetaldehyde

Figure 4.1. Chemical structure of the volatiles selected for behavioural assays.

Comparison between sexes for each dose showed no significant difference for α-ionone and 2-phenyl acetaldehyde at any amount of odorant. However, for β-ionone it seems that males respond with higher preference than females at 3.0 ng. Similarly, males showed higher preference than females at 30 ng.

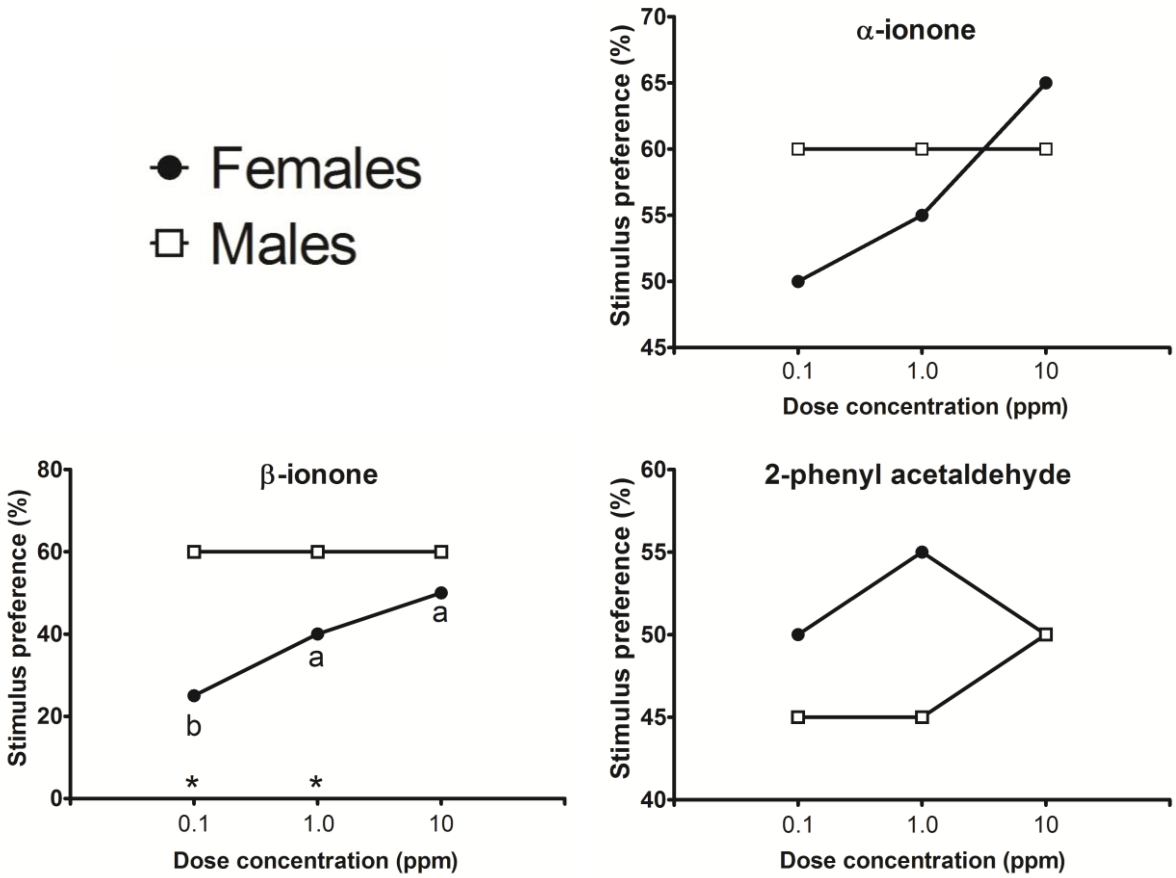


Figure 4.2. Olfactometric responses of *H. elegans* males and females to α-ionone, β-ionone

and 2-phenyl acetaldehyde in Y-tube olfactometer. Data were analyzed by G-test of goodness of fit, where asterisk (*) indicates significant difference between sex for a same dose ($P<0.05$, $N = 20$) and different letters indicate significant difference between doses. Letters are present in those graphs where a significant difference was obtained.

4.4. Discussion

It seems that low concentrations or amounts of stimulus, such as 3.0 or 30 ng, could play a role in the context of either females or males, since low concentrations seem to be closer to reality. For instance, high sensitivity of beetle's antennae has been reported by Hansson *et al*. (1999), where the scarab beetle *Phyllopertha diversa* shows a response threshold of 10 pg to green leaf volatiles (GLVs), such as (*Z*)-3-hexenyl acetate, (*E*)-2-hexenal and (*Z*)-3-hexenol, due to specific ORNs. These findings were supported by the presence of specific ORNs reported for the scarab beetle *Anomala octiescostata* closely related to *H. elegans* (Nikonov *et al*, 2002), in which single-sensillum recordings (SSRs) revealed an ORN specifically tuned to (*Z*)-3-hexenyl acetate. For *H. elegans* case, its distribution of sensilla has been described by Mutis *et al*. (2014), where sensilla closely related to chemoreception (i.e. placodea and coeloconica) is more abundant in males. This is consistent with previous EAG results, where male's antennae were more sensitive to β-ionone at 3.0 ng than females. Nevertheless, behavioural responses showed a low preference for β-ionone at the same amount by females. Although behavioural recordings seem discouraging, probably an age dependency of the scarab beetles played a role in the decision process, since they were taken randomly from *N. obliqua* forest and at different time points during their flight season. For example, early studies on age dependency of behavioural responses have been reported for the honeybee *Apis mellifera*, where 5-day old bees or younger showed the strongest response to the components of queen extracts, such as 9-keto-2-(*E*)-decenoic acid, 9-hydroxy-2-(*E*)-decenoic acid and *p*-hydroxybenzoic acid methyl ester (Pham-Delegue *et al*., 1991). Similarly, Schiestl and Ayasse (2000) reported that cuticle extracts from certain females of the ground-nesting bee *Andrena nigroaenea* were attractive for males. Subsequently, the same authors determined the status of ovaries by dissection, and concluded that attractive females were the youngest and likely virgins. Considering factors

like the mentioned above, it has been proposed that the use of only olfactometer experiments to test insect behaviour could result in a simplified and limited view of the decision process (Ballhorn and Kautz, 2013). Thus, it seems that a set of experiments should be performed to approximate the complex behaviour of insects in nature, such as a feeding test.

From the two terpenoids, β-ionone has been identified as highly attractive for both males and females of the scarab beetle *A. transvaalensis* (Donaldson *et al*. 1990). On the other hand, 2-phenyl acetaldehyde has been reported as part of a chemically active mixture acting as kairomone, where the scarab beetle *A. octiescostata* is attracted to dandelion *Taraxacum officinale* by a mixture of volatiles, such as (*Z*)-3-hexenyl acetate, benzaldehyde, 2-phenyl acetaldehyde and phenyl acetonitrile, among others (Leal *et al*. 1994d). Hitherto, *in silico* and experimental binding assays along with EAG recordings, strongly suggest α-ionone, β-ionone and 2-phenyl acetaldehyde as candidate's semiochemicals. However, it is still unclear where these compounds are present in the context of *H. elegans*. For instance, a special behaviour has been highlighted for this beetle, where after copulation females fly to either grass or cereal crops to deposit their fertilized eggs (Artigas, 1994), being this where the importance of the beetle as underground pest lies. Thus, it is thought that such behaviour could be driven by volatile compounds emitted by crops, suggesting that the three chemicals may play a role at that stage of the life cycle of *H. elegans*, considering that β-ionone and 2-phenyl acetaldehyde have been reported in red clover *Trifolium pratense* (Figueiredo *et al*. 2007) and as the major component of fresh turf *Lolium perenne* (Hong *et al*. 2013), respectively.

Although more evidence is necessary to establish a behavioural role of α-ionone, β-ionone and 2-phenyl acetaldehyde on either *H. elegans* females or males, their electrophysiological activity agrees with previous binding assays on HeleOBP1, showing that these three chemicals are not only transported by this protein, but they are also active at OR level. Finally, our preliminary behavioural results are indicating that β-ionone at 3.0 ng is acting as a moderately attractive odorant for males. Considering these findings plus previous results from reverse chemical ecology approaches, β-ionone could play a key role

at some point of the flight season of males *H. elegans*. However, further experiments should be performed to establish the specific role of this compound.

4.5. Acknowledgements

We would like to thank Dr. Ricardo Ceballos from Instituto de Investigaciones Agropecuarias (INIA), Quilamapu, Laboratorio de Ecología Química, Chillán for his valuable help during the development and analysis of electroantennographic recordings. This work was supported by CONICYT (21110933).

Chapter 5

General discussion, concluding remarks

and further perspectives

5.1. General discussion

In the present thesis the correlation between *in silico* approaches and experimental methods has been evaluated to identify high affinity volatiles and, subsequently, putative semiochemicals. This study was carried out with two approaches: (1) *in silico* approach based on homology modelling, molecular dynamics and molecular docking to elucidate ligands with high binding affinity to an odorant-binding protein (OBP) of *H. elegans*, and (2) experimental methods such as fluorescence binding assays, electroantennography (EAG) and behavioural bioassays to determine the empirical affinity of the selected ligands and their bioactivity on *H. elegans*. Thus, an OBP was identified from antennae of male and female *H. elegans*. The determination of its primary structure (i.e. amino acid sequence) led us to construct a 3D model of the protein. Our findings indicated a protein structure similar to classical reported OBPs, which consists of six α-helices and three disulfide bridges as main features of these insect proteins. Likewise, volatiles as ligands were first trapped and identified from leaves of *N. obliqua*, the main host plant of *H. elegans*, and other candidate compounds were also selected based on their role as semiochemicals for scarab beetles.

Since their discovery, OBPs have been extensively studied showing several properties and mechanisms of ligand binding. The key role that has been given to these proteins makes them essential for olfactory processes in insects, from mate finding to the search of host plants. Despite other proteins are present in the olfactory system (i.e. antennae and sensilla) of insects, such as CSPs, ODEs, SNMPs, IRs and ORs; OBPs are the most studied so far. The main function of OBPs is the transport of hydrophobic molecules (i.e. ligands) across the lymph of olfactory organs in insects. This function has offered the opportunity to discover new semiochemicals, which could be used in pest control strategies. For instance, the properties of six OBPs in the aphids *Acyrthosiphon pisum* and *Myzus persicae* have been studied. These aphids are considered agricultural pests and their alarm pheromone, (*E*)-β-farnesene, a difficult chemical to be synthesized. Therefore, OBPs have been used as targets to discover or design new potent repellents. For these aphids, it seems that OBP3 can bind strongly (*E*)-β-farnesene and differentiate those ligands that could not act as repellents (Qiao *et al*., 2009). Later, the same authors reported that OBP7 could act in

a similar way (Sun *et al*., 2012). Their findings provided the basis to propose other ligands as repellents, such as farnesol, 3,7-dimethyloctyl acetate and geranyl acetate because of their strong binding to OBP3 and OBP7.

The use of chemical ecology to discover semiochemicals has been adopted as a reverse process, which is currently called reverse chemical ecology and highlighted by using OBPs as targets. This approach has been well performed on some insects from Lepidoptera and Diptera. For example, an OBP, called CquiOBP1, was used to identify active volatiles for the mosquito *Culex quinquefasciatus* using conventional techniques and molecular approaches. Thus, in vitro binding assays using CquiOBP1 and gas chromatography-electroantennographic detection (GC-EAD) along with behavioural assays in field suggested that trimethylamine (TMA) and nonanal are a strong attractive blend for mosquitoes (Leal *et al*., 2008). However, less research has considered beetles for such purpose. Hitherto, *H. oblita* has been the most studied scarab beetle in which OBPs and their binding properties have been investigated (Deng *et al*., 2012; Wang *et al*., 2013; Zhuang *et al*., 2014). In Chile, *H. elegans* is a typical beetle belonging to Rutelinae subfamily, which acquires a significant importance due to the damage caused into cereals and grass crops, where it feeds on roots during its larval stage. Likewise, the adult stage of *H. elegans* is characterized by feeding on leaves of trees such as *Nothofagus* species, in special *N. obliqua*. Considering the economic importance of *H. elegans*, its difficult control and the lack of upbringing methods for research, the scarab beetle represents a suitable methodological model for the application of reverse chemical ecology using both computational and experimental techniques.

Our first attempt was the identification of volatiles emitted by leaves of *N. obliqua* due to the typical mating behaviour of not only *H. elegans*, but also other scarab beetles, such as *P. diversa* and *Anomala* species, where host plant volatiles have been suggested to play an important role (Ruther *et al*. 2000; Reinecke *et al*. 2002; Quiroz *et al*. 2007). It has been proposed that males are able to recognize females while they are eating through a sexual kairomone, which seems to be released by the attacked plant. Thus, volatile analysis from *N. obliqua* leaves showed an abundant emission of sesquiterpenes. However, a few

terpenes have been reported as semiochemical for beetles. An example is caryophyllene that acts as attractant for males of the *Harmonia axyridis* (Coccinelidae) beetle (Verheggen *et al*., 2007). On the other hand, a green leaf volatile (GLV) ((*Z*)-3-hexenyl acetate) was identified in *N. obliqua* as a molecule commonly emitted by plants, especially when they have been damaged (Dicke *et al*., 1990; Matsui *et al*., 2012). The presence of these compounds suggests us that the olfactory system of *H. elegans* could have a nervous system tuned with high specificity to some odorants. An example is the high sensitivity to GLVs such as (*Z*)-3-hexenyl acetate, (*E*)-2-hexenal and (*Z*)-3-hexenol of specific ORNs in *P. diversa* with a response threshold of 10 pg for each compound (Hansson *et al*., 1999). Similar findings were reported for *A. octiescostata* by Nikonov *et al*. (2002); in them single sensillum recordings (SSRs) revealed a specific ORN tuned to (*Z*)-3-hexenyl acetate. Therefore, the presence of sesquiterpenes and especially (*Z*)-3-hexenyl acetate in the profile of volatiles emitted by *N. obliqua* provided a suitable start point to test their role in the olfactory system of *H. elegans*. Moreover, we included some other compounds reported in literature that are semiochemicals for other scarab beetles.

A cDNA synthesis using degenerated primers and DNA sequencing revealed the presence of one OBP in antennae of male and female *H. elegans*. The complete sequence of the OBP showed six Cys residues, which are the main feature of insect OBPs. Thus, BLASTp analysis indicated that the OBP, called HeleOBP1, shared around 30% of sequence identity with some crystallized insect OBPs, such as AgamOBP1 and CquiOBP1. These crystal structures were used actually as templates to construct a model for HeleOBP1 by homology modelling. The resulting 3D structure showed three disulfide bridges in a globular arrangement, in which hydrophobic residues were mainly positioned inside the protein. However, some polar residues could be present, such as Thr101, His102, Tyr105 and Tyr113. The presence of the conserved disulfide bridges and six α-helices provides the characteristics to classify HeleOBP1 as a classic OBP, according to Zhou (2010) and Fan *et al*. (2011).

Binding simulations using molecular docking onto the 3D model of HeleOBP1 revealed among the host plant volatiles that some terpenes were able to bind to the protein

(e.g. myrcene and β-linalool) as well as (*Z*)-3-hexenyl acetate. However, no strong interactions such as hydrogen bonds were found. On the other hand, scarab beetle-related compounds showed better interactions than host plant volatiles. For instance, close interactions with Tyr105 and Tyr113 were found for aromatic ligands, such as cinnamyl alcohol, phenylacetonitrile and 2-phenyl acetaldehyde. Similarly, two terpenoids (α-ionone and β-ionone) showed strong interactions, including probable π-π interactions with Tyr105 and Tyr113, and a free binding energy around -6.9 kcal mol^{-1}. It is likely that π-π interactions have an important role when these residues participate. For example, these types of interactions could be formed when Tyr111 in the binding pocket of HoblOBP1 interacts with cinnamaldehyde and β-ionone (Zhuang *et al*. 2014). It is supposed that Tyr113 in HeleOBP1 could play a similar role for the binding of β-ionone considering that the residue is well aligned to the Tyr111 of HoblOBP1. Recent successful examples of our computational reverse chemical ecology approach have shown promising results. An example is the precise prediction of a putative semiochemical for fruit fly *Bactrocera dorsalis* using its GOBP, BdorGOBP. After words, 25 chemicals were screened for their binding to BdorGOBP by molecular docking coupled to molecular dynamics and followed by tryptophan fluorescence quenching (a type of fluorescence binding assay without fluorophore) along with behavioural assays, methyl eugenol was shown as a potent attractant for *B. dorsalis* (P D *et al*., 2014). More recently, the key participation of some residues was determined in the OBP of the moth *Spodoptera litura* (SlitOBP1) using the same computational approach. It was reported that Phe12 and Ile52 play an important role during the binding of several compounds to SlitOBP1, which was corroborated after potential hydrophobic and π-π interactions were interrupted when Ile52 and Phe12 were mutated to Ala, respectively (Yi *et al*., 2015). Likewise, the same authors reported a consistency between molecular docking and experimental binding assays, in which certain functional groups such as ketones and aliphatic compounds showed a similar binding profile in both *in silico* and experimental methods.

In an attempt to elucidate how correlated the *in silico* data with empirical experiments are, we performed competitive binding assays by fluorescence. Thus, the used recombinant HeleOBP1 showed a broad specificity from host plant volatiles to scarab

beetle-related compounds. None of the host plant volatiles nor putative sex pheromones could strongly displace the fluorophore (1-NPN) used in the binding assays. The low binding affinity to the volatiles tested suggested that HeleOBP1 could not be specific or its true ligand(s) is(are) not present in the profile of volatiles released by *N. obliqua*. For example, 15 OBPs have been identified from genome of the aphid *A. pisum* (Zhou *et al*., 2010). However, it has been found that only OBP3 and OBP7 can specifically bind the alarm pheromone (*E*)-β-farnesene, whereas the rest of the proteins show broad selectivity (Qiao *et al*., 2009; Sun *et al*., 2012). In scarab beetles, a limited number of OBPs have been identified, and the only Coleopteran genome known to date comes from *T. Castaneum* indicating the presence of up to 50 OBPs. Nevertheless, construction of a full-length cDNA library for the scarab beetle *H. oblita* revealed 4 OBPs, from which OBP3 showed more specific binding affinities than OBP4 (Wang *et al*., 2013). Surprisingly, HeleOBP1 showed some selectivity for 2-phenyl acetaldehyde, α-ionone and β-ionone with binding dissociation constants in the micromolar range, which is usually the range where OBPs bind molecules (Zhou, 2010). When molecular docking findings were matched with these three ligands, a good correlation was found. Thus, the *in silico* method successfully rendered the difference between the two terpenoids and 2-phenyl acetaldehyde showing a more stable ligand-HeleOBP1 complex. Taking this into account, subsequent studies were focused on these three ligands.

The main findings from EAG and behavioural assays comprise the role that β-ionone could play at low concentrations. For instance, high sensitivity of scarab beetles antennae has been reported for *P. diversa* and *A. octiescostata* (Hansson *et al*., 1999; Nikonov *et al*., 2002). For *H. elegans* case, its distribution of sensilla has been described by Mutis *et al*. (2014), where sensilla closely related to chemoreception (i.e. placodea and coeloconica) is more abundant in males. This is consistent with our EAG results, where male's antennae were more sensitive to β-ionone at 3.0 ng than females. Likewise, behavioural responses showed a moderate preference for β-ionone by males, which is contrary to the low preference shown by females. It is worth mentioning that an age dependency could play a role in the decision process, since they were taken randomly from *N. obliqua* forest and at different time points during their flight season. For example, early

studies on the honeybee *Apis mellifera* revealed that 5-day old bees or younger can respond stronger to the components of queen extracts, such as 9-keto-2-(*E*)-decenoic acid, 9-hydroxy-2-(*E*)-decenoic acid and *p*-hydroxybenzoic acid methyl ester than older bees (Pham-Delegue *et al*., 1991). Similarly, it has been reported that cuticle extracts from the youngest and likely virgin females of the ground-nesting bee *Andrena nigroaenea* are attractive for males (Schiestl and Ayasse, 2000). Despite the mentioned above, a promising correlation among molecular docking, fluorescence binding assays, EAG and behavioural recordings suggest β-ionone as a potential semiochemical for *H. elegans*.

Although 2-phenyl acetaldehyde represents an interesting molecule to be studied considering its binding to HeleOBP1 and abundant presence in fresh turf *Lolium perenne* (Hong *et al*., 2013), experimental techniques such as EAG and behavioural assays suggested us not to consider this molecule as relevant for *H. elegans*. Something similar could be concluded for the analogue α-ionone. Instead our findings led us to focus the attention to β-ionone. In plants, this molecule is characterized as a carotenoid cleavage product, which acts as an aromatic compound (Walter *et al*., 2010). The role of this terpenoid as semiochemical has been reported for several insects from Hymenoptera, Diptera and Coleoptera, in which the molecule acts mainly as attractant, but also as kairomone and pheromone. For instance, β-ionone was reported as attractant for the American leafminer *Liriomyza sativae* (Wei *et al*., 2005). Likewise, β-ionone was reported as attractant for the orchid bee *Euglossa mandibularis*, where males are highly attracted by this terpenoid (Faria and Zanella, 2015). Another study have revealed that a derivative of β-ionone, dihydro-β-ionone, acts as a kairomone for the fly *Rhagoletis pomonella*, which along with 3-methylbutan-1-ol, 4,8-dimethyl-1,3(*E*),7-nonatriene and butyl hexanoate form a blend from the host plant *Catraegus spp*. that elicits high levels of upwind flight (Nojima *et al*., 2003). For Coleoptera, early research on the leaf beetle *Diabrotica cristata* showed β-ionone as an attractant compound (Lampman and Metcalf, 1988). Similarly, β-ionone has been identified as highly attractive for both males and females of the scarab beetle *A. transvaalensis* (Donaldson *et al*. 1990). Finally, the potential roles of β-ionone and HeleOBP1 suggested by this thesis are summarized in a scheme proposed in Figure 5.1.

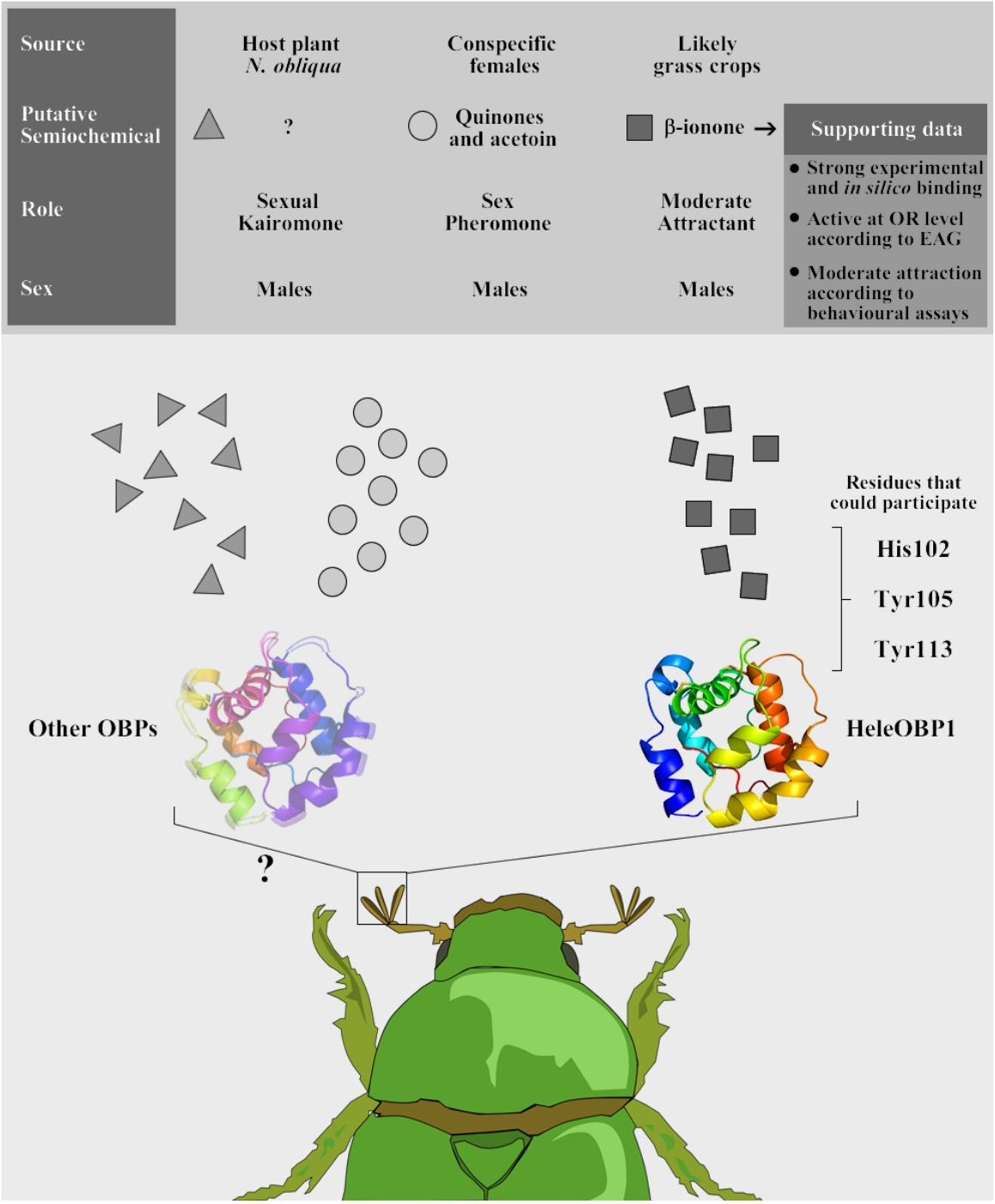


Figure 5.1. Proposed scheme for the role of β-ionone, HeleOBP1 and potential other OBPs in the olfactory process of *H. elegans*. Question mark (?) indicates identification and participation of OBPs and volatiles not determined yet. Volatiles from *N. obliqua*, conspecific females and likely grass crops are indicated as triangles, circles and squares, respectively.

5.2. Concluding remarks

For this work, experimental and *in silico* approaches were performed with the aim of developing a computational reverse chemical ecology method. The suitable complement of experimental techniques by *in silico* methods led us to study three specific candidates to semiochemical. Despite 2-phenyl acetaldehyde and α-ionone showed good binding to HeleOBP1 and appreciable activity by EAG, our efforts were focused on β-ionone due to better results in terms of binding and bioactivity.

As a typical aromatic compound from secondary metabolism in plants, it is probable that β-ionone plays a specific role during the flight season of males *H. elegans* as this thesis proposes. However, it is still unclear where this chemical is present in the context of *H. elegans* some studies suggests that red clover *T. pratense* could be a source of β-ionone. Considering this, a contradictory process is highlighted. It has been proposed that females fly to crops to deposit their fertilized eggs after copulation in *N. obliqua* trees. Nevertheless, our results suggest no preference for β-ionone at low concentration by females, whereas males have shown a major sensitivity to the chemical. Therefore, an overall state in the adulthood of males *H. elegans* can be summarized in such a way that a certain number of other OBPs could be tuned to both host plant volatiles driving the flight of males to *N. obliqua* and sex pheromones that allow copulation. On the other hand, already identified and characterized HeleOBP1, could be aimed to the detection of volatiles from secondary plants, such as *T. pratense*, in the way that β-ionone-sensitive males can direct its flight to *N. obliqua* trees and not to crops where gravid females should be found.

In relation to females, their flight to grass crops even with a probable presence of β-ionone could be explained because they are sensitive not only to volatiles, but also to other factors of host plants, such as non-volatile secondary metabolites, surface structure, tissue toughness and water content. This could ensure a host plant with high nutritive value for females and their offspring. Similar to males, other OBPs tuned to host plant volatiles could be present in the olfactory system of females since these beetles feed on leaves of *N. obliqua*.

Although further experiments are necessary to determine the role of the semiochemical-candidate compound found in this work, the promising correlation between experimental and *in silico* results provides a strong starting point for developing computational reverse chemical ecology. This method was able to show that the hypothesis raised in this work is true for HeleOBP1, which showed selectivity for a compound reported as semiochemical for other scarab beetles of Rutelinae subfamily. On the other hand, these results suggest that HeleOBP1 does not participate in host selection or mate finding through a sexual kairomone since none of the tested host plant volatiles showed strong affinity. Likewise, the still unknown sex pheromone for *H. elegans* makes the role of HeleOBP1 as a putative PBP difficult to predict. Finally, our research and the reported literature strongly suggest that this approach can help in the search of new potent semiochemicals for pest control.

5.3. Further perspectives

In Chile, *H. elegans* is highlighted as an important underground pest. Its control has considered biological and chemical methods. However, the lack of efficiency and specificity of these methods has led to the search of other strategies. Chemical ecology studies resulted in only one significant attractive blend for males of *H. elegans* (i.e. 1,4-benzoquinone and essential oil from *N. obliqua*). Despite the mentioned above, neither pheromone nor kairomone have been identified yet. Therefore, further experiments should be focused on the identification of the supposed highly attractive semiochemicals, such as a sexual kairomone or sex pheromone considering the complement between conventional and computational reverse chemical ecology.

Although the results presented here indicate *in silico* approaches as suitable tools in reverse chemical ecology, some limitations have been described specifically for homology modelling. For *H. elegans* case, the lack of a crystallized OBP structure from beetle make the construction of a 3D model a difficult task. Once an experimental 3D structure is solved, a higher sequence identity (i.e. over 40%) will allow the construction of more

precise models. On the other hand, the incorporation of site-directed mutagenesis could provide both an accurate validation of 3D protein models and the role of residues in the binding site of proteins, such as His102, Tyr105 and Tyr113 in HeleOBP1. Thus, our results suggest that while correct 3D protein models are generated, a large number of candidate compounds could be tested *in silico*. Such approximation would provide from a large range of putative semiochemicals, a few amount of them to be experimentally tested with the subsequent saving in both time and resources. Therefore, computational reverse chemical ecology could serve as the first screening prior olfactometry, wind tunnel and field assays, where bioactive compounds can be elucidated and used for pest control with the design of specific traps.

References

Affonso, R., Guimarães, A.P., Oliveira, A.A., Slana, G.B.C. and França, T.C.C. 2013. Applications of molecular modelling in the design of new insect repellents targeting the odorant binding protein of *Anopheles gambiae*. J. Brazil. Chem. Soc. 24, 473-482.

Aguilera, A., Cisternas, E., Gerding, M., and Norambuena, H. 1996. Plagas de las praderas, in: Ruiz, I. (ed.), Praderas para Chile. Santiago, Chile, Offset, pp. 309 – 339.

Aguilera, A. 2007. Plagas de la murta o murtilla: Sistemas de control. INIA Tierra Adentro. 72, 10-14.

Allen, J.E. and Wanner, K.W. 2011. Asian corn borer pheromone binding protein 3, a candidate for evolving specificity to the 12-tetradecenyl acetate sex pheromone. Insect Biochem. Molec. 41, 141-149.

Alonso, H., Bliznyuk, A.A. and Gready, J.E. 2006. Combinding docking and molecular dynamic simulations in drug design. Med. Res. Rev. 26, 531-568.

Alzari, P.M., Berglund, H., Berrow, N.S., Blagova, E., Busso, D., Cambillau, C., Campanacci, V., Christodoulou, E., Eiler, S., Fogg, M.J., Folkers, G., Geerloi, A., Hart, D., Haouz, A., Herman, M.D., Macieira, S., Nordlund, P., Perrakis, A., Quevillon-Cheruel, S., Tarandeau, F., van Tilbeurgh, H., Unger, T., Luna-Vargas, M.P.A., Velarde, M., Willmanns, M., Owens, R.J. 2006. Implementation of semi-automated cloning and prokaryotic expression screening: the impact of SPINE. Acta Crystallogr. Sect. D-Biol. Crystallogr. 62, 1103-1113.

Arnold K., Bordoli L., Kopp J. and Schwede, T. 2006. The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling. Bioinformatics. 22, 195-201.

Artigas, J. 1994. *Hylamorpha elegans* (Burmeister). Pololo San Juan verde (San juan green beetle), in: Universidad de Concepción (eds.), Entomología económica. Insectos de interés agrícola, forestal, médico y veterinario. Editorial Universidad de Concepción, Concepción, Chile, pp. 344 – 346.

Aslanidis, C., de Jong, P.J. 1990. Ligation-independent cloning of PCR products (LIC-PCR). Nucleic Acids Res. 18, 6069-6074.

Ballhorn, D.J. and Kautz, S. 2013. How useful are olfactometer experiments in chemical ecology research?. Commun. Integr. Biol. 6, e24787.

Ban, L., Scaloni, A., D'Ambrosio, C., Zhang, L., Yan, Y. and Pelosi, P. 2003. Biochemical characterization and bacterial expression of an odorant-binding protein from *Locusta migratoria*. Cell. Mol. Life Sci. 60, 390-400.

Bembenek, S.D., Tounge, B.A. and Reynolds, C.H. 2009. Ligand efficiency and fragment-based drug discovery. Drug Discov. Today. 14, 278-283.

Berendsen, H.J.C., van der Spoel, D. and van Drunen, R. 1995. GROMACS: A message-passing parallel molecular dynamics implementation. Comput. Phys. Commun. 91, 43-56.

Bette, S., Breer, H. and Krieger, J. 2002. Probing a pheromone binding protein of the silkmoth *Antheraea polyphemus* by endogenous tryptophan fluorescence. Insect Biochem. Moles. 32, 241-246.

Biessmann, H., Andronopoulou, E., Biessmann, M.R., Douris, V., Dimitratos, S.D., Eliopoulos, E., Guerin, P.M., Iatrou, K., Justice, R.W., Kröber, T., Marinotti, O., Tsitoura, P., Woods, D.F. and Walter, M.F. 2010. The *Anopheles gambiae* odorant binding protein 1 (AgamOBP1) mediates indole recognition in the antennae of female mosquitoes. PLoS ONE, 5, e9471. doi:10.1371/journal.pone.0009471.

Bordner, A.J. 2012. Force fields for homology modelling. *Homology Modelling: Methods* and *Protocols* (ed. by A.J.W. Orry and R. Abagyan), pp. 83-106. Molsoft L.L.C., San Diego.

Bordoli, L. and Schwede, T. 2012. Automated protein structure modelling with SWISS-MODEL Workspace and the protein model portal. *Homology Modelling: Methods* and *Protocols* (ed. by A.J.W. Orry and R. Abagyan), pp. 107-136. Molsoft L.L.C., San Diego.

Briand, L., Eloit, C., Nespoulous, C., Bézirard, V., Huet, J.-C., Henry, C., Blon, F., Trotier, D. and Pernollet, J.-C. 2002. Evidence of an odorant-binding protein in the human olfactory mucus: Location, structural characterization, and odorant-binding properties. Biochemistry-US. 41, 7241-7252.

Brimau, F., Cornard, J.-P., Le Danvie, C., Lagant, P., Vergoten, G., Grebert, D., Pajot, E. and Meillour, P.N.-L. 2010. Binding specificity of recombinante odorant-binding protein isoforms is driven by phosphorylation. J. Chem. Ecol. 36, 801-813.

Brooks, B.R., Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S. and Karplus, M. 1983. CHARMM: a program for macromolecular energy minimization and dynamics calculations. J. Comput. Chem. 4, 187-217.

Campanacci, V., Krieger, J., Bette, S., Sturgis, J.N., Lartigue, A., Cambillau, C., Breer, H. and Tegoni, M. 2001. Revisiting the specificity of *Mamestra brassicae* and *Antheraea polyphemus* pheromone-binding proteins with a fluorescence binding assay. J. Biol. Chem. 276, 20078-20048.

Cavasotto, C.N. and Phatak, S.S. 2009. Homology modelling in drug discovery: current trends and applications. Drug Dicov. Today. 14, 676-683.

Chang, C. and Swaan, P.W. 2006. Computational approaches to modelling drug transporters. Eur. J. Pharm. Sci. 27, 411-424.

Chu, W-.T., Zhang, J-.L., Zheng, Q-.C., Chen, L., Wu, Y-.J., Xue, Q. and Zhang, H-.X. 2013. Constant pH molecular dynamics (CpHMD) and molecular docking studies of CquiOBP1 pH-induced ligand releasing mechanism. J. Mol. Model. 19, 1301-1309.

Damberger, F.F., Ishida, Y., Leal, W.S. and Wüthrich, K. 2007. Structural basis of ligand binding and release in insect pheromone-binding proteins: NMR structure of *Antheraea polyphemus* PBP1 at pH 4.5. J. Mol. Biol. 373, 811-819.

Danty, J., and Olfos, M. 2012. Boletín de Trigo: Febrero 2012. ODEPA, Ministerio de Agricultura, Gobierno de Chile.

de Jong, L.A.A., Uges, D.R.A., Franke, J.P. and Bischoff, R. 2005. Receptor-ligand binding assays: Technologies and applications. J. Chromatogr. 829, 1-25.

Deng S., Yin J., Zhong T., Cao Y., Li K. 2012. Function and immunocytochemical localization of two novel odorant-binding proteins in olfactory sensilla of the scarab beetle *Holotrichia oblita* Faldermann (Coleoptera: Scarabaeidae). Chem. Senses. 37, 141-150.

Dicke, M., Sabelis, M.W., Takabayashi, J., Bruin, J. and Posthumus, M.A. 1990. Plant strategies of manipulating predator-prey interactions through allelochemicals: Prospects for application in pest control. J. Chem. Ecol. 16: 3091 – 3118.

Donaldson, J.M.I., McGovern, T.P. and Ladd, Jr.T.L. 1990. Floral attractants for Cetoniinae and Rutelinae (Coleoptera: Scarabaeidae). J. Econ. Entomol. 83, 1298-1305.

Dundas J., Ouyang Z., Tseng J., Binkowski A., Turpaz Y., Liang J. 2006. CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. Nucleic Acids Res. 34, W116-8.

Durrant, J.D. and McCammon, J.A. 2011. Molecular dynamics simulations and drug discovery. BMC Biol. 9, DOI:10.1186/1741-7007-9-71.

Edgar, R.C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792-1797.

Eswar, N., Marti-Renom, M.A., Webb, B., Madhusudhan, M.S., Eramian, D., Shen, M., Pieper, U. and Sali, A. 2006. Comparative protein structure modelling with MODELLER. *Current Protocols in Bioinformatics* (ed. by A. Baxevanis), pp. 5.6.1-5.6.30. John Wiley and Sons, Inc., New York.

Fan, H. and Mark, A.E. 2004. Refinement of homology-based protein structures by molecular dynamics simulation techniques. Prot. Sci. 13, 211-220.

Fan J., Francis F., Liu Y., Chen J.L., Cheng D.F. 2011. An overview of odorant-binding protein functions in insect peripheral olfactory reception. Genet. Mol. Res. 10, 3056-3069.

Faria, L.R.R. and Zanella, F.C.V. 2015. Beta-ionone attracts *Euglossa mandibularis* (Hymenoptera, Apidae) males in western Paraná forests. Rev. Bras. Entomol. doi:10.1016/j.rbe.2015.05.003.

Fernandez P., Hilker M. 2007. Host plant location by Chrysomelidae. Basic Appl. Ecol. 8, 97-116.

Figueiredo R., Rodrigues A.I., do Céu Costa M. 2007. Volatile composition of red clover (*Trifolium pratense* L.) forages in Portugal: The influence of ripening stage and ensilage. Food Chem. 104, 1445-1453.

Gajendrarao, P., Krishnamoorthy, N., Sakkiah, S., Lazar, P. and Lee, K.W. 2010. Molecular modelling study on orphan human protein CYP4A22 for identification of potential ligand binding site. J. Mol. Graph. Model. 28, 524-532.

Giganti H., and Dapoto G. 1990. Coleópteros de los bosques nativos del Departamento Aluminé (Neuquén-Argentina). Bosque. 11, 37-44.

Ginalski, K. 2006. Comparative modelling for protein structure prediction. Curr. Opin. Struc.l Biol. 16, 172-177.

Gomez-Diaz, C., Reina, J.H., Cambillau, C. and Benton, R. 2013. Ligands for pheromone-sensing neurons are not conformationally activated odorant binding proteins. PLoS BIOL. 11, e1001546. doi:10.1371/journal.pbio.1001546.

Gong, Y., Pace, T.C.S., Castillo, C., Bohne, C., O'Neill, M. and Plettner, E. 2009. Ligand-interaction kinetics of the pheromone-binding protein from the gypsy moth, *L. dispar*: Insights into the mechanism of binding and release. Chem. Biol. 16, 162-172.

Gong, Y., Tang, H., Bohne, C. and Plettner, E. 2010. Binding conformation and kinetics of two pheromone-binding proteins from the gypsy moth *Lymantria dispar* with biological and nonbiological ligands. Biochemistry-US. 49, 793-801.

Gräter, F., de Groot, B.L., Jiang, H. and Grubmüller, H. 2006. Ligand-release pathways in the pheromone-binding protein of *Bombyx mori*. Structure. 14, 1567-1576.

Gräter, F., Xu, W., Leal, W. and Grubmüller, H. 2006. Pheromone discrimination by the pheromone-binding protein of *Bombyx mori*. Structure. 14, 1577-1586.

Grolli, S., Merli, E., Conti, V., Scaltriti, E. and Ramoni, R. 2006. Odorant binding protein has the biochemical properties of a scavenger for 4-hydroxy-2-nonenal in mammalian nasal mucosa. FEBS Journal. 273, 5131-5142.

Große-Wilde, E., Svatoš, A. and Krieger, J. 2006. A pheromone-binding protein mediates the bombykol-induced activation of a pheromone receptor *in vitro*. Chem. Senses. 31, 547-555.

Gu, S., Wang, W., Wang, G., Zhang, X., Guo, Y., Zhang, Z., Zhou, J.-J. and Zhang, Y. 2011. Functional characterization and immunolocalization of odorant binding protein 1 in the Lucerne plant bug, *Adelphocoris lineolatus* (Goeze). Arch. Insect Biochem. Physiol. 77, 81-98.

Guo, H., Huang, L., Pelosi, P. and Wang, C. 2012. Three pheromone-binding proteins help segregation between two *Helicoverpa* species utilizing the same pheromone components. Insect Biochem. Mol. Biol. 42, 708-716.

Hang, N.T. 2008. Comparison of multiple sequence alignment programs in practise. A thesis submitted to The Bioinformatics Research Center (BiRC), University of Århus. Requirements for the degree of Master of Bioinformatics.

Hansen, S.B., Sulzenbacher, G., Huxford, T., Marchot, P., Taylor, P., and Bourne, Y. 2005. Structures of *Aplysia* AChBP complexes with nicotinic agonists and antagonists reveal distinctive binding interfaces and conformations. The EMBO Journal. 24, 3635-3646.

Hansson, B.S., Larsson, M.C. and Leal, W.S. 1999. Green leaf volatile-detecting olfactory receptor neurons display very high sensitivity and specificity in a scarab beetle. Physiol. Entomol. 24, 121-126.

He, X., Tzotzos, G., Woodcock, C., Pickett, J.A., Hooper, T., Field, L.M. and Zhou, J.-J. 2010. Binding of the general odorant binding protein of *Bombyx mori* BmorGOBP2 to the moth sex pheromone components. J. Chem. Ecol. 36, 1293-1305.

Hekmat-Scafe, D.S., Scafe, C.R., McKinney, A.J. and Tanouye, M.A. 2002. Genome-wide analysis of the odorant-binding protein gene family in *Drosophila melanogaster*. Genome Res. 12, 1357-1369.

Hong, S., Kang, W., Su, Y. and Guo, Y. 2013. Analysis of trace-level volatile compounds in fresh turf crop (*Lolium perenne* L.) by gas chromatography quadrupole time-of-flight mass spectrometry. Chin. J. Chem. 31, 1329-1335.

Honson, N., Johnson, M.A., Oliver, J.E., Prestwich, G.D. and Plettner, E. 2003. Structure-activity studies with pheromone-binding proteins of the gypsy moth, *Lymantria dispar*. Chem. Senses. 28, 479-489.

Honson, N., Gong, Y. and Plettner, E. 2005. Structure and function of insect odorant and pheromone-binding proteins (OBPs and PBPs) and chemosensory-specific proteins (CSPs). Recent Adv. Phytochem. 39, 268-277.

Honson, N. and Plettner, E. 2006. Disulfide connectivity and reduction in pheromone-binding proteins of the gypsy moth, *Lymantria dispar*. Naturwissenschaften. 93, 267-277.

Hooper, A.M., Dufour, S., He, X., Muck, A., Zhou, J.-J., Almeida, R., Field, L.M., Svatoš, A. and Pickett, J. 2009. High-throughput ESI-MS analysis of binding between the *Bombyx mori* pheromone-binding protein BmorPBP1, its pheromone components and some analogues. Chem. Commun. 14, 5725-5727.

Horst, R., Damberger, F., Luginbühl, P., Güntert, P., Peng, G., Nikonova, L., Leal, W.S. and Wüthrich, K. 2001. NMR structure reveals intramolecular regulation mechanism for pheromone binding and release. PNAS. 98, 14374-14379.

Ishida Y., Chiang V., Haverty M. and Leal W.S. 2002. Odorant-binding protein from a primitive termite. J. Chem. Ecol. 28, 1887-1893.

Ishida, Y. and Leal, W.S. 2002. Cloning of putative odorant-degrading enzyme and integumental esterase cDNAs from the wild silkmoth, *Antheraea polyphemus*. Insect Biochem. Mol. Biol. 32, 1775-1780.

Iturriaga-Vásquez, P., Carbone, A., García-Beltrán, O., Livingstone, P.D., Biggin, P.C., Cassels, B.K., Wonnacott, S., Zapata-Torres, G. and Bermudez, I. 2010. Molecular determinants for competitive Inhibition of α4β2 nicotinic acetylcholine receptors. Mol. Pharm. 78, 366-375.

Jacquin-Joly, E., Vogt, R.G., Francois, M.C. and Nagnan-Le Meillour, P. 2001. Functional and expression pattern analysis of chemosensory proteins expressed in antennae and pheromonal gland of *Mamestra brassicae*. Chem. Senses. 26, 833-844.

Jacquin-Joly, E. and Merlin, C. 2004. Insects olfactory receptors: contributions of molecular biology to chemical ecology. J. Chem. Ecol. 30, 2359-2397.

Jiang, Q., Wang, W., Zhang, Z. and Zhang, L. 2009. Binding specificity of locust odorant binding protein and its key binding site for initial recognition of alcohols. Insect Biochem. Mol. Biol. 39, 440-447.

Kaissling, K.-E. 2013. Kinetics of olfactory responses might largely depend on the odorant-receptor interaction and the odorant deactivation postulated for flux detectors. J. Comp. Physiol. A. DOI 10.1007/s00359-013-0812-z.

Katre, U.V., Mazumder, S., Prusti, R.K. and Mohanty, S. 2009. Ligand binding turns moth pheromone-binding protein into a pH sensor. J. Biol. Chem. 284, 32167-32177.

Klein, U. 1987. Sensillum-lymph proteins from antennal olfactory hairs of the moth *Antheraea polyphemus* (Saturniidae). Insect Biochem. 17, 1193-1204.

Klein C. and Waterhouse D.F. 2000. The distribution and importance of arthropods associated with agriculture and forestry in Chile. ACIAR Monograph, Canberra.

Klusak, V., Havlas, Z., Rulíšek, L., Vondrášek, J. and Svatoš, A. 2003. Sexual attraction in the silkworm moth: Nature of binding of bombykol in pheromone binding protein-an ab initio study. Chem. Biol. 10, 331-340.

Kolb, P., Ferreira, R.S., Irwin, J.J. and Shoichet, B.K. 2009. Docking and chemoinformatic screens for new ligands and targets. Curr. Opin. Biotech. 20, 429-436.

Kowcun, A., Honson, N. and Plettner, E. 2001. Olfaction in the gypsy moth, *Lymantria dispar*: Effect of pH, ionic strength, and reductants on pheromone transport by pheromone-binding proteins. J. Biol. Chem. 276, 44770-44776.

Krieger, J., von Nickisch-Rosenegk, E., Mameli, M., Pelosi, P. and Breer, H. 1996. Binding proteins from the antennae of *Bombyx mori*. Insect Biochem. Mol. Biol. 26, 297-307.

Kruse, S.W., Zhao, R., Smith, D.P. and Jones, D.N.M. 2003. Structure of a specific alcohol-binding site defined by the odorant binding protein LUSH from *Drosophila melanogaster*. Nat. Struc. Biol. 10, 694-700.

Lagarde, A., Spinelli, S., Tegoni, M., He, X., Field, L., Zhou, J.-J. and Cambillau, C. 2011. The crystal structure of odorant binding protein 7 from *Anopheles gambiae* exhibits an outstanding adaptability of its binding site. J. Mol. Biol. 414, 401-412.

Lampman, R.L. and Metcalf, R.L. 1988. The comparative response of *Diabrotica* species (Coleoptera: Chrysomelidae) to volatile attractants. Environ. Entomol. 17, 644-648.

Lanfranco D., Rojas E., Ríos R. and Ruíz C. 2001. Insect Defoliators of *Nothofagus obliqua* (Roble) in South Chile: Two Years Monitoring Species and Their Damage. In: Liebhold AM, McManus ML, Otvos IS, Fosbroke SLC (eds) *Proceedings: integrated*

management and *dynamics of forest defoliating insects*, pp. 91-103. U.S. Department of Agriculture, Forest Service, Northeastern Research Station.

Larsson, P., Wallner, B., Lindahl, E. and Elofsson, A. 2008. Using multiple templates to improve quality of homology models in automated homology modelling. Prot. Sci. 17, 990-1002.

Lartigue, A., Gruez, A., Spinelli, S., Rivière, S., Brossut, R., Tegoni, M. and Cambillau, C. 2003. The crystal structure of a cockroach pheromone-binding protein suggests a new ligand binding and release mechanism. J. Biol. Chem. 278, 30213-30218.

Laskowski R.A., MacArthur M.W., Moss D.S. and Thornton J.M. 1993. PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. 26, 283-291.

Laughlin, J.D., Ha, T.S., Jones, D.N.M. and Smith, P. 2008. Activation of pheromone-sensitive neurons is mediated by conformational activation of pheromone-binding protein. Cell. 133, 1255-1265.

Lautenschlager, C., Leal, W.S. and Clardy, J. 2005. Coil-to-helix transition and ligand release of *Bombyx mori* pheromone-binding protein. Biochem. Biophys. Res. Commun. 335, 1044-1050.

Leach, A.R. 2001. Protein structure prediction, sequence analysis and protein folding. *Molecular modelling principles* and *applications* (ed. by A.R. Leach), pp. 509-558. Pearson Education Limited, Edinburgh Gate, England.

Leal, W.S., Ono, M., Hasegawa, M., Sawada, M. 1994d. Kairomone from dandelion, *Taraxacum officinale*, attractant for scarab beetle *Anomala octiescostata*. J. Chem. Ecol. 20, 1697-1704.

Leal WS. 1998. Chemical ecology of phytophagous scarab beetles. Annu. Rev. Entomol. 43, 39-61.

Leal, W.S. 2003. Proteins that make sense. *Insect Pheromone Biochemistry* and *Molecular Biology* (ed. by G.J. Blomquist and R.G. Vogt), pp. 447-476. Elsevier, Oxford.

Leal, W.S. 2005. Pheromone Reception. Top. Curr. Chem. 240, 1-36.

Leal, W.S., Barbosa, R.M.R., Xu, W., Ishida, Y., Syed, Z., Latte, N., Chen, A.M., Morgan, T.I., Cornel, A.J., and Furtado, A. 2008. Reverse and Conventional Chemical Ecology Approaches for the Development of Oviposition Attractants for *Culex* Mosquitoes. PLoS ONE, 3: doi:10.1371/journal.pone.0003045.

Leal, W.S. 2013. Odorant reception in insects: roles of receptors, binding proteins, and degrading enzymes. Annu. Rev. Entomol. 58, 373-391.

Leite, N.R., Krogh, R., Xu, W., Ishida, Y., Lulek, J., Leal, W.S. and Olive, G. 2009. Structure of an odorant-binding protein from the mosquito *Aedes aegypti* suggests a binding pocket covered by a pH-sensitive "Lid". PLoS ONE. 4, e8006. doi:10.1371/journal.pone.0008006.

Lescop, E., Briand, L., Pernollet, J. and Guittet, E. 2009. Structural basis of the broad specificity of a general odorant-binding protein from Honeybee. Biochemistry-US. 48, 2431-2441.

Levinthal, C. 1966. Molecular model-building by computer. Scientific American. 214, 42-52.

Li, X., Lu, D., Liu, X., Zhang, Q. and Zhou, X. 2011. Ultrastructural characterization of olfactory sensilla and immunolocalization of odorant binding and chemosensory proteins from an ectoparasitoid *Scleroderma guani* (Hymenoptera: Bethylidae). Int. J. Biol. Sci. 7, 848-868.

Li, H-.L., Zhang, L-.Y., Zhuang, S-.L., Ni, C-.X. and Shang, H-.W. 2013. Fluorescence investigation on the interaction of a prevalent competitive fluorescent probe with entomic odorant binding protein. Spectrosc. Lett. 46, 527-534.

Liu, N., He, P. and Dong, S. 2012. Binding properties of pheromone-binding protein 1 from the common cutworm *Spodoptera litura*.Comp. Biochem. Phys. B. 161, 295-302.

Löbel, D., Jacob, M., Völkner, M. and Breer, H. 2002. Odorants of different chemical classes interact with distinct odorant binding protein subtypes. Chem. Senses. 27, 39-44.

Maida, R., Steinbrecht, A., Ziegelberger, G. and Pelosi, P. 1993. The pheromone binding protein of *Bombyx mori*: Purification, characterization and immunocytochemical localization. Insect Biochem. Mol. Biol. 23, 243-253.

Maida, R., Krieger, J., Gebauer, T., Lange, U. and Ziegelberger, G. 2000. Three pheromone-binding proteins in olfactory sensilla of the two silkmoth species *Antheraea polyphemus* and *Antheraea pernyi*. Eur. J. Biochem. 267, 2899-2908.

Maida, R., Ziegelberger, G. and Kaissling, K.-E. 2003. Ligand binding to six recombinant pheromone-binding proteins of *Antheraea polyphemus* and *Antheraea pernyi*. J. Comp. Phys. B. 173, 565-573.

Mangoni, M., Roccatano, D. and Di Nola, A. 1999. Docking of flexible ligands to flexible receptors in solution by molecular dynamics simulation. Proteins. 35, 153-162.

Martí-Renom, M.A., Stuart, A.C., Fiser, A., Sánchez, R., Melo, F. and Šali, A. 2000. Comparative protein structure modelling of genes and genomes. Annu. Rev. Biophys. Biomol. Struc. 29, 291-325.

Matsui, K., Sugimoto, K., Mano, J., Ozawa, R. and Takabayashi, J. 2012. Differential metabolisms of green leaf volatiles in injured and intact parts of a wounded leaf meet distinct ecophysiological requirements. PLoS ONE. 7: e36433. doi:10.1371/journal.pone.0036433.

McCammon, J.A., Gelin, B.R. and Karplus, M. 1977. Dynamics of folded proteins. Nature. 267, 585-590.

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N. and Teller, A.H. 1953. Equation of state calculations by fast computing machines. J. Chem. Phys. 21, 1087-1092.

Morris, G.M., Goodsell, D.S., Halliday, R.S., Huey, R., Hart, W.E., Belew, R.K. and Olson, A.J. 1998. Automated docking using a lamarckian genetic algorithm and an empirical binding free energy function. J. Comput. Chem. 19, 1639-1662.

Morris G.M., Huey R., Lindstrom W., Sanner M.F., Belew R.K., Goodsell D.S. and Olson AJ. 2009. Autodock 4 and AutoDockTools4: automated docking with selective receptor flexibility. J. Comput. Chem. 16, 2785-2791.

Mutis, A., Parra, L., Manosalva, L., Palma, R., Candia, O., Lizama, M., Pardo, F., Perich, F. and Quiroz, A. 2010. Electroantennographic and behavioural responses of adults of raspberry weevil *Aegorhinus superciliosus* (Coleoptera: Curculionidae) to odors released from conspecific females. Environ. Entomol. 39, 1276-1282.

Mutis, A., Palma, R., Parra, L., Alvear, M., Isaacs, R., Morón, M. and Quiroz, A. 2014. Morphology and distribution of sensilla on the antennae of *Hylamorpha elegans* Burmeister (Coleoptera: Scarabaeidae). Neotrop. Entomol. 43, 260-265.

Nespoulous, C., Briand, L., Delage, M.-M., Tran, V. and Pernollet, J.-C. 2004. Odorant binding and conformational changes of a rat odorant-binding protein. Chem. Senses. 29, 189-198.

Nikonov, A.A., Parker, J.M. and Caprio, J. 2002. Odorant-induced olfactory receptor neural oscillations and their modulation of olfactory bulbar responses in the cannel catfish. J. Neurosci. 22, 2352-2362.

Nojima, S., Linn, C.Jr., Morris, B., Zhang, A. and Roelofs, W. 2003. Identification of host fruit volatiles from hawthorn (Crataegus spp.) attractive to hawthorn-origin *Rhagoletis pomonella* flies. J. Chem. Ecol. 29, 321-336.

Notredame, C., Higgins, D.G. and Heringa, J. 2000. T-Coffee: A novel method for fast and accurate multiple sequence alignment. J. Mol. Biol. 302, 205-217.

Orry, A.J.W. and Abagyan, R. 2012. Preparation and refinement of model protein-ligand complexes. *Homology Modelling: Methods* and *Protocols* (ed. by A.J.W. Orry and R. Abagyan), pp. 351-373. Molsoft L.L.C., San Diego.

P D, K.J., Kempraj, V., Aurade, R.M., Roy, T.K., K S, S. and Verghese, A. 2014. Computational reverse chemical ecology: Virtual screening and predicting behaviourally active semiochemicals for *Bactrocera dorsalis*. BMC Genom. 15, 209.

Paas, Y., Devillers-Thiéry, A., Teichberg, V.I., Changeux, P. and Eisenstein, M. 2000. How well can molecular modelling predict the crystal structure: the case of the ligand-binding domain of glutamate receptors. Trends Pharmacol. Sci. 21, 87-92.

Palma R., Mutis A., Manosalva L., Ceballos R. and Quiroz A. 2012. Behavioural and electrophysiological responses of *Hylastinus obscurus* to volatiles released from the roots of *Trifolium pratense* L. J. Soil Sci. Plant Nutr. 12, 183-193.

Paramasivan, R., Sivaperumal, R., Dhananjeyan, K.J., Thenmozhi, V. and Tyagi, B.K. 2007. Prediction of 3-dimensional structure of salivary odorant-binding protein-2 of the

mosquito *Culex Quinquefasciatus*, the vector of human lymphatic filariasis. In Silico Biol. 7, 1-6.

Pelosi, P. and Maida, R. 1990. Odorant-binding proteins in vertebrates and insects: Similarities and possible common function. Chem. Senses. 15, 205-215.

Pelosi, P. 2001. The role of perireceptor events in vertebrate olfaction. Cell. Mol. Life Sci. 58, 503-509.

Pelosi, P., Calvello, M. and Ban, L. 2006. Diversity of odorant-binding proteins and chemosensory proteins in insects. Chem. Senses. 30, 291-292.

Pesenti, M.E., Spinelli, S., Bezirard, V., Briand, L., Pernollet, J-.C., Campanacci, V., Tegoni, M. and Cambillau, C. 2009. Queen bee pheromone binding protein pH-induced domain swapping favors pheromone release. J. Mol. Biol. 390, 981-990.

Pham-Delegue, M.-H., Trouiller, J., Bakchine, E., Roger, B. and Masson, C. 1991. Age dependency of worker bee response to queen pheromone in a four-armed olfactometer. Insectes Sociaux. 38, 283-292.

Phillips, J.C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R.D., Kale, L. and Schulten, K. 2005. Scalable molecular dynamics with NAMD. J. Comput. Chem. 26, 1781-1802.

Plettner E., Lazar J., Prestwich E.G., Prestwich G.D. 2000. Discrimination of pheromone enantiomers by two pheromone binding proteins from the gypsy moth *Lymantria dispar*. Biochemistry. 39, 8953-8962.

Pophof, B. 2004. Pheromone-binding proteins contribute to the activation of olfactory receptor neurons in the silkmoths *Antheraea polyphemus* and *Bombyx mori*. Chem. Senses. 29, 117-125.

Qiao, H., Tuccori, E., He, X., Gazzano, A., Field, L., Zhou, J.-J. and Pelosi, P. 2009. Discrimination of alarm pheromone (*E*)-β-farnesene by aphid odorant-binding proteins. Insect Biochem. Mol. Biol. 39, 414-419.

Qiao, H., He, X., Schymura, D., Ban, L., Field, L., Dani, F.R., Michelucci, E., Caputo, B., della Torre, A., Iatrou, K., Zhou, J.-J., Krieger, J. and Pelosi, P. 2011. Cooperative interactions between odorant-binding proteins of *Anopheles gambiae*. Cell. Mol. Life Sci. 68, 1799-1813.

Quiroz A., Fuentes-Contreras E., Ramírez C.C., Russell G.B. and Niemeyer H.M. 1999. Host-plant chemicals and distribution of *Neuquenaphis* on *Nothofagus*. J. Chem. Ecol. 25, 1043-1054.

Quiroz, A., Palma, R., Etcheverría, P., Navarro, V., and Rebolledo, R. 2007. Males of *Hylamorpha elegans* Burmeister (Coleoptera: Scarabaeidae) Are Attracted to Odors Released from Conspecific Females. Environ. Entomol. 36, 272 – 280.

Rarey, M., Kramer, B., Lengauer, T. and Klebe, G. 1996. A fast flexible docking method using an incremental construction algorithm. J. Mol. Biol. 261, 470-489.

Raval, A., Piana, S., Eastwood, M.P., Dror, R.O. and Shaw, D.E. 2012. Refinement of protein structure homology models via long, all-atom molecular dynamics simulations. Proteins. 80, 2071-2079.

Ravna, A.W. and Sylte, I. 2012. Homology Modelling of transporter proteins (carriers and ion channels). *Homology Modelling: Methods* and *Protocols* (ed. by A.J.W. Orry and R. Abagyan), pp. 281-299. Molsoft L.L.C., San Diego.

Reinecke A., Ruther J., Tolasch T., Francke W. and Hilker M. 2002. Alcoholism in cockchafers: orientation of male *Melolontha melolontha* towards green leaf alcohols. Naturwissenschaften. 89, 265-269.

Reynolds C.H., Bembenek S.D. and Tounge B.A. 2007. The role of molecular size in ligand efficiency. Bioorg. Med. Chem. 17, 4258-4261.

Rodríguez, M., France, A., and Gerding, M. 2004. Evaluación de dos ceptas del hongo *Metarhizium anisopliae* var. *anisopliae* (Metsh.) para el control de larvas de gusano blanco *Hylamorpha elegans* Burm. (Coleoptera: Scarabaeidae). Agricultura Técnica. 64, 17 – 24.

Rusconi, B., Maranhao, A.C., Fuhrer, J.P., Krotee, P., Choi, S.H., Grun, F., Thireou, T., Dimitratos, S.D., Woods, D.F., Marinotti, O., Walter, M.F. and Eliopoulos, E. 2012. Mapping the *Anopheles gambiae* odorant binding protein 1 (AgamOBP1) using modelling techniques, site directed mutagenesis, circular dichroism and ligand binding assays. Biochim. Biophys. Acta. 1824, 947-953.

Ruther, J., Reinecke, A., Thiemann, K., Tolasch, T., Francke, W., and Hilker, M. 2000. Mate finding in the forest cockchafer, *Melolontha hippocastani*, mediated by volatiles from plants and females. Physiol. Entomol. 25, 172 – 179.

Sakkiah, S., Arooj, M., Kumar, M.R., Eom, S.H. and Lee, K.W. 2013. Identification of inhibitor binding site in human Sirtuin 2 using molecular docking and dynamics simulations. PLoS ONE. 8, e51429. doi:10.1371/journal.pone.0051429.

Sánchez, R., Pieper, U., Melo, F., Eswar, N., Martí-Renom, M.A., Madhusudhan, M.S., Mirković, N. and Šali, A. 2000. Protein structure modelling for structural genomics. Nat. Struc. Biol. 7, 986-990.

Sánchez-Gracia, A. and Rozas, J. 2008. Divergent evolution and molecular adaptation in the *Drosophila* odorant-binding protein family: inferences from sequence variation at the *OS-E* and *OS-F* genes. BMC Evol. Biol. 8, 323.

Sandler, B.H., Nikonova, L., Leal, W.S. and Clardy, J. 2000. Sexual attraction in the silkworm moth: structure of the pheromone-binding protein-bombykol complex. Chem. Biol. 7, 143-151.

Schiestl, F.P. and Ayasse, M. 2000. Post-mating odor in females of the solitary bee, And*rena nigroaenea* (Apoidea, Andrenidae), inhibits male mating behaviour. Behav. Ecol. Sociobiol. 48, 303-307.

Schwede, T., Sali, A., Eswar, N. and Peitsch, M.C. 2007. Protein structure modelling. *Computational Structural Biology – Methods* and *Applications* (ed. by T. Schwede and M.C. Peitsch), pp. 3-35. World Scientific., Danvers.

Scott, W.R.P., Hünenberger, P.H., Tironi, I.G., Mark, A.E., Billeter, S.R., Fennen, J., Torda, A.E., Huber, T., Krüger, P. and van Gunsteren, W.F. 1999. The GROMOS biomolecular simulation program package. J. Phys. Chem. A. 103, 3596-3607.

Shi, J., Blundell, T.L. and Mizuguchi, K. 2001. FUGUE: Sequence-structure homology recognition using environment-specific substitution tables and structure-dependent gap penalties. J. Mol. Biol. 310, 243-257.

Siciliano P, He XL, Woodcock C, Pickett JA, Field LM, Birkett MA, Kalinova B, Gomulski LM, Scolari F, Gasperi G, Malacrida AR, Zhou J-J (2014) Identification of pheromone components and their binding affinity to the odorant binding protein CcapOBP83a-2 of the Mediterranean fruit fly, *Ceratitis capitata*. *Insect Biochemistry* and *Molecular Biology* **48**, 51-62.

Sokkar, P., Mohandass, S. and Ramachandran, M. 2011. Multiple templates-based homology modelling enhances structure quality of AT1 receptor: validation by molecular dynamics and antagonist docking. J. Mol. Model. 17, 1565-1577.

Spinelli, S., Lagarde, A., Iovinella, I., Legrand, P., Tegoni, M., Pelosi, P. and Cambillau, C. 2012. Crystal structure of *Apis mellifera* OBP14, a C-minus odorant-binding protein, and its complexes with odorant molecules. Insect Biochem. Mol. Biol. 42, 41-50.

Sun, Y., Qiao, H., Ling, Y., Yang, S., Rui, C., Pelosi, P. and Yang, X. 2011. New analogues of (*E*)-β-farnesene with insecticidal activity and binding affinity to aphid odorant-binding proteins. J. Agr. Food Chem. 59, 2456-2461.

Sun, Y., Huang, L., Pelosi, P. and Wang, C. 2012. Expression in antennae and reproductive organs suggests a dual role of an odorant-binding protein in two sibling *Helicoverpa* species. PLoS ONE. 7, e30040. doi:10.1371/journal.pone.0030040.

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. and Pelosi, P. 2012. Two odorant-binding proteins mediate the behavioural response of aphids to the alarm pheromone (*E*)-β-farnesene and structural analogues. PLoS ONE. 7, e32759. doi:10.1371/journal.pone.0032759.

Sun, Y-.L., Huang, L-.Q., Pelosi, P. and Wang, C-.Z. 2013. A lysine at the C-terminus of and odorant-binding protein is involved in binding aldehyde pheromone components in two *Helicoverpa* species. PLoS ONE. 8, e55132. doi:10.1371/journal.pone.0055132.

Tegoni, M., Pelosi, P., Vincent, F., Spinelli, S., Campanacci, V., Grolli, S., Ramoni, R. and Cambillau, C. 2000. Mammalian odorant binding proteins. Biochim. Biophys. Acta. 1482, 229-240.

Tegoni, M., Companacci, V. and Cambillau, C. 2004. Structural aspects of sexual attraction and chemical communication in insects. Trends Biochem. Sci. 29, 257-264.

Thompson, J.D., Higgins, D.G. and Gibson, T.J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673-4680.

Tsitsanou, K.E., Thireou, T., Drakou, C.E., Koussis, K., Keramioti, M.V., Leonidas, D.D., Eliopoulos, E., Iatrou, K. and Zographos, S.E. 2012. *Anopheles gambiae* odorant binding protein cristal complex with the synthetic repellent DEET: implications for structure-based design of novel mosquito repellents. Cell. Mol. Life Sci. 69, 283-297.

Tsuchihara, K., Fujikawa, K., Ishiguro, M., Yamada, T., Tada, C., Ozaki, K. and Ozaki, M. 2005. An odorant-binding protein facilitates odorant transfer from air to hydrophilic surroundings in the blowfly. Chem. Senses. 30, 559-564.

Venthur, H., Mutis, A., Zhou, J.-J., and Quiroz, A. 2014. Ligand binding and homology modelling of insect odorant-binding proteins. Physiol. Entomol. 39, 183-198.

Verheggen, F.J., Fagel, Q., Heuskin, S., Lognay, G., Francis, F., and Haubruge, E. 2007. Electrophysiological and behavioral responses of the multicolores asian lady beetle, Harmonia axyridis Pallas, to sesquiterpene semiochemicals. J. Chem. Ecol. 33: 2148 – 2155.

Vieira, F.G. and Rozas, J. 2011. Comparative genomics of the odorant-binding and chemosensory protein gene families across the arthropoda: Origin and evolutionary history of the chemosensory system. Genome, Bio. Evol. 3, 476-490.

Vogt, R.G. and Riddiford, L.M. 1981. Pheromone binding and inactivation by moth antennae. Nature. 293, 161-163.

Vogt, R.G., Rybczynski, R. and Lerner, M.R. 1991. Molecular cloning and sequencing of general odorant-binding proteins GOBP1 and GOBP2 from the tobacco hawk moth *M*and*uca sexta*: Comparisons with other insect OBPs and their signal peptides. J. Neurosci. 11, 2972-2984.

Vuts J., Imrei Z., Birkett M.A., Pickett J.A., Woodcock C.M. and Tóth M. 2014. Semiochemistry of the Scarabaeoid Li, K., Yin, J., Cao, Y., 2013. Potential cooperations between odorant-binding proteins of the scarab beetle *Holotrichia oblita* Faldermann (Coleoptera: Scarabaeidae). PLoS ONE 8, doi: 10.1371/journal.pone.0084795.

Walter, M.H., Floss, D.S. and Strack, D. 2010. Apocarotenoids: Hormonos, mycorrhizal metabolites and aroma volatiles. Planta. 232, 1-17.

Wang, B., Guan, L., Zhong, T., Li, K., Yin, J. and Cao, Y. 2013. Potential cooperations between odorant-binding proteins of the scarab beetle *Holotrichia oblita* Faldermann (Coleoptera: Scarabaeidae). PLoS ONE. 8, doi: 10.1371/journal.pone.0084795.

Wang, S.Y., Gu, S.H., Han, L., Guo, Y.Y., Zhou, J.-J. and Zhang, Y.J. 2013. Specific involvement of two amino acid residues in cis-nerolidol binding to odorant-binding protein 5 Alin OBP5 in the alfalfa plant bug, *Adelphocoris lineolatus* (Goeze). Insect Mol. Biol. 22, 172-182.

Wei, M., Deng, X. and Du, J. 2005. Analysis and identification of *Liriomyza sativae*-attractants from cowpea and kidney bean volatiles. Ying Yong Sheng Tai Xue Bao. 16, 907-910.

Wei, Y., Brandazza, A. and Pelosi, P. 2008. Binding of polycyclic aromatic hydrocarbons to mutants of odorant-binding protein: A first step towards biosensors for environmental monitoring. Biochim. Biophys. Acta. 1784, 666-671.

Weiner, P.K. and Kollman, P.A. 1981. AMBER: Assisted Model Building with Energy Refinement. A general program for modelling molecules and their interactions. J. Comput. Chem. 2, 287-303.

Wojtasek H., Picimbon J., Leal W. 1999. Identification and Cloning of Odorant Binding Proteins from the Scarab Beetle *Phyllopertha diversa*. Biochem. Biophys. Res. Commun. 263, 832 – 837.

Wu, G., Robertson, D.H., Brooks III, C.L. and Vieth, M. 2003. Detailed analysis of grid-based molecular docking: A case study of CDOCKER-A CHARMm-based MD docking algorithm. J. Comput. Chem. 24, 1549-1562.

Xu, P.X., Zwiebel, L.J. and Smith, D.P. 2003. Identification of a distinct family of genes encoding atypical odorant-binding proteins in the malaria vector mosquito, *Anopheles gambiae*. Insect Mol. Biol. 12, 549-560.

Xu, W., Xu, X., Leal, W.S. and Ames, J.B. 2011. Extrusion of the C-terminal helix in navel orangeworm moth pheromone-binding protein (AtraPBP1) controls pheromone binding. Biochem. Bioph. Res. Co. 404, 335-338.

Yin, J., Feng, H., Sun, H., Xi, J., Cao, Y. and Li, K. 2012. Functional analysis of general odorant binding protein 2 from the meadow moth, *Loxostege sticticalis* L. (Lepidoptera: Pyralidae). PLoS ONE, 7, e33589. doi:10.1371/journal.pone.0033589.

Yi, X., Zhang, Y., Wang, P., Qi, J., Hu, M. and Zhong, G. 2015. Ligands binding and molecular simulation: The potential investigation of a biosensor based on an insect odorant binding protein. Int. J. Biol. Sci. 11, 75-87.

Yu, F., Zhang, S., Zhang, L. and Pelosi, P. 2009. Intriguing similarities between two novel odorant-binding proteins of locusts. Biochem. Bioph. Res. Co. 385, 369-374.

Yu, Y., Ma, F., Cao, Y., Zhang, J., Zhang, Y., Duan, S., Wei, Y., Zhu, S. and Chen, N. 2012. Structural and functional difference of pheromone binding proteins in discriminating chemicals in the gypsy moth, *Lymantria dispar*. Int. J. Biol. Sci. 8, 979-991.

Yu, Y. and Plettner, E. 2013. Enantiomer and conformer recognition of (+) and (-)-disparlure on their analogs by the pheromone binding protein of the gypsy moth, *Lymantria dispar*. Bioorgan. Med. Chem. 21, 1811-1822.

Zhang, T., Mei, X., Feng, J., Berg, B.G., Zhang, Y. and Guo, Y. 2012. Characterization of three pheromone-binding proteins (PBPs) of *Helicoverpa armigera* (Hübner) and their binding properties. J. Insect. Physiol. 58, 941-948.

Zhang, T., Wang, W., Gu, S., Zhang, Z., Wu, K., Zhang, Y. and Guo, Y. 2012. Structure, binding characteristics and 3D model prediction of a newly identified odorant-binding protein from the cotton bollworm, *Helicoverpa armigera* (Hübner). J. Integr. Agr. 11, 430-438.

Zhong, T., Yin, J., Deng, S., Li, K. and Cao, Y. 2012. Fluorescence competition assay for the assessment of green leaf volatiles and *trans*-β-farnesene bound to three odorant-binding proteins in the wheat aphid *Sitobion avenae* (Fabricius). J. Insect. Physiol. 58, 771-781.

Zhou, J.-J., Huang, W., Zhang, G., Pickett, J. and Field, L. 2004. "Plus-C" odorant-binding protein genes in two *Drosophila* species and the malaria mosquito *Anopheles gambiae*. Gene. 327, 117-129.

Zhou, J.-J., Kan, Y., Antoniw, J., Pickett, J.A. and Field, L.M. 2006. Genome and EST analyses and expression of a gene family with putative functions in insect chemoreception. Chem. Senses. 13, 453-465.

Zhou, J.-J., Robertson, G., He, X., Dufour, S., Hooper, A.M., Pickett, J.A., Keep, N.H. and Field, L.M. 2009. Characterisation of *Bombyx mori* odorant-binding proteins reveals that a

general odorant-binding protein discriminates between sex pheromone components. J. Mol. Biol. 389, 529-545.

Zhou, J.-J. 2010. Odorant-binding proteins in insects. Vitamins and Hormones (ed. by G. Litwack), pp. 241-272. Academic Press, Burlington.

Zhuang, X., Wang, Q., Wang, B., Zhong, T., Cao, Y., Li, K. and Yin, J. 2014. Prediction of the key binding site of odorant-binding protein of *Holotrichia oblita* Faldermann (Coleoptera: Scarabaeidae). Insect Mol. Biol. 23, 381-390.

Zubkov, S., Gronenborn, A.M., Byeon, I.L. and Mohanty, S. 2005. Structural consequences of the pH-induced conformational switch in *A. polyphemus* pheromone-binding protein: Mechanisms of ligand release. J. Mol. Biol. 354, 1081-1090.

Appendices

Appendix 3.1

Composition of volatiles from leaves of *N. obliqua* trapped by SPME and Porapak Q

Compound	Mass spectra data (m/z)*	KI Exp.[a]	KI Lib.[b]	Reliability of identification[c]
β-Myrcene		985	985	1
(*Z*)-3-hexenyl acetate		989	989	1
Unknown 1	43, 57, 85, 99, 142	1029		
β-Ocimene		1041	1041	1
β-Linalool		1086	1086	1
Unknown 2	41, 51, 69, 79, 94, 107, 135, 150	1106		
Dodecane		1199	1200	1
Sesquiterpene	41, 43, 55, 77, 93, 105, 119, 121, 161, 204	1375	1375	2
Sesquiterpene	41, 51, 67, 81, 91, 105, 123, 161, 204	1383	1383	2
Tetradecane		1399	1400	1
α-Gurjunene		1408	1411	1
Caryophyllene		1416	1416	1
Aromadendrene		1457	1455	1
Sesquiterpene	41, 53, 67, 79, 81, 93, 105, 107, 121, 133, 135, 148, 175, 189, 204	1480	1480	2
Sesquiterpene	43, 55, 67, 69, 79, 95, 105, 109, 121, 133, 147, 161, 189, 204	1580	1575	2

[a] Kovats indices experimental

[b] Kovats indices library

[c] Reliability of identification is indicated as: (1) when the method comprises a comparison with mass spectra, KI and matching with commercial standard; (2) when the method comprises only a comparison with mass spectra and KI according to data of literature.

*(m/z) data for mass spectrum were included for those unidentified compound.

Appendix 3.2

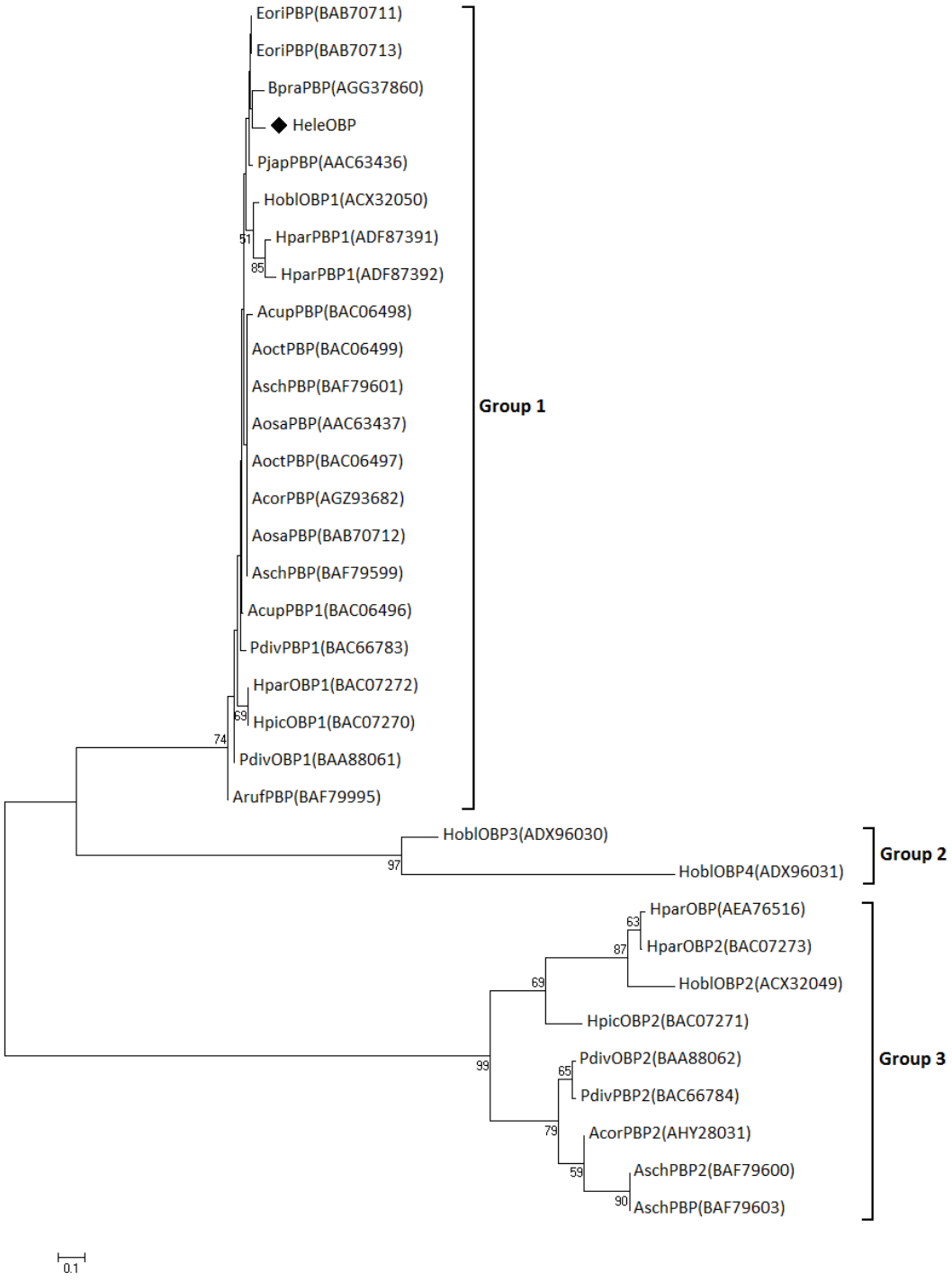
Oligonucleotide primers designed for cDNA cloning and sub-cloning of the OBP in *H. elegans*.

Primer name	Sequence
PdivOBP_1	5'-GARGARATGGARGARCT-3'
PdivOBP_2	5'-GARGARATGGARGARTTR-3'
OBP 5'RACE (GSP1)	5'-CACTTCCTCATTACAGGCTCAG-3'
OBP 5'RACE (GSP2)	5'-CACGATCATATAGTTCTTGGAATCC-3'
OBP LIC fwd	5'-TACTTCCAATCCATGAGCGAGGAGATGGAGGAGTTAG-3'
OBP LIC rev	5'-TATCCACCTTTACTGTTACACGATCATATAGTTCTTG-3'

The restriction site for BsaI is underlined including start codon for OBP LIC fwd and stop codon for OBP LIC rev.

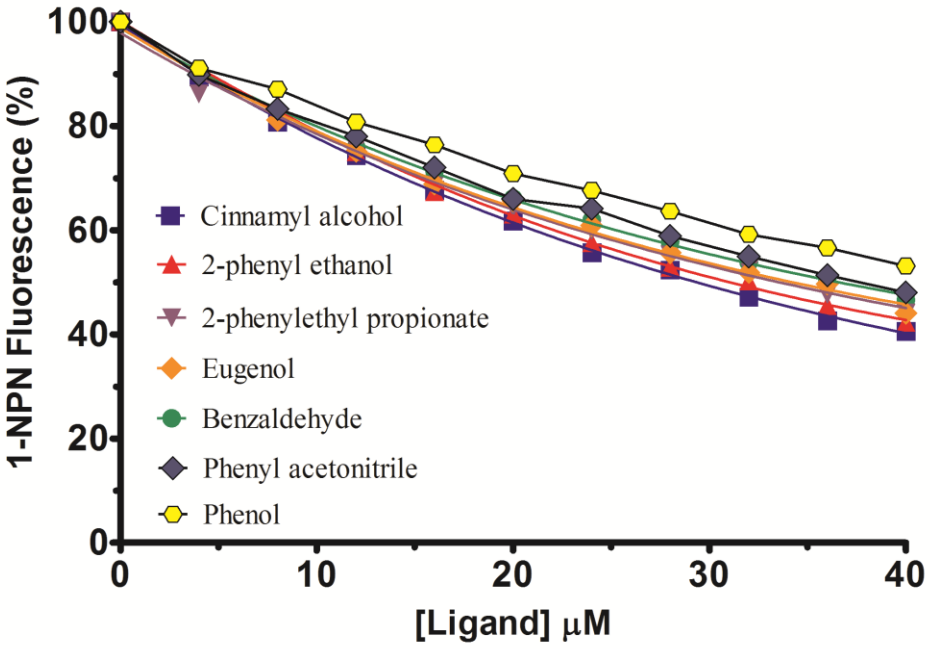
Appendix 3.3

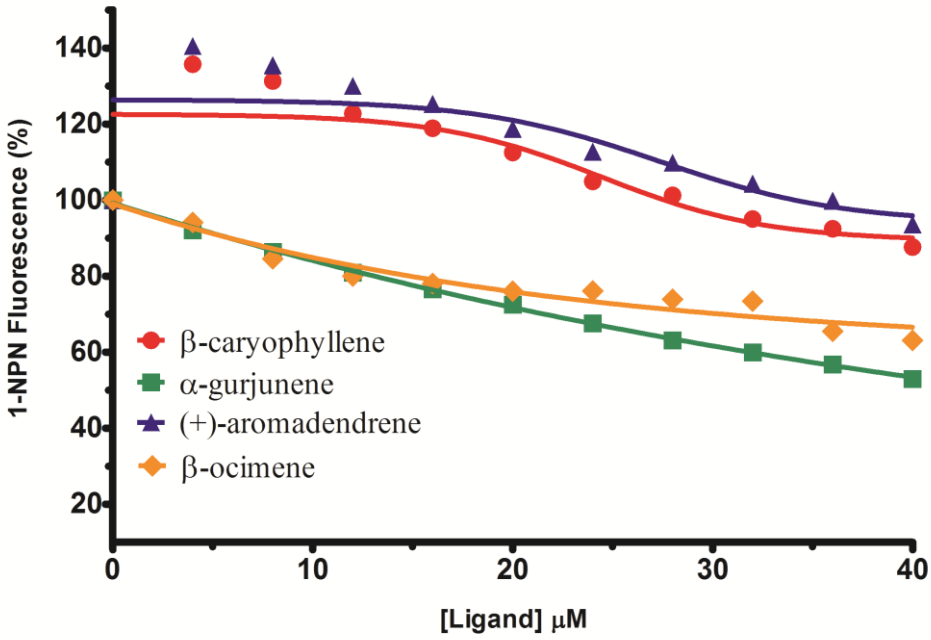
Phylogenetic tree with scarab OBPs available in NCBI (http://www.ncbi.nlm.nih.gov/). The OBPs with the closest evolutionary relationships were labelled as Group 1, 2 and 3. "◆" indicates HeleOBP1 (accession number: AGM37951).

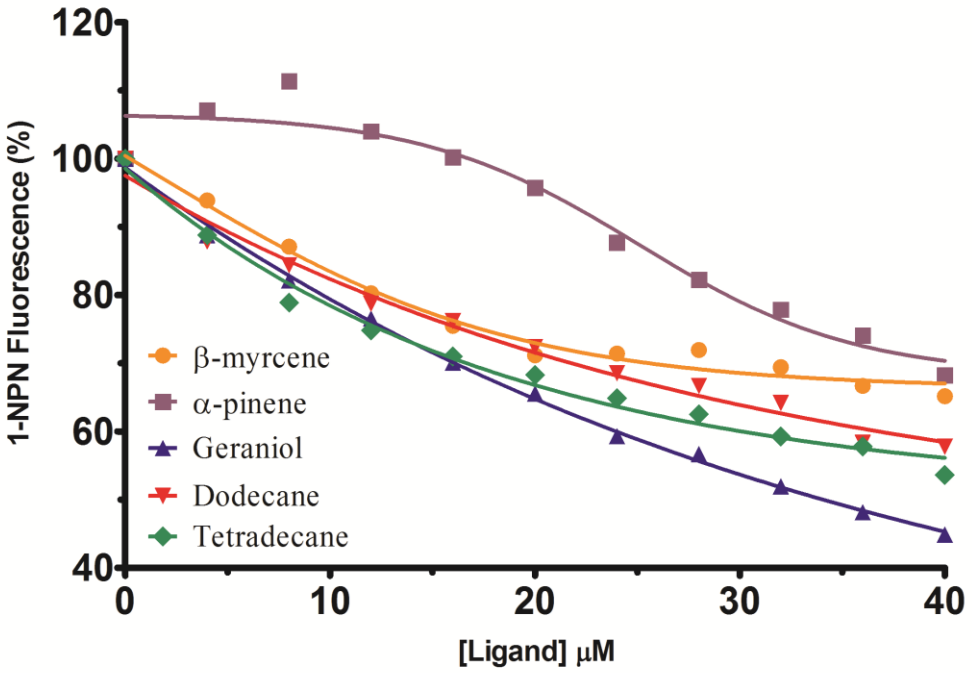


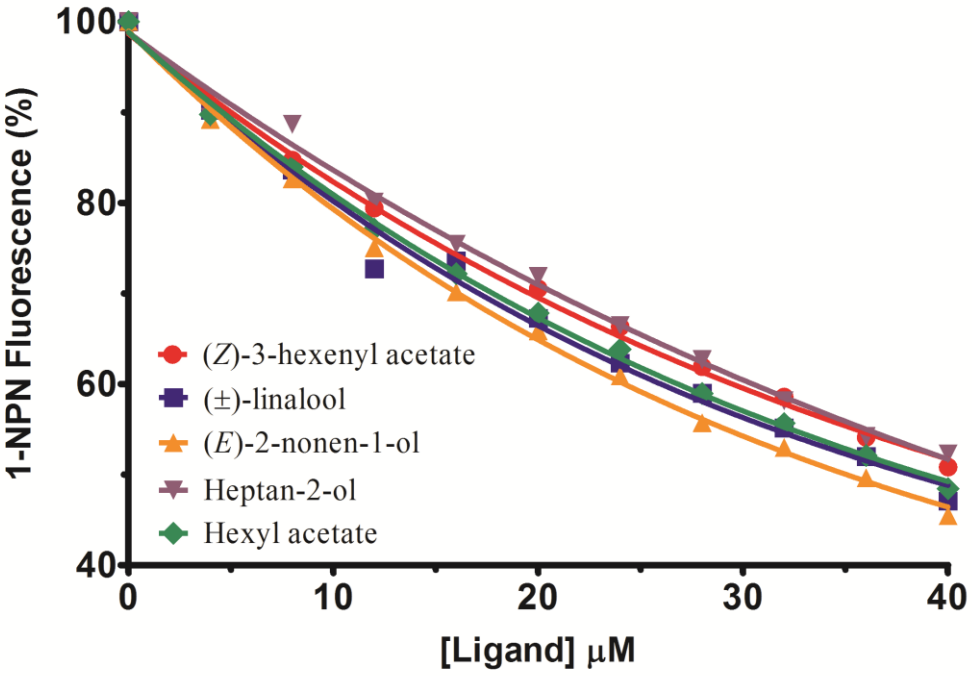
Appendix 3.4

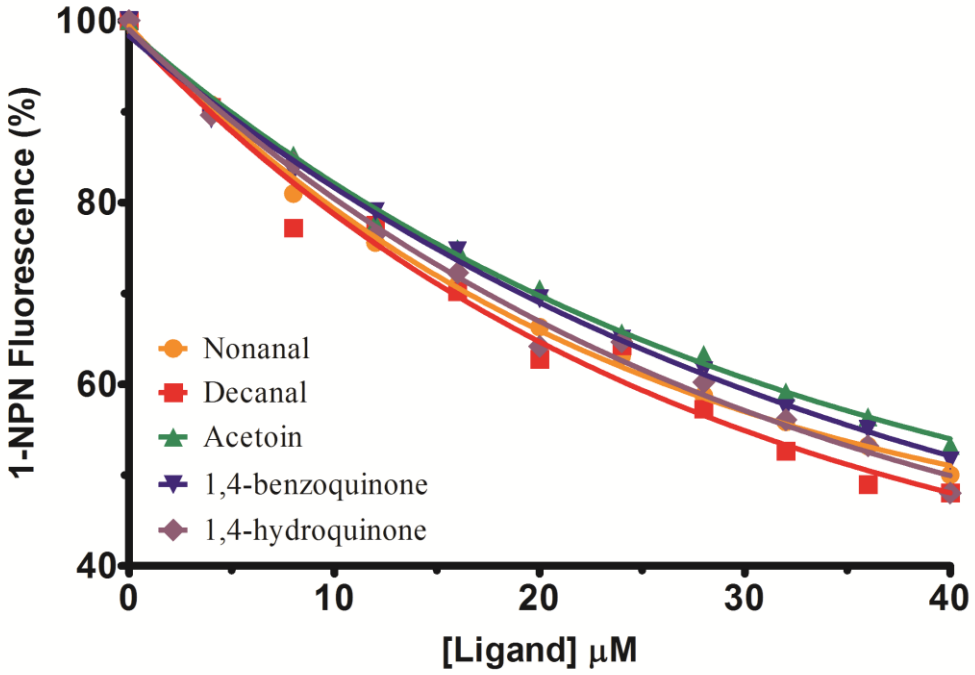
Competitive binding of HeleOBP1 to selected odorants. A 2 µM solution of the protein plus 1-NPN was titrated with 1 mM solutions of each ligand in methanol to final concentrations of 4-40 µM.





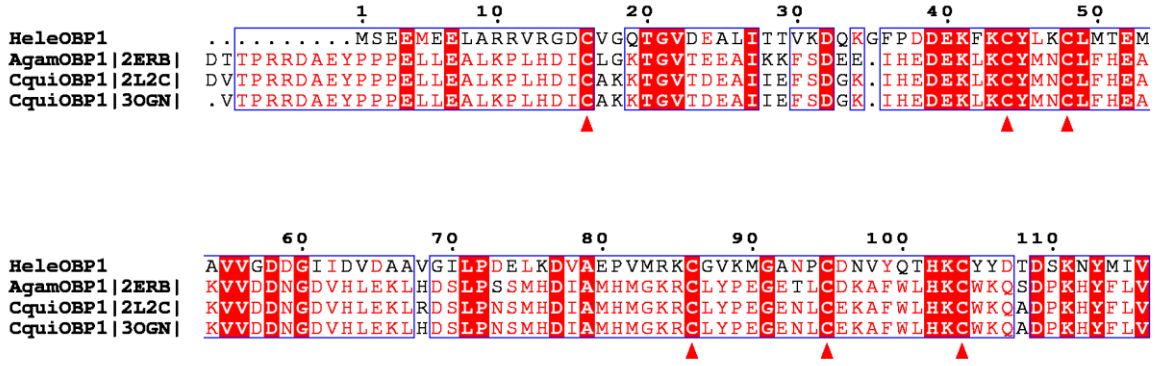

120
100
80
60
40
1-NPN Fluorescence (%)
β-myrcene
α-pinene
Geraniol
Dodecane
Tetradecane
0
10
20
30
40
[Ligand] μM


100
80
60
40
1-NPN Fluorescence (%)
(Z)-3-hexenyl acetate
(±)-linalool
(E)-2-nonen-1-ol
Heptan-2-ol
Hexyl acetate
0
10
20
30
40
[Ligand] μM


100
80
60
40
1-NPN Fluorescence (%)
0
10
20
30
40
[Ligand] μM
Nonanal
Decanal
Acetoin
1,4-benzoquinone
1,4-hydroquinone

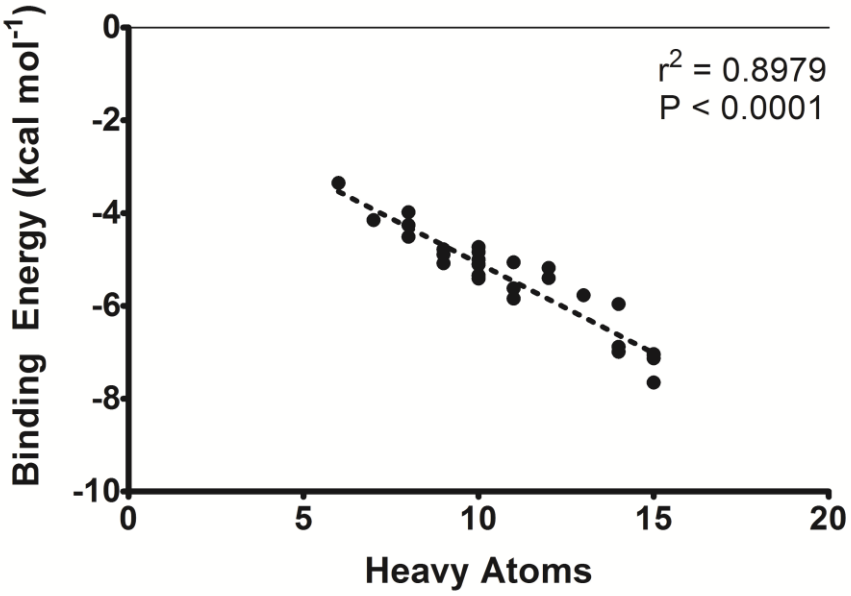
Appendix 3.5

Multiple sequence alignment among HeleOBP1, AgamOBP1 (PDB code: 2ERB) and CquiOBP1 (PDB code: 2L2C and 3OGN). Alignment was carried out by Multalin and ESPript without signal peptides. Identical residues are highlighted in white letters with a red background. Amino acids with similar physic-chemical properties are shown in red letters. Alignment positions are framed in blue if the corresponding residues are identical or similar. Red arrows indicate cysteine residues.



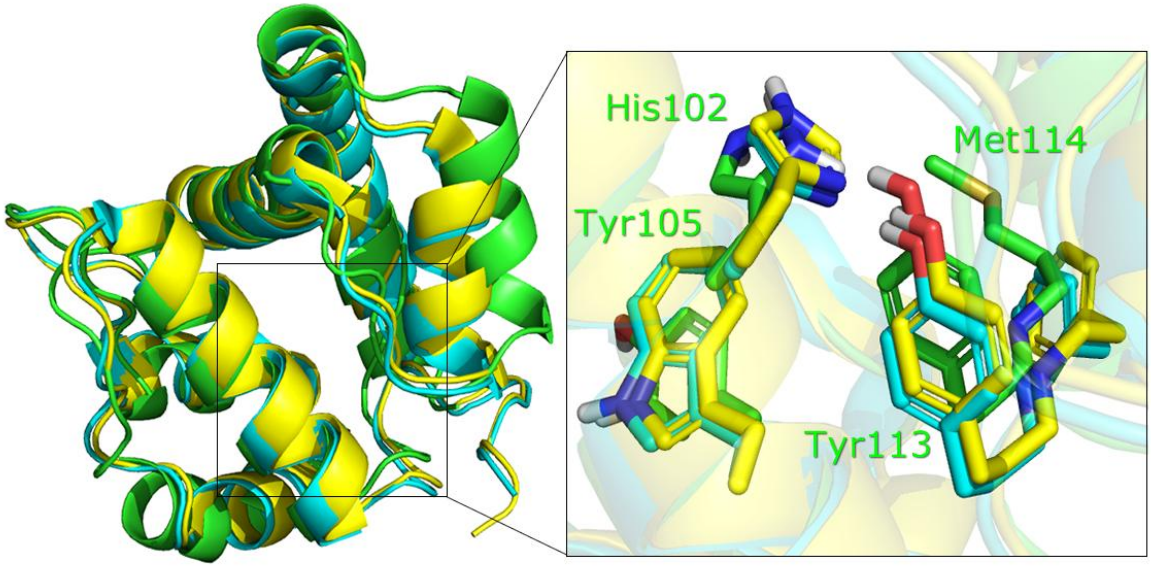
Appendix 3.6

Correlation between free binding energy and molecular size as heavy atoms (non-hydrogens) from molecular docking.



Appendix 3.7

Super-imposed 3D structures of HeleOBP1 (green), AgamOBP1 (PDB code: 2ERB) (light blue) and CquiOBP1 (PDB code: 3OGN) (yellow). Right scheme indicates conserved amino acids between the OBP structures (target and templates). Residues from HeleOBP1 (i.e. green sticks) are labelled.



Appendix 3.8

In silico binding affinities expressed as free binding energy and fit quality (FQ) for tested ligands by molecular docking.

Ligand	Free binding energy (kcal mol^{-1})	Fit Quality (FQ)
β-caryophyllene	-7.13	0.82
α-gurjunene	-7.05	0.81
(+)-aromadendrene	-7.04	0.81
β-ocimene	-5.41	0.75
α-pinene	-5.35	0.74
Geraniol	-5.62	0.73
α-ionone	-6.88	0.81
β-ionone	-6.99	0.82
β-myrcene	-5.11	0.71
(±)-linalool	-5.84	0.76
Dodecane	-5.40	0.67
Tetradecane	-5.96	0.70
(*E*)-2-nonen-1-ol	-5.10	0.71
Heptan-2-ol	-4.31	0.76
(*Z*)-3-hexenyl acetate	-5.00	0.69
Hexyl acetate	-4.84	0.67
Nonanal	-4.73	0.66
Decanal	-5.06	0.66
Acetoin	-3.35	1.52
Phenol	-4.15	0.96
Benzaldehyde	-4.51	0.80
1,4-benzoquinone	-4.26	0.75
1,4-hydroquinone	-3.98	0.70
Eugenol	-5.18	0.65
2-phenyl acetaldehyde	-4.78	0.73
2-phenyl ethanol	-4.89	0.74
2-phenylethyl propionate	-5.77	0.69
2-phenyl acetonitrile	-5.08	0.77
Cinnamyl alcohol	-5.34	0.74