Physiological Entomology (2014) **39**, 183–198 DOI: 10.1111/phen.12066

REVIEW ARTICLE

Ligand binding and homology modelling of insect odorant-binding proteins

HERBERT VENTHUR[1,2,3], ANA MUTIS[2], JING-JIANG ZHOU[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 and [3]Department of Biological Chemistry and Crop Protection, Rothamsted Research, Harpenden, U.K.

Abstract. This review describes the main characteristics of odorant-binding proteins (OBPs) for homology modelling and presents a summary of structure prediction studies on insect OBPs, along with the steps involved and some limitations and improvements. The 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 considered to play a functional role for recognition, desorption, scavenging, protection and transportation of hydrophobic molecules (odourants) across an aqueous environment (lymph) to olfactory receptor neurones (ORNs) located in sensilla, the main olfactory units of insect antennae. Lepidopteran pheromone-binding proteins, a subgroup of OBPs, are characterized by remarkable structural features, in which high sequence identities (approximately 30%) among these OBPs and a large number of available templates can facilitate the prediction of precise homology models. Approximately 30 studies have been performed on insect OBPs using homology modelling as a tool to predict their structures. Although some of the studies have assessed ligand-binding affinity using structural information and biochemical measurements, 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.

Key words. Chemical ecology, homology modelling, ligand binding, molecular docking, molecular dynamic simulation, odorant-binding protein.

Introduction

Host-seeking, oviposition behaviour and mating of insects are governed mainly by odour perception through sensory organs such as antennae. These organs contain a well described olfactory system (Jacquin-Joly & 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 neurones (ORNs) with their dendrites bathed in sensillum lymph. These ORNs modulate ion potentials across their plasma membrane and 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 odour perception and discrimination. The main olfactory proteins include odorant-binding proteins (OBPs), chemosensory proteins (CSPs), odorant-degrading enzymes, sensory neurone membrane proteins, olfactory co-receptors (formerly OR83b receptors) and olfactory receptors (ORs). It is considered that volatile odorant molecules such as pheromones are transported by OBPs or CSPs across sensillum lymph to ORs located

Correspondence: Jing-Jiang Zhou, Department of Biological Chemistry and Crop Protection, Rothamsted Research, Harpenden, Herts, AL5 2JQ, U.K. Tel.: +44 1582 763133; e-mail: jing-jiang.zhou@rothamsted.ac.uk; and Andrés Quiroz, Universidad de La Frontera, Temuco, Casilla 54-D, Chile. Tel.: +56 045 2732419; e-mail: andres.quiroz@ufrontera.cl

in dendritic membrane of ORNs in sensilla (Vogt & 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). Subsequent to the first OBP being identified by Vogt & Riddiford (1981), many more have been characterized. These proteins are soluble, with a molecular weight of 13–16 kDa, and are abundant in sensillum lymph. Some evidence suggests that OBPs, particularly members of the pheromone-binding protein (PBP) subgroup, can bind odorants selectively; for example, the pheromone (*E*,*Z*)-10,12-hexadecadienol (bombykol) is a specific ligand for the *Bombyx mori* PBP (BmorPBP1) (Sandler *et al.*, 2000) and (*E*,*Z*)-6,11-hexadecadienal is the pheromone component for *Antheraea polyphemus* PBP (ApolPBP1). However, OBPs may also bind to a wide range of odourant molecules (Honson *et al.*, 2005; Zhou, 2010). The above papers summarize all of the early work on OBP/ligand interactions that were determined experimentally, as well as the various assays that are still used today to study the ligand binding of OBPs. More studies are necessary to understand fully the selectivity of OBPs during odour recognition, although some important parameters of odourant molecules are known, such as length of carbon chain, functional groups, location of methyl ends and the specific position of unsaturation, as well as hydrophobic characteristics (Zhou, 2010; Yu & Plettner, 2013). These provide useful information for structural homology modelling.

Ligands (e.g. odourant molecules) bind to proteins through physicochemical forces, which are noncovalent interactions, such as hydrogen bonds, ionic bonds and van der Waals interactions, amongst others. Different amino acids contribute to the interactions according to their structures and positions in the ligand-binding site of the proteins. The most common approach for understanding these interactions in relation to the potency and specificity of ligand binding is to obtain the three-dimensional (3D) structure of the binding proteins bound with ligands (Ravna & 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 the optimal techniques so far for the analysis of ligand–protein interactions based on structural information as a result of their high precision (high resolution), although they are expensive and time consuming. Moreover, it is not always possible to obtain protein crystals in the presence of testing ligands. Computational approaches can contribute as complementary tools for 3D structure prediction (Paas *et al.*, 2000). One of them is comparative or homology modelling, a method that is based on the assumption that protein folds within a structural family are more conserved than the corresponding primary sequences. This structural prediction approach is based on evolutionary relationships between a target protein and template proteins for which a 3D structure exists, as determined by crystallographic or NMR experiments.

OBPs are present at high concentrations (up to 10 mM) in sensillum lymph (Klein, 1987). There is increased interest in OBPs and other olfactory proteins of invertebrates, in the hope that detailed knowledge of their structures and functionality may provide leads for the prevent human diseases transmitted by insects, and alternative pest control strategies through manipulation of 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 together with robust biological assays to study ligand–OBP interactions, as well as to research 'super-ligands' for insect behavioural manipulation. Leal (2005) proposes 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 for using these chemicals in pest management. An advantage of this approach is that insect OBPs are structurally different from and have no sequence homology with vertebrate OBPs (also called lipocalins). However, insect OBPs are functionally similar to vertebrate OBPs. Insect OBPs have mainly α-helical domains, whereas vertebrate OBPs have only β-strands and a short α-helix. Nevertheless, both insect and vertebrate OBPs have conserved disulphide bridges and are small soluble proteins. Ligand binding assays indicate that both insect and vertebrate OBPs bind to a wide range of volatile molecules, with dissociation constants of either a micro- or millimolar 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 recordings show 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 in the nanomolar range was reported by Katre *et al.* (2009). Although many studies have used homology modelling in structure-based drug discovery, only a few modelling studies have been performed on insects OBPs. As a result of the high similarity of OBPs across lepidopteran 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 (MD) simulations to obtain the best structural models in terms of energy. Thus, the present review aims to present current knowledge of computer-assisted protein modelling by homology to predict the ligand-binding affinities, focusing on lepidopteran OBPs.

Materials and methods

The classification and functions of insect OBPs are reviewed, and current theories concerning the mechanisms of binding and release of air-borne ligands to the olfactory receptors in the chemo-sensilla on insect antennae are described. Furthermore, the steps and limitations of OBP structure prediction by homology modelling are delineated, and proposals are made for some improvements to the predictions. Finally, the ligand-binding affinity predictions of homology models including methods and interactions are summarized.

OBPs: classification, function and mechanisms

Classification. Lepidopteran OBPs can be divided into three classes based on their amino acid sequences (Fig. 1) and their

structural characteristics. Thus, OBPs that bind to pheromones are referred as pheromone-binding proteins (PBPs), such as BmorPBP1 (Krieger *et al.*, 1996). General odorant-binding proteins (GOBPs), such as GOBP1 and GOBP2, which are present in both females and males of tobacco hawk moth *Manduca sexta* (Vogt *et al.*, 1991), are proposed to bind to general odorants. Less well described are antennal-binding proteins X (ABPx), which are first reported in *B. mori* (Krieger *et al.*, 1996), although their ligand binding specificity has only been studied recently (He *et al.*, 2010). As shown in Fig. 1, an alternative classification for OBPs consists of 'Classic' OBPs, 'Plus-C' OBPs, 'Minus-C' OBPs and 'Atypical' OBPs (Hekmat-Scafe *et al.*, 2002; Zhou *et al.*, 2004). The main sequence differences are the number of cysteine (Cys) residues and their conservation. According to this, 'Classic' OBPs have six conserved Cys residues at specific positions and include PBPs, GOBPs and ABPx, whereas 'Plus-C' OBPs have two additional Cys and a conserved proline (Hekmat-Scafe *et al.*, 2002; Zhou *et al.*, 2004). 'Minus-C' OBPs have less than six Cys and 'Atypical' OBPs are characterized by having six Cys residues as in the 'Classic' OBPs but with additional cysteines in the C-terminal region (Xu *et al.*, 2003). For OBPs with the six conserved Cys residues, there is also a classification according to chain length and C-terminus length (Tegoni *et al.*, 2004): (i) long-chain OBPs with approximately 140 amino acids (e.g. OBPs of the moths *B. mori* and *A. polyphemus*); (ii) medium-chain OBPs with approximately 120 amino acids (e.g. some OBPs of the mosquito *Anopheles gambiae* and the bee *Apis mellifera*); and (iii) short-chain OBPs with approximately 100 amino acids (e.g. OBPs from the cockroach *Leucophaea maderae*). More recently, a novel subclass of OBPs, according to their C-termini and 3D structure, called C8 OBP class, was proposed by Lagarde *et al.* (2011). The findings for the *A. gambiae* OBP7 (AgamOBP7) suggest a more evolved protein as a result of an increase in structural complexity with eight Cys residues, four disulphide bridges and a C-terminus slightly longer than Classic OBPs. Moreover, it is proposed that AgamOBP7 may have evolved from Classic OBPs, which served as a basal group of OBPs according to phylogenetic analyses (Vieira & Rozas, 2011). The focus in the present review is on Classic OBPs because they all have well conserved characteristics such as six Cys residues, three disulphide bridges and an average molecular weight of approximately 14 kDa. Furthermore, most of the well-studied OBPs such as PBPs and GOBPs of lepidopteran species and the OBP (LUSH) for the pheromone (11-*Z*-vaccenyl acetate) of *Drosophila melanogaster* are Classic OBPs.

Function and specificity. In insects, perception of the environment, such as host plants, prey and potential mates is guided mainly by chemical signals, which are termed semiochemicals. If a semiochemical is a small hydrophobic molecule, it cannot easily cross a polar environment, such as the insect sensillum lymph. It is considered that OBPs are the key component to solubilize and transport these molecules across the lymph to reach the ORs. For hydrophobic ligands such as lepidopteran sex pheromones at least, it has become clear that OBPs play an important role in ligand capturing and transport to achieve the ligand–OR interaction, as well as in contributing

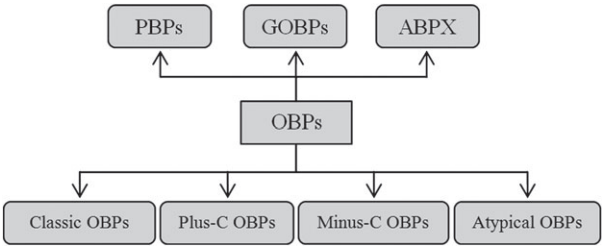

Fig. 1. Classification scheme of insect odorant binding proteins (OBPs) according to their primary protein sequence. ABPX, antennal-binding protein X; GOBP, general odorant-binding protein; PBP, pheromone-binding protein.

to the subsequent signal transduction (Kaissling, 2013). However, experimental evidence suggests that there are three primary functions for OBPs, and these may play dynamic roles in ligand selectivity: (i) ligand scavenging, responsible for the protection of ORs from saturation, which was proposed using a kinetic model for *Lymantria dispar* PBP2 (LdisPBP2) (Gong *et al.*, 2009); (ii) ligand desorption from the cuticular wax layer of olfactory pores to the lymph, as suggested for LdisPBP1 and LdisPBP2 (Kowcun *et al.*, 2001); and (iii) ligand recognition, as proposed for *D. melanogaster* OBP (LUSH) acting as an activator of ORNs tuned to pheromones (Laughlin *et al.*, 2008). By contrast, Gomez-Diaz *et al.* (2013) report 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 may not have a direct role as the activator of ORNs. It is also proposed that OBPs protect odourant molecules from degradation by odorant-degrading enzymes during their transport and transfer to the ORs (Ishida & Leal, 2002; Leal, 2005, 2013).

When semiochemicals are acting as pheromones, there is an intraspecific interaction between insects through associated PBPs. For example, the 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 is 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, a process that 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 demonstrate that an OR response can be achieved with bombykol solubilized with BmorPBP1 or with the organic solvent dimethyl sulphoxide (DMSO). For bombykal, BmorPBP1 does not substitute for DMSO, indicating that BmorPBP1 has a higher specificity to bombykol. They postulate that, as in *Antheraea pernyi* and *A. polyphemus* (Maida *et al.*, 2000), a specialized PBP for

bombykal may exist in *B. mori*. The interaction with ORs and other downstream events is not yet well understood. Apart from lepidopteran PBPs, the ligand specificity of insect OBPs has so far not been demonstrated conclusively. Nevertheless, the sex pheromone specificity of lepidopteran PBPs has provided interesting opportunities to study protein–ligand interactions by using OBPs as attractive targets for homology modelling and ligand screening. Currently, competitive binding assays using *N*-phenyl-1-naphthylamine as a fluorescence probe are used to study the ligand binding affinity to OBPs (Lescop *et al.*, 2009; Zhou, 2010). Only a few studies have provided the binding affinities in connection with structural data. Hydrogen bonds appear to have an important role in specific interactions for the recognition of some ligands, as reported for LUSH and its alcohol-binding site (Kruse *et al.*, 2003), and for other lepidopteran OBPs (Zhou *et al.*, 2009). The first attempt to correlate binding affinity with structural data was made in a study by Sandler *et al.* (2000), in which BmorPBP1 was crystallized as a complex with bombykol. It was reported that Ser56 forms a hydrogen bond with the hydroxyl of bombykol, which plays a role in the specificity of the PBP. However, a more in-depth study was performed on the bombykol-BmorPBP complex by Klusak *et al.* (2003), using *ab initio* methods, which suggests 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 indicate that bombykol could adopt two conformations (A and B). A single hydrogen bond with Ser56 is formed in conformation A, whereas two hydrogen bonds, involving Ser56 and Met61, are formed in conformation B. BmorPBP1 is shown to bind to analogues of bombykol, when measured with Chip-assisted high-throughput electrospray ionization mass spectrometry analysis, where the compounds having 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). Similarly, Campanacci *et al.* (2001) report that PBP1 of *Mamestra brassicae* (MbraPBP1) and ApolPBP1 is unable to discriminate pheromones from certain other compounds. By competitive binding assays using 1-aminoanthracene, fatty acids show significant binding affinity. Zhou *et al.* (2009) and He *et al.* (2010) report that BmorPBP1 binds well to both bombykol and bombykal, whereas a GOBP (BmorGOBP2) binds differently to these compounds. It is suggested that the binding of bombykol to BmorGOBP2 involves hydrogen bonding to Arg110 rather than to Ser56 as found for BmorPBP1. In addition, the hydroxyl group of bombykol may form an additional hydrogen bond with Glu98 via a water molecule, as predicted through MD simulations (Gräter *et al.*, 2006a,b), whereas bombykal cannot form such additional hydrogen bond. This explains the ligand discrimination by BmorGOBP2 at the structural level. Interestingly, Zhou *et al.* (2009) also demonstrate by crystallography that 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) shows numerous contacts with *N*,*N*-diethyl-3-methylbenzamide (DEET) at its binding site (e.g. van der Waals interactions). Hydrogen bonds are formed via water molecules with Trp114, Cys95 and Gly92 (Tsitsanou *et al.*, 2012). Although DEET appears to have a specific interaction with the OBP, it was recently reported that eugenyl acetate could be a better repellent than DEET as a result of a better affinity with the binding site of AgamOBP1, as shown by molecular modelling studies (Affonso *et al.*, 2013). However, binding assays and robust behavioural bioassays are necessary to corroborate those findings.

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

Ligand binding and release by lepidopteran OBPs. It has been proposed that, once the bombykol-BmorPBP1 complex reaches the proximity of dendritic membrane, a pH-induced conformational change occurs to release the ligand. Thus, studies have been conducted with crystal structures of OBPs to determine the key features involved in the pH-dependent change. For example, for the moth *B. mori*, it is suggested that the pheromone is released to the ORs through a conformational change of BmorPBP1 by acid pH in the proximity of the ORNs membrane. A significant structural change in the BmorPBP1 is denoted by the shifting of a long C-terminus, which is a characteristic feature of OBPs that function as PBPs in Lepidoptera, from an extended form to an α-helix by this change in pH (from 6.5 to 4.5) (Horst *et al.*, 2001; Lautenschlager *et al.*, 2005; Leal, 2005). Thus, at acidic pH near the surface of dendritic membrane, there is 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 in the lymph space is unknown. Therefore, pH-dependent conformational changes should be considered as an interesting phenomenon of insect OBPs to which no clear function has yet been attributed. The mechanisms of ligand binding and release are not always dependent on the C-terminus of the OBPs. Zhou *et al.* (2009) report a significant difference in the C-terminal conformation between BmorPBP1 and BmorGOBP2. BmorPBP1 has an extended C-terminus occupying its binding site, whereas BmorGOBP2 has a longer C-terminus, which forms an α-helix and does not cover the binding site, nor does it participate in ligand binding. In addition, there are no significant conformational changes among six different ligand–BmorGOBP2 complexes. Furthermore, in a study of pheromone binding to the PBP1 of the moth *Amyelois transitella* (AtraPBP1), Xu *et al.* (2011) record how the C-terminus controls this binding according to pH. The pheromone (*Z*11, *Z*13)-hexadecadienal

binds more strongly to the AtraPBP1 at neutral pH than at acid pH. However, the deletion of the C-terminus of AtraPBP1 increases the pheromone-binding affinity by 100-fold at pH 5.0 and 1.5-fold at pH 7.0. There is probably an occupation of the binding site of AtraPBP1 by its 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 *Helicoverpa assulta* show a decrease in the binding affinity at acid pH. The mutant-type PBPs without C-terminal segment have 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) report that the *A. mellifera* PBP (ASP1) has a contradictory conformational change compared with the observed in BmorPBP1. Thus, ligands bind to this protein at low pH (4.0) and the release occurs at pH 7.0. However, binding is also possible at neutral pH; therefore, an uncommon conformational change was proposed. 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) as suggested by Lartigue *et al.* (2003). This mechanism, in contrast to that for 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 undertaken for *L. dispar* LdisPBP2 using fluorescence binding assays and tryptophan anisotropy measurements to obtain the association in timescale and multimerization, respectively (Gong *et al.*, 2009). The results indicate that LdisPBP2 binds (+)- and (−)-disparlure in both a rapid step and a slow step. The slow step could be more important for obtaining an active ligand for ORs as a result of an internal binding, which results in a more stable ligand–OBP complex. By contrast, the rapid step appears to be related to binding at an external site, where the initial interactions between ligand and OBP occur. Moreover, within a few hours, there is an increased presence of LdisPBP2 as dimer, which has a smaller binding capacity compared with monomer. Subsequently, a more detailed binding mechanism was proposed. It appears 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 from an external binding site to a different internal one (Gong *et al.*, 2010). In Fig. 2, a schematic representation is provided that summarizes the most outstanding research on lepidoptera OBPs, and their ligand binding and release mechanisms.

Structure prediction of OBPs by homology modelling. Approximately 4 years ago, Zhou (2010) reported that 47 crystal structures of insect OBPs were deposited in the Protein Data Bank (PDB) (http://www.rcsb.org/pdb). At the beginning of 2014, a search with 'odorant binding protein' as keywords returned the structures of OBPs in only eight insect species, despite several hundred OBPs having been reported in an 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 lower resolution of 1.5–2.5 Å and three 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 both OBPs and PBPs in some cases. These crystal structures were solved under different conditions, such as ligand–OBP complex, apo-OBP form and at different pHs. The PBPs from moths often have seven α-helices and three conserved disulphide 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 example, early studies on homology modelling with OBPs were carried out by Honson & Plettner (2006) aiming to determine the three-dimensional arrangement of disulphide 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 disulphide bridge is C2–C5 (i.e. Cys50–Cys108), which is the most easily reduced disulphide bridge by cyanylation using tris(2-carboxyethyl)phosphine, 2-mercaptoethanol and dithiothreitol. The structural models of these proteins suggest that the easy reduction of C2–C5 is the result of a steric property of Cys side chains that are more exposed. Moreover, the conformation of C2–C5 could be influenced strongly by electronic effects from a nearby aspartic acid (Asp106), which stabilizes C5 through a proton transfer.

Steps, limitations and improvements within the homology modelling. Knowledge of 3D protein structures allows an understanding of molecular mechanisms and evolutionary relationships, as well as detailed information about binding sites and ligand–protein interactions. The prediction of 3D protein structure from the 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 starting with studies 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 structural modelling of proteins performs a comparison between the amino acid sequence of target and template(s). The templates are stored in the PDB (http://www.rcsb.org/pdb/home/home.do) and are available using a designation of four digits called the PDB code. An example is the alignment between *Acyrthosiphon pisum* OBP3 as target, and *L. maderae* PBP with code '1ORG

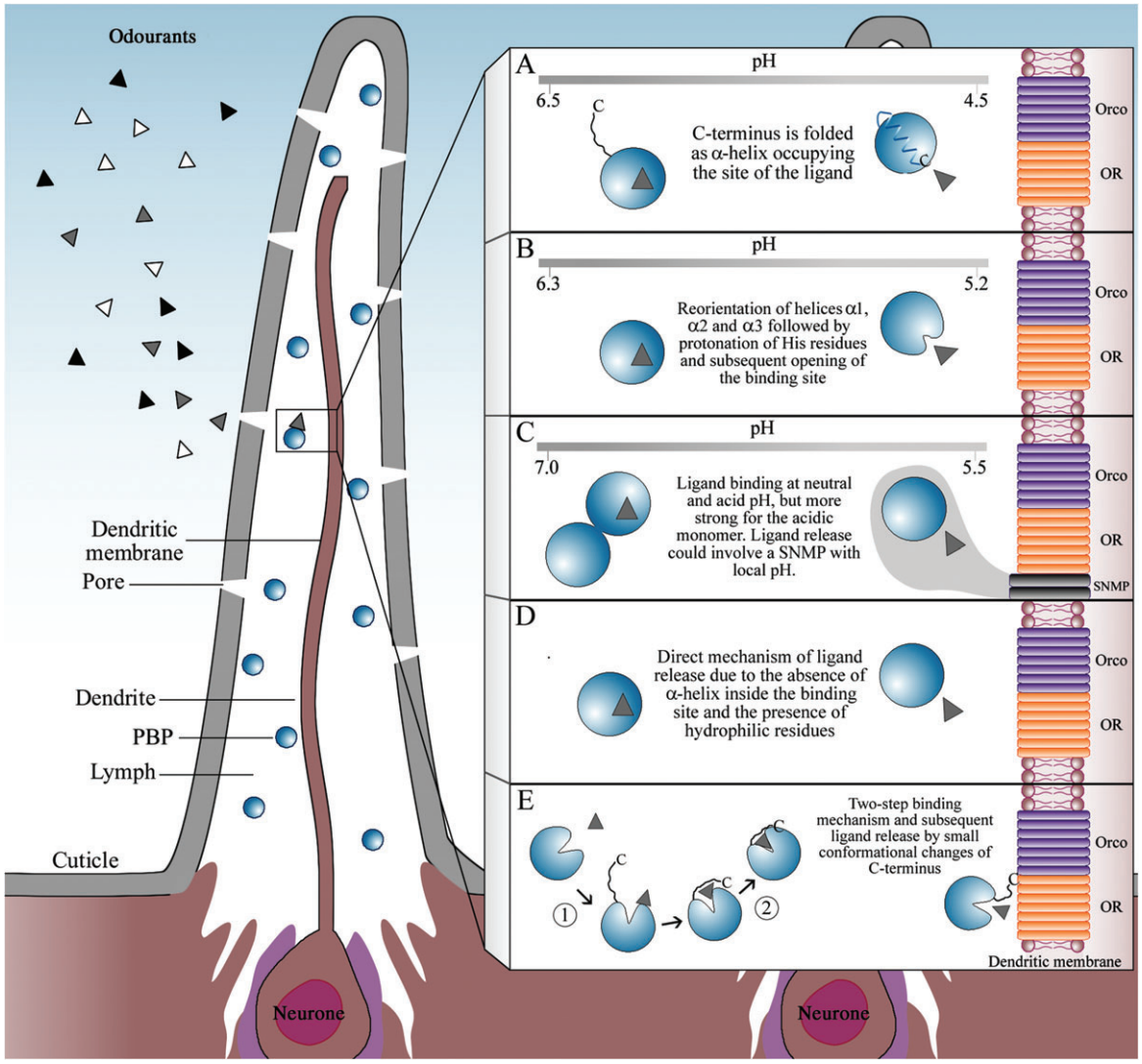

Fig. 2. Schematic representation of the mechanisms of ligand release of moth pheromone-binding proteins (PBPs). Background section represents a commonly occurring multipore sensillum present in moths, such as *Bombyx mori* and *Antheraea polyphemus*. From the external environment, different odorants (△, ▲ and ▲) 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 *B. mori* PBP (BmorPBP) (Horst *et al.*, 2001; Lautenschlager *et al.*, 2005; Leal, 2005). (B) Ligand binding and release mechanism proposed for *A. polyphemus* PBP (ApolPBP1) (Zubkov *et al.*, 2005). (C) Uncommon ligand binding and release mechanism proposed for *Apis mellifera* ASP1 (Pesenti *et al.*, 2009). (D) Ligand binding and release mechanism proposed for *Leucophaea maderae* PBP (LmaPBP) (Lartigue *et al.*, 2003). (E) Ligand binding and release mechanism proposed for *Lymantria dispar* PBPs (LdisPBPs) (Gong *et al.*, 2009, 2010). Inactivation or degradation (by odorant-degrading enzymes) is not represented in this scheme because this is not part of the main discussion of this review. OR, olfactory receptor; Orco, OR co-receptor; SNMP, Sensory neuron membrane protein.

chain A' as template, for the modelling of aphid OBPs (Qiao *et al.*, 2009). Similarly, other studies have used the same template to predict the 3D structure of the OBPs of *A. pisum* and *Sitobion avenae* (Sun *et al.*, 2011; Sun *et al.*, 2012a; Zhong *et al.*, 2012) and of the plant bug *Adelphocoris lineolatus* (Wang *et al.*, 2013a). Moreover, a template can be used to model not only a full-length, but also a part of target (Chang & Swaan, 2006). For example, for modelling of *Culex quinquefasciatus* OBP2, Paramasivan *et al.* (2007) used two templates. The first 15 residues are modelled with the 1DI1 template of *Penicillium roqueforti* synthase and the remainder with 1R5R of *A. mellifera* ASP1. Such templates are selected based on the folding of the target by the GENTHREADER fold recognition server (http://bioinf.cs.ucl.ac.uk/psipred//?program=psipred). Another 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 processes, both of which offer friendly platforms for 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, the server offers the capability to build homology models within a workspace through a web connection (Arnold *et al.*, 2006). On the other hand, the MODELLER software must be downloaded and installed. Although MODELLER calculates a model automatically, it must be provided with an alignment between the target and template, as well as user scripts as commands. Once the model has been built, other software can be used to visualize the model, such as PYMOL (http://www.pymol.org), RASMOL (http://www.rasmol.org), CHIMERA (http://www.cgl.ucsf.edu/chimera/) and SWISS-PDBVIEWER (http://www.spdbv.vital-it.ch/), amongst others. Homology modelling studies of insect OBPs to date are listed in Table 1, including templates and the software used. Apart from homology modelling, there are other types of computing approaches in protein structure predictions, according to several reviews on the subject (Sánchez *et al.*, 2000; Schwede *et al.*, 2007; Ravna & Sylte, 2012). These include threading methods, *ab initio* methods and integrative or hybrid methods. However, homology modelling is the most precise and reliable method currently being used (Bordoli & Schwede, 2012; Ravna & Sylte, 2012). There are numerous reviews on homology modelling that have a similar content structure, each highlighting the steps, precision and limitations of the technique, together with an apparently endless list of available modelling software. Therefore, below, the important concepts in homology modelling are only summarized, with a focus on the ligand affinities of insect OBPs.

Protein structure modelling based on homology between target and template consists of four main steps: template identification, target-template alignment, model building and model refinement and validation (Fig. 3). Template identification is the first step in obtaining a homology model. Such identification is usually carried out with basic local alignment search tool BLAST (http://www.ncbi.nlm.nih.gov/BLAST/). This tool provides regions of similarity and an identity percentage of nucleotide or protein sequences between target and templates. To determine a ligand–protein interaction, a template must have some characteristics, such as high resolution, sharing regions or substructures with target, and be bound as holo-protein (i.e. ligand–protein interaction as biochemically active complex) (Orry & Abagyan, 2012). The percentage sequence identity is an important indicator for establishing the best template. Thus, 30% or more sequence homology is considered good to obtain a precise model. If the sequence identity is less than 30%, the quality of the model decreases (Schwede *et al.*, 2007). Based on the above, the sequence identities of the targets LdisPBP1 and LdisPBP2, with the template BmorPBP (1DQE), are 61% and 48%, respectively (Honson *et al.*, 2003), which has provided a good homology model for both PBPs. However, the study by Honson *et al.* (2003) does 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/), comprising software that evaluates the stereochemical quality of the predicted structure. A special case is reported by Gu *et al.* (2011), in which a reliable homology model for the lucerne plant bug *A. lineolatus* OBP (AlinOBP1) is based on a low sequence identity of 16.8% with the template bombykol-BmorPBP1 complex (1DQE). In this instance, the homology recognition software FUGUE (http://www.tardis.nibio.go.jp/fugue/) was used, which comprises a sequence-structure alignment to find common folds even with low sequence similarity between target and template (Shi *et al.*, 2001). Thus, a *Z*-score is related with a 99% 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 approximately 1 Å root-mean square deviation (RMSD) of the Cα atomic coordinates 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 using software such as CLUSTALW (Thompson *et al.*, 1994), MUSCLE (Edgar, 2004) and T-COFFEE (Notredame *et al.*, 2000), which comprise online bioinformatic tools, where, according to Hang (2008), CLUSTALW (http://www.ebi.ac.uk/Tools/msa/clustalw2/) and MUSCLE (http://www.ebi.ac.uk/Tools/msa/muscle/) have higher speeds in running time compared with T-COFFEE (http://www.ebi.ac.uk/Tools/msa/tcoffee/) for a large number of sequences. The third step of homology modelling is the building of main structures including the core modelling of a structurally conserved region in proteins, loop modelling, and side chain and backbone modelling (Leach, 2001; Ravna & Sylte, 2012). There are two main approaches for building the structure model, as highlighted by Leach (2001); Schwede *et al.* (2007) and Ravna & Sylte (2012). The first approach is called the 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. The second approach for homology model building is 'satisfaction of spatial restraints'. This is an optimization strategy, in which restraints such as atomic distances, solvent and torsion angles are considered. Once an initial protein has been modelled, the refinement must be carried out. 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 MD simulations using force fields, is to obtain a structural model with minimum energy. Force fields estimate the energy associated with each possible conformation of the protein structures modelled to achieve the correct covalent geometry, avoid atomic overlaps, select the nearest conformation to the native structure and assess model quality (Bordner, 2012). The MD approach consists of the simulation of protein motion according to Newton's laws of motion to obtain the

Table 1. Summary of homology modelling studies to determine odorant-binding protein structures of insects.

OBP	Species	Template(s)	Protein Data Bank code	Software	References
OBP	*Locusta migratoria*	*Apis mellifera* PBP	3BFH	DISCOVERY STUDIO 2.0	Jiang *et al.* (2009)
OBP2, OBP3 and OBP7	*Sitobion avenae*	*Leucophaea maderae* PBP	1ORG	SWISS-MODEL	Zhong *et al.* (2012)
PBP1	*Spodoptera litura*	*Antheraea polyphemus* and *Bombyx mori* PBP	1IS8, 1QWV	MODELLER	Liu *et al.* (2012)
OBP	*Acyrthosiphon pisum*	*Leucophaea maderae* PBP	1ORG	SWISS-MODEL	Qiao *et al.* (2009)
OBP4	*Anopheles gambiae*	*Anopheles gambiae* OBP1	2ERB	SWISS-MODEL	Qiao *et al.* (2011)
OBP1, OBP2 and OBP3	*Helicoverpa armigera*	*Bombyx mori* PBP	1DQE	DISCOVERY STUDIO 2.0	Zhang *et al.* (2012a)
OBP5	*Helicoverpa armigera*	*Aedes aegypti* OBP1	3K1E	DISCOVERY STUDIO 2.0	Zhang *et al.* (2012b)
OBP2 and OBP3	*Locusta migratoria*	*Anopheles gambiae* OBP1	2ERB, 1DQE	SWISS-MODEL	Yu *et al.* (2009)
OBP	*Phormia regina*	*Bombyx mori* PBP	–	INSIGHT II	Tsuchihara *et al.* (2005)
OBP3	*Acyrthosiphon pisum*	*Leucophaea maderae* PBP	1ORG	SWISS-MODEL	Sun *et al.* (2011)
OBP10	*Helicoverpa armigera*, *Helicoverpa assulta*	*Culex quinquefasciatus* OBP	3OGN	SWISS-MODEL	Sun *et al.* (2012b)
OBP3 and OBP7	*Acyrthosiphon pisum*	*Leucophaea maderae* PBP and *Drosophila melanogaster* LUSH	1ORG, 3B6X	SWISS-MODEL	Sun *et al.* (2012a)
OBP15 and OBP21	*Apis mellifera*	*Apis mellifera* OBP14	3RZS	MODELLER	Spinelli *et al.* (2012)
OBP	*Drosophila melanogaster*	*Anopheles gambiae* OBP1	2ERB	SWISS-MODEL	Sánchez-Gracia & Rozas (2008)
OBP1	*Anopheles gambiae*	*Anopheles gambiae* OBP1	2ERB, 3N7H	SWISS-MODEL	Rusconi *et al.* (2012)
OBP2	*Culex quinquefasciatus*	*Apis mellifera* ASP1 and *Penicillium roqueforti* synthase	1R5R, 1DI1	MODELLER	Paramasivan *et al.* (2007)
OBP1 and OBP2	*Scleroderma guani*	*Pyrococcus horikoshii* PH1010 and *Apis mellifera* ASP1	3D76, 3BJH	SWISS-MODEL	Li *et al.* (2011)
PBP1	*Lymantria dispar*	*Bombyx mori* PBP	1DQE	SWISS-MODEL	Honson *et al.* (2003)
PBP1 and PBP2	*Lymantria dispar*	*Bombyx mori* PBP	1DQE	SWISS-MODEL	Honson & Plettner (2006)
OBP1	*Adelphocoris lineolatus*	*Bombyx mori* PBP	1DQE	DISCOVERY STUDIO 2.0	Gu *et al.* (2011)
PBP1	*Mamestra brassicae*	*Bombyx mori* PBP	1DQE	TURBO-FRODO	Campanacci *et al.* (2001)
OBP1	*Anopheles gambiae*	*Anopheles gambiae* OBP1	2ERB	–	Biessmann *et al.* (2010)
OBP1	*Locusta migratoria*	*Bombyx mori* PBP	1DQE	INSIGHT/HOMOLOGY	Ban *et al.* (2003)
PBP3	*Ostrinia* spp.	*Bombyx mori* GOBP2, *Antheraea polyphemus* PBP and *Bombyx mori* PBP	2WCJ, 1QWV, 1DQE	MODELLER	Allen & Wanner (2011)
PBP1, PBP2 and PBP3	*Helicoverpa armigera* and *Helicoverpa assulta*	*Bombyx mori* PBPs	1DQE, 2FJY	SWISS-MODEL	Guo *et al.* (2012)
OBP5	*Adelphocoris lineolatus*	*Drosophila melanogaster* LUSH	1OOI	DISCOVERY STUDIO 2.0	Wang *et al.* (2013b)
PBP1 and PBP2	*Lymantria dispar*	*Antheraea polyphemus* PBP	1QWV	DISCOVERY STUDIO 3.0	Yu *et al.* (2012)
PBP1 and PBP2	*Lymantria dispar*	*Bombyx mori* GOBP2	2WCJ	SWISS-MODEL	Yu & Plettner (2013)
OBP2	*Apis cerana*	*Apis mellifera* OBP2	1TUJ	SWISS-MODEL	Li *et al.* (2013)
OBP7	*Helicoverpa armigera*	*Anopheles gambiae* OBP1	3N7H	SWISS-MODEL	Sun *et al.* (2013)
OBP37 and OBP39	*Aedes albopictus*	*Culex quinquefasciatus OBP1 and Aedes aegypti OBP1*	2L2C, 3K1E	SWISS-MODEL	Deng *et al.* (2013)
OBP3 and OBP4	*Holotrichia oblita*	*Anopheles gambiae* OBP20	3VB1	SWISS-MODEL	Wang *et al.* (2013a)
OBP1	*Holotrichia oblita*	*Anopheles gambiae* OBP1	2ERB	SWISS-MODEL	Zhuang *et al.* (2014)
ASP2	*Apis cerana*	–	–	I-TASSER	Lu *et al.* (2014)

ASP, *A. mellifera*; OBP, odorant-binding protein; PBP, pheromone-binding protein.

energy of a particular protein conformation through calculations by force fields, which is usually performed with the software AMBER (Weiner & Kollman, 1981), GROMACS (Berendsen *et al.*, 1995) and NAMD (Phillips *et al.*, 2005). It is noteworthy that MD simulations are used not only for refinement of homology models, but also to determine structural changes in proteins. For example, MD simulation has been used to assess the binding dynamics and specificity of insect PBPs. Thus, Gräter *et al.* (2006b) report a similar binding affinity for bombykol and bombykal to BmorPBP1 through MD simulations that considered 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 shows 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) provide a putative pH-induced ligand-releasing mechanism, which is predicted using MD simulation. Their findings suggest that the OBP1 of *C. quinquefasciatus* (CquiOBP1) releases the mosquito oviposition pheromone at low pH as a result of the cleavage of hydrogen bonds in the binding site. These simulations have been performed on crystal structures of OBPs. However, Yu *et al.* (2012) report a constant pH MD simulation on the modelled structures of *L. dispar* PBPs (LdisPBPs) at both pH 7.3 and pH 5.5. Subsequently, the findings were corroborated by circular dichroism, where conformational changes are revealed (Yu *et al.*, 2012).

Overall, MD simulation involves classical mechanics based on Newtonian physics to reduce the computational cost (Durrant & McCammon, 2011). By contrast to quantum mechanics,

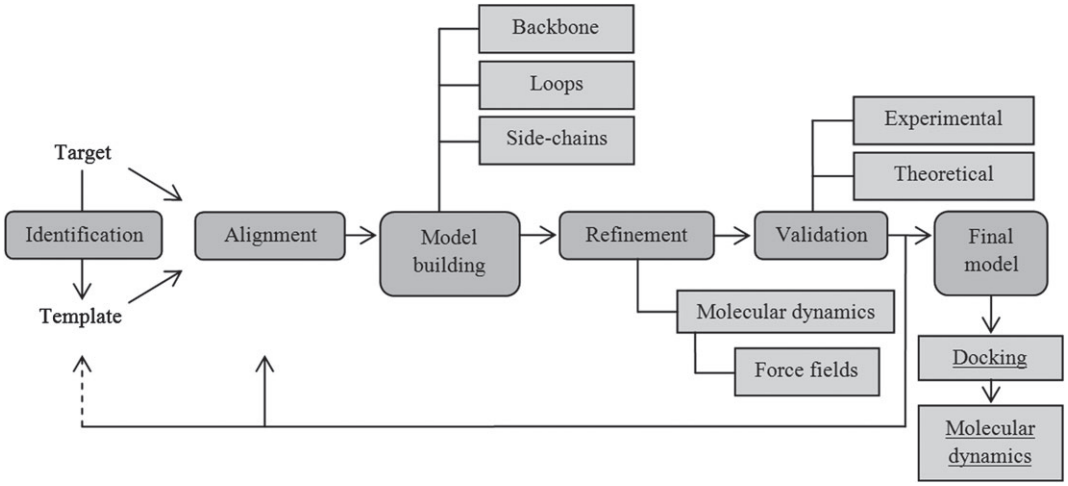

Fig. 3. General building scheme for protein structures by homology modelling, as applied to insect odorant-binding proteins to date. Underlined section indicates an optional step when ligand binding is studied.

classical mechanics or molecular mechanics (MM) works to find the global minimum of free energy in a protein structure to describe a conformation, considering atoms and bonds as single particles and springs, respectively (Bordner, 2012). 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). Similarly, a refined homology model of *Locusta migratoria* OBP1 (LmigOBP1) was obtained by Jiang *et al.* (2009) using CHARMM force field software for macromolecular energy minimization and dynamics calculation (Brooks *et al.*, 1983). Despite these studies on the refinement of predicted OBPs structures, there is no information available on how long these simulations were required to complete. Currently, MD at long-timescale (e.g. nanoseconds) appears to be useful for refinement of homology models from proteins of small or medium size (Fan & Mark, 2004; Raval *et al.*, 2012), such as insect OBPs.

Model validation is a crucial step to ensure its quality. There are two approaches for model validation: experimental and theoretical. The experimental approach compares the results from modelling studies with biological observations, and establishes the model quality. If the biological results are in accordance with the homology model, it is considered correct (Chang & Swaan, 2006; Ravna & Sylte, 2012). The theoretical method for 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 described above are usually repeated iteratively until the best homology model is obtained.

Insect OBPs are flexible and this property is likely to be very important for ligand recognition patterns and function. 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 below. It is important to note that homology models and experimental protein structures are only a snapshot of a specific conformation. Therefore, all results that are determined represent only one state of the protein without information about its dynamic. Thus, studies on drug discovery have been directed towards determining the stability of the ligand–protein complexes (Alonso *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). Five well known inhibitors (suramin, mol-6, sirtinol, 67 and nf675) were docked to SIRT2. The plotted RMSD values reveal 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 being highlighted as the most reliable and precise method for protein structure prediction, there are usually three limitations. These are template selection, precise target-template alignment and imprecisions in refinement. The limitations are considered the main sources of errors in homology modelling (Arnold *et al.*, 2006; Larsson *et al.*, 2008). Selection of an appropriate template depends on the evolutionary relationship with the target and, hence, structural and functional divergence from the target. To overcome the imprecisions from template selection and sequence alignment, the ideal approach is to select the template with the highest sequence identity ($\geq 30\%$) to target, use MSA with manual intervention if necessary and use software or servers such as FUGUE and genTHREADER for low sequence identities. Every 2 years, the Critical Assessment (http://www.predictioncenter.org/index.cgi) of Protein Structure Prediction (CASP) centre assesses and reports the

progress of modelling techniques. If the template selection for a full-length target is difficult because the sequence identity is too low (less than 30%) or the template resolution is too low, the modelling of some conserved regions is feasible. For example, homology modelling of α4β2 nicotinic acetylcholine receptor (nAChR) with homopentameric acetylcholine binding protein (AChBP) from the sea slug *Aplysia californica* as template gave a sequence identity of 18–20%. Despite such low identity, the shared features between the target nAChR and template AChBP, such as a similar fold and highly conserved residues in the binding site (Hansen *et al.* (2005), allowed an nAChR model to be predicted using AChBP as a suitable template (Iturriaga-Vásquez *et al.*, 2010).

It is important to obtain precise alignment between targets and templates. Distant evolutionary relationships, specific regions that do not align reliably and MSA programmes themselves could lead to imprecision in the alignment. Arnold *et al.* (2006) note that, when the imprecision of a target-template alignment cannot be determined, visual and manual intervention can improve the model quality significantly. If the precision of a sequence alignment decreases, the percentage of equivalent Cα atoms for the superimposed target and template structures increases, as measured by RMSD. Raval *et al.* (2012) note that the results of homology modelling have a low resolution by imprecise refinement during MD simulation with errors reflected in the values of 3 Å for atomic coordinates. The timescale of MD application and the precision of force fields could also be the source of errors. However, the main failure of MD within the refinement process appears to be caused by force field errors (Raval *et al.*, 2012). A restricted conformational sampling for the initial homology model was proposed 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.

Ligand binding affinity to homology models of OBPs: methods and interaction. Many types of biological effects in vertebrates and invertebrates are dependent on ligand–protein interactions. The transport function of OBPs is carried out through interactions between ligands and the binding sites of these proteins. As mentioned previously, homology modelling, as a quick and inexpensive method, allows the prediction of the binding site in protein structures. Once a good quality homology model is obtained, it is possible to simulate 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, a step that can be achieved by searching for ligands in databases (e.g. PubChem; http://www.pubchem.ncbi.nlm.nih.gov/) or by drawing them *de novo* using appropriate software, such as CHEM3D (http://www.cambridgesoft.com) or SPARTAN (http://www.wavefun.com). Then, molecular docking is applied over a sufficient number of times 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), MD docking (McCammon *et al.*, 1977; Mangoni *et al.*, 1999), genetic algorithms (Morris *et al.*, 1998) and the ligand fragment-based method (Rarey *et al.*, 1996). The use of these docking algorithms depends on the success of its application for a particular protein family. An example is the study of the affinity of volatile compounds to the LmigOBP1 (Jiang *et al.*, 2009). The specific binding between LmigOBP1 and pentadecanol (the ligand with the highest affinity) was carried out using CDOCK (Wu *et al.*, 2003) through a MD algorithm. Docking simulations in this work suggest that asparagine (Asn74) is a key amino acid in the binding site of LmigOBP1. Another approach 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 docking simulations on BmorGOBP2, using a genetic algorithm with AUTODOCK (http://www. autodock.scripps.edu/). Their findings show that two hydrogen bonds are formed with Arg110 and Glu98 in bombykol–BmorGOPB2 complexes, which is consistent with the 3D structure of the bombykol–BmorGOPB2 complex reported by Zhou *et al.* (2009). Furthermore, using seven analogue ligands, docking simulations predict that (10*E*,12*Z*)-hexadecadienyl acetate has a 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 of high-throughput screening (HTS), allowing hit rates that are 10- to 1000-fold higher than HTS. Besides the ability to predict the presence of cavities in proteins for ligand binding, molecular docking methods can also indicate amino acid residues that form the cavities. Likewise, they provide a fine selection of ligands from extensive libraries, thus allowing a deeper analysis of noncovalent interactions such as van der Waals interactions. For example, Jiang *et al.* (2009) report van der Waals interactions for hydrophobic residues in the binding pocket of LmigOBP1. They suggest that these interactions can occur through benzene rings even with Tyr109 and Tyr117 as hydrophilic residues. To date, only 12 studies that have performed molecular docking on the homology models of insect OBPs, and which include the relevant information about key residues for ligand binding, have been published (Honson *et al.*, 2003; Jiang *et al.*, 2009; Biessmann *et al.*, 2010; Rusconi *et al.*, 2012; Yu et al., 2012; Zhang *et al.*, 2012; Li *et al.*, 2013; Wang *et al.*, 2013b; Yu & Plettner, 2013; Lu *et al.*, 2014; Jayanthi, P.D.K. *et al.*, 2014; Zhuang *et al.*, 2014) (Table 2). Although different software has been used for docking, the most recent studies appear to be coinciding in their development of detailed analysis, with more parameters being included, such as energy minimizations, theoretical evaluations and more sophisticated algorithms. An example is the comprehensive study performed by Jayanthi, P.D.K. *et al.* (2014) for the rapid screening

Table 2. Molecular docking on homology models of insect odorant-binding proteins.

OBP	Ligands	Software	Reference
Anopheles gambiae OBP1	Indole	QUANTA-CHARMM	Biessmann *et al.* (2010)
Locusta migratoria OBP	Pentadecanol	CDOCK	Jiang *et al.* (2009)
Helicoverpa armigera OBP1, OBP2 and OBP3	(*Z*)-9-Hexadecenal and (*Z*)-11-hexadecenal	–	Zhang *et al.* (2012a)
Lymantria dispar PBP1	Aziridine	–	Honson *et al.* (2003)
Anopheles gambiae OBP1	*N*-Phenyl-1-naphylamine (generic ligand)	AUTODOCK 4.0	Rusconi *et al.* (2012)
Lymantria 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)
Lymantria dispar PBP1 and PBP2	(+) and (−) disparlure	MOLECULAR OPERATING ENVIRONMENT (MOE)	Yu & Plettner (2013)
Apis cerana OBP2	*N*-Phenyl-1-naphylamine (generic ligand)	MOLEGRO VIRTUAL DOCKER 4.2	Li *et al.* (2013)
Adelphocoris lineolatus OBP5	*cis*-Nerolidol	CDOCK	Wang *et al.* (2013b)
Holotrichia oblita OBP1	Hexyl benzoate, β-ionone, cinnamaldehyde and myrcene	AUTODOCK 4.0	Zhuang *et al.* (2014)
Apis 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)

OBP, odorant-binding protein; PBP, pheromone-binding protein.

of active semiochemicals using a computational reverse chemical ecology approach on a GOBP of the oriental fruit fly *Bactrocera dorsalis*.

Although the docking method is a powerful tool for predicting ligand–protein interactions, other experimental methods are being used in conjunction with homology modelling, such as fluorescence competitive binding assays (FBA), radioactively labelled ligand (RLL) studies and site-directed mutagenesis. Fluorescence competitive binding assays and RLL comprise two technologies for receptor-ligand binding investigations. By employing radio-isotopic labels such as ^{3}H, RLL can be used in the determination of receptor distribution and subtypes, screening of ligands and quantification. However, this type of 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 the affinity to a wide range of compounds of insect OBPs (Zhou *et al.*, 2004; de Jong *et al.*, 2005). Fluorescence competitive binding assays provide a ligand screening with the use of fluorophores (commonly 1-*N*-phenylnaphthylamine) and quantification through displacement of the fluorophore by ligands. An example is the FBA approach used by Zhang *et al.*, 2012b for 113 compounds, using homology modelling as a complement. Although these binding assays indicate that (*E*)-β-farnesene, ethyl butyrate, ethyl heptanoate and acetic acid 2-methylbutyl ester are the best ligands for *H. armigera* OBP5, the homology 3D model reveals 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 examination of ligand–protein interactions, the role of nonconserved residues in the binding sites and ligand binding modes (Cavasotto & Phatak, 2009). The ligand binding affinity can be altered through site-directed mutagenesis, as in *A. gambiae* OBP1 (Rusconi *et al.*, 2012). Also, physicochemical properties can be changed through mutations of the residues in the binding site, such as Phe59, Met84, His111, Trp114, Tyr122, Phe123 and Leu124. Thus, the mutated OBPs exhibit a lower binding affinity in the binding assays. More recently, Zhuang *et al.* (2014) report 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, it was possible to demonstrate the role of Tyr111 in the binding of β-ionone, myrcene, hexyl benzoate and cinnamaldehyde.

Conclusions

Research into insect OBPs has clearly progressed and is ready to offer alternative approaches for ligand screening to identify novel semiochemicals and analogues in the hope for their eventual use in pest management. As a result of intense studies on the sensory system of insects, OBPs as semiochemical carrier proteins are attractive targets for ligand screening, which is the core target of reverse chemical ecology. The absence of crystal structures of OBPs for a broader range of insect species has limited progress thus far. Homology modelling can provide information about the 3D structure of these proteins and predict the amino acid residues that could be involved in the formation of ligand–protein complexes. It is important to understand, and be aware of, the various factors that may 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 lepidopteran OBPs can provide precise homology models and, along with molecular docking (as a virtual method

to predict best ligands), may represent a great leap forward in the search for super-ligands to manipulate insect behaviours. Finally, it is worth noting that 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 carried out and super-ligands need to be screened with insects to identify potential pest control agents.

Acknowledgements

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

References

Affonso, R., Guimarães, A.P., Oliveira, A.A. *et al.* (2013) Applications of molecular modeling in the design of new insect repellents targeting the odorant binding protein of *Anopheles gambiae*. *Journal of the Brazilian Chemical Society*, **24**, 473–482.

Allen, J.E. & Wanner, K.W. (2011) Asian corn borer pheromone binding protein 3, a candidate for evolving specificity to the 12-tetradecenyl acetate sex pheromone. *Insect Biochemistry and Molecular Biology*, **41**, 141–149.

Alonso, H., Bliznyuk, A.A. & Gready, J.E. (2006) Combinding docking and molecular dynamic simulations in drug design. *Medicinal Research Reviews*, **26**, 531–568.

Arnold, K., Bordoli, L., Kopp, J. & Schwede, T. (2006) The SWISS-MODEL Workspace: a web-based environment for protein structure homology modelling. *Bioinformatics*, **22**, 195–201.

Ban, L., Scaloni, A., D'Ambrosio, C. *et al.* (2003) Biochemical characterization and bacterial expression of an odorant-binding protein from *Locusta migratoria*. *Cellular and Molecular Life Sciences*, **60**, 390–400.

Berendsen, H.J.C., van der Spoel, D. & van Drunen, R. (1995) GROMACS: a message-passing parallel molecular dynamics implementation. *Computer Physics Communications*, **91**, 43–56.

Bette, S., Breer, H. & Krieger, J. (2002) Probing a pheromone binding protein of the silkmoth *Antheraea polyphemus* by endogenous tryptophan fluorescence. *Insect Biochemistry and Molecular Biology*, **32**, 241–246.

Biessmann, H., Andronopoulou, E., Biessmann, M.R. *et al.* (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 modeling. *Homology Modeling: Methods and Protocols* (ed. by A. J. W. Orry and R. Abagyan), pp. 83-106. Molsoft L.L.C., San Diego, California.

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

Briand, L., Eloit, C., Nespoulous, C. *et al.* (2002) Evidence of an odorant-binding protein in the human olfactory mucus: location, structural characterization, and odorant-binding properties. *Biochemistry*, **41**, 7241–7252.

Brimau, F., Cornard, J.-P., Le Danvie, C. *et al.* (2010) Binding specificity of recombinante odorant-binding protein isoforms is driven by phosphorylation. *Journal of Chemical Ecology*, **36**, 801–813.

Brooks, B.R., Bruccoleri, R.E., Olafson, B.D. *et al.* (1983) CHARMM: a program for macromolecular energy minimization and dynamics calculations. *Journal of Computational Chemistry*, **4**, 187–217.

Campanacci, V., Krieger, J., Bette, S. *et al.* (2001) Revisiting the specificity of *Mamestra brassicae* and *Antheraea polyphemus* pheromone-binding proteins with a fluorescence binding assay. *Journal of Biological Chemistry*, **276**, 20078–20084.

Cavasotto, C.N. & Phatak, S.S. (2009) Homology modeling in drug discovery: current trends and applications. *Drug Discovery Today*, **14**, 676–683.

Chang, C. & Swaan, P.W. (2006) Computational approaches to modeling drug transporters. *European Journal of Pharmaceutical Sciences*, **27**, 411–424.

Chu, W.-T., Zhang, J.-L., Zheng, Q.-C. *et al.* (2013) Constant pH molecular dynamics (CpHMD) and molecular docking studies of CquiOBP1 pH-induced ligand releasing mechanism. *Journal of Molecular Modeling*, **19**, 1301–1309.

Damberger, F.F., Ishida, Y., Leal, W.S. & 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. *Journal of Molecular Biology*, **373**, 811–819.

Deng, Y., Yan, H., Gu, J. *et al.* (2013) Molecular and functional characterization of odorant-binding protein genes in an invasive vector mosquito, *Aedes albopictus*. *PLoS ONE*, **8**, e68836. DOI:10.1371/journal.pone.0068836.

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

Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, **32**, 1792–1797.

Eswar, N., Marti-Renom, M.A., Webb, B. *et al.* (2006) Comparative protein structure modeling with MODELLER. *Current Protocols in Bioinformatics* (ed. by A. Baxevanis), pp. 5.6.1-5.6.30. John Wiley & Sons, Inc., New York, New York.

Fan, H. & Mark, A.E. (2004) Refinement of homology-based protein structures by molecular dynamics simulation techniques. *Protein Science*, **13**, 211–220.

Gajendrarao, P., Krishnamoorthy, N., Sakkiah, S. *et al.* (2010) Molecular modeling study on orphan human protein CYP4A22 for identification of potential ligand binding site. *Journal of Molecular Graphics and Modelling*, **28**, 524–532.

Ginalski, K. (2006) Comparative modeling for protein structure prediction. *Current Opinion in Structural Biology*, **16**, 172–177.

Gomez-Diaz, C., Reina, J.H., Cambillau, C. & Benton, R. (2013) Ligands for pheromone-sensing neurons are not conformationally activated odorant binding proteins. *PLoS Biology*, **11**, e1001546. DOI:10.1371/journal.pbio.1001546.

Gong, Y., Pace, T.C.S., Castillo, C. *et al.* (2009) Ligand-interaction kinetics of the pheromone-binding protein from the gypsy moth, *L. dispar*: insights into the mechanism of binding and release. *Chemistry & Biology*, **16**, 162–172.

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

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

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

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

Grolli, S., Merli, E., Conti, V. *et al.* (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.

Gu, S., Wang, W., Wang, G. *et al.* (2011) Functional characterization and immunolocalization of odorant binding protein 1 in the lucerne plant bug, *Adelphocoris lineolatus* (Goeze). *Archives of Biochemistry and Physiology*, **77**, 81–98.

Guo, H., Huang, L., Pelosi, P. & Wang, C. (2012) Three pheromone-binding proteins help segregation between two *Helicoverpa* species utilizing the same pheromone components. *Insect Biochemistry and Molecular Biology*, **42**, 708–716.

Hang, N.T. (2008) Comparison of multiple sequence alignment programs in practise. Master of Bioinformatics Thesis, University of Århus.

Hansen, S.B., Sulzenbacher, G., Huxford, T. *et al.* (2005) Structures of *Aplysia* AChBP complexes with nicotinic agonists and antagonists reveal distinctive binding interfaces and conformations. *EMBO Journal*, **24**, 3635–3646.

He, X., Tzotzos, G., Woodcock, C. *et al.* (2010) Binding of the general odorant binding protein of *Bombyx mori* BmorGOBP2 to the moth sex pheromone components. *Journal of Chemical Ecology*, **36**, 1293–1305.

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

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

Honson, N., Johnson, M.A., Oliver, J.E. *et al.* (2003) Structure–activity studies with pheromone-binding proteins of the gypsy moth, *Lymantria dispar*. *Chemical Senses*, **28**, 479–489.

Honson, N., Gong, Y. & Plettner, E. (2005) Structure and function of insect odorant and pheromone-binding proteins (OBPs and PBPs) and chemosensory-specific proteins (CSPs). *Recent Advances in Phytochemistry*, **39**, 268–277.

Hooper, A.M., Dufour, S., He, X. *et al.* (2009) High-throughput ESI-MS analysis of binding between the *Bombyx mori* pheromone-binding protein BmorPBP1, its pheromone components and some analogues. *Chemical Communications*, **14**, 5725–5727.

Horst, R., Damberger, F., Luginbühl, P. *et al.* (2001) NMR structure reveals intramolecular regulation mechanism for pheromone binding and release. *Proceedings of the National Academy of Sciences of the United States of America*, **98**, 14374–14379.

Ishida, Y. & Leal, W.S. (2002) Cloning of putative odorant-degrading enzyme and integumental esterase cDNAs from the wild silkmoth, *Antheraea polyphemus*. *Insect Biochemistry and Molecular Biology*, **32**, 1775–1780.

Iturriaga-Vásquez, P., Carbone, A., García-Beltrán, O. *et al.* (2010) Molecular determinants for competitive Inhibition of α4β2 nicotinic acetylcholine receptors. *Molecular Pharmacology*, **78**, 366–375.

Jacquin-Joly, E. & Merlin, C. (2004) Insects olfactory receptors: contributions of molecular biology to chemical ecology. *Journal of Chemical Ecology*, **30**, 2359–2397.

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

Jayanthi, P.D.K., Kempraj, V., Aurade, R.M. *et al.* (2014) Computational reverse chemical ecology: virtual screening and predicting behaviorally active semiochemicals for *Bactrocera dorsalis*. *BMC Genomics*, **15**, 209. (also cited as P.D. *et al.*, 2014)

Jiang, Q., Wang, W., Zhang, Z. & Zhang, L. (2009) Binding specificity of locust odorant binding protein and its key binding site for initial recognition of alcohols. *Insect Biochemistry and Molecular Biology*, **39**, 440–447.

de Jong, L.A.A., Uges, D.R.A., Franke, J.P. & Bischoff, R. (2005) Receptor-ligand binding assays: technologies and applications. *Journal of Chromatography B*, **829**, 1–25.

Kaissling, K.-E. (2013) Kinetics of olfactory responses might largely depend on the odorant-receptor interaction and the odorant deactivation postulated for flux detectors. *Journal of Comparative Physiology A*, **199**, 879–896.

Katre, U.V., Mazumder, S., Prusti, R.K. & Mohanty, S. (2009) Ligand binding turns moth pheromone-binding protein into a pH sensor. *Journal of Biological Chemistry*, **284**, 32167–32177.

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

Klusak, V., Havlas, Z., Rulíšek, L. *et al.* (2003) Sexual attraction in the silkworm moth: nature of binding of bombykol in pheromone binding protein-an ab initio study. *Chemistry & Biology*, **10**, 331–340.

Kolb, P., Ferreira, R.S., Irwin, J.J. & Shoichet, B.K. (2009) Docking and chemoinformatic screens for new ligands and targets. *Current Opinion in Biotechnology*, **20**, 429–436.

Kowcun, A., Honson, N. & Plettner, E. (2001) Olfaction in the gypsy moth, *Lymantria dispar*: effect of pH, ionic strength, and reductants on pheromone transport by pheromone-binding proteins. *Journal of Biological Chemistry*, **276**, 44770–44776.

Krieger, J., von Nickisch-Rosenegk, E., Mameli, M. *et al.* (1996) Binding proteins from the antennae of *Bombyx mori*. *Insect Biochemistry and Molecular Biology*, **26**, 297–307.

Kruse, S.W., Zhao, R., Smith, D.P. & Jones, D.N.M. (2003) Structure of a specific alcohol-binding site defined by the odorant binding protein LUSH from *Drosophila melanogaster*. *Nature Structural Biology*, **10**, 694–700.

Lagarde, A., Spinelli, S., Tegoni, M. *et al.* (2011) The crystal structure of odorant binding protein 7 from *Anopheles gambiae* exhibits an outstanding adaptability of its binding site. *Journal of Molecular Biology*, **414**, 401–412.

Larsson, P., Wallner, B., Lindahl, E. & Elofsson, A. (2008) Using multiple templates to improve quality of homology models in automated homology modeling. *Protein Science*, **17**, 990–1002.

Lartigue, A., Gruez, A., Spinelli, S. *et al.* (2003) The crystal structure of a cockroach pheromone-binding protein suggests a new ligand binding and release mechanism. *Journal of Biological Chemistry*, **278**, 30213–30218.

Laughlin, J.D., Ha, T.S., Jones, D.N.M. & 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. & Clardy, J. (2005) Coil-to-helix transition and ligand release of *Bombyx mori* pheromone-binding protein. *Biochemical and Biophysical Research Communications*, **335**, 1044–1050.

Leach, A.R. (2001) Protein structure prediction, sequence analysis and protein folding. *Molecular Modeling Principles and Applications* (ed. by A. R. Leach), pp. 509-558. Pearson Education Limited, U.K.

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, U.K.

Leal, W.S. (2005) Pheromone Reception. *Topics in Current Chemistry*, **240**, 1–36.

Leal, W.S. (2013) Odorant reception in insects: roles of receptors, binding proteins, and degrading enzymes. *Annual Review of Entomology*, **58**, 373–391.

Leite, N.R., Krogh, R., Xu, W. *et al.* (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. & Guittet, E. (2009) Structural basis of the broad specificity of a general odorant-binding protein from honeybee. *Biochemistry*, **48**, 2431–2441.

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

Li, X., Lu, D., Liu, X. *et al.* (2011) Ultrastructural characterization of olfactory sensilla and immunolocalization of odorant binding and chemosensory proteins from an ectoparasitoid *Scleroderma guani* (Hymenoptera: Bethylidae). *International Journal of Biological Sciences*, **7**, 848–868.

Li, H.L., Zhang, L.Y., Zhuang, S.L. *et al.* (2013) Fluorescence investigation on the interaction of a prevalent competitive fluorescent probe with entomic odorant binding protein. *Spectroscopy Letters*, **46**, 527–534.

Liu, N., He, P. & Dong, S. (2012) Binding properties of pheromone-binding protein 1 from the common cutworm *Spodoptera litura*. *Comparative Biochemistry and Physiology - Part B*, **161**, 295–302.

Lu, Y., Li, H., Zhuang, S. *et al.* (2014) Olfactory biosensor using odorant-binding proteins from honeybee: ligands of floral odors and pheromones detection by electrochemical impedance. *Sensors and Actuators B: Chemical*, **193**, 420–427.

Löbel, D., Jacob, M., Völkner, M. & Breer, H. (2002) Odorants of different chemical classes interact with distinct odorant binding protein subtypes. *Chemical Senses*, **27**, 39–44.

Maida, R., Steinbrecht, A., Ziegelberger, G. & Pelosi, P. (1993) The pheromone binding protein of *Bombyx mori*: purification, characterization and immunocytochemical localization. *Insect Biochemistry and Molecular Biology*, **23**, 243–253.

Maida, R., Krieger, J., Gebauer, T. *et al.* (2000) Three pheromone-binding proteins in olfactory sensilla of the two silkmoth species *Antheraea polyphemus* and *Antheraea pernyi*. *European Journal of Biochemistry*, **267**, 2899–2908.

Maida, R., Ziegelberger, G. & Kaissling, K.-E. (2003) Ligand binding to six recombinant pheromone-binding proteins of *Antheraea polyphemus* and *Antheraea pernyi*. *Journal of Comparative Physiology B*, **173**, 565–573.

Mangoni, M., Roccatano, D. & 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. *et al.* (2000) Comparative protein structure modeling of genes and genomes. *Annual Review of Biophysics and Biomolecular Structure*, **29**, 291–325.

McCammon, J.A., Gelin, B.R. & Karplus, M. (1977) Dynamics of folded proteins. *Nature*, **267**, 585–590.

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N. & Teller, A.H. (1953) Equation of state calculations by fast computing machines. *Journal of Chemical Physics*, **21**, 1087–1092.

Morris, G.M., Goodsell, D.S., Halliday, R.S. *et al.* (1998) Automated docking using a lamarckian genetic algorithm and an empirical binding free energy function. *Journal of Computational Chemistry*, **19**, 1639–1662.

Nespoulous, C., Briand, L., Delage, M.-M. *et al.* (2004) Odorant binding and conformational changes of a rat odorant-binding protein. *Chemical Senses*, **29**, 189–198.

Notredame, C., Higgins, D.G. & Heringa, J. (2000) T-Coffee: a novel method for fast and accurate multiple sequence alignment. *Journal of Molecular Biology*, **302**, 205–217.

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

Paas, Y., Devillers-Thiéry, A., Teichberg, V.I. *et al.* (2000) How well can molecular modeling predict the crystal structure: the case of the ligand-binding domain of glutamate receptors. *Trends in Pharmacological Sciences*, **21**, 87–92.

Paramasivan, R., Sivaperumal, R., Dhananjeyan, K.J. *et al.* (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 Biology*, **7**, 1–6.

Pelosi, P. (2001) The role of perireceptor events in vertebrate olfaction. *Cellular and Molecular Life Sciences*, **58**, 503–509.

Pelosi, P., Calvello, M. & Ban, L. (2006) Diversity of odorant-binding proteins and chemosensory proteins in insects. *Chemical Senses*, **30**, 291–292.

Pesenti, M.E., Spinelli, S., Bezirard, V. *et al.* (2009) Queen bee pheromone binding protein pH-induced domain swapping favors pheromone release. *Journal of Molecular Biology*, **390**, 981–990.

Phillips, J.C., Braun, R., Wang, W. *et al.* (2005) Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry*, **26**, 1781–1802.

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

Qiao, H., Tuccori, E., He, X. *et al.* (2009) Discrimination of alarm pheromone (*E*)-β-farnesene by aphid odorant-binding proteins. *Insect Biochemistry and Molecular Biology*, **39**, 414–419.

Qiao, H., He, X., Schymura, D. *et al.* (2011) Cooperative interactions between odorant-binding proteins of *Anopheles gambiae*. *Cellular and Molecular Life Sciences*, **68**, 1799–1813.

Rarey, M., Kramer, B., Lengauer, T. & Klebe, G. (1996) A fast flexible docking method using an incremental construction algorithm. *Journal of Molecular Biology*, **261**, 470–489.

Raval, A., Piana, S., Eastwood, M.P. *et al.* (2012) Refinement of protein structure homology models via long, all-atom molecular dynamics simulations. *Proteins*, **80**, 2071–2079.

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

Rusconi, B., Maranhao, A.C., Fuhrer, J.P. *et al.* (2012) Mapping the *Anopheles gambiae* odorant binding protein 1 (AgamOBP1) using modeling techniques, site directed mutagenesis, circular dichroism and ligand binding assays. *Biochimica et Biophysica Acta*, **1824**, 947–953.

Sakkiah, S., Arooj, M., Kumar, M.R. *et al.* (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. *et al.* (2000) Protein structure modeling for structural genomics. *Nature Structural Biology*, **7**, 986–990.

Sánchez-Gracia, A. & 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 Evolutionary Biology*, **8**, 323.

Sandler, B.H., Nikonova, L., Leal, W.S. & Clardy, J. (2000) Sexual attraction in the silkworm moth: structure of the pheromone-binding protein-bombykol complex. *Chemistry & Biology*, **7**, 143–151.

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

Scott, W.R.P., Hünenberger, P.H., Tironi, I.G. *et al.* (1999) The GROMOS biomolecular simulation program package. *Journal of Physical Chemistry A*, **103**, 3596–3607.

Shi, J., Blundell, T.L. & Mizuguchi, K. (2001) FUGUE: sequence-structure homology recognition using environment-specific substitution tables and structure-dependent gap penalties. *Journal of Molecular Biology*, **310**, 243–257.

Sokkar, P., Mohandass, S. & Ramachandran, M. (2011) Multiple templates-based homology modeling enhances structure quality of AT1 receptor: validation by molecular dynamics and antagonist docking. *Journal of Molecular Modeling*, **17**, 1565–1577.

Spinelli, S., Lagarde, A., Iovinella, I. *et al.* (2012) Crystal structure of *Apis mellifera* OBP14, a C-minus odorant-binding protein, and its complexes with odorant molecules. *Insect Biochemistry and Molecular Biology*, **42**, 41–50.

Sun, Y., Qiao, H., Ling, Y. *et al.* (2011) New analogues of (*E*)-β-farnesene with insecticidal activity and binding affinity to aphid odorant-binding proteins. *Journal of Agricultural and Food Chemistry*, **59**, 2456–2461.

Sun, Y.F., De Biasio, F., Qiao, H.L. *et al.* (2012a) 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., Huang, L., Pelosi, P. & Wang, C. (2012b) 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.L., Huang, L.Q., Pelosi, P. & 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. *et al.* (2000) Mammalian odorant binding proteins. *Biochimica et Biophysica Acta*, **1482**, 229–240.

Tegoni, M., Companacci, V. & Cambillau, C. (2004) Structural aspects of sexual attraction and chemical communication in insects. *Trends in Biochemical Sciences*, **29**, 257–264.

Thompson, J.D., Higgins, D.G. & 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 Research*, **22**, 4673–4680.

Tsitsanou, K.E., Thireou, T., Drakou, C.E. *et al.* (2012) *Anopheles gambiae* odorant binding protein cristal complex with the synthetic repellent DEET: implications for structure-based design of novel mosquito repellents. *Cellular and Molecular Life Sciences*, **69**, 283–297.

Tsuchihara, K., Fujikawa, K., Ishiguro, M. *et al.* (2005) An odorant-binding protein facilitates odorant transfer from air to hydrophilic surroundings in the blowfly. *Chemical Senses*, **30**, 559–564.

Vieira, F.G. & 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 Biology and Evolution*, **3**, 476–490.

Vogt, R.G. & Riddiford, L.M. (1981) Pheromone binding and inactivation by moth antennae. *Nature*, **293**, 161–163.

Vogt, R.G., Rybczynski, R. & Lerner, M.R. (1991) Molecular cloning and sequencing of general odorant-binding proteins GOBP1 and GOBP2 from the tobacco hawk moth *Manduca sexta*: comparisons with other insect OBPs and their signal peptides. *Journal of Neuroscience*, **11**, 2972–2984.

Wang, B., Guan, L., Zhong, T. *et al.* (2013a) Potential cooperations between odorant-binding proteins of the scarab beetle *Holotrichia oblita* Faldermann (Coleoptera: Scarabaeidae). *PLoS ONE*, **8**, e84795. DOI:10.1371/journal.pone.0084795.

Wang, S.Y., Gu, S.H., Han, L. *et al.* (2013b) 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 Molecular Biology*, **22**, 172–182.

Wei, Y., Brandazza, A. & Pelosi, P. (2008) Binding of polycyclic aromatic hydrocarbons to mutants of odorant-binding protein: a first step towards biosensors for environmental monitoring. *Biochimica et Biophysica Acta*, **1784**, 666–671.

Weiner, P.K. & Kollman, P.A. (1981) AMBER: assisted model building with energy refinement. A general program for modeling molecules and their interactions. *Journal of Computational Chemistry*, **2**, 287–303.

Wu, G., Robertson, D.H., Brooks, C.L. III & Vieth, M. (2003) Detailed analysis of grid-based molecular docking: a case study of CDOCKER-A CHARMm-based MD docking algorithm. *Journal of Computational Chemistry*, **24**, 1549–1562.

Xu, P.X., Zwiebel, L.J. & Smith, D.P. (2003) Identification of a distinct family of genes encoding atypical odorant-binding proteins in the malaria vector mosquito, *Anopheles gambiae*. *Insect Molecular Biology*, **12**, 549–560.

Xu, W., Xu, X., Leal, W.S. & Ames, J.B. (2011) Extrusion of the C-terminal helix in navel orangeworm moth pheromone-binding protein (AtraPBP1) controls pheromone binding. *Biochemical and Biophysical Research Communications*, **404**, 335–338.

Yin, J., Feng, H., Sun, H. *et al.* (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.

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

Yu, F., Zhang, S., Zhang, L. & Pelosi, P. (2009) Intriguing similarities between two novel odorant-binding proteins of locusts. *Biochemical and Biophysical Research Communications*, **385**, 369–374.

Yu, Y., Ma, F., Cao, Y. *et al.* (2012) Structural and functional difference of pheromone binding proteins in discriminating chemicals in the gypsy moth, *Lymantria dispar*. *International Journal of Biological Sciences*, **8**, 979–991.

Zhang, T., Mei, X., Feng, J. *et al.* (2012a) Characterization of three pheromone-binding proteins (PBPs) of *Helicoverpa armigera* (Hübner) and their binding properties. *Journal of Insect Physiology*, **58**, 941–948.

Zhang, T., Wang, W., Gu, S. *et al.* (2012b) Structure, binding characteristics and 3D model prediction of a newly identified odorant-binding protein from the cotton bollworm, *Helicoverpa armigera* (Hübner). *Journal of Integrative Agriculture*, **11**, 430–438.

Zhong, T., Yin, J., Deng, S. *et al.* (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). *Journal of Insect Physiology*, **58**, 771–781.

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

Zhou, J.-J., Huang, W., Zhang, G. *et al.* (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. *et al.* (2006) Genome and EST analyses and expression of a gene family with putative functions in insect chemoreception. *Chemical Senses*, **13**, 453–465.

Zhou, J.-J., Robertson, G., He, X. *et al.* (2009) Characterisation of *Bombyx mori* odorant-binding proteins reveals that a general

odorant-binding protein discriminates between sex pheromone components. *Journal of Molecular Biology*, **389**, 529–545.

Zhuang, X., Wang, Q., Wang, B. *et al.* (2014) Prediction of the key binding site of odorant-binding protein of *Holotrichia oblita* Faldermann (Coleoptera: Scarabaeidae). *Insect Molecular Biology*, **23**, 381–390. DOI: 10.1111/imb.12088.

Zubkov, S., Gronenborn, A.M., Byeon, I.L. & Mohanty, S. (2005) Structural consequences of the pH-induced conformational switch in *A. polyphemus* pheromone-binding protein: mechanisms of ligand release. *Journal of Molecular Biology*, **354**, 1081–1090.

Accepted 1 June 2014