A novel mechanism of ligand binding and release in the odorant binding protein 20 from the malaria mosquito Anopheles gambiae

ARTICLE *in* PROTEIN SCIENCE · JANUARY 2013

Impact Factor: 2.85 · DOI: 10.1002/pro.2179 · Source: PubMed

4 AUTHORS, INCLUDING:

Emma Joanna Murphy
University of Colorado Boulder
14 PUBLICATIONS **207** CITATIONS
SEE PROFILE

Hannah Edlin
University of Colorado
1 PUBLICATION **9** CITATIONS
SEE PROFILE

David N M Jones
University of Colorado
41 PUBLICATIONS **1,427** CITATIONS
SEE PROFILE

THE PROTEIN SOCIETY

A novel mechanism of ligand binding and release in the odorant binding protein 20 from the malaria mosquito *Anopheles gambiae*

Brian P. Ziemba,[1] Emma J. Murphy,[1] Hannah T. Edlin,[1] and David N. M. Jones[1,2]*

[1]Department of Pharmacology, University of Colorado School of Medicine, Aurora, Colorado 80045
[2]Program in Structural Biology and Biochemistry, University of Colorado School of Medicine, Aurora, Colorado 80045

Received 15 September 2012; Revised 9 October 2012; Accepted 10 October 2012
DOI: 10.1002/pro.2179
Published online 18 October 2012 proteinscience.org

Abstract: *Anopheles gambiae* mosquitoes that transmit malaria are attracted to humans by the odor molecules that emanate from skin and sweat. Odorant binding proteins (OBPs) are the first component of the olfactory apparatus to interact with odorant molecules, and so present potential targets for preventing transmission of malaria by disrupting the normal olfactory responses of the insect. AgamOBP20 is one of a limited subset of OBPs that it is preferentially expressed in female mosquitoes and its expression is regulated by blood feeding and by the day/night light cycles that correlate with blood-feeding behavior. Analysis of AgamOBP20 in solution reveals that the apo-protein exhibits significant conformational heterogeneity but the binding of odorant molecules results in a significant conformational change, which is accompanied by a reduction in the conformational flexibility present in the protein. Crystal structures of the free and bound states reveal a novel pathway for entrance and exit of odorant molecules into the central-binding pocket, and that the conformational changes associated with ligand binding are a result of rigid body domain motions in α-helices 1, 4, and 5, which act as lids to the binding pocket. These structures provide new insights into the specific residues involved in the conformational adaptation to different odorants and have important implications in the selection and development of reagents targeted at disrupting normal OBP function.

Keywords: odorant binding protein; *Anopheles gambiae*; X-ray crystallography; NMR spectroscopy

Introduction

Female mosquitoes of the *Anopheles gambiae* species are the primary vector for transmission of *Plasmodium falciparum* malaria and are attracted to humans by odors that emanate from incubated human sweat.[1–4] In insects, the detection of airborne odorants occurs primarily in the olfactory sensilla, which are highly compartmentalized units, each tuned to respond to either a single odorant or a subset of odorant molecules. Each sensillum encases the dendrites of between one and three olfactory neurons, each of which typically expresses a single

Abbreviations: 6MH: 6-methyl-5-heptene-2-one; OBP: odorant-binding protein; PBP: pheromone-binding protein.

Additional Supporting Information may be found in the online version of this article.

Brian P. Ziemba's current address is Department of Chemistry and Biochemistry, University of Colorado Boulder, Boulder CO 80309, USA.

Grant sponsor: NIH; Grant number: DC008834 (to D.N.M.J.).

*Correspondence to: David N. M. Jones, Department of Pharmacology, University of Colorado School of Medicine, 12801 East 17th Ave., MS-8303 Aurora, CO 80045.
E-mail: david.jones@ucdenver.edu

specific odorant receptor. Support cells at the base of the sensillum secrete accessory factors, including odorant binding proteins (OBPs)[5] and odorant degrading enzymes[6] into the lymph fluid that surrounds the dendrites.[7] The OBPs are one of the first protein components of the olfactory system to interact with the odorant molecules in the lymph and function to solubilize hydrophobic odorants in the lymph, protect them from degradation, and transport them to the olfactory receptors.[8,9]

While many odorant receptors can respond to direct activation by odor molecules,[10–12] the presence of OBPs can dramatically increase sensitivity and selectivity for a particular odorant. In the moth *Antheraea polyphemus*, the pheromone binding protein (PBP) ApolPBP2 is required for activation of the odorant receptors at low concentrations of pheromone.[13] While in *Drosophila melanogaster*, mutations to the pheromone binding protein LUSH can lead to direct activation of the pheromone receptors in the absence of any pheromone.[14] Structural studies of the activating mutant of LUSH reveal that it adopts the same conformation as the wild-type protein bound to pheromone suggesting that a specific structural change induced by binding of pheromone may be required for activity.[14]

A. gambiae has approximately fifty five OBPs,[15] and a number of these are expressed preferentially in female mosquitoes, and their expression correlates with the life cycle of the mosquito including blood feeding.[16–19] A recent analysis of circadian and diel (light/dark) regulated gene expression[20] has identified AgamOBP20 as one of a limited subset of OBPs (OBPs 3, 17, 20, and 47) whose expression is regulated by light/dark cycles that correlate with well established patterns of host-seeking and biting behaviors, suggesting that this subset of OBPs are likely to play prominent roles in regulating olfactory responses to human derived odors.[20] Consequently, these proteins are attractive targets to disrupt normal olfactory signaling and subsequent transmission of malaria. Indeed, recent studies have shown that DEET and other repellents may directly target OBPs.[21]

We have initiated structural studies to elucidate the ligand binding properties of OBPs implicated as potential regulators of the host-seeking behavior of the mosquito, and to understand how ligand binding affects AgamOBP20 structure. Here we present the structure of AgamOBP20 in both the free and ligand-bound forms. The structures of a number of PBPs and OBPs[22–37] have revealed a common scaffold of six alpha helices surrounding a predominantly hydrophobic central pocket. OBPs differ predominantly in the length of their C-terminal tail (long, medium, and short[38]) and AgamOBP20 is a member of the medium-length group of OPBs typified by the Dmel OBP LUSH,[14,23] *Apis mellifera* (Amel) ASP1,[30] AgamOBP1,[27] and AgamOBP4.[36] NMR spectroscopy reveals that AgamOBP20 exhibits significant conformational heterogeneity in the absence of odorants and that binding of odorant molecules results in a significant reduction in the conformational flexibility. The crystal structures of the free and bound forms of the protein reveal that this conformational heterogeneity is associated with semirigid domain motions of helices 1, 4, and 5, which function to regulate access to the central ligand-binding pocket. The structure reveals a novel entrance to the central binding pocket, and so these results have implications in the selection and design of novel compounds designed to interfere with normal OBP function.

Results

Cloning and expression of AgamOBP20

Analysis of the amino acid sequence of AgamOBP20 (AGAP005208) using the Signal IP server v4.0 failed to unambiguously identify the location of the signal peptide sequence.[39,40] As AgamOBP20 has little sequence conservation with other OBPs outside the core structured domain (Supporting Information Fig. S1) we generated an expression construct based on the closest OBPs for which structures were available,[15] namely *Drosophila* OBP LUSH[14,23] and *A. gambiae* OBP4.[36] The final construct starts at Glu23 of the full-length sequence and contains three additional amino acids from the expression vector (Met-Thr-Val) at the N-terminus. Because of the ambiguity in defining the location of the signal peptide it is possible that there are additional (or fewer residues) at the N-terminus in the native protein. The protein prepared from this construct elutes as a monomer on size exclusion chromatography.

Apo-AgamOBP20 is highly dynamic in solution

The structure of the purified apo-protein was evaluated using two-dimensional ^{1}H-^{15}N heteronuclear Single Quantum Correlation (HSQC) NMR spectroscopy, which revealed the presence of significant conformational averaging on the micro- to millisecond timescale [Fig. 1(A)]. In the spectrum of the apoprotein there are only 64 well-resolved peaks out of an expected 115, while the remaining peaks are broad or are not observable. Although the tertiary structure of the protein is not well defined under these conditions, circular dichroism spectrum recorded under the same conditions shows the protein has a well-defined all alpha-helical secondary structure (Supporting Information Fig. S2).

NMR was used to screen a panel of odorants[10] that are major components of incubated human sweat in an effort to identify the cognate ligand(s) for AgamOBP20. Binding of a number of ligands, including indole [109 peaks, Fig. 1(B)], 6-methyl-5-

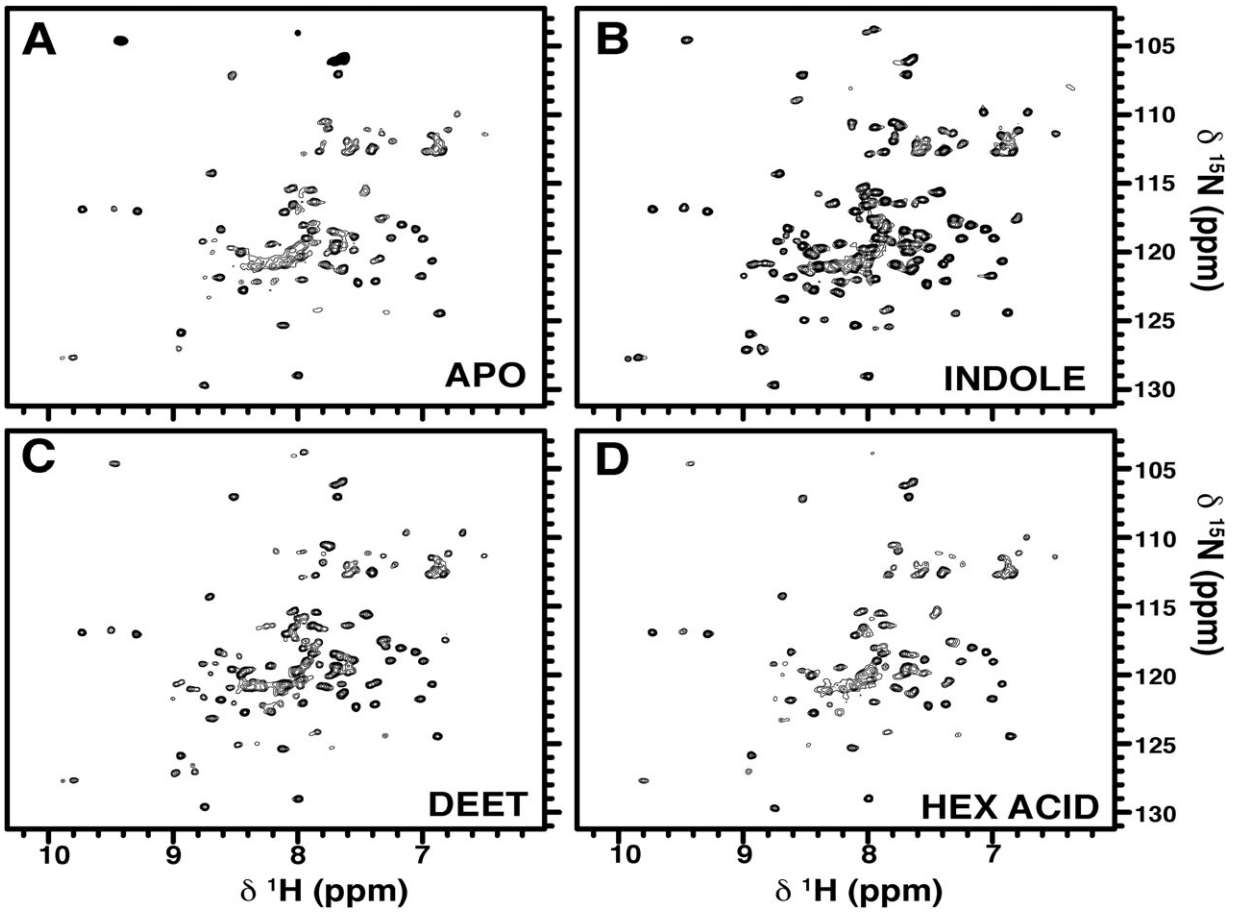

Figure 1. Binding of ligands to AgamOBP20 leads to a reduction in the conformational flexibility as judged by changes in the appearance of ^{1}H-^{15}N HSQC NMR spectra of the protein (A) apo AgamOBP20, (B) in the presence of Indole, (C) DEET, and (D) hexanoic acid (caproic acid). All spectra were collected at pH 6.9, in 20 m*M* sodium phosphate. The ligand concentration in each case was ~5 m*M*.

heptene-2-one (110 peaks, 6MH), and DEET [103 peaks, Fig. 1(C)] led to significant improvements in the appearance of the NMR spectrum, while other ligands, for example, hexanoic acid [Fig. 1(D)] failed to produce any significant changes. However, even for the best ligands there is still evidence of a degree of conformational heterogeneity, which suggest these odorants may not be optimal ligands. A number of ligands could not be screened by NMR because of limiting solubility. Subsequently, we selected those ligands judged to produce the greatest stabilization of the protein (based on the number of cross-peaks and improvements in peak widths in the NMR spectrum) for crystallographic studies.

Changes in pH have no effect on structure of AgamOBP20

It has previously been suggested that changes in pH in the vicinity of dendritic membranes trigger conformational changes in OBPs and this acts as a mechanism to regulate ligand binding and release.[22,26,27,41] AgamOBP20 does not contain any histidine residues, and therefore is unlikely to be sensitive to pH changes in the range of 5 to 7. To test this, circular dichroism spectra were collected over the pH range 7.5 to 4.5, which showed no effect on the overall secondary structure of the protein on its own (Supporting Information Fig. S2) or in the presence of the ligands 6MH or indole (data not shown). NMR spectroscopy was then used to determine if a change in the pH would prevent the conformational changes induced by the binding of ligand. NMR spectra recorded at pH 4.5 show similar changes in peak positions upon addition of indole as spectra recorded at pH 7.4 (Supporting Information Fig. S3) indicating that a shift in pH from 7.4 to 4.5 does not prevent binding of indole. Therefore, it appears unlikely that changes in pH induce conformational changes that regulate ligand binding to AgamOBP20.

Crystal structure of AgamOBP20

Initial crystallization screens of AgamOBP20-ligand complexes identified a set of conditions that gave rise to two separate crystal forms, one in the $P2_1\ 2_1\ 2_1$ space group and a second in the C1 2 1 space group. The $P2_1\ 2_1\ 2_1$ crystal form contains a single protein molecule in the asymmetric unit and the structure (PDB ID: 3V2L) was solved by molecular

Table I. *X-ray Data Collection and Refinement Statistics*

PDB ID	3V2L	4F7F	3VB1
Space group	P 2_1 2_1 2_1	C 1 2 1	P 2_1 2_1 2_1
Cell dimensions			
a, *b*, *c* (Å)	34.58, 38.23, 89.91	131.44,36.54, 96.68	34.92, 36.46, 92.09
α, β, γ (°)	90.0, 90.0, 90.0	90.00, 91.15, 90.00	90.0, 90.0, 90.0
Resolution	34.00–1.80 (1.85–1.80)	24.03–1.80(1.85–1.80)	32.65–2.00 (2.05–2.00)
Completeness (%)	97.1 (99.3)	86.3 (92.5)	98.9 (99.2)
Redundancy	4.13 (4.17)	3.1 (3.2)	4.4 (4.5)
I/sigma	7.7 (2.5)	8.5/0.5	15.1 (5.5)
R_{merge}	0.079 (0.265)	0.059 (0.257)	0.046 (0.198)
R_{work} (%)/R_{free} (%)	19.4/24.2	19.6/25.2	23.1/27.1
Ramachandran (plot) (%) (favorable, additional, generous, allowed)	94.5, 5.5, 0.0, 0.0	94.2, 5.8, 0.0, 0.0	95.4, 4.6, 0.0, 0.0
RMS deviations			
Bond lengths (Å)	0.012	0.010	0.014
Bond angles (°)	1.218	1.269	1.383

Data are from a single crystal. Values in parentheses are for highest-resolution shell.

replacement using the *Drosophila* OBP LUSH as an initial search model. The structure of the second crystal form was subsequently solved by molecular replacement (PDB ID: 4F7F) using the 3V2L structure as an initial search model. This latter crystal form contained four molecules in the asymmetric unit, which are present as two pseudo noncrystallographic dimers with very different packing interactions (described below). Ultimately, even though ligands were included in both the protein and precipitant solutions at concentrations that are demonstrated to bind and induce conformational changes in the protein, the presence of polyethylene glycol (PEG) in the crystallization solution was sufficient to displace any ligand.

Subsequently we identified a third crystal form of AgamOBP20 using ammonium sulfate as the precipitant. In this case crystals were obtained in the absence of ligand. This structure (PDB ID: 3VB1) was solved by molecular replacement using the structure of *A. gambiae* OBP4[36] as the initial search model. The final structure contained a single molecule within the asymmetric unit. Collection and refinement statistics for all crystals are listed in Table I, and coordinates and structure factors have been deposited with the Protein Data Bank.

Overall structure of AgamOBP20

All of the structures of AgamOBP20 presented here share the same overall fold in common with other classical medium-length OBPs. The C-terminal tail of AgamOBP20 inserts into the core of the protein to form one edge of the binding pocket [Fig. 2(A)], and the carboxyl group of the C-terminal proline (Pro120) makes a hydrogen bond with the hydroxyl of Tyr45 and also interacts with two arginine residues at positions 14 and 32 which indicates the carboxyl group of Pro120 is charged. Additional electrostatic interactions between Arg14 and Glu25, and Glu11 on the outside of the protein help to stabilize the conformation of this C-terminus/Arg electrostatic cluster.

Structure of the bound state of AgamOBP20

The binding pocket of AgamOBP20, is lined with predominantly hydrophobic residues: Met6, Met7, Gly10, Ile13, Met50, Met53, Ile70, Ile73, Met82, Leu86, Leu106, Leu107, Leu110, Phe117, Ile118, Phe119, and Pro120. Two polar residues, Thr55 in helix 3 and Ser66 in helix 4, are in close proximity on one edge of the cavity, and are positioned to potentially make multiple hydrogen bonds with polar groups of ligands. In the 3V2L structure, the central cavity has a total volume of ~550 Å^3 (CASTp[42]). In the 3V2L and 4F7F structures all molecules have a fragment of polyethylene glycol (PEG) bound in the central cavity (Fig. 2). In 3V2L, this PEG is bound in an elongated conformation [Fig. 2(B)], with one end exiting the cavity through an opening formed by Met6, Met7 and Gly10 in helix 1, Arg32 in the loop between helices 4 and 5, and Ile118 and Pro120 in the C-terminal tail.

In the 4F7F structure (Supporting Information Figs. S4 and S5), there are four molecules in the asymmetric unit that are very similar to each other (average pairwise RMSD = 0.59 Å) and to the bound structure 3V2L (RMSD = 0.51 Å). Differences between these molecules are relatively small, and are located predominantly in the N-terminus of helix 1, helix 4, helix 2, and the loop between helices 2 and 3 (Supporting Information Fig. S5). These latter changes can be explained by differences in crystal packing as the four molecules pack to form two noncrystallographic dimers. The two molecules labeled C and D (magenta and orange in Supporting Information Fig. S5A) form one noncrystallographic dimer that are related by a 180° rotation. The

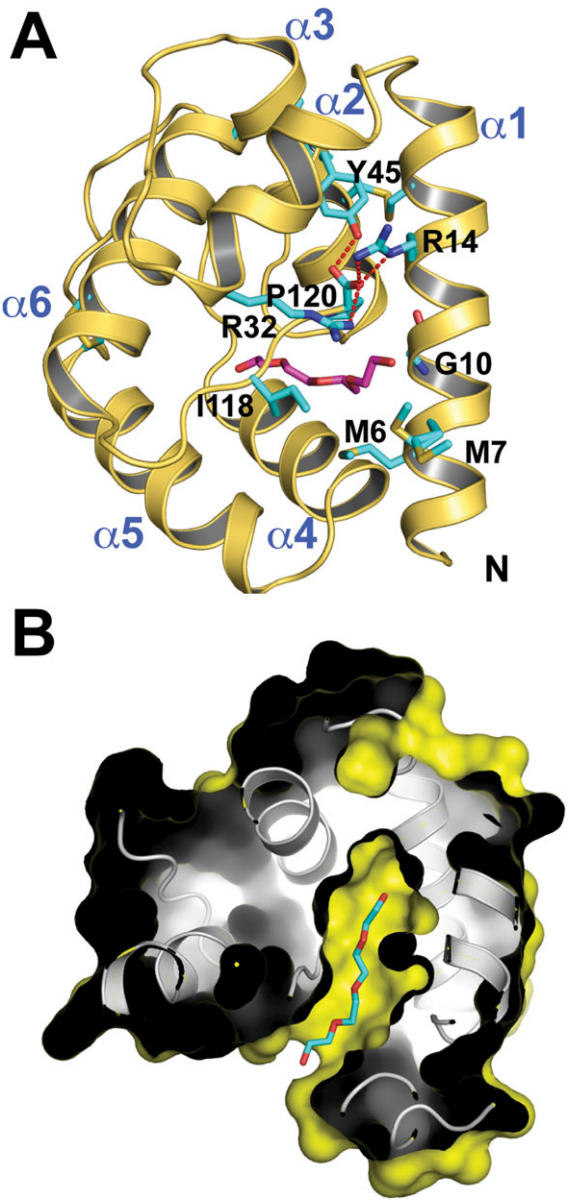

Figure 2. Crystal Structure of the PEG bound state of AgamOBP20 (A) Ribbon diagram of AgamOBP20 (PDB: 3V2L). The C-terminal Pro120 forms interactions with Arg32 and Arg14. A single PEG molecule (purple) binds in the central cavity and exits through an opening formed by helix 1, helix 4, and the C-terminal tail. (B) Cut-away view of the surface plot of AgamOBP20 showing the PEG molecule (cyan) bound in an elongated conformation.

interface between these two monomers is formed by residues from helix 1, helix 2 and in the C-terminal tail, and buries 730 Å^2 of surface area (PISA[43]). In this dimer, a single fragment of a PEG molecule is shared between the 2 monomers. In the second dimer, formed by chains A and B (cyan and green in Supporting Information Fig. S5B), the interface is formed by residues in helices 1, 3, and 4 of chain A, and helices 1 and 2, and the C-terminal tail of chain B (helices 1, 3, and 4 of chain B make a similar set of interactions with a symmetry related molecule from chain A). This interface buries 630 Å^2 of surface area. In this dimer, a PEG molecule exits from the pocket of chain B and packs at the interface between the two monomers (Supporting Information Fig. S5A), while the PEG molecule in monomer A extends into solvent. Neither of the observed dimers (or other packing interactions present with symmetry related molecules) is predicted to form stable quaternary structures in solution.[43] This is in agreement with size exclusion chromatography where the purified refolded protein elutes as a monomer (not shown).

Structure of apo-AgamOBP20 reveals movement of key helices

The 3VB1 structure solved in the absence of any ligand is significantly expanded and more open than

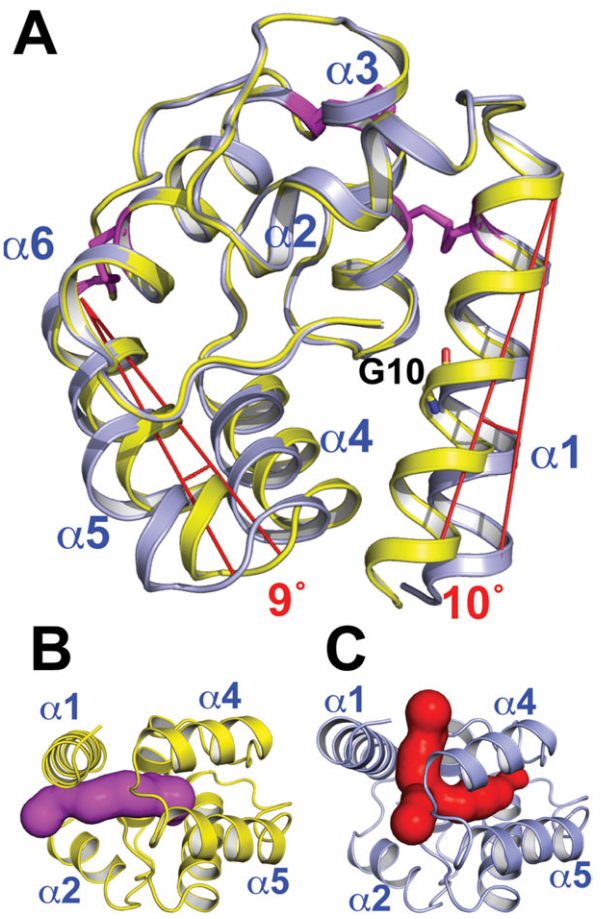

Figure 3. Comparison of apo and bound states of AgamOBP20 reveal conformational changes in helices 1, 4, and 5 (A) In the apo state (blue), helices 1, 4, and 5 are rotated out from the main body of the protein compared with the bound state (yellow) which leads to an enlargement of the central binding cavity and formation of a second opening to the cavity. (B) and (C) Space filling plots of the central cavity and entrances to the central pocket in (B) the bound state of AgamOBP20 and (C) the apo state. Cavity analysis was performed using Caver.[44]

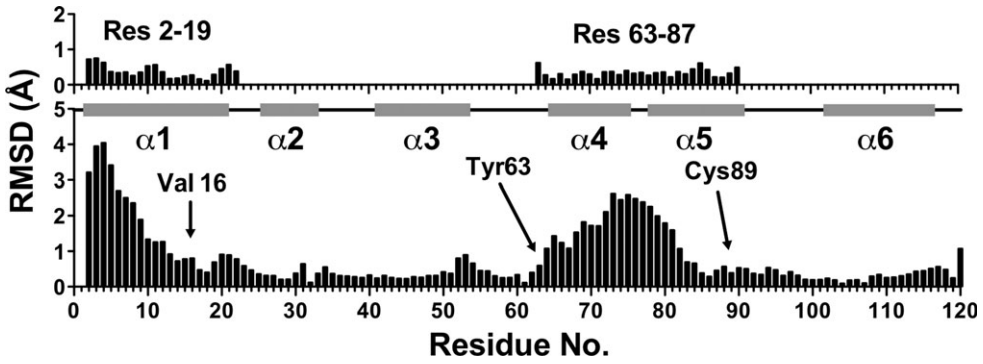

Figure 4. Analysis of per residue differences between free and bound states of AgamOBP20 reveal conformational changes are a result of rigid body motions. (Bottom) Per residue RMSD after a global superposition of the apo and bound states of AgamOBP20. (Top) Results of superposition on residues 2–19 alone, and for residues 63–87 alone. The positions of helices 1-6 are indicated on the central line by grey boxes.

the bound form of the protein [Fig. 3(A,C)]. Helix 1 is rotated out by 10° from the main body of the protein such that residues at the N-terminus of this helix are displaced by ~3.5 Å relative to their position in the bound state [Fig. 3(A)]. Simultaneously, residues 64-82 in helices 4 and 5 are rotated by ~9° in the opposite direction from the rotation of helix 1 and this results in displacements of between 2.4 and 3.2 Å of the residues in the loop between helices 4 and 5 [Fig. 3(A)]. These motions lead to a significant expansion in the volume of the central cavity (~855 Å^3 compared with ~550 Å^3), an increase in the size of the main opening to the pocket, and formation of a second opening in the face of the protein formed by helices 1, 3, and 4 [Fig. 3(C)]. There is evidence of increased conformational heterogeneity compared to the bound form, and this includes a second conformation for residues 64 to 82 that is similar to the conformation observed in the bound state, while residues 114 to 118 are also significantly less well defined than the rest of the protein.

A global superposition of the apo and bound states of AgamOBP20 provides evidence that the conformational rearrangements in helices 1, 4, and 5 are a result of semirigid domain motions. A plot of the per residue RMSD[45] illustrates the displacement of helices 1, 4, and 5 (Fig. 4). However, when these regions are considered separately they superimpose upon each other with low RMSDs (0.36 Å for residues 2–17 and 0.28 Å for residues 63–87) (Fig. 4, top). This implies that the differences between the free and bound states of helix1, and particularly of helices 4 and 5, are a result of domain reorientations that maintain the relative orientations of helices 4 and 5 to each other, rather than as a result of localized structural perturbations along the length of these helices. This data further suggests that the major structural changes arise as a result of hinge motions in the vicinity of residues Val16/Cys17 in helix 1, Tyr63 and Asp64 in helix 4, and Cys89 and Arg90 in helix 5.

Comparison of AgamOBP20 with other known structures

The structure of the bound state of AgamOBP20 is most similar to the structure of the constitutively active D118A mutant of the *Drosophila* OBP LUSH[14] (PDB ID: 2QDI, Dali *Z*-score[46,47] of 19.8 and backbone RMSD of 1.54 Å) and to the indole bound form of AgamOBP4[36] (PDB ID: 3Q8I, *Z*-score 19.7 and RMSD 1.7 Å). The LUSH D118A mutant structure is also bound to a fragment of PEG, but in contrast to OBP20, the PEG molecule exits the pocket through an opening on the front face of the protein formed by residues from helices 1, 3, and 4 (Supporting Information Fig. S6).

The structure of the apo form of OBP20 is most similar to that of the Honey bee *Apis mellifera* ASP5 (PDB ID: 3RZ2) (Dali *Z*-score 18.6, RMSD = 1.6 Å). Helices 2, 3, 4, and 6 of AgamOBP20 and ASP5 superimpose well (RMSD = 0.91 Å) while helix 5 of ASP5 is moved out from the main body of the protein compared to AgamOBP20 [Fig. 5(A)]. In contrast, helix 1 in ASP5 is found in the same orientation as helix 1 of the bound state of AgamOBP20 [Fig. 5(B)]. As a result of these changes the binding pocket of ASP5 is significantly expanded compared with that seen in AgamOBP20. This allows ASP5 to accommodate two molecules of *N*-butyl-benzene sulfonamide, a common softening agent used in the manufacture of laboratory plasticware.[48] Further analysis suggests that helices 4 and 5 of ASP5 may also move as a semi rigid body, as residues 65 to 92 can be superimposed almost exactly on the corresponding regions of AgamOBP20 (residues 62–89) both in the apo state (RMSD 0.73 Å) (Supporting Information Fig. S7A) and the PEG bound state (RMSD 0.78 A) (Supporting Information Fig. S7B).

Discussion

Our current findings reveal that binding of odorants to AgamOBP20 results in a significant reduction in

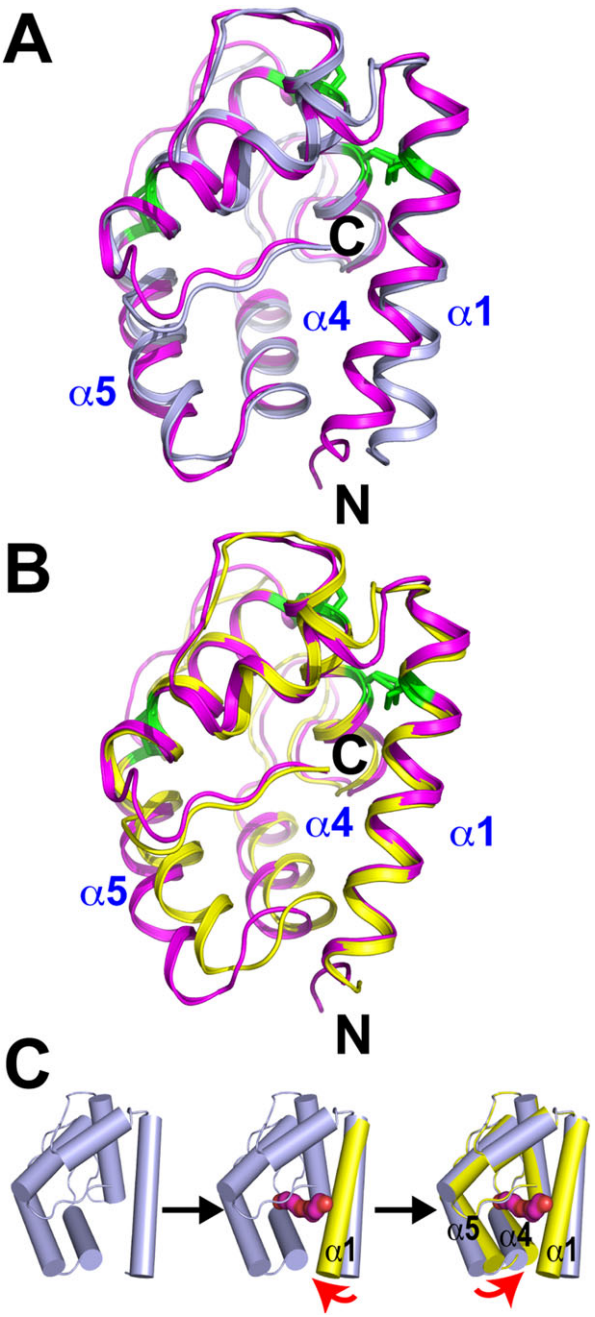

Figure 5. Comparison of AgamOBP20 structure with *Apis melifera* ASP5. (A) Superposition of apo-AgamOBP20 (blue) and ASP5 (magenta) shows that helices 4 and 5 of ASP5 adopt a similar orientation to that found in the apo state of AgamOBP20. (B) Superposition of ASP5 with the bound state of AgamOBP20 (yellow) reveals that helix 1 of ASP5 adopts the same conformation seen in the bound state of AgamOBP20. In both panels the positions of disulfides are shown in green. (C) Proposed mechanism for conformational changes induced by ligand binding to AgamOBP20. Binding of ligand leads to shift in position of helix-1 followed by closure of helices 4 and 5 over the pocket. Positions in the apo protein are shown in gray and the bound protein is shown in yellow. An interactive view is available in the electronic version of the article.

the conformational heterogeneity present in solution, and this appears to be a result of semirigid domain motions that act to regulate access of ligands to the central binding pocket. The structure of *Apis mellifera* ASP5, which shares features with both the closed and open forms of AgamOBP20, suggests that similar conformational rearrangements may also regulate ligand binding in this protein. Therefore, we hypothesize that the structures of AgamOBP20 and ASP5 represent snapshots of the conformational transitions that occur to helices 1, 4, and 5 upon ligand binding. While overall, ASP5 has the most open and accessible binding pocket, helix 1 in this structure adopts the same conformation seen in the more compact ligand-bound state of AgamOBP20. Therefore, we hypothesize that entry of ligand into the pocket is followed first by closing of helix 1 over the pocket, followed by a shift of helices 4 and 5, which act as a clamp at the base of the pocket to hold the ligand in place [Fig 5(C)]. Alternatively, this may simply represent the structural plasticity of the region encompassing helices 4 and 5 to adapt to ligands of different size.

In both AgamOBP20 and ASP5 the main entrance to the ligand binding pocket is located between the C-terminal tail and the N-terminal region of helix 1, with a second smaller opening centered in the face formed by helix 1 and helices 3 and 4. This contrasts with other classical medium-length OBPs, that include AgamOBP1,[27] AgamOBP4,[36] *Aedes egyptii* OBP1,[26] *Culex quinquefasciatus* OBP1,[25] and *Apis mellifera* ASP1[48] where the primary entrance to the pocket is in the face of the protein formed by helices 1, 3, and 4, and in some cases an additional entrance is located between helices 4 and 5 near the N-terminus of helix 4. Alternatively, other studies have suggested that the loop between helices 3 and 4 forms a lid to the binding pocket.[26] In some cases there is evidence of conformational rearrangements of the C-terminal tail in response to ligand binding (e.g. in ASP1[48]), but here the tail does not appear to control access to the binding pocket. In AgamOBP20 the exit from the central cavity between helix 1 and the C-terminal tail is made possible by the presence of a glycine at residue 10 [Fig. 3(A)], while in ASP5 this residue is an alanine. In contrast, in both LUSH and AgamOBP1 the structurally homologous residue is a leucine, which completely blocks this exit from the cavity.

Conformational heterogeneity in helices 4 and 5 has also been observed in *A. gambiae* OBP7,[34] a novel type of OBP that contains 8 cysteines and an additional C-terminal helix that forms a wall to the binding pocket and so is not as directly comparable to AgamOBP20 as ASP5. In a series of structures of AgamOBP7-ligand complexes,[34] many residues in positions homologous to helices 4 and 5 are disordered. However, even though these two helices

exhibit the greatest conformational heterogeneity, in most of these structures of AgamOBP7 the primary entrance to the binding pocket is again in the face of the protein formed by helices 1, 3, and 4. However, in one structure of AgamOBP7 (3R1V) where helices 4 and 5 are defined by the electron density, a second opening is observed between helices 1, 5, and the C-terminal helix, in a similar position to that seen with AgamOBP20. This may reflect the fact that OBPs exhibit significant promiscuity in their ligand binding capabilities, and that different odorants and pheromones are capable of inducing different conformational changes within the same OBP.[14,32,34,38,49,50]

It is widely accepted that OBPs play important roles in the transport and delivery of odorants to olfactory receptors. In addition there is increasing evidence that, in some cases, OBPs play a more direct role in the activation of these receptors.[13,14,51,52] It is possible that in these latter cases, the activity of OBPs may be regulated by the presence of extensive conformational flexibility present in the apo-protein. While a certain degree of conformational flexibility must be required to allow ligands to access the central binding pocket in almost all OBPs, a number of OBPs particularly LUSH.[14,23,53] AgamOBP4[36] and Amel ASP2,[32] exhibit significantly higher levels of conformational heterogeneity. In the case of LUSH, mutations in the protein that mimic the conformational changes induced by binding of pheromones can activate olfactory receptors in the absence of pheromones.[14] While binding of indole to AgamOBP4, induces an even more dramatic shift in the conformational equilibrium that is required to stabilize the binding site for AgamOBP1.[36] In *Apis mellifera* ASP2 binding of an opportunistic ligand to also led to a significant reduction in the conformational heterogeneity.[32] However, this was not to the same extent as seen with LUSH[23,53] and AgamOBP4.[36] It is likely that this heterogeneity would be further reduced in the presence of a cognate ligand. High levels of conformational flexibility may be important to tightly control responses to specific ligands such as pheromones, but less important in responses to more general food odors.

Ligand dependent structural transitions have been well documented in the case of moth PBPs.[22,54–59] However, these PBPs are fundamentally different from the medium length OBPs, containing a significant longer C-terminal tail, which forms an additional helix that occupies the central binding pocket in the absence of pheromone. It has been proposed that a change in pH in the vicinity of dendritic membranes induces a conformational change that can displace bound pheromone, releasing it to activate the pheromone receptors. AgamOBP20 contains no histidine residues and changes in pH over the range of 5 to 7.5 have no discernible effect on the structure in solution, and do not appear to inhibit the conformational ordering induced by the binding of tested odorants.

At present little is known about what odorants are recognized by AgamOBP20 *in vivo*. Multiple studies have established that its expression strongly correlates with blood seeking behaviors and changes in response to blood feeding.[16,18–20] Therefore, it is potentially a key modulator of mosquito-human interactions and so presents an attractive target to disrupt normal olfactory responses to human-derived odors. Knowledge of how odorant binding proteins adapt to their ligands is important for defining the molecular features that can be targeted as part of rationale discovery efforts. The structures presented here established that there is significant conformational flexibility in the helix 4/5 regions. In contrast, helices 2, 3, and 6 and most of the C-terminal tail which contribute a large number of residues to the binding pocket are well ordered in both the apo-state and the PEG-bound form of the protein. In the six separate protein molecules in the three crystal structures we have presented here, the only significant differences that we observe are in the positions of the side chains of Met53, Thr55, and Ser66. This suggests that there is a well-defined interaction surface within the pocket that can aid in the discovery of specific ligands. Defining the nature of the binding sites, and understanding the conformational transitions involved in ligand binding and release are key steps in this discovery process.

Materials and Methods

Protein expression and purification

A. gambiae AgamOBP20 was subcloned from an *A. gambiae* antennal cDNA library (provided by Dr. Larry Zweibel) into the *NdeI* and *BamH1* restriction of the pET-13a vector.[60] Protein was expressed in *Escherichia coli* BL21(DE3) cells, isolated and refolded from the insoluble fraction as previously described.[23,61] The resulting protein was purified by anion exchange chromatography using DEAE resin (BioRad), followed by cation exchange chromatography (Resource S, GE Lifesciences) and size exclusion chromatography (Superdex S-75). The folded state of the protein was confirmed using circular dichroism spectroscopy. For NMR experiments, AgamOBP20 was expressed in M9 minimal media supplemented with 2g/L D-glucose and 1 g/L ^{15}N NH_4Cl (>98%) (Sigma-Aldrich).

Circular dichroism

Experiments were performed on a Jasco J815 spectrapolarimeter in the CU School of Medicine Biophysical Core facility. The protein concentration was 5 μ*M* and ligand concentrations used were 1 m*M*, unless otherwise listed.

NMR spectroscopy

All experiments were performed at 25°C on a Varian INOVA 600 MHz NMR spectrometer equipped with a cryoprobe. Experiments used 150 to 200 μ*M* protein in 20 m*M* sodium phosphate, pH 7.4. Ligands were added directly to protein samples to a nominal concentration of 5 m*M* and incubated for at least 2 hr. Changes in sample pH were accomplished by titrating with 0.5*M* HCl and monitored with a micro sample pH probe (Sigma-Aldrich) at 25°C.

Crystallization

Crystallization conditions for AgamOBP20-ligand complexes were screened using the sitting drop vapor diffusion method in 96-well plates using custom designed in-house screens and yielded diffraction quality crystals using 29% PEG 3350, 0.1*M* sodium citrate at pH 5.5. In all cases, ligands (indole, hexanoic acid, 6-methyl-5-hepten-2-one, butanol, and 4-methyl phenol dissolved in buffer) were added to the protein and well solutions to a final concentration of 5 m*M*. For the apo-protein initial crystals were found in the Hampton's Crystal Screen 2 buffer #14 (Hampton Research Inc.), and after optimization, diffraction quality crystals were obtained using a protein solution of 8 mg mL^{-1} mixed in a 1:1 ratio with well solution containing 0.2*M* potassium sodium tartrate, 0.1*M* sodium citrate, pH 5.57, and 1.9*M* ammonium sulfate and grown at 4°C.

X-ray data collection and analysis

X-ray diffraction data was collected at the Molecular Biology Consortium Beamline 4.2.2 at the Advanced Light Source, Berkeley California, and processed used D*trek.[62] The 3V2L structure was solved by molecular replacement using PHASER[63] implemented in CCP4 and used the structure of the LUSH-butanol complex (PDB ID 1OOH) as an initial search model and refined using Refmac Version 5.0.0,[64] with manual rebuilding using COOT.[65] The3BVI structure was solved by molecular replacement in the same way, but used the structure of AgamOBP4 (PDB ID 3Q8I) as an initial search model. The 4F7F structure was solved using PHASER[63] implemented within Phenix,[66] using the 3V2L structure as an initial search model, and was refined using Phenix and rebuilt using COOT. All figures were prepared usingPymol.[67]

Accession numbers

Structure of apo AgamOBP20 (PDB ID: 3BV1), structure of AgamOBP20 bound to PEG in P $2_12_12_1$ space group (PDB ID: 3V2L), structure of AgamOBP20-PEG complex in C1 2 1 space group (PDB ID: 4F7F) have been deposited in the RCSB PDB (available at: http://www.rcsb.org/pdb).

Acknowledgments

The authors thank Dr. Jay Nix for assistance with beamline X-ray data collection and Dr. Larry Zweibel (Vanderbilt University) for mosquito cDNA libraries.

References

1. Braks MAH, Takken W (1999) Incubated human sweat but not fresh sweat attracts the malaria mosquito *Anopheles gambiae* sensu stricto. J Chem Ecol 25: 663–672.
2. Braks MA, Anderson RA, Knols BG (1999) Infochemicals in mosquito host selection: human skin microflora and Plasmodium parasites. Parasitol Today 15: 409–413.
3. Cork A, Park KC (1996) Identification of electrophysiologically-active compounds for the malaria mosquito, *Anopheles gambiae*, in human sweat extracts. Med Vet Entomol 10:269–276.
4. Meijerink J, Braks MAH, Brack AA, Adam W, Dekker T, Posthumus MA, Van Loon JJA (2000) Identification of olfactory stimulants for *Anopheles gambiae* from human sweat samples. J Chem Ecol 26:1367–1382.
5. Vogt RG, Riddiford LM (1981) Pheromone binding and inactivation by moth antennae. Nature 293:161–163.
6. Vogt RG, Riddiford LM, Prestwich GD (1985) Kinetic properties of a sex pheromone-degrading enzyme: the sensillar esterase of *Antheraea polyphemus*. Proc Natl Acad Sci USA 82:8827–8831.
7. Vogt RG, Biochemical diversity of odor detection: OBPs, ODEs and SNMPs. In: Blomquist GJ, Vogt RG, Ed. (2003) Insect pheromone biochemistry and molecular biology. London: Elsevier Academic Press, pp391–446.
8. Ronderos DS, Smith DP (2009) Diverse signaling mechanisms mediate volatile odorant detection in Drosophila. Fly 3:290–297.
9. Pelosi P, Zhou JJ, Ban LP, Calvello M (2006) Soluble proteins in insect chemical communication. Cell Mol Life Sci 63:1658–1676.
10. Carey AF, Wang G, Su C-Y, Zwiebel LJ, Carlson JR (2010) Odorant reception in the malaria mosquito *Anopheles gambiae*. Nature 464:66–71.
11. Kreher SA, Kwon JY, Carlson JR (2005) The molecular basis of odor coding in the Drosophila larva. Neuron 46:445–456.
12. Hallem EA, Ho MG, Carlson JR (2004) The molecular basis of odor coding in the *Drosophila* antenna. Cell 117:965–979.
13. Forstner M, Breer H, Krieger J (2009) A receptor and binding protein interplay in the detection of a distinct pheromone component in the silkmoth *Antheraea polyphemus*. Int J Biol Sci 5:745–757.
14. Laughlin JD, Ha TS, Jones DN, Smith DP (2008) Activation of pheromone-sensitive neurons is mediated by conformational activation of pheromone-binding protein. Cell 133:1255–1265.
15. Xu PX, Zwiebel LJ, Smith DP (2003) Identification of a distinct family of genes encoding atypical odorant-binding proteins in the malaria vector mosquito, *Anopheles gambiae*. Insect Mol Biol 12:549–560.
16. Biessmann H, Nguyen QK, Le D, Walter MF (2005) Microarray-based survey of a subset of putative olfactory genes in the mosquito *Anopheles gambiae*. Insect Mol Biol 14:575–689.
17. Justice RW, Dimitratos S, Walter MF, Woods DF, Biessmann H (2003) Sexual dimorphic expression of putative

antennal carrier protein genes in the malaria vector *Anopheles gambiae*. Insect Mol Biol 12:581–594.
18. Marinotti O, Nguyen QK, Calvo E, James AA, Ribeiro JM (2005) Microarray analysis of genes showing variable expression following a blood meal in *Anopheles gambiae*. Insect Mol Biol 14:365–373.
19. Marinotti O, Calvo E, Nguyen QK, Dissanayake S, Ribeiro JMC, James AA (2006) Genome-wide analysis of gene expression in adult *Anopheles gambiae*. Insect Mol Biol 15:1–12.
20. Rund SSC, Hou TY, Ward SM, Collins FH, Duffield GE (2011) Genome-wide profiling of diel and circadian gene expression in the malaria vector *Anopheles gambiae*. Proc Natl Acad Sci USA 108:E421–E430.
21. Tsitsanou KE, Thireou T, Drakou CE, Koussis K, Keramioti MV, Leonidas DD, Eliopoulos E, Iatrou K, Zographos SE (2012) Anopheles gambiae odorant binding protein crystal complex with the synthetic repellent DEET: implications for structure-based design of novel mosquito repellents. Cell Mol Life Sci 69:283–297.
22. Horst R, Damberger F, Luginbuhl P, Guntert P, Peng G, Nikonova L, Leal WS, Wuthrich K (2001) NMR structure reveals intramolecular regulation mechanism for pheromone binding and release. Proc Natl Acad Sci USA 98:14374–14379.
23. Kruse SW, Zhao R, Smith DP, Jones DN (2003) Structure of a specific alcohol-binding site defined by the odorant binding protein LUSH from *Drosophila melanogaster*. Nat Struct Biol 10:694–700.
24. Sandler BH, Nikonova L, Leal WS, Clardy J (2000) Sexual attraction in the silkworm moth: structure of the pheromone-binding-protein-bombykol complex. Chem Biol 7:143–151.
25. Mao Y, Xu X, Xu W, Ishida Y, Leal WS, Ames JB, Clardy J (2010) Crystal and solution structures of an odorant-binding protein from the southern house mosquito complexed with an oviposition pheromone. Proc Natl Acad Sci USA 107:19102–19107.
26. Leite NR, Krogh R, Xu W, Ishida Y, Iulek J, Leal WS, Oliva G (2009) Structure of an odorant-binding protein from the mosquito *Aedes aegypti* suggests a binding pocket covered by a pH-sensitive "Lid". PLoS One 4: e8006.
27. Wogulis M, Morgan T, Ishida Y, Leal WS, Wilson DK (2006) The crystal structure of an odorant binding protein from *Anopheles gambiae*: evidence for a common ligand release mechanism. Biochem Biophys Res Commun 339:157–164.
28. Pesenti ME, Spinelli S, Bezirard V, Briand L, Pernollet JC, Campanacci V, Tegoni M, Cambillau C (2009) Queen bee pheromone binding protein pH-induced domain swapping favors pheromone release. J Mol Biol 390:981–990.
29. Pesenti ME, Spinelli S, Bezirard V, Briand L, Pernollet JC, Tegoni M, Cambillau C (2008) Structural basis of the honey bee PBP pheromone and pH-induced conformational change. J Mol Biol 380:158–169.
30. Lartigue A, Gruez A, Briand L, Blon F, Bezirard V, Walsh M, Pernollet JC, Tegoni M, Cambillau C (2004) Sulfur single-wavelength anomalous diffraction crystal structure of a pheromone-binding protein from the honeybee *Apis mellifera* L. J Biol Chem 279:4459–4464.
31. Lartigue A, Gruez A, Spinelli S, Riviere S, Brossut R, Tegoni M, Cambillau C (2003) The crystal structure of a cockroach pheromone-binding protein suggests a new ligand binding and release mechanism. J Biol Chem 278:30213–30218.
32. Lescop E, Briand L, Pernollet J-C, Guittet E (2009) Structural basis of the broad specificity of a general odorant-binding protein from honeybee. Biochemistry 48:2431–2441.
33. Spinelli S, Lagarde A, Iovinella I, Legrand P, Tegoni M, Pelosi P, Cambillau C (2012) Crystal structure of Apis mellifera OBP14, a C-minus odorant-binding protein, and its complexes with odorant molecules. Insect Biochem Mol Biol 42:41–50.
34. Lagarde A, Spinelli S, Tegoni M, He X, Field L, Zhou JJ, Cambillau C (2011) The crystal structure of odorant binding protein 7 from *Anopheles gambiae* exhibits an outstanding adaptability of its binding site. J Mol Biol 414:401–412.
35. Lagarde A, Spinelli S, Qiao H, Tegoni M, Pelosi P, Cambillau C (2011) Crystal structure of a novel type of odorant-binding protein from *Anopheles gambiae*, belonging to the C-plus class. Biochem J 437:423–430.
36. Davrazou F, Dong E, Murphy EJ, Johnson HT, Jones DN (2011) New insights into the mechanism of odorant detection by the malaria-transmitting mosquito *Anopheles gambiae*. J Biol Chem 286:34175–34183.
37. Zhou JJ, Robertson G, He X, Dufour S, Hooper AM, Pickett JA, et al. (2009) Characterisation of *Bombyx mori* odorant-binding proteins reveals that a general odorant-binding protein discriminates between sex pheromone components. J Mol Biol 389:529–545.
38. Tegoni M, Campanacci V, Cambillau C (2004) Structural aspects of sexual attraction and chemical communication in insects. Trends Biochem Sci 29:257–264.
39. Emanuelsson O, Brunak S, von Heijne G, Nielsen H (2007) Locating proteins in the cell using TargetP, SignalP and related tools. Nat Protoc 2:953–971.
40. Petersen TN, Brunak S, von Heijne G, Nielsen H (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods 8:785–786.
41. Wojtasek H Leal W (1999) Conformational change in the pheromone-binding protein from *Bombyx mori* induced by pH and by interaction with membranes. J Biol Chem 274:30950–30956.
42. Dundas J, Ouyang Z, Tseng J, Binkowski A, Turpaz Y, Liang J (2006) CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. Nucleic Acids Res 34:W116–118.
43. Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state. J Mol Biol 372: 774–797.
44. Petrek M, Otyepka M, Banas P, Kosinova P, Koca J, Damborsky J (2006) CAVER: a new tool to explore routes from protein clefts, pockets and cavities. BMC Bioinformatics 7:316–324.
45. Diamond R (1992) On the multiple simultaneous superposition of molecular structures by rigid body transformations. Protein Sci 1:1279–1287.
46. Holm L, Rosenstrom P (2010),Dali server: conservation mapping in 3D. Nucleic Acids Res 38:W545–W549.
47. Holm L, Sander C (1995) Dali: a network tool for protein structure comparison. Trends Biochem Sci 20: 478–480.
48. Pesenti ME, Spinelli S, Bezirard V, Briand Lc, Pernollet J-C, Tegoni M, Cambillau C (2008) Structural basis of the honey bee PBP pheromone and pH-induced conformational change. J Mol Biol 380:158–169.
49. Mohl C, Breer H, Krieger J (2002) Species-specific pheromonal compounds induce distinct conformational changes of pheromone binding protein subtypes from *Antheraea polyphemus*. Invert Neurosci 4:165–174.
50. Campanacci V, Lartigue A, Hallberg BM, Jones TA, Giudici-Orticoni MT, Tegoni M, Cambillau C (2003) Moth chernosensory protein exhibits drastic

conformational changes and cooperativity on ligand binding. Proc Natl Acad Sci USA 100:5069–5074.
51. Swarup S, Williams TI, Anholt RRH (2011) Functional dissection of odorant binding protein genes in *Drosophila melanogaster*. Genes Brain Behavior 10:648–657.
52. Harada E, Nakagawa J, Asano T, Taoka M, Sorimachi H, Ito Y, Aigaki T, Matsuo T (2012) Functional evolution of duplicated odorant-binding protein genes, Obp57d and Obp57e, in Drosophila. PLoS One 7: e29710.
53. Bucci BK, Kruse SW, Thode AB, Alvarado SM, Jones DN (2006) Effect of n-alcohols on the structure and stability of the *Drosophila* odorant binding protein LUSH. Biochemistry 45:1693–1701.
54. Damberger F, Nikonova L, Horst R, Peng G, Leal WS, Wüthrich K (2000) NMR characterization of a pH-dependent equilibrium between two folded solution conformations of the pheromone-binding protein from *Bombyx mori*. Protein Sci 9:1038–1041.
55. Damberger FF, Ishida Y, Leal WS, Wüthrich K (2007), Structural basis of ligand binding and release in insect pheromone-binding proteins: NMR structure of *Antheraea polyphemus* PBP1 at pH 4.5. J Mol Biol 373: 811–819.
56. Lee D, Damberger FF, Peng GH, Horst R, Guntert P, Nikonova L, Leal WS, Wuthrich K (2002) NMR structure of the unliganded *Bombyx mori* pheromone-binding protein at physiological pH. FEBS Lett 531: 314–318.
57. Mohanty S, Zubkov S, Gronenborn AM (2004) The solution NMR structure of *Antheraea polyphemus* PBP provides new insight into pheromone recognition by pheromone-binding proteins. J Mol Biol 337:443–451.
58. Zubkov S, Gronenborn AM, Byeon IJ, Mohanty S (2005) Structural consequences of the pH-induced conformational switch in *A. polyphemus* pheromone-binding protein: mechanisms of ligand release. J Mol Biol 354:1081–1090.
59. Katre UV, Mazumder S, Prusti RK, Mohanty S (2009) Ligand binding turns moth pheromone-binding protein into a pH sensor: effect on the *Antheraea polyphemus* PBP1 conformation. J Biol Chem 284:32167–32177.
60. Studier FW, Rosenburg AH, Dunn JJ, Dubendorff JW (1990) Use of T7 RNA polymerase to direct expression of cloned genes. Methods Enzymol 185:61–89.
61. Thode AB, Kruse SW, Nix JC, Jones DN (2008) The role of multiple hydrogen-bonding groups in specific alcohol binding sites in proteins: insights from structural studies of LUSH. J Mol Biol 376:1360–1376.
62. Pflugrath JW (1999) The finer things in X-ray diffraction data collection. Acta Crystallogr D Biol Crystallogr 55:1718–1725.
63. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ (2007) Phaser crystallographic software. J Appl Cryst 40:658–674.
64. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. Acta Cryst D53:240–255.
65. Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr 60:2126–2132.
66. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr 66:213–221.
67. DeLano WL (2002) The PyMol molecular graphics system. San Carlos, CA:DeLano Scientific.