



# The Crystal Structure of Odorant Binding Protein 7 from *Anopheles gambiae* Exhibits an Outstanding Adaptability of Its Binding Site

Amandine Lagarde<sup>1</sup>, Silvia Spinelli<sup>1</sup>, Mariella Tegoni<sup>1</sup>, Xiaoli He<sup>2</sup>, Lin Field<sup>2</sup>, Jing-Jiang Zhou<sup>2</sup> and Christian Cambillau<sup>1\*</sup>

<sup>1</sup>Architecture et Fonction des Macromolécules Biologiques, UMR 6098 CNRS and Universités of Marseille, 13288 Marseille cedex 9, France

<sup>2</sup>Biological Chemistry Department, Rothamsted Research, Harpenden, Herts AL5 2JQ, UK

Received 24 June 2011;  
received in revised form  
29 September 2011;  
accepted 6 October 2011  
Available online  
12 October 2011

Edited by R. Huber

## Keywords:

odorant binding protein;  
crystal structure;  
*Anopheles gambiae*;  
olfaction;  
crystal structure

*Anopheles gambiae* (Agam) targets human and animals by using its olfactory system, leading to the spread of *Plasmodium falciparum*, the malaria vector. Odorant binding proteins (OBPs) participate to the first event in odorant recognition and constitute an interesting target for insect control. OBPs interact with olfactory receptors to which they deliver the odorant molecule. We have undertaken a large-scale study of proteins belonging to the olfactory system of Agam with in mind of designing strong olfactory repellants. Here, we report the expression, three-dimensional structures and binding properties of AgamOBP07, a member of a new structural class of OBPs, characterized by the occurrence of eight cysteines. We showed that AgamOBP07 possesses seven  $\alpha$ -helices and four disulfide bridges, instead of six  $\alpha$ -helices and three disulfide bridges in classical OBPs. The extra seventh helix is located at the surface of the protein, locked by the fourth disulfide bridge, and forms a wall of the internal cavity. The binding site of the protein is mainly hydrophobic, elongated and open and is able to accommodate elongated ligands, linear or polycyclic, as suggested also by binding experiments. An elongated electron density was observed in the internal cavity of the purified protein, belonging to a serendipitous ligand. The structure of AgamOBP07 in complex with an azo-bicyclic model compound reveals that a large conformational change in the protein has reshaped its binding site, provoking helix 4 unfolding and doubling of the cavity volume.

© 2011 Elsevier Ltd. All rights reserved.

\*Corresponding author. AFMB, 163 Avenue de Luminy  
Case 932, 13288 Marseille Cedex 09, France. E-mail address:  
[cambillau@afmb.univ-mrs.fr](mailto:cambillau@afmb.univ-mrs.fr).

Abbreviations used: PBP, pheromone binding protein; OBP, odorant binding protein; GOBP, general odorant binding protein; ESRF, European Synchrotron Radiation Facility; OR, olfactory receptor; 1-NPN, *N*-phenyl-1-naphthylamine; MALS, multi-angle static light scattering; RI, refractometry; SEC, size-exclusion chromatography; PA, palmitic acid; AZO, 4-hydroxy-4'-isopropyl-azobenzene.

## Introduction

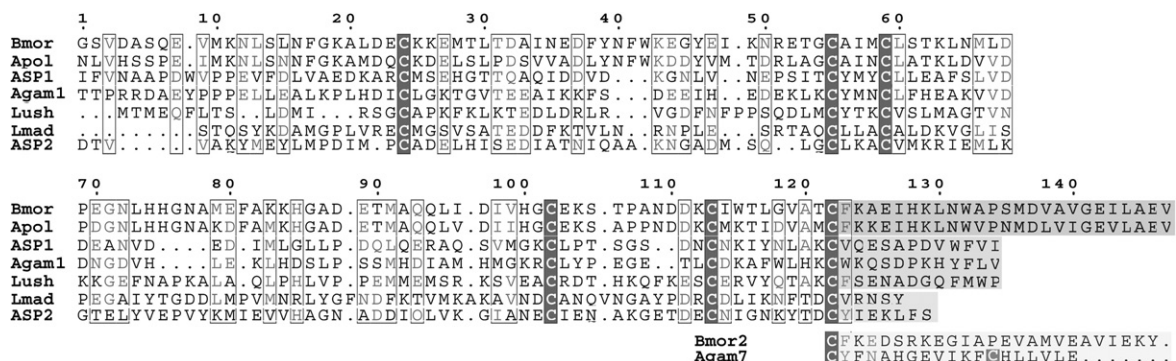
Olfaction in insects relies on the olfactory neurons located at the base of a large number of sensilla in antennae. All the components necessary for translating a chemical signal into an electric stimulus are present in a given sensillum. The axon of an olfactory neuron harbors olfactory receptors (ORs) that are bathed in the sensillar lymph<sup>1</sup> containing a large concentration of ~13- to 17-kDa proteins, the odorant binding proteins (OBPs).<sup>2,3</sup> ORs can be

stimulated by odorant molecules, directly or indirectly, *via* the OBPs.<sup>4–6</sup> OBPs that bind and convey the signals of pheromones are called pheromone binding proteins (PBPs), while those that bind and convey the signals of odorants molecules are called general odorant binding proteins (GOBPs).<sup>7</sup> It has been shown that *Drosophila melanogaster* (Dmel) PBP (named Lush) changes conformation upon pheromone binding, a trigger for OR activation.<sup>8</sup> In contrast, *Antheraea polyphemus* (Apol, a lepidopteran) ORs can directly bind pheromones, but with less specificity than when they are presented by a specific PBP.<sup>9</sup> Furthermore, a 1000-fold higher concentration of pheromone is needed to achieve the same effect on the OR when the specific PBP is not present.<sup>9</sup>

Several PBPs or GOBPs have been structurally characterized.<sup>10–16</sup> The remarkable similarities among many of them, such as a six- $\alpha$ -helix core, an internal cavity and three disulfide bridges, have made it possible to define a “classical OBPs” group. However, an in-depth analysis of their structures revealed differences at their C-termini that have functional implications<sup>17</sup> (Fig. 1). PBPs from Lepidoptera, such as those of *Bombyx mori* (Bmor)<sup>10</sup> or Apol,<sup>20</sup> are the longest classical OBPs. They possess a C-terminal segment of 15 residues with an  $\alpha$ -helical fold, buried inside the binding site at acidic pH but released outside at neutral pH.<sup>10,11,20</sup> OBPs from Dmel,<sup>12</sup> *Apis mellifera* (Amel)<sup>13</sup> or *Anopheles gambiae* (Agam)<sup>14</sup> are shorter and possess an elongated C-terminal segment buried inside the protein core, forming a wall of the internal cavity. The last residue is tightly bound by two hydrogen bonds of its carboxylic moiety to two residues of the protein core.<sup>14</sup> In Amel PBP, a mechanism of domain swapping, depending upon ligand and pH, has

been described.<sup>21,22</sup> A third subclass is even shorter since the C-terminus stops at the opening of the cavity, as in the PBP of the cockroach *Leucophaea maderae*<sup>15</sup> or in ASP2 from Amel.<sup>16</sup> We have named these three kinds of classical OBPs long, medium and short C-terminal subclasses.<sup>17</sup> Recently, another classical OBP, BmorGOBP2, has been described with a different C-terminal arrangement (Fig. 1) in which a last and seventh helix is present and located outside the protein core.<sup>23</sup>

The genomes of several insects are now available. Their analysis revealed the occurrence of a large number of olfactory proteins, ORs, OBPs and chemosensory proteins.<sup>24</sup> Besides the classical OBPs class, large classes of nonclassical OBPs have been identified. We have shown that AmelOBP14, a member of the minus-C class<sup>25</sup> with only four cysteines, has indeed only two disulfide bridges but shares a fold close to that of classical OBPs.<sup>26</sup> In AmelOBP14, the C-terminus folds as a seventh helix and forms a wall between the exterior and the internal cavity. High affinity and a certain specificity have been observed toward odorant molecules, in particular with eugenol (the main component of the odor of cloves). The third disulfide bridge of classical OBPs was re-introduced in AmelOBP14 by engineering an appropriate double cysteine mutant.<sup>26</sup> Another group of OBPs possesses a longer chain and 12 or more cysteines, the plus-C OBPs class.<sup>27</sup> We have solved the crystal structures of a member of this class, AgamOBP47.<sup>28</sup> We show that AgamOBP47 has six disulfide bridges and exhibits a new fold consisting of a classical OBP core, which is decorated with large loops and N- and C-terminal extensions forming a new open binding site on either face of the OBP. Finally, another large class encompasses OBPs with a double domain of



**Fig. 1.** Sequence alignment of several OBPs illustrating the importance of their C-terminus. Sequences have been aligned with Multalin (<http://multalin.toulouse.inra.fr/multalin/>)<sup>18</sup> and ESPrpt.<sup>19</sup> The Cys residues are depicted in white with a gray background. Partially conserved residues are boxed. The disulfide bridges are numbered 1–3 (classical OBPs). The C-termini, after the last Cys residue, are drawn on a gray background and allowed to distinguish three subclasses of classical OBPs. From top to bottom, the long C-termini of classical lepidopteran OBPs (which fold inside the binding site at low pH), the medium-length C-termini (which are elongated and form a wall of the binding site) and the short C-termini (stop at the binding site entrance). Bmor OBP2 and AgamOBP07 C-terminal sequences are provided for comparison with the three classes of classical OBPs mentioned above.

**Table 1.** Dissociation constants obtained for AgamOBP07 with 1-NPN and some ligands

	$K_{d1}$ ( $\mu\text{M}$ )	$K_{d2}$ ( $\mu\text{M}$ )	% fluorescence decrease
1-NPN	$5.4 \pm 0.4$	—	—
AZO	$0.12 \pm 0.02$	$1.9 \pm 0.2$	95
<i>p</i> - <i>t</i> -Butyl-phenylbenzoate	$0.08 \pm 0.01$	$3.9 \pm 0.2$	74
<i>n</i> -Butylcinnamate	$0.07 \pm 0.01$	$2.9 \pm 0.5$	72
<i>p</i> - <i>t</i> -Butylbenzophenone	$0.06 \pm 0.03$	$2.7 \pm 0.4$	84
2-Iodophenyl- <i>p</i> - <i>t</i> -butylbenzoate	$0.46 \pm 0.17$	$4.4 \pm 0.5$	79

classical OBPs,<sup>29</sup> which we term here “double OBPs.” These OBPs are similar in sequence and probably structure with mosquitoes (Agam, *Aedes*) D7 salivary proteins.<sup>30</sup>

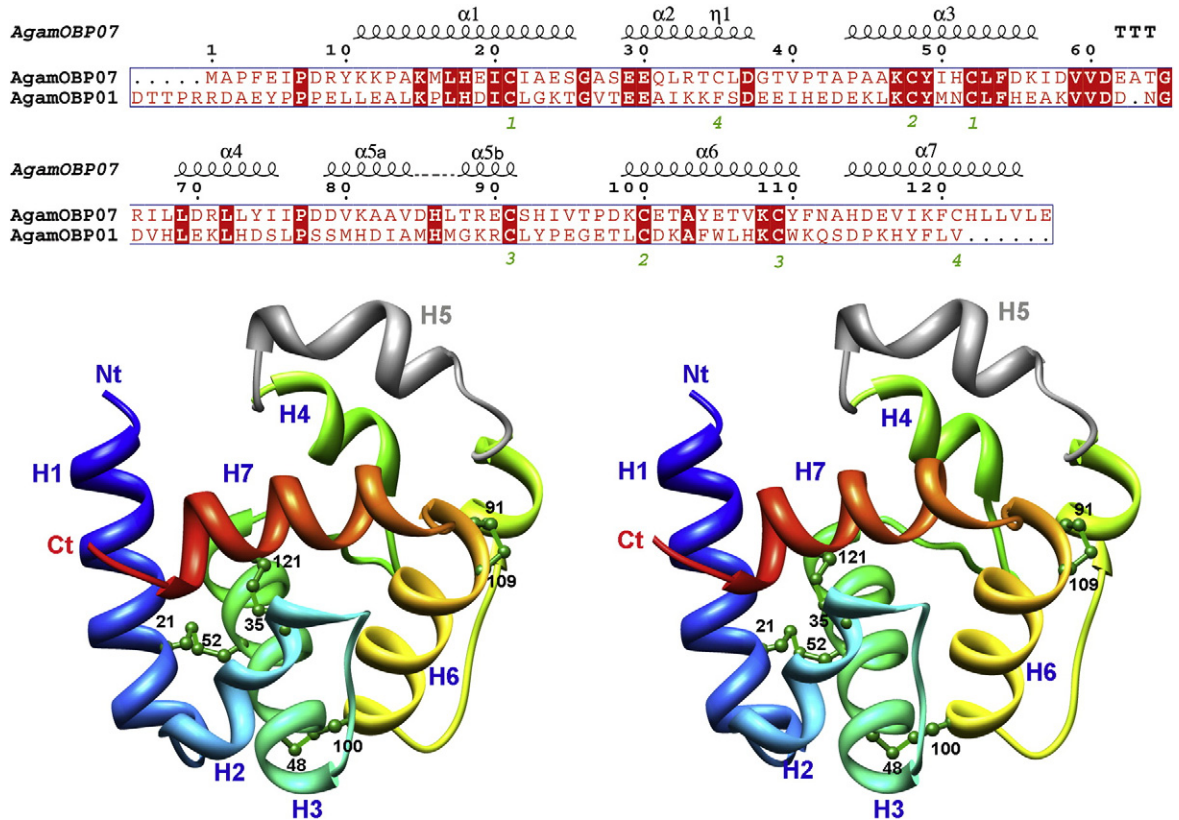
Besides, many OBPs do not fit in these large OBP classes, among which is AgamOBP07, a representative of a small group of OBPs containing cysteines. Here, we report the cloning, expression, characterization, binding properties and X-ray structure of AgamOBP07, the first member of a new class of OBPs. AgamOBP07 possesses eight cysteines and is slightly longer than medium C-terminal classical OBPs. Several occurrences of this OBP are observed in insect genomes. AgamOBP07 was found to be

expressed in all tissues, but the highest level was observed in the antennae (J.-J.Z., unpublished results). We suggest that AgamOBP07 is a member of an emerging new small class of OBPs that we term here the C8 OBP class.

## Results

### Protein characterization in solution

The protein oligomeric state was characterized in solution using HPLC with multi-angle laser light



**Fig. 2.** Sequence alignment and structure of AgamOBP07 from *A. gambiae*. (a) Sequences of AgamOBP07 and AgamOBP01 aligned with Multalin<sup>18</sup> and ESPript.<sup>19</sup> The secondary structures (top line) have been assigned according to AgamOBP07 crystal structure. The conserved residues are displayed in white on a red background. The disulfide bridges are numbered 1–4. (b) Ribbon representation with rainbow coloring mode, from blue (N-terminus) to red (C-terminus). The seven  $\alpha$ -helices have been labeled (H1–H7) as well as the Cys residues forming the four disulfide bridges (black numbers). Helix 5 is in gray and modeled from the closest OBP structure.



scattering/UV/refractometry (RI)-coupled detectors (Wyatt, Santa-Barbara, USA). The results point to a monomeric state, with a measured mass of 17 kDa (14.3 kDa theoretical mass) (Fig. S1). The CD spectra indicate that the protein is folded and contains mostly  $\alpha$ -helices (66%) followed by turns (12%) and random coil (16%), as expected from secondary structure prediction. The integrity of the protein was checked using SDS Coomassie blue gels. The protein after gel filtration, the protein from a crystallization drop without crystals and the protein from washed and dissolved crystals display bands with the same displacement.

### Ligands binding

We delipidated with lipidex AgamOBP07 before to perform binding assays, since the presence of a bound serendipitous ligand (see below) led to significant misvaluations of the binding constants.<sup>31</sup> We have measured the affinity of AgamOBP07 for a fluorescent reporter, *N*-phenyl-1-naphthylamine (1-NPN). Fluorescence at 410 nm increased when 1-NPN was added to the protein (Fig. S2a). The  $K_d$  calculated with the software package Prism was  $5.4 \pm 0.4 \mu\text{M}$ . We further used a 1:10 complex of AgamOBP07:1-NPN to assay the affinity of several ligands. This assay was based on the capacity of these ligands to displace 1-NPN, leading to a decrease in fluorescence. Our results indicated that several ligands were able to displace 1-NPN with significant fluorescence decrease and constants comprised between  $0.017 \mu\text{M}$  and  $0.66 \mu\text{M}$

(Fig. S2b and Table 1). In particular, 4-hydroxy-4'-isopropyl-azobenzene (AZO) yielded a  $K_d$  of  $0.12 \mu\text{M}$  and was able to chase up to 95% of 1-NPN at a concentration of  $20 \mu\text{M}$ . Noteworthy, 1-NPN data indicated a unique site of binding, while, with the other ligands, a better fit was obtained with two consecutive binding events. This fact may indicate that a second weaker site of binding was still available (or formed) after binding of one ligand molecule.

### Overall structure of AgamOBP07

The three-dimensional structure of AgamOBP07 was solved by molecular replacement using AgamOBP01 (2ERB<sup>14</sup>) as starting model with data collected at  $2.1 \text{ \AA}$  resolution from a  $P2_1$  crystal form. AgamOBP07 and AgamOBP01 share only 23% identity, but AgamOBP07 has only a unique residue insertion, relative to AgamOBP01, except for the differences at the N- and C-termini (Fig. 2a). AgamOBP07 crystals contain two molecules in the asymmetric unit. A second crystal form, crystallized under the same conditions, belongs to the  $P1$  space group and diffracted up to  $1.85 \text{ \AA}$  resolution. Molecular replacement using a monomer from the  $P2_1$  crystal form returned six molecules in the asymmetric unit (Table 2). After refinement, the final  $R/R_{\text{free}}$  values were 18.5/23.7% and 21.7/23.7%, respectively (Table 2). The rmsd values between the structures of AgamOBP07 (Form  $P1$ ) and AgamOBP01, the model used for molecular replacement, are  $2.5 \text{ \AA}$

**Table 2.** Data collection and refinement statistics

	$P2_1$ form	$P1$ form	$P2_1$ AZO complex
<i>Data collection</i>			
Beamline	BM14 (ESRF)	ID14-1 (ESRF)	Proxima 1 (Soleil)
Space group	$P2_1$	$P1$	$P2_1$
Cell parameters			
$a, b, c$ (Å)	46.7, 46.5, 57.4	46.8, 58.0, 78.1	46.7, 42.0, 57.95
$\beta$ (°)	103.8	85.9	106.0
$\alpha, \gamma$ (°)		69.2, 75.55	
Wavelength (Å)	0.97625	0.9782	
Resolution limits <sup>a</sup> (Å)	55.8–2.1 (2.2–2.1)	39.3–1.85 (1.95–1.85)	32.0–2.19 (2.25–2.19)
$R_{\text{merge}}$ <sup>a</sup> (%)	9.1 (19.3)	8.5 (51.1)	7.8 (64)
No. of observations <sup>a</sup>	72,623 (10,817)	130,934 (19,318)	57,178 (4150)
No. of unique reflections <sup>a</sup>	14,055 (2052)	59,725 (8815)	11,865 (923)
Mean( $I$ )/sd( $I$ ) <sup>a</sup>	22.1 (8.8)	13.4 (2.4)	12.2 (2.4)
Completeness <sup>a</sup> (%)	99.8 (99.3)	94.2 (95.0)	99.0% (99.1)
Multiplicity <sup>a</sup>	5.2 (5.3)	2.2 (2.2)	4.8 (4.5)
<i>Refinement</i>			
Resolution <sup>a</sup> (Å)	55.8–2.1 (2.3–2.1)	38.8–1.85 (1.9–1.85)	31.3–2.19 (2.4–2.19)
No. of reflections <sup>a</sup>	13,831 (2730)	59,719 (4227)	11,779 (2785)
No. of protein/ligands/solvent atoms	1751/36/191	5060/108/415	1874/36/44
No. of test set reflections	995 (7.2%)	3016 (5%)	640 (5%)
$R_{\text{work}}/R_{\text{free}}$ (%)	27.2/28.9	21.7/23.7	24.3/25.5
rmsd bonds (Å)/angles (°)	0.009/1.33	0.010/1.13	0.009/1.27
$B$ -factors: protein/C16/water (Å <sup>2</sup> )	21.73/24.8/42.5	28.8/61.7/36.3	39.6/38/45
Ramachandran: favored/allowed (%)	95.8/4.2	99.0/1.0	94.2/5.8

<sup>a</sup> Refers to the highest-resolution bin.

of the 121 residues in common and 1.5 Å of 96 residues when using a distance cutoff of 3 Å. The two crystal forms of AgamOBP07 provide thus eight independent monomers, allowing significant comparisons. The rmsd values calculated between all pairs of the six monomers in the  $P1$  space group range from 0.3 Å to 0.5 Å, while the rmsd value is 0.77 Å between the two  $P2_1$  monomers. When the two  $P2_1$  monomers are compared to the 6  $P1$  monomers, rmsd values of 0.5–0.7 Å are measured, indicating that the largest difference occurs between the two  $P2_1$  monomers.

The cloned protein is 127 residues long and possesses 8 cysteine residues (Fig. 2a). The chain was built in the electron density map with the exception of residues 1–9/12 at the N-termini and residues ~78–88 (Fig. 2b). The borders of these segments can vary depending on the crystallographic monomer. Residues 78–88 do not belong to a loop in the AgamOBP04 structure, but to helix 5 (Fig. 2b). Since helices are not the best candidates for providing disordered segments, we checked the integrity of the protein in the crystal (see above). Our analysis showed that the protein is intact, confirming that this segment is present but not ordered in the crystal. Furthermore, some uninterpretable bulbs of density are observed along the putative track of this segment, another clue confirming the disorder hypothesis. Finally, the  $P2_1$  crystal form in complex with an azobicyclic compound reveals that this segment is present and ordered in the crystal (see below). We could, however, model easily the missing segment by utilizing the corresponding peptide from AgamOBP01 and mutating its sequence to that of AgamOBP07. The borders of the segment fit well with the AgamOBP07 structure, and the sequence change did not introduce steric clashes (Fig. S3). The helix covers a side of AgamOBP07 core and might contribute slightly to the binding site wall (Fig. S3). We further used this model, in parallel with the authentic structure for our analysis.

AgamOBP07 protein core has the overall fold of classical OBPs and is mostly helical (Fig. 2b). In AgamOBP07, however, seven  $\alpha$ -helices are observed, which differs from the six  $\alpha$ -helices observed in all classical OBPs such as Bmori PBP (at neutral pH) or Lush (Fig. 1). The seventh C-terminal helix (114–127) is a feature of AgamOBP07 comparable only to that observed in BmorGOBP02.<sup>23</sup> However, in contrast with BmorGOBP02, AgamOBP07 seventh helix is part of the internal cavity wall. Three disulfide bridges (21/52, 48/100 and 91/109) are observed at similar positions in space as compared to their positions in the classical OBPs. In AgamOBP07, a fourth disulfide bridge (35/121) is formed that links the seventh C-terminal helix to the core of the protein.

It has been reported that OBPs may form functional dimers.<sup>32,33</sup> Equivocal data have been

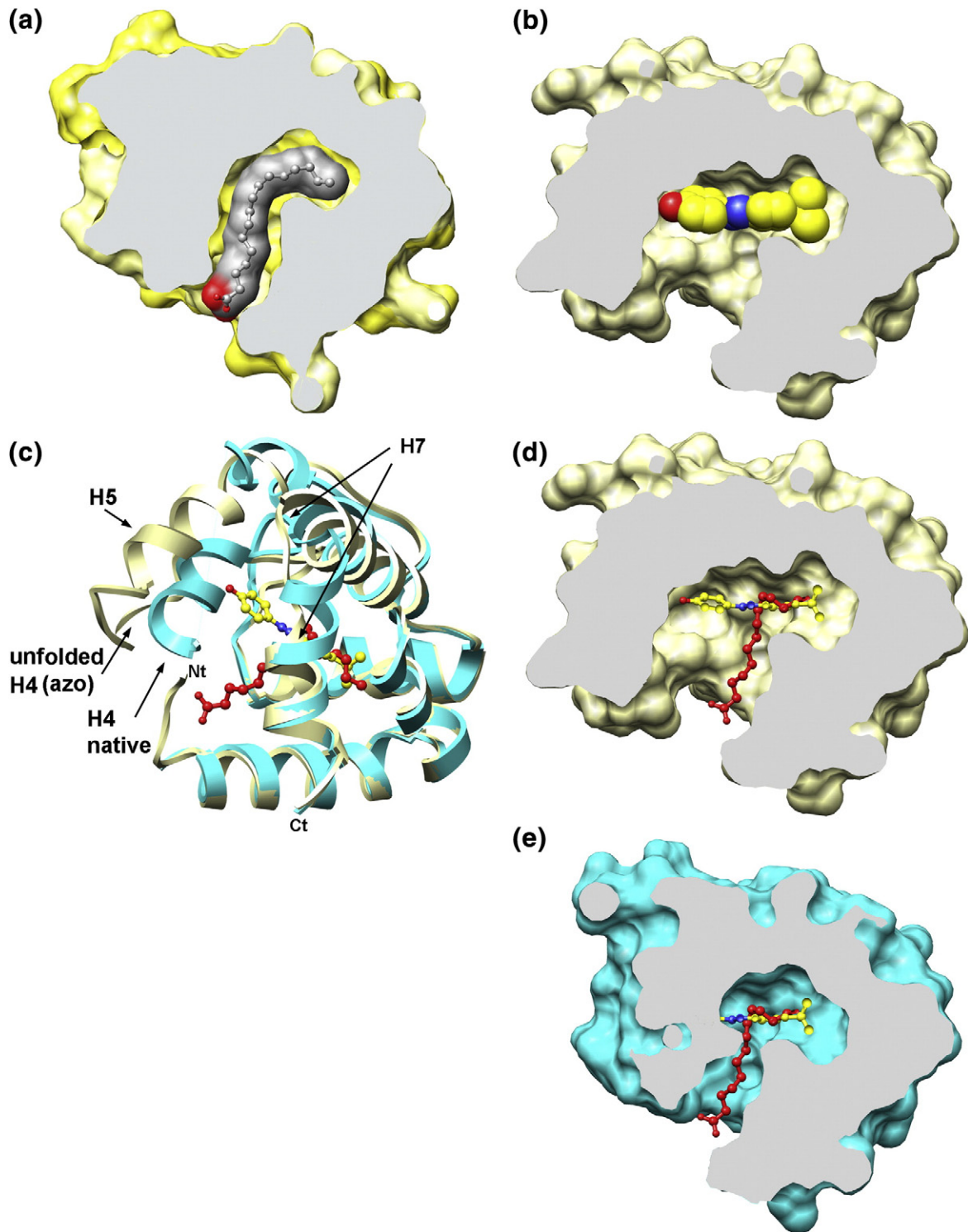
reported, however, that do not allow to conclude firmly for or against dimerization of OBPs in their biological context. In the crystal, dimers have been observed in the case of AgamOBP01, but the authors do not conclude that this observation may also apply to solutions, due to an insufficient contact area.<sup>15</sup> In the case of AgamOBP07, we observe a monomer of 17 kDa using multi-angle static light scattering (MALS)/dynamic light scattering and RI coupled on-line with a size-exclusion chromatography (SEC) column (MALS/UV/RI/SEC) (Fig. S1). The samples are in this case at ~0.3 mM concentration. In our crystallization experiments, the samples are at ~0.6 mM. The PISA server<sup>34</sup> did not identify any putative dimer in both crystal forms.

### AgamOBP07 binding site and its serendipitous ligand

AgamOBP07 possesses an elongated binding site with a volume of  $800 \pm 60 \text{ \AA}^3$  (CASTp<sup>35</sup>). We observed an elongated electron density in the structures obtained from AgamOBP07 crystals of the purified protein, in both space groups, and in all monomers (eight occurrences; Fig. S4). This elongated density from a serendipitous ligand was best modeled with a 16-carbon-long fatty acid [palmitic acid (PA)], although the exact nature of this ligand remains still unknown (Fig. 3a and Fig. S4). This fatty acid fits perfectly with the shape of the cavity, in which it buries  $275 \text{ \AA}^2$  of water-accessible surface of a total of  $650 \pm 20 \text{ \AA}^2$  (Fig. 3a). It contacts only aliphatic or aromatic hydrophobic residues (Table 3).

### Complex of AgamOBP07 with AZO

We have obtained crystals of AgamOBP07 co-crystallized with AZO. These crystals took several weeks to appear, had a different morphology from the previous ones and were yellow colored, indicating that formation of a complex with AZO might have occurred. These crystals belong to the space group  $P2_1$ , as the crystal form 2 of the purified protein, but with a ~4.5-Å shorter crystal  $b$ -axis length, leading to a decrease in solvent content. The OBPs packing in the crystals of the purified and the complexed protein are similar, however. We have collected a data set at 2.19 Å resolution and solved the structure by molecular replacement from the unsoaked protein structure (Table 2). We found drastic changes in the structure as compared to the two previous structures from the purified protein (Fig. 3b and c). Firstly, the chain could be traced from 10/11 to 133 continuously, despite two areas with weaker electron density in subunit A around residues 83–85 and 120–122. The first half of the molecule (11–67) adopts a structure very close to that of the native structure, with rmsd values of 0.64 Å on average. The second half (68–133) deviates



**Fig. 3.** Molecular surface of AgamOBP07 and bound ligands. (a) Surface of the native protein in complex with the serendipitous ligand, modeled as PA. (b) Surface of the complex with the AZO compound, in the same orientation as in (a). (c) Ribbon views of the structure of the native protein (blue) in complex with the serendipitous ligand, (PA, red balls-and-sticks) superimposed to the structure of the complex with the AZO compound (yellow, ligand in color-coded balls-and-sticks). (d) Surface view of the AZO complex structure with the AZO ligand (atom color coded) and the serendipitous ligand (red) represented. Note that the serendipitous ligand position is compatible with the AZO complex binding site structure. (e) Surface view of the native structure with the AZO ligand (atom color coded) and the serendipitous ligand (red) represented. Note that the AZO position is not compatible with the native binding site structure.

**Table 3.** List of the residues forming the binding site of AgamOBP07 in the native/palmitic and the AZO complexed forms and of the residues involved in interactions with the ligand

	AZO binding site	AZO contacts	Native binding site	PA contacts
Phe4	•			
Glu5	•			
Ile6	•			
Pro7	•			
Tyr10	•		•	
Pro13	•		•	•
Ala14	•		•	
Leu17	•		•	•
Leu32	•		•	
Cys35	•	•	•	•
Val40	•		•	
Pro41	•	•	•	•
Ala46			•	
Tyr49	•	•	•	•
Ile50	•	•	•	
Leu53	•		•	•
Phe54	•		•	•
Ile57	•		•	•
Leu69	•	•	•	
Leu72	•			
Ile75	•			
Ile76	•			
Asp78	•			
Asp79	•			
Val80	•			
Ala83	•			
Val84	•	•		
Leu87	•			
Val107	•		•	•
Tyr110	•		•	
Phe111	•	•	•	
His114	•	•	•	
Glu116	•		•	
Val117	•	•	•	
Lys119	•			
Phe120	•	•	•	•
Cys121	•	•	•	
Leu123	•		•	•
Leu124	•		•	•
Val125	•		•	

Cutoff is 3.8 Å. The binding site residues have been identified with CASTp, and the contacting residues, with CCP4.

significantly ( $\sim 1.9$  Å on average) with peaks of 3–4 Å at residues 70, 77, 96, 117 and 120 (Fig. S5). These fairly large conformational changes result from an opening of the binding cavity, the volume of which doubled compared to the native structure, with values of 1640 and 1460 Å<sup>3</sup> for subunits A and B, respectively. The ligand is visible in both subunits but is better defined in subunit B. The ligand molecules cover 342 Å<sup>2</sup> and 332 Å<sup>2</sup> of 1000 Å<sup>2</sup> and 930 Å<sup>2</sup> in subunits A and B, respectively. Because AZO is a non-flexible, almost linear compound that cannot fit in the bent binding site present in the native protein harboring the PA ligand, the cavity opening itself is probably the

result of AZO pushing helix 4 and unfolding it (Fig. 3c–e). As a consequence, helix 5 is stabilized and becomes visible in density probably due to better contacts with the rest of the protein, contacts themselves resulting from helix 4 unfolding (Fig. 3b and c).

The number of residues forming the binding site in the AZO complex is much larger than that in the PA complex, as defined by CASTp (Table 3). All but one of the residues that form the PA binding site are present in that of the AZO complex. These residues are mostly hydrophobic, but main-chain atoms of some hydrophilic residues contribute to the cavity wall. Binding interactions are exclusively formed by van der Waals contacts for both PA and AZO complexes.

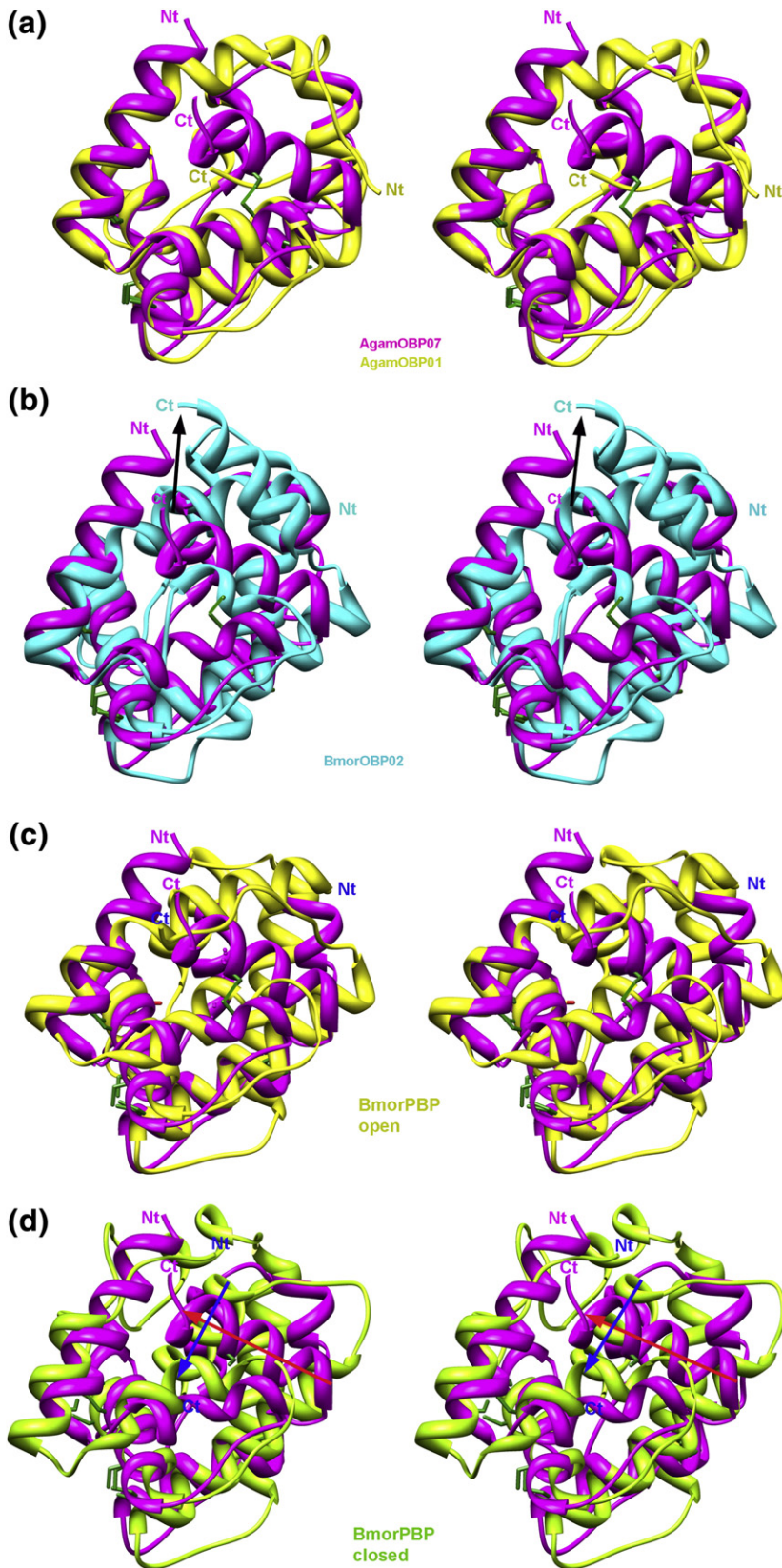
### Comparison of AgamOBP07 with other OBPs of known structures

The OBP closest to native AgamOBP07 is AgamOBP01, with an rmsd of 2.3 Å and a Z-score of 11.8, followed by Amel ASP1 with the same rmsd but a lower DALI Z-score of 11.4. Dmel Lush follows, with an rmsd of 2.8 Å ( $Z=9.4$ ), the best-characterized OBP from dipterans,<sup>8,12</sup> a member of the medium-length classical OBPs. Noteworthy, when comparing the AgamOBP07 structure in complex with AZO, Amel ASP1 has the best rmsd (2.0 Å;  $Z=15.5$ ), followed by AgamOBP01 (2.1 Å;  $Z=13.9$ ) and Dmel Lush (2.4 Å;  $Z=12.4$ ).

When comparing Agam OBP01 and OBP07, we noticed that the  $\alpha$ -helices are well aligned (Fig. 4a) and the disulfide bridges superimpose well, explaining why this structure could be used successfully in molecular replacement. The largest differences occur at the chain termini. AgamOBP01 has a longer N-terminus consisting, before helix 1, in an elongated stretch followed by a kinked  $\alpha$ -helical segment. The C-terminus is elongated and forms a wall of the cavity. The C-terminus of AgamOBP07 is located at a similar position but forms a long helix, a wall of the ligand binding cavity.

The fit of AgamOBP07 and BmorGOBP2 is less good than that with the members of the medium C-terminal classical OBPs. The rmsd between the two structures is 3.2 Å with a DALI Z-score of 6.4. BmorGOBP2 has also an  $\alpha$ -helix at its C-terminus (Fig. 4b).<sup>23</sup> Its N-terminus is also longer and mostly  $\alpha$ -helical. Noteworthy, the C-terminal helix of BmorGOBP2 does not occupy a similar position as that of AgamOBP07: it is displaced by  $\sim 15$  Å, does not form a wall of the cavity and is external to the protein core. We therefore cannot consider AgamOBP07 and BmorGOBP2 as belonging to the same subclass since the C-terminal helix plays a very different role in both proteins. BmorGOBP2 is most probably a variation of the long C-terminal classical OBPs from lepidopteran.





**Fig. 4.** Stereo ribbon views of AgamOBP07 (purple) structure superimposed with selected OBPs. (a) Comparison with Agam OBP1 (yellow). (b) Comparison with Bmor OBP2 (blue). The black arrow underlined the differences in the seventh loop positions. (c) Comparison with Bmor PBP open form obtained at neutral pH (yellow). (d) Comparison with Bmor PBP closed form obtained at acidic pH (green). The red arrows show the main axis and orientation of AgamOBP07 last helix, while the blue arrow shows the main axis and orientation of BmorPBP closed last internal helix.



We have then compared AgamOBP07 with the best-documented member of the long C-terminal classical OBPs, the PBP from *B. mori*. The superposition score is also mediocre, with an rmsd of 3.2 Å ( $Z=7.1$ ; 1DQE). The structure of BmorPBP was first solved at neutral pH, in complex with *Bombyx*-specific pheromone, bombykol.<sup>10</sup> Bombykol was found to occupy the central cavity, while, after the sixth helix, the C-terminus wrapped around the protein surface (Fig. 4c). The first helix of BmorPBP is longer and tilted significantly toward the protein core. The main axes of the other helices are also displaced by 2–4 Å with respect to that of AgamOBP07 helices. The closed form was obtained at acidic pH and solved first by NMR<sup>11</sup> and later by X-ray diffraction.<sup>36,37</sup> The core of the closed form exhibits about the same features as those of the open form. The two noticeable exceptions are the N-terminus, which does not form a long helix, and the C-terminus, which folds into a seventh helix located inside the bombykol binding cavity of the open form (Fig. 4d). This seventh buried helix indeed has a totally different topology compared to the seventh helix of AgamOBP07, which forms a wall of the binding site and is bound by a disulfide bridge to the protein core.

## Discussion

### How to classify AgamOBP07?

From the comparisons above, it is clear that the AgamOBP07 C-terminus is very different topologically and functionally from those of BmorGOBP2 and BmorPBP. The next question is to determine whether AgamOBP07 is an isolated case or if it belongs to a new OBP class. In order to answer this question, we performed a sequence search of AgamOBP07-like OBPs in Agam genome. This search did not return any protein with eight cysteines. The closest Agam OBPs returned were successively OBP17, OBP01, OBP03, OBP23, OBP24 and others, all belonging to classical OBPs. However, a close look at Agam OBPs revealed that AgamOBP15 possesses a C-terminus with the same length as that of AgamOBP07, with an extra cysteine residue. Its position, however, is not conserved, nor is the other cysteine that forms the disulfide bridge in AgamOBP07. However, a sequence search in all available insect genomes returned six orthologous OBPs with eight cysteines at conserved positions and with sequence identities comprised between 99% and 76% (Fig. S6). They all belong to the Diptera order and Culicidae family and to the sub-families of Anophelinae, *Culex* and *Aedes*. We propose therefore that AgamOBP07 is a representative of a small subclass of OBPs that we propose to name the Cys8 subclass.

## Evolutionary aspects

Noteworthy, this subclass of OBP was not found in the more ancient insect species such as Hymenoptera, Coleoptera and Lepidoptera or in the Drosophilidae family. Culicidae and Drosophilidae families separated ~250–210 million years ago (early Triassic) after which the Cys8 family evolved, probably from classical OBPs. Classical OBPs, in contrast, are localized at the root of insect families, since they are found in locusts and aphids. However, the minus-C class OBPs was found to be more abundant in the more ancient insects, which led us to postulate that they might be the most ancestral OBPs from which classical OBPs might have evolved.<sup>26</sup> In contrast, plus-C OBPs, being found only in more recent insects, might have evolved from classical OBPs. It seems therefore that the OBP driving force in evolution is toward introducing a larger number of disulfide bridges and more complexity. In this context, AgamOBP07 would form a novel evolutionary branch, developing in parallel with the larger plus-C OBP class.

## Materials and Methods

### AgamOBP07 gene cloning

The adult mosquitoes *A. gambiae* s.s. were kindly provided by Dr. Mark Rowland of the London School of Hygiene and Tropical Medicine, London, UK. On arrival, the insects were immediately frozen in liquid nitrogen and then stored at  $-80^{\circ}\text{C}$  until use. Total RNA was isolated using Trizol reagent (Invitrogen Life Technologies) following the manufacturer's instructions. Then, poly(A<sup>+</sup>)-RNA was separated using the poly-ATtract mRNA Isolation System IV (Promega, Madison, WI). First-strand cDNAs were synthesized with oligo(dT)<sub>15</sub> primers (Promega) and SuperScript™ III Reverse Transcriptase RNase H<sup>-</sup> (Invitrogen Life Technologies) and used as the PCR template. The primers for the mature protein were designed based on AgamOBP07 sequence (GenBank No: AY146742) flanked by NdeI and EcoRI restriction sites in the forward and reverse primers, respectively (5'-catatgGCCCCATTCGAAATACCGGAC-3' and 5'-cttaagTCACTCCAGCACTAGTAGGTG-3'). The PCR products were cloned into the pGEM T easy vector (Promega) and then subcloned into the bacterial expression vector pET17b (Novagen, Darmstadt, Germany) between the NdeI and EcoRI restriction sites. The inserts of the resultant construct pET17b-Ag7 in the resulting plasmids were sequenced to confirm that they encoded the correct mature proteins.

### Protein expression and purification

The pET17b-Ag7 plasmid with the correct insert was transformed into BL21(DE3)pLysS *Escherichia coli* cells, and protein synthesis was induced at OD<sub>600</sub> of 0.5–0.8 with IPTG (4 mM final concentration) at 37 °C for 3 h. All

proteins were found to be expressed as inclusion bodies, and solubilization was performed by denaturation in urea/DTT, renaturation and extensive dialysis at  $-4^{\circ}\text{C}$  in a cold room for 12 h in 20 mM Tris buffer, pH 7.4. The recombinant proteins were purified by two rounds of anion-exchange chromatography with a HiPrep 16/40 column (GE Healthcare, Hatfield, UK) filled with DE-52 resin (Whatman, Kent, UK), followed by several runs of gel filtration on a Superdex S75 Hiload 26/60 column (GE Healthcare). A MonoQ column was also used at the final stage of purification. The purified proteins were stored at  $-20^{\circ}\text{C}$  in 20 mM Tris-HCl buffer, pH 7.4.

The integrity of the protein was analyzed by migration on a microcapillary chip (Caliper LabChip GX II; Life Sciences) with SDS and Coomassie blue. CD experiments were recorded using 0.1 mg/ml AgamOBP07 in 10 mM NaHPO<sub>4</sub>, pH 7.5, on a Jasco J-810 spectropolarimeter.

The monodispersity of purified AgamOBP07 was analyzed by a combination of UV spectrophotometry, MALS and RI, coupled on-line with an analytical SEC column, as already described.<sup>38</sup> A 15-ml KW-803 column (Shodex) was used on an Alliance HPLC 2695 system (Waters). The protein concentration was adjusted to 6 mg/ml. A 50 mM Tris (pH 7.5) and 50 mM NaCl buffer and a flow of 0.5 ml min were used for column equilibration and running.

### Fluorescence experiments

AgamOBP07 was delipidated with lipidex 1000 before to perform binding assays using the procedure described by Oldham *et al.*<sup>31</sup> Data have been processed by the Prism/GraphPad software. Fluorescence experiments on AgamOBP07 with 1-NPN were carried out on a Varian Eclipse spectrofluorimeter using a quartz cuvette in a right-angle configuration. The interactions were monitored by recording 1-NPN fluorescence upon addition of 1-NPN aliquots. The excitation wavelength was 337 nm, and emission spectra were recorded at 350–450 nm. The excitation slit was 5 nm, while the emission slit was 10 nm. A moving-average smoothing procedure was applied, with a window of 3 nm. Titrations were carried out at  $20^{\circ}\text{C}$  with 0.5  $\mu\text{M}$  protein in 50 mM Tris, pH 8.0. The fluorescence intensities at the maximum of emission (400–410 nm) for different concentrations of ligand were corrected for the buffer contribution before plotting and further analysis. The affinity was estimated by plotting the increase in fluorescence intensity (1-NPN addition) or the decrease in fluorescence intensity (competing ligands displacing 1-NPN). The  $K_d$  values were estimated using Prism 4 (GraphPad Software, Inc.) by nonlinear regression for one (1-NPN) or two sites of binding.

The binding of ligands to AgamOBP07 was measured by adding aliquots of competing ligand dissolved in dimethyl sulfoxide to a solution containing 1-NPN 5  $\mu\text{M}$  with 0.5  $\mu\text{M}$  protein. Fluorescence intensities at the maximum of emission (400–410 nm) were determined for different concentrations of competitor and were corrected before further data analysis by subtracting the background of 1-NPN and ligand in the cuvette.

### Crystallization of AgamOBP07

Purified AgamOBP07 was concentrated to 20 mg/ml in 15 mM Tris (pH 8.0) and 25 mM NaCl buffer, using an

Amicon 10-kDa-cutoff concentrator. Screening experiments were performed with several commercial kits, using the sitting-drop method and the nanodrop technology in Greiner plates.<sup>39,40</sup> The first crystals appeared in Stura Footprint Screens (Molecular Dimension). After optimization, more crystals grew in 33 mM Hepes (pH 7.25–8.25), 0.4–1 M KCl, 6.6 mM sodium borate and 0.83 M sodium citrate tribasic. They reached a final size of  $250\ \mu\text{m} \times 150\ \mu\text{m} \times 50\ \mu\text{m}$ . Crystals were frozen for data collection in a buffer containing 1.5 M sodium citrate tribasic. AgamOBP07 was co-crystallized with AZO, by adding 5 molar equivalents of ligand in the protein sample. Yellow-colored crystals grew the under same conditions as the native form and were frozen in a buffer containing 1.5 M sodium citrate tribasic.

### Structures determination and refinement

A first data set was collected at  $2.1\ \text{\AA}$  resolution on the native protein at BM14 [European Synchrotron Radiation Facility (ESRF), Grenoble], integrated using XDS<sup>41</sup> and scaled with XSCALE.<sup>41</sup> The structure was solved by molecular replacement with MOLREP<sup>42</sup> using AgamOBP1 (entry 2ERB<sup>14</sup>) as starting model, with the  $P2_1$  space group, and two molecules in the asymmetric unit ( $V_m = 2.16\ \text{\AA}^3/\text{Da}$  and 43% solvent). Another data set was collected at  $1.85\ \text{\AA}$  at ID14-1 (ESRF) with a crystal grown under identical conditions as the previous one, but belonging to the  $P1$  space group. Molecular replacement with Phaser,<sup>43</sup> using the  $P2_1$  model, led to the positioning of six molecules in the asymmetric unit ( $V_m = 2.28\ \text{\AA}^3/\text{Da}$  and 46% solvent). The complex with the AZO compound was collected at Proxima 1 (Synchrotron Soleil, Saint-Aubin, France). The crystal belongs to the  $P2_1$  space group of the native form, but with a shorter  $b$ -axis, and contains also two molecules in the asymmetric unit ( $V_m = 2.04\ \text{\AA}^3/\text{Da}$  and 40% solvent). Structure refinement was performed with AutoBUSTER<sup>44</sup> using TLS parameters and autoNCS, alternated with rebuilding using Coot<sup>45</sup> (Table 2). Figures were made with PyMOL<sup>†46</sup> or Chimera.<sup>47</sup>

### Accession codes

The structure coordinates have been deposited to the Protein Data Bank at the Research Collaboratory for Structural Bioinformatics<sup>‡</sup> with accession codes 3R1O (native  $P2_1$ ), 3R1P (native  $P1$ ) and 3R1V (AZO complex).

### Acknowledgements

This work was supported in part by the PACA Region, by the Marseille-Nice-Genopole® and by the FP7 EU project Enaromatic (FP7/2007-2013, Grant Agreement No. FP7-222927).

<sup>†</sup> <http://pymol.sourceforge.net/>

<sup>‡</sup> <http://www.rcsb.org/pdb/>

We are grateful to Dr. Pierre Legrand for help with data collection and structure determination and to the Synchrotron Soleil and European Synchrotron Facility (Grenoble, France) for beamtime allocation.

## Supplementary Data

Supplementary data to this article can be found online at [doi:10.1016/j.jmb.2011.10.005](https://doi.org/10.1016/j.jmb.2011.10.005)

## References

1. Krieger, J., Mameli, M. & Breer, H. (1997). Elements of the olfactory signaling pathways in insect antennae. *Invertebr. Neurosci.* **3**, 137–144.
2. Vogt, R. G. & Riddiford, L. M. (1981). Pheromone binding and inactivation by moth antennae. *Nature*, **293**, 161–163.
3. Krieger, J., Ganssle, H., Raming, K. & Breer, H. (1993). Odorant binding proteins of *Heliothis virescens*. *Insect Biochem. Mol. Biol.* **23**, 449–456.
4. Vosshall, L. B. & Stocker, R. F. (2007). Molecular architecture of smell and taste in *Drosophila*. *Annu. Rev. Neurosci.* **30**, 505–533.
5. Sato, K., Pellegrino, M., Nakagawa, T., Vosshall, L. B. & Touhara, K. (2008). Insect olfactory receptors are heteromeric ligand-gated ion channels. *Nature*, **452**, 1002–1006.
6. Wicher, D., Schafer, R., Bauernfeind, R., Stensmyr, M. C., Heller, R., Heinemann, S. H. & Hansson, B. S. (2008). *Drosophila* odorant receptors are both ligand-gated and cyclic-nucleotide-activated cation channels. *Nature*, **452**, 1007–1011.
7. Stengl, M., Zufall, F., Hatt, H. & Hildebrand, J. G. (1992). Olfactory receptor neurons from antennae of developing male *Manduca sexta* respond to components of the species-specific sex pheromone *in vitro*. *J. Neurosci.* **12**, 2523–2531.
8. Laughlin, J. D., Ha, T. S., Jones, D. N. & Smith, D. P. (2008). Activation of pheromone-sensitive neurons is mediated by conformational activation of pheromone-binding protein. *Cell*, **133**, 1255–1265.
9. Forstner, M., Breer, H. & Krieger, J. (2009). A receptor and binding protein interplay in the detection of a distinct pheromone component in the silkworm *Antheraea polyphemus*. *Int. J. Biol. Sci.* **5**, 745–757.
10. Sandler, B. H., Nikonova, L., Leal, W. S. & Clardy, J. (2000). Sexual attraction in the silkworm moth: structure of the pheromone-binding-protein-bombbykol complex. *Chem. Biol.* **7**, 143–151.
11. Horst, R., Damberger, F., Luginbuhl, P., Guntert, P., Peng, G., Nikonova, L. *et al.* (2001). NMR structure reveals intramolecular regulation mechanism for pheromone binding and release. *Proc. Natl Acad. Sci. USA*, **98**, 14374–14379.
12. Kruse, S. W., Zhao, R., Smith, D. P. & Jones, D. N. (2003). Structure of a specific alcohol-binding site defined by the odorant binding protein LUSH from *Drosophila melanogaster*. *Nat. Struct. Biol.* **10**, 694–700.
13. Lartigue, A., Gruez, A., Briand, L., Blon, F., Bezirard, V., Walsh, M. *et al.* (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.
14. Wogulis, M., Morgan, T., Ishida, Y., Leal, W. S. & Wilson, D. K. (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.
15. 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.
16. 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.
17. Tegoni, M., Campanacci, V. & Cambillau, C. (2004). Structural aspects of sexual attraction and chemical communication in insects. *Trends Biochem. Sci.* **29**, 257–264.
18. Corpet, F. (1988). Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res.* **16**, 10881–10890.
19. Gouet, P., Robert, X. & Courcelle, E. (2003). ESPript/ENDscript: extracting and rendering sequence and 3D information from atomic structures of proteins. *Nucleic Acids Res.* **31**, 3320–3323.
20. Damberger, F. F., Ishida, Y., Leal, W. S. & Wuthrich, 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.
21. Pesenti, M. E., Spinelli, S., Bezirard, V., Briand, L., Pernollet, J. C., Campanacci, V. *et al.* (2009). Queen bee pheromone binding protein pH-induced domain swapping favors pheromone release. *J. Mol. Biol.* **390**, 981–990.
22. Pesenti, M. E., Spinelli, S., Bezirard, V., Briand, L., 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.
23. Zhou, J. J., Robertson, G., He, X., Dufour, S., Hooper, A. M., Pickett, J. A. *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.
24. Campanacci, V., Lartigue, A., Hallberg, B. M., Jones, T. A., Giudici-Orticoni, M. T., Tegoni, M. & Cambillau, C. (2003). Moth chemosensory protein exhibits drastic conformational changes and cooperativity on ligand binding. *Proc. Natl Acad. Sci. USA*, **100**, 5069–5074.
25. Foret, S. & Maleszka, R. (2006). Function and evolution of a gene family encoding odorant binding-like proteins in a social insect, the honey bee (*Apis mellifera*). *Genome Res.* **16**, 1404–1413.
26. Spinelli, S., Lagarde, L., Iovinella, I., Legrand, P., Tegoni, M., Pelosi, P. and Cambillau, C. Crystal structure of *Apis mellifera* OBP 14, a C-minus odorant-binding protein, and its complexes with odorant molecules. *Insect Biochem. Mol. Biol.* In press. [doi:10.1016/j.jmb.2011.10.005](https://doi.org/10.1016/j.jmb.2011.10.005).



27. Zhou, J. J., Huang, W., Zhang, G. A., Pickett, J. A. & Field, L. M. (2004). "Plus-C" odorant-binding protein genes in two *Drosophila* species and the malaria mosquito *Anopheles gambiae*. *Gene*, **327**, 117–129.
28. 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+ class. *Biochem. J.* **437**, 423–430.
29. Hekmat-Safe, D. S., Scape, C. R., McKinney, A. J. & Tanouye, M. A. (2002). Genome-wide analysis of the odorant-binding protein gene family in *Drosophila melanogaster*. *Genome Res.* **12**, 1357–1369.
30. Calvo, E., Mans, B. J., Ribeiro, J. M. C. & Andersen, J. F. (2009). Multifunctionality and mechanism of ligand binding in a mosquito antiinflammatory protein. *Proc. Natl Acad. Sci. USA*, **106**, 3728–3733.
31. Oldham, N. J., Krieger, J., Breer, H. & Svatos, A. (2001). Detection and removal of an artefact fatty acid from the binding site of recombinant *Bombyx mori* pheromone-binding protein. *Chem. Senses*, **26**, 529–531.
32. Qiao, H., He, X., Schymura, D., Ban, L., Field, L., Dani, F. R. *et al.* (2011). Cooperative interactions between odorant-binding proteins of *Anopheles gambiae*. *Cell. Mol. Life Sci.* **68**, 1799–1813.
33. Biessmann, H., Andronopoulou, E., Biessmann, M. R., Douris, V., Dimitrats, S. D., Eliopoulos, E. *et al.* (2010). The *Anopheles gambiae* odorant binding protein 1 (AgamOBP1) mediates indole recognition in the antennae of female mosquitoes. *PLoS One*, **5**, e9471.
34. Krissinel, E. & Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797.
35. 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–W118.
36. Lautenschlager, C., Leal, W. S. & Clardy, J. (2005). Coil-to-helix transition and ligand release of *Bombyx mori* pheromone-binding protein. *Biochem. Biophys. Res. Commun.* **335**, 1044–1050.
37. Lautenschlager, C., Leal, W. S. & Clardy, J. (2007). *Bombyx mori* pheromone-binding protein binding nonpheromone ligands: implications for pheromone recognition. *Structure*, **15**, 1148–1154.
38. Veessler, D., Blangy, S., Siponen, M., Vincentelli, R., Cambillau, C. & Sciara, G. (2009). Production and biophysical characterization of the CorA transporter from *Methanosarcina mazei*. *Anal. Biochem.* **388**, 115–121.
39. Sulzenbacher, G., Gruez, A., Roig-Zamboni, V., Spinelli, S., Valencia, C., Pagot, F. *et al.* (2002). A medium-throughput crystallization approach. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **58**, 2109–2115.
40. Lartigue, A., Gruez, A., Briand, L., Pernollet, J. C., Spinelli, S., Tegoni, M. & Cambillau, C. (2003). Optimization of crystals from nanodrops: crystallization and preliminary crystallographic study of a pheromone-binding protein from the honeybee *Apis mellifera* L. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **59**, 919–921.
41. Kabsch, W. (2010). Xds. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **66**, 125–132.
42. Vagin, A. & Teplyakov, A. (1997). MOLREP: an Automated Program for Molecular Replacement. *J. Appl. Crystallogr.* **30**, 1022–1025.
43. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674.
44. Blanc, E., Roversi, P., Vornrhein, C., Flensburg, C., Lea, S. M. & Bricogne, G. (2004). Refinement of severely incomplete structures with maximum likelihood in BUSTER-TNT. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **60**, 2210–2221.
45. Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **60**, 2126–2132.
46. DeLano, W. The PyMOL Molecular Graphics System (<http://pymol.sourceforge.net/>), DeLano Scientific LLC, San Carlos, CA, USA.
47. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. (2004). UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612.