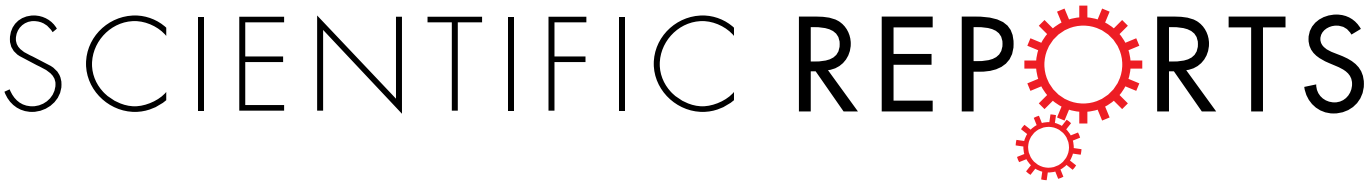

OPEN

Received: 25 April 2017
Accepted: 16 November 2017
Published online: 04 December 2017

Key site residues of pheromone-binding protein 1 involved in interacting with sex pheromone components of *Helicoverpa armigera*

Kun Dong[1,2], Hong-Xia Duan[3], Jing-Tao Liu[1], Liang Sun[1,4], Shao-Hua Gu[1], Ruo-Nan Yang[1], Khalid Hussain Dhiloo[1,5], Xi-Wu Gao[2], Yong-Jun Zhang[1] & Yu-Yuan Guo[1]

Pheromone binding proteins (PBPs) are widely distributed in insect antennae, and play important roles in the perception of sex pheromones. However, the detail mechanism of interaction between PBPs and odorants remains in a black box. Here, a predicted 3D structure of PBP1 of the serious agricultural pest, *Helicoverpa armigera* (HarmPBP1) was constructed, and the key residues that contribute to binding with the major sex pheromone components of this pest, (*Z*)-11- hexadecenal (*Z*11-16:Ald) and (*Z*)-9- hexadecenal (*Z*9-16:Ald), were predicted by molecular docking. The results of molecular simulation suggest that hydrophobic interactions are the main linkage between HarmPBP1 and the two aldehydes, and four residues in the binding pocket (Phe12, Phe36, Trp37, and Phe119) may participate in binding with these two ligands. Then site-directed mutagenesis and fluorescence binding assays were performed, and significant decrease of the binding ability to both *Z*11-16:Ald and *Z*9-16:Ald was observed in three mutants of HarmPBP1 (F12A, W37A, and F119A). These results revealed that Phe12, Trp37, and Phe119 are the key residues of HarmPBP1 in binding with the *Z*11-16:Ald and *Z*9-16:Ald. This study provides new insights into the interactions between pheromone and PBP, and may serve as a foundation for better understanding of the pheromone recognition in moths.

Pheromones perception is crucial for insects to seek out sexual partners[1]. For Lepidoptera species, sex pheromones blends are normally produced at accurate proportion and emitted by females to attract conspecific males for mating[2,3]. Such specific perception benefits from male moths' sophisticated olfactory system including numerous antennal sensilla, especially the sensilla trichodea, which are sensitive to different sex pheromone components[4,5]. The pheromone detection in male moths is initiated when pheromone molecules enter the lymph of trichoid sensilla through multipores[2,6] and it is widely accepted that several different groups of olfactory proteins, such as pheromone-binding proteins (PBPs), chemosensory proteins (CSPs), sensory neuron membrane proteins (SNMPs), odorant receptors (ORs) and ionotropic receptors (IRs) are involved in the process of pheromone detection[7,8].

Pheromone binding proteins are small (15–17 kDa) water-soluble proteins which present in the sensillar lymph with extremely high concentrations (up to 10 mM)[9,10]. These proteins are thought to solubilize, capture and transport hydrophobic pheromone molecules across the aqueous sensillar lymph to the pheromone receptors (PRs)[11,12]. The first *PBP* gene was identified in the silkworm, *Antheraea polyphemus*[13], then their orthologous genes have been identified in many Lepidoptera species[14]. Further research revealed that PBPs may specifically

[1]State Key Laboratory for Biology of Plant Diseases and Insect Pests, Institute of Plant Protection, Chinese Academy of Agricultural Sciences, Beijing, 100193, China. [2]Department of Entomology, China Agricultural University, Beijing, 100193, China. [3]College of Science, China Agricultural University, Beijing, 100193, China. [4]Key Laboratory of Tea Biology and Resources Utilization, Ministry of Agriculture, Tea Research Institute, Chinese Academy of Agricultural Sciences, Hangzhou, 310008, China. [5]Department of Entomology, Faculty of Crop Protection, Sindh Agriculture University Tandojam, Tandojam, Pakistan. Kun Dong and Hong-Xia Duan contributed equally to this work. Correspondence and requests for materials should be addressed to Y.-J.Z. (email: yjzhang@ippcaas.cn)

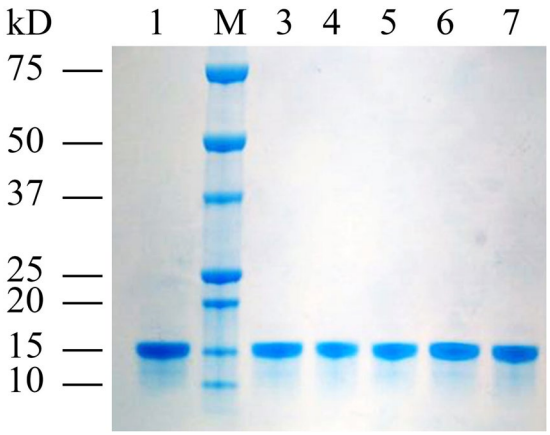

Figure 1. Sodium dodecyl sulphate polyacrylamide gel electrophoresis analysis of recombinant protein HarmPBP1 and mutant proteins. 1: wild-type HarmPBP1; 2: protein molecular weight marker; 3: mutant F12A; 4: mutant F36A; 5: mutant W37A; 6: mutant F119A; 7: mutant Q64A.

recognize distinct pheromone components and enhance the sensitivity of PRs in response to pheromones[15,16]. Because of the high sensitivity to pheromone components, PBPs are often served as the molecular targets to design the attractants of moths or other insect species[17,18].

It is well accepted that insect PBPs play important roles in pheromone perception[7,9]. However, the detail interaction mechanism between pheromones and PBPs is still unknown. Many three-dimensional (3-D) structures of insect PBPs have been solved in the crystal forms or in solution since the structure of BmorPBP/bombykol complex was reported[19–23]. Most insect PBPs exhibit series of identical structure characteristics including six or seven α-helices, three strictly conserved disulfide bridges, and a hydrophobic binding pocket. However, structure diversity is also observed and such differences make insect PBPs show different cavity shapes and openings to accommodate distinct ligands[19,23–27]. Various studies suggested that lepidopteran PBPs existed pH-dependent conformational change associated with significant decrease in affinity at low pH values[18,19,21,22,28–31]. The C-terminals of moth PBPs fold into an additional α-helix and enter the binding pocket to occupy the corresponding pheromone-binding sites at acid pH, whereas at neutral pH, the additional helix withdraws from the binding pocket and made it available for pheromone binding[19,24]. Other insect PBPs with short C-terminals, such as the LmaPBP in cockroach, could not form the additional helix but make a lid to cover the binding pocket, and such 'lid' would also affect the binding between PBPs to ligands[23]. All the research revealed that insect PBPs own diverse mechanisms in ligand binding and release, and such mechanisms relate closely to the structures of PBPs. It also suggested that the structural study at molecular level should be helpful in understanding of the action mode and binding specificity between pheromones and PBPs.

In recent years, the interactions between ligands and insect PBPs have been proposed based on the diversity of key residues. Many amino acids have been identified as the critical residues for ligands binding[19,25,32]. In moth species, the structure of BmorPBP/bombykol complex revealed that Ser56 forms a specific hydrogen bond between bombykol and BmorPBP[19], and in *A. polyphemus*, Asn53 had been confirmed to be the key site in specific recognition of acetate[25]. Besides, the structure of LUSH/cVA complex in *Drosophila melanogaster* showed that cVA forms two polar interactions with Ser52 and Thr57 in the binding pocket[32].

The cotton bollworm, *Helicoverpa armigera*, is one of the most serious agriculture pests worldwide and cause great damage to cotton and other crops[33]. This insect utilize *Z*11-16:Ald and *Z*9-16:Ald as the primary components of the pheromone blend[3]. Previously, three PBP genes, *HarmPBP1-3* have been identified and the results of fluorescence-binding assay revealed that HarmPBP1 equally bind the two principal pheromone components with strong affinities[34,35]. HarmPBP1 may play key roles in the pheromone perception of *H. armigera*. In the present study, we built a 3D model of the HarmPBP1 structure to predict the potential binding sites by homology modeling and molecular docking. The binding roles of these predicted residues were further investigated by site-directed mutagenesis and fluorescence binding assays. This work will help to deeply understand the interaction between HarmPBP1 and sex pheromone components in *H.armigera*.

Results

Expression of recombinant HarmPBP1. The coding region of HarmPBP1 was sub-cloned into an *E. coli* expression vector pET-32a/TEV and confirmed by PCR and sequencing. The protein expression was induced for 12 h by adding IPTG (1.0 mM) into the cell culture. The induced and non-induced cells were solicited and the crude inclusion body and supernatant were analyzed by SDS-PAGE. It was found that the recombinant HarmPBP1 was expressed in both supernatant and inclusion body. Then, the supernatant was collected and purified by His-Trap affinity columns (GE Healthcare, USA) followed by removal of the his-tag with TEV Protease. SDS-PAGE analysis indicated that the molecular weight of the final purified HarmPBP1 was about 15kD (Fig. 1), which is consistent with the theoretical molecular weight calculated by a computer pI/Mw online program (http://web.expasy.org/compute_pi/).

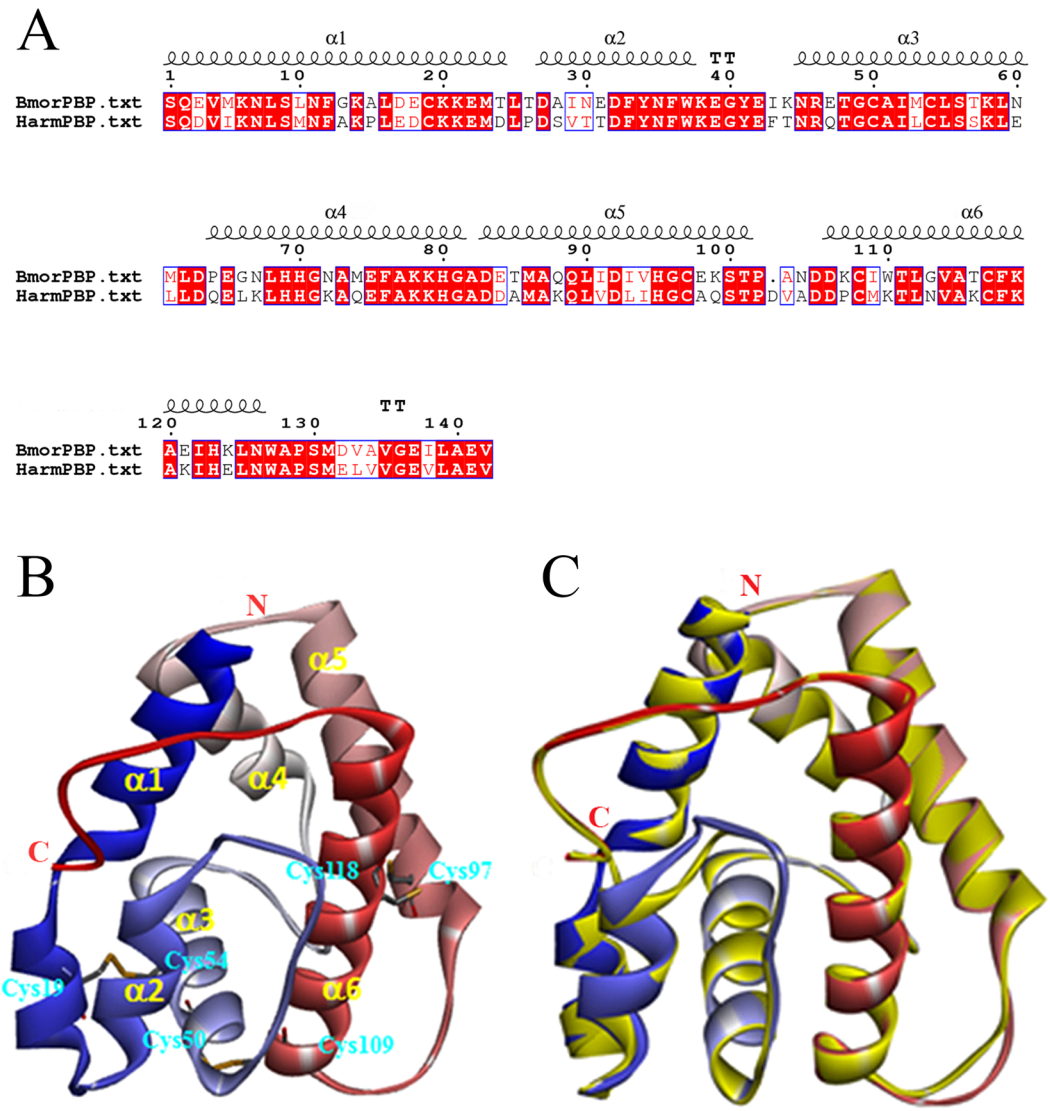

Figure 2. 3-D structure model of HarmPBP1. (**A**) Sequence alignment of HarmPBP1 and BmorPBP. (**B**) Predicated 3-D model of HarmPBP1. Helices, N-terminal (N) and C-terminal (**C**) are labelled. Disulphide bridges are coloured yellow. (**C**) The alignment plot of the target protein HarmPBP1 and the template protein BmorPBP (yellow).

3-D structure modeling and molecular docking. From a BLAST search in the Protein Data Bank (PDB), four structurally determined OBPs, *Bombyx mori* PBP (BmorPBP), *Amyelois transitella* PBP (AtraPBP1), *Antheraea polyphemus* PBP (ApolPBP) and *Bombyx mori* OBP (BmorGOBP2) were selected to share sequence similarities with HarmPBP1. The total sequence identity between the target protein (HarmPBP1) and the template protein (BmorPBP) is 67% (Fig. 2A). Thus, to guarantee the quality of the homology model, BmorPBP with the high level of sequence identity was used as a template to construct the 3D structure of HarmPBP1.

The overlap between 3D model of HarmPBP1 and template showed a high similarity of 0.828, which revealed that the overall conformation of target protein is very similar to the template (Fig. 2B,C). The predicted 3D structure demonstrated that HarmPBP1 is a "classical PBP". Six α-helices were located between residues Ser1-Asp24 (α1), Asp27-Trp37 (α2), Asn45-Glu60 (α3), Gln64-Gly81 (α4), Asp83-Thr101 (α5), and Asp107-Asn127 (α6). Four antiparallel helices (α1, α4, α5 and α6) converge to form the hydrophobic binding pocket. The converging ends of the helices formed the narrow end of the pocket, and the opposite end of the pocket is capped by α3 (Fig. 2B). Disulphide bonds and helix–helix packing enforce the organization of the helices. Three pairs of disulfide bridges are observed between Cys19-Cys54, Cys50-Cys109, and Cys97-Cys118 (Fig. 2B). In this model, most of the amino acid residues that formed the pocket were hydrophobic, such as phenylalanine, tryptophan, alanine, valine, leucine, and isoleucine.

To further investigate the potential key residues in HarmPBP1, *Z*11-16:Ald and *Z*9-16:Ald were selected to dock with the 3D model. The docking results showed that both the two ligands are consistent in orientation, and

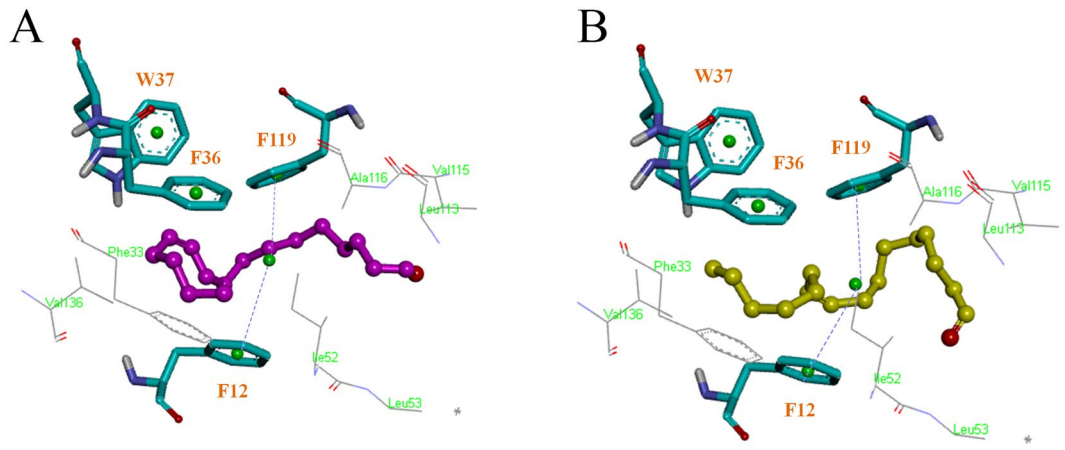

Figure 3. Docking results of *Z*11-16:Ald with the model (**A**), and *Z*9-16:Ald with the model (**B**).

highly overlapped in the same tunnel of binding pocket. Moreover, the oxygen of *Z*11-16:Ald and *Z*9-16:Ald are located in the similar position of binding cavity (Figure S1).

Widely hydrophobic interactions have been observed as the main linkage between HarmPBP1 and the two aldehydes. All the hydrophobic residues included phenylalanine12 (Phe12), phenylalanine33 (Phe33), phenylalanine36 (Phe36), tryptophan37 (Trp37), Isoleucine52 (Ile52), Leucine53 (Leu53), Leucine113 (Leu113), Valine115 (Val115), Alanine116 (Ala116), phenylalanine119 (Phe119), and Valine136 (Val136) with less than 7.0 Å distances to *Z*11-16:Ald and *Z*9-16:Ald (Fig. 3). Amongst of these residues, Phe12 and Phe119 showed a sandwich-like pose to locate the binding conformation of ligands. Such hydrophobic contacts were much favorable to the binding between protein and ligands due to the nonpolar aromatic ring of Phe residue. Trp37 and Phe36 also provided certain nonpolar binding effects on the ligands with different sidechain.

Site-directed mutagenesis and binding characterization of mutants. Based on the 3-D structure modeling and molecular docking described above, combined with an X-ray structure of the HarmPBP1/*Z*9-16:Ald complex (unpublished data), we predicted that four residues (Phe12, Phe36, Trp37, and Phe119) may play important roles in ligand binding. To verify the importance of such residues, the alanine scanning mutagenesis modeling have been performed, and the binding free energy for *Z*11-16:Ald and the wild-type (WT) or four mutants of HarmPBP1 were calculated (Table S1). Mutants F12A and F119A showed significant differences on binding to *Z*11-16:Ald from the WT. Meanwhile, W37A also showed a certain effect on the binding of *Z*11-16:Ald. However, the binding free energy of *Z*11-16:Ald and F36A changed only slightly compare to that of *Z*11-16:Ald and WT.

All the four residues were mutated to alanine, respectively, by using a site-directed mutagenesis kit. In addition, Gln64, a randomly selected residue on the loop between helices α3 and α4, was mutated to alanine as a control. The recombinant mutants F12A, F36A, W37A, F119A, and Q64A were expressed and purified as described above. The purified proteins were also checked by SDS-PAGE (Fig. 1). It was showed that the expression levels of mutants were apparently the same as that of wild-type HarmPBP1.

Fluorescence binding assays were performed in a reaction system at pH7.4. Probed by 1-NPN, the maximum emission wavelengths of F12A, F36A, W37A, F119A and Q64A were in the range of 390–410 nm, which are similar to that of HarmPBP1 (400 nm). The dissociation constant (Kd) of F12A/1-NPN, F36A/1-NPN, W37A/1-NPN, F119A/1-NPN, Q64A/1-NPN, and HarmPBP1/1-NPN complexes were 2.1 ± 0.17, 4.05 ± 0.69, 1.87 ± 0.18, 1.96 ± 0.16, 2.39 ± 0.19 and $1.79 \pm 0.14\,\mu M$, respectively (Figure S2). These results revealed that the Kd values of all mutants were closed to that of wild type HarmPBP1.

The affinities of all mutants with *Z*11-16:Ald and *Z*9-16:Ald were also investigated by fluorescence binding assays (Fig. 4). The results showed that compared to the wild-type HarmPBP1, each of the four mutants, F12A, F36A, W37A and F119A showed a different degree of decline in their binding capacities to the sex pheromone compounds, whereas, there was almost no change in the binding ability of Q64A with the two ligands. Three mutants, F12A, W37A and F119A had lower affinities to both *Z*11-16:Ald and *Z*9-16:Ald than that of F36A (Table 2). Compared to the wild-type HarmPBP1 ($0.67 \pm 0.05\,\mu M$ to *Z*11-16:Ald and $0.56 \pm 0.05\,\mu M$ to *Z*9-16:Ald, separately), F119A showed the most dramatic decrease in binding capacity, with the dissociation constant (Ki) of $5.11 \pm 0.47\,\mu M$ to *Z*11-16:Ald and $4.61 \pm 0.33\,\mu M$ to *Z*9-16:Ald, respectively. F12A had a four to five fold decline in its affinity, with the Ki of $3.10 \pm 0.13\,\mu M$ to *Z*11-16:Ald and $3.06 \pm 0.22\,\mu M$ to *Z*9-16:Ald. W37A also showed decrease in binding ability to *Z*11-16:Ald and *Z*9-16:Ald, with Ki of $1.95 \pm 0.08\,\mu M$ and $1.49 \pm 0.13\,\mu M$, respectively. However, compared to the wild-type HarmPBP1, F36A demonstrated only a slight decline in binding to *Z*9-16:Ald. Thus, three amino acids, Phe12, Trp37, and Phe119 in the binding pocket of HarmPBP1, should be the key residues which involved in the binding of *Z*11-16:Ald and *Z*9-16:Ald.

Discussion

PBPs are known to bind and transport hydrophobic pheromone molecules across the sensillum lymph to PRs, and enhance the sensitivity of PRs to sex pheromones[13,14,16,36–39]. It was also reported that PBPs could specifically bind distinct pheromone components[11,15,40], and such binding specificity was attributed to the spatial structure

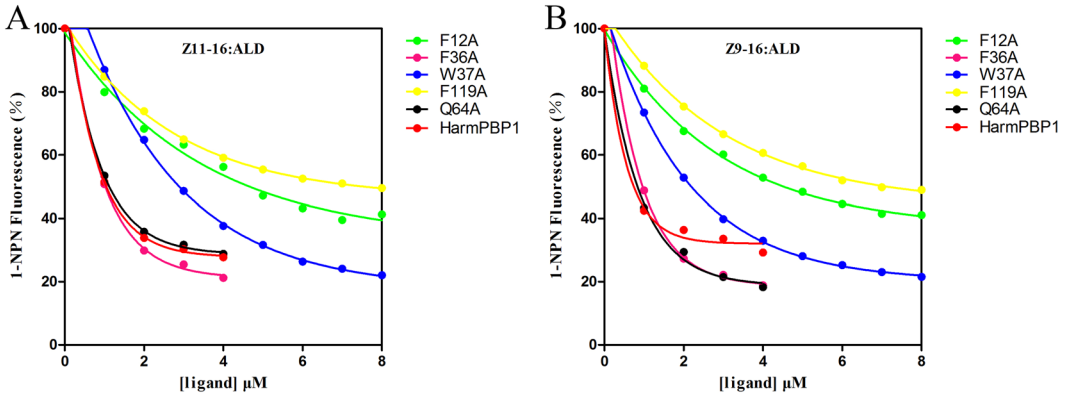

Figure 4. Competitive binding curves of *Z*11-16:Ald and *Z*9-16:Ald to the wild-type and mutants of HarmPBP1. (**A**) Binding curves of *Z*11-16:Ald to wild-type HarmPBP1 and all mutants. (**B**) Binding curves of *Z*9-16:Ald to wild-type HarmPBP1 and all mutants.

Primer name	Sequence (5′-3′)
For recombinant proteins expression	
HarmPBP1–forward	GGCCATGGCGTCGCAAGATGTTATTA[a]
HarmPBP1- reverse	GGAAGCTTTTAGACTTCGGCCAAG[a]
For site-directed mutagenesis	
F12A-forward	CCTCTCTATGAATGCCGCTAAGCCCTTAG[b]
F12A-reverse	CTAAGGGCTTAGCGGCATTCATAGAGAGG[b]
F36A-forward	CTTCTACAACGCCTGGAAGGAAGGC[b]
F36A-reverse	GCCTTCCTTCCAGGCGTTGTAGAAG[b]
W37A-forward	CTACAACTTCGCGAAGGAAGGCTAC[b]
W37A-reverse	GTAGCCTTCCTTCGCGAAGTTGTAG[b]
F119A-forward	GGCCAAGTGCGCCAAGGCCAAGATA[b]
F119A-reverse	TATCTTGGCCTTGGCGCACTTGGCC[b]
Q64A-forward	GCTACTGGACCAGGAGCTCAAGC[b]
Q64A-reverse	GCTTGAGCTCCTGGTCCAGTAGC[b]

Table 1. Primers used in this study. [a]"__"represent the restriction sites, [b]"__" represent the mutation sites.

of proteins and ligands, especially their specific interactions[41]. As a result, clarifying the structure of insect PBPs should be helpful in better understanding of their binding mechanisms and biological roles in pheromone perception. In previous study, some crystal structures of lepidopteran PBPs have been solved by NMR or X-ray diffraction[19,21,22]. However, the structures of *H. armigera* OBP/PBPs are still lack.

Three PBPs of *H. armigera* have been reported in our previous study[35]. The results of fluorescence binding assay showed that HarmPBPs could specifically bind to different pheromone components of *H. armigera*[34,35,42]. The main composition of *H. armigera* pheromone blend contain two hexadecane, *Z*11-16:Ald and *Z*9-16:Ald[43]. Both *Z*11-16:Ald and *Z*9-16:Ald own similar size of the carbon chain, and HarmPBP1 showed stronger affinities to these two aldehydes than to other minor components[34,35]. Therefore, we decided to predict the structure of HarmPBP1 by using 3D homology modeling, and *Z*11-16:Ald and *Z*9-16:Ald were selected as suitable ligands to dock with this structure.

From a BLAST research in the PDB, BmorPBP1 (1DQE) with most sequence similarity (67% identify) to HarmPBP1 was selected as the template to build a 3D homology structure of HarmPBP1. Subsequent docking results revealed that the binding cavity of HarmPBP1 is mainly formed by hydrophobic residues, and *Z*11-16:Ald and *Z*9-16:Ald are well overlapped in the binding packet (Figure S1). Widely hydrophobic interaction was observed to contribute the binding between protein and ligands, but no hydrogen action was found in this structure. Actually, although hydrogen bonds have been confirmed to be the primary link between proteins and ligands in several insect OBPs[44–47], there are still some OBPs that only form hydrophobic interactions or van der Waals interactions[48,49]. In the docking structure of HarmPBP1, Phe12 and Phe119 are located on the two sides of the ligands, respectively, and the molecular plane of ligands is sandwiched by these two residues with their aromatic rings parallel (Fig. 3). Such sandwich-like pose contributes to solidify the binding conformation of ligands, so we suspected that Phe12 and Phe119 should be the important binding sites. Phe36 and Trp37 are close to the ligands, which may also play roles in the formation of hydrophobic interactions. Hence, we predicted that four active sites, Phe12, Phe119, Phe36 and Trp37, were possibly responsible for the ligand binding of HarmPBP1. The alanine scanning mutagenesis modeling was later performed to verify such prediction. The results showed that mutants F12A and F119A were of remarkable difference in binding to *Z*11-16:Ald from the wild-type of

Proteins	Z11-16:Ald		Z9-16:Ald	
	IC_{50} (μM)	Ki (μM)	IC_{50} (μM)	Ki (μM)
HarmPBP1	1.03 ± 0.08	0.67 ± 0.05a	0.87 ± 0.08	0.56 ± 0.05a
F12A	4.69 ± 0.33	3.10 ± 0.13c	4.63 ± 0.27	3.06 ± 0.22c
F36A	1.04 ± 0.06	0.76 ± 0.08a	0.98 ± 0.15	0.72 ± 0.11a
W37A	2.91 ± 0.12	1.95 ± 0.08b	2.22 ± 0.19	1.49 ± 0.13b
F119A	7.66 ± 0.56	5.11 ± 0.47d	6.91 ± 0.49	4.61 ± 0.33d
Q64A	1.12 ± 0.21	0.77 ± 0.13a	0.84 ± 0.06	0.58 ± 0.03a

Table 2. Binding abilities of HarmPBP1 and mutants to *Z*11-16:Ald and *Z*9-16:Ald. Data represent the mean values ± S.E.M of three independent replicates. Different letters within the same column mean that the values were significantly different ($P < 0.05$). IC_{50}, ligand concentration displacing 50% of the fluorescence intensity of the protein/N-phenyl-1-naphthylamine complex; Ki, dissociation constant.

HarmPBP1, suggesting that these two residues of HarmPBP1 should be important on the ligand binding. W37A also showed a certain effect on the binding with *Z*11-16:Ald, indicating its potential contribution to the ligand binding. F36A demonstrated a slight change on the binding free energy of *Z*11-16:Ald, which suggested that this residue might not vital to the ligand binding.

Further site-directed mutagenesis and fluorescence binding assays were performed to characterize the binding abilities of the four mutants of HarmPBP1. A random mutation, Q64A was set as one of the control. The results of binding tests revealed that Q64A had no difference in affinity to *Z*11-16:Ald and *Z*9-16:Ald compared with the wild-type protein, which suggested that non-specific mutation could not affect the interactions between proteins and ligands. Both the single amino acid mutants, F12A and F119A could not efficiently bind to *Z*11-16:Ald and *Z*9-16:Ald. A possible explanation is that ligands cannot remain in the binding cavity due to the loss of the hydrophobic interactions between ligands and residues. Ligands are sandwiched by Phe12 and Phe119 with their aromatic rings, and such stable binding conformation was broken when any of these two residues was mutated to alanine. As a result, we suggested that Phe12 and Phe119 play the key roles in the ligand-binding of HarmPBP1. Mutant W37A showed a certain decrease in affinity to *Z*11-16:Ald and *Z*9-16:Ald due to the changes of hydrophobic interaction between the mutant and ligands. Thus, W37 is also an important binding site of HarmPBP1. Another mutant F36A, however, showed nearly no change in its binding ability to *Z*11-16:Ald and *Z*9-16:Ald. Therefore, we suspected that Phe36 may not be involved in the binding with *Z*11-16:Ald and *Z*9-16:Ald, or may participate in the binding with other ligands. All the four residues are highly conserved in lepidopteran PBPs and most GOBPs[19,25,35], but only Phe12 and Phe119 contribute significantly to bind with the *Z*11-16:Ald and *Z*9-16:Ald. Interestingly, these two residues also play important roles in the binding process between BmorPBP1 and Bombykol[19]. Moreover, in SlitOBP1, the mutants of Phe12 and Phe118 result in lower docking scores to all tested chemicals in the simulation of site-direct mutagenesis, and the recombinant mutant Phe12 could not bind to all the ligands which exhibit good affinities to the wild-type protein[50]. Such results suggest that some conserved hydrophobic residues, such as Phe12 and Phe119, may be responsible for non-specific binding among different lepidopteran OBPs. On the other hand, strictly conserved Phe36 had been confirmed to be the key residue of LdisPBP1 in binding with its pheromone and analogues[51]. However in the current study, the affinity of mutant F36A to *Z*11-16:Ald and *Z*9-16:Ald showed nearly no change compared with the wild-type protein. In view of such difference, we speculated that beside the amino acids which contribute to non-specific binding, some other residues should be the key sites in binding with specific components in lepidopteran OBPs. And it is important and interesting to further clarify such functional difference between the conserved residues in the binding pocket.

Our data indicated that multiple hydrophobic interactions play the key roles in the ligand binding of HarmPBP1. It was also revealed that besides the NMR or X-ray diffraction of protein–ligand complexes, molecular docking and the mutant binding assay could be a potential and effective tool to further analyze the molecular mechanisms of ligand-protein interactions. Moreover, the results of this study may serve as a foundation for future studies on integrated pest management through manipulating the pheromone detection of target insects.

Methods

Insects. A colony of *H. armigera* was maintained in the laboratory of the Institute of Plant Protection, Chinese Academy of Agricultural Sciences. Larvae were reared on an artificial diet, and the conditions were maintained at 26 ± 1 °C, 60% ± 5% RH, and L 14 h: D 10 h. After emergence, adult moths were fed with 10% honey solution. Antennae were removed from three days old male moths and were immediately stored in liquid nitrogen till to use.

RNA extraction and cDNA synthesis. Total RNA was isolated from antennae samples by SV Total RNA Isolation System (Promega, Madison, USA) following the manufacturer's protocol. The integrity of the RNA was checked by using 1.2% agarose gel electrophoresis and quantified using a ND-1000 spectrophotometer (NanoDrop, Wilmington, DE, USA) at OD260 nm. The high concentration (>800 ng/μL) of the total RNA showed that the high quality of the RNA sample meet the standard of reverse transcriptase reaction. The first strand cDNA was synthesized using the SuperScript™ III Reverse Transcriptase System (Invitrogen, Carlsbad, CA, USA).

Expression and purification of recombinant HarmPBP1. The full sequence of *HarmPBP1* was identified from *H. armigera* antennal cDNA library in our previously work[42]. The sequence encoding mature *HarmPBP1*

was amplified by PCR with gene-specific primers (Table 1). The PCR product was purified and sub-cloned into pGEM-T vector (Promega, Madison, USA). Target sequence was excised with *Nco* I and *Hind* III and then cloned into pET-32a/TEV vector (Novagen, Germany) with T4 DNA ligase. The correct recombinant plasmid pET/HarmPBP1 was transformed to BL21 (DE3) competent cells. Cells were incubated at 37 °C until OD_{600} reached 0.6–0.8, and the proteins were expressed after induction with 0.2 mM IPTG for 12 h. Cells were harvested by centrifugation at 7000 rpm for 20 min, and precipitate was re-suspended with 1 × phosphate-buffered saline (PBS). After ultrasonic, cells were centrifugalized at 16000 rpm for 20 min, then inclusion bodies and supernatant was collected and checked by 15% SDS-polyacrylamide gel electrophoresis (SDS-PAGE) analysis. The supernatant was filtered with a 0.22 μm ultrafiltration and purified by two rounds of Ni ion affinity chromatography (GE Healthcare,USA), and the His-tag was removed with Tobacco Etch Virus (TEV) protease (GenScript, Nanjing, China). The highly purified proteins were desalted through extensive dialysis. The size and purity of recombinant HarmPBP1 were confirmed by 15% SDS-PAGE analysis.

3D structure modeling and molecular docking. The 3D model of HarmPBP1 were built with a template of BmorPBP1 (1DQE) by using On-line Swiss-model software (https://www.swissmodel.expasy.org/). The binding cavity was predicted with an automobile mode by SYBYL 7.3 software. The molecular conformations of *Z*11-16:Ald and *Z*9-16:Ald were constructed by Sketch mode and optimized using the Tripos force field and Gasteiger-Hückel charge. The Surflex-Dock module of SYBYL 7.3 was employed to perform the molecular docking modeling[52]. The binding cavity was set as "Automatic" and the Total Score was used to evaluate the binding affinity between ligands and protein[53]. All molecular modeling between putative HarmPBP1 protein and ligands were conducted on the Silicon Graphics® (SGI) Fuel Workstation (Silicon Graphics International Corp., CA, USA).

Simulation of Site-directed mutagenesis and the expression of mutants. The alanine scanning mutagenesis modeling were performed by the AMBER 14 package[54] to verify the predicted key binding sites, and the binding free energy between the active site and *Z*11-16:Ald was calculated by the MM-GBSA method[55].

Four mutations of HarmPBP1, F12A (mutating phenylalanine to alanine at position 12), F36A (mutating phenylalanine to alanine at position 36), W37A (mutating tryptophan to alanine at position 37) and F119A (mutating phenylalanine to alanine at position 119) were generated by using the Quick-change lightning site-directed mutagenesis kit (Stratagene, USA), and a random mutation, Q64A (mutating glutanine to alanine at position 64) was set as control. The pGEM-T Easy/HarmPBP1 construct was used as template, and the specific primers designed for mutations were also listed in Table 1. The PCR conditions were 95 °C for 30 s, followed by 18 cycles of 95 °C for 30 s, 60 °C for 1 min, 68 °C for 4 min. Valid mutants were sub-cloned into pGEM-T easy vector (Promega, USA). Same expression vector and competent cells were used as the HarmPBP1. The recombinant mutant protein prokaryotic expression and purification were conducted as mentioned above.

Fluorescence binding assays. Fluorescence binding assays were conducted on the F-380 fluorescence spectrophotometer (Gangdong Sci. & Tech, Tianjin, China) in a 1-cm light path quartz cuvette to further investigate the binding abilities of the principal pheromone components of *H. armigera*, *Z*11-16:Ald and *Z*9-16:Ald, to mutants. The fluorescent probe N-phenyl- 1-naphthylamine (1-NPN) was dissolved in methanol to yield a 1 mM stock solution. Both of the excitation and emission slit widths were 10 nm. Fluorescence of 1-NPN was excited at 337 nm and the emission spectra were recorded between 390 and 490 nm. *Z*11-16:Ald and *Z*9-16:Ald were purchased from Sigma-Aldrich (purity >98%). All chemicals used in this study were dissolved in HPLC purity grade methanol. Fluorescence measurements were performed according to Gu *et al.*[11]. Dissociation constants of the competitors were calculated from the corresponding IC_{50} (the ligand concentration displacing 50% of the NPN fluorescence intensity of the HarmPBP1/1-NPN complex) values, using the equation: $Ki = [IC_{50}]/(1 + [1\text{-NPN}]/K_{1\text{-NPN}})$, where [1-NPN] is the free concentration of 1-NPN and $K_{1\text{-NPN}}$ is the dissociation constant of the HarmPBP1/1-NPN complex.

References

1. Leffingwell, J. C. Olfaction–Update No. 5. *Leffingwell reports.* **2**, 1–34 (2002).
2. Kaissling, K. Chemo-electrical transduction in insect olfactory receptors. *Annu Rev Neurosci.* **9**, 121–145 (1986).
3. Wang, H. L., Zhao, C. H. & Wang, C. Z. Comparative study of sex pheromone composition and biosynthesis in *Helicoverpa armigera*, *H. assulta* and their hybrid. *Insect Biochem Mol Biol.* **35**, 575–583, https://doi.org/10.1016/j.ibmb.2005.01.018 (2005).
4. Field, L. M., Pickett, J. A. & Wadhams, L. J. Molecular studies in insect olfaction. *Insect Mol Biol.* **9**, 545–551 (2000).
5. Takanashi, T. *et al.* Unusual response characteristics of pheromone-specific olfactory receptor neurons in the Asian corn borer moth. *Ostrinia furnacalis. J Exp Biol.* **209**, 4946–4956, https://doi.org/10.1242/jeb.02587 (2006).
6. Steinbrecht, R. A. Pore structures in insect olfactory sensilla: A review of data and concepts. *Int J Insect Morphol Embryol.* **26**, 229–245, https://doi.org/10.1016/s0020-7322(97)00024-x (1997).
7. Leal, W. S. Odorant reception in insects: roles of receptors, binding proteins, and degrading enzymes. *Annu Rev Entomol.* **58**, 373–391, https://doi.org/10.1146/annurev-ento-120811-153635 (2013).
8. Zhang, J., Walker, W. B. & Wang, G. Pheromone reception in moths: from molecules to behaviors. *Prog Mol Biol Transl Sci.* **130**, 109–128, https://doi.org/10.1016/bs.pmbts.2014.11.005 (2015).
9. Vogt, R. G., Prestwich, G. D. & Lerner, M. R. Odorant-binding-protein subfamilies associate with distinct classes of olfactory receptor neurons in insects. *J Neurobiol.* **22**, 74–84, https://doi.org/10.1002/neu.480220108 (1991).
10. Klein, U. Sensillum-lymph proteins from antennal olfactory hairs of the moth *Antheraea polyphemus* (Saturniidae). *Insect Biochem.* **17**, 1193–1204, https://doi.org/10.1016/0020-1790(87)90093-x (1987).
11. Gu, S. H., Zhou, J. J., Wang, G. R., Zhang, Y. J. & Guo, Y. Y. Sex pheromone recognition and immunolocalization of three pheromone binding proteins in the black cutworm moth *Agrotis ipsilon*. *Insect Biochem Mol Biol.* **43**, 237–251, https://doi.org/10.1016/j.ibmb.2012.12.009 (2013).
12. Sun, M., Liu, Y. & Wang, G. Expression patterns and binding properties of three pheromone binding proteins in the diamondback moth. *Plutella xyllotella. J Insect Physiol.* **59**, 46–55, https://doi.org/10.1016/j.jinsphys.2012.10.020 (2013).
13. Vogt, R. G. & Riddiford, L. M. Pheromone binding and inactivation by moth antennae. *Nature* **293**, 161–163 (1981).

14. Vogt, R. G., Grosse-Wilde, E. & Zhou, J. J. The Lepidoptera Odorant Binding Protein gene family: Gene gain and loss within the GOBP/PBP complex of moths and butterflies. *Insect Biochem Mol Biol.* **62**, 142–153, https://doi.org/10.1016/j.ibmb.2015.03.003 (2015).
15. Grosse-Wilde, E., Svatos, A. & Krieger, J. A pheromone-binding protein mediates the bombykol-induced activation of a pheromone receptor *in vitro*. *Chem senses.* **31**, 547–555, https://doi.org/10.1093/chemse/bjj059 (2006).
16. Forstner, M., Breer, H. & Krieger, J. 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 (2009).
17. Leal, W. S. *et al*. Reverse and conventional chemical ecology approaches for the development of oviposition attractants for Culex mosquitoes. *PloS One.* **3**, e3045, https://doi.org/10.1371/journal.pone.0003045 (2008).
18. Leal, W. S. *et al*. Olfactory proteins mediating chemical communication in the navel orangeworm moth. *Amyelois transitella*. *PloS One.* **4**, e7235, https://doi.org/10.1371/journal.pone.0007235 (2009).
19. Sandler, B. H., Nikonova, L., Leal, W. S. & Clardy, J. Sexual attraction in the silkworm moth: structure of the pheromone-binding-protein-bombykol complex. *Chem Biol.* **7**, 143–151, https://doi.org/10.1016/S1074-5521(00)00078-8 (2000).
20. Tegoni, M., Campanacci, V. & Cambillau, C. Structural aspects of sexual attraction and chemical communication in insects. *Trends Biochem Sci.* **29**, 257–264, https://doi.org/10.1016/j.tibs.2004.03.003 (2004).
21. Xu, X. *et al*. NMR structure of navel orangeworm moth pheromone-binding protein (AtraPBP1): implications for pH-sensitive pheromone detection. *Biochemistry.* **49**, 1469–1476, https://doi.org/10.1021/bi9020132 (2010).
22. Zubkov, S., Gronenborn, A. M., Byeon, I. J. & Mohanty, S. Structural consequences of the pH-induced conformational switch in *A.polyphemus* pheromone-binding protein: mechanisms of ligand release. *J Mol Biol.* **354**, 1081–1090, https://doi.org/10.1016/j.jmb.2005.10.015 (2005).
23. Lartigue, A. *et al*. The crystal structure of a cockroach pheromone-binding protein suggests a new ligand binding and release mechanism. *J Biol Chem.* **278**, 30213–30218, https://doi.org/10.1074/jbc.M304688200 (2003).
24. Horst, R. *et al*. NMR structure reveals intramolecular regulation mechanism for pheromone binding and release. *Proc Natl Acad Sci USA* **98**, 14374–14379, https://doi.org/10.1073/pnas.251532998 (2001).
25. Mohanty, S., Zubkov, S. & Gronenborn, A. M. The solution NMR structure of Antheraea polyphemus PBP provides new insight into pheromone recognition by pheromone-binding proteins. *J Mol Biol.* **337**, 443–451, https://doi.org/10.1016/j.jmb.2004.01.009 (2004).
26. Pesenti, M. E. *et al*. Structural basis of the honey bee PBP pheromone and pH-induced conformational change. *J Mol Biol.* **380**, 158–169, https://doi.org/10.1016/j.jmb.2008.04.048 (2008).
27. Pesenti, M. E. *et al*. Queen bee pheromone binding protein pH-induced domain swapping favors pheromone release. *J Mol Biol.* **390**, 981–990, https://doi.org/10.1016/j.jmb.2009.05.067 (2009).
28. Kowcun, A., Honson, N. & Plettner, E. Olfaction in the gypsy moth, *Lymantria dispar*: effect of pH, ionic strength, and reductants on pheromone transport by pheromone-binding proteins. *J Biol Chem.* **276**, 44770–44776, https://doi.org/10.1074/jbc.M104688200 (2001).
29. Damberger, F. F., Ishida, Y., Leal, W. S. & Wuthrich, K. 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, https://doi.org/10.1016/j.jmb.2007.07.078 (2007).
30. Liu, N. Y. *et al*. Two subclasses of odorant-binding proteins in *Spodoptera exigua* display structural conservation and functional divergence. *Insect Mol Biol.* **24**, 167–182, https://doi.org/10.1111/imb.12143 (2015).
31. Liu, N. Y., Liu, C. C. & Dong, S. L. Functional differentiation of pheromone-binding proteins in the common cutworm *Spodoptera litura*. *Comp Biochem Physiol A Mol Integr Physiol.* **165**, 254–262, https://doi.org/10.1016/j.cbpa.2013.03.016 (2013).
32. Laughlin, J. D., Ha, T. S., Jones, D. N. & Smith, D. P. Activation of pheromone-sensitive neurons is mediated by conformational activation of pheromone-binding protein. *Cell.* **133**, 1255–1265, https://doi.org/10.1016/j.cell.2008.04.046 (2008).
33. Guo, Y. Y. Progress in the researches on migration regularity of cotton bollworm and relationships between the pest and its host plants. *Acta Entomol Sinica.* **40**, 1–6 (1997).
34. Guo, H., Huang, L. Q., Pelosi, P. & Wang, C. Z. Three pheromone-binding proteins help segregation between two *Helicoverpa* species utilizing the same pheromone components. *Insect Biochem Mol Biol.* **42**, 708–716, https://doi.org/10.1016/j.ibmb.2012.06.004 (2012).
35. Zhang, T. T. *et al*. Characterization of three pheromone-binding proteins (PBPs) of *Helicoverpa armigera* (Hubner) and their binding properties. *J Insect Physiol.* **58**, 941–948, https://doi.org/10.1016/j.jinsphys.2012.04.010 (2012).
36. Steinbrecht, R. A., Ozaki, M. & Ziegelberger, G. Immunocytochemical localization of pheromone-binding protein in moth antennae. *Cell Tissue Res.* **270**, 287–302, https://doi.org/10.1007/bf00328015 (1992).
37. Pelosi, P., Mastrogiacomo, R., Iovinella, I., Tuccori, E. & Persaud, K. C. Structure and biotechnological applications of odorant-binding proteins. *Appl Microbiol Biotechnol.* **98**, 61–70, https://doi.org/10.1007/s00253-013-5383-y (2014).
38. Grosse-Wilde, E., Gohl, T., Bouche, E., Breer, H. & Krieger, J. Candidate pheromone receptors provide the basis for the response of distinct antennal neurons to pheromonal compounds. *Eur J Neurosci.* **25**, 2364–2373, https://doi.org/10.1111/j.1460-9568.2007.05512.x (2007).
39. Chang, H. T. *et al*. Pheromone binding proteins enhance the sensitivity of olfactory receptors to sex pheromones in *Chilo suppressalis*. *Sci Rep.* **5**, 13093, https://doi.org/10.1038/srep13093 (2015).
40. Mohl, C., Breer, H. & Krieger, J. Species-specific pheromonal compounds induce distinct conformational changes of pheromone binding protein subtypes from *Antheraea polyphemus*. *Invert Neurosci.* **4**, 165–174, https://doi.org/10.1007/s10158-002-0018-5 (2002).
41. Pelosi, P., Zhou, J. J., Ban, L. P. & Calvello, M. Soluble proteins in insect chemical communication. *Cell Mol Life Sci.* **63**, 1658–1676, https://doi.org/10.1007/s00018-005-5607-0 (2006).
42. Zhang, T. T., Gu, S. H., Wu, K. M., Zhang, Y. J. & Guo, Y. Y. Construction and analysis of cDNA libraries from the antennae of male and female cotton bollworms *Helicoverpa armigera* (Hubner) and expression analysis of putative odorant-binding protein genes. *Biochem Biophys Res Commun.* **407**, 393–399, https://doi.org/10.1016/j.bbrc.2011.03.032 (2011).
43. Kehat, M. & Dunkelblum, E. Behavioral response of male *Heliothis armigera* (Lepidoptera: Noctuidae) moths in a flight tunnel to combinations of components identified from female sex pheromone glands. *J Insect Behav.* **3**, 75–83, https://doi.org/10.1007/BF01049196 (1990).
44. Thode, A. B., Kruse, S. W., Nix, J. C. & Jones, D. N. 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, https://doi.org/10.1016/j.jmb.2007.12.063 (2008).
45. Zhou, J. J. *et al*. 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, https://doi.org/10.1016/j.jmb.2009.04.015 (2009).
46. Sun, Y. L., Huang, L. Q., Pelosi, P. & Wang, C. Z. A lysine at the C-terminus of an odorant-binding protein is involved in binding aldehyde pheromone components in two *Helicoverpa* species. *PloS One.* **8**, e55132 (2013).
47. Yin, J. *et al*. Three amino acid residues of an odorant-binding protein are involved in binding odours in *Loxostege sticticalis* L. *Insect Mol Biol.* **24**, 528–538, https://doi.org/10.1111/imb.12179 (2015).
48. Wogulis, M., Morgan, T., Ishida, Y., Leal, W. S. & Wilson, D. K. The crystal structure of an odorant binding protein from *Anopheles gambiae*: evidence for a common ligand release mechanism. *Biochem Biophys Res Commun.* **339**, 157–164, https://doi.org/10.1016/j.bbrc.2005.10.191 (2006).

49. Mao, Y. *et al*. 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, https://doi.org/10.1073/pnas.1012274107 (2010).
50. Yi, X. *et al*. Ligands binding and molecular simulation: the potential investigation of a biosensor based on an insect odorant binding protein. *Int J Biol Sci.* **11**, 75–87, https://doi.org/10.7150/ijbs.9872 (2015).
51. Ma, F., Xu, Y. X., Qin, H. & Luo, Y. H. Key sites residues involved in interacting with chemicals of pheromone-binding proteins from *Lymantria dispar*. *J Appl Entomol.* **138**, 733–742, https://doi.org/10.1111/jen.12143 (2015).
52. Jain, A. N. Surflex: fully automatic flexible molecular docking using a molecular similarity-based search engine. *J Med Chem.* **46**, 499–511, https://doi.org/10.1021/jm020406h (2003).
53. Jain, A. N. Scoring noncovalent protein-ligand interactions: a continuous differentiable function tuned to compute binding affinities. *J Comput Aided Mol Des.* **10**, 427–440 (1996).
54. Case, D. A. *et al*. AMBER 14, University of California, San Francisco (2014).
55. Miller, B. R. *et al*. MMPBSA.py: An Efficient Program for End-State Free Energy Calculations. *J Chem Theory Comput.* **8**, 3314–3321, https://doi.org/10.1021/ct300418h (2012).

Acknowledgements

This work was supported by the National Natural Science Foundation of China (31471778, 31672038, 31621064 and 31772176). This manuscript has been edited by the native English-speaking experts of Elsevier Language Editing Services.

Author Contributions

Conceived and designed the experiments: K.D., H.X.D., Y.J.Z., X.W.G. and Y.Y.G. Sample collection: K.D., J.T.L., L.S. and R.N.Y. Performed the experiments: K.D., J.T.L., L.S., S.H.G. and Khalid Hussain Dhiloo. Data analysis: K.D., S.H.G., H.W.L. and Y.J.Z. Wrote the paper: K.D., Y.J.Z. Manuscript revision: Y.Y.G., Y.J.Z., X.W.G.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-017-17050-5.

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.