

SCIENTIFIC REPORTS



OPEN

Structural insights into *Cydia pomonella* pheromone binding protein 2 mediated prediction of potentially active semiochemicals

Zhen Tian*, Jiyuan Liu* & Yalin Zhang

Received: 02 October 2015

Accepted: 08 February 2016

Published: 01 March 2016

Given the advantages of behavioral disruption application in pest control and the damage of *Cydia pomonella*, due progresses have not been made in searching active semiochemicals for codling moth. In this research, 31 candidate semiochemicals were ranked for their binding potential to *Cydia pomonella* pheromone binding protein 2 (CpomPBP2) by simulated docking, and this sorted result was confirmed by competitive binding assay. This high predicting accuracy of virtual screening led to the construction of a rapid and viable method for semiochemicals searching. By reference to binding mode analyses, hydrogen bond and hydrophobic interaction were suggested to be two key factors in determining ligand affinity, so is the length of molecule chain. So it is concluded that semiochemicals of appropriate chain length with hydroxyl group or carbonyl group at one head tended to be favored by CpomPBP2. Residues involved in binding with each ligand were pointed out as well, which were verified by computational alanine scanning mutagenesis. Progress made in the present study helps establish an efficient method for predicting potentially active compounds and prepares for the application of high-throughput virtual screening in searching semiochemicals by taking insights into binding mode analyses.

For most insects, such fundamental behaviors as mating, predation, and oviposition as well as avoiding threats are controlled by small molecule signals. Former research has revealed that olfaction system played a significant role in detecting chemical signals^{1,2}. In the olfaction system, to reach the membrane containing pheromone receptors, hydrophobic pheromones have to cross the aqueous sensillum lymph surrounding the dendrites of neuronal cells³. The crossing of this aqueous barrier is thought to be assisted by pheromone binding proteins (PBPs)⁴⁻⁶.

Insect PBPs, as a class of odorant binding proteins, are small soluble proteins rich in the lymph of pheromone-sensitive sensilla, in the Sensilla Trichodea of Lepidoptera, the concentration can reach 10–20 mM^{4,7}. Since their first identification by binding to radio-labelled pheromone in the giant silk moth *Antheraea polyphemus*, PBPs have been identified in various lepidopteran species^{4,8-10}.

Olfaction systems are of high sensitivity and specificity¹¹. It could be estimated, according to wind tunnel experiments, that about 200 molecules were sufficient enough to elicit a behavioral response; EAG experiments also showed that a single molecule of pheromone can be transduced to an electrical signal¹². Recent experimental evidence, both from behavioral and molecular biology, may support a more specific role of PBPs in detecting and recognizing semiochemicals¹³⁻¹⁵. Insects can correctly recognize pheromones from a mass of physiologically irrelevant chemical compounds, but even a minimal modification can result in the inactivity of pheromones^{16,17}. However, *in vitro* binding experiments of PBPs indicated that some lepidopteran PBPs like BmorPBP from *Bombyx mori*, ApolPBP from *Antheraea polyphemus* and LdisPBP from *Lymantria dispar* exhibited high affinity to more than one ligand, pheromones included^{18,19}. To some degree, this promiscuity could be explained, since sacrificing specificity for sensitivity is reasonable in the early events of olfaction.

The binding and release mechanism of PBPs is always an important issue. Lepidopteran PBPs were suggested to have two conformations in a pH/ligand-dependent manner. The A-form (closed or free form) characterized by the seventh α -helix formed by the C-terminus, was observed in the absence of pheromone and/or at low pH, whereas the B-form (open or bound form) characterized by the unstructured C-terminal tail, was detected in

Key Laboratory of Plant Protection Resources & Pest Management of the Ministry of Education, College of Plant Protection, Northwest A&F University, Yangling 712100, Shaanxi, China. *These authors contributed equally to this work. Correspondence and requests for materials should be addressed to Y.Z. (email: yalinzh@nwsuaf.edu.cn)

the presence of ligand at high pH^{20–22}. These results in combination with the lower pH generated by negatively charged membrane surfaces led to the promotion of a pH-induced ligand-releasing mechanism, following which the BmorPBP released the bound bombykol when encountering the lower pH around the receptors¹⁷. A similar phenomenon was also observed in ApolPBP²³. Additionally, Katre *et al.* pointed out that the C-terminus in ApolPBP1, forming the seventh α -helix, played an important role in ligand binding and/or locking the ligand in the binding pocket as well²¹.

Even though several protein classes, including odorant receptors (ORs), odorant binding proteins (OBPs), sensory neuron membrane proteins (SNMPs) and odorant degrading enzymes (ODEs), have been reported to be involved in insect odorant perception^{24–27}, the best studied OBPs were thought as appealing targets for structure-based prediction of physiologically active semiochemicals. Since the promotion of reverse chemical ecology, several attempts have been made either to screen odorants or to study ligand-binding mechanism on the basis of modeling OBP structure^{28–30}.

The codling moth *Cydia pomonella*, is a quarantine insect causing great harm to a variety of pome fruits every year in China. Today, pesticide applications are being called for diminishing. Therefore to control codling moth, the use of odorants for behavioral disruption is the most often alternative to pesticides³¹. However, very little is known about the chemoreception system of the codling moth. These days, studies on the codling moth PBPs mainly focused on searching new PBPs³², no reports on functional analysis and structure of CpomPBPs have been found.

In this article, the *Cydia pomonella* pheromone binding protein 2 (CpomPBP2) was cloned and expressed to study its biochemical characterizations. More importantly, in order to construct a rapid way for searching semiochemicals, virtual screening and an *in vitro* binding assay were used in combination with reference to the notion of “reverse chemical ecology”^{28,30,31}. The 3D structure of CpomPBP2 was initially modeled and used for molecular docking to screen semiochemicals. Competitive binding assay was applied to test the prediction accuracy of our computational method. Moreover, on the account of preparing for the application of high-throughput virtual screening, we also tried to find out key factors affecting affinity by analyzing the binding modes of the CpomPBP2-ligand complex.

Results and Discussion

CpomPBP2 protein characteristics and tissue distribution. To analyze the phylogenetic relationship of CpomPBP2 with other lepidopteran insects, we constructed a phylogenetic tree comprising 28 lepidopteran PBPs from different families (see Supplementary Fig. S1). CpomPBP2 was phylogenetically closest to GmolPBP2 from *Grapholitha molesta* and LglyPBP2 from *Leguminivora glycinivorella*, both of which belong to the Olethreutidae. This result was in line with the classification of *Cydia pomonella* (Tortricidae: Olethreutinae) in Lepidoptera.

In this research, to make CpomPBP2 and TPBP2 express in the supernatant, the expression systems PET28a(+)/CpomPBP2 and PET28a(+)/TPBP2 were transformed into the Rosetta-gami 2 (DE3) competent cells. Compared with strains like BL21 (DE3), Rosetta gami 2 promotes the expression of soluble protein. The expected bands were detected at about 20 KD and 18 KD (see Supplementary Fig. S2C), as the molecular weights of CpomPBP2 and TPBP2 were 16.5 KD and 14.5 KD, and the PET28a(+) tag was about 4 KD. By using a Ni²⁺-NTA agarose gel column (Transgen, China), the target recombinant protein was purified and analyzed on 15% SDS-PAGE gel. The presence of the additional N-terminal sequence containing the His tag slightly increased molecular weight and modified the isoelectric point of the recombinant CpomPBP2, but did not affect the functional property of ligand binding^{33,34}.

Reverse-transcription PCR (RT-PCR) was applied to check the distribution pattern of CpomPBP2 expression. The objective product was specifically detected in male and female antennae (see Supplementary Fig. S2A), and this result has been verified by western blot analysis of native CpomPBP2 (see Supplementary Fig. S2B). This is also in accordance with former reports in other lepidopteran insects^{9,10,35,36}.

3D structure modeling. As shown in Fig. 1, CpomPBP2 shared 50% sequence identity with the selected template (PDB ID: 1DQE). After 100 CpomPBP2 models calculation by MODELLER, the CpomPBP2 model with the lowest DOPE assessment score was selected. 3D quality of the best model was assessed by the Ramachandran Plot (see Supplementary Fig. S3). The result revealed that 99.3% (141/142) of all residues were in favored (98%) regions, 100.0% (142/142) of all residues were in allowed (>99.8%) regions, and there were no outliers. The best model of CpomPBP2 was also verified by Profile 3D³⁷. According to Supplementary Fig. S4, 88.89% of the residues had an average 3D-1D score ≥ 0.2 , and at least 80% of the amino acids have scored ≥ 0.2 in the 3D/1D profile. All of these parameters suggest that the 3D structure of CpomPBP2 is rational and can be used for further virtual screening.

The 3D structure of CpomPBP2, as shown in Fig. 1, was formed by a roughly conical arrangement of six α helices connected by loops. Four antiparallel helices ($\alpha 1$, $\alpha 4$, $\alpha 5$ and $\alpha 6$) converged to form a large flask-shaped binding pocket with the narrow end being open and the opposite end being capped by the $\alpha 3$ helix.

Furthermore, conformation of the best predicted model of CpomPBP2 was checked by intrinsic fluorescence analyses^{18,38}. The corresponding emission spectra of CpomPBP2 peaked at 337 nm when excited at 250 nm (see Supplementary Fig. S5). Former researches showed that the fluorescence maximum for free tryptophan in water and Trp residues in unstructured peptides was reached at 350 nm and shifted to shorter wavelengths in more hydrophobic environment¹⁸. In the present study, the detected peak at 337 nm was consistent with a mixture of hydrophobic and solvent exposed environments of the two Trp residues. This was in accordance with the predicted results when modeling the CpomPBP2 sequence on the crystal structure of BmorPBP-bombykol complex with one (Trp37) in a hydrophobic environment and the other (Trp127) being more solvent exposed.

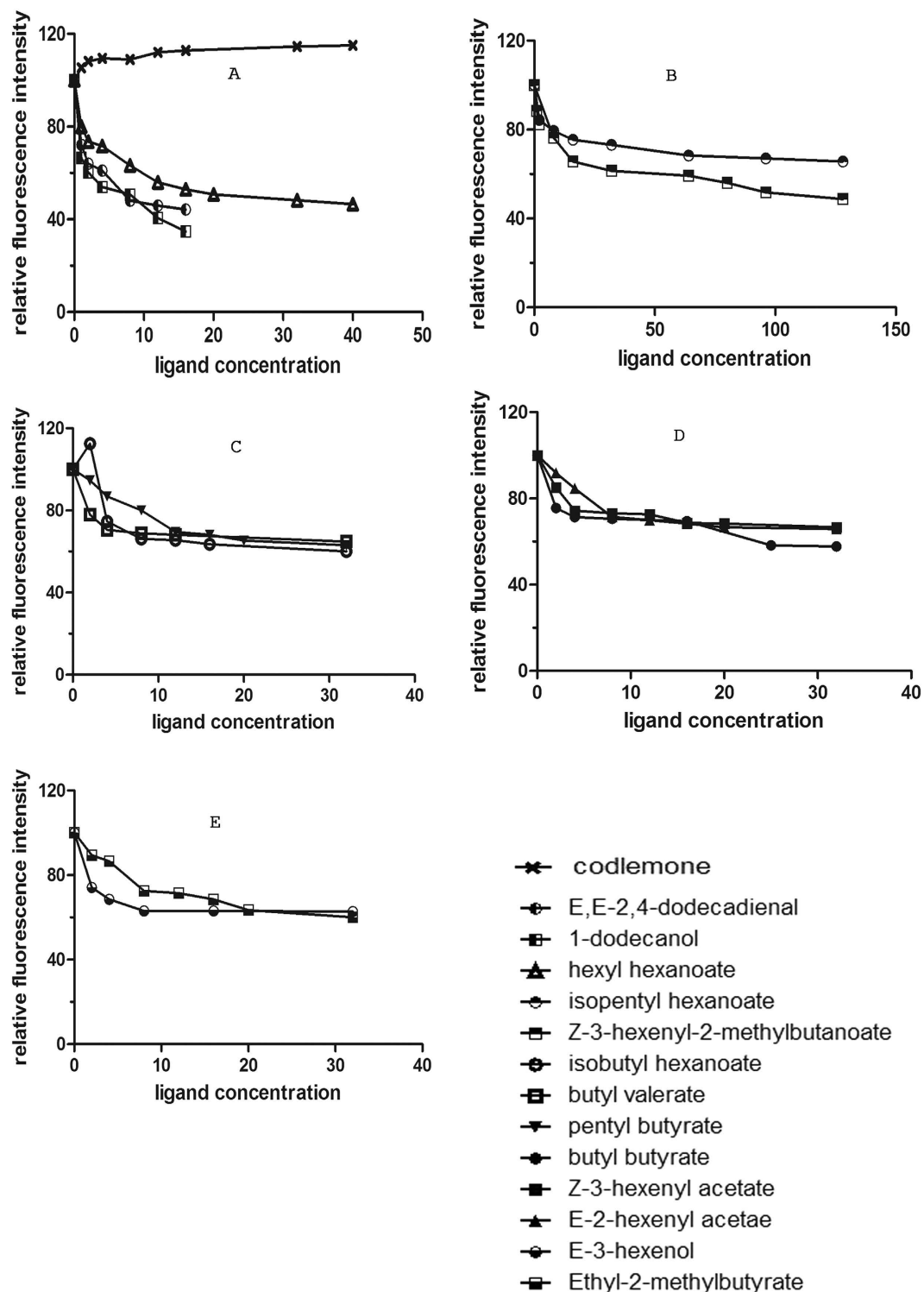


Figure 2. Competitive binding curves of some ligands to CpomPBP2. Solutions (50 mM Tris-HCl pH 6.8) containing CpomPBP2 and 1-NPN both at 2 μM, were titrated with increasing amounts of competing ligands. (A) Ligands applied consisted of 12C-Skeleton odorants, (B) 10C-Skeleton odorants, (C) 9C-Skeleton odorants, (D) 8C-Skeleton odorants and (E) 6C-Skeleton odorants.

The dissociation constants above exhibited poor discrimination of CpomPBP2 to molecules, CpomPBP2 is never an individual example in this aspect^{19,44}. This promiscuity could be regarded as a form of olfaction sensitivity. After all, in the early events of olfaction, sacrificing certain specificity for sensitivity is reasonable and necessary¹⁷. Our results also provide evidence for the concept of a two-step filter in which specific detection of pheromone is mediated by both PBPs and pheromone receptors^{44–46}. Some reports demonstrated that the specific conformation of PBP-ligand complex also has a role in pheromone discrimination^{18,44,47}. In the case of *Drosophila*

Name	Structure	ΔG (KJ/mol)	S(hbond)	S(lipo)	H(rot)	K_D (μ M)
1-Dodecanol		-28.57	1.27	207.74	2.13	4.14 ± 0.29
E,E-2,4-Dodecadienal		-27.46	0.95	192.24	1.44	6.99 ± 0.72
Hexyl Hexanoate		-26.23	0	211.40	1.56	16.20 ± 0.84
Z-3-Hexenyl-2-Methylbutanoate		-24.74	0	196.37	1.45	97.03 ± 2.26

Table 1. Predicted Gibson free energy change (ΔG), individual energy terms and dissociation constants (K_D) for ligands binding to CpomPBP2.

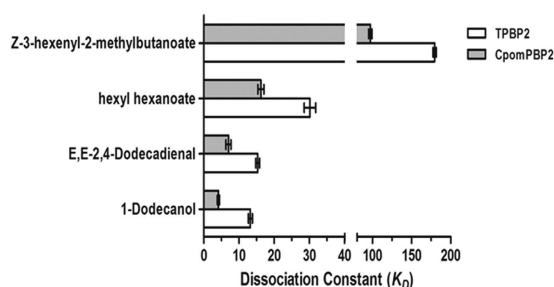


Figure 3. Effects of C-terminus on the binding of ligands to CpomPBP2. Deletion of the C-terminus of CpomPBP2 largely decreased the affinity between ligands and protein.

LUSH, only the specific pheromone can trigger the appropriate conformational change, allowing it to be recognized by the specific receptor, even though LUSH was able to bind different nonspecific ligands with similar affinity to that of the specific one¹⁴.

Binding mode and binding free energy analyses. In order to further investigate the characteristics of binding at the structural level, four active compounds (Table 1) were selected to perform an analysis for binding mode and binding free energy. As shown in Fig. 4A,C, Hexyl Hexanoate and Z-3-Hexenyl-2-Methylbutanoate possessed similar binding modes, and were bound in the large hydrophobic pocket whose component residues including Leu8, Phe12, Trp37, Ile52, Leu61, Leu68, Phe76, Leu90, Val91, Ile94, Phe117 and Leu134 were involved in binding to these two compounds (see Fig. 4B,D). Both of these compounds formed polar interactions with residues Thr9 and His74 as well. By comparison, we found that Hexyl Hexanoate was provided with a larger hydrophobic contact area than Z-3-Hexenyl-2-Methylbutanoate, especially at the site of Leu90 and Phe117 (see Fig. 4A,B). This phenomenon corresponded to the individual energy terms derived from binding free energy analysis shown in Table 1, the compound Hexyl Hexanoate possessed a more favorable hydrophobic interaction energy contribution [S(lipo)] than Z-3-Hexenyl-2-Methylbutanoate, leading to more binding free energy change and higher affinity (lower K_D) of CpomPBP2-Hexyl Hexanoate complex.

Different from the two compounds above, 1-Dodecanol and E,E-2,4-Dodecadienal formed hydrogen bonds and charged interactions within the binding pocket of CpomPBP2 as shown in Fig. 4E,G, even though similar polar interactions with Thr9 and Thr110 were also found in CpomPBP2-Dodecanol and CpomPBP2-Dodecadienal complexes. For E,E-2,4-Dodecadienal, a hydrogen bond was established between the oxygen atom derived from the carbonyl group of the compound and the NH atom from the side chain of Arg109 with an O-N distance of 3.0 Å (see Fig. 4G,H). Meanwhile, charged interaction between Arg109 and E,E-2,4-Dodecadienal was also detected (see Fig. 4G,H). In the CpomPBP2-Dodecanol complex, the hydroxyl group of 1-Dodecanol formed a hydrogen bond with the side chain of Arg109 with the distance 2.9 Å (see Fig. 4E,F). What quite different from the complex formed by CpomPBP2 and E,E-2,4-Dodecadienal was that a second hydrogen bond was established between the hydroxyl group and the main chain of Ala66 whose distance was 1.6 Å (see Fig. 4E,F). Moreover, the hydroxyl group of 1-Dodecanol formed positive and negative charged interactions with Lys67 and Glu98. These were in perfect agreement with the binding free energy changes and individual energy terms for 1-Dodecanol and E,E-2,4-Dodecadienal listed in Table 1. The CpomPBP2-Dodecanol complex exhibited over 1 KJ/mol more binding free energy change than CpomPBP2-Dodecadienal complex. The individual energy contributions in Table 1 showed that, corresponding to the relatively stronger hydrogen bond interactions in 1-Dodecanol, the hydrogen bond energy item contribution [S(hbond)] of 1-Dodecanol (1.27) was much greater than that of E,E-2,4-Dodecadienal (0.95). Interestingly, by comparing the chemical scaffold between 1-Dodecanol and E,E-2,4-Dodecadienal, we found that the saturated aliphatic chain group derived from 1-Dodecanol (207.74) was provided with a larger hydrophobic energy item contribution [S(lipo)] than the olefin group in E,E-2,4-Dodecadienal (192.24).

All of these findings suggested that both hydrogen bond energy item and hydrophobic interaction energy item had something to do with the binding affinity of these four compounds to the active pocket of CpomPBP2. As shown in Table 1 and Supplementary Table S1, It seemed that compound owning long aliphatic chain group

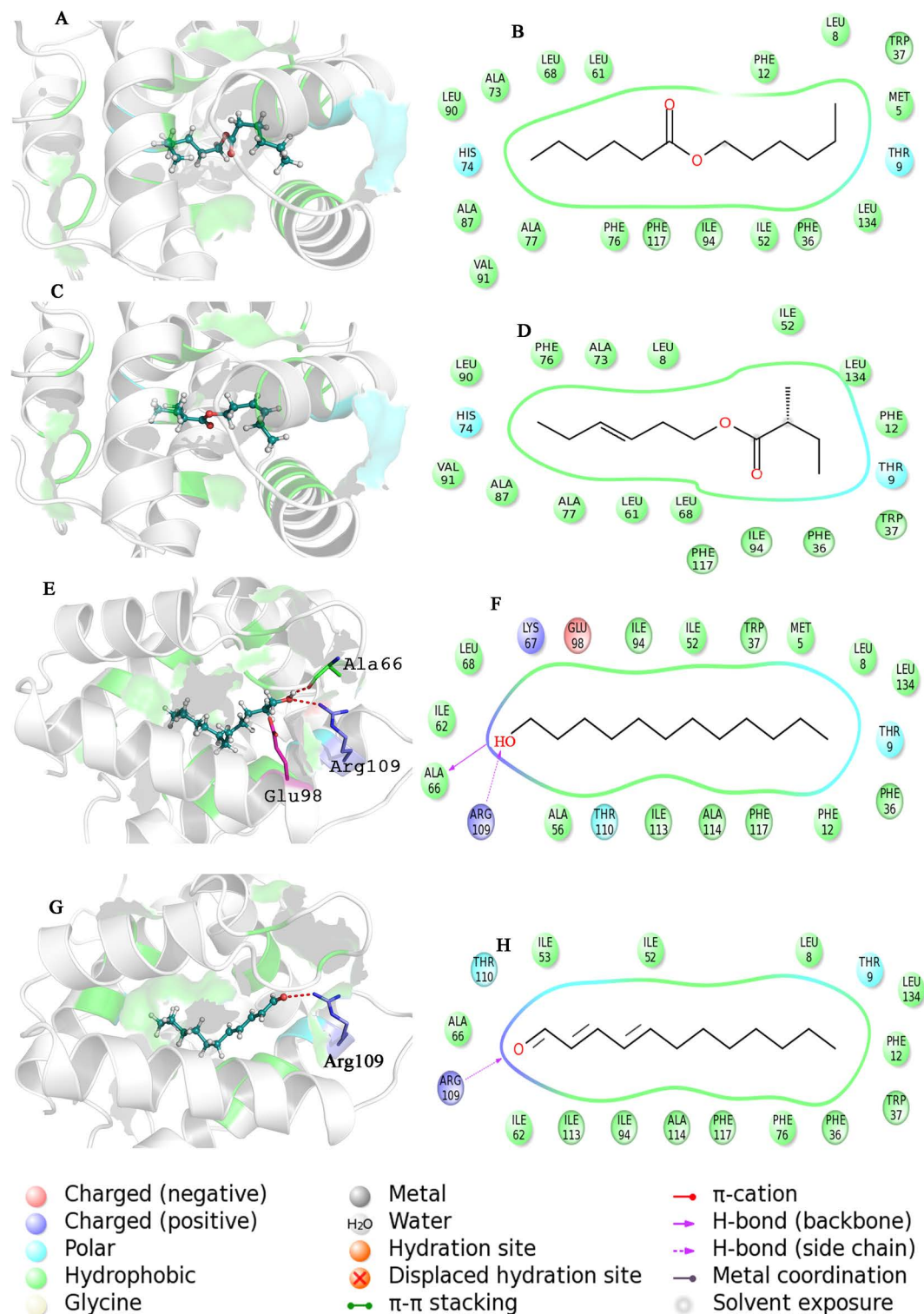


Figure 4. The cartoon representation of CpomPBP2 3D model around the binding site of all odorants. The binding model and interaction diagram of CpomPBP2 with Hexyl hexanoate (A,B), with Z-3-hexenyl-2-methylbutanoate (C,D), with 1-Dodecanol (E,F) and with E,E-2,4-dodecadienal (G,H). The four odorants are presented as stick and sphere model. Cyan C, White H, Red O. the red dotted lines show hydrogen bonds among the atoms from amino acid residues and odorants.

was provided with remarkable hydrophobic interaction for the binding. It was assumed that shorter chain may result in smaller surface area of interaction and poorer binding ability²⁹. So compounds of saturated 12C-skeleton tended to exhibit higher affinity to CpomPBP2 than shorter/unsaturated-chain compounds.

Computational alanine scanning mutagenesis. To verify the key residues identified in the simulation docking procedure, the CpomPBP2/1-Dodecanol complex with the highest affinity was subjected to the computational alanine scanning (CAS) mutagenesis⁴⁸. According to the per-residue energy decomposition, four key residues Phe12, Glu98, Arg109 and Ile113 (Figure S8), whose side chains contributed more than 1 kcal/mol to the energy change, were chosen for mutation to alanine based on the 10ns molecular dynamics trajectories.

The CAS results of the key residues were shown in table S2. The mutation of Arg109 to Ala caused the largest change in binding free energy ($\Delta\Delta G_{\text{bind}} = 6.25$ kcal/mol). The binding free energy also dropped dramatically when the residue Glu98 was mutated to alanine ($\Delta\Delta G_{\text{bind}} = 5.34$ kcal/mol). The $\Delta\Delta G_{\text{bind}}$ lower than 4 kcal/mol was observed on mutating the residues Phe12 ($\Delta\Delta G_{\text{bind}} = 3.17$ kcal/mol) and Ile113 ($\Delta\Delta G_{\text{bind}} = 2.66$ kcal/mol).

By reference to the definition of hot-spots and warm-spots^{48,49}, two residues (Arg109 and Glu98) met requirements of hot-spots, whereas the other two fitted with warm-spots. The CAS method achieved an overall success rate of 80% and an 82% success rate in residues whose alanine mutation caused an increase in the binding free energy > 2.0 kcal/mol (warm- and hot-spots)⁴⁸. Our CAS results could be regarded as verification to the results of simulation docking.

Conclusions

By taking virtual screening and *in vitro* binding assay together, we provided an efficient and viable method for screening active semiochemicals. Of the 13 ligands applied, results of both methods (*in vitro* binding assay and molecular docking) indicated that 1-Dodecanol was most favored by the binding pocket of CpomPBP2, with E,E-2,4-Dodecadienal, Hexyl Hexanoate and Z-3-Hexenyl-2-Methylbutanoate less favored. The agreement between the results of competitive binding assay and virtual screening suggested the high prediction accuracy of the computational method we applied.

Binding mode analyses revealed much information on the interactions between ligands and CpomPBP2. Key residues involved in interacting with these four ligands were pointed out. The computational alanine scanning mutagenesis also confirmed the confidence of key residues identified in binding mode analyses. For CpomPBP2-ligand complex, we found that hydroxyl group with higher S(hbond) contributed more to binding free energy change than carbonyl group, saturated aliphatic chain of 12-C skeleton owned larger hydrophobic interaction than unsaturated or shorter chain, this characteristic made ligands of long saturated chain tend to exhibit higher affinity to CpomPBP2 than shorter/unsaturated ones. Considering individual energy terms and the sequence of affinity in Table 1 in combination, it could be estimated that hydrogen bond and hydrophobic interaction were key factors in determining the binding affinity of ligands to CpomPBP2. This work is of guiding importance to potentially active semiochemicals screening and makes high-throughput virtual screening become available.

Methods

Sample collection and RNA extraction. Codling moths *C. pomonella* were reared at 27 °C, in 16h light : 8h dark on an artificial diet in the laboratory. The antennae of 3d-old male and female adults were excised at the base and immediately transferred into tubes immersed in liquid nitrogen. The preparation of all other samples including the egg, the 1st to 5th instar larvae, pupae, head (without antennae), thorax, abdomen, leg and wing were carried out in the same way. Prepared samples were stored at -80 °C. The total RNA of each sample was isolated according to the manual of RNAsiso Plus (TaKaRa, Japan). First-strand cDNA was synthesized by reference to the handbook of RevertAid First Strand cDNA Synthesis Kit (Thermo, USA) and employed as templates for latter PCR amplification.

Phylogenetic analysis. To conduct the phylogenetic analysis of CpomPBP2 (GenBank: JQ776635), 28 lepidopteran PBPs available in NCBI were downloaded and aligned using ClustalX, MbraPBP1 (*Mamestra brassicae*, AAC05702), MbraPBP2 (*M. brassicae*, AAC05701), BmorPBP (Bombyx mori, NM_001044029.1), SnonPBP2 (*Sesamia nonagrioides*, AAS49922), AipsPBP2 (*Agrotis ipsilon*, AAX85459), HvirPBP2 (*Heliothis virescens*, CAL48346), HassPBP3 (*H. assulta*, ABB91374), MsexPBP2 (*Manduca sexta*, AAF16710), AtraPBP2 (*Amyetois transitella*, ACX47892), LglyPBP2 (*Leguminivora glycinivorella*, AEO91540), GmolPBP2 (*Grapholitha molesta*, AHZ89398), AhetPBP2 (*Atrijuglans hetaohei*, AKA27976), PxylPBP2 (*Plutella xylostella*, AGH13203), DplePBP3 (*Danaus plexippus*, EHJ71308), SinfPBP2 (*Sesamia inferens*, AEX58642), SinfPBP3 (*Sesamia. inferens*, KF960746), SexigPBP2 (*Spodoptera exigua*, AAS55551), SexigPBP3 (*Spodoptera exigua*, ACY78413), CsupPBP2 (*Chilo suppressalis*, ACJ07123), ApolPBP2 (*Antheraea polyphemus*, AJ277266), AperPBP2 (*Antheraea pernyi*, X96860), SlituPBP2 (*Spodoptera litura*, ABK41048), XcniPBP (*Xestia cnigrum*, AGS41497), HzeaPBP (*Helicoverpa zea*, AF090191), HvirPBP (*Heliothis virescens*, AFI25170) and HarmPBP1 (*Helicoverpa armigera*, AEB54585). The phylogenetic tree was constructed by the MEGA4 program using the neighbor-joining method and bootstrapping sampled 1000 times.

Distribution pattern of CpomPBP2. Tissue distribution of CpomPBP2 was assessed by RT-PCR with cDNA templates from different tissues of female and male moths, cDNA templates from egg, the 1st to the 5th instar larvae and pupae were also subjected to RT-PCR analysis. Specific primers, PBP2F and PBP2R (see Table 2), were employed. For testing the integrity of the cDNA templates, a control primer pair (see Table 2) from the coding region of the *Cydia pomonella* β -actin gene (GenBank: KC832921) was used.

CpomPBP2 expression and purification. CpomPBP2 gene containing endonuclease restriction sites was cloned (EPIF and EPIR as primers) and incorporated into PET-28a(+), and the generated constructs were transformed into *E. coli* Rosetta gami 2 (DE3) strains. After 18hrs' induction with 0.6 mM IPTG under the condition of 16 °C, 160rpm, purification of the recombinant CpomPBP2 was initiated by centrifugation harvesting of 1L CpomPBP2-expressing DE3 cultures. Periplasmic fractions were prepared according to the osmotic shock

Purpose/Primer Name	Sequence (5'-3')
Tissue Distribution of CpomPBP2(RT-PCR)	
PBP2F	ATGGCGGCCGCCGCGAAATGG
PBP2R	CTACGACGGCTTGACTTCAGT
ActinF	TCCGGCATGTGCAAGGCCGGT
ActinR	GTCCCAGTTGTGACGATGCC
CpomPBP2 Expression	
EP2F	CGGGATCCATGTCGGCGGAGATTATGAAAA
EP2R	CCAAGCTTCTACGACGGCTTGACTTCAGTCA
TPBP2 (C-terminus Truncated CpomPBP2) Expression	
TP2F	CGGGATCCATGTCGGCGGAGATTATG
TP2R	CGAAGCTTCTACGCCCACTTGAGCCC

Table 2. Oligonucleotide Primers used for cloning and expression of pheromone binding protein 2 (CpomPBP2) from *Cydia pomonella*.

procedure in the manual (Novagen). These obtained periplasmic fractions were centrifuged at 12000 g for 30 min, the supernatants were loaded onto Ni²⁺-NTA sepharose gel columns (Transgen, China) and eluted following the manufacturer's instructions. Purified CpomPBP2 was analyzed by SDS-PAGE and dialyzed against 10 mM PBS, pH7.4.

C-terminal tail truncated form of CpomPBP2 (TPBP2) was also expressed by removing the last 17 amino acids PSMETVLEEVMTEVKPS. A pair of primers, TP2F and TP2R (see Table 2), were applied to obtain the expression-ready gene. The following steps were the same as mentioned above.

Western blot analysis. Polyclonal antibodies against CpomPBP2 were produced by injecting rabbits. To test the expression pattern of CpomPBP2, Total protein extracted from tissues including antenna, head, thorax, abdomen, leg and wing of 3d-old moths by homogenizing in 20 mM Tris-HCl (pH7.4). Extracted proteins of the same quantity were separated by 15% SDS-PAGE, and the proteins were transferred to NC membranes by a semi-dry transfer cell (Biorad, USA). After being blocked for 3h with 5% skimmed milk in TBST (0.05% Tween-20 in TBS), the NC membrane was incubated for 2h with the polyclonal antibody against CpomPBP2 at a dilution of 1:8000, several washes with TBST followed. Thereafter, the second antibody, goat anti-rabbit IgG conjugated with HRP (Jackson, USA), was applied at a dilution of 1:10000, the immunoreaction bands were detected using an ECM kit (Booster, China).

Structure modeling. The program Modeller 9.10 was applied to construct the 3D molecular model of CpomPBP2⁵⁰. After searching for the PDB95 database with the amino acid sequence of CpomPBP2 being a probe, the crystal structure of the BmorPBP-bombykol complex from *Bombyx mori* (PDB ID: 1DQE, Chain A, resolution 1.8 Å) was selected as a template on the basis of the crystallographic R-factor (21.8%), the sequential identity (50%) and the pH state (pH7.0). The 3D model of CpomPBP2 was generated and refined using the automodel and loopmodel modules in the Modeller program respectively. The obtained model was also subjected to the GA341 and discrete optimized energy (DOPE) scores to measure the relative stability of CpomPBP2 conformation. The credible structure of CpomPBP2 with the lowest DOPE energy was selected and its quality was assessed by using MolProbity to identify the rationality of the stereochemistry for the structure⁵¹.

Intrinsic fluorescence detection. Intrinsic fluorescence analysis was taken to verify the reliability of the simulated 3D model of CpomPBP2¹⁸. Fluorometric analyses were performed using a Hitachi F-4500 spectrofluorimeter in a 1 cm light path fluorimeter quartz cuvette. To measure the intrinsic tryptophan fluorescence, an excitation wavelength of 250 nm was chosen and the emission scans were recorded from 280 to 400 nm with a slit width of 5 nm. Spectra were recorded with 2 μM CpomPBP2 in 50 mM Tris-HCl of different pH value at 25 °C^{16,20,33}.

Virtual screening based on molecular docking simulations. To obtain the suitable binding poses of CpomPBP2 with the 31 compounds (see Supplementary Table S1) composed of codling moth pheromones and host volatiles reported hitherto^{52,53}, and to rank their binding affinity, virtual screening based on molecular docking was investigated by the program GOLD 5.3⁵⁴. To find out the best set of docking parameters and to ensure reliability of docking results, the bombykol derived from the crystal structure of the PBP-bombykol complex was first docked into the binding site of the CpomPBP2 3D model. The C25 atom coordinates of the bombykol were defined as the centroid of the binding site with 10 Å radius sphere. 3D structures of all compounds were sketched using Maestro version (Schrodinger Inc.) and optimized 2000 steps in Amber12 with the GAFF force field⁵⁵. After being performed 5000 steps minimization in Amber12 with the ff99SB force field⁵⁶, the 3D model of CpomPBP2 was selected as a receptor for docking simulations. For binding pose prediction, ChemPLP was suggested to be superior to other scoring functions in the GOLD program³⁹. Hence the ChemPLP score was employed to obtain the most accurate binding modes for the 31 candidate compounds. It should be noted that all docking simulations

were done at pH7.0. Interaction diagram was presented by the program Maestro version 10.1 and visualization of the structures was performed by PyMol 1.3r1 edu⁵⁷.

Competitive binding assay. On the basis of virtual screening and commercial availability, 13 typical ligands were purchased from Aladdin (China), Sigma-Aldrich (USA) and TCI (Japan) to be subjected to competitive binding assay, the major sex pheromone component codlemone (E8, E10-12:OH) was also employed as a control. Fluorescence emission spectra were recorded on a Hitachi F-4500 spectrofluorimeter in a 1 cm light path fluorimeter quartz cuvette. To get the dissociation constant (K_D) between 1-NPN and CpomPBP2, 2 μ M CpomPBP2 dissolved in 50 mM Tris-HCl pH6.8/pH5.0 was added with 1 mM 1-NPN to final concentrations of 0.5–20 μ M. Purchased ligands were then added with increasing concentrations in solutions containing 2 μ M CpomPBP2 and 2 μ M 1-NPN. The excitation wavelength was decided to be 337 nm and the maximum emission wavelength at 410 nm was recorded. All ligands including 1-NPN were dissolved in GC grade methanol (Aladdin, China).

Ligands screened above were employed to test their affinity to TPBP2 by reference to competitive binding assay. Solutions containing 2 μ M TPBP2 were titrated with 1-NPN to final concentrations of 1–32 μ M and competitive binding under pH6.8 and pH5.0 was performed following above procedures described.

Dissociation constants (K_D) calculation. It was assumed that CpomPBP2 was 100% active and the binding was 1 : 1 protein : ligand at saturation. In this article, Graphpad Prism software (Graphpad Software, Inc.) was applied to analyze the obtained data. For determination of dissociation constants, the intensity value corresponding to the maximum fluorescence emission was plotted against the concentration of 1-NPN, and bound ligands were evaluated from the values for fluorescence intensity. The curve was linearized using Scatchard Plots to verify the confidence of K_D . Dissociation constants of the competitors were calculated from the corresponding IC_{50} value by reference to the following equation (1):

$$K_D = [IC_{50}]/(1 + [1 - NPN]/K_{1-NPN}) \quad (1)$$

In this equation, $[IC_{50}]$ stands for the ligand concentration where the ligand quenching the fluorescence intensity of 1-NPN to 50%, $[1-NPN]$ and K_{1-NPN} mean the free concentration of 1-NPN and the K_D of the CpomPBP2/1-NPN complex respectively.

Binding energy calculations. To compare the binding affinity of these semiochemicals to CpomPBP2 3D model, the top docking pose for each analog corresponding to the ChemPLP score was rescored by the Chemscore to measure affinity data by ranking according to the chemscore delta value^{58,59}. Chemscore estimates the binding free energy ΔG according to equation (2):

$$\Delta G_{binding} = \Delta G_O + \Delta G_{hbond} + \Delta G_{lipo} + \Delta G_{rot} \quad (2)$$

The Chemscore function in our work can be written in the form:

$$\Delta G = - 5.4800 + - 3.3400 * S (hbond) + - 0.1170 * S (lipo) + 2.5600 * H (rot) \quad (3)$$

Each component of this equation is the product of a term dependent on the magnitude of a particular physical contribution to free energy (e.g. hydrogen bonding).

Computational alanine scanning mutagenesis. Based on the competitive binding assay, the complex formed by CpomPBP2 and the ligand with the highest affinity was subjected to the computational alanine scanning (CAS) method⁴⁸ to verify the reliability of the key amino acid residues identified by the simulation docking. The CAS was performed according to our former reports^{39,60}.

References

- Kaissling, K. E., Strausfeld, C. Z. & Rumbo, E. R. Adaption processes in insect olfactory receptors: mechanisms and behavioral significance. *Ann N Y Acad Sci* **510**, 104–112 (1987).
- Breer, H., Raming, K. & Krieger, J. Signal recognition and transduction in olfactory neurons. *BBA-Mol Cell Res* **1224**, 277–287, doi: 10.1016/0167-4889(94)90201-1 (1994).
- Wojtasek, H. & Hansson, B. S. Attracted or repelled? - a matter of two neurons, one pheromone binding protein, and a chiral center. *Biochem Biophys Res Commun* **250**, 217–222 (1998).
- Vogt, R. G. & Riddiford, L. M. Pheromone binding and inactivation by moth antennae. *Nature* **293**, 161–163 (1981).
- Horst, R. *et al.* NMR structure reveals intramolecular regulation mechanism for pheromone binding and release. *PNAS* **98**, 14374–14379, doi: 10.1073/pnas.251532998 (2001).
- Leal, W. S. In *Chemistry of Pheromones and Other Semiochemicals II Vol. 240 Topics in Current Chemistry* (ed S. Schulz) 1–36 (2005).
- Klein, U. Sensillum-lymph proteins from antennal olfactory hairs of the moth *Anthearea polyphemus*. *Insect Biochem Mol Biol* **17**, 1193–1204 (1987).
- Abraham, D., Lofstedt, C. & Picimbon, J. F. Molecular characterization and evolution of pheromone binding protein genes in *Agrotis* moths. *Insect Biochem Mol Biol* **35**, 1100–1111, doi: 10.1016/j.ibmb.2005.05.002 (2005).
- de Santis, F. *et al.* Molecular cloning and *in situ* expression patterns of two new pheromone-binding proteins from the corn stemborer *Sesamia nonagrioides*. *J Chem Ecol* **32**, 1703–1717, doi: 10.1007/s10886-006-9103-2 (2006).
- Xiu, W. M., Zhou, Y. Z. & Dong, S. L. Molecular characterization and expression pattern of two pheromone-binding proteins from *Spodoptera litura* (Fabricius). *J Chem Ecol* **34**, 487–498, doi: 10.1007/s10886-008-9452-0 (2008).
- Krieger, J. & Breer, H. Olfactory reception in invertebrates. *Science* **286**, 720–723, doi: 10.1126/science.286.5440.720 (1999).
- Kaissling, K. E. A quantitative model of odor deactivation based on the redox shift of the PBP in moth antennae. *Ann N Y Acad Sci* **30**, 320–322 (1998).
- Xu, P. X., Atkinson, R., Jones, D. N. M. & Smith, D. P. *Drosophila* OBP LUSH is required for activity of pheromone-sensitive neurons. *Neuron* **45**, 193–200, doi: 10.1016/j.neuron.2004.12.031 (2005).

14. Laughlin, J. D., Ha, T. S., Jones, D. N. M. & Smith, D. P. Activation of pheromone-sensitive neurons is mediated by conformational activation of pheromone-binding protein. *Cell* **133**, 1255–1265, doi: 10.1016/j.cell.2008.04.046 (2008).
15. Forstner, M., Breer, H. & Krieger, J. 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 (2009).
16. Leal, W. S. *et al.* Kinetics and molecular properties of pheromone binding and release. *PNAS* **102**, 5386–5391, doi: 10.1073/pnas.0501447102 (2005).
17. Damberger, F. F., Michel, E., Ishida, Y., Leal, W. S. & Wüthrich, K. Pheromone discrimination by a pH-tuned polymorphism of the *Bombyx mori* pheromone-binding protein. *PNAS* **110**, 18680–18685, doi: 10.1073/pnas.1317706110 (2013).
18. Bette, S., Breer, H. & Krieger, J. Probing a pheromone binding protein of the silkworm *Antheraea polyphemus* by endogenous tryptophan fluorescence. *Insect Biochem Mol Biol* **32**, 241–246, doi: 10.1016/s0965-1748(01)00171-0 (2002).
19. Yu, Y. X. *et al.* Structural and functional difference of pheromone binding proteins in discriminating chemicals in the Gypsy moth, *Lymantria dispar*. *Int J Biol Sci* **8**, 979–991, doi: 10.7150/ijbs.4557 (2012).
20. Lautenschlager, C., Leal, W. S. & Clardy, J. Coil-to-helix transition and ligand release of *Bombyx mori* pheromone-binding protein. *Biochem Biophys Res Commun* **335**, 1044–1050, doi: 10.1016/j.bbrc.2005.07.176 (2005).
21. Katre, U. V., Mazumder, S. & Mohanty, S. Structural insights into the ligand binding and releasing mechanism of *Antheraea polyphemus* pheromone-binding protein 1: Role of the C-terminal tail. *Biochem* **52**, 1037–1044, doi: 10.1021/bi301393v (2013).
22. Katre, U. V., Mazumder, S., Prusti, R. K. & Mohanty, S. Ligand binding turns moth pheromone-binding protein into a pH sensor EFFECT ON THE ANTERAEA POLYPHEMUS PBP1 CONFORMATION. *J Biol Chem* **284**, 32167–32177, doi: 10.1074/jbc.M109.013383 (2009).
23. Damberger, F. F., Ishida, Y., Leal, W. S. & Wüthrich, 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, doi: 10.1016/j.jmb.2007.07.078 (2007).
24. Sato, K. *et al.* Insect olfactory receptors are heteromeric ligand-gated ion channels. *Nature* **452**, 1002–U1009, doi: 10.1038/nature06850 (2008).
25. Zhou, J. J. Odorant-binding proteins in insects. *Vitam Horm Pheromones* **83**, 241–272, doi: 10.1016/S0083-6729(10)83010-9 (2010).
26. Vogt, R. G., Riddiford, L. M. & Prestwich, G. D. Kinetic properties of a sex pheromone-degrading enzyme: The sensillar esterase of *Antheraea polyphemus*. *PNAS* **82**, 8827–8831 (1985).
27. Ishida, Y. & Leal, W. S. Cloning of putative odorant-degrading enzyme and integumental esterase cDNAs from the wild silkworm, *Antheraea polyphemus*. *Insect Biochem Mol Biol* **32**, 1775–1780 (2002).
28. Jayanthi, P. D. K. *et al.* Computational reverse chemical ecology: Virtual screening and predicting behaviorally active semiochemicals for *Bactrocera dorsalis*. *BMC Genomics* **15**, doi: 10.1186/1471-2164-15-209 (2014).
29. Li, D. Z. *et al.* Structure-based analysis of the ligand-binding mechanism for the DhelOBP21, a C-minus odorant binding protein, from *Dastarcus helophoroides* (Fairmaire; Coleoptera: Bothrideridae). *Int J Biol Sci* **11**, 1281–1295, doi: 10.7150/ijbs.12528 (2015).
30. Leal, W. S. *et al.* Reverse and conventional chemical ecology approaches for the development of oviposition attractants for *Culex* mosquitoes. *PLoS One* **3**, doi: 10.1371/journal.pone.0003045 (2008).
31. Witzgall, P., Stelinski, L., Gut, L. & Thomson, D. Codling moth management and chemical ecology. *Ann Rev Entomol* **53**, 503–522 (2008).
32. Garczynski, S. F. *et al.* Application of *Cydia pomonella* expressed sequence tags: Identification and expression of three general odorant binding proteins in codling moth. *Insect Sci* **20**, 559–574, doi: 10.1111/j.1744-7917.2012.01560.x (2013).
33. Wojtasek, H. & Leal, W. S. Conformational change in the pheromone-binding protein from *Bombyx mori* induced by pH and by interaction with membranes. *J Biol Chem* **274**, 30950–30956, doi: 10.1074/jbc.274.43.30950 (1999).
34. Campanacci, V. *et al.* Revisiting the specificity of *Mamestra brassicae* and *Antheraea polyphemus* pheromone-binding proteins with a fluorescence binding assay. *J Biol Chem* **276**, 20078–20084, doi: 10.1074/jbc.M100713200 (2001).
35. Maida, R., Krieger, J., Gebauer, T., Lange, U. & Ziegelberger, G. Three pheromone-binding proteins in olfactory sensilla of the two silkworm species *Antheraea polyphemus* and *Antheraea pernyi*. *Eur J Biochem* **267**, 2899–2908, doi: 10.1046/j.1432-1327.2000.01303.x (2000).
36. Newcomb, R. D., Sirey, T. M., Rassam, M. & Greenwood, D. R. Pheromone binding proteins of *Epiphyas postvittana* (Lepidoptera: Tortricidae) are encoded at a single locus. *Insect Biochem Mol Biol* **32**, 1543–1554, doi: 10.1016/s0965-1748(02)00075-9 (2002).
37. Bowie, J. U., Luthy, R. & Eisenberg, D. A method to identify protein sequences that fold into a known three-dimensional structure. *Science* **253**, 164–170 (1991).
38. Eftink, M. R. In *Fluorescence Spectroscopy Vol. 278 Methods in Enzymology* (eds L. Brand & M. L. Johnson) 221–257 (1997).
39. Yang, X. Q., Liu, J. Y., Li, X. C., Chen, M. H. & Zhang, Y. L. Key amino acid associated with acephate detoxification by *Cydia pomonella* carboxylesterase based on molecular dynamics with alanine scanning and site-directed mutagenesis. *J Chem Inf Model* **54**, 1356–1370 (2014).
40. Chen, X., Liu, J. & Zhang, Y. Cantharidin impedes the activity of protein serine-threonine phosphatase in *Plutella xylostella*. *Mol Biosyst* **10**, 240–250 (2014).
41. Gong, Y. *et al.* Ligand-interaction kinetics of the pheromone-binding protein from the Gypsy moth, *L. dispar*: Insights into the mechanism of binding and release. *Chem Biol* **16**, 162–172, doi: 10.1016/j.chembiol.2009.01.005 (2009).
42. Xu, W., Xu, X. Z., Leal, W. S. & Ames, J. B. Extrusion of the C-terminal helix in navel orangeworm moth pheromone-binding protein (AtraPBP1) controls pheromone binding. *Biochem Biophys Res Commun* **404**, 335–338, doi: 10.1016/j.bbrc.2010.11.119 (2011).
43. 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, doi: 10.1016/s1074-5521(00)00078-8 (2000).
44. Gräter, F., Xu, W., Leal, W. & Grubmüller, H. Pheromone discrimination by the pheromone-binding protein of *Bombyx mori*. *Structure* **14**, 1577–1586, doi: 10.1016/j.str.2006.08.013 (2006).
45. Prestwich, G. D., Du, G. & Laforest, S. How is pheromone specificity encoded in proteins. *Chem Senses* **20**, 461–469 (1995).
46. Steinbrecht, R. A., Laue, M., Maida, R. & Ziegelberger, G. Odorant-binding proteins and their role in the detection of plant odours. *Entomol Exp Appl* **80**, 15–18 (1996).
47. 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, doi: 10.1007/s10158-002-0018-5 (2002).
48. Moreira, I. S., Fernandes, P. A. & Ramos, M. J. Computational alanine scanning mutagenesis - An improved methodological approach. *J Comp Chem* **28**, 644–654, doi: 10.1002/jcc.20566 (2007).
49. Pons, J., Rajpal, A. & Kirsch, J. F. Energetic analysis of an antigen/antibody interface: Alanine scanning mutagenesis and double mutant cycles on the HyHEL-10/lysozyme interaction. *Protein Sci* **8**, 958–968 (1999).
50. Šali & Blundell, T. L. Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* **234**, 779 (1993).
51. Chen, V. B. & Arendall, W. B. Molprobit: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D* **66**, 12 (2010).
52. Ansebo, L. *et al.* Antennal and behavioural response of codling moth *Cydia pomonella* to plant volatiles. *JEN* **128**, 488–493, doi: 10.1111/j.1439-0418.2004.00878.488-493 (2004).
53. Witzgall, P. *et al.* Identification of further sex pheromone synergists in the codling moth, *Cydia pomonella*. *Entomol Exp Appl* **101**, 131–141 (2001).
54. Jones, G., Willett, P., Glen, R. C., Leach, A. R. & Taylor, R. Development and validation of a genetic algorithm for flexible docking. *J Mol Biol* **267**, 727–748 (1997).

55. Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A. & Case, D. A. Development and testing of a general amber force field. *J Comput Chem* **25**, 1157–1174 (2004).
56. Hummer, G., Rasaiah, J. C. & Noworyta, J. P. Water conduction through the hydrophobic channel of a carbon nanotube. *Nature* **414**, 188–190 (2001).
57. Delano, W. L. (2002). PyMOL molecular graphics system, education version 1.3.0.4, Schrödinger, LLC. New York, NY, USA.
58. Baxter, C. A., Murray, C. W., Clark, D. E., Westhead, D. R. & Eldridge, M. D. Flexible docking using tabu search and an empirical estimate of binding affinity. *Proteins: Structure, Function, and Bioinformatics* **33**, 367–382 (1998).
59. Liu, J., Yang, X. & Zhang, Y. Characterization of a lambda-cyhalothrin metabolizing glutathione S-transferase CpGSTd1 from *Cydia pomonella* (L.). *Appl Microbiol Biotech* **98**, 8947–8962 (2014).
60. Liu, J. Y., Chen, X. E. & Zhang, Y. L. Insights into the key interactions between human protein phosphatase 5 and cantharidin using molecular dynamics and site-directed mutagenesis bioassays. *Sci Rep* **5**, doi: 10.1038/srep12359 (2015).

Acknowledgements

We are grateful to Dr. John Richard Schrock from Emporia State University (Emporia, KS, USA) for his critical reading of this manuscript. We also thank Dr. Zhengwei Wu for his help in insect rearing. This research was supported by the Special Fund for Agro-scientific Research in the Public Interest of China (No. 200903042-03).

Author Contributions

Y.L.Z., Z.T. and J.Y.L. conceived the project and designed the experiment. Z.T. and J.Y.L. performed the experiments and wrote the paper. Y.L.Z. supervised the study and contributed reagents/materials: All authors contributed to data analysis.

Additional Information

Supplementary information accompanies this paper at <http://www.nature.com/srep>

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Tian, Z. *et al.* Structural insights into *Cydia pomonella* pheromone binding protein 2 mediated prediction of potentially active semiochemicals. *Sci. Rep.* **6**, 22336; doi: 10.1038/srep22336 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>