Reverse chemical ecology: Olfactory proteins from the giant panda and their interactions with putative pheromones and bamboo volatiles

Jiao Zhu[a,1], Simona Arena[b,1], Silvia Spinelli[c,d], Dingzhen Liu[e], Guiquan Zhang[f], Rongping Wei[f], Christian Cambillau[c,d], Andrea Scaloni[b], Guirong Wang[a,2], and Paolo Pelosi[a,2]

[a]State Key Laboratory for Biology of Plant Diseases and Insect Pests, Institute of Plant Protection, Chinese Academy of Agricultural Sciences, Beijing 100193, China; [b]Proteomics & Mass Spectrometry Laboratory, Institute for the Animal Production System in the Mediterranean Environment-National Research Council, 80147 Napoli, Italy; [c]Architecture et Fonction des Macromolécules Biologiques (UMR 6098), CNRS, 13288 Marseille Cedex 09, France; [d]Architecture et Fonction des Macromolécules Biologiques (UMR 6098), Aix-Marseille University, 13288 Marseille Cedex 09, France; [e]Key Laboratory of Biodiversity Science and Ecological Engineering of Ministry of Education, School of Life Sciences, Beijing Normal University, Beijing 100875, China; and [f]Key Laboratory for Reproduction and Conservation Genetics of Endangered Wildlife of Sichuan Province, China Conservation and Research Center for the Giant Panda, Wolong, Sichuan 623006, China

Edited by Jerrold Meinwald, Cornell University, Ithaca, NY, and approved September 26, 2017 (received for review June 26, 2017)

The giant panda *Ailuropoda melanoleuca* belongs to the family of Ursidae; however, it is not carnivorous, feeding almost exclusively on bamboo. Being equipped with a typical carnivorous digestive apparatus, the giant panda cannot get enough energy for an active life and spends most of its time digesting food or sleeping. Feeding and mating are both regulated by odors and pheromones; therefore, a better knowledge of olfaction at the molecular level can help in designing strategies for the conservation of this species. In this context, we have identified the odorant-binding protein (OBP) repertoire of the giant panda and mapped the protein expression in nasal mucus and saliva through proteomics. Four OBPs have been identified in nasal mucus, while the other two were not detected in the samples examined. In particular, AimelOBP3 is similar to a subset of OBPs reported as pheromone carriers in the urine of rodents, saliva of the boar, and seminal fluid of the rabbit. We expressed this protein, mapped its binding specificity, and determined its crystal structure. Structural data guided the design and preparation of three protein mutants bearing single-amino acid replacements in the ligand-binding pocket, for which the corresponding binding affinity spectra were measured. We also expressed AimelOBP5, which is markedly different from AimelOBP3 and complementary in its binding spectrum. By comparing our binding data with the structures of bamboo volatiles and those of typical mammalian pheromones, we formulate hypotheses on which may be the most relevant semiochemicals for the giant panda.

odorant-binding proteins | chemical communication | X-ray structure | proteomics | giant panda

The giant panda *Ailuropoda melanoleuca* is endemic of China and was formerly classified as an endangered species, now as a vulnerable species, but its population has remained rather stable, although very low, during the last centuries (1). Its phylogenetic classification has been a matter of debate for some time, but molecular genetic studies have recently shown that this species belongs to Ursidae, of which it represents an ancestral branch together with the spectacled bears, *Tremarctos*, and the sloth bear (1–3). The diets of these species are different from those of carnivorous bears: the giant panda is fully herbivorous, the spectacled bears are mainly herbivorous, and sloth bears feed on termites, fruits, and vegetables. The giant panda also shares with spectacled bears and sloth bears the absence of hibernation, an important characteristic that differentiates these species from other Ursidae (4).

The obligate bamboo diet of the giant panda, which is not compatible with its carnivorous digestive system, is barely sufficient to provide the energy required for an active life, likely accounting for the slow movements and long periods of rest typical of this species (5). It has been also suggested that the reduced size of the brain, liver, and kidneys of the giant panda relative to other mammals could be a measure to further reduce the use of its limited energies (6). The vulnerability of the giant panda as a species is increased by their limited reproduction rate, usually with only a single offspring every other year.

Both feeding habits and reproductive activity strongly rely on chemical signals. Therefore, a study of olfaction and chemical communication in the giant panda can shed light on the molecular mechanisms that are responsible for the unique and anomalous diet of this species compared with other Ursidae. At the same time, understanding the molecular mechanisms of chemical communication mediating courtship and mating could explain the low reproduction rate and may suggest strategies to increase the survival rate of the species. In this context, we have focused our study on odorant-binding proteins (OBPs), a class of soluble proteins involved in olfaction as carriers of hydrophobic odorants and pheromones. Vertebrate OBPs (7–11) belong to

Significance

The giant panda, an endangered species and a popular emblem, still conceals puzzling unexplored aspects. It shares with bears, to which it is evolutionary related, a carnivorous digestive system but follows a strictly herbivorous diet. The low energy obtained from such poor food accounts for its slow movements and probably, a reduced reproductive activity. Feeding and mating are regulated by olfaction, still poorly investigated in this species at the molecular level. Here, we describe two odorant-binding proteins with complementary affinities to different chemical classes and present the 3D structure of one of them. In a reverse chemical ecology approach, which could be adopted for other vertebrates, we use ligand-binding data to suggest putative structures of still unknown sex pheromones.

Author contributions: D.L., G.W., and P.P. designed research; J.Z., S.A., and S.S. performed research; G.Z., R.W., and G.W. contributed new reagents/analytic tools; J.Z., S.A., S.S., D.L., C.C., A.S., G.W., and P.P. analyzed data; D.L. provided biological samples; and C.C., A.S., and P.P. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The 3D structure of AimelOBP3 reported in this paper has been deposited in the Protein Data Bank, www.wwpdb.org (PDB ID code 5NGH).

See Commentary on page 12094.

[1]J.Z. and S.A. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: grwang@ippcaas.cn or ppelosi.obp@libero.it.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1711437114/-/DCSupplemental.

a superfamily of carrier proteins named lipocalins (12), which includes serum retinol-binding protein, responsible for delivering retinol in the whole body (13); milk β-lactoglobulin, having a still uncertain function; fatty acid-binding protein; and other proteins involved in organism development and differentiation (14). Vertebrate OBPs share with lipocalins a compact structure made of eight antiparallel β-sheets and a short segment of α-helix close to the protein C terminus (15, 16). Several pieces of evidence strongly suggest that, in mammals, OBPs are specific carriers for pheromones (17). The most compelling facts are their sites of production outside the olfactory mucosa in the vomeronasal organ and the nasal respiratory epithelium and the occurrence of the same or very similar OBPs in the nose and in biological glands and fluids releasing specific pheromones. The best examples of this fact are the major urinary proteins (MUPs) of rodents (18, 19) and the salivary lipocalins (SALs) of the pig (20–22). In both cases, the same proteins are produced in the nose as well as in the liver (in rodents) or in the salivary glands (in the pig). Indeed, it has been shown that, when secreted outside the nose, OBPs are loaded with specific pheromones (20, 23), clearly suggesting a common function in releasing these chemical messengers in the environment.

The genome of the giant panda has recently been sequenced (24, 25), but its annotation is not complete. Thus, preliminary information is available on OBPs and other lipocalins, but no experimental work on such proteins has been reported. Animal behavior studies and chemical analysis of specific secretions have indicated urine and the perianal gland secretions as the biological fluids responsible for carrying semiochemicals. Courtship and mating as well as competition between males are likely mediated by specific pheromones. Scent marks carried by urine and perianal secretions are utilized by both sexes to advertise their presence and status (26). In female urine, short-chain fatty acids seem to be predominant (27, 28), while male perianal glands contain medium and long linear aldehydes as well as a number of long-chain fatty acids together with a variety of other chemicals (29, 30). At present, it is not clear which volatiles might be the best putative semiochemicals.

This work provides a contribution to the study of chemical communication in the giant panda through a structural and functional characterization of its OBPs. In our study, we followed a reverse chemical ecology approach to suggest likely structures for the still unknown sex pheromone through the study of structural and functional characteristics of their binding proteins. In particular, of the six OBPs present in the databases, we identified four in the nose of the animal; the two most abundant ones were produced in recombinant form. We then obtained the crystallographic structure of one of them and the structural model of the other one, and we investigated the ligand specificity and mode of binding of both OBPs, also using selected mutant recombinant products. Finally, we formulate hypotheses on likely pheromone candidates based on the structures of the best ligands.

Results

Sequence Analysis. Starting from genome sequencing results on the giant panda, we analyzed genes present in the National Center for Biotechnology Information (NCBI) database when searching for animal lipocalin homologs. After discarding sequence data for redundant, very short, or very long entries, we obtained information for a total of 36 nonredundant lipocalins, which have sequences that are compared in the tree shown in Fig. S1. They belong to different subgroups, including retinoid-binding proteins, fatty acid-binding proteins, β-lactoglobulins, and other lipocalins. Six of these sequences were classified as OBPs based on comparison with their orthologs in other mammalian species. Additional sequence analysis and comparison with genomic data highlighted few mistakes regarding the identification of introns and ORFs, which were corrected. A sequence alignment of resulting OBPs is presented in Fig. 1. As

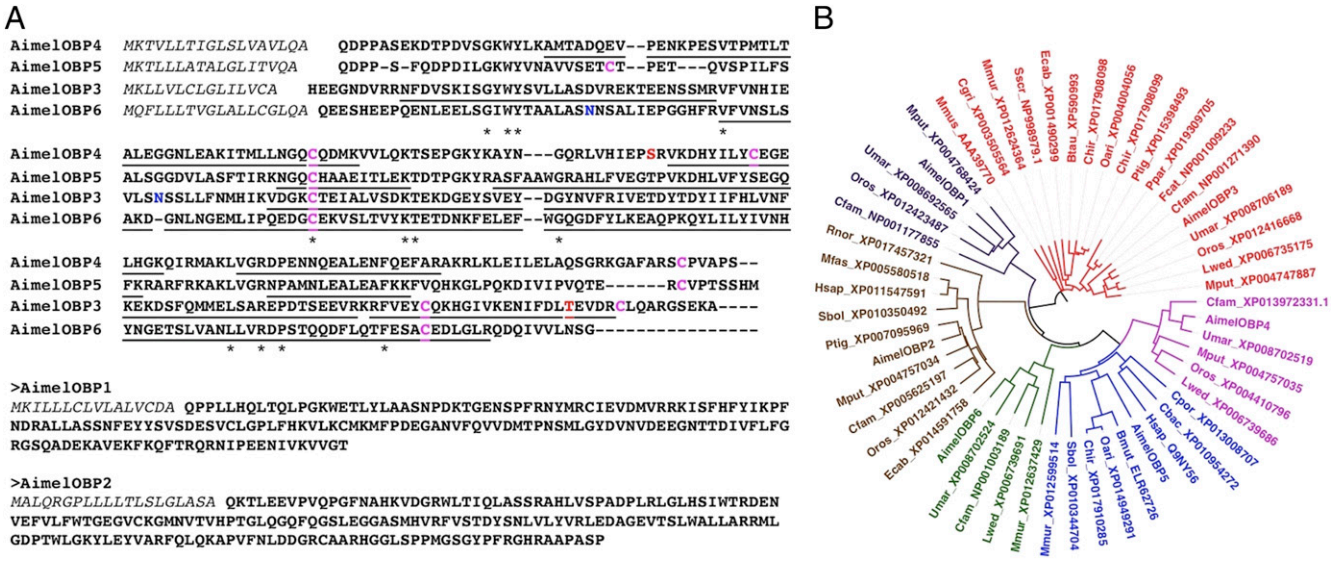


Fig. 1. (*A*) Alignment of the six OBPs identified in the currently available database. Crude sequences were compared with genomic sequences and edited for likely errors deriving from wrong identifications of introns. The alignment is shown only for AimelOBPs 1–4, which have sequences that represent a rather homogeneous group. AimelOBP1 and AimelOBP2, which are much more divergent, were not experimentally detected in our samples. Segments covered by our proteomic analysis are underlined. *N*-linked glycosylated sites are in blue, phosphorylated sites are in red, and cysteines are in magenta. (*B*) Phylogenetic tree constructed with the six AimelOBPs and their closest orthologs from other mammalian species. The proteins clearly segregate into six groups, each containing one OBP of the giant panda. Bmut, *Bos mutus*; Btau, *Bos taurus*; Cbac, *Camelus bactrianus*; Cfam, *Canis familiaris*; Cgri, *Cricetus griseus*; Chir, *Capra hircus*; Cpor, *Cavia porcellus*; Ecab, *Equus caballus*; Fcat, *Felis catus*; Hsap, *Homo sapiens*; Lwed, *Leptonychotes weddellii*; Mfas, *Macaca fascicularis*; Mmur, *Microcebus murinus*; Mmus, *Mus musculus*; Mput, *Mustela putorius*; Oari, *Ovis aries*; Oros, *Odobenus rosmarus*; Ppar, *Panthera pardus*; Ptig, *Panthera tigris*; Rnor, *Rattus norvegicus*; Sbol, *Saimiri boliviensis*; Sscr, *Sus scrofa*; Umar, *Ursus maritimus*.

expected, these proteins are divergent between each other (13–26% of identical residues, except for AimelOBP4 and AimelOBP5 sharing 40% of their amino acids) as well as with their orthologs from other species.

A sequence comparison of the giant panda OBPs with counterparts from other mammals showed some similarities, thus suggesting specific functions in chemical communication (Fig. 1). In particular, AimelOBP3 is about 54% identical with pig SALs, which are responsible for carrying the sex pheromones androstenone and androstenol in the saliva of the boar as well as for detecting them in the nose (20–22). They belong to a subgroup including the rodent MUPs (18, 31), the hamster aphrodisin (32), and rabbit seminal protein OBP3 (33), which are all involved in the release of specific pheromones (11). Instead, AimelOBP4 is more similar to Von Ebner's gland proteins, which are reported in tear and saliva of mammals (34, 35) and endowed with bacteriostastic function, other than putative but not experimentally shown roles in semiochemical transport (36, 37). AimelOBP5 is most similar to the human nose OBP1, with 43% identical amino acids (38, 39). Regarding AimelOBP1, -2, and -6, we could not identify orthologs in other mammalian species, as identities at the amino acid level barely exceed 20%, with the exception of AimelOBP1 and pigOBP1 (34% identity). In general, we observe that bear OBPs present the best sequence matches, in agreement with the currently accepted assignment of the giant panda to the family of Ursidae.

Proteomic Analysis of Body Fluids. To ascertain the occurrence of OBPs in giant panda biological fluids associated with chemical communication, we analyzed samples of nasal mucus and saliva. Total proteins were resolved by SDS/PAGE, excised from the gel, subjected to trypsinolysis, and analyzed for their digests by nanoLC-ESI-Q-Orbitrap-MS/MS. Several OBPs were detected in both secretions.

The complete results of proteomic analysis of *A. melanoleuca* crude nasal mucus and saliva are reported in Datasets S1 and S2, while Fig. 2 summarizes the OBPs detected in these biological fluids and their migration areas within SDS/PAGE. We managed to map large sequence regions in AimelOBP3, AimelOBP4, AimelOBP5, and AimelOBP6, while we did not find traces of AimelOBP1 or of AimelOBP2 (Figs. 1 and 2). All proteins detected in nasal mucus were also found in the saliva, probably as the result of a biological fluid exchange between nasal cavity and mouth. This was confirmed by Western blot analysis for AimelOBP3, which is present at high concentration in the nasal mucus, but only in traces in two samples of saliva (Fig. S2*A*).

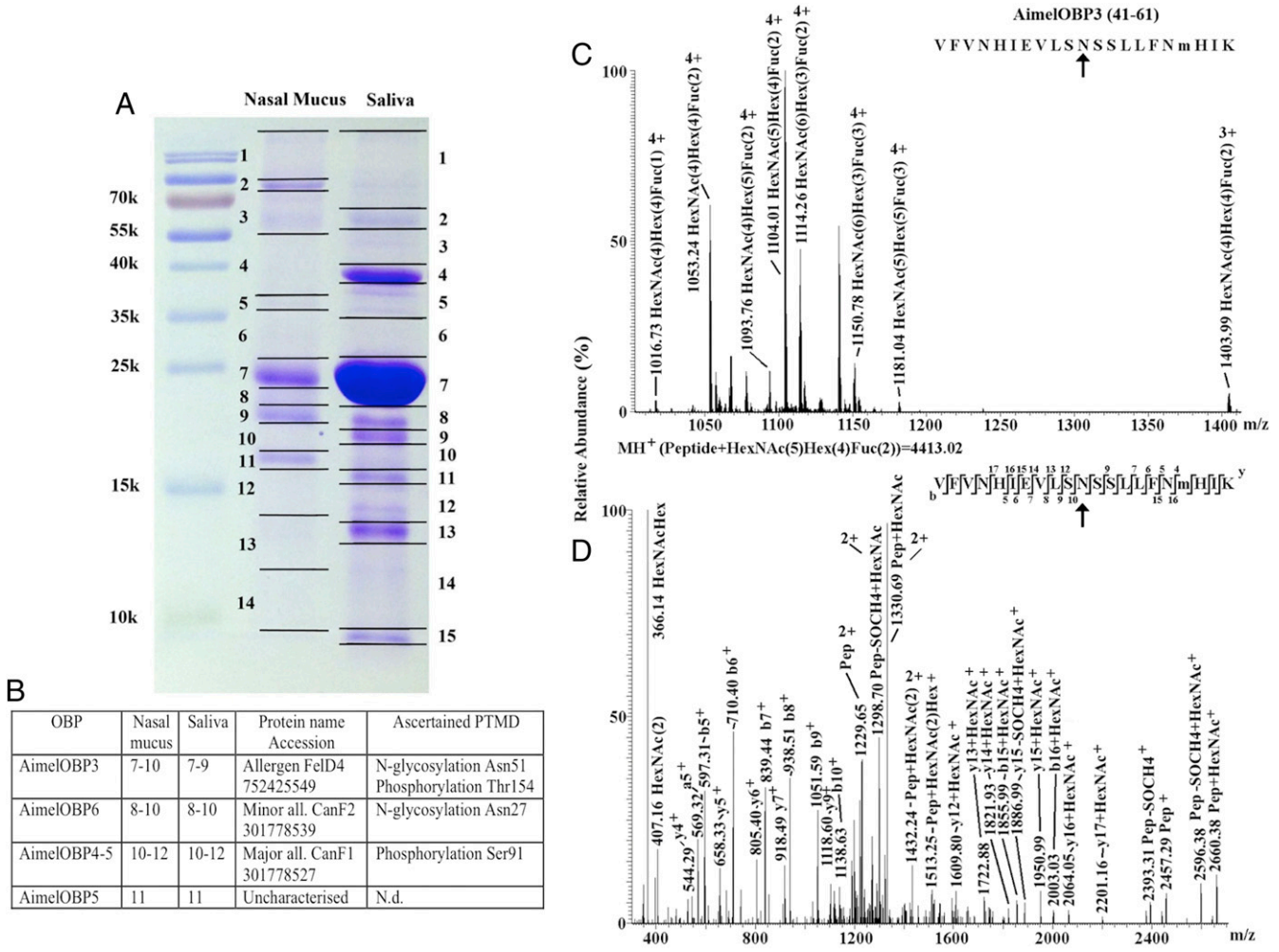
Proteomics also revealed posttranslational modifications (PTMs) present in each protein species. In particular, sequence analysis using NetNGlyc 1.0 software predicted the presence of two *N*-linked glycosylation sites in AimelOBP3 (at Asn36 and Ans51 of the mature protein), the second of which was actually found to be modified with a complex-type *N*-linked glycan moiety. Fig. 2 shows the MS and MS/MS spectra of some coeluting glycopeptides detected in the AimelOBP3 digest. The complete list of the *N*-linked glycopeptides detected in AimelOBP3 is reported in Table S1. *N*-linked glycosylation of AimelOBP3 was also evidenced by digestion of this protein with *N*-glycosidase and analysis of the resulting products by SDS/PAGE and Western blot (Fig. S2*C*). No glycosylation was predicted for AimelOBP4, AimelOBP5, and AimelOBP6. Nevertheless, a modified peptide with an *N*-linked glycan chain was detected in AimelOBP6, which resulted in modification at Asn27. Glycosylation has been observed for other mammalian OBPs (20, 40). Their most likely function could be to increase the solubility of these proteins, present at high concentrations in body fluids.

Finally, NetPhos 3.1 software predicted several sites of potential phosphorylation in AimelOBPs. Proteomic analysis showed actual phosphorylation of AimelOBP3 and AimelOBP4 at Thr154 and Ser91, respectively. Fragmentation spectra of corresponding phosphopeptides are reported in Fig. S3. Nonphosphorylated peptide counterparts were also detected. Modified sites in AimelOBPs are indicated in Figs. 1 and 2. Phosphorylation of mammalian OBPs has been reported in the past (41) and suggested to be a way of modifying the binding specificity of the protein. However, the function of this modification on OBPs still remains to be experimentally shown.

Ligand Binding Studies on AimelOBP3 and AimelOBP5. Based on their abundance in the nasal mucus and on similarities with proteins of chemical communication in other mammals, we decided to functionally characterize AimelOBP3 and AimelOBP5 by using ligand-binding assays (33). We expressed AimelOBP3 and AimelOBP5 in a bacterial system using synthetic genes because of the difficulties in obtaining samples of fresh tissues from the giant panda. The recombinant proteins were purified by anion-exchange chromatography on DE-52 and Mono-Q columns and used for production of polyclonal antisera, X-ray crystallography, and ligand-binding experiments. We measured the protein affinity toward 40 natural compounds in competitive binding experiments by using *N*-phenyl-1-naphthylamine (1-NPN) as a fluorescent reporter. Results are summarized in Dataset S3 and Fig. 3, while all experimental data are reported in Figs. S4 and S5. Both proteins bind the fluorescent probe with high yields and good affinities (Fig. 3*A*). The selected potential ligands (Fig. 3*B*) belong to two classes of structurally unrelated compounds. The first is a collection of plant volatiles, several of which have been identified in bamboo leaves, the exclusive diet of the giant panda. The second group is a series of long-chain aldehydes, acids, and other derivatives, which might include putative semiochemicals. Being that the pheromones of the giant panda are still unknown, we have tested chemicals reported as semiochemicals for other mammals or for insects. AimelOBP3 showed good affinity to both natural terpenoids and long-chain unsaturated aldehydes, these latter being pheromone components for several Lepidoptera. A structurally related alcohol and an acetate as well as a number of linear fatty acids did not bind this protein. On the contrary, AimelOBP5 showed strong affinity to fatty acids in a size- and structure-dependent fashion (Fig. 3 *D* and *E*), while it exhibited weak or no binding to the aldehydes and to most plant volatiles (Fig. 3*B*). We can incidentally observe that the binding curves of some fatty acids (Fig. 3*D*) exhibit a peculiar behavior, decreasing at low concentrations of the ligands and increasing at concentrations higher than 4 μM. This phenomenon has been previously reported and attributed to formation of micelles when the concentration of the ligand is higher than its critical micelle concentration. Such micelles can encapsulate molecules of 1-NPN, thus enhancing the emitted fluorescence (42, 43).

The idea that insects and mammals might share structurally related or even identical chemicals as their pheromones is documented by several examples reported in the literature (44, 45). The simple reason behind this phenomenon is that Lepidopteran pheromones, most of them being unsaturated long-chain alcohols, aldehydes, or acetates, are synthesized from fatty acids, which are important components of the diet of insects as well as of mammals. Other than the well-known example of the elephant pheromone dodecenyl acetate (46), which is a pheromone component for several Lepidoptera, fatty acids have been reported to act as pheromones in sheep, cow, and buffalo (45) as well as in tiger, lion, and other felids (47).

Among the plant volatile compounds, citral, safranal, farnesol, β-ionone, and cedrol, all abundantly present in bamboo fresh leaves, exhibit optimal ligand properties. Particularly interesting is the high affinity measured with cedrol. In fact, this compound and its isomer epicedrol are highly represented in spring bamboo, while their levels in winter bamboo are strongly reduced (48, 49).



OBP	Nasal mucus	Saliva	Protein name Accession	Ascertained PTMD
AimelOBP3	7-10	7-9	Allergen FelD4 752425549	N-glycosylation Asn51 Phosphorylation Thr154
AimelOBP6	8-10	8-10	Minor all. CanF2 301778539	N-glycosylation Asn27
AimelOBP4-5	10-12	10-12	Major all. CanF1 301778527	Phosphorylation Ser91
AimelOBP5	11	11	Uncharacterised	N.d.

Fig. 2. Proteomic analysis of nasal mucus and saliva samples from *A. melanoleuca*. (*A*) SDS/PAGE of proteins from crude nasal mucus and saliva samples. Gel lanes were cut into discrete slices (as shown), which were further subjected to proteomic analysis for protein identification and PTMs assignment (Datasets S1 and S2). (*B*) Identified OBPs within each slice and corresponding PTMs. (*C*) MS analysis of some glycopeptides detected in the AimelOBP3 digest having a retention time of 46.75 min. (*C*) MS spectrum of *N*-linked glycated forms of the oxidized AilmeOBP3 peptide (41–61). Assignment to specific glycation structures was based on fragmentation data. (*D*) MS/MS spectrum of the ion at *m/z* 1,104.01 reported in *C*. Assigned peptide fragments are highlighted in the peptide sequence reported within the panel. The assigned *N*-linked glycan structure is also shown. The fragmentation spectrum also shows fragments caused by the characteristic loss of methanesulfenic acid from the side chain of oxidized Met derivative (−64 and −32 for singly and doubly protonated ions, respectively) (66). The latter derivative is reported as m; the peptide moiety is indicated with Pep. N.d., not determined; PTMD, posttranslational modification details.

Being that the native AimelOBP3 *N*-glycosylated and phosphorylated unlike the recombinant protein used in binding experiments, we decided to purify this protein directly from the animal nasal mucus and to measure its affinity toward a selection of the best ligands with the aim to verify whether PTMs might affect its binding specificity. The protein was obtained at a degree of purity satisfactory for our purpose (a single band visible on SDS/PAGE) by anion-exchange chromatography on Mono-Q and was further identified by Western blot analysis. Fig. S2 reports the results of the purification (Fig. S2*B*) and binding assays performed on fraction 8 of the chromatographic separation with a number of ligands (Fig. S2*D*). We observe that the native protein is not different in its binding properties from the recombinant OBP; thus, we can reasonably conclude that the glycan moiety of the native protein does not interfere substantially with binding. As for the role of AimelOBP3 phosphorylation, this issue remains an open question, since we were not able to evaluate the extent of this modification, having detected both phosphorylated and non-phosphorylated peptides that are known to present different ionization efficiencies.

3D Structure of AimelOBP3. The crystal structure of AimelOBP3 was solved by molecular replacement using the MUP4 [Protein Data Bank (PDB) ID code 3KFF) as a model and subsequently refined at 2.8-Å resolution (Table S2). The entire experiment was performed using a single crystal frozen at 100 K. The stereochemistry was analyzed with molprobity, which indicated that 94.4% of the residues are in the most favorable region and that 5.6% are in the additionally allowed region. The polypeptide chain is visible from residue 4 to residue 164. AimelOBP3 has a classical lipocalin fold (Fig. 4 *A* and *C*), with a β-barrel domain composed of nine β-strands (residues 17–124, 150–154) and an α-helix (130–143) flanking the β-barrel. The C-terminal segment (144–164) comprising the ninth β-strand of the barrel (residues 150–154) and an unstructured region (residues 155–164) follow the α-helix. A disulfide bridge (Cys66-Cys159) links β-strand 4 to the C-terminal segment.

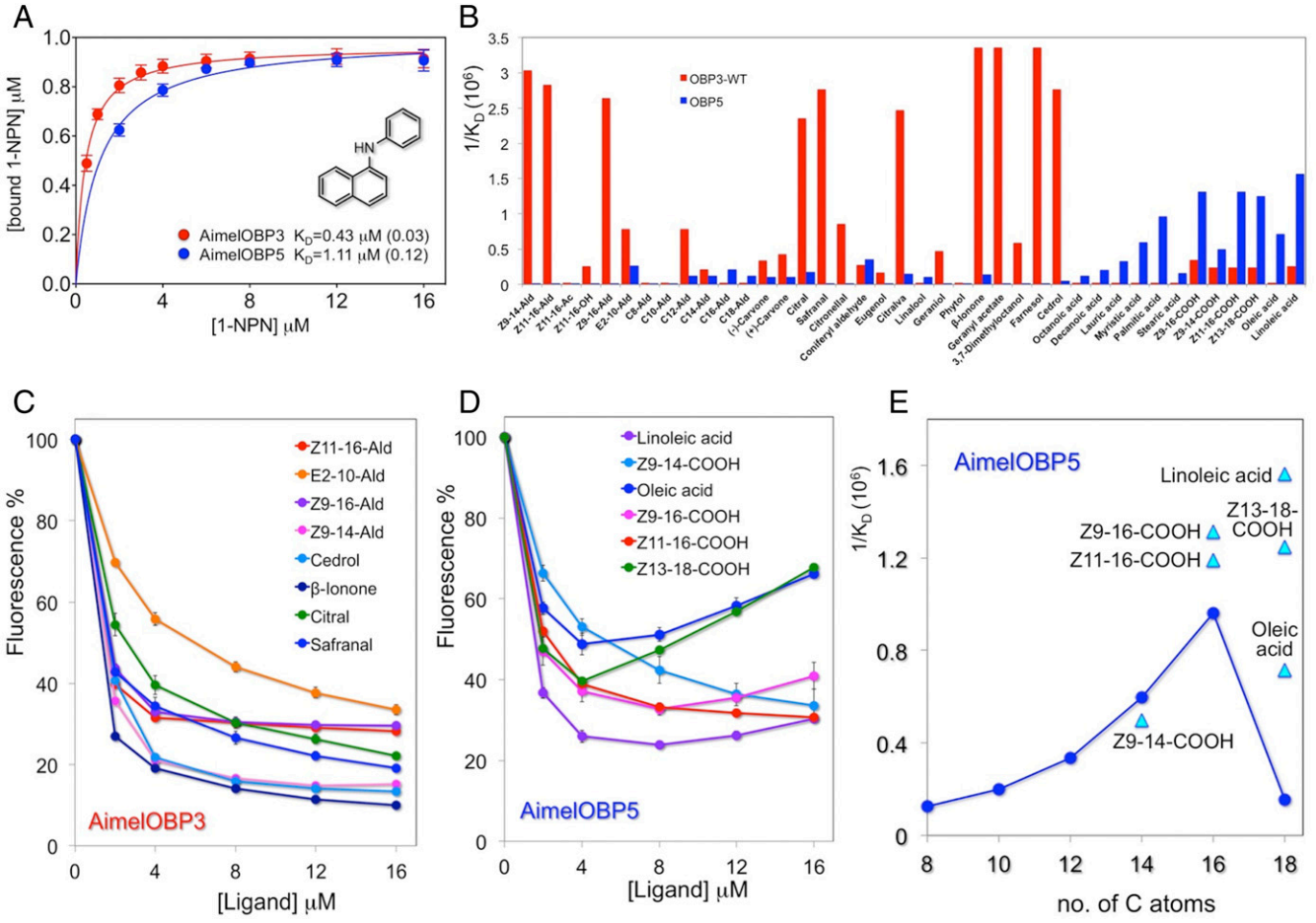


Fig. 3. Binding properties of AimelOBP3 and AimelOBP5. (*A*) Both proteins bind the fluorescent probe 1-NPN with good affinities. (*B*) Toward the 40 ligands tested, AimelOBP3 and AimelOBP5 exhibited markedly different and complementary spectra of binding. The first one is tuned to unsaturated long-chain aldehydes as well as to some bamboo leaves volatiles; the second one is rather specific for fatty acids. (*C* and *D*) Examples of competitive binding curves obtained with the two proteins using 1-NPN as fluorescence reporter. All binding curves are reported in Figs. S4 and S5. (*E*) The affinity of AimelOBP5 to fatty acids is length-dependent, with a peak at 16 carbon atoms. Unsaturated acids are better ligands than their saturated analogs.

The Buried Cavity and the Putative Channel. An electron density is visible in the AimelOBP3 cavity, indicating the presence of an unknown bound molecule. Its size and shape are compatible with trimethylamine *N*-oxide (TMAO) used for cryocooling. AimelOBP3 possesses an internal buried cavity with no access to solvent. The same feature was found in several other lipocalins, such as MUP (31), bovine and porcine OBPs (15, 16, 50), and human OBPIIa (51). Most of the residues that form the walls of the cavity are hydrophobic, with the exception of three polar amino acids (Asn90, Ser73, and Ser122) and two charged residues (Asp87 and Glu120) (Fig. 4*B*). In AimelOBP3, the volume of the cavity (392 Å^3) is in the middle of the range (300–500 Å^3) observed in other lipocalins. This cavity, however, is shielded from the solvent by only three residues: Asp87, Asn90, and Met39 (Fig. 4*B*).

Modeling and Ligand Binding of OBP5. Since AimelOBP5 exhibits quite different binding properties compared with AimelOBP3, we modeled its structure from human OBPIIa (PDB ID code 4RUN), which shares 66% of identical residues and 93% of similar residues, thus yielding a plausible model (Fig. 4*B*). Despite their different and complementary binding spectra, AimelOBP3 and AimelOBP5 exhibited very similar structures, as can be appreciated by superimposing the models on one another (Fig. 4*E*), with only a major structural difference: the segment 31–42 in AimelOBP3 is directed toward the protein interior, while the corresponding stretch is shorter in AimelOBP5 (residues 25–30) and follows a more external path. As a result, the cavity of AimelOBP3 is closed (Fig. 4*B*), while that of AimelOBP5 is open (Fig. 4*D*).

Design, Expression, and Ligand Binding of AimelOBP3 Mutants. Based on the crystallographic structure of AimelOBP3 and on docking simulations, we designed and prepared three mutants of this protein by replacing either Glu120 or Ser122 with Ala or otherwise, Asn90 with Leu. The recombinant proteins were purified by anion-exchange chromatography and used in binding experiments. All three mutants showed good affinity to 1-NPN, with dissociation constants similar to that of the WT (Fig. 5*A*). Competitive binding experiments were performed with the same set of ligands (linear aldehydes and plant volatiles) used for the WT protein, excluding the 12 fatty acids that were good ligands only for AimelOBP5. Each mutant showed different binding properties (Fig. 5*B*, Dataset S2, and Figs. S6–S8). Replacing Glu120 with Ala produced a major disruption in the binding properties of the protein. None of the good ligands of AimelOBP3-WT showed reasonable affinity for this mutant, suggesting that this amino acid substitution most likely affects the whole binding properties of this protein toward tested molecules, although the affinity to the fluorescent probe 1-NPN was barely modified.

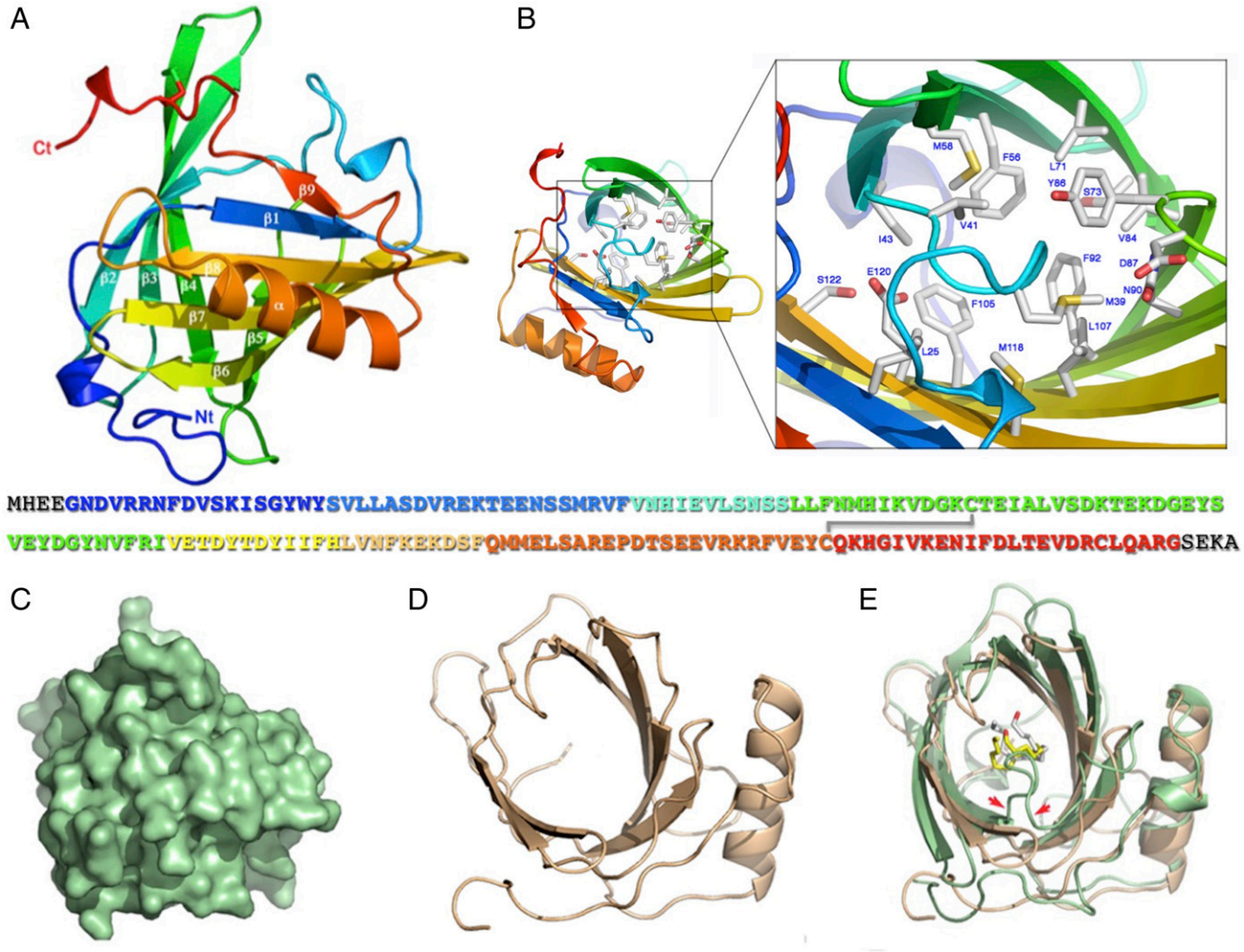


Fig. 4. Crystal structure of AimelOBP3. (*A*) Ribbon view is rainbow colored (blue to red), with numbering of the secondary structures. The sequence of the protein is provided with the same coloring mode. (*B*) AimelOBP3 internal cavity. Ribbon view of AimelOBP3 is rainbow colored. The side chains of the residues forming the cavity wall are shown and numbered in the enlarged view. (*C*) Surface representation of the AimelOBP3 structure. The binding cavity is covered, and the access to solvent is blocked by residues Asp87, Asn90, and Met39. (*D*) Model of AimelOBP5 based on the structure of human OBPIIa (PDB ID code 4RUN). (*E*) Superposition of AimelOBP3 (green) and AimelOBP5 (beige) in ribbon representation. The red arrows indicate the position of the loop 31–42 closing the binding cavity in AimelOBP3. The two proteins appear very similar in structure, despite their poor amino acid identity (17%) and their markedly different binding spectra.

When Ser122 was replaced with Ala, instead, we measured only limited effects; Z11-16:Ald became a weaker ligand, while the affinity of citral improved. Conversely, the binding properties of all other compounds were not appreciably modified. The third mutant (Asn90Leu) showed the most interesting behavior. Binding of linear aldehydes was strongly reduced, while affinities of terpenoids remained unchanged or slightly affected (Fig. 5*B*).

Ligand Binding Inside the Cavity. The strong binding of long-chain aldehydes to AimelOBP3 and the observation that their affinity was markedly and selectively reduced in Asn90Leu mutant prompted us to model the binding of two linear aldehydes in the cavity of the protein. Z11-16:Ald seemed to be an excellent ligand. From a structural viewpoint, it filled the binding pocket, establishing contacts with most of the cavity residues and accepting a hydrogen bond from Asn90 NH_2 moiety (Fig. 5 *C* and *D*). Z9-14:Ald also fit nicely inside AimelOBP3, and although not entirely filling the cavity, it also established the above-mentioned hydrogen bond with Asn90 NH_2 (Fig. S9). Whenever the aldehyde group in both derivatives was changed into the corresponding carboxylate and methyl ester counterparts, all compounds were still able to maintain the hydrogen bond reported above; however, they all fitted less properly within the binding pocket of AimelOBP3, as the additional atoms (hydroxyl or methoxy groups) clash with the cavity residues.

Discussion

The giant panda, with an obligate strict diet of bamboo and a carnivorous digestion system, lives on minimum energy at the brink of survival (3). Habitat fragmentation makes such situations worse, with lower food availability and high risk of inbreeding (52–55). Both diet and mating are mediated by chemical cues, and a better knowledge of how the giant panda chooses its food and finds its mate can suggest strategies to improve the life of these animals and protect the species.

In this report, we focus on OBPs, soluble proteins acting as carriers of pheromones to the olfactory and vomeronasal mucosa and releasers of pheromones in the environment. Their binding specificity, therefore, may pave the way for the identification of pheromones that are still unknown in the giant panda.

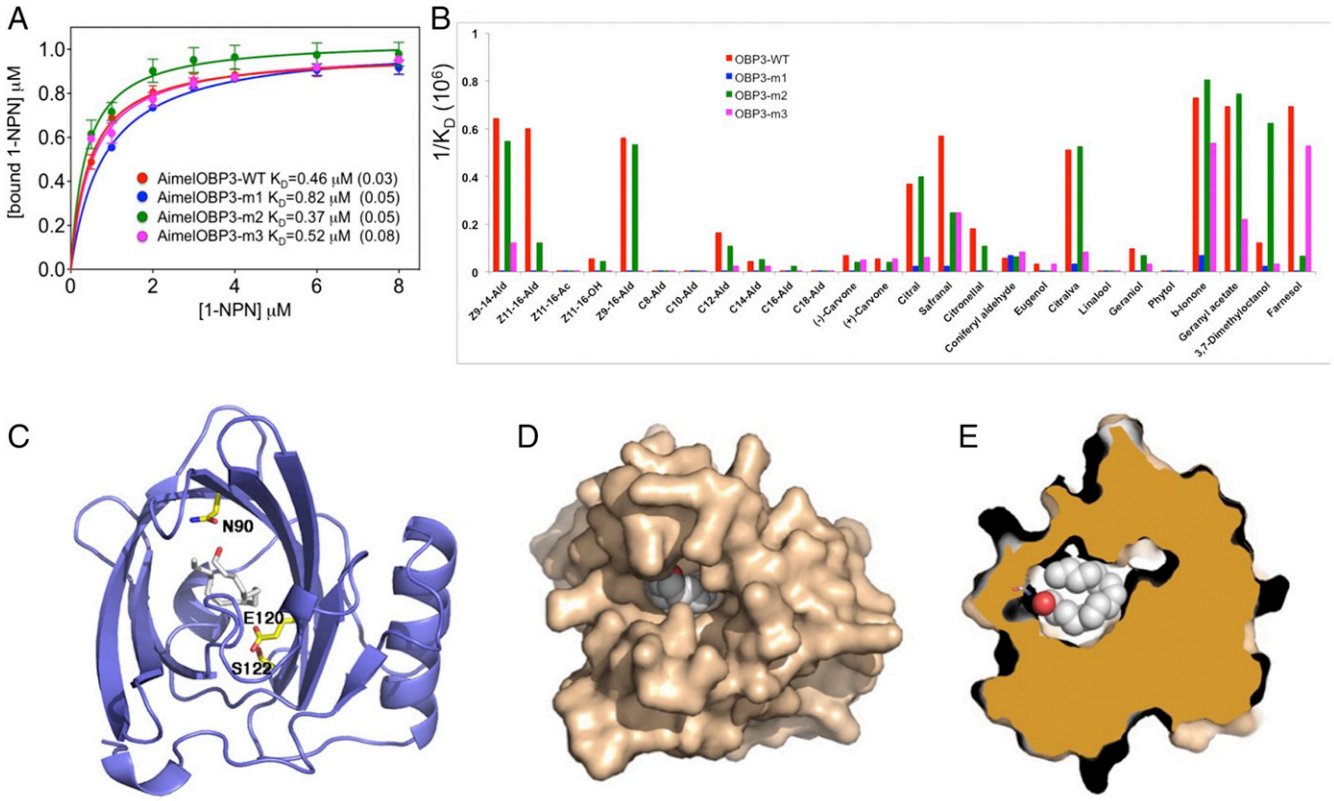


Fig. 5. Binding characteristics of AimelOBP3 mutants. (*A*) All three mutants prepared bind the fluorescent probe with good affinities, similar to that of the WT. (*B*) The three mutations affected the binding properties of AimelOBP3 in different ways: in Glu120Ala (m1), affinity to all ligands was strongly reduced, while in Ser122Ala (m2), the amino acid replacement had little effect on the performance of the protein. Mutant Asn90Leu (m3) showed the most interesting behavior, with selective reduction of the affinities to long-chain aldehydes. (*C*) Ribbon representation of the complex of AimelOBP3 with Z11-16:Ald. The protein has been slabbed at the level of the ligand. The three mutated residues (Asn90, Glu120, and Ser122) are shown in sticks inside the cavity. Red, O; white, C. (*D*) Space fill of the same complex. The side chains of Asp87, Asn90, and Met39 have been removed to show the cavity entry. (*E*) Slabbed view to show the ligand inside the cavity.

We found that AimelOBP3, highly abundant in the nasal mucus, is tuned to two classes of structurally unrelated compounds: plant volatiles and long-chain aldehydes. Among the plant volatiles showing best affinity to AimelOBP3 are several chemicals found in bamboo leaves, such citral, safranal, farnesol, β-ionone, and cedrol. This last compound is one of the best represented in bamboo leaves collected in the spring (when the mating season of the panda occurs), while its concentration drastically decreases in winter (48, 49). Linear aldehydes (ligands for AimelOBP3 as good as some plant volatiles) are common insect pheromones and might likely act as semiochemicals in the panda.

The other protein studied in this work, AimelOBP5, binds unsaturated fatty acids but not their corresponding aldehydes or the plant volatiles that instead represent the best ligands for AimelOBP3. Thus, the two proteins exhibit complementary spectra of binding.

We still do not know the structures of pheromone components in the panda. Analyses of perianal secretions, used by pandas to deposit scent marks, showed the presence of several fatty acids as well as some aldehydes (27, 29, 30). We can hypothesize on the basis of the best ligands found for our OBPs as well as on the information available for pheromone components in other mammals that both long-chain aldehydes and their corresponding carboxylic acids might represent suitable candidates.

In the case that long-chain unsaturated aldehydes prove to be the real pheromone components, then the presence of their corresponding carboxylic acids might be the result of spontaneous oxidation in the environment. If this is the case, we can venture and speculate that using AimelOBP3 and AimelOBP5 as two distinct olfactory channels to monitor aldehydes and fatty acids could provide the panda with a sort of clock to evaluate the age of the scent marks.

Our data provide some tools that might be useful for additional investigation of the chemical ecology of the giant panda and suggest putative structures for its pheromones. Moreover, the approach used in this work to search for pheromones by studying their binding proteins suggests a shortcut, which may have wider applications to other mammals and vertebrates. This method would prove particularly useful when dealing with species endangered or difficult to reach, for which it would be difficult to obtain enough biological samples or perform accurate behavioral observations.

Materials and Methods

Biological Material. Samples of nasal mucus and saliva were obtained at Yaan panda base and Dujiangyan base of the China Conservation and Research Center for the Giant Panda (CCRCGP) at Wolong, Sichuan, China. Collection of samples was performed during regular health examination and did not cause discomfort to the animals. Detailed information on panda management can be found elsewhere (56). In brief, animals were housed individually in pens consisting of an indoor house (6 × 4 m) and an outdoor yard with shrubs, climbing facilities, and a small pond, and they were fed on bamboo, shoots, panda bread (containing nutritional supplements), apples, and carrots. Sample collection was performed according to the regulations of CCRCGP and adhered to the Chinese Regulations and Standards for Captive Animals. All protocols for animal management were approved by the Institutional Animal Care and Use Committee of Beijing Normal University (CLS-EAW-2014-013).

Protein Extraction and Purification. Saliva and nasal mucus samples of *A. melanoleuca* were extracted with 50 mM Tris·HCl (pH 7.4) buffer and used

for SDS/PAGE and Western blotting experiments or purified through Mono-Q column (GE Healthcare Biosciences).

Proteomics. Gel slices were triturated, in-gel reduced, *S*-alkylated with iodacetamide, and digested with trypsin (57). Digest samples were desalted by μZip-TipC$_{18}$ using 50% (vol/vol) acetonitrile and 5% (vol/vol) formic acid as eluent. Resulting peptide mixtures were analyzed with a nanoLC-ESI-Q-Orbitrap MS/MS system consisting of an UltiMate 3000 HPLC RSLC nano system-Dionex coupled to a Q-Exactive^{Plus} mass spectrometer through a Nanoflex ion source (Thermo Fisher Scientific). Peptides were loaded on an Acclaim PepMap RSLC C18 column (150 mm × 75 μm i.d., 2-μm particles, 100-Å pore size; Thermo Fisher Scientific) and eluted with a gradient of solvent B [19.92/80/0.08 (vol/vol/vol) water/acetonitrile/formic acid] in solvent A [99.9/0.1 (vol/vol) water/formic acid] at a flow rate of 300 nL/min. The gradient of solvent B started at 3%, increased to 40% over 40 min, raised to 80% over 5 min, remained at this percentage for 4 min, and finally, returned to 3% in 1 min, at which it remained for an additional 20 min. The mass spectrometer operated in data-dependent mode using a full scan (*m/z* range 375–1,500, nominal resolution of 70,000) followed by MS/MS scans of the 10 most abundant ions. MS/MS spectra were acquired in a scan *m/z* range 200–2,000 using a normalized collision energy of 32%, an automatic gain control target of 100,000, a maximum ion target of 100 ms, and a resolution of 17,500. A dynamic exclusion value of 20 s was used.

Bioinformatics. MS and MS/MS raw data files were loaded into Proteome Discoverer v 2.1 software (Thermo Scientific) and searched with Mascot v 2.4.2 (Matrix Science) against a homemade *A. melanoleuca* protein database containing Uniprot and NCBI sequence entries (June 17, 2016). For PTMs discovery, we used Bionics 2.6.46 (Protein Metrics) and PEAKS Studio 8.0 (Bioinformatics Solutions) software. In all cases, we used the following search parameters: carbamidomethylation of Cys as a fixed modification and oxidation of Met, deamidation of Asn and Gln, pyroglutamate formation of Gln, phosphorylation of Ser/Thr/Tyr, and glycation of Asn with common mammalian *N*-linked glycans as variable modifications. Peptide mass tolerance was set to ±20 ppm, and the fragment mass tolerance was set to ±0.05 Da. Proteolytic enzyme and maximal number of missed cleavages were set to trypsin and three, respectively. Protein candidates assigned on the basis of at least two sequenced peptides with an individual peptide expectation value <0.05 (corresponding to a confidence level for peptide identification >95%) were considered confidently identified. Definitive peptide assignment was always associated with manual spectra visualization and verification. Results were filtered to 1% false discovery rate.

Plasmids and Reagents. Full-length genes encoding mature AimelOBP3 and AimelOBP5 were custom synthesized at Jinsirui Biotechnological Company. Plasmids were sequenced at Sheng Gong. All enzymes were from New England Biolabs. All other chemicals and reagents were purchased from Sigma-Aldrich and were of reagent grade.

Bacterial Expression. The custom synthesized cDNAs were amplified using specific primers bearing enzyme recognition sites (underlined) at both ends: AimelOBP3-Nde: AA<u>CATATG</u>CACGAGGAAGGTAACGAC; AimelOBP3-XhoI: AAA<u>CTCGAG</u>TTACGCTTTCTCGCTGCC; AimelOBP5-Nde: AA<u>CATATG</u>CAGGACCCGCCGAGCTT; AimelOBP5-Eco: AA<u>GAATTC</u>TTACATGTGGCTGCTGGT. After digestion, they were inserted into expression vector pET30a (Novagen). Protein expression was induced by isopropyl-β-D-thiogalactoside, and cells were grown for another 2 h. After sonication and centrifugation, recombinant proteins, which were mainly present as inclusion bodies, were dissolved in Tris buffer containing 8 M urea and 1 mM DTT and refolded by extensive dialysis against Tris buffer. Proteins were purified by anion-exchange chromatography on DE-52 (Whatman) followed by chromatography on Mono-Q (GE Healthcare Biosciences).

Synthesis of Mutants. Specific mutations were introduced into the gene encoding AimelOBP3 by PCR using the following primers: for AimelOBP3-m1: fw: ATGATGGCACTGAGCGC, rv: TTACGCTTTCTCGCTGCC; for AimelOBP3-m2: fw: GATGGAACTGGCCGCGCGTGA, rv: TTACGCTTTCTCGCTGCC; for AimelOBP3-m3: fw: T7, rv: AACACGAGATAGCCATCGTA. PCR conditions were (*i*) first step: 95 °C for 3 min; (*ii*) 35 cycles: 95 °C for 30 s, 60 °C for 30 s, and 72 °C for 1 min; and (*iii*) final step: 72 °C for 10 min. The amplified band was excised from the gel, extracted, and used for the second PCR using the following conditions: nine cycles at 95 °C for 30 s and 68 °C for 6 min and then, at 68 °C for 16 min. This second PCR product was digested with DpnI at 37 °C for 2 h and used to transform *Trans-T1 Escherichia coli* competent cells (Tiangen). Colonies containing the expected mutation were used for expression of recombinant proteins.

Fluorescence Binding Assays. Fluorescence was measured on a Horiba Scientific Fluoromax-4 spectrofluorometer using slits of 3, 4, or 5 nm according to the protein and a light path of 1 cm. The pure protein was dissolved with 50 mM Tris·HCl (pH 7.4) at a final concentration of 2 μM. The ligands, dissolved in methanol at the concentration of 1 mM, were added as aliquots to the protein solution. The fluorescent probe 1-NPN was excited at 337 nm, and emission was recorded between 380 and 450 nm. Competitive binding was measured by titration of a solution of both protein and 1-NPN at the concentration of 2 μM by adding aliquots of 1 mM methanol solution of ligand to final concentrations of 2–16 μM. Dissociation constants for 1-NPN were calculated using the software Graph Pad Prism. Dissociation constants of the competitors were calculated by the equation $K_d = IC_{50}/(1 + [1\text{-NPN}]/K_{1\text{-NPN}})$, where IC_{50} is the concentration of ligands halving the initial fluorescence value of 1-NPN, [1-NPN] is the free concentration of 1-NPN, and $K_{1\text{-NPN}}$ is the dissociation constant of the complex protein/1-NPN. Experiments were performed in triplicates, except for ligands showing not significant binding that were analyzed in single experiments.

Crystallization and X-Ray Diffraction of AimelOBP3. Diffraction-quality crystals of the AimelOBP3 were obtained by sitting-drop vapor diffusion at 277 K using a Mosquito robot (TTP Labtech) in 2.2 M ammonium sulfate and 0.2 M potassium nitrate at pH 7.0. Crystals were briefly soaked in TMAO before being flash-cooled in a nitrogen gas stream at 100 K. Crystals belong to the hexagonal space group $P6_522$ with unit cell dimensions a = b = 94 Å and c = 114.5 Å. Diffraction data were collected under standard cryogenic conditions on beamline Proxima 1 using a PILATUS 6M detector at the Soleil synchrotron (Saint Aubin, France); 1,000 images were collected with an oscillation step of 0.10° and 0.1-s exposure time. Data were integrated, scaled, and merged using the XDS package (58). The crystal structure of AimelOBP3 was determined from single-wavelength native diffraction experiments by molecular replacement using MOLREP (59), with the structure of major mouse urinary protein IV (3KFF) as the starting model. Refinement was performed with autoBUSTER (60) alternated with display modeling with COOT (61) (Table S2). Cavity analysis was performed with PISA (62). Figures were made with Pymol (63).

Modelization of AimelOBP3 Complexes. The ligand structures were constructed using the CCP4 tool Sketcher (64). They were fitted manually within the AimelOBP3 cavity using COOT (61) in the best position to avoid steric clashes and maximize favorable interactions. Geometry optimization was performed with REFMAC (65).

Modelization of OBP5. AimelOBP5 was modeled manually from the human OBPIIa structure (PDB ID code 4RUN) using COOT (61), and geometry optimization was performed with REFMAC (65).

ACKNOWLEDGMENTS. We thank the staff at the CCRCGP for their assistance in sample collection and animal management. We also thank the Soleil synchrotron for beam time allocation. This work was funded by Natural Science Foundation of China Grant 31472009 (to D.L.) and Opening Project Programme of State Key Laboratory for Biology of Plant Disease and Insect Pests Grant SKLOF201502 (to G.W.).

1. Hu YD, et al. (2016) Analysis of the cytochrome c oxidase subunit 1 (COX1) gene reveals the unique evolution of the giant panda. *Gene* 592:303–307.
2. Talbot SL, Shields GF (1996) A phylogeny of the bears (Ursidae) inferred from complete sequences of three mitochondrial genes. *Mol Phylogenet Evol* 5:567–575.
3. Wei F, et al. (2015) Progress in the ecology and conservation of giant pandas. *Conserv Biol* 29:1497–1507.
4. O'Brien SJ, Nash WG, Wildt DE, Bush ME, Benveniste RE (1985) A molecular solution to the riddle of the giant panda's phylogeny. *Nature* 317:140–144.
5. Nie Y, et al. (2015) ANIMAL PHYSIOLOGY. Exceptionally low daily energy expenditure in the bamboo-eating giant panda. *Science* 349:171–174.
6. Wang D (2015) Low daily energy expenditure enables giant pandas to survive on bamboo. *Sci China Life Sci* 58:925–926.
7. Pelosi P, Baldaccini NE, Pisanelli AM (1982) Identification of a specific olfactory receptor for 2-isobutyl-3-methoxypyrazine. *Biochem J* 201:245–248.
8. Pelosi P (1994) Odorant-binding proteins. *Crit Rev Biochem Mol Biol* 29:199–228.
9. Bignetti E, et al. (1985) Purification and characterisation of an odorant-binding protein from cow nasal tissue. *Eur J Biochem* 149:227–231.
10. Pevsner J, Trifiletti RR, Strittmatter SM, Snyder SH (1985) Isolation and characterization of an olfactory receptor protein for odorant pyrazines. *Proc Natl Acad Sci USA* 82:3050–3054.

11. Tegoni M, et al. (2000) Mammalian odorant binding proteins. *Biochim Biophys Acta* 1482:229–240.
12. Flower DR (1996) The lipocalin protein family: Structure and function. *Biochem J* 318: 1–14.
13. Monaco HL, Rizzi M, Coda A (1995) Structure of a complex of two plasma proteins: Transthyretin and retinol-binding protein. *Science* 268:1039–1041.
14. Flower DR (2000) Experimentally determined lipocalin structures. *Biochim Biophys Acta* 1482:46–56.
15. Bianchet MA, et al. (1996) The three-dimensional structure of bovine odorant binding protein and its mechanism of odor recognition. *Nat Struct Biol* 3:934–939.
16. Tegoni M, Ramoni R, Bignetti E, Spinelli S, Cambillau C (1996) Domain swapping creates a third putative combining site in bovine odorant binding protein dimer. *Nat Struct Biol* 3:863–867.
17. Pelosi P (2001) The role of perireceptor events in vertebrate olfaction. *Cell Mol Life Sci* 58:503–509.
18. Cavaggioni A, Mucignat-Caretta C (2000) Major urinary proteins, alpha(2U)-globulins and aphrodisin. *Biochim Biophys Acta* 1482:218–228.
19. Beynon RJ, Hurst JL (2004) Urinary proteins and the modulation of chemical scents in mice and rats. *Peptides* 25:1553–1563.
20. Marchese S, Pes D, Scaloni A, Carbone V, Pelosi P (1998) Lipocalins of boar salivary glands binding odours and pheromones. *Eur J Biochem* 252:563–568.
21. Loebel D, et al. (2000) Cloning, post-translational modifications, heterologous expression and ligand-binding of boar salivary lipocalin. *Biochem J* 350:369–379.
22. Spinelli S, Vincent F, Pelosi P, Tegoni M, Cambillau C (2002) Boar salivary lipocalin. Three-dimensional X-ray structure and androsterol/androstenone docking simulations. *Eur J Biochem* 269:2449–2456.
23. Bacchini A, Gaetani E, Cavaggioni A (1992) Pheromone binding proteins of the mouse, Mus musculus. *Experientia* 48:419–421.
24. Li R, et al. (2010) The sequence and de novo assembly of the giant panda genome. *Nature* 463:311–317.
25. Hu Y, et al. (2017) Comparative genomics reveals convergent evolution between the bamboo-eating giant and red pandas. *Proc Natl Acad Sci USA* 114:1081–1086.
26. Swaisgood RR, Lindburg DG, Zhou X (1999) Giant pandas discriminate individual differences in conspecific scent. *Anim Behav* 57:1045–1053.
27. Hagey L, MacDonald E (2003) Chemical cues identify gender and individuality in giant pandas (*Ailuropoda melanoleuca*). *J Chem Ecol* 29:1479–1488.
28. Dehnhard M, et al. (2006) Comparative endocrine investigations in three bear species based on urinary steroid metabolites and volatiles. *Theriogenology* 66:1755–1761.
29. Zhang J-X, et al. (2008) Potential chemosignals in the anogenital gland secretion of giant pandas, *Ailuropoda melanoleuca*, associated with sex and individual identity. *J Chem Ecol* 34:398–407.
30. Yuan H, et al. (2004) Anogenital gland secretions code for sex and age in the giant panda, Ailuropoda melanoleuca. *Can J Zool* 82:1596–1604.
31. Böcskei Z, et al. (1992) Pheromone binding to two rodent urinary proteins revealed by X-ray crystallography. *Nature* 360:186–188.
32. Mägert HJ, et al. (1995) cDNA sequence and expression pattern of the putative pheromone carrier aphrodisin. *Proc Natl Acad Sci USA* 92:2091–2095.
33. Mastrogiacomo R, et al. (2014) An odorant-binding protein is abundantly expressed in the nose and in the seminal fluid of the rabbit. *PLoS One* 9:e111932.
34. Glasgow BJ, Gasymov OK (2011) Focus on molecules: Tear lipocalin. *Exp Eye Res* 92: 242–243.
35. Redl B (2000) Human tear lipocalin. *Biochim Biophys Acta* 1482:241–248.
36. Schmale H, Holtgreve-Grez H, Christiansen H (1990) Possible role for salivary gland protein in taste reception indicated by homology to lipophilic-ligand carrier proteins. *Nature* 343:366–369.
37. Scalfari F, et al. (1998) Expression of a lipocalin in human nasal mucosa. *Comp Biochem Physiol B* 118:819–824.
38. Lacazette E, Gachon A-M, Pitiot G (2000) A novel human odorant-binding protein gene family resulting from genomic duplicons at 9q34: Differential expression in the oral and genital spheres. *Hum Mol Genet* 9:289–301.
39. Briand L, et al. (2002) Evidence of an odorant-binding protein in the human olfactory mucus: Location, structural characterization, and odorant-binding properties. *Biochemistry* 41:7241–7252.
40. Nagnan-Le Meillour P, Vercoutter-Edouart AS, Hilliou F, Le Danvic C, Lévy F (2014) Proteomic analysis of pig (*Sus scrofa*) olfactory soluble proteome reveals o-linked-N-acetylglucosaminylation of secreted odorant-binding proteins. *Front Endocrinol (Lausanne)* 5:202.
41. Nagnan-Le Meillour P, Le Danvic C, Brimau F, Chemineau P, Michalski JC (2009) Phosphorylation of native porcine olfactory binding proteins. *J Chem Ecol* 35: 752–760.
42. Leal GM, Leal WS (2014) Binding of a fluorescence reporter and a ligand to an odorant-binding protein of the yellow fever mosquito, *Aedes aegypti*. *F1000 Res* 3: 305.
43. Sun YF, et al. (2012) Two odorant-binding proteins mediate the behavioural response of aphids to the alarm pheromone (E)-ß-farnesene and structural analogues. *PLoS One* 7:e32759.
44. Wyatt TD (2014) *Pheromones and Animal Behavior: Chemical Signals and Signatures* (Cambridge Univ Press, Cambridge, UK).
45. Archunan G, Rajanarayanan S, Karthikeyan K (2014) Cattle pheromones. *Neurobiology of Chemical Communication*, ed Mucignat-Caretta C (CRC/Taylor & Francis, Boca Raton, FL), pp 461–468.
46. Rasmussen LE, Lee TD, Roelofs WL, Zhang A, Daves GD, Jr (1996) Insect pheromone in elephants. *Nature* 379:684.
47. Poddar-Sarkar M, Brahmachary RL (2014) Pheromones of tiger and other big cats. *Neurobiology of Chemical Communication*, ed Mucignat-Caretta C (CRC/Taylor & Francis, Boca Raton, FL), pp 407–460.
48. Chung MJ, Cheng SS, Lin CY, Chang ST (2012) Profiling of volatile compounds of *Phyllostachys pubescens* shoots in Taiwan. *Food Chem* 134:1732–1737.
49. Guo M, et al. (2015) Evaluating the environmental health effect of bamboo-derived volatile organic compounds through analysis the metabolic indices of the disorder animal model. *Biomed Environ Sci* 28:595–605.
50. Spinelli S, et al. (1998) The structure of the monomeric porcine odorant binding protein sheds light on the domain swapping mechanism. *Biochemistry* 37:7913–7918.
51. Schiefner A, Freier R, Eichinger A, Skerra A (2015) Crystal structure of the human odorant binding protein, OBPIIa. *Proteins* 83:1180–1184.
52. Zhu L, Zhang S, Gu X, Wei F (2011) Significant genetic boundaries and spatial dynamics of giant pandas occupying fragmented habitat across southwest China. *Mol Ecol* 20:1122–1132.
53. Zhu L, et al. (2013) Genetic consequences of historical anthropogenic and ecological events on giant pandas. *Ecology* 94:2346–2357.
54. Garbe JR, Prakapenka D, Tan C, Da Y (2016) Genomic inbreeding and relatedness in wild panda populations. *PLoS One* 11:e0160496.
55. Li Y, et al. (2017) Withered on the stem: Is bamboo a seasonally limiting resource for giant pandas? *Environ Sci Pollut Res Int* 24:10537–10546.
56. Bian X, et al. (2013) Exposure to odors of rivals enhances sexual motivation in male giant pandas. *PLoS One* 8:e69889.
57. Bortolussi G, et al. (2015) Impairment of enzymatic antioxidant defenses is associated with bilirubin-induced neuronal cell death in the cerebellum of Ugt1 KO mice. *Cell Death Dis* 6:e1739.
58. Kabsch W (2010) XDS. *Acta Crystallogr D Biol Crystallogr* 66:125–132.
59. Vagin A, Teplyakov A (2010) Molecular replacement with MOLREP. *Acta Crystallogr D Biol Crystallogr* 66:22–25.
60. Blanc E, et al. (2004) Refinement of severely incomplete structures with maximum likelihood in BUSTER-TNT. *Acta Crystallogr D Biol Crystallogr* 60:2210–2221.
61. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66:486–501.
62. Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* 372:774–797.
63. Schrödinger, LLC (2015) The PyMOL Molecular Graphics System, Version 1.5.0.4. Available at http://pymol.org. Accessed October 11, 2017.
64. Winn MD, et al. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* 67:235–242.
65. Murshudov GN, et al. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* 67:355–367.
66. Bachi A, Dalle-Donne I, Scaloni A (2013) Redox proteomics: Chemical principles, methodological approaches and biological/biomedical promises. *Chem Rev* 113: 596–698.

Supporting Information

Zhu et al. 10.1073/pnas.1711437114

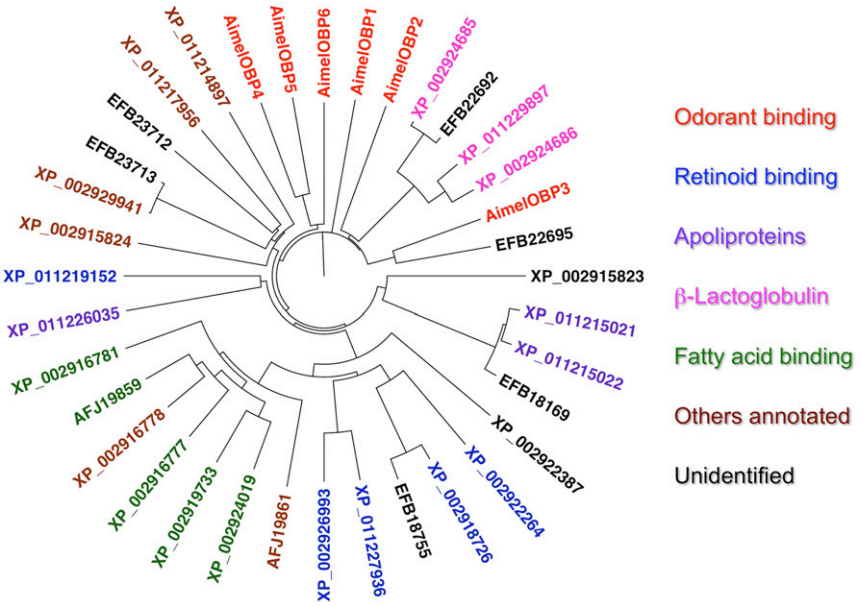


Fig. S1. Lipocalins of the giant panda. After discarding double entries, extremely long sequences, and short fragments, we have found 36 lipocalins of the giant panda *Ailuropoda melanoleuca* in the NCBI Protein database. Apart from the six OBPs described in this work, assignment to different subclasses is based on annotations in the NCBI database.

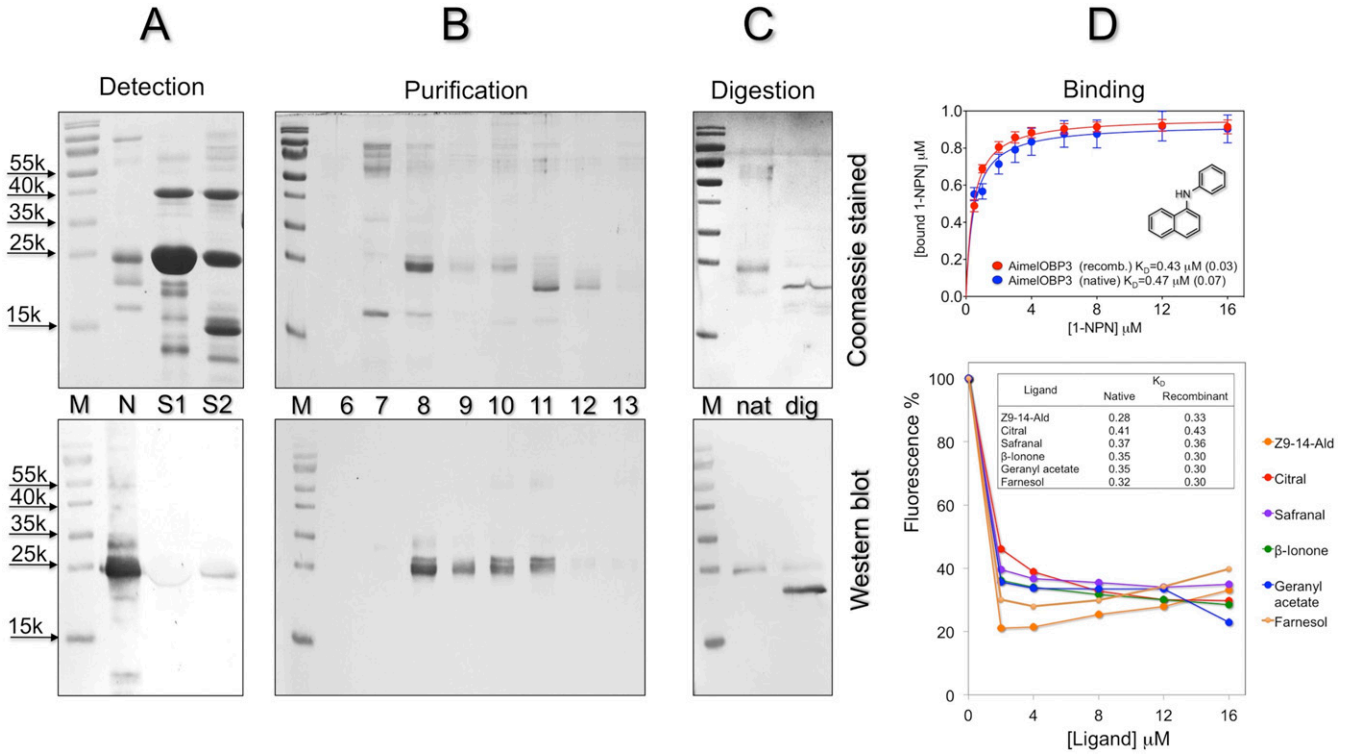


Fig. S2. (*A*) SDS/PAGE analysis and Western blot of crude samples of nasal mucus and saliva. AimelOBP3 appears to be highly represented in the nasal mucus but is barely detectable in two different samples of saliva. M, molecular weight markers; N, nasal mucus; S, saliva. (*B*) Purification of AimelOBP3 from nasal mucus on Mono-Q. The protein was identified by Western blot analysis and found to be mostly concentrated and relatively pure in fraction 8. (*C*) SDS/PAGE and Western blot of an aliquot from fraction 8 before and after digestion with *N*-glycosidase. After enzymatic treatment, the protein migrates with a lower apparent molecular weight, thus confirming the *N*-glycosylation detected by MS (Fig. 2). (*D*) Binding assays performed on AimelOBP3 purified from nasal mucus (fraction 8 from *B*). The native protein binds the fluorescent probe 1-NPN and few selected ligands with affinities not significantly different from those measured with the recombinant protein, thus indicating that the glycosylation does not affect the binding properties of this protein.

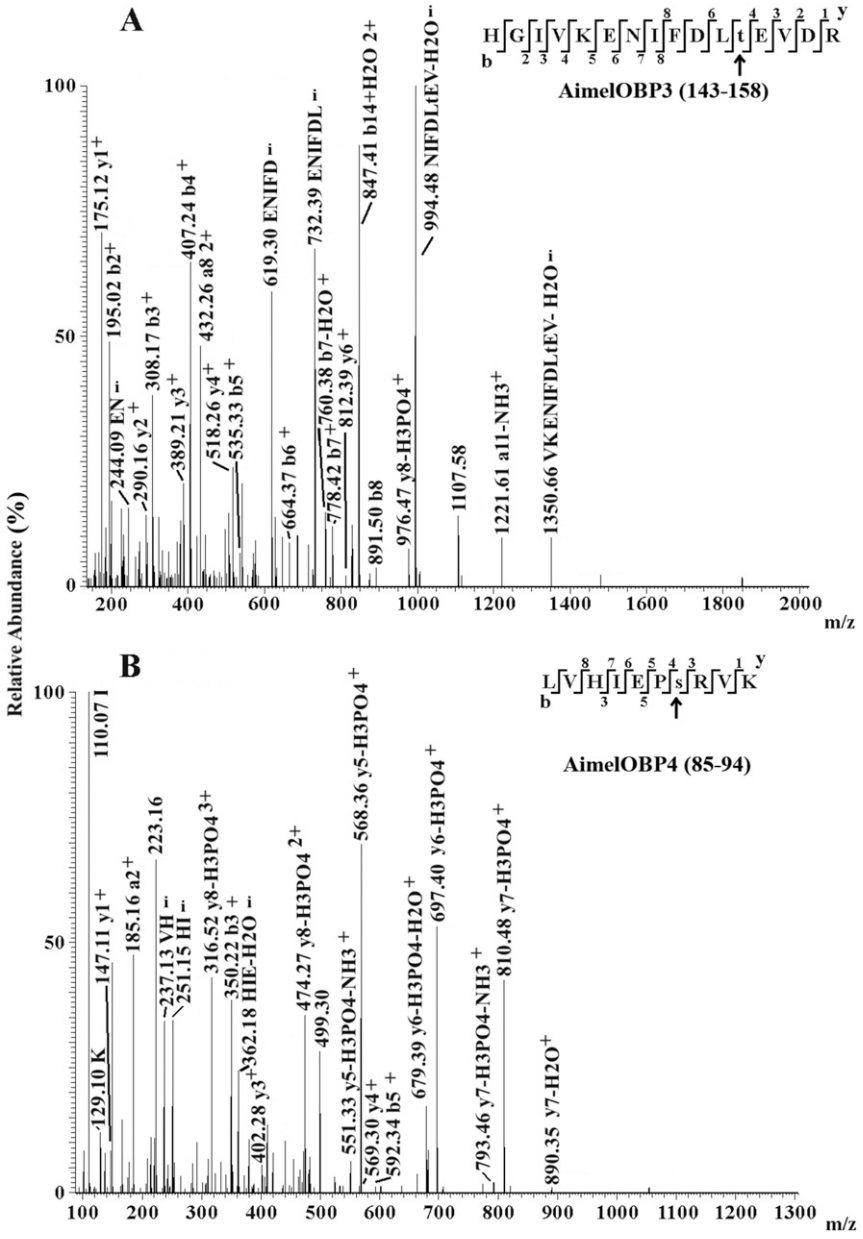


Fig. S3. Fragmentation spectra of phosphopeptides identified in AimelOBP3 and AimelOBP4. (*A*) MS/MS spectrum of the ion at *m/z* 655.98 present in the tryptic digest of AimelOBP3. Based on fragmentation data, phosphorylation was assigned to Thr154. (*B*) MS/MS spectrum of the ion at *m/z* 419.90 present in the tryptic digest of AimelOBP4. Based on fragmentation data, phosphorylation was assigned to Ser91. Assigned peptide fragments are highlighted in the peptide sequences reported within each panel.

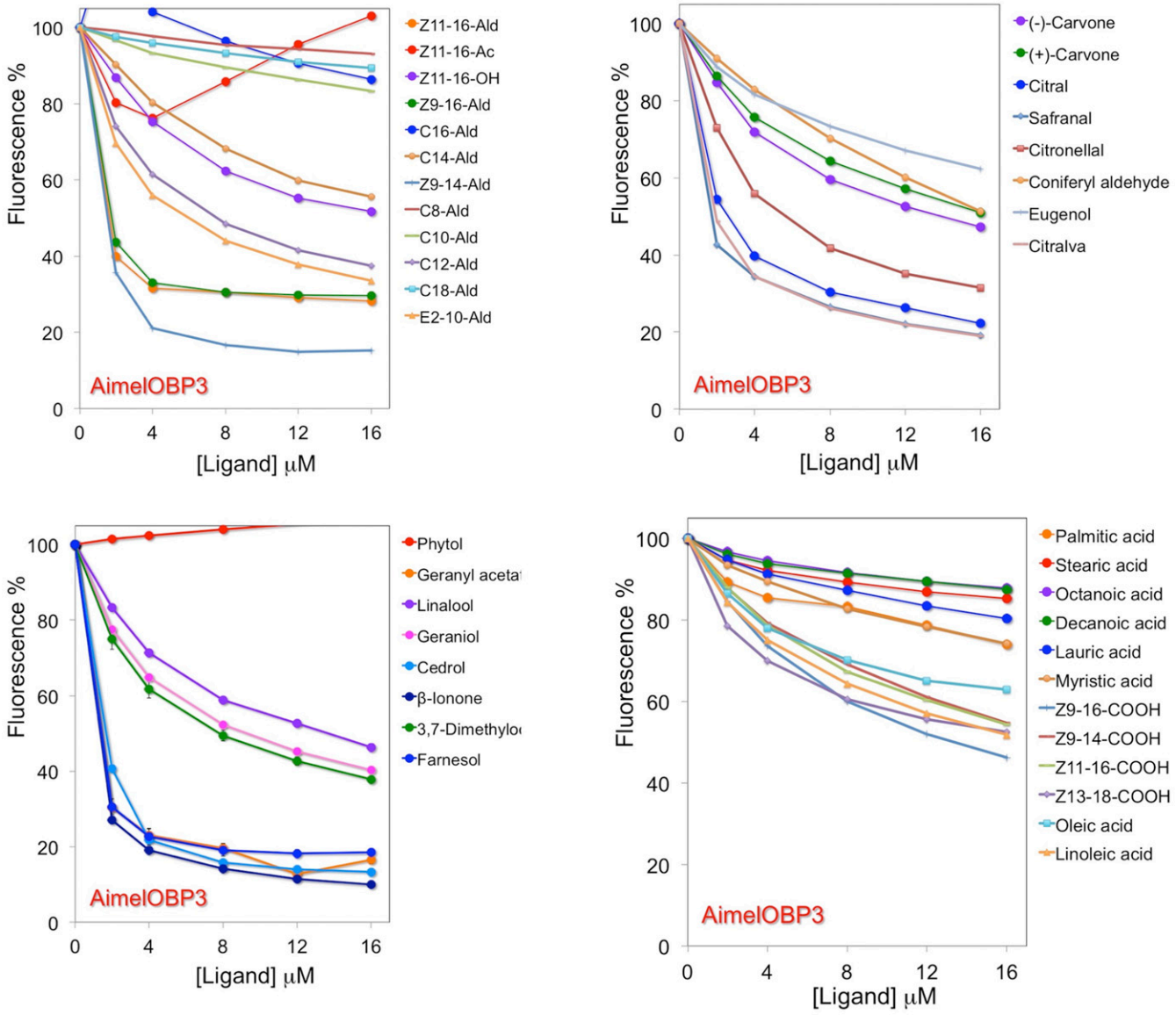


Fig. S4. Competitive binding curves of AimelOBP3-WT toward 40 ligands.

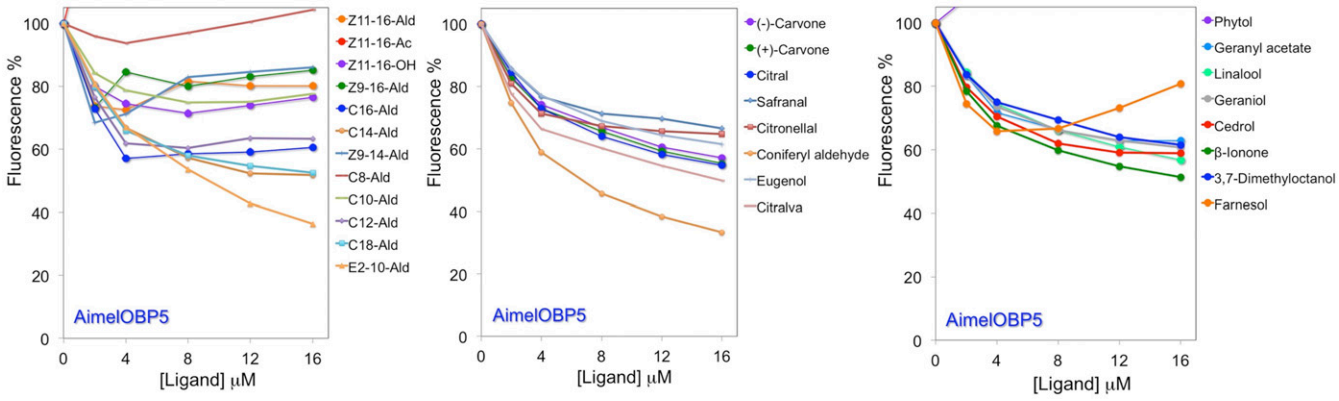


Fig. S5. Competitive binding curves of AimelOBP5 toward 28 ligands. The results relative to 12 fatty acids are reported in Fig. 3.

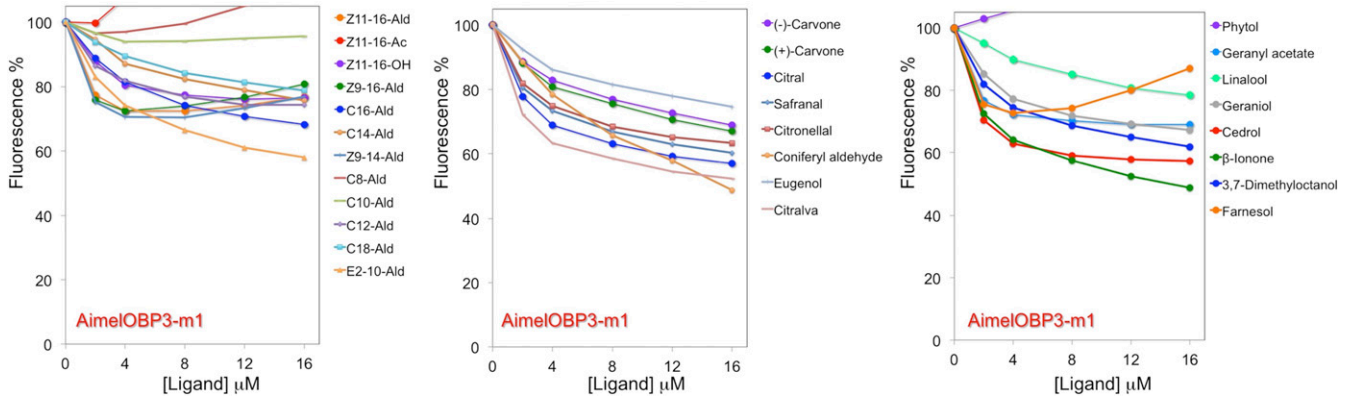


Fig. S6. Competitive binding curves of AimelOBP3-m1 (Glu120Ala) toward 28 ligands.

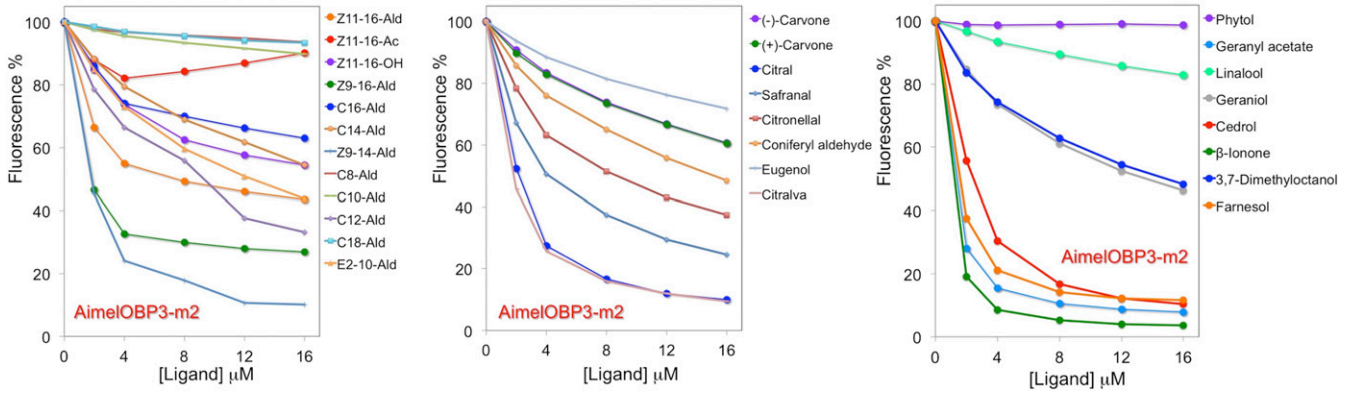


Fig. S7. Competitive binding curves of AimelOBP3-m2 (Ser122Ala) toward 28 ligands.

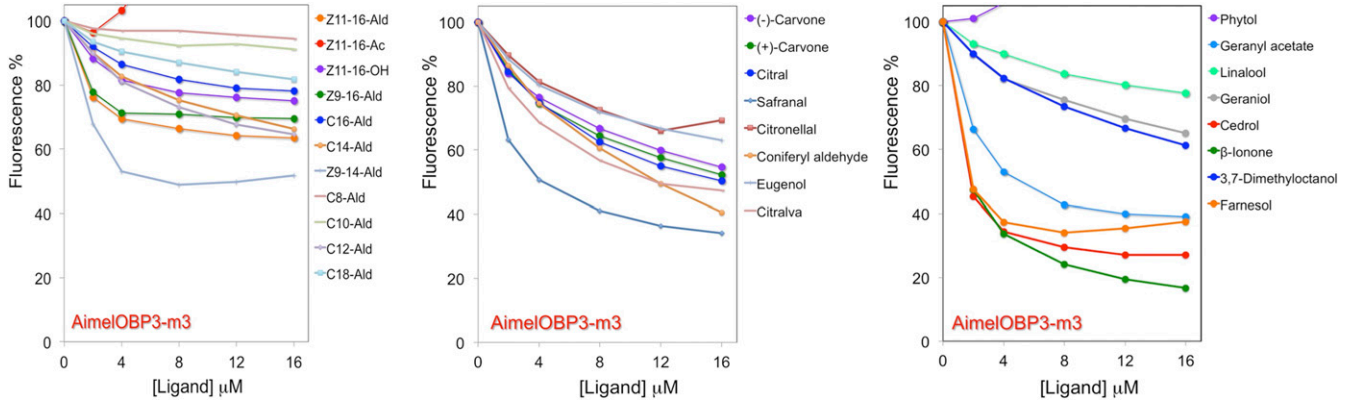


Fig. S8. Competitive binding curves of AimelOBP3-m3 (Asn90Leu) toward 28 ligands.

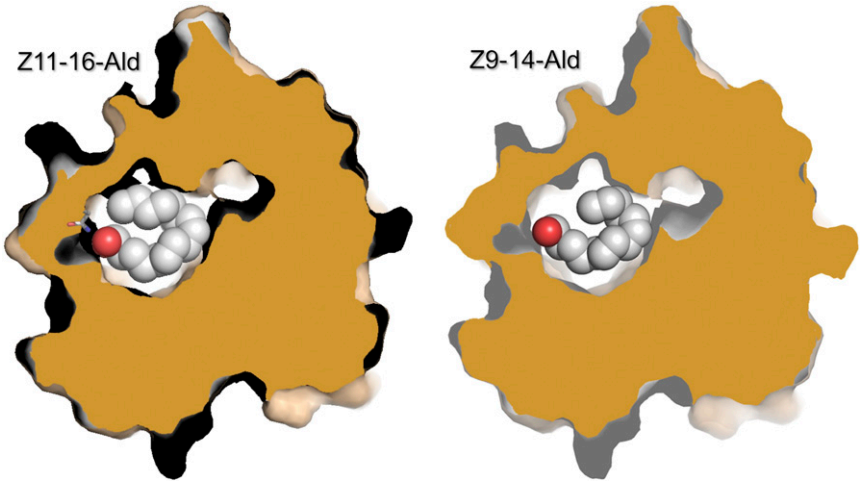


Fig. S9. Docking of two of the best ligands to AimelOBP3. Z11-16:Ald fits nicely inside the binding cavity, establishing a hydrogen bond with Asn90. The shorter aldehyde Z9-14:Ald also fits in the cavity but leaves some free space. The protein has been slabbed at the level of the ligand.

Table S1. *N*-linked glycopeptides identified in AimelOBP3 and AimelOBP6 digests from crude nasal mucus

Sequence	Modification(s)	*N*-linked glycans	*z*	Obs. *m/z*	Calc. *m/z*	Δppm	Obs. MH^+	Calc. MH^+
OBP3 N-glycosilation								
HIEVLSN[+1752.64460]SSLLFNM[+15.99492]HIK	N7(NGlyc/1752.6446); M14(Ox/15.9949)	HexNAc(4)Hex(4)Fuc(2)	4	938.4296	938.4285	1.15	3,750.6966	3,750.6923
VFVNHIEVLSN[+1606.58669]SSLLFNM[+15.99492]HIK	N11(NGlyc/1606.5867); M18(Ox/15.9949)	HexNAc(4)Hex(4)Fuc(1)	4	1,016.9800	1,016.7261	3.02	4,064.8982	4,063.8825
VFVNHIEVLSN[+1751.62420]SSLLFNM[+15.99492]HIK	N11(NGlyc/1751.6242); M18(Ox/15.9949)	HexNAc(4)Hex(4)NeuAc(1)	4	1,052.9869	1,052.9855	1.41	4,208.9260	4,208.9200
VFVNHIEVLSN[+1752.64460]SSLLFNM[+15.99492]HIK	N11(NGlyc/1752.6446); M18(Ox/15.9949)	HexNAc(4)Hex(4)Fuc(2)	4	1,053.4900	1,053.2406	−1.34	4,210.9382	4,209.9404
VFVNHIEVLSN[+1914.69743]SSLLFNM[+15.99492]HIK	N11(NGlyc/1914.6974); M18(Ox/15.9949)	HexNAc(4)Hex(5)Fuc(2)	4	1,094.0100	1,093.7538	4.93	4,373.0182	4,371.9933
VFVNHIEVLSN[+1955.72397]SSLLFNM[+15.99492]HIK	N11(NGlyc/1955.724); M18(Ox/15.9949)	HexNAc(5)Hex(4)Fuc(2)	4	1,104.0106	1,104.0104	0.22	4,413.0208	4,413.0198
VFVNHIEVLSN[+1996.75052]SSLLFNM[+15.99492]HIK	N11(NGlyc/1996.7505); M18(Ox/15.9949)	HexNAc(6)Hex(3)Fuc(2)	4	1,114.5200	1,114.2670	1.90	4,455.0582	4,454.0464
VFVNHIEVLSN[+2012.74544]SSLLFNM[+15.99492]HIK	N11(NGlyc/2012.7454); M18(Ox/15.9949)	HexNAc(6)Hex(4)Fuc(1)	4	1,118.2500	1,118.2658	−14.12	4,469.9782	4,470.0413
VFVNHIEVLSN[+2133.77171]SSLLFNM[+15.99492]HIK	N11(NGlyc/2133.7717); M18(Ox/15.9949)	HexNAc(5)Hex(6)Fuc(1)	4	1,148.2800	1,148.5223	7.40	4,590.0982	4,591.0675
VFVNHIEVLSN[+2142.80843]SSLLFNM[+15.99492]HIK	N11(NGlyc/2142.8084); M18(Ox/15.9949)	HexNAc(6)Hex(3)Fuc(3)	4	1,151.0100	1,150.7815	−19.45	4,601.0182	4,600.1043
VFVNHIEVLSN[+2246.81939]SSLLFNM[+15.99492]HIK	N11(NGlyc/2246.8194); M18(Ox/15.9949)	HexNAc(5)Hex(4)Fuc(2)NeuAc(1)	4	1,177.0300	1,176.7843	−4.34	4,705.0982	4,704.1152
VFVNHIEVLSN[+2263.83471]SSLLFNM[+15.99492]HIK	N11(NGlyc/2263.8347); M18(Ox/15.9949)	HexNAc(5)Hex(5)Fuc(3)	4	1,181.0402	1,181.0381	1.78	4,721.1390	4,721.1305
VFVNHIEVLSN[+2294.80414]SSLLFNM[+15.99492]HIK	N11(NGlyc/2294.8041); M18(Ox/15.9949)	HexNAc(5)Hex(6)NeuGc(1)	4	1,188.5400	1,188.7805	8.74	4,751.1382	4,752.1000
VFVNHIEVLSN[+2303.84086]SSLLFNM[+15.99492]HIK	N11(NGlyc/2303.8409); M18(Ox/15.9949)	HexNAc(6)Hex(3)Fuc(2)NeuGc(1)	4	1,191.2700	1,191.0396	−17.20	4,762.0582	4,761.1367
VFVNHIEVLSN[+2393.87255]SSLLFNM[+15.99492]HIK	N11(NGlyc/2393.8725); M18(Ox/15.9949)	HexNAc(7)Hex(6)	4	1,213.8000	1,213.5476	1.33	4,852.1782	4,851.1684
VFVNHIEVLSN[+1564.53974]SSLLFNM[+15.99492]HIK	N11(NGlyc/1564.5397); M18(Ox/15.9949)	HexNAc(3)Hex(4)NeuGc(1)	3	1,341.6000	1,341.2834	−13.30	4,022.7854	4,021.8356
VFVNHIEVLSN[+1914.69743]SSLLFNM[+15.99492]H	N11(NGlyc/1914.6974); M18(Ox/15.9949)	HexNAc(4)Hex(5)Fuc(2)	3	1,377.9400	1,377.6096	−2.94	4,131.8054	4,130.8142
VFVNHIEVLSN[+3055.04271]SSLLFNM[+15.99492]HIK	N11(NGlyc/3055.0427); M18(Ox/15.9949)	HexNAc(5)Hex(6)Fuc(1)NeuGc(3)	4	1,378.6100	1,378.8401	15.05	5,511.4182	5,512.3385
VFVNHIEVLSN[+1752.64460]SSLLFNM[+15.99492]HIK	N11(NGlyc/1752.6446); M18(Ox/15.9949)	HexNAc(4)Hex(4)Fuc(2)	3	1,404.3200	1,403.9850	0.39	4,210.9454	4,209.9404
VFVNHIEVLSN[+1996.75052]SSLLFNM[+15.99492]HIK	N11(NGlyc/1996.7505); M18(Ox/15.9949)	HexNAc(6)Hex(3)Fuc(2)	3	1,485.6600	1,485.3536	−18.92	4,454.9654	4,454.0464
VFVNHIEVLSN[+2263.83471]SSLLFNM[+15.99492]H	N11(NGlyc/2263.8347); M18(Ox/15.9949)	HexNAc(5)Hex(5)Fuc(3)	3	1,494.3300	1,493.9887	4.59	4,480.9754	4,479.9515
OBP6 N-glycosilation								
LASN[+1054.37004]NSALIEPGGHFR	N4(NGlyc/1054.37)	HexNAc(2)Hex(4)	3	913.0800	913.0818	−1.99	2,737.2254	2,737.2309

Assignment was obtained after searching MS and MS/MS raw data with Bionics 2.6.46 software. Reported structures are associated with peptide species assigned on the basis of their fragmentation spectra. Calc., calculated; Obs., observed.

Table S2. Data collection and refinement statistics of AimelOBP3

Data collection	
PDB	5NGH
Source	Soleil PX1
Space group	$P6_522$
Cell (Å; °)	a = b = 94.0, c = 114.5;α = β = 90.0, γ = 120
No. of monomers	1
Resolution limits (Å)	80–2.80 (2.97–2.8)
R_{merge}	0.19 (0.114)
CC1/2	0.998 (0.93)
Unique reflections	7,821 (1,218)
Mean ((I)/SD(I))	10.3 (1.0)
Completeness (%)	99.9 (99.8)
Multiplicity	10.4 (10.3)
Refinement	
Resolution (Å)	24.7–2.80 (3.13–2.80)
No. of reflections	7,784 (2,131)
No. of proteins/water/ions	1,272/16/16
Test set reflections	389
R_{work}/R_{free}	0.277/0.287 (0.38/0.40)
rmsd Bonds (Å)/angles (°)	0.010/1.26
B Wilson/B mean (Åe	104/93
Ramachandran: preferred/allowed/outliers (%)	94.4/5.6/0

Numbers in parentheses refer to the highest resolution bin.

Dataset S1. Proteomic data of crude nasal mucus from *A. melanoleuca*

Dataset S1

Dataset S2. Proteomic data of crude saliva from *A. melanoleuca*

Dataset S2

Dataset S3. Values of IC_{50} and dissociation constants (micromolar) for the ligands used with the five proteins described in this work

Dataset S3

For very weak ligands, if an IC_{50} value could not be extrapolated, an arbitrary value of 200 μM was assigned.