CMLS, Cell. Mol. Life Sci. 60 (2003) 390–400
1420-682X/03/020390-11

CMLS Cellular and Molecular Life Sciences

Research Article

Biochemical characterization and bacterial expression of an odorant-binding protein from *Locusta migratoria*

L. Ban [a], A. Scaloni [b], C. D'Ambrosio [b], L. Zhang [a], Y. Yan [a] and P. Pelosi [c,*]

[a] Department of Entomology, China Agricultural University, Beijing (China)
[b] Proteomics and Mass Spectrometry Laboratory-I.S.P.A.A.M., National Research Council, Naples (Italy)
[c] Dipartimento di Chimica e Biotecnologie Agrarie, University of Pisa, via S. Michele, 26124 Pisa (Italy), Fax + 39 050 571564, e-mail: ppelosi@agr.unipi.it

Received 23 November 2002; accepted 19 December 2002

Abstract. Analysis of soluble proteins from different body parts of *Locusta migratoria* revealed a fast-migrating component in native electrophoresis, unique to antennae of both sexes. N-terminal sequence analysis and cloning identified this protein as a member of the insect odorant-binding proteins, carrying a well-conserved six-cysteine motif. Mass spectrometry analysis confirmed the occurrence of two distinct polypeptide species determined by nucleotide sequencing and demonstrated that the cysteine residues are paired in an interlocked fashion. The protein was expressed in a bacterial system with yields of about 10 mg/l of culture, mostly present as inclusion bodies. However, this recombinant product was solubilized after disulfide reduction. Air oxidation yielded a species with all disulfides spontaneously formed as in the native counterpart. Both native and recombinant proteins migrated as a dimer in gel filtration chromatography. Ligand binding was measured, using N-phenyl-1-naphthylamine as the fluorescent probe; the affinity of other ligands was measured in competitive binding assays. The protein exhibited great resistance to thermal denaturation even following prolonged treatment at 100 °C. A structural model for this dimeric species was generated on the basis of its sequence homology with *Bombyx mori* pheromone-binding protein, whose three-dimensional structure has been resolved as an unbound species and in complex with its physiological ligand. This is the first report of an odorant-binding protein identified and characterized from Orthoptera.

Key words. Odorant-binding protein; pheromone; chemical communication; fluorescence binding assay; *Locusta migratoria*.

Soluble proteins play important roles in chemical communication, both in vertebrates and insects [1, 2]. These polypeptides are present in unusually high concentrations in the mucus of vertebrates and in the sensillar lymph of insects. They are known under the general name of odorant-binding proteins (OBPs), and share small size (15–20 kDa), an acidic nature, and the property of reversibly binding odorants and pheromones. However, vertebrate OBPs share no structural similarity with those of insects, the former showing the typical β barrel folding of lipocalins [3, 4], the latter being mainly composed of α-helical domains [5]. Insect OBPs are also characterized by six conserved cysteines, establishing three interlocked disulfide bridges [6, 7]. A second family of putative OBPs, structurally different from both vertebrate and insect OBPs, is highly expressed in the lymph of insect chemosensilla [8–12]. These polypeptides, called chemosensory proteins (CSPs) are smaller than OBPs (12 kDa) and contain only four cysteines linked by two

* Corresponding author.

non-interlocked disulfide bridges [10]. Very recently, two novel classes of putative odorant- and pheromone-binding proteins (PBPs) have been reported in *Drosophila*, specifically expressed in the fore tarsi of male individuals [13].

Insect OBPs are further divided into PBPs, general OBPs (GOBPs), and antennal binding protein X (ABPX). Members of each sub-class have been described in several insect species from most orders, each species expressing a few isoforms of each sub-class. Only the three-dimensional structure of *Bombyx mori* PBP has been resolved and the position of its ligand, bombykol, identified inside the binding cavity. It consists mainly of α-helical domains that confer high stability on the molecule [5, 14]. OBPs have been mainly described in Lepidoptera, but recently, examples from other insect orders have been reported, such as Diptera [8, 9, 15, 16], Hymenoptera [17–20], and Coleoptera [21, 22]. The only examples of OBPs from hemimetabolous insect are the so-called LAP of *Lygus lineolaris*, belonging to the order of Hemiptera [23, 24], and the very recently reported OBPs of the termite *Zootermopsis nevadensis* [25]. Beside the six conserved cysteines, OBPs of different insect orders share only a few residues and appear to be more divergent than CSPs.

These polypeptides, especially PBP, were designated binding proteins on the basis of ligand-binding assays with radioactively labeled pheromones [26]. More recently, the use of fluorescent binding assays at the equilibrium point has allowed accurate evaluation of the relative dissociation constants [27]. Similar studies with GOBPs and ABPX are lacking and their function in chemoreception is mainly suggested by the specific expression of these proteins in the lymph of chemosensilla.

The physiological function of OBPs, in both insects and vertebrates is still under investigation. Are they essential for correct performance of the olfactory system and what role do they play in the recognition of chemical stimuli? Recently, two reports have indicated that the absence of OBPs in insects can affect their behavior toward odors and pheromones. The so called 'lush' protein, a member of the OBP family in *Drosophila*, is required for correct orientation of the flies. Individuals not expressing this protein are attracted by high concentrations of ethanol, otherwise avoided by normal flies [28]. In fire ants, the lack of an OBP produces colonies with multiple queens, a fact probably related to their impaired ability to perceive the queen's pheromone [20].

Here we report the first identification of an orthopteran OBP, its cloning, and structural and functional characterization. A bacterial expression procedure was optimized to generate amounts of recombinant product suitable for further structural investigations.

Materials and methods

Purification and identification of loc-OBP1

A crude antennal extract from *Locusta migratoria manilensis* was chromatographed on a 1 × 30 cm Superose-12 column, using 50 mM Tris-HCl, pH 7.4 as the eluent. Fractions containing low-molecular-weight (14–20 kDa) proteins, as evaluated by SDS-PAGE, were further fractionated by preparative electrophoresis in native conditions, using a Mini-Prep Cell (Bio-Rad, Hercules, Calif.). Fractions were analyzed by native electrophoresis and the putative OBP was identified as the fastest migrating band. For N-terminal sequencing, protein samples were blotted onto PVDF membrane and subjected to automated Edman degradation, using a Procise 491 protein sequencer (Applied Biosystems, Foster City, Calif.) equipped with a 140C microgradient apparatus and a 785A UV detector (Applied Biosystems) for the automated identification of phenylthiohydantoin-amino acids.

RNA extraction and cDNA synthesis

Total RNA was extracted from the antennae of male and female adult *L. migratoria*, using the Trizol reagent kit (GIBCO BRL, Paisley, U. K.), according to the manufacturer's protocol. cDNA was prepared from total RNA by reverse transcription, using 200 units of Moloney murine leukemia virus (MMLV) reverse transcriptase (GIBCO BRL) and 0.5 µg of oligo dT_{12-18} (Sigma) in a 50-µl total volume. The mixture also contained 0.5 mM of each dNTP (Pharmacia Biotech, Uppsala, Sweden), 75 mM KCl, 3 mM $MgCl_2$, 10 mM DTT, and 0.1 mg/ml BSA in 50 mM Tris-HCl, pH 8.3. The reaction mixture was incubated at 37 °C for 60 min and the product was directly used for PCR amplification or stored at –20 °C.

Polymerase chain reaction

Aliquots of 1 µl crude cDNA were amplified in a Bio-Rad Gene Cycler thermocycler, using 2.5 units of *Thermus aquaticus* DNA polymerase (Sigma), 1 mM of each dNTP (Pharmacia Biotech), 1 µM of each PCR primer, 50 mM KCl, 2.5 mM $MgCl_2$ and 0.1 mg/ml BSA in 10 mM Tris-HCl, pH 8.3, containing 0.1 % Triton X-100. The 32mer degenerate primer, GAY GTN AAY ATG AAR CT, designed on the first six amino acid residues of the OBP sequence, was used at the 5′ end, together with an oligo-dT_{18}V at the 3′ end. After a denaturing step at 95 °C for 5 min, the reaction was performed for 35 cycles (95 °C for 1 min, 44 °C for 1 min, 68 °C for 1 min), followed by a final step of 7 min at 72 °C.

Cloning and sequencing

The 600-bp PCR product was purified by excision of the band from the agarose gel and elution, using the Qiaex II gel extraction kit (Qiagen, Hilden, Germany), and ligated

into a pGEM vector. After transformation of *Escherichia coli* XL-1 Blue competent cells with the ligation product, positive colonies were selected by PCR, using the plasmid primers SP6 and T7 and grown in LB/ampicillin medium. DNA was extracted and sequenced. The plasmid (25 ng) and the insert were ligated by incubating in a 1 : 5 molar ratio overnight at room temperature. The ligated product was used to transform XL-1 Blue competent cells. After incubation in LB medium for 1 h at 37 °C, the cells were plated on LB/agar, containing ampicillin, IPTG and X-Gal. White colonies were then grown in liquid LB/ampicillin and analyzed for the presence of the right insert by PCR, using the plasmid primers SP6 and T7, followed by 2% agarose electrophoresis. Nucleotide sequences of both strands of the cDNA clones were determined from double-stranded plasmid DNA using the Perkin Elmer-Applied Biosystems Dye Deoxy Terminator Cycle Sequencer Kit and an Applied Biosystems 310 automated sequencer at the ENEA laboratory (Casaccia, Rome, Italy).

Cloning of loc-OBP1 in the expression vector

The DNA encoding loc-OBP1-B was obtained by PCR from the pGEM plasmid containing the appropriate sequence, using the following primers: TAC ATA TGG ACG TGA ACA TGA AAC TT at the 5′ end and AAT TGA ATT CCT ATT AGT CAC CAA TGG TTG CAT C at the 3' end. Sequences were designed to include an *Nde*I restriction site with the ATG codon in the 5′ primer and an *Eco*RI restriction site in the 3′ primer. Amplified DNA was again inserted into a pGEM vector that was used to transform *E. coli* XL-1 Blue competent cells. Colonies were tested for the presence of the insert by PCR, using the plasmid primers SP6 and T7, DNA was extracted, purified and digested with *Nde*I and *Eco*RI. The excised 400-bp fragment was purified from the agarose electrophoretic gel and cloned into pET5b (Novagen, Darmstadt, Germany), previously linearized with the same enzymes. The resulting plasmid, pET-loc-OBP1-B, was sequenced and shown to encode the mature loc-OBP1-B.

Production of recombinant loc-OBP1-B

For protein expression, *E. coli* BL21(DE3) cells were transformed with the pET-loc-OBP1 plasmid. Single colonies were grown overnight in 10 ml Luria-Bertani/Miller broth containing 100 mg/l ampicillin. The culture was diluted 1 : 100 with fresh medium and grown at 37 °C until the absorbance at 600 nm reached 0.6 AU. At this stage, 0.4 mM isopropyl thio-β-D-galactoside was added to the culture to induce expression. After 3 h at 37 °C, or alternatively overnight at room temperature, the cells were harvested by centrifugation, resuspended in 50 mM Tris-HCl pH 7.4, 300 mM NaCl, and 1 mM PMSF, and lysed by sonication. The recombinant OBP was present at this stage in the pellet as inclusion bodies and was solubilized by incubation with 4 M urea, 1 mM DTT at 90 °C for 10 min. After prolonged dialysis at air atmosphere, the protein was purified by two chromatographic steps on anion-exchange resin DE-52, using a gradient of 0 – 0.5 M NaCl in Tris-HCl buffer, followed by gel filtration on Superose-12. Fractions were analyzed by SDS-PAGE for the presence of protein contaminants, and by UV spectroscopy to evaluate the amount of DNA coeluted with the protein.

Molecular mass characterization

The apparent molecular mass for native and recombinant loc-OBP samples was determined by gel filtration chromatography using the FPLC System (Pharmacia), with a Superose-12 column (1 × 30 cm), and eluted with 50 mM Tris-HCl, pH 7.4, at a flow rate of 0.2 ml/min. The following proteins were used as molecular mass standards: cytochrome c (12,300 Da), bovine carbonic anhydrase (29,000 Da), and bovine serum albumin (66,000 Da).

Enzymatic digestion

Protein samples were digested with trypsin (Sigma) in 0.4% ammonium bicarbonate, pH 6.5 at 37 °C for 18 h, using an enzyme/substrate ratio of 1 : 50 (w/w). Digest aliquots were directly subjected to a desalting/concentration step on µZipTipC_{18} (Millipore, Bedford, Mass.) before matrix-assisted laser desorption ionization mass spectrometry (MALDI-MS) analysis or separated by reverse-phase HPLC on a Vydac C_{18} column 218TP52 (250 × 1 mm), 5 µm, 300 Å pore size (The Separation Group, Hesperia, Calif.), using a linear gradient from 5 to 60% of acetonitrile in 0.1% TFA over 60 min, at a flow rate of 60 µl/min. Individual components were collected manually and analyzed by MALDI-MS.

Mass spectrometry

MALDI mass spectra were recorded using a Voyager DE-PRO MALDI-TOF mass spectrometer (Applied Biosystems). A mixture of analyte solution, α-cyano-4-hydroxycinnamic acid, bovine insulin, and horse apomyoglobin was applied to the sample plate and dried in vacuo. Mass calibration was performed using the molecular ions from these proteins and the matrix as internal standards.

Fluorescence measurements

Emission fluorescence spectra were recorded on a Jasco FP-750 instrument at 25 °C in a right-angle configuration with a 1-cm light path quartz cuvette and 5-nm slits for both excitation and emission. The protein was dissolved in 50 mM Tris-HCl, pH 7.4, while ligands were added as 1 mM methanol solutions.

Binding assays

To measure the affinity of the fluorescent ligands N-phenyl-1-naphthylamine (1-NPN) and 1-amino-an-

thracene (1-AMA) to loc-OBP1, a 2-µM solution of protein in 50 mM Tris-HCl, pH 7.4, was titrated with aliquots of 1 mM ligand solution in methanol to a final concentration of 1–20 µM. The affinity of other ligands was measured in competitive binding assays, using 1-NPN as the fluorescent reporter at 1 or 2 µM concentration.
For determining binding constants, the intensity values corresponding to the maximum fluorescence emission were plotted against free ligand concentrations. Bound ligand was evaluated from the values of fluorescence intensity assuming the protein was 100% active, with a stoichiometry of 1 : 1 protein:ligand at saturation. The curves were linearized using Scatchard plots. Dissociation constants of the competitors were calculated from the corresponding IC_{50} values, using the equation: $K_d = [IC_{50}]/(1+[1\text{-NPN}]/K_{1\text{-NPN}})$, [1-NPN] being the free concentration of 1-NPN and $K_{1\text{-NPN}}$ being the dissociation constant of the complex OBP1/1-NPN.

Temperature stability measurements

One milliliter aliquots of 2 µM OBP in Tris-HCl were kept for 1, 5, 10, and 20 min at 100 °C. After cooling to room temperature, 2 µl of 1 mM 1-NPN in methanol was added and fluorescence spectra were recorded between 380 and 460 nm, using an excitation wavelength of 337 nm.

Molecular modeling

A three-dimensional model of dimeric *L. migratoria* OBP1 was constructed on the basis of the crystallographic structure of dimeric OBP from *B. mori* (PDB code 1DQE) taken from the Protein Data Bank, Brookhaven National Laboratory (Upton, N. Y.). Sequence analysis was performed using the BLASTP2 algorithm in the EXNRL-3D database. Structure-based sequence alignment was obtained using the program AMPS [29] with secondary-structure-dependent gap penalties [30]. Computer modeling was performed on a Silicon Graphics O2 workstation. A model was constructed with the Insight/Homology program package (Biosym). Several cycles of constrained energy minimization regularized the structure and geometrical parameters. The final structure was validated using the WHATCHECK program [31]. Graphical representation was prepared using the Swiss-PDBviewer [32].

Results

Identification of locust OBP1

Electrophoretic analysis in native conditions of extracts from different parts of the insect body showed the presence of a fast-migrating band only in the antennal extracts of both sexes (fig. 1A). After gel filtration on Superose-12, the low-molecular-weight components were further fractionated by native preparative electrophoresis. The fractions obtained were then again analyzed by native electrophoresis (fig. 1B) and Western blot (not shown), using a polyclonal antiserum against CSPs of *Schistocerca gregaria*. The bands present in fractions 7 and 8 differed from those of the other fractions by their higher mobility and lack of reaction with the antiserum. They could not be assigned to the CSP family and were considered as putative OBPs. N-terminal sequencing of both bands, after blotting onto PVDF membrane, yielded the same sequence, DVNMKLTGRIMDAAKEVDHT, confirming this hypothesis. In both samples, during Edman degrada-

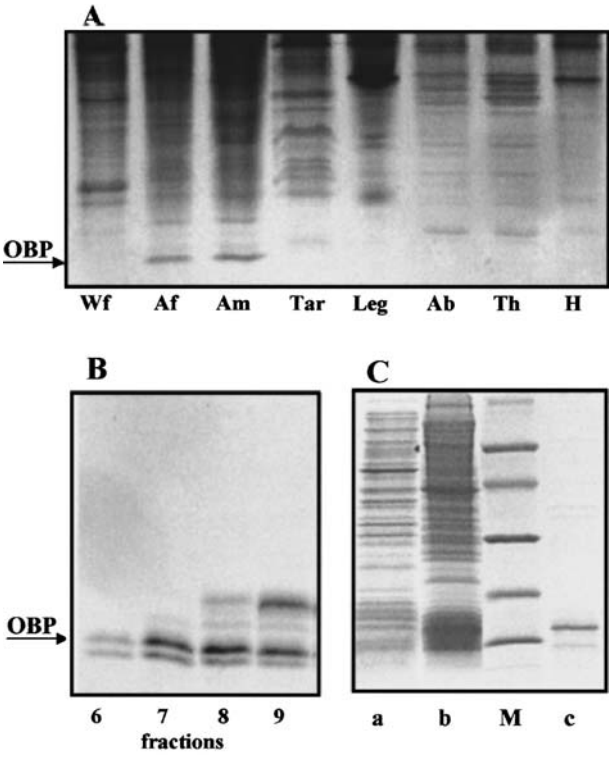

Figure 1. Electrophoretic analysis of *L. migratoria* OBP1. (*A*) Electrophoretic analysis in native conditions of extracts from different parts of the body of *L. migratoria*. The antennal extracts of both sexes contained a fast-migrating band absent in the other samples. Wf, female wings; Af, female antennae; Am, male antennae; Tar, tarsi; Leg, legs; Ab, abdomen; Th, thorax; H, head. (*B*) Purification of OBP1 from the antennae of *L. migratoria*. The antennal extract was separated by gel filtration on a Superose-12 column (not shown) and the low-molecular-weight fraction was further fractionated by preparative electrophoresis in native conditions. The figure shows the electrophoretic analysis in native conditions of the fractions obtained. OBP1 was the first protein to be eluted from the preparative gel, mainly in fractions 7 and 8. (*C*) Bacterial expression of locust OBP1-B. Lane a, crude bacterial pellet before induction; lane b, crude bacterial pellet after addition of 0.4 mM IPTG; lane c, sample solubilized after sonication and extraction of the pellet with 4 M urea and 1 mM DTT. OBP1 represents the major band and was further purified by ion exchange chromotography on a DE-52 column. Molecular weight markers (lane M) were BSA (66 kDa), ovalbumin (45 kDa), carbonic anhydrase (29 kDa), and lactalbumin (14 kDa).

tion at the 19th cycle, phenylthiohydantoin-Asn was also detected.

Cloning and deduced amino acid sequence

A degenerate primer, designed on the first six amino acids of the sequence was used, together with an oligo-dTV, to amplify by PCR sequences from an antennal cDNA sample. A band of about 600 bp was obtained, which was ligated into a pGEM vector. After transformation of XL-1 Blue competent cells, colonies containing the insert were selected by PCR, using the plasmid primers SP6 and T7, and grown in LB/ampicillin medium. DNA was then extracted and sequenced. Analysis of five colonies obtained from two different ligations gave two sequences, differing for the single amino acid substitution His19→Asn. This result confirmed the data obtained by direct sequencing of the native loc-OBP1 and highlighted the occurrence of two isoforms (named A and B) or a single-nucleotide polymorphism phenomenon for this protein.

In both cases, the open reading frame encoded a mature protein of 130 residues, followed by two stop codons and a non-translated region of 200 bases (sequence available from GenBank, accession number AF542076). The deduced protein sequence showed low similarity to a number of insect OBPs from different orders. In particular, the six-cysteine motif, a typical feature of insect OBPs, appeared to be conserved. The best similarity was found with dipteran and coleopteran OBPs. Figure 2A shows the alignment of locust OBP1-A with some representative proteins of the same class.

MALDI-MS analysis of the purified species allowed verification of the deduced sequence. In fact, the spectrum revealed the occurrence of two distinct components with an averaged molecular mass of 14,832.7 Da and 14,808.9 Da, respectively, in perfect agreement with the expected values (14,831.5 Da for loc-OBP1-A and 14,808.5 Da for loc-OBP1-B; data not shown).

Bacterial expression and purification of recombinant protein

One of the isoforms of locust OBP1, bearing Asn19, was expressed in *E. coli*, using a pET5b vector following the protocol described in Materials and methods. The recombinant protein contained an additional Met at the N terminus, as the sole modification with respect to the native species. SDS-PAGE analysis of a sample of the bacterial pellet showed a major protein band being synthesized after induction with IPTG with yields of about 10 mg/l of culture and migrating with an apparent molecular mass of 16 kDa (fig. 1C). After sonication, OBP remained in the pellet, but could be efficiently solubilized by incubation with 4 M urea and 5 mM DTT. The reduced protein was purified by two chromatographic steps on anion-exchange resin DE-52, followed by gel filtration on Superose 12. Refolding and oxidation of cysteines was obtained by spontaneous air oxidation, occurring during prolonged dialysis between the purification steps. The correct nature of the recombinant OBP1 was confirmed by MALDI-MS. The spectrum revealed the occurrence of a component with an averaged molecular mass of 14,940.7 Da, in perfect agreement with the expected value (14,939.5 Da; data not shown).

Gel filtration also provided a measurement of the aggregation state for native and recombinant loc-OBP1 under native conditions. In both cases, an apparent molecular mass of 30 kDa was observed, demonstrating the dimeric nature of the protein in solution.

Disulfide determination in native and recombinant loc-OBP1

A MALDI-MS mapping approach was used to investigate the disulfide pairing of the native protein and to verify the nature of the reoxidized cysteine residues in the recombinant one. Fig. 3A shows the spectrum obtained for the native loc-OBP1 following digestion with trypsin. The signals at m/z 4131.6, 4153.4, 4740.3, and 4761.9 were associated with the peptides (16–22) + (52–80) and (16–22) + (52–85) of both isoforms, thus demonstrating the occurrence of an S-S bond connecting Cys21 and Cys52. This disulfide was confirmed by the additional presence in the spectrum of the signals at m/z 1890.9, 1913.4, 3673.6, and 4281.6 that were ascribed to the peptides (16–22) + (52–60), (20–22) + (52–80), and (20–22) + (52–85), originating from the aspecific hydrolysis at His19 and Leu60. Furthermore, the intense MH^+ signal at *m/z* 4981.9 was assigned to a three-peptide cluster involving fragments (35–51), (86–91) and (92–110) connected by two disulfide bridges between Cys48, Cys90, Cys99, and Cys108. Finally, the signal at 3411.7 was interpreted as arising from peptides (35–51) and (92–103) linked by an S-S bond between Cys48-Cys99. This peak originated from the peptide at m/z 4981.9 following aspecific hydrolysis at Tyr103. The nature of all disulfide-containing peaks was confirmed by MALDI-MS analysis of the peptide mixture following reduction with DTT and Edman degradation analysis of these components following HPLC purification. These data definitively proved that the disulfides pairing in *L. migratoria* OBP1 is Cys21-Cys52, Cys48-Cys99 and Cys90-Cys108, identical to the pattern previously reported for OBP and PBP from *B. mori* [6, 7].

MALDI-MS analysis of the recombinant OBP tryptic digest showed the same signals present in the spectrum of the natural species (fig. 3). In this case, as expected, the peaks originating from the second isoform occurring in the natural sample were absent. Signals associated with peptides originating from an incorrect pairing of the cysteine residues were not detected, demonstrating the correct refolding process of the recombinant OBP1.

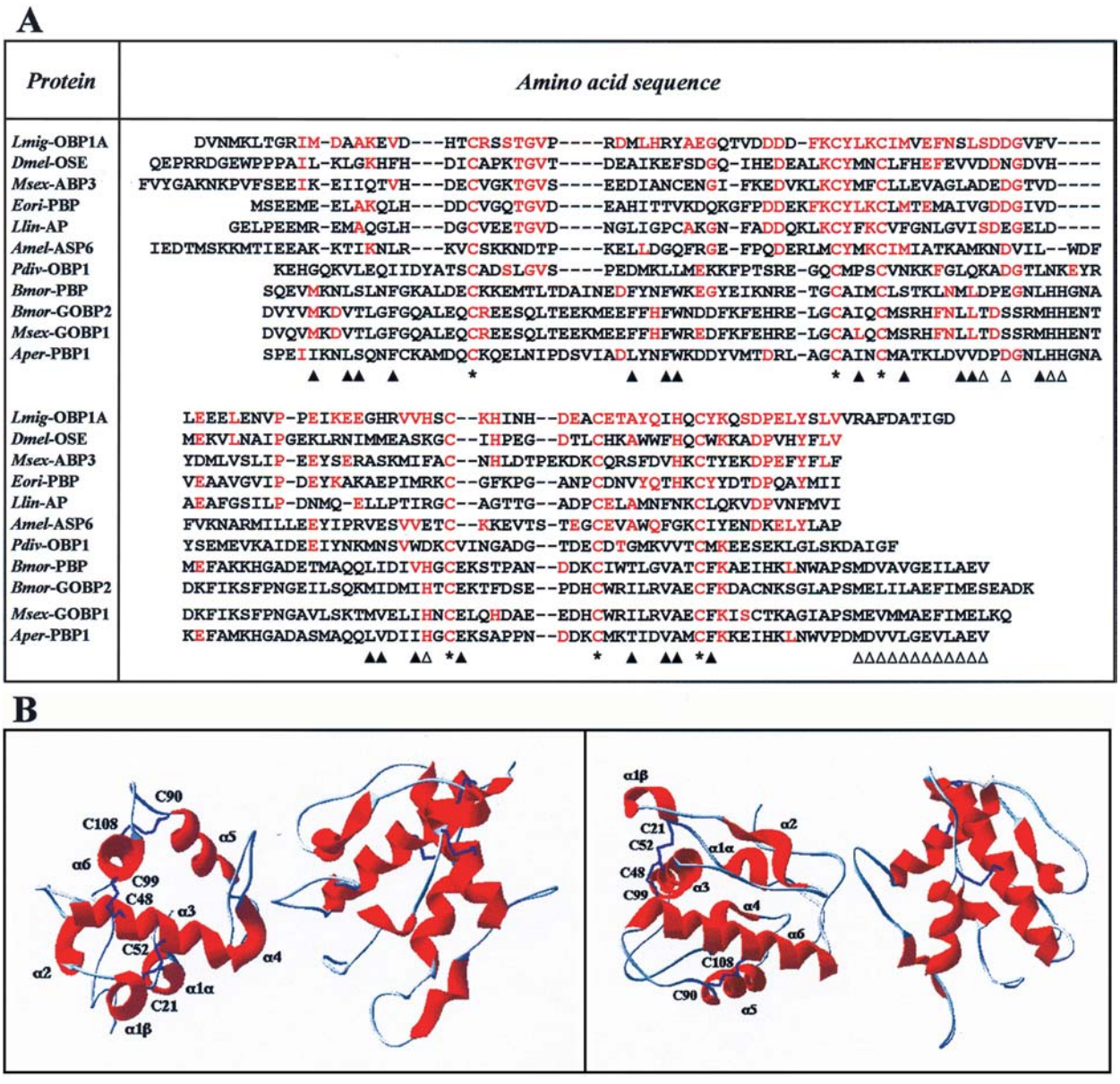

Figure 2. Sequence analysis and molecular modeling of locust OBP1. (*A*) Amino acid sequence of locust OBP1-A aligned with those of other insect OBPs, representative of different orders. Residues conserved in all sequences are marked by asterisks; those common to *Lmig*-OBP1-A are shown in red. Amino acids involved in ligand binding or modulating the pH-induced conformational change according to the structure of the *Bmor*-PBP/bombykol complex are marked by ▲ and △, respectively [5, 14]. *Lmig*, *Locusta migratoria*; *Dmel*, *Drosophila melanogaster* (OSE, 22% identity); Msex, *Manduca sexta* (ABP3, 20% identity; GOBP1, 12% identity); *Eori*, *Exomala orientalis* (PBP, 30% identity); *Llin*, *Lygus lineolaris* (LAP, 22% identity); *Amel*, *Apis melliphera* (ASP6, 22% identity); *Pdiv*, *Phylloperta diversa* (OBP1, 16% identity); *Bmor*, *Bombyx mori* (PBP, 13% identity; GOBP2, 11% identity); *Aper*, *Antheraea pernyi* (PBP1, 12% identity). (*B*) A three-dimensional model of dimeric locust OBP1. Two different views of the dimer are reported. Helices are numbered and shown in red; disulfides are colored in blue; the residues involved in S-S bridges are indicated.

Fluorescence binding assays

The fluorescent reporter 1-NPN, when excited at the wavelength of 337 nm, produces a very weak fluorescent peak at 480 nm. In the presence of loc-OBP1, a strong blue shift was observed, with an intense fluorescence peak at 407 nm. The intensity of this peak was used to measure the amount of 1-NPN bound to the protein (fig. 4A). Such binding was saturable with a single population of binding sites, without apparent cooperativity effect, and a dissociation constant of 1.67 µM (fig. 4B). Binding to 1-AMA was also saturable, but the blue shift and the increase in fluorescence intensity were too small for reliable evaluation of the binding constant. Interestingly, both 2-NPN and 2-AMA, isomers of the above ligands, failed to show significant changes in their spectra in the presence of OBP.

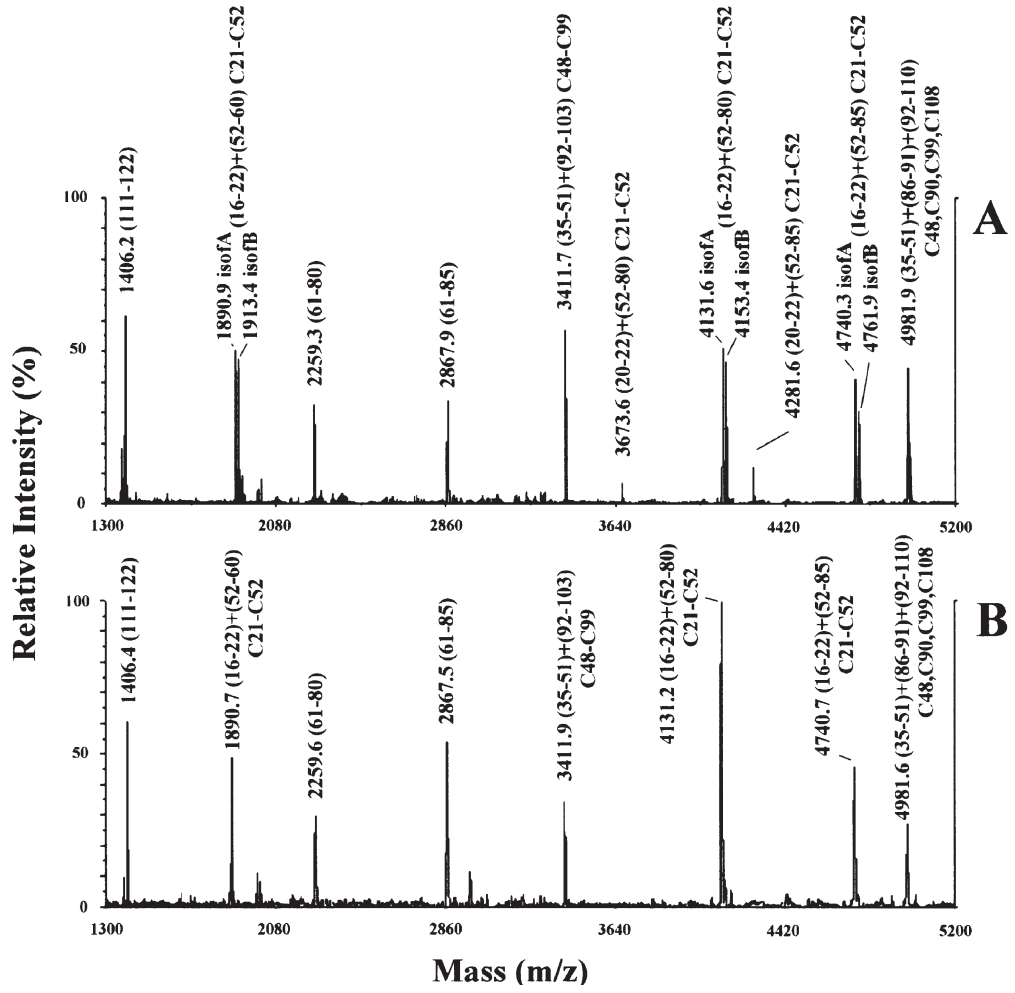

Figure 3. MALDI-MS analysis of native (*A*) and recombinant (*B*) loc-OBP1 following digestion with trypsin. Signals corresponding to the disulfide-bridged peptides are shown; each signal is assigned to the corresponding peptide pair and the two cysteine residues involved in the S-S bond are indicated.

We also decided to investigate the thermal stability of loc-OBP1 by measuring the intensity of the emission peak at 407 nm. Spectra were therefore recorded with the protein incubated at temperatures up to 100 °C for up to 20 min, as described in Materials and methods. All the samples, after cooling to room temperature, fully recovered their binding activity.

To investigate the specificity of the binding site of OBP1, we used a number of potential ligands either in direct fluorescent assays or in competitive binding experiments. A first series included several fluorescent amines and phenols. 1-Naphthol and 2-naphthol, employed with major urinary proteins [33], lipocalins, which are structurally similar to OBPs, as well as both naphthylamines did not show any change in their fluorescence spectra in the presence of loc-OBP1. The absence of any shift in their emission peaks could indicate a poor or no affinity of these compounds for the protein, but could also be the result of binding not accompanied by modification of their optical properties.

As the best ligand for loc-OBP1, among those tested so far, we used 1-NPN fluorescent probe in competitive binding experiments together with several potential ligands. The choice of ligands included phenylacetonitrile, identified as a pheromone component of the related species *S. gregaria* [34] and oleoamide, which was found as the major compound bound to CSPs purified from the wings of *L. migratoria* [L. P. Ban, A. Scaloni, A. Brandazza, S. Angeli, L. Zhang, Y. H. Yan, et al., unpublished data]. Apart from these two compounds of potential ecological interest, the other competitors were randomly chosen, as representative of different types of chemical and odor classses. Aliphatic derivatives included alcohols (2-hexanol, 3,7-dimethyloctanol, and 1-dodecanol), ketones (2-undecanone and 6-undecanone), 1-aminododecane, decanal, and palmitic acid. Aromatic compounds were represented by butyrophenone, eugenol, thymol, diphenylamine, α-amylcinnamaldehyde, and N-octadecyl-1-aminoanthracene. Most of the compounds exam-

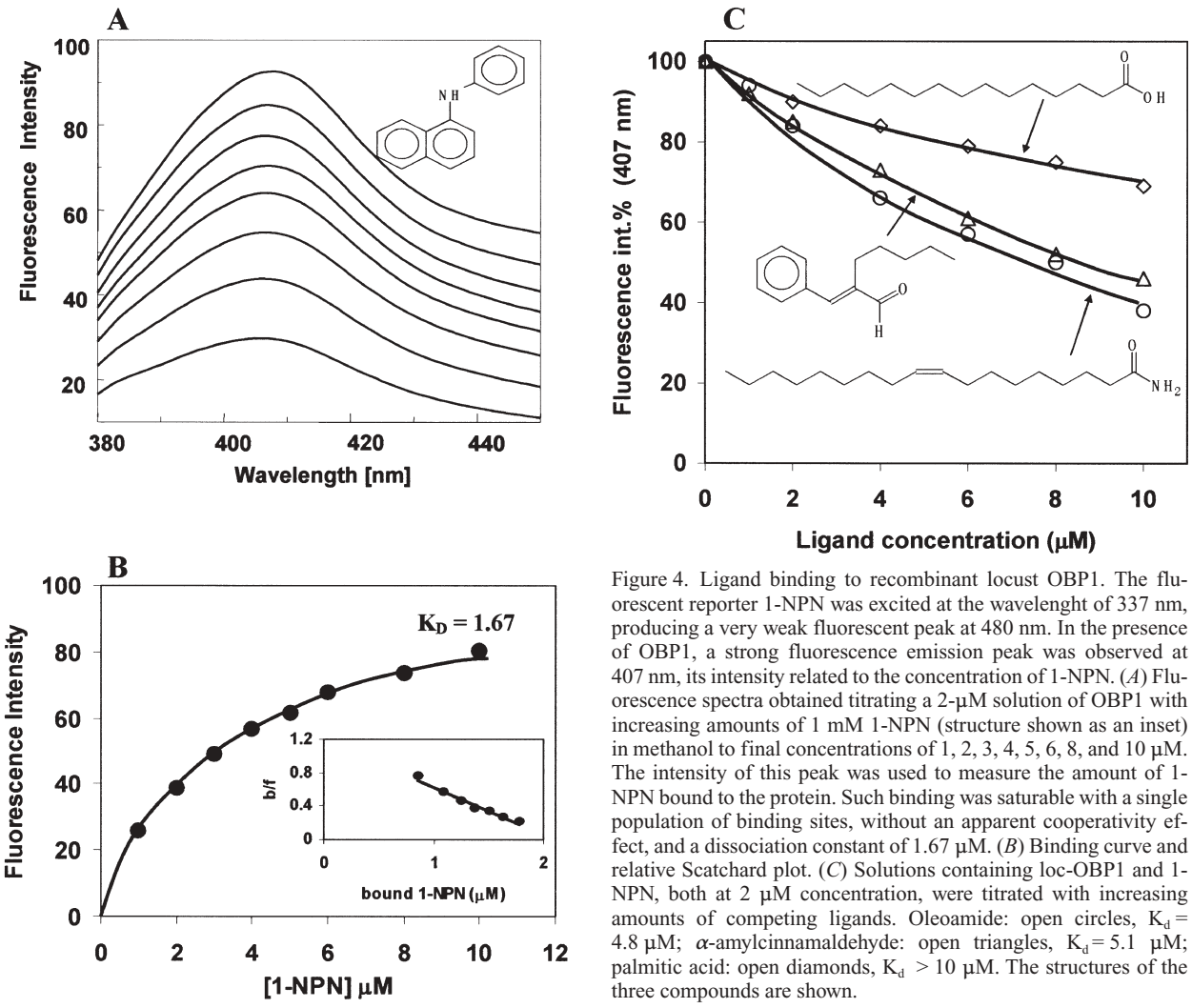

Figure 4. Ligand binding to recombinant locust OBP1. The fluorescent reporter 1-NPN was excited at the wavelenght of 337 nm, producing a very weak fluorescent peak at 480 nm. In the presence of OBP1, a strong fluorescence emission peak was observed at 407 nm, its intensity related to the concentration of 1-NPN. (*A*) Fluorescence spectra obtained titrating a 2-μM solution of OBP1 with increasing amounts of 1 mM 1-NPN (structure shown as an inset) in methanol to final concentrations of 1, 2, 3, 4, 5, 6, 8, and 10 μM. The intensity of this peak was used to measure the amount of 1-NPN bound to the protein. Such binding was saturable with a single population of binding sites, without an apparent cooperativity effect, and a dissociation constant of 1.67 μM. (*B*) Binding curve and relative Scatchard plot. (*C*) Solutions containing loc-OBP1 and 1-NPN, both at 2 μM concentration, were titrated with increasing amounts of competing ligands. Oleoamide: open circles, K_d = 4.8 μM; α-amylcinnamaldehyde: open triangles, K_d = 5.1 μM; palmitic acid: open diamonds, K_d > 10 μM. The structures of the three compounds are shown.

ined were inefficient in displacing the fluorescent probe 1-NPN from its complex with loc-OBP1. The best ligands proved to be oleoamide and α-amylcinnamaldehyde, with dissociation constant of 4.8 and 5.1 μM, respectively. Palmitic acid, decanal, and N-octadecyl-1-aminoanthracene were able to displace only 20–30% of bound 1-NPN, when present at a 10 μM concentration. Figure 4C reports the competition curves for the first three compounds mentioned above.

Molecular modeling

A three-dimensional model of the dimeric protein was constructed on the basis of the crystallographic structure of the dimeric *B. mori* PBP/bombykol complex. Sequence identity between the two polypeptides was about 13%. In addition, the locust protein presented a longer N terminus and a shorter C-terminal tail with respect to the *B. mori* counterpart. However, the overall structure of the two dimers was similar and the dimerization interface consisted of homologous regions (fig. 2B). In particular, the polypeptide segments 58–69 (monomer 1) and 117–122 (monomer 2) were present at the dimerization interface in *L. migratoria* OBP1. Moreover, the loc-OBP1 monomer was similar to the *B. mori* PBP monomer, with a tightly packed arrangement of six α helices consisting of the residues 12–17 (α1a), 21–26 (α1b), 31–34 (α2), 44–57 (α3), 68–72 (α4), 81–89 (α5) and 97–116 (α6). This packing was stabilized by conserved disulfides determining a tightening of α1 and α6 over α3 (Cys21-Cys52, Cys48-Cys99) and α5 over α6 (Cys90-Cys108). Different van der Waals interactions also contributed to enforce packing between helices. Loop regions between α1a-α1b, α1β-α2, α3-α4, and α5-α6 were shorter in the locust protein with respect to *B. mori* PBP, and a scaffold of four antiparallel helices (α1, α3, α5, and α6) formed the ligand-binding site that is closed, on one side, by helix α3. Residues in loc-OBP1 hypothetically involved in binding to the ligand are Leu60 and Val87 (conserved with respect to *B. mori* PBP), and Ile10, Ala13, Ala14, Val17, Met31, Arg34,

Tyr35, Tyr49, Leu50, Met54, Phe57, Glu70, Ile105, His106, and Tyr109. All these amino acids presented their side chains oriented toward the inner cavity of the protein and able to interact with the hydrophobic ligand. Changes in the nature of the amino acid at position 19 did not modify the overall structure of the protein and probably should not influence loc-OBP1 binding specificity, being an amino acid located far from the binding pocket.

Discussion

We have reported the first OBP identified and characterized from Orthoptera. As expected, its sequence similarity with other insect OBPs is poor (between 11 and 30% identity), but significant for assigning this protein to the OBPs family. This finding is further confirmed by the presence of six conserved cysteines, establishing three disulfide bridges in the same fashion as in the OBPs of *B. mori* [6, 7] and *Apis mellifera* [18].

Only two isoforms, differing by a single substitution (His19→Asn), were detected both at the protein and RNA level in *L. migratoria*. However, at this time we cannot exclude in this animal the presence of other OBPs with different N-terminal sequences. In fact, in other insect orders, such as Lepidoptera, Diptera, and Imenoptera, several sub-classes of OBPs have been reported, each represented by a few sequences. Therefore, we can hypothezise that in Orthoptera also, more OBP types remain to be discovered. On the other hand, CSPs are represented in Orthoptera by a great number of sequences, belonging to several different sub-families [10, 35]. Locust OBP1 was easily expressed in good yields in a bacterial system but, unlike CSPs, was produced entirely in its insoluble form. The protein was solubilized in the presence of high concentrations of urea, but it precipitated again during dialysis. The only way to obtain loc-OBP1 in its soluble form was to reduce its disulfide bridges. However, by air oxidation, the denatured protein reestablished its correct cysteine pairing, as demonstrated by MALDI-MS analysis. Moreover, the renatured protein showed binding activity with a saturable curve and a single type of binding site. The best ligand, among the fluorescent probes tested, proved to be 1-NPN, also shown as a good ligand for CSPs from *S. gregaria* [36], as well as for the mouse urinary proteins [37]. The recombinant loc-OBP1 also bound 1-AMA, previously employed with other insect OBPs [27], but less efficiently.

Competitive binding experiments provided some indication of the specificity of this OBP. Since locust pheromones have not yet been identified, we tested phenylacetonitrile, one of the aggregation pheromone components of the desert locust *S. gregaria* [34]. This compound, like several other randomly chosen among different chemical classes, was completely ineffective. Also medium- and long-chain aliphatic compounds, such as 1-dodecanol, 1-dodecyl amine, and palmitic acid, proved to be poor ligands. In contrast, oleoamide was the only compound able to displace more than 50% of 1-NPN from loc-OBP1, when used at concentrations below 10 µM. This molecule was found to be strongly associated with another class of binding proteins that we recently purified from the wings of *L. migratoria* (Ban et al., unpublished data). Comparison of the structural features of these ligands does not establish a clear relationship between molecular structure and their binding efficiency to loc-OBP1. However, large size seems to be required for effective binding, as indicated by the better affinity of oleoamide (18 carbon atoms) than palmitic acid (16 carbon atoms) or 1-dodecanol. The fluorescent probe 1-NPN also contains 18 carbons, suggesting that long-chain compounds, such as oleoamide and palmitic acid, should adopt a non-extended conformation when interacting with the protein binding site. The hypothesis that loc-OBP1 could bind long-chain derivatives was not further tested with compounds such as alcohols and hydrocarbons, due to their limited solubility in water.

A model of locust OBP1, built on the basis of the structure of *B. mori* PBP, allows formulation of some hypotheses on the mode of action of this polypeptide, compared with its lepidopteran protein homologue. Recently, comparative analysis of the bombykol/PBP and unbound PBP structures suggested that a pH-induced conformational change in this protein can modulate ligand binding by destabilizing salt bridges between Asp63, Glu65, His69, His70, His95, and Glu98 [14, 38]. Furthermore, pH regulation of protein function should be associated with a structural mechanism involving a conformational change of the C-terminal dodecapeptide segment, which is in an extended conformation and located on the protein surface in the pheromone complex, but forms a regular helix located in the bombykol-binding site in the protein devoid of the ligand. Therefore, physiological pheromone binding and release should be modulated by the pH difference between the bulk sensillar lymph and membrane surface environment. These structural elements, generally well conserved in lepidopteran OBPs, are not present in many homologous proteins from other insect orders (fig. 2A). In fact, the C-terminal end is often absent and the charged residues controlling this conformational variation are not present in these species, supporting the idea of a different mechanism for the regulation of ligand binding. In this respect, locust OBP1 with its long C terminus with two conserved charged residues seems more similar to the *B. mori* PBP. Therefore, on this basis, a similar conformational change should also be hypothesized for this protein, probably modulated by other amino acids present in the polypeptide chain. A more detailed description of the binding mechanism of locust OBP1 will be gained after identifying the structure of its physiological ligand.

Acknowledgements. L. B. thanks the European Chemoreception Research Organization for a travel grant to Italy and the University of Pisa for sponsoring her visit. Work was partly supported by PRIN Project, MURST, Italy, by the Foundation for University Key Lab Teacher of the China Ministry of Education, and by grants from the Italian National Research Council.

1 Pelosi P. (1998) Odorant-binding proteins: structural aspects. Ann. N. Y. Acad. Sci. **855:** 281–293
2 Steinbrecht R. A. (1998) Odorant-binding proteins: expression and function. Ann. N. Y. Acad. Sci. **855:** 323–332
3 Bianchet M. A., Bains G., Pelosi P., Pevsner J., Snyder S. H., Monaco H. L. et al. (1996) The three dimensional structure of bovine odorant-binding protein and its mechanism of odor recognition. Nat. Struct. Biol. **3:** 934–939
4 Tegoni M., Ramoni R., Bignetti E., Spinelli S. and Cambillau C. (1996) Domain swapping creates a third putative combining site in bovine odorant binding protein dimer. Nat. Struct. Biol. **3:** 863–867
5 Sandler B. H., Nikonova L., Leal W. S. and Clardy J. (2000) Sexual attraction in the silkworm moth: structure of the pheromone-binding-protein-bombykol complex. Chem. Biol. **7:** 143–151
6 Leal W. S., Nikonova L. and Peng G. (1999) Disulfide structure of the pheromone binding protein from the silkworm moth, *Bombyx mori*. FEBS Lett. **464:** 85–90
7 Scaloni A., Monti M., Angeli S. and Pelosi P. (1999) Structural analyses and disulfide-bridge pairing of two odorant-binding proteins from *Bombyx mori*. Biochem. Biophys. Res. Commun. **266:** 386–391
8 McKenna M. P., Hekmat-Scafe D. S., Gaines P. and Carlson J. R. (1994) Putative Drosophila pheromone-binding proteins expressed in a subregion of the olfactory system. J. Biol. Chem. **269:** 16340–16347
9 Pikielny C. W., Hasan G., Rouyer F. and Rosbach M. (1994) Members of a family of *Drosophila* putative odorant-binding proteins are expressed in different subsets of olfactory hairs. Neuron **12:** 35–49
10 Angeli S., Ceron F., Scaloni A., Monti M., Monteforti G., Minnocci A. et al. (1999) Purification, structural characterization, cloning and immunocytochemical localization of chemoreception proteins from *Schistocerca gregaria*. Eur. J. Biochem. **262:** 745–754
11 Marchese S., Angeli S., Andolfo A., Scaloni A., Brandazza A., Mazza M. et al. (2000) Soluble proteins from chemosensory organs of *Eurycantha calcarata* (Insecta, Phasmatodea). Insect Biochem. Mol. Biol. **30:** 1091–1098
12 Jacquin-Joly E., Vogt R. G., Francois M. C. and Nagnan-Le Meillour P. (2001) Functional and expression pattern analysis of chemosensory proteins expressed in antennae and pheromonal gland of *Mamestra brassicae*. Chem. Senses **26:** 833–844
13 Xu A., Park S. K., D'Mello S., Kim E., Wang Q. and Pikielny W. (2002) Novel genes expressed in subsets of chemosensory sensilla on the front legs of male *Drosophila melanogaster*. Cell Tissue Res. **307:** 381–392
14 Horst R., Damberger F., Luginbuhl P., Guntert P., Peng G., Nikonova L. et al. (2001) NMR structure reveals intramolecular regulation mechanism for pheromone binding and release. Proc. Natl. Acad. Sci. USA **98:** 14374–14379
15 Biessmann H., Walter M. F., Dimitratos S. and Woods D. (2002) Isolation of cDNA clones encoding putative odourant binding proteins from the antennae of the malaria-transmitting mosquito, *Anopheles gambiae*. Insect Mol. Biol. **11:** 123–132
16 Ishida Y., Cornel A. J. and Leal W. S. (2002) Identification and cloning of a female antenna-specific odorant-binding protein in the mosquito *Culex quinquefasciatus*. J. Chem. Ecol. **28:** 867–871
17 Danty E., Arnold G., Huet J. C., Huet D., Masson C. and Pernollet J. C. (1998) Separation, characterization and sexual heterogeneity of multiple putative odorant-binding proteins in the honeybee *Apis mellifera* L. (Hymenoptera: Apidea). Chem. Senses **23:** 83–91
18 Briand L., Nespoulous C., Huet J. C. and Pernollet J. C. (2001) Disulfide pairing and secondary structure of ASP1, an olfactory-binding protein from honeybee (*Apis mellifera* L). J. Pept. Res. **58:** 540–545
19 Briand L., Nespoulous C., Huet J. C., Takahashi M. and Pernollet J. C. (2001) Ligand binding and physico-chemical properties of ASP2, a recombinant odorant-binding protein from honeybee (*Apis mellifera* L.). Eur. J. Biochem. **268:** 752–760
20 Krieger M. J. and Ross K. G. (2002) Identification of a major gene regulating complex social behavior. Science **295:** 328–332
21 Wojtasek H., Hansson B. S. and Leal W. S. (1998) Attracted or repelled? A matter of two neurons, one pheromone binding protein, and a chiral center. Biochem. Biophys. Res. Commun. **250;** 217–222
22 Wojtasek H., Picimbon J. F. and Leal W. S. (1999) Identification and cloning of odorant binding proteins from the scarab beetle *Phyllopertha diversa*. Biochem. Biophys. Res. Commun. **263:** 832–837
23 Dickens J. C., Callahan F. E., Wergin W .P., Murphy C. A. and Vogt R. G. (1998) Odorant-binding proteins of true bugs: generic specificity, sexual dimorphism, and association with subsets of chemosensory sensilla. Ann. N. Y. Acad. Sci. **855:** 306–310
24 Vogt R. G., Callahan F. E., Rogers M. E. and Dickens J. C. (1999) Odorant binding protein diversity and distribution among the insect orders, as indicated by LAP, an OBP-related protein of the true bug *Lygus lineolaris* (Hemiptera, Heteroptera). Chem. Senses **24:** 481–495
25 Ishida Y., Chiang V. P., Michael I. H. and Leal W. S. (2002) Odorant-binding proteins from a primitive termite. J. Chem. Ecol. **28:** 1887–1893
26 Vogt R. G. and Riddiford L. M. (1981) Pheromone binding and inactivation by moth antennae. Nature **293:** 161–163
27 Campanacci V., Krieger J., Bette S., Sturgis J. N., Lartigue A., Cambillau C. et al. (2001) Revisiting the specificity of *Mamestra brassicae* and *Antheraea polyphemus* pheromone-binding proteins with a fluorescence binding assay. J. Biol. Chem. **276:** 20078–20084
28 Kim M. S., Repp A. and Smith D. P. (1998) LUSH odorant-binding protein mediates chemosensory responses to alcohols in *Drosophila melanogaster*. Genetics **150:** 711–721
29 Barton G. J. (1990) Protein multiple sequence alignment and flexible pattern matching. Methods Enzymol. **183:** 403–428
30 Barton G. J. and Sternberg M. J. E. (1987) Evaluation and improvements in the automatic alignment of protein sequenes. Protein Eng. **1:** 89–94
31 Doreleijers J. F., Vriend G., Raves M. L. and Kaptein R. (1999) Validation of nuclear magnetic resonance structures of proteins and nucleic acids: hydrogen geometry and nomenclature. Proteins **37:** 404–416
32 Guex N. and Peitsch M. C. (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis **18:** 2714–2723
33 Sartor G., Pagani R., Ferrari E., Sorbi R. T., Cavaggioni A., Cavatorta P. et al. (2001) Determining the binding capability of the mouse major urinary proteins using 2-naphthol as a fluorescent probe. Anal. Biochem. **292:** 69–75
34 Schmidt G. H. (1997) Pheromones in *Schistocerca gregaria* (Forsk): the present situation. In: New Strategies in Locust Control, pp. 313–321, Krall S., Peveling R. and Badiallo D. (eds), Birkhauser, Basel

35 Picimbon J. F., Dietrich K., Breer H. and Krieger J. (2000) Chemosensory proteins of *Locusta migratoria* (Orthoptera: Acrididae). Insect Biochem. Mol. Biol. **30:** 233–241
36 Ban L. P., Zhang L., Yan Y. H. and Pelosi P. (2002) Binding properties of a locust's chemosensory protein. Biochem. Biophys. Res. Commun. **293:** 50–54
37 Marie A. D., Veggerby C., Robertson D. H. L., Gaskell S. J., Hubbard S. J., Martinsen L. et al. (2001) Effect of polymorphisms on ligand binding by mouse major urinary proteins. Protein Sci. **10:** 411–417
38 Damberger F., Nikonova L., Horst R., Peng G., Leal W. S. and Wuthrich K. (2000) NMR characterization of a pH-dependent equilibrium between two folded solution conformations of the pheromone-binding protein from *Bombyx mori*. Protein Sci. **9:** 1038–1041