Solution structure of a chemosensory protein from the moth *Mamestra brassicae*

Amor MOSBAH¹, Valérie CAMPANACCI¹, Audrey LARTIGUE, Mariella TEGONI, Christian CABBILLAU² and Hervé DARBON²

AFMB, UMR 6098-CNRS and Universités Aix-Marseille I & II, 31 Chemin J. Aiguier, 13402 Marseille Cedex 20, France

Chemosensory proteins (CSPs) are believed to be involved in chemical communication and perception. A number of such proteins, of molecular mass ~ 13 kDa, have been isolated from different sensory organs of a wide range of insect species. Several CSPs have been identified in the antennae and proboscis of the moth *Mamestra brassicae*. CSPMbraA6, a 112-amino-acid antennal protein, has been expressed in a soluble form in large quantities in the *Escherichia coli* periplasm. NMR structure determination of CSPMbraA6 has been performed with ¹H- and ¹⁵N-labelled samples. The calculated structures present an average root mean square deviation about the mean structure of 0.63 Å for backbone atoms and 1.27 Å for all non-hydrogen atoms except the 12 N-terminal residues. The protein is well

folded from residue 12 to residue 110, and consists of a non-bundle α -helical structure with six helices connected by $\alpha\alpha$ loops. It has a globular shape, with overall dimensions of 32 Å $\times$ 28 Å $\times$ 24 Å. A channel is visible in the hydrophobic core, with dimensions of 3 Å $\times$ 9 Å $\times$ 21 Å. In some of the 20 solution structures calculated, this channel is closed either by Trp-94 at one end or by Tyr-26 at the other end; in some other solutions, this channel is closed at both ends. Binding experiments with 12-bromododecanol indicate that the CSPMbraA6 structure is modified upon ligand binding.

Key words: conformational change, membrane, NMR, pheromone-binding protein.

INTRODUCTION

In moths, the organs devoted to olfactory perception have been identified in antennae as the sensilla trichodea and basiconica, the former being tuned to the perception of pheromones [1]. Typically, lepidopteran pheromones are volatile hydrophobic compounds comprising 12–20 carbon atoms. Due to the hydrophobic character of pheromones and odours, a general scheme was proposed in which transport of the volatile molecule from the air/sensilla interface to the olfactory receptor requires a carrier, which could also discriminate among different volatile molecules [2–6].

A class of small proteins with an average molecular mass of 13 kDa has been identified in the antennae of *Drosophila melanogaster* and in several sensorial organs from a wide range of species of the insect order [7–19]; they were therefore called chemosensory proteins (CSPs). CSPs do not belong to the pheromone-binding protein (PBP) or general odorant-binding protein class, since they contain only four conserved cysteine residues that form two disulphide bridges, and share no sequence similarity with them. Two classes of CSP have been reported in the literature which differ in their primary sequence [16–18]. CSPs have been proposed to be involved in CO₂ detection [8], in chemical signal transmission in regenerating legs [10] and in chemoperception (olfaction and taste), due to their localization in antennae, tarsi and labrum [14,20]. They may also play a role in the transport of hydrophobic chemicals (volatile or not) from air or water to olfactory or taste receptors, in a similar way to PBPs. However, the exact physiological role of CSPs has still to be identified. In the moth *Mamestra brassicae*, several CSPs have been identified in the proboscis [20] and in the antennae [21]. *M. brassicae* CSPs have been shown to bind several components of the pheromonal blend, and therefore are believed to have a

function analogous to that of PBPs [20]. In the proboscis, however, a putative role as odour or taste carriers has been assigned to CSPs.

We have expressed CSPMbraA6, originating from the antennae of *M. brassicae*, as a recombinant protein in the periplasm of *Escherichia coli* [22]. The NMR resonance assignment of CSPMbraA6 has already been published [22]. In the present paper, we describe the three-dimensional NMR structure of CSPMbraA6, which displays a novel fold, and ¹⁵N heteronuclear single-quantum coherence (HSQC)-based binding experiments with a surrogate ligand, 12-bromododecanol.

MATERIAL AND METHODS

Production of CSPMbraA6

A detailed description of the molecular cloning, expression, purification and characterization of CSPMbraA6 has been reported previously [22].

NMR spectroscopy

Experimental restraints

NMR spectral analysis was performed using XEASY software [23]. The identification of amino acid spin systems and the sequential assignment were performed by using the standard strategy described by Wüthrich [24]. ¹H and ¹⁵N chemical shift assignments for CSPMbraA6 have already been described [22]. Nuclear Overhauser effects (nOes) were collected from a NOESY spectrum recorded with a mixing time of 100 ms (checked for the absence of spin diffusion). The integration of nOe data using

Abbreviations used: CNS, crystallography and NMR system; CSP, chemosensory protein; HSQC, heteronuclear single-quantum coherence; LTP, lipid transfer protein; nOe, nuclear Overhauser effect; PBP, pheromone-binding protein; RMSD, root mean square deviation.

¹ These authors contributed equally to this work.

² To whom correspondence should be addressed (e-mail cambillau@afmb.cnrs-mrs.fr and darbon@afmb.cnrs-mrs.fr).

Figure 1 Repartition of intra-residue (white), sequential (light grey), medium-range (dark grey) and long-range (black) nOes

manual integration in the XEASY software [23] made it possible to determine peak volumes and to convert them into distance restraints using the CALIBA routine of the DIANA package [25], according to calculated curves, with a function of $1/r^6$ for intraresidual and sequential nOes involving HN, H α and H β , as well as medium- and long-range nOes involving backbone protons, and a function of $1/r^4$ for others. A set of restraints was thus obtained with upper distance limits defined for each nOe. The lower limit was systematically set to 1.8 Å.

Structure calculations

Distance geometry calculations were performed using the variable target function program DIANA [25]. A preliminary set of 1000 structures was initiated that included only intraresidual and sequential upper-limit distances. From these, the 500 best structures were kept for a second round, including medium-range distances, and the resulting 250 best structures were then used in

a third round with the whole set of upper-limit restraints. Starting from the 100 best structures, according to their target function, a REDAC (redundant dihedral angle constraints) [25] strategy was used in the last step in order to include the dihedral restraints together with the additional distance restraints coming out from hydrogen bonds. Final energy refinement was achieved by CNS (crystallography and NMR system) [26]. Visual analysis of the resulting structures was carried out with TURBO [27], as well as with MOLMOL [28]. Root mean square deviation (RMSD) data were obtained using MOLMOL and DYANA [29]. Analysis of the quality of the structure was performed using PROCHECK-NMR [30]. Structural comparison of CSPMbraA6 with the structures of proteins already deposited in the protein data bank was performed using DALI [31]. The atomic coordinates of the 20 best CSPMbraA6 structures have been deposited in the Protein Data Bank at RCSB (<http://www.rcsb.org/pdb/>) as entry 1K19.

RESULTS

NMR spectral analysis and sequential assignment of recombinant CSPMbraA6

The comparative analysis of COSY and TOCSY spectra recorded in water at 296 K gave the spin system signature of the protein and is described elsewhere [22]. The spin systems were connected sequentially using the NOESY spectra [22]. We did not use stereospecific assignments, as most of the results obtained were ambiguous and therefore useless.

Experimental restraints

We have calculated the nOe-based structure of CSPMbraA6 using a set of restraints comprising 1484 nOes (482 intraresidue, 331 sequential, 280 medium-range and 301 long-range restraints), which were converted into distance restraints according to their intensities. The intra- and inter-residual restraints, as well as the medium-range ones, are uniformly distributed along the

Figure 2 Stereo view of the 20 best structures of CSPMbraA6 superimposed for the best fit

Only C α atoms are displayed.

sequence, which reveals a high density of secondary structures (Figure 1). The long-range restraints mainly involve hydrophobic and aromatic residues. Amide protons still giving rise to nOe cross-peaks after 5 h of exchange were considered as being slowly exchangeable and therefore engaged in hydrogen bonds with carbonyl oxygen partners that were determined by visual analysis of the preliminary calculated structures and confirmed by the hydrogen bonds proposed by the DIANA software [25]. Hydrogen bonds were translated into distance restraints ($d_{\text{NH-O}} = 2.0 \text{ \AA}$ and $d_{\text{N-O}} = 3.0 \text{ \AA}$ for slowly exchangeable NH) and included in the structure calculation. During the course of nOe collection, we obtained the X-ray structure of CSPMbraA6 [32], which allowed us to confirm the conformation of helix D, for which no medium-range nOes involving Gly-65 were detected. We added a hydrogen bond, i.e. NH69–CO65, that was not detected by the exchange experiment. We also solved one nOe assignment ambiguity. In fact, due to resonance degeneracy, this nOe was previously assigned as H α 30–H β 166 and subsequently corrected to H α 30–H γ 38.

Structure calculation

The structure calculation of CSPMbraA6 was achieved using the default parameters of DIANA [25]. Purging the final experimental distance restraints from the meaningless restraints led to 12.5 nOe restraints per residue on average. In addition, the final step of the calculation contained 84 hydrogen-bond restraints and six disulphide-bond restraints. Altogether, the final experimental set corresponds to 13.2 restraints per residue on average. Starting from the 50 best solutions obtained, a final energy refinement was achieved by CNS [27] with quadratic van der Waals repulsion terms and square well quadratic potential terms for inter-proton distance included in the force field. CNS was used without a simulated solvent force field, as usual for refinement against experimental data. The van der Waals hard-sphere radii were set to 0.78 times the CHARMM values. The final 20 structures were obtained after 2000 cycles of restrained Powell energy minimization. The best-fit superimposition (Figure 2) of backbone atoms for the 20 models gives RMSD values of $0.63 \pm 0.39 \text{ \AA}$ and $1.27 \pm 0.40 \text{ \AA}$ if all non-hydrogen atoms were included except those in the N-terminal region (Glu-1 to Asn-12) (Table 1).

A stereo-pair representation of the best-fit superimposition of the backbone atoms for the 20 converged structures is shown in Figure 2. Analysis of local RMSD values (not shown) indicates that the precision of the calculation is fairly constant all along the sequence, except for the N-terminal region and the region comprising residues 51–57. All of the solutions have a good non-bonded contact and good covalent geometry, as indicated by the low values of CNS energy terms (Table 1). The correlation with the experimental data shows no nOe-derived distance violation greater than $0.43 \text{ \AA}$ (two violations between $0.2 \text{ \AA}$ and $0.5 \text{ \AA}$).

Analysis of the Ramachandran plot shows 90% of residues in the most favoured region and 10% of residues in the generously allowed region (PROCHECK nomenclature [30]). The RMSD distribution along the protein polypeptide backbone shows that the structure is well defined in all regions except the N-terminal region (residues 1–13) and the region comprising residues 51–57. The largest RMSDs were observed for the N-terminal region (Glu-1–Asp-13; $< 2 \text{ \AA}$) and the region from Ile-51 to Lys-57 ($< 1 \text{ \AA}$). The global fold is well defined, because of the presence of a large number of medium-range restraints [$1 < |i-j| < 5$] (280 observations) and long-range restraints [$|i-j| < 4$] (301 observations). Selection of the 20 best solution structures was done after CNS refinement according to the number of residual nOe

Table 1 Structural statistics for the 20 best structures of CSPMbraA6

(DG) indicates values for the final 20 CSPMbraA6 structures obtained by distance geometry and energy minimization. [DG] indicates values for the mean structure obtained by averaging the co-ordinates of the individual distance geometry structures best fitted to each other. The square-well nOe (E_{nOe}) was calculated with a force constant of $314 \text{ kJ} \cdot \text{mol}^{-1}$ ($75 \text{ kcal} \cdot \text{mol}^{-1}$). The square-well torsional angle E_{tor} was calculated with a force constant of $837 \text{ kJ} \cdot \text{mol}^{-1}$ ($200 \text{ kcal} \cdot \text{mol}^{-1}$). The quadratic van der Waals term was calculated with a force constant of $16.7 \text{ kJ} \cdot \text{mol}^{-1}$ ($4 \text{ kcal} \cdot \text{mol}^{-1}$), and the van der Waals radii were set to 0.78 times the standard CHARMM values. RMSD values were obtained by best fitting the backbone atom co-ordinates for residues 12–112 of the 20 individual structures, and were calculated with respect to the mean structure. The values given for the backbone atoms and all heavy atoms represent the average pairwise values $\pm$ S.D.

(A)		
Parameter	Number/value	
Restraints for structure calculation		
Total restraints	1484	
Total nOe restraints	1394	
Intraresidue	482	
Sequential	331	
Medium range	280	
Long range	301	
Hydrogen-bond restraints	84	
Disulphide-bond restraints	6	
RMSD (Å)		
Backbone (C, Cα, N)	0.63 ± 0.39	
All heavy atoms	1.27 ± 0.4	
(B)		
Parameter	(DG)	[DG]
Energy (kJ/mol)		
Total	996.3	1071.4
Bonds	51.3	58.0
Angles	342.9	357.0
Improper	46.8	52.2
van der Waals (repel)	359.0	378.6
N0e	196.3	221.5
Number of violations from ideal value		
Bonds > 0.05 Å	0	0
Angles > 5°	0	0
Improper > 5°	0	0
Van der Waals < 1.6 Å	0.25	1
Dihedral > 30°	88.5	94
Maximum of nOe violations (Å)	0.43	0.54
RMSD from ideal value		
Bonds	0.0027	0.0027
Angles	0.42	0.43
Improper	0.29	0.29
Dihedral	29.61	29.69
RMSD from restraints		
nOe	0.021	0.023

violations and the quality of the geometry estimated on the sight of energy values (see Table 1).

Overall structure description

The protein has the overall shape of a globular protein of dimensions $32 \text{ \AA} \times 28 \text{ \AA} \times 24 \text{ \AA}$ (Figure 3). The molecular structure of CSPMbraA6 consists of a non-bundle α -helical structure made up of six amphipathic helices connected by $\alpha\alpha$ loops (Figure 3). The helices comprise residues Leu-13–Ala-18 (A), Arg-20–Val-30 (B), Ser-37–Ala-51 (C), Glu-60–Ile-76 (D), Glu-79–Lys-86 (E) and Trp-94–Ala-105 (F). The N- and C-terminal regions can be considered as extended regions, although they are less ordered. Turn types were determined according to the classification of Baldomero, as described in [33]. The first turn

Figure 3 Ribbon drawing of the averaged nOe-minimized CSPMbraA6 structure

The α -helices are labelled A–F (see the text). Residues at the beginning and end of helices are numbered, as are the cysteine residues involved in the disulphide bridges and the aromatic residues whose side chains are proposed to close the mouths of the channel.

is an $\alpha\alpha$ loop 1.1.4 comprising residues 19–20, the second is an $\alpha\alpha$ loop 5.1.1 comprising residues 31–36, the third is an $\alpha\alpha$ loop 5.1.1 comprising residues 51–59, the fourth is an $\alpha\alpha$ loop 2.1.1 comprising residues 78–79, the fifth is a $\alpha\alpha$ loop comprising residues 88–93, and the last is a $\alpha\beta$ loop 0.1.2 comprising residues 105–106.

In the case of the CSP of *Schistocerca gregaria*, it has been demonstrated that the four cysteine residues are involved in two disulphide bridges (Cys-29–Cys-38 and Cys-57–Cys-60) [14], which is mirrored in the present structure of CSPMbraA6. In the latter, the disulphide bridges close small loops, and involve cysteines 29 and 36 and cysteines 55 and 58 (Figure 3). These disulphide bridges seem to have little rigidifying effect on the loops.

As expected from the similarities between NMR and X-ray structures [32], comparing the structure of CSPMbraA6 with those contained in the DALI protein database [31] resulted in the detection of 37 motifs with Z scores of > 2 but with RMSD values of > 3 Å. However, these proteins have very weak sequence identity with CSPMbraA6. Visualization of the superimposed motifs with CSPMbraA6 indicates that this similarity is not significant, since the helical folds are not comparable. Therefore the structure of CSPMbraA6 should be considered as pertaining to a novel protein fold.

Protein surface

The surface of CSPMbraA6 is constituted mainly of polar amino acids, especially lysine and glutamic acid residues. The whole sequence comprises 14 lysine, six arginine, seven aspartic acid and 16 glutamic acid residues, i.e. 38% of the total number of residues. All of these polar residues are exposed to the solvent. In the complete set of NMR solutions, it appears that most of these

Figure 4 Molecular surface of CSPMbraA6 in two opposite orientations showing the two mouths of the channel

Hydrophilic surfaces are dark grey, and hydrophobic surfaces are light grey.

surface side chains are mobile. We can detect two hydrophobic regions at the surface of CSPMbraA6. One is localized at the C-terminal region and contains residues Gly-106, Ile-107, Val-108 and Ile-109 (Figure 4A). This hydrophobic patch goes through the molecule from the C-terminal surface down to a hydrophobic channel. The second region is located at the entrance of the channel, and encompasses residues Ile-11, Leu-13, Ile-16, Leu-22, Val-24 and Tyr-26 (Figure 4B).

Hydrophobic channel

All amphipathic helices direct their hydrophobic surface to the centre of CSPMbraA6, thus forming a hydrophobic channel. The relative positions of these hydrophobic helical sides with respect to each other are ensured by numerous long-range constraints. This hydrophobic channel crosses the molecule from the hydrophobic surface containing Tyr-26 (Figure 4B) to the surface containing Trp-94, Tyr-67 and Tyr-98, and is slightly bent. The channel has therefore two mouths: one comprising the hydrophobic region around Tyr-26, and the other (more hydrophilic) comprising three aromatic residues, Tyr-67, Trp-94 and Tyr-97, surrounded by the charged residues Glu-63, Lys-64, Arg-

Figure 5 HSQC spectra recorded for CSPMbraA6 alone (red) and in complex with 12-bromododecanol (black)

The molar concentration of 12-bromododecanol was two times that of CSPMbraA6.

68 and Lys-97. The side chains of the two aromatic residues Tyr-26 and Trp-94 appear to be highly mobile, as shown by their high RMSD values. Due to these mobilities, the channel is open at both ends in four of the 20 solutions, closed at both ends in 10 solutions, closed by the Tyr-26 side chain in four solutions and closed by Trp-94 for two solutions.

12-Bromododecanol binding experiment

Since 12-bromododecanol binds well to CSPMbraA6 [32], we compared HSQC spectra recorded for CSPMbraA6 free and in complex with 12-bromododecanol (Figure 5). Most of the nitrogen and amide proton chemical shifts were affected by ligand binding. Keeping in mind that the chemical shift values are directly related to the chemical environment, and therefore to the local conformation of the protein, it becomes clear that a large conformational change occurs in CSPMbraA6 upon ligand binding.

DISCUSSION

The three-dimensional structures of members of several classes of small lipid transport proteins have been solved to date. Lipocalins are among the best characterized [34]; they are formed of a β -barrel, with an internal cavity accommodating the hydrophobic ligand. Several complexes of fatty acid-binding proteins in complex with fatty acids have known three-dimensional structures, which show fatty acids accommodated in a cavity inside the barrel, and presenting a bent conformation [35].

In contrast, CSPMbraA6 has an all- α -helical structure. Several other proteins capable of accommodating lipids also have an all-helical structure, i.e. PBPs, lipid transfer proteins (LTPs) and the B1 and B2 proteins, which are secreted in the tubular accessory glands of the adult male mealworm beetle (*Tenebrio molitor*). The structure of PBP from *Bombyx mori* has revealed the

presence of a closed pheromone-containing cavity delineated by its six helices [36]. The bombykol pheromone, an unsaturated alcohol, presents a bent conformation, reminiscent of that of fatty acids in fatty acid-binding proteins [35]. LTPs are small proteins (around 100 amino acids) that facilitate the transfer of lipids through membranes; the structures of LTPs, solved by NMR [37–39] and X-ray crystallography [40], reveal four helices delineating a cavity extending through the entire molecule. B1 and B2 proteins share common biochemical features with CSPs, and have been proposed to be lipid carriers that are able to keep hydrophobic compounds in solution in the aqueous seminal fluid [41]. NMR studies of THP12, a protein from *Tenebrio molitor* homologous to B1/B2, revealed the non-bundle structure of six α -helices and demonstrated that THP12 binds fatty acids (nonanoic acid and octanoic acid) as well as its specific pheromones, 4-methylnonanol and ergosterol [42]. However, the relative arrangements of the helices of all of these proteins differ greatly. Furthermore, whereas PBPs and LTPs possess an internal cavity for ligand accommodation, THP12 exhibits elongated grooves located at the protein surface. Indeed, such long grooves are able to bind aliphatic ligands in an elongated rather than a bent conformation, in contrast with proteins with a binding site in a closed cavity.

CSPMbraA6 has an internal hydrophobic channel which can be open or closed according to the orientation of two aromatic side chains. However, the overview of the HSQC spectra recorded for the CSPMbraA612-bromododecanol complex indicates that the conformational change is greater than re-orientation of a few side chains, and might involve some re-orientation of the secondary structure. This can be related to the intrinsic mobility of CSPMbraA6, illustrated by the relatively poor quality of the fit of helix C. Moreover, we can clearly detect a large chemical shift change of the indole nitrogen of Trp-94, which is in accordance with its mobility. We can hypothesize that CSPMbraA6 binds to membranes possibly through an interaction of the polar mouth of the hydrophobic channel. The slightly bent channel might then accommodate linear hydrophobic molecules, but also larger molecules such as those identified by Ban et al. [43], since protein flexibility seems to be very high.

We thank Rainer Wechselberger and Olivier Bornet for technical assistance. This study was supported in part by the PACA region (no. 9811/2177), by the EU BIOTECH Structural Biology project (BIO4-98-0420 OPTIM) and by the Large Scale Facility program of the EU. A.M. is the recipient of a Ph.D. fellowship from CNRS.

REFERENCES

- Steinbrecht, R. A., Ozaki, M. and Ziegelberger, G. (1992) Immunocytochemical localization of pheromone-binding-protein in moth antennae. *Cell Tissue Res.* **270**, 287–302
- Vogt, R. G., Riddiford, L. M. and Prestwich, G. D. (1985) Kinetic properties of a sex pheromone-degrading enzyme: the sensillar esterase of *Antheraea polyphemus*. *Proc. Natl. Acad. Sci. U.S.A.* **82**, 8827–8831
- Pelosi, P. (1996) Perireceptor events in olfaction. *J. Neurobiol.* **30**, 3–19
- Steinbrecht, R. A., Laue, M. and Ziegelberger, G. (1995) Immunolocalization of pheromone-binding protein and general odorant-binding protein in olfactory sensilla of the silk moths *Antheraea* and *Bombyx*. *Cell Tissue Res.* **282**, 203–217
- Pace, U. and Lancet, D. (1986) Olfactory GTP-binding protein: signal-transducing polypeptide of vertebrate chemosensory neurons. *Proc. Natl. Acad. Sci. U.S.A.* **83**, 4947–4951
- Kurihara, K. and Koyama, N. (1972) Mechanism of bitter taste reception: interaction of bitter compounds with monolayers of lipids from bovine circumvallate papillae. *Biochem. Biophys. Res. Commun.* **48**, 30–34
- Jacquin-Joly, E., Vogt, R. G., François, 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

- 8 Maleszka, R. and Stange, G. (1997) Molecular cloning, by a novel approach, of a cDNA encoding a putative olfactory protein in the labial palps of the moth *Cactoblastis cactorum*. *Gene* **202**, 39–43
- 9 Robertson, H. M., Martos, R., Sears, C. R., Todres, E. Z., Walden, K. K. and Nardi, J. B. (1999) Diversity of odourant binding proteins revealed by an expressed sequence tag project on male *Manduca sexta* moth antennae. *Insect Mol. Biol.* **8**, 501–518
- 10 Nomura, A., Kawasaki, K., Kubo, T. and Natori, S. (1992) Purification and localization of p10, a novel protein that increases in nymphal regenerating legs of *Periplaneta americana* (American cockroach). *Int. J. Dev. Biol.* **36**, 391–398
- 11 Adams, M. D., Celniker, S. E., Holt, R. A., Evans, C. A., Gocayne, J. D., Amanatides, P. G., Scherer, S. E., Li, P. W., Hoskins, R. A., Galle, R. F. et al. (2000) The genome sequence of *Drosophila melanogaster*. *Science* **287**, 2185–2195
- 12 Dyanov, H. M. and Dzitoeva, S. G. (1995) Method for attachment of microscopic preparations on glass for *in situ* hybridization, PRINS and *in situ* PCR studies. *Biotechniques* **18**, 822–826
- 13 Pikielny, C. W., Hasan, G., Rouyer, F. and Rosbash, M. (1994) Members of a family of *Drosophila* putative odorant-binding proteins are expressed in different subsets of olfactory hairs. *Neuron* **12**, 35–49
- 14 Angeli, S., Ceron, F., Scaloni, A., Monti, M., Monteforti, G., Minnocci, A., Petacchi, R. and Pelosi, P. (1999) Purification, structural characterization, cloning and immunocytochemical localization of chemoreception proteins from *Schistocerca gregaria*. *Eur. J. Biochem.* **262**, 745–754
- 15 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
- 16 Marchese, S., Angeli, S., Andolfo, A., Scaloni, A., Brandazza, A., Mazza, M., Picimbon, J., Leal, W. S. and Pelosi, P. (2000) Soluble proteins from chemosensory organs of *Eurycantha calcarata* (Insecta, Phasmatodea). *Insect Biochem. Mol. Biol.* **30**, 1091–1098
- 17 Nagnan-Le Meillour, P., Cain, A. H., Jacquin-Joly, E., François, M. C., Ramachandran, S., Maida, R. and Steinbrecht, R. A. (2000) Chemosensory proteins from the proboscis of *Mamestra brassicae*. *Chem. Senses* **25**, 541–553
- 18 Mameli, M., Tuccini, A., Mazza, M., Petacchi, R. and Pelosi, P. (1996) Soluble proteins in chemosensory organs of phasmids. *Insect Biochem. Mol. Biol.* **26**, 875–882
- 19 McKenna, M. P., Hekmat-Safes, 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
- 20 Picimbon, J. F., Dietrich, K., Angeli, S., Scaloni, A., Krieger, J., Breer, H. and Pelosi, P. (2000) Purification and molecular cloning of chemosensory proteins from *Bombyx mori*. *Arch. Insect Biochem. Physiol.* **44**, 120–129
- 21 Bohbot, J., Sobrio, F., Lucas, P. and Nagnan-Le Meillour, P. (1998) Functional characterization of a new class of odorant-binding proteins in the moth *Mamestra brassicae*. *Biochem. Biophys. Res. Commun.* **253**, 489–494
- 22 Campanacci, V., Mosbah, A., Bornet, O., Wechselberger, R., Jacquin-Joly, E., Cambillau, C., Darbon, H. and Tegoni, T. (2001) Chemosensory protein from the moth *Mamestra brassicae*. Expression and secondary structure from ¹H and ¹⁵N NMR. *Eur. J. Biochem.* **268**, 4731–4739
- 23 Bartels, C. H., Xia, T.-H., Billeter, M., Güntert, P. and Wüthrich, K. (1995) The program XEASY for computer-supported NMR spectral analysis of biological macromolecules. *J. Biomol. NMR* **5**, 1–10
- 24 Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York
- 25 Güntert, P. and Wüthrich, K. (1991) Improved efficiency of protein structure calculations from NMR data using the program DIANA with redundant dihedral angle constraints. *J. Biomol. NMR* **1**, 447–456
- 26 Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kustleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Panun, N. S. et al. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr.* **D54**, 905–921
- 27 Roussel, A. and Cambillau, C. (1989) Silicon Graphics Geometry Partner Directory pp. 77–78. Silicon Graphics, Mountain View, CA
- 28 Koradi, R., Billeter, M. and Wüthrich, K. (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graphics* **14**, 29–32
- 29 Güntert, P., Mumenthaler, C. and Wüthrich, K. (1997) Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J. Mol. Biol.* **273**, 283–298
- 30 Laskowski, R. A., Rullmann, J. A., MacArthur, M. W., Kaptein, R. and Thornton, J. M. (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* **8**, 477–486
- 31 Holm, L. and Sander, C. (1998) Protein folds and families: sequence and structure alignments. *Nucleic Acids Res.* **26**, 316–319
- 32 Lartigue, A., Campanacci, V., Roussel, A., Larsson, A. M., Jones, T. A., Tegoni, M. and Cambillau, C. (2002) X-ray structure and ligand binding study of a moth chemosensory protein. *J. Biol. Chem.* **277**, 32094–32098
- 33 Oliva, B., Bates, P. A., Querol, E., Aviles, F. X. and Sternberg, J. E. (1997) An automated classification of the structure of protein loops. *J. Mol. Biol.* **266**, 814–830
- 34 Flower, D. R., North, A. C. and Sansom, C. E. (2000) The lipocalin protein family: structural and sequence overview. *Biochim. Biophys. Acta* **1482**, 9–24
- 35 Balendiran, G. K., Schnutgen, F., Scapin, G., Borchers, T., Xhong, N., Lim, K., Godbout, R., Spener, F. and Sacchettini, J. C. (2000) Crystal structure and thermodynamic analysis of human brain fatty acid-binding protein. *J. Biol. Chem.* **275**, 27045–27054
- 36 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
- 37 Campanacci, V., Krieger, J., Bette, S., Sturgis, J. N., Lartigue, A., Cambillau, C., Breer, H. and Tegoni, M. (2001) Revisiting the specificity of *Mamestra brassicae* and *Antheraea polyphemus* pheromone-binding proteins with a fluorescence binding assay. *J. Biol. Chem.* **276**, 20078–20084
- 38 Gincel, E., Simorre, J. P., Caille, A., Marion, D., Ptak, M. and Vovelle, F. (1994) Three-dimensional structure in solution of a wheat lipid-transfer protein from multidimensional ¹H-NMR data. A new folding for lipid carriers. *Eur. J. Biochem.* **226**, 413–422
- 39 Gomar, J., Petit, M. C., Sodano, P., Sy, D., Marion, D., Kader, J. C., Vovelle, F. and Ptak, M. (1996) Solution structure and lipid binding of a nonspecific lipid transfer protein extracted from maize seeds. *Protein Sci.* **5**, 565–577
- 40 Lerche, M. H., Kragelund, B. B., Bech, L. M. and Poulsen, F. M. (1997) Barley lipid-transfer protein complexed with palmitoyl CoA: the structure reveals a hydrophobic binding site that can expand to fit both large and small lipid-like ligands. *Structure* **5**, 291–306
- 41 Shin, D. H., Lee, J. Y., Hwang, K. Y., Kim, K. K. and Suh, S. W. (1995) High-resolution crystal structure of the non-specific lipid-transfer protein from maize seedlings. *Structure* **3**, 189–199
- 42 Paesen, G. C. and Happ, G. M. (1995) The B proteins secreted by the tubular accessory sex glands of the male mealworm beetle, *Tenebrio molitor*, have sequence similarity to moth pheromone-binding proteins. *Insect Biochem. Mol. Biol.* **25**, 401–408
- 43 Ban, L., Zhang, L., Yan, Y. and Pelosi, P. (2002) Binding properties of a locust's chemosensory protein. *Biochem. Biophys. Res. Commun.* **293**, 50–54

Received 2 August 2002; accepted 9 September 2002

Published as BJ Immediate Publication 9 September 2002, DOI 10.1042/BJ20021217