X-ray Structure and Ligand Binding Study of a Moth Chemosensory Protein

Audrey Lartigue, Valérie Campanacci, Alain Roussel, Anna M. Larsson, T. Alwyn Jones, Mariella Tegoni, and Christian Cambillau

From the ‡Architecte et Fonction des Macromolécules Biologiques, UMR 6098, CNRS and Universités d’Aix-Marseille I and II, 31 chemin J. Aiguier, 13402 Marseille Cedex 20, France and ¶Department of Cell and Molecular Biology, Uppsala University, Biomedical Centre, Box 596, S-75124, Uppsala, Sweden

Chemosensory proteins (CSPs) are believed to be involved in chemical communication and perception. Such proteins, of M, 13,000, have been isolated from several sensory organs of a wide range of insect species. Several CSPs have been identified in the antennae and proboscis of the moth Mamestra brassicae. One of them, CSPMbraA6, a 112-amino acid antennal protein, has been expressed in large quantities and is soluble in the Escherichia coli periplasm. X-ray structure determination has been performed in parallel with ligand binding assays using tryptophan fluorescence quenching. The protein has overall dimensions of 25 × 30 × 32 Å and exhibits a novel type of α-helical fold with six helices connected by α-α loops. A narrow channel extends within the protein hydrophobic core. Fluorescence quenching with brominated alkyl alcohols or fatty acids and modeling studies indicates that CSPMbraA6 is able to bind such compounds with C12–18 alkyl chains. These ubiquitous proteins might have the role of extracting hydrophobic linear compounds (pheromones, odors, or fatty acids) dispersed in the phospholipid membrane and transporting them to their receptor.

Two classes of highly soluble and very abundant proteins of ~150 amino acids have been detected in sensilla of Lepidoptera, both containing 6 conserved cysteines forming three disulfide bridges. The first class, that of GOBPs, is equally distributed in both sexes, whereas the second class, that of PBPs, is mainly present in males. A third class of small proteins (average M, 13,000) has been identified in antennae from Drosophila melanogaster and in antennae and several sensorial organs (tarsi, labrum) from a wide range of species of the insect order (2–11). These proteins have been proposed to be involved in CO2 detection (3), in chemical signal transmission in regenerating legs (5), or in chemo-perception (either olfaction or taste (9, 12)), and they were therefore called chemosensory proteins (CSPs). CSPs are shorter (110–115 amino acids) than PBP or GOBP, contain only 4 conserved cysteines forming two disulfide bridges (9), and share no sequence homology with them. 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 as other transport proteins such as GOBPs or 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 (12) and in the antennae (13). M. brassicace CSPs have been shown to bind several components of the pheromonal blend and therefore might also have a function analogous to that of PBPs (12). In the proboscis, however, a putative role of odor or taste carriers has been assigned to CSPs (12).

To improve our knowledge of these important insect proteins, we have expressed in the Escherichia coli periplasm CSPMbraA6 originating from the antennae of M. brassicace (14). CSPMbraA6 NMR preliminary assignment and crystallization have already been published (14, 15). The present paper is the first report of a CSP structure and of its preliminary functional characterization. We describe the three-dimensional x-ray structure of CSPMbraA6, which displays a novel fold, and fluorescence binding studies with 12-bromododecanol and brominated fatty acids. CSPMbraA6 is able to bind these compounds with good affinity. Based on these results, we hypothesize its putative function as a lipid carrier.

MATERIALS AND METHODS

Production of CSPMbraA6—A detailed description of the molecular cloning, expression, purification, and characterization of CSPMbraA6 has previously been reported (14). Expression of selenomethionine-substituted CSPMbraA6 was performed using the methionine pathway inhibition method (16). Briefly, bacteria were grown to midlog phase in minimal medium before addition of the following amino acids: lysine, phenylalanine, and threonine at 100 mg/ml, isoleucine, leucine, and valine at 50 mg/ml, and selenomethionine at 50 mg/ml. Induction thereafter carried on for 30 min in the same conditions as described above for the expression of mutants. Mass spectrometry was performed to confirm the substitution of the unique Met in CSPMbraA6.

X-ray Structure Determination—Crystals of CSP2 were obtained using two different conditions (15). Crystals of form 1 belong to the P41212 space group with unit cell dimensions of a = b = 49.4 Å, c = 79.6 Å and contain 1 molecule per asymmetric unit whereas crystals of form 2 belong to a monoclinic space group P21 with unit cell dimensions of a = 47.6 Å, b = 49.7 Å, c = 50.3 Å, β = 110.1° and contain 2 molecules per asymmetric unit. Crystals of selenomethionyl protein were obtained in an identical fashion as the native crystals but needed seeding using small native crystals. For data collection, crystals were flash frozen at 100 K in their mother liquor without cryoprotectant.
**RESULTS AND DISCUSSION**

**Overall Structure Description**—CSPMbraA6 has an overall globular shape with dimensions of $25 \times 30 \times 32 \text{ Å}$ and consists of six helices connected by $\alpha$-$\alpha$ loops (Fig. 1a). In crystal form 1 and in the monomer B of crystal form 2, the polypeptide chain is visible from residue 11 to residue 112, whereas in monomer A of crystal form 2, the chain starts at residue 4. The remaining N-terminal residues are probably disordered in the crystal. The helices correspond to residues 5 (or 12)–18 (helix A), 20–30 (B), 38–53 (C), 60–76 (D), 78–88 (E), and 93–105 (F) (Fig. 1a). Two disulfide bridges close small loops because they are formed by cysteines 29 and 36 and cysteines 55 and 58 (Fig. 1a), similar to what was proposed for the CSP of Schistocerca gregaria (9).

No comparable fold has been found in the Protein Data Bank using DALI (28), indicating that the six helices are arranged in an original way. Helices A and B (residues 6–30) as well as helices D and E (residues 60–88) form two V-shaped structures, with opening angles of ~60° (Fig. 1, b and c). The planes defined by the two V-shaped structures are parallel and about 12 Å apart. Helix C is perpendicular to these two planes and positioned in between the four ends of the two V-shaped structures (Fig. 1, b and c). The final helix (F) is located packed against the external face of the D–E helices and does not take part in the core assembly. Aside from the N termini, the Co chains of the different protein forms or monomers are very similar. A RMSD of 0.44 Å is observed between the Co atoms common to monomers A and B of form 2, and a RMSD of 0.64 Å is observed between the Co atoms of crystal form 1 and monomer A of crystal form 2. A few significant differences are observed in some side chain orientations, however (see below).

The sequence of CSPMbraA6 contains 16 Glu, 7 Asp, 14 Lys, and 6 Arg residues, accounting for 39% of the total amino acid content. All of these charged residues are located at the protein surface. Aromatic and aliphatic residues, accounting for 30% of the total amino acid content, form the core of the protein and the walls of the internal channel (see below); only a few of these hydrophobic residues are located near the protein surface. The high content of charged residues and the absence of hydrophobic patch at the protein surface explain well the monomeric nature of CSPMbraA6 and its outstanding solubility.

**Hydrophobic Channel**—A narrow channel, starting from the surface region between residues 6–10 (helix A) and residues 62–68 (helix D), extends 14 Å within the core of monomer A of form 2 between the two V-shaped structures (Fig. 2, a and b). Six ordered water molecules are visible in this channel at hydrogen bonding distances from each other (Fig. 2, a and b). They contact, from outside to inside, Arg-68, Asn-10, Asn-61,
Leu-13, Asp-9, His-46, Glu-62, Gly-65, Ala-66, Leu-43, Val-69, Leu-47, and Tyr-26. One side of the channel is formed by helix A (residues 5–18) and is stabilized by a crystal contact between Tyr-8 and Glu-39 of a symmetry-related molecule. The side chain of Tyr-26 forms the bottom of the channel, preventing its continuity with a nearby internal cavity (Fig. 2, a and b).

In monomer B, the disorder of the peptidic segment 1–10 results in a large opening at the position of the beginning of the channel in monomer A (Fig. 2, c and d). NMR studies indicate that this segment is also disordered in solution. The Tyr-26 side chain is rotated by 100° around the γ1 angle toward the protein surface, and the position of its hydroxyl group in monomer A is occupied by the side chain of Leu-43. As a result of these side chain rotations, the channel is much shorter, and a

FIG. 1. X-ray structure of CSPMbraA6. a, stereo view of the Ca tracing and disulfide bridges arrangement. Helices are indicated as a–f. b, representation of the helical V-shaped fold of CSPMbraA6 (helices A–E). c, 90° rotation compared with b (helices A–F). Views were prepared with SPOCK (29) for a and Turbo-Frodo (23) for b and c.

FIG. 2. Molecular surface of CSPMbraA6. a, monomer A with its six bound water molecules in the channel and Tyr-26 (ball-and-sticks, pink) blocking the communication with the rest of the channel (blue arrows). b, 90° view. c, monomer B with two bound water molecules in the short channel and Tyr-26 (ball-and-sticks, pink) orientated toward protein surface. The internal cavity is identified with blue arrows. d, 90° view. Views were prepared with SPOCK (29).

\[2\] A. Mosbah, V. Campanacci, A. Lartigue, M. Tegoni, C. Cambillau, and H. Darbon, submitted for publication.
closed cavity starts in monomer B where the channel stops in monomer A.

We hypothesize that, upon ligand binding, the Tyr-26 side chain might be rotated toward the protein surface as in monomer B. Consequently, applying to monomer A the rotation of Tyr-26 observed in monomer B makes it possible to form a continuous internal channel starting from the surface and about 20 Å deep. This narrow and elongated channel seems suitable for alkyl compound binding, a hypothesis tested below.

**Ligand Binding Solution Studies**—As AMA has been widely used in binding assays with mammalian lipocalins (30) and insect PBPs (31), we initially attempted to use this fluorescent probe to determine the affinity of various ligands to CSPMbraA6. Unfortunately, no fluorescence increase was observed upon addition of AMA to a solution of CSPMbraA6 (data not shown), indicating that AMA does not bind in the internal hydrophobic channel. This was not surprising considering the bulkiness of the anthracene moiety and the small radius of CSPMbraA6 channel.

As bromine has been reported as an excellent long range quencher of tryptophan fluorescence (32), we decided to perform binding assays with linear compounds bearing a bromine atom: 12-bromo-dodecanol, 15-bromo-pentadecanoic acid, and 9-bromo-stearic acid (named C12Br, C15Br, and C18Br, respectively). In the presence of saturating concentrations of these compounds, the intrinsic tryptophan fluorescence of CSPMbraA6 was quenched at 26, 34, and 30%, respectively (Fig. 3, A–C), which suggests that at least one of the two tryptophans interacts with the ligand. This decrease in fluorescence intensity was associated with a blue shift of the emission maximum, from 343 to 339 nm with C12Br and to 332 nm with C15Br and C18Br, indicating that the tryptophan environment becomes probably more hydrophobic upon ligand binding. To determine the affinity of CSPMbraA6 for the three ligands, the fluorescence quenching at 342 nm was measured as a function of ligand concentration, and the dissociation constant was estimated by non-linear regression of the binding curve (Fig. 3, A–C). The K_d values determined were 0.90, 1.60, and 0.35 μM, respectively, in the range of values observed for PBP-pheromone complexes (31, 33).

**Ligand Binding Modeling**—The x-ray structure and the positive binding assays prompted us to model the interaction between the ligands and CSPMbraA6. We performed a small rotation of the side chain of Tyr-26, which extends the length of the hydrophobic channel to about 20 Å inside the protein (see above). The three brominated ligands used in fluorescence experiments were modeled and docked in the cavity with the OH or COOH moieties pointing out of the channel. In this orientation, C12Br is fully imbedded in the channel while the carboxylic moieties of the C15Br and C18Br reach the protein surface and the bulk solvent (Fig. 3D). The three ligands fit well in the narrow channel in an elongated way for most of the alkyl chain, a kink being induced at the level of carbons 5–6 by a change in the channel direction. These ligand positions with the hydrophilic group turned to the outside are also in agreement with the hydrophobic character of the channel. The bromine atoms of the ligands are closer to the Trp-81 indole rings (8, 8, and 11 Å, respectively) than to Trp-94 (16, 16, and 15 Å, respectively). This orientation accounts well for the fluorescence of Trp-81 being mostly quenched.

**Comparison with Other Lipid Transport Proteins**—The three-dimensional structures of several classes of small lipid transport proteins have been solved to date and have diverse structural frameworks. Two families (34–36) employ anti-parallel β-barrels to form internal, variable shaped internal cavities that accommodate the often hydrophobic ligands. By

![Fig. 3. Complexes of CSPMbraA6 with lipidic ligands.](image)

Tryptophan fluorescence quenching curves with C18Br (A), C15Br (B), and C12Br (C). The tryptophan fluorescence intensities (corrected for dilution) at the emission maxima are plotted as a function of the quencher concentration. The quenching spectra obtained upon addition of ligand are shown in the insets. D, molecular model of the complex between CSPMbraA6 and C15Br. The protein is represented in ribbon representation with a transparent surface. The ligand is represented with spherical atoms (red, O; yellow, C; blue, Br). View made with SPOCK (29).
changing the residues inside of such barrels, a wide range of specificity can be achieved.

In contrast, CSPMbraA6 has an all α-helical structure. Several other proteins capable of accommodating lipids also have an all helical structure: LTPs (37, 38), the B1 and B2 proteins (39), and PBPs (40, 41). PBPs and LTPs possess an internal cavity, as in the case of CSPs. The structure and binding properties of an insect chemosensory protein are reported here for the first time. Our results clearly point to a lipid transport function for CSPMbraA6. Lipids are poorly soluble in the fluids from antenna or other organs containing CSPMbraA6 but are most probably dispersed in the phospholipidic membrane. We hypothesis therefore that CSPMbraA6 should interact with membranes to extract pheromones or other lipidic compounds dispersed in the hydrophobic membrane matrix and bring them to their specific target. Whether the N-terminal segment is the trigger of this interaction in a comparable way as the C-terminal helix of PBP remains to be explored.

Acknowledgments—The ESRF (Grenoble, France) is acknowledged for beam time allocation. We thank Philippe Cantau and Eric Blanc for technical assistance.

REFERENCES
1. Steinbrecht, R. A., Laue, M. & Zieglerberger, G. (1995) Cell Tissue Res. 282, 203–217
2. Jacquin-Joly, E., Vogt, R. G., François, M.-C. & Nagan-Le Meillour, P. (2001) Chem. Senses 26, 833–844
3. Maleszka, R. & Stange, G. (1997) Gene (Amst.) 209, 39–43
4. Robertson, H. M., Martos, R., Sears, C. R., Todres, E. Z., Walden, K. K. & Nardi, J. B. (1999) Insect Biochem. Mol. Biol. 8, 501–518
5. Nomura, A., Kawasaki, K., Kubo, T. & Natori, S. (1992) Int. J. Dev. Biol. 36, 391–398
6. Adams, M. D. et al. (2000) Science 287, 2185–2195
7. Dyanov, H. M. & Dzitoeva, S. G. (1995) BioTechniques 18, 822–824, 826
8. Piskelny, C. W., Hassan, G., Rouyer, F. & Rosbash, M. (1994) Neuron 12, 35–49
9. Angeli, S., Ceron, F., Scaloni, A., Menzi, M., Monteforte, G., Minnocci, A., Petaec, R. & Pelosi, P. (1999) Eur. J. Biochem. 262, 745–754
10. Pimbinon, J. F., Dietrich, K., Breer, H. & Krieger, J. (2000) Insect Biochem. Mol. Biol. 30, 233–241
11. Marchese, S., Angeli, S., Andolfi, A., Scaloni, A., Brandazza, A., Maza, M., Pimbinon, J., Leal, W. S. & Pelosi, P. (2000) Insect Biochem. Mol. Biol. 30, 1091–1098
12. Nagan-Le Meillour, P., Cain, A. H., Jacquin-Joly, E., François, M. C., Ramachandran, S., Maida, R. & Steinbrecht, R. A. (2000) Chem. Senses 25, 541–553
13. Bohbot, J., Sobrio, F., Lucas, P. & Nagnan-Le Meillour, P. (1998) Biochem. Biophys. Res. Commun. 253, 489–494
14. Campanacci, V., Musbah, A., Bornet, O., Wechselberger, R., Jacquin-Joly, E., Cambillau, C., Darbon, H. & Tegoni, M. (2001) Eur. J. Biochem. 268, 4731–4739
15. Campanacci, V., Spinelli, S., Lartigue, A., Lewandowski, C., Brown, K., Tegoni, M. & Cambillau, C. (2001) Acta Crystallogr. Sect. D Biol. Crystallogr., 57, 137–139
16. Doublie, S. (1997) Methods Enzymol. 276, 523–530
17. Paesen, G. C. & Happ, G. M. (1995) Insect Biochem. Mol. Biol. 25, 30, 137–139