Solution Structure of a Chemosensory Protein from the Desert Locust *Schistocerca gregaria*[†,‡]

Simona Tomaselli,[§] Orlando Crescenzi,[§] Domenico Sanfelice,[§] Eiso AB,[||,⊥] Rainer Wechselberger,[||] Sergio Angeli,[#] Andrea Scaloni,[△] Rolf Boelens,[||] Teodorico Tancredi,[◇] Paolo Pelosi,[#] and Delia Picone*[,§]

Department of Chemistry, University of Naples Federico II, 80126 Napoli, Italy, Department of NMR Spectroscopy, Bijvoet Center for Biomolecular Research, Utrecht University, 3584 CH Utrecht, The Netherlands, Department of Agricultural Chemistry and Biotechnology, University of Pisa, 56124 Pisa, Italy, Proteomics and Mass Spectrometry Laboratory, ISPAAM, National Research Council, 80147 Napoli, Italy, and Institute of Biomolecular Chemistry, National Research Council, Napoli, Italy

Received May 19, 2006; Revised Manuscript Received July 7, 2006

ABSTRACT: Chemical stimuli, generally constituted by small volatile organic molecules, are extremely important for the survival of different insect species. In the course of evolution, insects have developed very sophisticated biochemical systems for the binding and the delivery of specific semiochemicals to their cognate membrane-bound receptors. Chemosensory proteins (CSPs) are a class of small soluble proteins present at high concentration in insect chemosensory organs; they are supposed to be involved in carrying the chemical messages from the environment to the chemosensory receptors. In this paper, we report on the solution structure of CSP*sg*4, a chemosensory protein from the desert locust *Schistocerca gregaria*, which is expressed in the antennae and other chemosensory organs. The 3D NMR structure revealed an overall fold consisting of six α-helices, spanning residues 13−18, 20−31, 40−54, 62−78, 80−90, and 97−103, connected by loops which in some cases show dihedral angles typical of *β*-turns. As in the only other chemosensory protein whose structure has been solved so far, namely, CSP from the moth *Mamestra brassicae*, four helices are arranged to form a V-shaped motif; another helix runs across the two V's, and the last one is packed against the external face. Analysis of the tertiary structure evidenced multiple hydrophobic cavities which could be involved in ligand binding. In fact, incubation of the protein with a natural ligand, namely, oleamide, produced substantial changes to the NMR spectra, suggesting extensive conformational transitions upon ligand binding.

Chemical signals regulate most aspects of insect life. Insects have developed very sensitive and sophisticated systems to detect and correctly recognize different molecules present in the environment. Receptors are distributed on the surface of sensory organs, such as antennae, tarsi, and mouth parts, housing dendrites of chemosensory neurons. G-protein-coupled receptors, spanning the neuronal membrane with seven α-helical segments, have been generally recognized as the biochemical elements responsible for recognition of olfactory and taste stimuli (*1*, *2*).

Before interacting with their cognate membrane-bound receptors, chemical stimuli, mainly consisting of small hydrophobic molecules, have to cross an aqueous barrier. According to a current view, this process is assisted by small soluble proteins capable of reversibly binding a great variety of these organic molecules (*3*−*10*). These proteins are classified in two main groups, based mainly on sequence analysis, namely, odorant/pheromone binding proteins (OBPs/PBPs)[1] and chemosensory proteins (CSPs). Their specific roles in chemoreception have not been entirely elucidated, although it has been demonstrated that olfactory receptors (e.g., those expressed in heterologous systems) can correctly recognize odorant molecules even in the absence of OBPs or CSPs (*11*, *12*). Other experiments seem to indicate that LUSH, a member of *Drosophila* OBPs, is required for both behavioral and electrophysiological response to the pheromone vaccenyl acetate (*13*).

Whatever their functions are, the presence of PBPs and CSPs seems crucial for these important physiological processes; their biosynthesis in exceptionally large amounts requires a significant investment of energy by the insect and suggests an important role, both for the survival of the individual and for the conservation of the species (*10*). Moreover, the extremely high concentrations of OBPs and CSPs in the sensillar lymph (around 10 mM) (*14*) do not

[†] Financial support was from the Italian MIUR (FIRB RBNE03B8KK).

[‡] The coordinates have been deposited into the PDB together with the NMR data. The accession codes are 2GVS and 7184, respectively.

* Address correspondence to this author. Fax: + 39 081 674409. Tel: + 39 081 674406. E-mail: delia.picone@unina.it.

[§] University of Naples Federico II.

[||] Utrecht University.

[⊥] Current address: Gorlaeus Laboratories, Einsteinweg 55, 2333 CC Leiden, The Netherlands.

[#] University of Pisa.

[△] Proteomics and Mass Spectrometry Laboratory, ISPAAM, National Research Council.

[◇] Institute of Biomolecular Chemistry, National Research Council.

[1] Abbreviations: CSP, chemosensory protein; OBP, odorant binding protein; PBP, pheromone binding protein; CSP*sg*4, CSP4 from *Schistocerca gregaria*; CSPMbra, CSP from *Mamestra brassicae*; HSQC, heteronuclear single-quantum coherence; TOCSY, total correlated spectroscopy; NOESY, nuclear Overhauser enhancement spectroscopy; NOE, nuclear Overhauser effect.

10.1021/bi060998w CCC: $33.50
Published on Web 08/09/2006

allow odorants and other semiochemicals to travel from the outer environment to the dendritic membrane without continually impinging into these proteins (*10*).

One of the approaches to understand the role of OBPs/PBPs and CSPs in chemoreception relies on a detailed knowledge of their three-dimensional structure and their mode of binding, including possible conformational changes induced by formation of complexes with their ligands. The structures of several insect OBPs have been resolved by X-ray diffraction and/or by NMR spectroscopy (*15*−*23*). By contrast, the three-dimensional fold has been described for only a single CSP, namely, CSPMbraA6 from the moth *Mamestra brassicae*, both in the crystal and in solution (*24*−*26*). Both OBPs and CSPs from insects are mainly consisting of α-helical domains, packed in very compact and stable structures although different in their folding motives. Both classes of proteins feature a hydrophobic binding pocket, which in some cases has been shown to contain putative endogenous ligands (*27*, *28*).

In proteins of both classes, major conformational changes have been observed as a consequence of ligand binding. In the PBP of *Bombyx mori*, the C-terminus is unstructured at neutral pH, and the natural pheromone ligand, bombykol, is bound with high affinity (*15*, *16*). However, at acidic pH values the C-terminus folds into an α-helical segment that fills the binding pocket, thus displacing the bombykol molecule (*17*). This behavior has been invoked to support a biochemical model in which the PBP binds the pheromone from the environment and then releases it in the more acidic environment of the membrane, directly onto the olfactory receptor. Similarly, a conformational change associated with ligand binding has also been observed in CSPMbraA6, which has been reported to bind three molecules of 12-bromododecanol, swelling extensively in this process (*24*).

Although both OBPs and CSPs share common structural features, they seem to differ widely in ligand specificities. Thus, a deeper understanding of their three-dimensional structure, particularly focused on their ligand binding sites, may shed light on the molecular mechanisms underlying insect behavior. In this paper, we report on the solution structure of CSP*sg*4, which is expressed in the antennae and other chemosensory organs of the desert locust *Schistocerca gregaria*. This protein has been shown to fill the outer lymph of one-pore contact chemosensilla, thus suggesting its possible involvement in taste or more generally in contact chemoreception (*29*).

MATERIALS AND METHODS

Protein Expression and Characterization. Details of the cDNA cloning and of the recombinant CSP*sg*4 expression have been described previously (*30*). CSP*sg*4 was heterologously expressed using the pET-Schi 10 plasmid in the *Escherichia coli* BL21(DE3) strain. To prepare the ^{15}N-, ^{15}N,^{13}C-, and ^{15}N,10% ^{13}C-labeled proteins, bacteria were grown on M9 minimal media containing as sole nitrogen and carbon source 1 g/L $^{15}NH_4Cl$, 1 g/L $^{15}NH_4Cl$ and 3 g/L [^{13}C]glucose or 1 g/L $^{15}NH_4Cl$ and 0.3 g/L [^{13}C]glucose, respectively. The protein was purified by a three-step procedure that combines two sequential ion-exchange chromatographic steps, followed by Vydac C_4 reverse-phase high-performance liquid chromatography (*30*). In particular, a rough purification was initially obtained by loading the cell lysate directly on a CM-Sepharose Fast Flow column (Pharmacia), equilibrated with 25 mM Tris-HCl, pH 7.6. The flow-through was then applied on a DEAE-Sepharose Fast Flow column (Pharmacia) equilibrated in the same buffer and eluted with a linear gradient of NaCl. The final purification step was a reverse-phase HPLC, as already reported (*30*). Protein samples were verified in purity and isotopic content by MALDI-TOF-MS analysis using a Voyager-DE PRO mass spectrometer (Applied Biosystems, Framingham, MA), as previously described (*30*). The NMR samples were 1.2 mM protein in 50 mM sodium phosphate buffer, pH 6.8, and 5% D_2O; protein concentration was estimated by UV (*30*).

Spectral Assignment. NMR spectra were acquired at 300 K on Bruker spectrometers operating at 500, 600, 700, and 750 MHz, equipped with triple resonance gradient probes. Data were processed with NMRPipe (*31*); visualization of spectra, peak-picking, and analysis were done in NMRView (*32*). Sequence assignments were performed using three-dimensional (3D) HNCA (*33*), HNCACB (*34*), CBCA(CO)-NH (*35*), and HNCO (*33*) triple resonance experiments, in combination with a 3D TOCSY- and NOESY-[^{1}H,^{15}N]-HSQC, all collected at 500 MHz. Side chain carbons were assigned from a 700 MHz (H)CCH-TOCSY (*36*) spectrum, while side chain protons were assigned from a 700 MHz H(C)CH-TOCSY (*36*) spectrum and a 500 MHz 2D TOCSY acquired in D_2O on an unlabeled sample. The stereospecific assignment of Cδ of leucines and Cγ of valines was performed by analyzing the 750 MHz [^{1}H,^{13}C]-HSQC spectrum of ^{15}N,10% ^{13}C-labeled CSP*sg*4, as described (*37*). NOE-based distance restraints were extracted from a 500 MHz 2D NOESY spectrum, a 500 MHz 3D NOESY-[^{1}H,^{15}N]-HSQC spectrum, and a 700 MHz 3D NOESY-[^{1}H,^{13}C]-HSQC spectrum, all with a mixing time of 100 ms. A total of 127 dihedral angle constraints were derived from chemical shift values using TALOS (*31*); additional restraints were obtained from 47 $^3J_{HNH\alpha}$ coupling constants extracted from a 500 MHz quantitative *J*-correlated HNHA experiment (*38*).

Structure Calculation. Structure calculations were performed by simulated annealing in torsion angle space using the CYANA 2.0 package (*39*), which implements an efficient protocol for structure calculation/automated assignment of NOEs (*40*). The standard annealing protocol was used with 20000 steps of torsion angle dynamics; in each of seven cycles, 100 structures were calculated, and the 20 with the lowest target function were used in the next stage. Thirty-six NOEs were assigned manually and imposed in all stages of the protocol, while all other NOE assignments accrued in the course of the automatic procedure, resulting in a total of 1931 upper distance bounds. In the final run, 200 structures were computed, and the 40 with lowest target function were refined by 4000 steps of restrained minimization in the more realistic AMBER99 force field (*41*), using an implicit solvation model (*42*, *43*) to account for the influence of solvent. The AMBER-minimized structures were sorted by increasing restraint violation, and the first 20 were selected as the representative bundle. Quality checks were done in PROCHECK-NMR (*44*), while MOLMOL (*45*) was used for structure visualization and analysis. Surface pockets and internal cavities were identified and measured with the CASTp protocol (*46*, *47*).

Titration Experiments. Ligand binding was followed by measuring changes in the 1D proton spectrum and the 2D [^{1}H-^{15}N]-HSQC of a sample containing 1 mM ^{15}N-labeled CSP*sg*4. Prescored amounts of oleamide (99.9%, Sigma-Aldrich), each corresponding to 1 molar equiv, were dissolved in methanol and taken to dryness in vials. After reference spectra of the free protein were collected, the sample was transferred in turn to one of the vials, incubated for 30 min at 315 K, cooled at 300 K, placed again in the NMR tube, and examined. Samples with a 1:1, 2:1, and 3:1 oleamide:protein molar ratio were examined. A final titration point was obtained by incubation of the sample with a 12 molar excess of oleamide.

RESULTS

Resonance Assignment. All of the peaks that appear in the [^{1}H,^{15}N]-HSQC spectrum of CSP*sg*4 (available as Supporting Information, Figure 1S) were assigned unambiguously; the backbone assignment was 99% complete, since neither the backbone nor the side chain of D9 could be assigned, possibly due to strong overlap with other signals. The resonance assignment was 98.5% complete for carbons, 98.1% for nitrogens, and 99.5% for protons. The methyl groups of all 5 valines and 11 out of 12 leucines (except L107) have been stereospecifically assigned. The 3 prolines have been all assigned to the trans conformation, based on Cβ and Cγ chemical shift values (*48*). The $^3J_{HNH\alpha}$ coupling constants, as well as the dihedral intervals computed by the program TALOS (*49*) based on the chemical shifts values of Cα and CO, were used as additional constraints during structure calculation.

Three-Dimensional Structure of CSPsg4. Table 1 lists the constraints used for structure calculations, together with the statistical analysis of the final structures. A total of 1931 upper limit distance restraints as well as 127 TALOS and 47 $^3J_{HNH}\alpha$ coupling restraints were used, corresponding to an average of 19 constraints per residue. After restrained energy minimization in the AMBER99 force field (*41*) in the presence of an implicit solvent model (*42*), the final ensemble showed only minor restraint violations. The structure was very well defined: the RMSD of the structured regions (residues 13−103, vide infra) was 0.24 Å for the backbone and 0.72 Å for all heavy atoms. 99.8% of all residues were located in the most favored and additionally allowed regions of Ramachandran plot, and only 0.2% fell in the generously allowed regions, these latter being all located in loops or terminals. The equivalent resolution of the structured region, as assessed by PROCHECK-NMR (*44*) based on Ramachandran plot quality and on hydrogen bond energies, was of 1.5−1.6 Å. Figure 1A shows the bundle of the 20 lowest energy structures.

The structure is characterized by an overall fold consisting of six α-helices connected by loops. The helices were identified by analysis of the backbone dihedrals, with support from the STRIDE (*50*) and PROMOTIF (*51*) protocols, and were named on the basis of a comparison with the secondary structure elements of CSPMbraA6 (*24*, *25*); in CSP*sg*4, they span residues 13−18 (helix A), 20−31 (B), 40−54 (C), 62−78 (D), 80−90 (E), and 97−103 (F). A stereoview of the lowest energy structure in the bundle is reported in Figure 1B. In several models, the beginning of helix C was slightly irregular at the site of residue D41. Moreover, due to the presence of a proline at position 50, a certain distortion was also evident in the middle of helix C. Residues 36−37, 58−59, and 92−93 showed dihedral angles and NOE patterns indicative of a type I β-turn; in most structures, residues 95 and 96 also folded into a type I β-turn, whereas a type I′ β-turn was present between residues 56 and 57. As in the case of CSPMbraA6, the four helices, A, B, D, and E, are arranged to form a V-shaped motif. Helix C runs across the two V's, while helix F is packed against the external face of the V formed by helices D and E (Figure 1B).

Table 1: Summary of Residual Constraint Violations, Energies, and Statistics for the Final CSP*sg*4 Structure Ensemble (20 Structures)

unambiguous restraints	
intraresidue	457
sequential	507
medium range[a]	615
long range	352
total n. of contacts	1931
other restraints	
TALOS dihedral angles	127
$^3J_{HNH\alpha}$ couplings (φ angles)	47
av n. of distance restraint violations (Å)	
$0.1 < d \leq 0.2$	9.7 ± 2.1
$0.2 < d$	0
av max violation (Å)	0.16 ± 0.02
av n. of torsional restraint violations (deg)	
$5 < \theta \leq 10$	2.2 ± 1.1
$10 < \theta$	0
av max violation (deg)	6.6 ± 1.0
av AMBER energies (kcal/mol)	
E (AMBER)[b]	-4289.5 ± 15.5
E (distance constraint)	16.5 ± 0.6
E (torsional constraint)	1.0 ± 0.2
E (van der Waals)	-780.0 ± 7.1
E (total)	-4272.0 ± 15.7
Ramachandran plot	
most favored region (%)	99.8
generously allowed region (%)	0.2
RMSD from mean (Å)	
backbone atoms (N, Cα, C′)[c]	0.24
all heavy atoms[c]	0.72

[a] $2 < |i - j| < 4$. [b] The force constants for the distance (torsional) restraints were 32 kcal mol^{-1} A^{-2} (16 kcal mol^{-1} rad^{-2}). Errors are given as ±1 standard deviation. [c] Residues L13−R103.

The arrangement of secondary structure elements in the fold of CSP*sg*4 corresponds rather closely to that observed in the NMR structure of CSPMbraA6, although helix B and helix E are longer by one and three residues, respectively, while helix F is five residues shorter. Figure 2A shows that the best alignment of CSPMbraA6 and CSP*sg*4 is obtained when a gap of two amino acids is inserted between residues 32 and 33 of CSPMbraA6. It is also evident that the amino acid sequences of the two proteins are very similar, the number of identical residues corresponding to 52%.

For a quantitative comparison of the 3D structures of CSPMbraA6 and CSP*sg*4, we decided to rely on the X-ray structure of CSPMbraA6 (*25*) (PDB ID 1KX9). This structure features longer helices, particularly helix F, with respect to both the solution structure of CSPMbraA6 solved by the same group (*24*) (PDB ID 1K19) and CSP*sg*4. These differences are presumably related to an increased flexibility/disorder in solution, and admittedly they introduce an element of complication in the comparison between the solution structure of CSP*sg*4 and the X-ray structure of CSPMbraA6. However, the alternative of using the solution structure of CSPMbraA6 as a reference was still less tempting, given its

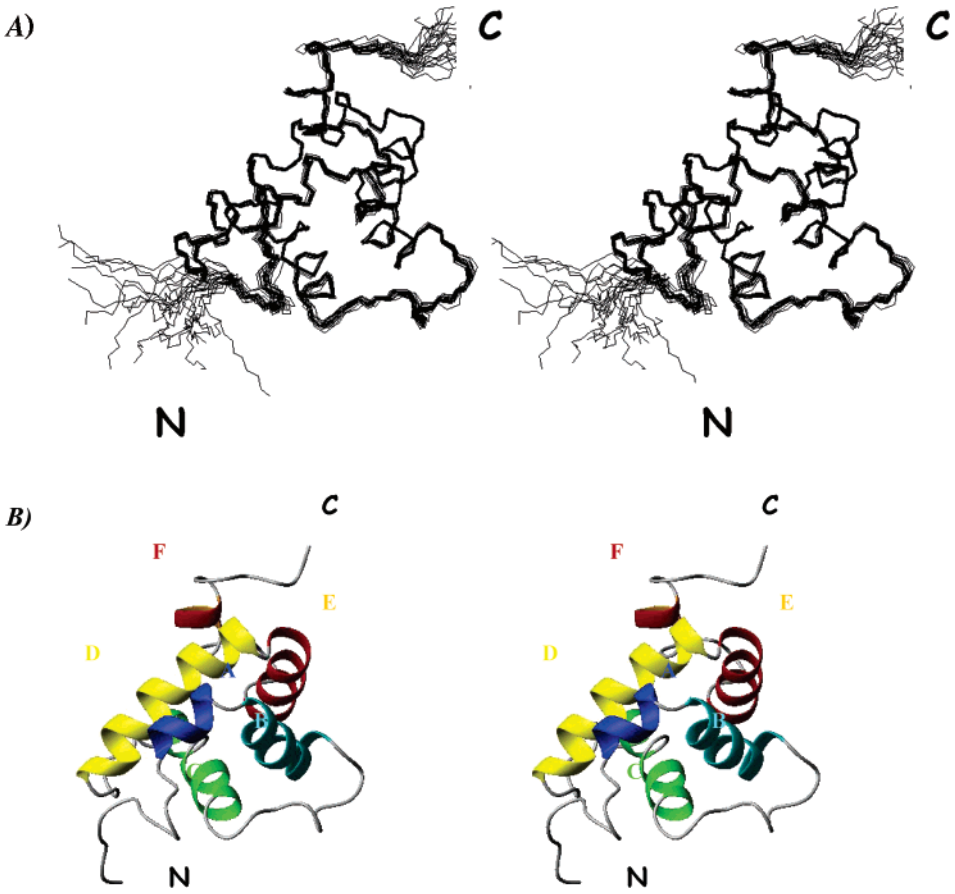

FIGURE 1: Stereoview of the solution structure of CSP*sg*4 after AMBER minimization. Panel A: Bundle of the 20 lowest energy structures superimposed for backbone atoms of residues 13−103. The RMSD was 0.24 Å. Panel B: Ribbon representation showing the helices named on the basis of a comparison with the secondary structure elements of CSPMbraA6 (*24*, *25*). They correspond to residues 13−18 (helix A, dark blue), 20−31 (helix B, light blue), 40−54 (helix C, green), 62−78 (helix D, yellow), 80−90 (helix E, orange), and 97−103 (helix F, red).

poor resolution, which probably contributed to the large backbone RMSD observed with respect to both CSPMbraA6 in the solid state (3.40 Å for residues 13−101) and CSP*sg*4 (3.49 Å for residues 13−31, 36−103).

Superimposition of the backbone of CSP*sg*4 (residues 13−31, 36−103) with monomer A of the crystal form 1 (residues 13−31, 34−101) (*25*) resulted in an RMSD of 2.21 Å, which confirms the visually apparent close similarity between the two proteins (Figure 2B). Interestingly, a still lower backbone RMSD (2.09 Å) is obtained when comparing our structure with that of the complex CSPMbraA6−bromododecanol determined by X-ray crystallography (*26*) (PDB ID 1N8V; see Figure 2C), which features an extensive swelling of the internal cavity and partial unfolding of the N-terminal helix. As an aid for the interpretation of binding experiments (vide infra), several pockets and cavities were identified in the structure of CSP*sg*4 by means of specific analysis tools (*46*, *47*). Three of these cavities are depicted in Figure 3 with different colors: in the reference structure, the largest pocket (red, 290 Å^3, computed using Connolly's molecular surface) features a single large mouth and is mostly made up of hydrophobic residues belonging to helices A, C, and D, as well as to the N-terminal; however, across the members of the bundle there is remarkable variability both in the volume and in the exact position of this pocket. By contrast, the second largest pocket is an internal, closed cavity (green, 225 Å^3) and is quite well conserved in the whole bundle. Adjacent to it, but extending to the protein surface, is the third pocket shown (blue, 165 Å^3), which again shows appreciable modifications in other individual bundle structures. At this stage of analysis, several other pockets identified appear less interesting, because of their much smaller volumes and/or lack of reproducibility within the bundle.

Binding Studies. Despite their overall very high structural similarity, CSP*sg*4 and CSPMbraA6 show specific affinities for different ligands (*25*, *52*). To better understand the structural features that determine this specific selectivity in molecular recognition, we followed by NMR a titration of CSP*sg*4 with oleamide, the main endogenous ligand of a CSP in the very closely related species *Locusta migratoria* (*28*). Chemical shift perturbations are very sensitive even to subtle conformational effects and represent a suitable method to study weak interactions, such as those expected for binding and delivery of chemical stimuli. One-dimensional ^{1}H NMR spectra acquired at 1:1, 2:1, and 3:1 ligand:protein molar ratios are reported as Supporting Information (Figure 2S). A comparison with the spectrum of the free protein shows very clearly that, upon addition of oleamide, the intensity of the two signals in the extreme regions, i.e., $H^{\gamma 13}$ of I76 at high field and $H^{\epsilon 1}$ of W83 at low field, is significantly decreased, suggesting that these two residues are involved in oleamide binding. Figure 4 shows the [^{1}H−^{15}N]-HSQC spectrum acquired upon prolonged incubation at 310 K in the presence of an excess of oleamide: the substantial

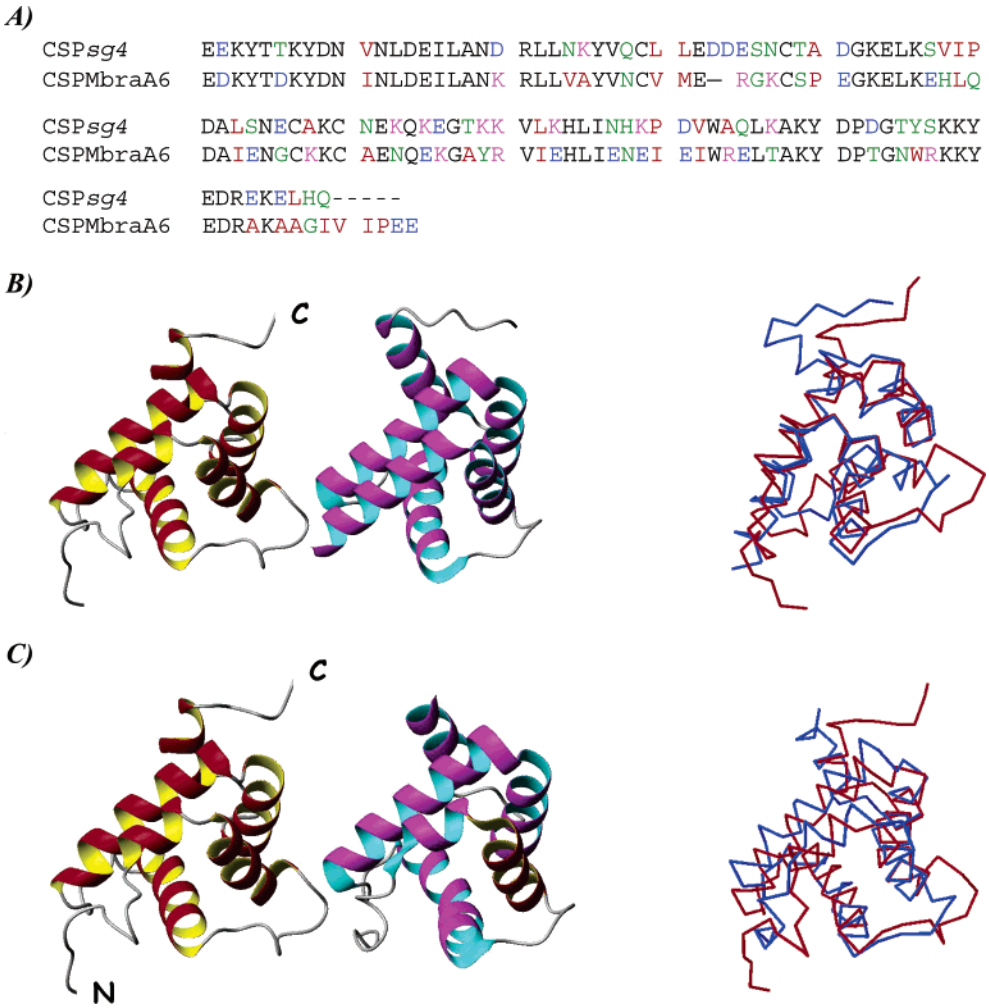

FIGURE 2: Structural comparison of CSP*sg*4 and CSPMbraA6. Panel A: Alignment of amino acid sequences. Panel B: NMR structure of CSP*sg*4 (first structure of the bundle, left) and monomer A of the crystal form 1 of CSPMbraA6 (PDB ID 1KX9, center). The overlay of the same structures (RMSD = 2.21 Å) is reported as a line representation on the right side (with CSP*sg*4 in red and CSPMbraA6 in blue). Panel C: NMR structure of CSP*sg*4 (first structure of the bundle, left) and bromododecanol-bound CSPMbraA6 (PDB ID 1N8V, center). The overlay of the same structures (RMSD = 2.09 Å) is reported as a line representation on the right side (with CSP*sg*4 in red and CSPMbraA6 in blue).

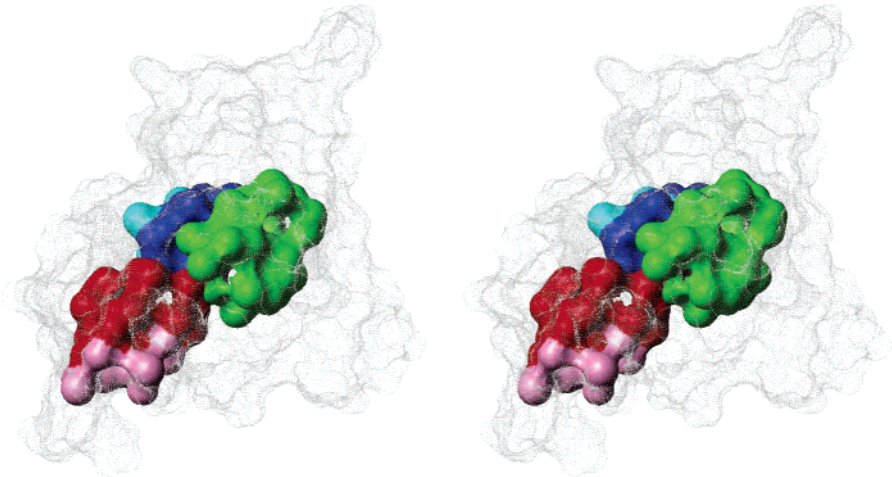

FIGURE 3: Selected surface pockets and internal cavities of CSP*sg*4. The protein (dashed gray surface) is shown in the same orientation as in panel B of Figure 2; the pockets are highlighted in color, with their mouth regions in a lighter shade. Pockets: 1, red; 2, green;t 3, blue.

differences with respect to the free protein suggest the occurrence of an extensive conformational transition. However, [^{1}H−^{15}N]-HSQC spectra corresponding to the intermediate stages (data not shown) display extensive peak doubling and broadening, typical of chemical processes occurring on an intermediate/slow exchange regime, which prevents a straightforward identification of all the residues involved in binding site.

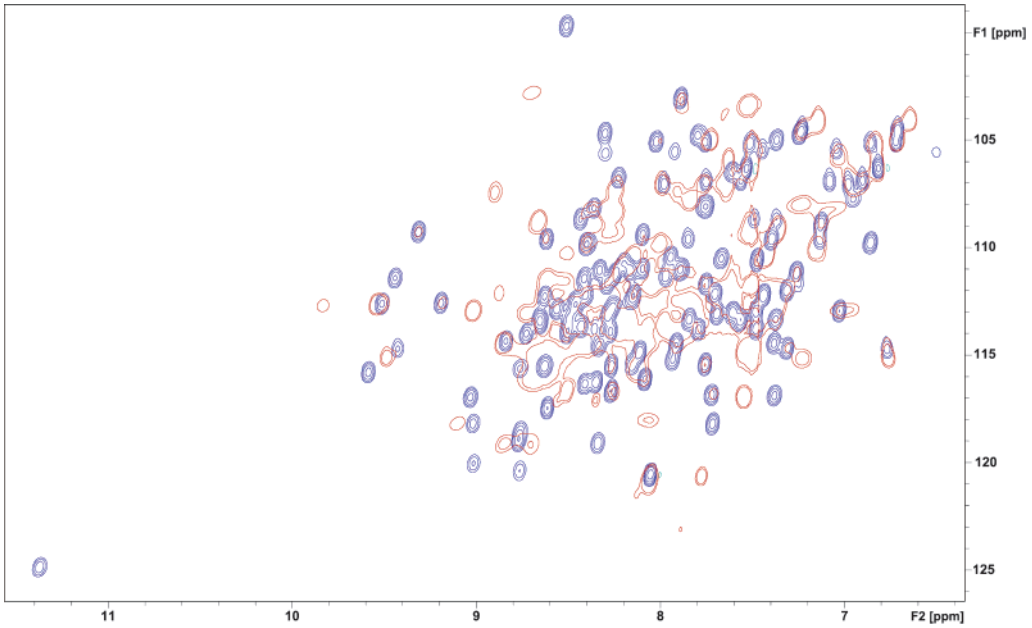

FIGURE 4: Overlay of the 600 MHz [^{1}H−^{15}N]-HSQC spectra of free (red) and oleamide-bound (blue) CSP*sg*4.

DISCUSSION

In this study, we have reported the solution structure of CSP*sg*4 from *S. gregaria*, as determined by NMR spectroscopy. Our results show a general fold very close to the CSP of *M. brassicae*, the only other structure available for proteins of this class (Figure 2) (*25*). The observed structural similarity between the orthopteran and lepidopteran proteins is hardly surprising, given their sequence homology (Figure 2). In the case of the two above-mentioned CSPs, the sequence identity corresponds to 52%, but in general CSPs are well conserved even across very distant insect orders. This finding suggests that the difference in ligand binding may be determined by the presence of a few specific amino acids located in the protein hydrophobic cavities. On the other hand, even OBPs, the other class of soluble binding proteins in insect chemoreception, show a very conserved folding, as documented by six structures solved so far, despite a poor sequence similarity, limited, in some cases, to the motif of six cysteines (*53*).

Despite the high overall structural similarity, several differences were evident between CSP*sg*4 and CSPMbraA6 that may account for protein specificities toward the natural ligand. First, the structure of CSP*sg*4 is less ordered than the X-ray structure of CSPMbraA6 taken as a reference. This could be partly attributed to the solution state, although the backbone overlap is definitely better with the X-ray structure of CSPMbraA6 than with its NMR structure. Interestingly, the closest fit of backbone atoms (2.09 Å) corresponds to the overlay of the first structure in the bundle of the NMR structures of CSP*sg*4 with the bromododecanol-bound form of CSPMbraA6 (PDB ID 1N8V), which is remarkably swollen in the unbound protein and features a more disordered N-terminal helix.

Other differences were observed at the level of surface pockets and internal cavities, which were significantly larger in our protein. The bromododecanol-bound form of CSPMbraA6 features a largely predominant pocket, which accommodates three ligand molecules, and results from the coalescence of several distinct pockets/cavities of the free form, accompanied by extensive rearrangement. As a matter of fact, the extent of rearrangement is such that it would be difficult to recognize a preformed ligand binding cavity in the structure of the free form. By analogy, one can surmise that even in the case of CSP*sg*4 the pocket/cavity features of the free form will be profoundly altered by ligand binding. This may well explain the far-reaching NMR changes that occur on interaction with oleamide. The type of peak doubling and broadening that accompanies the intermediate stages of titration with oleamide is characteristic of a ligand exchange kinetics slow to intermediate on the chemical shift time scale. Since the NMR chemical shifts do not exhibit a continuous change over the course of a protein−ligand titration, a straightforward transfer of the assignments of the free form onto the bound form is not possible, and more demanding *de novo* assignment strategies will be required to clarify the structure of the bound form. However, the NMR titration data clearly point to a specific interaction with oleamide that reaches a final plateau characterized by a well-defined, although somewhat broadened, [^{1}H−^{15}N]-HSQC spectrum.

Oleamide was the only endogenous ligand isolated from a pool of CSPs purified from the wings of the very closely related locust species *L. migratoria* (*28*). In addition to other straight-chain amides, the same compound was also identified as a natural ligand of OBPs from the wings of the paper wasp *Polistes dominulus* (*27*). The ecological role of oleamide, and its effect on the locust, is still unclear; similarly, the knowledge on putative locust pheromones is very poor and fragmentary. Simple aromatic derivatives, such as phenylacetonitrile and substituted phenols and anisoles, have been isolated from the feces of *S. gregaria* and suggested as potential aggregation pheromones for this species. In the past, we failed to measure any appreciable binding of these compounds to CSPs of *S. gregaria* and to the close species *L. migratoria*, nor to OBPs of the latter species (*28*, *54*). On the contrary, our binding experiments indicated some large aromatic compounds as strong ligands,

namely, various phthalates and α-amylcinnamaldehyde; such compounds were not found among natural semiochemicals (*52*, *54*).

If the physiological role of OBPs in insect olfaction is still a matter of discussion, there is not even a wide agreement on CSPs involvement in the chemosensory system. However, their high production and secretion into the lymph of chemosensilla, together with the demonstration of their binding activity to small organic molecules, strongly suggest that these polypeptides could play a role similar to that of OBPs in chemoreception (*10*). Moreover, protein members associated with nonsensory organs, such as the CSPs isolated from the wings of locusts, could be involved in the gradual and controlled release of volatile pheromones. In this context, they could perform a function similar to that of the urinary, salivary, or vaginal proteins described in several mammals (*55*). The elucidation of the three-dimensional structure of different members of the CSP family will help in understanding functional interactions between these proteins and their ligands.

ACKNOWLEDGMENT

The 700 and 750 MHz 2D and 3D spectra were recorded at the SON NMR Large Scale Facility in Utrecht, which is funded by the Access to Research Infrastructures program of the European Union (HPRI-CT-2001-00172). We thank the CIMCF of the University of Naples Federico II where some preliminary NMR experiments were carried out.

SUPPORTING INFORMATION AVAILABLE

Figure 1S, [^{1}H−^{15}N]-HSQC spectrum of CSP*sg*4 acquired at 300 K, showing the assignment of the amide cross-peaks, and Figure 2S, 1D proton NMR spectra of CSP*sg*4 upon addition of the ligand at a 1:1, 1:2, and 1:3 molar ratio (traces B, C, and D, respectively), with trace A reporting the spectrum of the free protein. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Clyne, P. J., Warr, C. G., Freeman, M. R., Lessing, D., Kim, J., and Carlson, J. R. (1999) A novel family of divergent seven-transmembrane proteins: candidate odorant receptors in *Drosophila*, *Neuron 22*, 327−338.
2. Vosshall, L. B., Amrein, H., Morozov, P. S., Rzhetsky, A., and Axel, R. (1999) A spatial map of olfactory receptor expression in the *Drosophila* antenna, *Cell 96*, 725−736.
3. Pelosi, P., and Maida, R. (1995) Odorant-binding proteins in insects, *Comp. Biochem. Physiol. 111B*, 503−514.
4. Pelosi, P. (1998) Odorant-binding proteins: structural aspect, *Ann. N.Y. Acad. Sci. 855*, 281−293.
5. Steinbrecht, R. (1998) Odorant-binding proteins: expression and function, *Ann. N.Y. Acad. Sci. 855*, 323−332.
6. Picimbon, J. F. (2003) Biochemistry and evolution of OBP and CSP proteins, in *Insect Pheromone Biochemistry and Molecular Biology* (Blomquist, G. J., and Vogt, R. G., Eds.) pp 539−566, Elsevier Academic Press, London.
7. Vogt, R. G. (2003) Biochemical diversity of odor detection: OBPs, ODEs and SNMPs, in *Insect Pheromone Biochemistry and Molecular Biology* (Blomquist, G. J., and Vogt, R. G., Eds.) pp 391−446, Elsevier Academic Press, London.
8. Vogt, R. G. (2005) Molecular basis of pheromone detection in insects, in *Comprehensive Insect Physiology, Biochemistry, Pharmacology and Molecular Biology. Volume 3. Endocrinology* (Gilbert, L. I., Iatro, K., and Gill, S., Eds.) pp 753−804, Elsevier, London.
9. Wanner, K. W., Willis, L. G., Theilmann, D. A., Isman, M. B., Feng, Q., and Plettner, E. (2004) Analysis of the insect OS-D-like gene family, *J. Chem. Ecol. 30*, 889−911.
10. Pelosi, P., Zhou, J. J., Ban, L. P., and Calvello, M. (2006) Soluble proteins in insect chemical communication, *Cell. Mol. Life Sci. 63* (in press.)
11. Wetzel, C. H., Behrendt, H. J., Gisselmann, G., Stortkuhl, K. F., Hovemann, B,. and Hatt, H. (2001) Functional expression and characterization of a *Drosophila* odorant receptor in a heterologous cell system, *Proc. Natl. Acad. Sci. U.S.A. 98*, 9377−9380.
12. Hallem, E. A., Fox, A. N., Zwiebel, L. J., and Carlson, J. R. (2004) Olfaction: mosquito receptor for human-sweat odorant, *Nature 427*, 212−213.
13. Xu, P., Atkinson, R., Jones, D. N., and Smith, D. P. (2005) *Drosophila* OBP LUSH is required for activity of pheromone-sensitive neurons, *Neuron 45*, 193−200.
14. Klein, U. (1987) Sensillum-lymph proteins from antennal olfactory hairs of the moth *Antheraea polyphemus* (Saturniidae), *Insect Biochem. 17*, 1193−1204.
15. Lee, D., Damberger, F. F., Peng, G., Horst, R., Guntert, P., Nikonova, L., Leal, W. S., and Wüthrich, K. (2002) NMR structure of the unliganded *Bombyx mori* pheromone-binding protein at physiological pH, *FEBS Lett. 531*, 314−318.
16. 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.
17. Horst, R., Damberger, F., Luginbuhl, P., Guntert, P., Peng, G., Nikonova, L., Leal, W. S., and Wüthrich K. (2001) NMR structure reveals intramolecular regulation mechanism for pheromone binding and release, *Proc. Natl. Acad. Sci. U.S.A. 98*, 14374−14379.
18. Lartigue, A., Gruez, A., Spinelli, S., Riviere, S., Brossut, R., Tegoni, M., and Cambillau, C. (2003) The crystal structure of a cockroach pheromone-binding protein suggests a new ligand binding and release mechanism, *J. Biol. Chem. 278*, 30213−30218.
19. Kruse, S. W., Zhao, R., Smith, D. P., and Jones, D. N. (2003) Structure of a specific alcohol-binding site defined by the odorant binding protein LUSH from *Drosophila melanogaster*, *Nat. Struct. Biol. 10*, 694−700.
20. Mohanty, S., Zubkov, S., and Gronenborn, A. M. (2004) The solution NMR structure of *Antheraea polyphemus* PBP provides new insight into pheromone recognition by pheromone-binding proteins, *J. Mol. Biol. 337*, 443−451.
21. Lartigue, A., Gruez, A., Briand, L., Blon, F., Bezirard, V., Walsh, M., Pernollet, J. C., Tegoni, M., and Cambillau, C. (2004) Sulfur single wavelength anomalous diffraction crystal structure of a pheromone-binding protein from the honeybee *Apis mellifera* L., *J. Biol. Chem. 279*, 4459−4464.
22. Wogulis, M., Morgan, T., Ishida Y., Leal, W. S., and Wilson, D. K. (2005) The crystal structure of an odorant binding protein from *Anopheles gambiae*: Evidence for a common ligand release mechanism, *Biochem. Biophys. Res. Commun. 339*, 157−164.
23. Zubkov, S., Gronenborn, A. M., Byeon, I. J., and Mohanty, S. (2005) Structural consequences of the pH-induced conformational switch in *A. polyphemus* pheromone-binding protein: mechanisms of ligand release, *J. Mol. Biol. 354*, 1081−1090.
24. Mosbah, A., Campanacci, V., Lartigue, A., Tegoni, M., Cambillau, C., and Darbon, H. (2003) Solution structure of a chemosensory protein from the moth *Mamestra brassicae*, *Biochem. J. 369*, 39−44.
25. 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.
26. Campanacci, V., Lartigue, A., Hallberg, B. M., Jones, T. A., Giudici-Orticoni, M. T., Tegoni, M., and Cambillau, C. (2003) Moth chemosensory protein exhibits drastic conformational changes and cooperativity on ligand binding, *Proc. Natl. Acad. Sci. U.S.A. 100*, 5069−5074.
27. Calvello, M., Guerra, N., Brandazza, A., D'Ambrosio, C., Scaloni, A., Dani, F. R., Turillazzi, S., and Pelosi, P. (2003) Soluble proteins of chemical communication in the social wasp *Polistes dominulu*, *Cell. Mol. Life Sci. 60*, 1933−1943.
28. Ban, L. P., Scaloni, A., Brandazza, A., Angeli, S., Zhang, L., Yan, Y. H., and Pelosi, P. (2003). Chemosensory proteins of *Locusta migratoria*, *Insect Mol. Biol. 12*, 125−134.
29. 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.
30. Picone, D., Crescenzi, O., Angeli, S., Marchese, S., Brandazza, A., Ferrara, L., Pelosi, P., and Scaloni, A. (2001) Bacterial expression and conformational analysis of a chemosensory protein from *Schistocerca gregaria*, *Eur. J. Biochem. 268*, 4794−4801.
31. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes, *J. Biomol. NMR 6*, 277−293.
32. Johnson, B. A., and Blevins, R. A. (1994) NMRView: A computer program for the visualization and analysis of NMR data, *J. Biomol. NMR 4*, 603−614.
33. Kay, L. E., Ikura, M., Tschudin, R., and Bax, A. (1990) Three-dimensional triple-resonance spectroscopy of isotopically enriched proteins, *J. Magn. Reson. 89*, 496−514.
34. Wittekind, M., and Mueller, L. (1993) HNCACB, a high-sensitivity 3D NMR experiment to correlate amide-proton and nitrogen resonances with the alpha- and beta-carbon resonances in proteins, *J. Magn. Reson. B 101*, 201−205.
35. Grzesieck, S., and Bax, A. (1992) Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR, *J. Am. Chem. Soc. 114*, 6291−6293.
36. Kay, L. E., Xu, G. Y., Singer, A. U., Muhandiram, D. R., and Kay, J. D. F. (1993) A gradient-enhanced HCCH-TOCSY experiment for recording side-chain ^{1}H and ^{13}C correlations in H_2O samples of proteins, *J. Magn. Reson. 101*, 333−337.
37. Hilty, C., Wider, G., Fernandez, C., and Wüthrich, K. (2003) Stereospecific assignments of the isopropyl methyl group of the membrane protein OmpX in DHPC micelles, *J. Biomol. NMR 27*, 377−382.
38. Bax, A., Vuister, G. W., Grzesiek, S., Delaglio, F., Wang, A. C., Tschudin, R., and Zhu, G. (1994) Measurement of homo- and heteronuclear J couplings from quantitative J correlation, *Methods Enzymol. 239*, 79−105.
39. 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.
40. Güntert, P. (2003) Automated NMR protein structure calculation, *Prog. NMR Spectrosc. 43*, 105−125.
41. Wang, J., Cieplak, P., and Kollman, P. A. (2000) How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules, *J. Comput. Chem. 24*, 1999−2012.
42. Hawkins, G. D., Cramer, C. J., and Truhlar, D. G. (1995) Pairwise solute screening of solute charges from a dielectric medium, *Chem. Phys. Lett. 246*, 122−129.
43. Hawkins, G. D., Cramer, C. J., and Truhlar, D. G. (1996) Parameterized models of aqueous free energies of solvation based on pairwise descreening of solute atomic charges from a dielectric medium, *J. Phys. Chem. 100*, 19824−19839.
44. Laskowski, R. A., Rullmann, J. A. C., 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.
45. Koradi, R., Billeter, M., and Wüthrich, K. (1996) MOLMOL: a program for display and analysis of macromolecular structures, *J. Mol. Graphics 14*, 51−55.
46. Binkowski, T. A., Naghibzadeh, S., and Liang, J. (2003) CASTp: Computed atlas of surface topography of proteins, *Nucleic Acids Res. 31*, 3352−3355.
47. Liang, J., Edelsbrunner, H., and Woodward, C. (1998) Anatomy of protein pockets and cavities: measurement of binding site geometry and implications for ligand design, *Protein Sci. 7*, 1884−1897.
48. Dorman, E. D., and Bovey, F. A. (1973) Proton coupled carbon-13 magnetic resonance spectra. The simple amides, *J. Org. Chem. 38*, 2379−2383.
49. Cornilescu, G., Delaglio, F., and Bax, A. (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology, *J. Biomol. NMR 13*, 289−302.
50. Frishman, D., and Argos, P. (1995) Knowledge-based secondary structure assignment, *Proteins 23*, 566−579.
51. Hutchinson, E. G., and Thornton, J. M. (1996) PROMOTIF—A program to identify and analyze structural motifs in proteins, *Protein Sci. 5*, 212−220.
52. 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.
53. Tegoni, M., Campanacci, V., and Cambillau, C. (2004) Structural aspects of sexual attraction and chemical communication in insects, *Trends Biochem. Sci. 29*, 257−264.
54. Ban, L. P., Scaloni, A., D'Ambrosio, C., Zhang, L., Yan, Y. H., and Pelosi, P. (2003) Biochemical characterisation and bacterial expression of an odorant-binding protein from *Locusta migratoria*, *Cell. Mol. Life Sci. 60*, 390−400.
55. Pelosi, P. (2001) The role of perireceptor events in vertebrate olfaction, *Cell. Mol. Life Sci. 58*, 503−509.

BI060998W