NMR Solution Structure of α-Conotoxin ImI and Comparison to Other Conotoxins Specific for Neuronal Nicotinic Acetylcholine Receptors[†,‡]

Jessica P. Rogers,[§] Peter Luginbühl,[§] Gregory S. Shen,[||] R. Tyler McCabe,[||] Raymond C. Stevens,*[,§] and David E. Wemmer*[,§]

Department of Chemistry, University of California, Berkeley, California 94720, and Cognetix Inc., Salt Lake City, Utah 84108

Received November 4, 1998; Revised Manuscript Received December 11, 1998

ABSTRACT: α-Conotoxins, peptides produced by predatory species of *Conus* marine snails, are potent antagonists of nicotinic acetylcholine receptors (nAChRs), ligand-gated ion channels involved in synaptic transmission. We determined the NMR solution structure of the smallest known α-conotoxin, ImI, a 12 amino acid peptide that binds specifically to neuronal α7-containing nAChRs in mammals. Calculation of the structure was based on a total of 80 upper distance constraints and 31 dihedral angle constraints resulting in 20 representative conformers with an average pairwise rmsd of 0.44 Å from the mean structure for the backbone atoms N, C^{α}, and C′ of residues 2−11. The structure of ImI is characterized by two compact loops, defined by two disulfide bridges, which form distinct subdomains separated by a deep cleft. Two short 3_{10}-helical regions in the first loop are followed by a C-terminal *β*-turn in the second. The two disulfide bridges and Ala 9 form a rigid hydrophobic core, orienting the other amino acid side chains toward the surface. Comparison of the three-dimensional structure of ImI to those of the larger, 16 amino acid α-conotoxins PnIA, PnIB, MII, and EpI—also specific for neuronal nAChRs—reveals remarkable similarity in local backbone conformations and relative solvent-accessible surface areas. The core scaffold is conserved in all five conotoxins, whereas the residues in solvent-exposed positions are highly variable. The second helical region, and the specific amino acids that the helix exposes to solvent, may be particularly important for binding and selectivity. This comparative analysis provides a three-dimensional structural basis for interpretation of mutagenesis data and structure−activity relationships for ImI as well other neuronal α-conotoxins.

Over the past 50 million years, marine snails in the genus *Conus* have developed venoms containing a remarkable array of small peptide neurotoxins, highly specific for different subtypes of ligand- and voltage-gated ion channels in the central and peripheral nervous systems. These peptide toxins exhibit potent pre- and postsynaptic activity not only in the species' natural prey—including worms, mollusks, and fish—but in mammals as well (*1*). Venom from each *Conus* species contains a variety of biologically active peptides, many of which are 10−30 amino acids in length and rich in disulfide bonds (*2*). These conotoxins are translated as larger preproproteins with distinct conserved and hypervariable regions (*2*−*5*). The conserved regions include the N-terminal signal sequence and propeptide regions as well as the C-terminal cysteine residues, while the loops between cysteine residues comprise the hypervariable regions and contribute to the pharmacological diversity (*3, 4*). The conserved regions may influence the formation of specific disulfide bonds and the folding in the C-terminal region, which is proteolytically cleaved to form the small, stable mature toxin (*2*−*5*).

Classification of the known disulfide-rich conotoxins based on their activity and primary structure results in 10 groups (reviewed in refs *6* and *7*): α- and αA-conotoxins are competitive nicotinic acetylcholine receptor (nAChR) antagonists, whereas *ψ*-conotoxins are noncompetitive nAChR antagonists; *μ*-, *μ*O-, and *δ*-conotoxins target different sites on the sodium channel, with separate effects of inhibition and activation; *ω*-, *κ*-, and *κ*A-conotoxins inhibit voltage-gated calcium and potassium channels, respectively; *σ*-conotoxins inhibit the 5-HT_3 (serotonin) receptor. Thus far, only four primary patterns of disulfide bonding have been found in the conotoxins (Figure 1a), resulting in three general topologies of two, three, or four loops. Beyond that, great variation in loop length and amino acid sequence is found in all classes of the conotoxins, even among those which target the same receptor. The basis of the conotoxins' ability to distinguish and exhibit preferential affinity for subtypes of their target molecules lies in this variable loop structure, which may be explored at both the primary and tertiary levels.

The present study focuses on the structures of the α-conotoxins, which target subtypes of either skeletal muscle

[†] This work was supported by the Director, Office of Energy Research and Office of Biological and Environmental Research, U.S. Department of Energy, under Contract DE-AC03-76SF00098 and through instrumentation grants from the U.S. Department of Energy (DE-FG05-86ER75281) and the National Science Foundation (DMB 86-09305 and BBS 87-20134). D.E.W. was supported by the Miller Foundation during the period of this work. J.P.R. was supported by National Institutes of Health Molecular Biophysics Training Grant GM08295.

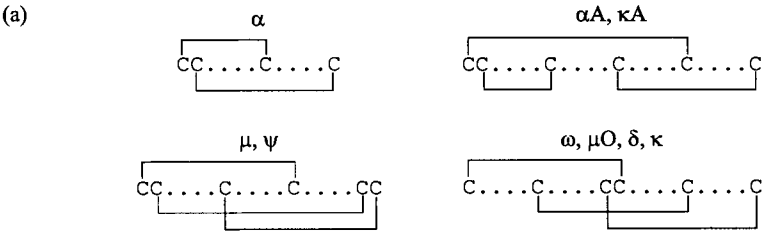
[‡] Coordinates, chemical shifts, and conformational restraints for α-conotoxin ImI have been deposited in the Brookhaven Protein Data Bank under the accession code 1IM1.

* To whom correspondence should be addressed. E-mail: R.C.S., stevens@adrenaline.berkeley.edu; D.E.W., dewemmer@lbl.gov.

[§] University of California, Berkeley.

[||] Cognetix Inc.

10.1021/bi9826254 CCC: $18.00
Published on Web 03/13/1999



(b)

Species	α–Conotoxin	Sequence	Specificity	Reference
C. imperialis (worm-hunting)	ImI	GCCSDPRCA--W-R-C	α7	(*24, 25*)
C. episcopatus (mollusk-hunting)	EpI	GCCSDPRCNMNNPDYC	α3β2	(*74*)
C. pennaceus (mollusk-hunting)	PnIA	GCCSLPPCAANNPDYC	α3β2	(*75, 69*)
C. pennaceus (mollusk-hunting)	PnIB	GCCSLPPCALSNPDYC	α7	(*75, 69*)
C. magus (fish-hunting)	MII	GCCSNPVCHLEHSNLC	α3β2	(*29*)

FIGURE 1: (a) Disulfide bonding patterns for the known classes of conotoxins. The pattern for the *σ*-conotoxins, which contain five disulfide bonds, has not yet been determined. (b) Sequences of the α-conotoxins specific for neuronal nAChRs. The C-terminus of each sequence is amidated; Tyr 15 of EpI is sulfated. Subtype specificities are given for mammalian (rat) nAChRs. References are given for the report of the sequence, followed by the report of the specificity, when separate.

or neuronal nAChRs, each composed of five membrane-spanning subunits. Muscle-type nAChRs contain two α1, one *β*1, one *δ*, and one *γ* or *ϵ* subunit (*8, 9*), whereas neuronal nAChRs contain combinations of α (α2−α9) and *β* (*β*2−*β*4) subunits (*10, 11*). Three-dimensional structures have been determined previously for six α-conotoxins: the muscle-specific GI (*12*−*16*) and MI (*17*) and the neuronal-specific PnIA (*18*), PnIB (*19*), MII (*20, 21*), and EpI (*22*). These structures demonstrate the general stability and compactness of the α-conotoxin fold and emphasize the variety of activities achievable by amino acid sequence variation on a conserved scaffold. In particular, we are interested in those α-conotoxins that target neuronal nAChRs (Figure 1b). Subunit composition in the neuronal nAChRs is more variable than in the muscle-type, resulting in physiologically distinct receptors within the neuronal subtype which the α-conotoxins are able to bind and inhibit selectively. As neuronal nAChRs are significant in synaptic transmission and neurotransmitter release (*11*) and are affected in a number of neuropsychiatric and neurodegenerative disorders (*23*), the structural basis for α-conotoxin binding is of direct biological and pharmacological consequence.

We present here the solution structure of α-conotoxin ImI, expanding the small group of conotoxins with well-characterized structure and function. ImI is the first conotoxin isolated from a vermivorous cone and exhibits novel biochemical and pharmacological properties (*24*). Consisting of only 12 amino acids, incorporating two disulfide linkages, ImI is the smallest α-conotoxin yet identified. The peptide acts as a competitive antagonist of α-bungarotoxin-sensitive, α7-containing neuronal nAChRs in mammals, with higher selectivity than α-bungarotoxin, which also binds muscle-type receptors (*25, 26*). It is active on human neuronal α-bungarotoxin-sensitive nAChRs, as evidenced by studies in human neuroendocrine carcinoma cells (*27*), thus displaying wide phylogenetic range. In addition, ImI has been shown to discriminate between kinetically distinct receptors, selectively blocking subtypes which are rapidly desensitized (*28*). A detailed comparison of the structure of ImI to those of other conotoxins specific for neuronal nAChRs, with attention to primary sequence comparisons and mutagenesis data, identifies several features of the structures that contribute to α-conotoxin binding and specificity.

EXPERIMENTAL PROCEDURES

Peptide Synthesis and Sample Preparation. α-Conotoxin ImI was synthesized using standard Fmoc chemistry on an Advanced Chemtech 357 FBS peptide synthesizer and purified by reversed-phase HPLC (procedure modified from ref *29*). The two disulfides were formed sequentially, 2−8 by potassium ferricyanide oxidation of the thiols and 3−12 by iodine oxidation of the bis *S*-acetamidomethyl derivatives. Peptide purity and oxidation state were assessed by electrospray-ionization mass spectrometry on a Hewlett-Packard 5989A quadrupole instrument equipped with an Analytica ion source; peptide identity was also confirmed by HPLC coelution with the native toxin.

A 2 mM NMR sample (pH 3.0) was prepared in either 100% D_2O or a mixed solvent of 90% H_2O/10% D_2O, with the concentration based on a molar absorption coefficient at 280 nm of 5750 M^{-1} cm^{-1} (*30*). For D_2O experiments, the peptide was lyophilized and redissolved in 100% D_2O.

NMR Spectroscopy. ^1H NMR spectra were obtained at 298 K on Bruker DRX 300 and DRX 500 spectrometers in the pure phase absorption mode using the States−TPPI method of quadrature detection (*31*). For ^1H resonance assignment a DQF-COSY spectrum (*32*) was recorded in D_2O, and TOCSY (*33*) and ROESY (*34, 35*) spectra were recorded in 90% H_2O/10% D_2O. After zero filling, the DQF-COSY data size was 4096 × 2048 points, resulting in digital resolutions of 0.73 and 1.46 Hz/point in ω_1 and ω_2, respectively. The TOCSY experiment incorporated a DIPSI-2 (*36*) sequence with a 10 kHz rf field and a mixing time of 70 ms. The ROESY experiment was taken with the carrier frequency positioned at 6 ppm, using a 5 kHz CW spin-lock field during

Table 1: Quantitative Description of the 20 DYANA Conformers Used To Represent the Solution Structure before and after Energy Refinement with the Program OPAL[a]

parameter	DYANA	OPAL
DYANA target function (Å^2)[b]	0.15 ± 0.05 (0.09...0.24)	
AMBER energy (kcal/mol)	−290.05 ± 18.50 (−326.29...−250.06)	−455.06 ± 15.48 (−485.52...−426.97)
van der Waals energy	84.35 ± 15.32 (61.66...118.24)	−13.01 ± 3.13 (−18.64...−7.34)
electrostatic energy	−475.09 ± 15.32 (−507.62...−449.20)	−519.37 ± 16.30 (−547.98...−491.42)
residual NOE distance constraint violations (Å)		
no. > 0.1 Å	3.30 ± 0.46 (3.00...4.00)	0.00 ± 0.00 (0.00...0.00)
sum	0.90 ± 0.10 (0.72...1.14)	1.00 ± 0.09 (0.83...1.19)
maximum	0.18 ± 0.02 (0.15...0.21)	0.09 ± 0.00 (0.08...0.10)
residual dihedral angle constraint violations (deg)		
no. > 2.5°	0.05 ± 0.22 (0.00...1.00)	0.00 ± 0.00 (0.00...0.00)
sum	0.21 ± 0.67 (0.00...3.11)	5.72 ± 1.68 (3.08...8.82)
maximum	0.21 ± 0.65 (0.00...3.00)	1.73 ± 0.24 (1.15...2.09)
rmsd values (Å)[c]		
backbone (1−12)	0.56 ± 0.24 (0.26...1.13)	0.63 ± 0.21 (0.35...1.07)
all heavy atoms	1.38 ± 0.36 (0.86...2.14)	1.45 ± 0.34 (1.03...2.28)
backbone (2−11)	0.31 ± 0.14 (0.15...0.58)	0.44 ± 0.14 (0.25...0.69)
backbone (2−11) + best-defined side chains	0.43 ± 0.09 (0.27...0.59)	0.56 ± 0.10 (0.40...0.75)

[a] The numbers given are the average ± the standard deviation for the group, with individual minimum and maximum values given in parentheses. [b] The target function is not given for the energy-minimized conformers, as they do not have the ECEPP/2 standard covalent geometry (*72, 73*). [c] Rmsd values presented are relative to the mean coordinates. The best-defined side chains—taken as those with global displacement of the side chain heavy atoms less than 1.5 Å—are C2, C3, S4, D5, P6, C8, and A9, upon superimposition of the backbone atoms for residues 2−11.

the 100 ms mixing period. Data sizes for these experiments after zero filling were 2048 × 1024 points, resulting in digital resolutions of 3.18 Hz/point in ω_1 and 6.36 Hz/point in ω_2.

Processing of all the spectra was performed with the program PROSA (*37*). The residual water signal was reduced using the convolution method of Marion et al. (*38*), and baseline distortions were corrected using the IFLAT procedure (*39*). Prior to Fourier transformation, the time domain data were apodized using a sine bell window function shifted by $\pi/2$ (*40*). Peak picking, spin system identification, and volume integration of the ROESY cross-peaks were performed with the interactive program XEASY (*41*). $^3J_{HN\alpha}$ scalar coupling constants were measured from a 1D ^{1}H spectrum, and $^3J_{\alpha\beta}$ coupling constants for residues containing $C^\beta H_2$ groups were measured from the DQF-COSY spectrum.

Resonance Assignment. An initial ^{1}H resonance assignment was determined automatically using the program GARANT (*42, 43*). The resonance assignment was then corrected and completed using spin system identification from the COSY and TOCSY spectra in conjunction with sequential NOE connectivities identified in the ROESY spectrum (*44*).

Structure Calculation and Characterization. Upper distance constraints were calculated from ROESY cross-peak volumes using the CALIBA procedure (*45*), and stereospecific assignments and dihedral angle constraints were determined from coupling constants and upper distance constraints using the HABAS procedure (*46*). Disulfide bond constraints were added as established previously (*47*). The structure calculations were performed using simulated annealing in torsion angle space as implemented in the program DYANA (*48*). Additional stereospecific assignments based on the preliminary structure were made using the GLOMSA procedure (*45*). Restrained energy minimization was performed using the AMBER all-atom force field (*49*) with the program OPAL (*50*), which includes pseudoenergy terms for distance and dihedral angle constraints (*51, 52*). Distance constraints for the chemically determined disulfide bridges, and the ψ and χ^2 dihedral angle constraints, for which no direct spin−spin coupling information was available, were removed from the set of constraints used during energy refinement. The energy minimization was executed after immersion of each DYANA conformer in a 6 Å thick shell of explicit water molecules, using a constant dielectric permittivity for the electrostatic interactions. For each conformer a maximum of 2500 steps of conjugate gradient minimization was performed, after which a small total gradient was attained.

The program MOLMOL (*53*) was used to analyze the structure in terms of rmsd values (*54*), atom displacements (*55, 56*), hydrogen bonds (*57*), regular secondary structures (*58*), solvent-accessible surface areas (*59, 60*), and local electrostatic potential (*61*). The criteria for identification of a hydrogen bond were a maximal proton−acceptor distance of 2.4 Å and a maximal angle of 35° between the donor−proton bond and the line connecting the donor and acceptor heavy atoms in at least half of the conformers. Regular secondary structure elements were identified on the basis of hydrogen bond patterns, according to DSSP standards. The solvent-accessible surface areas (SASAs) were calculated with a solvent radius of 1.4 Å. MOLMOL was also used for display of the contact surface (*62, 63*) and preparation of the color figures.

Structure Comparison. The solution structure of ImI was compared to the crystal structures of PnIA and PnIB (PDB accession codes 1PEN and 1AKG, respectively) and to the solution structure of MII (PDB accession code 1M2C; made available to us prior to release by K.-J. Shon). Since the coordinates for the EpI structure have not yet been released by the Protein Data Bank, our comparison to EpI was limited to structural features described by Hu et al. (*22*). For consistency, hydrogen bonds and regular secondary structure elements for all compared structures were identified using the stringent criteria described above for ImI.

RESULTS

Resonance Assignment and Structure Calculation. A complete chemical shift list is given in the Supporting Information. A total of 255 ROESY cross-peaks were

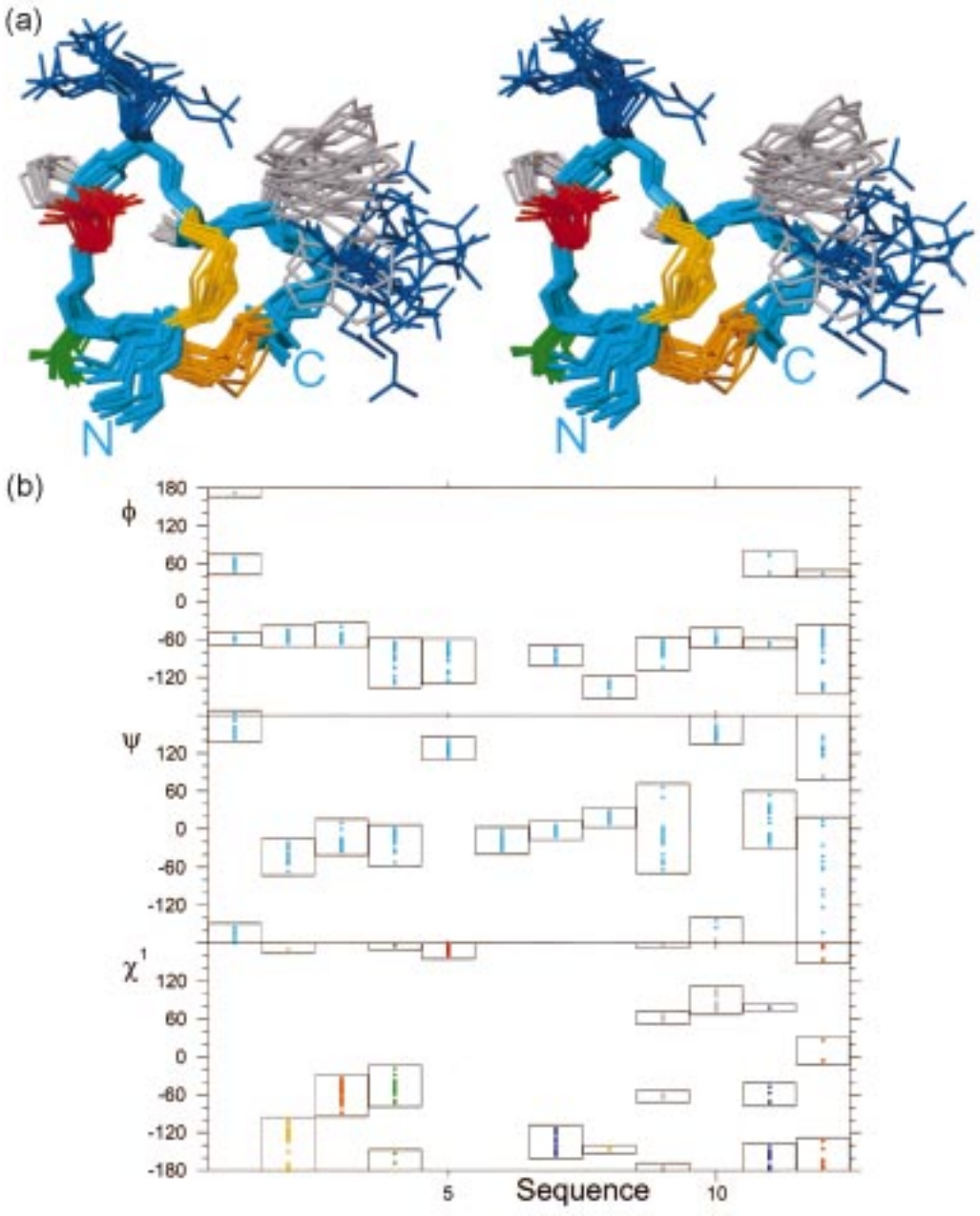


FIGURE 2: (a) Stereoview of the 20 energy-refined conformers representing the ImI solution structure. The backbone is in cyan, with hydrophobic side chains shown in gray, polar side chains in green, positively charged side chains in blue, and negatively charged side chains in red. The 2−8 disulfide bridge is shown in yellow and the 3−12 in orange. (b) Plot of the backbone ϕ, ψ, and side chain χ^1 dihedral angles versus the amino acid sequence. The dihedral angle observed in each individual conformer is represented by a dot, and the ranges are shown by boxes. Coloring of the backbone and side chains is as in (a).

assigned and integrated, with concomitant cycles of structure calculation for evaluation of distance and angle constraint violations as well as assignment of additional peaks based on the preliminary structure. Calibration of these cross-peaks, followed by an initial screening in DYANA (discarding constraints which were duplicate, nonconstraining, or constraining a fixed distance), resulted in a total of 80 meaningful upper distance constraints. Six $^3J_{HN\alpha}$ and eight $^3J_{\alpha\beta}$ coupling constants were used together with intraresidual and sequential upper distance constraints to generate 46 constraints for the ϕ, ψ, χ^1, and χ^2 dihedral angles. A total of 40 conformers were calculated, from which the 20 with the lowest residual target function were selected for restrained energy minimization. The resulting conformers contained no significant violations of any constraint and were used to represent the solution structure of ImI. A quantitative characterization of the structure is given in Table 1.

Characterization of the Structure. The solution structure of α-conotoxin ImI is shown in Figure 2a and is characterized by two compact loops of four and three residues, defined by the two disulfide bonds. Despite the involvement of Cys 12 in a disulfide bond, the backbone conformation is least well defined at the chain termini. Interestingly, the first loop is better defined and more structured than the second loop. The first loop contains two regions with 3_{10}-helical character, supported by $O_i'-HN_{i+3}$ hydrogen bonds observed for Gly 1−Ser 4, Cys 2−Asp 5, and Asp 5−Cys 8. In contrast, no regular secondary structure elements are found in the second loop, which forms a β-turn with a hydrogen bond between Cys 12 HN and Ala 9 O′. The Trp 10 and Arg 11 side chains

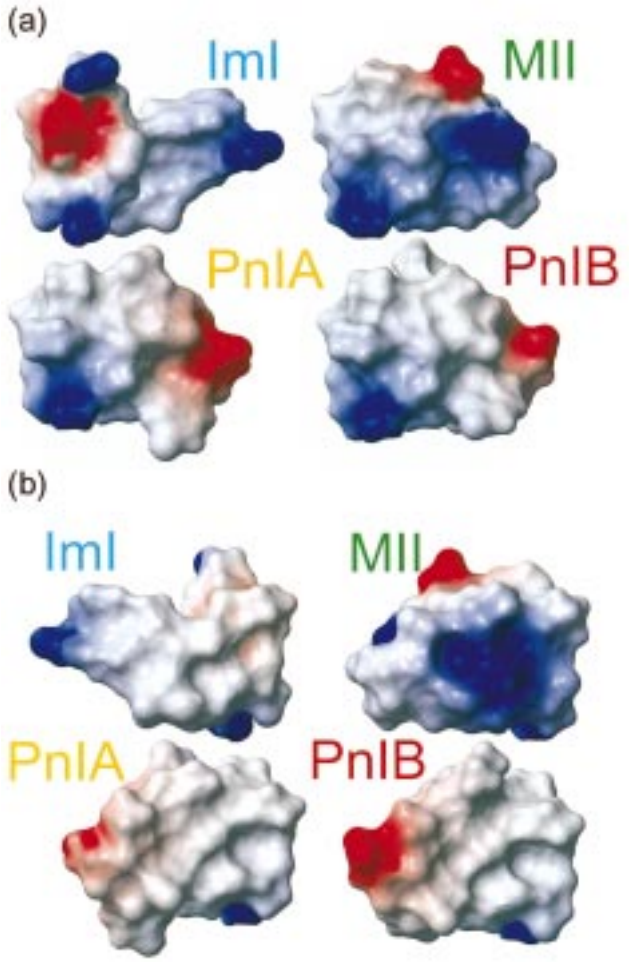


FIGURE 3: Solvent-accessible surfaces of ImI, MII, PnIA, and PnIB colored according to local electrostatic potential (a) with orientation as in Figure 2a and (b) rotated 180° about the vertical axis. Red indicates a negative charge and blue a positive charge. Hydrogen atoms were included in the calculation of both the surface and the potential. For ImI and MII, the surface is shown for the conformer with the smallest rmsd to the mean structure. The MII structure was solved at low pH; thus the surface shows positive charges for the two histidines although this may not be representative of the protonation state at physiological pH.

in the second loop display the most conformational flexibility, whereas side chains for residues in the first loop are well defined. χ^1 dihedral angles indicate a gauche$^-$ conformation for Cys 3, a trans or gauche$^-$ conformation for Ser 4, and a trans conformation for Asp 5 (Figure 2b). The side chain of Asp 5 is further constrained by the hydrogen bond between Asp 5 δO and Arg 7 HN. This hydrogen bond and the extended backbone conformation suggest that Asp 5 functions as an N-cap to further stabilize the 3_{10}-helical turn from Pro 6 to Cys 8 (*64*−*66*). This function is supported by the observation that Asp 5 is the only amino acid residue in ImI for which H$^\alpha$ displays a positive secondary structure induced shift (*67*) and is consistent with a clear preference of 2.1:1 for Asp at the N-cap position, and 2.6:1 for Pro at the (N1) helix-initiator position (*65*).

Figure 3 shows the contact surface of the ImI conformer with the smallest rmsd to the mean structure, colored according to the local electrostatic potential. The two disulfide bridges and Ala 9 compose the core of the molecule, providing a rigid scaffold from which the other side chains are oriented toward the solvent. The two loops form two subdomains connected by Cys 8 and Ala 9, which make contacts to residues in both loops. The two subdomains are separated by a deep cleft, where the 2−8 disulfide bridge forms the left-hand wall, the 3−12 disulfide bridge forms the right-hand wall, and the Trp 10 side chain forms the top. The first loop carries most of the charged entities, which are the positively charged N-terminus and the side chain of Arg 7 as well as the sole negatively charged side chain of Asp 5. One δO of Asp 5 is effectively buried upon formation of the aforementioned side chain−backbone hydrogen bond. The only charge found in the second loop is the positively charged side chain of Arg 11. The front side of ImI (in the orientation of Figures 2a and 3a) is highly charged, whereas the side chains of Cys 3, Pro 6, Ala 9, and Cys 12 form a large hydrophobic region on the back side (Figure 3b).

Structure Comparison. Four known neuronal α-conotoxin structures are shown in Figure 4, where the backbone atoms N, C$^\alpha$, and C′ of residues 1−9, 11, and 12 in ImI and of residues 1−9, 14, and 16 in PnIA/B and MII have been superimposed for minimal rmsd. The amino acid sequence of the first loop is 75% identical in all four conotoxins and displays a very similar profile of relative solvent-accessible surface areas (Figure 5). The only significant difference in this profile is at position 7, where the charged Arg side chain in ImI has a large SASA and the hydrophobic side chains in PnIA/B and MII are buried from the solvent. In PnIA and PnIB, the hydrophobic Leu 5 replaces the hydrophilic Asp 5 and Asn 5 side chains found in ImI and MII and thus cannot form the N-cap. Superposition of the residues in the second loop was based on similarities in the SASA between Ala 9, Arg 11, and Cys 12 in ImI, Ala 9, Asp 14, and Cys 16 in PnIA/B, and His 9, Asn 14, and Cys 16 in MII. Additionally, the local backbone conformations of the toxins at position 11 in ImI and position 14 in PnIA/B and MII are very similar (Figure 4a); thus the longer side chain of Arg 11 is in close proximity to Asp 14 and Asn 14 and directly compensates for the smaller second loop size of ImI (Figure 4b). Although the Trp 10 side chain in ImI is closest to Asn 12 in PnIA/B and His 12 in MII, it has no clear counterpart in those structures. These considerations result in a structure-based revision of the sequence alignment of ImI with PnIA, PnIB, and MII (Figures 1b and 6).

Figure 6 provides a survey of hydrogen bonds, secondary structure elements, and residues with low solvent accessibility in the ImI, PnIA, PnIB, and MII structures. All four structures show two helical regions in the first loop. The first of these regions is a 3_{10}-helical turn from Cys 2 to Ser 4, supported by two $O_i'-HN_{i+3}$ hydrogen bonds. MII has only one of these hydrogen bonds, and thus the 3_{10}-helix is not identified, although it is described in the structure by Shon et al. (*20*). Whereas ImI shows one more short 3_{10}-helix from Pro 6 to Cys 8 (described above), PnIA, PnIB, and MII are able to form longer α-helices, which extend into the second loop. Interestingly, Pro 6 is the only highly conserved amino acid residue in the α-conotoxins apart from the cystines and is responsible for helix initiation. In PnIA and PnIB, the α-helix is very regular, supported by two $O_i'-HN_{i+4}$ hydrogen bonds and ending with one residue in local 3_{10}-helical conformation, as indicated by an $O_i'-HN_{i+3}$ hydrogen bond. For PnIA and PnIB a side chain−backbone hydrogen bond is also observed for Asn 12−Cys 8, suggesting a C-cap (*65, 66*). The two α-helical hydrogen bonds present in PnIA and PnIB are

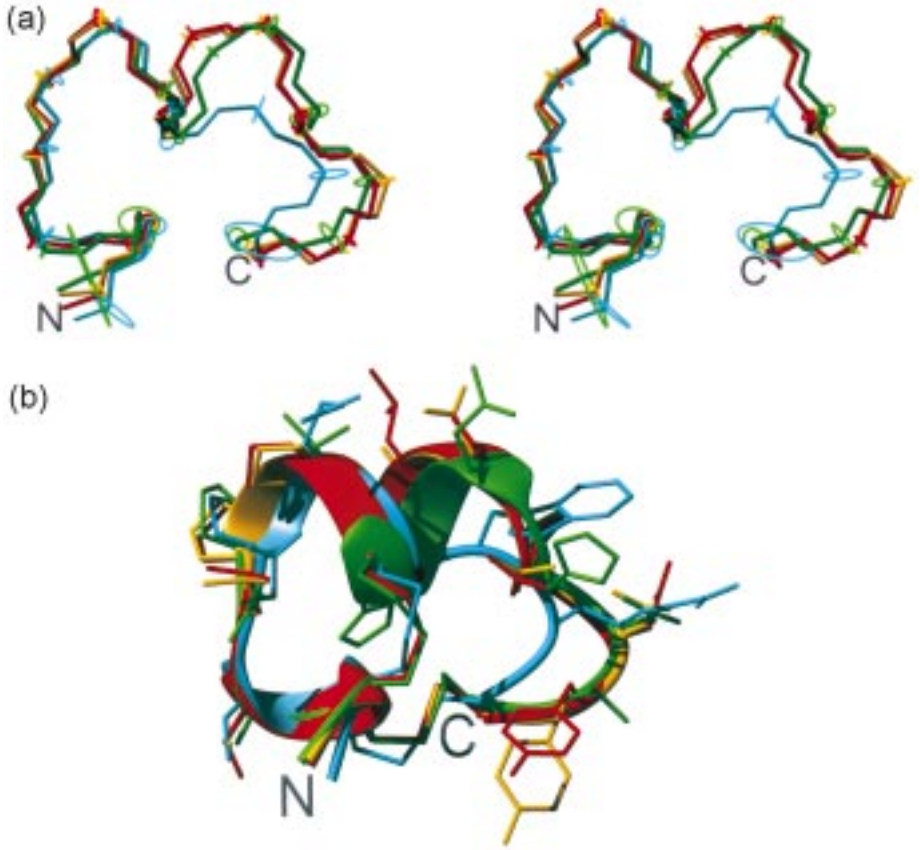


FIGURE 4: (a) Stereoview of the polypeptide backbone of ImI, PnIA, PnIB, and MII in cyan, yellow, red, and green, respectively. The backbone atoms N, C^{α}, and C′ of residues 1−9, 11, and 12 in ImI were superimposed with those of residues 1−9, 14, and 16 in PnIA, PnIB, and MII. Circles at the C^{α} position of each residue indicate the global displacement *D* for the NMR structures (ImI and MII) and the crystal structures (PnIA and PnIB), where the relation between the *B*-factor of the crystal structures and *D* is given by $B = 8\pi^2 D^2$ (*55*, *56*). (b) Ribbon representation of the polypeptide backbone, illustrating secondary structure features and relative side chain positions. For ImI and MII, the structure is shown for the conformer with the smallest rmsd to the mean structure. Coloring is as in (a).

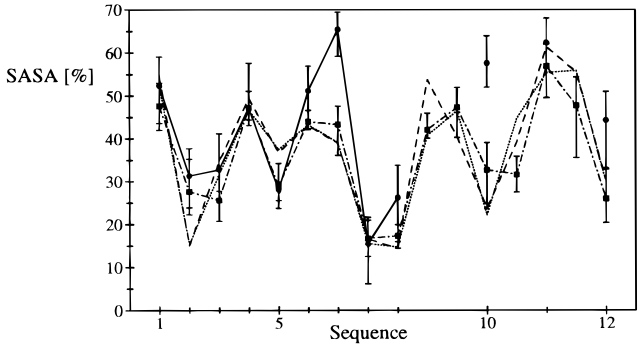


FIGURE 5: Plot of the average solvent-accessible surface area relative to the total surface area for individual residues of ImI, PnIA, PnIB, and MII. Numbering of the amino acid sequence is for ImI; sequence alignment is as described in the text. Solid lines indicate the relative SASA for ImI, dotted lines for PnIA, dashed lines for PnIB, and dotted-dashed lines for MII. Bars indicating the range of values seen among NMR conformers are shown with circles for ImI and squares for MII.

absent in MII, which shows an alternative hydrogen bond pattern, resulting in a distorted α-helix. Due to this pattern the α-helix in MII is identified only from Cys 8 to Glu 11, although local backbone conformation and formation of an N-cap by Asn 5 [as described by Shon et al. (*20*)] suggest that the α-helix starts at Pro 6, as in the other structures. Following the α-helix, two consecutive β-turns are observed in PnIA, PnIB, and MII, with corresponding hydrogen bonds. In the shorter second loop of ImI, only one β-turn can be formed. The 3_{10}-helical turn from Cys 2 to Ser 4, the α-helix from Pro 6 to Asn 11, and the two consecutive β-turns have also been described for the structure of EpI (*22*). The sequence requirements for N- and C-caps terminating the α-helix are present in EpI as well (Figure 1b). The disulfide bridges and Ala 9 (ImI, PnIA/B) and His 9 (MII) are buried from the solvent in all four structures, indicating that the molecules share a common hydrophobic core. The only additional buried residues are Asp 5 in ImI, Asn 5 in MII, and Asn 12 in PnIA/B, which participate in side chain−backbone hydrogen bonds and function as N- and C-caps, respectively.

Figure 3 compares the contact surface and charge distribution of ImI with PnIA, PnIB, and MII. The deep cleft observed in ImI is filled by the side chains of Asn 12 and Tyr 15 in PnIA/B and His 12 and Leu 15 in MII. In PnIA and PnIB, a positive charge and a negative charge are located at the bottom of the molecule, in the orientation shown. The positively charged N-terminus is common to all four structures, whereas the negatively charged Asp 14 side chain in PnIA and PnIB corresponds to the positively charged Arg 11 in ImI. The top surface of PnIA and PnIB is uncharged, where additional charges are found in ImI and MII. As described previously, ImI has the negatively and positively charged side chains of Asp 5 and Arg 7, respectively, in the first loop. MII, like PnIA and PnIB, has its sole negatively

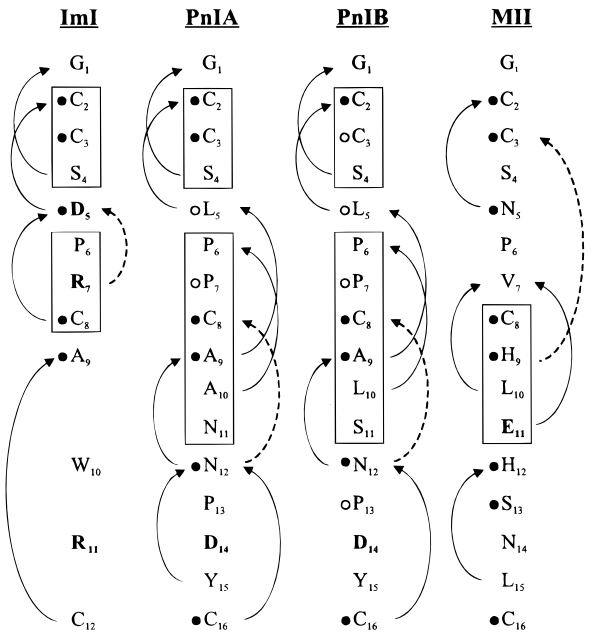


FIGURE 6: Schematic comparison of structural features found in α-conotoxins ImI, PnIA, PnIB, and MII. The sequences are presented from top to bottom; residues with charged side chains are in bold print, and helices are indicated by boxes. Backbone−backbone hydrogen bonds are represented by arrows pointing from the donor residue to the acceptor residue. The 3_{10}-helical $O_i'-HN_{i+3}$ hydrogen bonds are shown on the left of the individual sequences; all others are shown on the right. Hydrogen bonds between backbone and side chain atoms are identified by arrows with dashed lines. Filled circles indicate residues with low relative solvent accessibility (<33%), and open circles indicate residues with intermediate solvent accessibility (33−40%). Sequence alignment is as described in the text.

charged side chain in the second loop; since Glu 11 of MII occurs earlier in the sequence than Asp 14 of PnIA and PnIB, the negative charge is located on the top rather than the bottom surface. The structure of MII was solved at low pH and therefore shows two additional positively charged histidine side chains; however, the protonation state of the imidazole rings at physiological pH is unknown. With the possible exception of MII, the back sides of all four toxins are largely hydrophobic (Figure 3b). The amino acid sequence of EpI is identical in the first loop to that of ImI and is very similar in the second loop to that of PnIA and PnIB. The fold and shape of EpI, as well as the electrostatic charge in the second loop, are thus very similar to those of PnIA and PnIB (*22*), while the conformation and surface potential in the first loop are like that of ImI.

DISCUSSION

α-Conotoxin ImI, the smallest known α-conotoxin, is a potent antagonist of the α7 subtype of neuronal nAChRs. As such, it is a powerful biological tool which may be used to probe the complex architecture of the ligand-gated ion channel and possibly to control the functional contributions of certain nAChR subtypes. ImI is also an excellent example of a minimal functional protein and may stand as a paradigm for small subtype-specific ligands. Elucidation of the structural features contributing to ImI binding to the nAChR will give insight into the requirements for a minimal structure and may be useful for structure-based design. The three-dimensional structure of ImI and comparison to the structures of related α-conotoxins form a basis for evaluating structure−activity relationships in these molecules and, subsequently, for understanding their selectivity and modes of binding.

The five neuronal α-conotoxins—ImI, EpI, PnIA, PnIB, and MII—share several common structural features. The local backbone conformations and the SASAs for individual residues are similar for all five toxins, as is the rigid hydrophobic core. Pro 6 is highly conserved in all the known neuronal α-conotoxins and occupies the helix-initiator position, stabilizing the helix and promoting regular secondary structure. Furthermore, Asp 5 in ImI and EpI and Asn 5 in MII function as helix-stabilizing N-caps. Additional features must therefore confer specificity in these molecules. Within the conserved disulfide framework characteristic of the α-conotoxins the loop length is a principal factor determining the overall fold of the molecule. Loop sizes of three residues in the first loop and five in the second (3,5-type) are typically found in the known muscle-specific α-conotoxins, whereas loop sizes of four residues in the first loop and seven in the second (4,7-type) are typically found in the known neuronal specific α-conotoxins. Exceptions to this rule exist, however, as the muscle-specific EI has 4,7-type loops and the neuronal-specific ImI has 4,3-type loops. Overall electrostatic charge has also been implicated as a factor determining specificity, with the muscle-specific α-conotoxins generally exhibiting positive net charges and the neuronal-specific α-conotoxins exhibiting neutral or negative net charges (*22*). Again, ImI is an exception to this classification, with a clear +2 net charge.

For ImI in particular, Asp 5, Pro 6, Arg 7, and Trp 10 have been shown by analysis of conservative mutations to be important for binding affinity to α7 receptors (*68*). The three-dimensional structure of ImI indicates that Asp 5 and Pro 6 function as N-cap and helix-initiator, respectively; thus their contribution to binding may be due to their distinct structural roles rather than to direct interactions with the receptor as previously proposed (*68*). In the case of Asp 5, the negative surface charge may also be of consequence. The significance of the N-cap and Pro 6 suggests that the formation of the second helical region is important for binding. As a result of helix formation, Arg 7 and Trp 10 in ImI and the residues at positions 10 and 11 in PnIA, PnIB, MII, and EpI are solvent exposed, whereas Cys 8 and the residues at position 9 in all five molecules are buried, contributing to the core (Figure 6). These solvent-accessible amino acid side chains may be directly involved in interactions with the receptor. Indeed, the side chains of Arg 7 and Trp 10 significantly affect ImI binding affinity (*68*). Moreover, residues 10 and 11 represent the only sequence differences between PnIA and PnIB, which target different nAChR subtypes, indicating that this region may be critical for binding specificity as well as affinity. This impression is supported by the observation that a single amino acid mutation of Ala 10 in PnIA to the Leu 10 found in PnIB shifts PnIA specificity to that of PnIB (*69*). Although Trp 10 of ImI does not correspond exactly to residues 10 and 11 in the other four toxins, it is located on approximately the same face of the molecule (Figure 3a) and may contribute similarly to specificity.

The comparison of the structure of ImI to those of EpI, MII, PnIA, and PnIB highlights several features. First, it does not appear that the charged residues themselves determine

binding specificity in the neuronal α-conotoxins, nor does it appear that a different loop size alone precludes a certain activity or specificity. Formation of the second helical region, however, does seem to be critical for binding affinity and specificity. Additionally, the deep cleft in the second subdomain of ImI, filled by either Tyr or Leu residues in the other four neuronal α-conotoxins, may be important. Nonetheless, the specificity of the α-conotoxins for various neuronal nAChR subtypes cannot be explained conclusively by their particular structural features. The relatively high sequence and conformational diversity found within a class of conotoxins such as the α-conotoxins likely derives from the high diversity found in the genus, which has undergone very rapid recent speciation (*70*), and is influenced by the multiple biotic interactions of each *Conus* species (*7*). It may well be that each individual α-conotoxin, while binding to the same macrosite on the receptor and eliciting the same pharmacological effect, interacts with a unique subset of microsites within that region (*71*). We have described specific conserved and variable structural features of the neuronal α-conotoxins which are important for activity. Further establishment of structure−activity relationships for this family of conotoxins will require a combination of this knowledge with more extensive structural characterization of the target sites on the neuronal nAChRs themselves.

ACKNOWLEDGMENT

We thank D. King of the Howard Hughes Medical Institute for analysis of ImI by mass spectrometry and K.-J. Shon of Case Western Reserve University for providing the coordinates of the MII structure.

SUPPORTING INFORMATION AVAILABLE

^{1}H chemical shifts, TOCSY and ROESY spectra, survey of sequential and medium-range NOEs, and plot of NOE constraints versus the sequence for α-conotoxin ImI. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Olivera, B. M., Gray, W. R., Zeikus, R., McIntosh, J. M., Varga, J., Rivier, J., de Santos, V., and Cruz, L. J. (1985) *Science 230*, 1338−1343.
2. Olivera, B. M., Rivier, J., Clark, C., Ramilo, C. A., Corpuz, G. P., Abogadie, F. C., Mena, E. E., Woodward, S. R., Hillyard, D. R., and Cruz, L. J. (1990) *Science 249*, 257−263.
3. Woodward, S. R., Cruz, L. J., Olivera, B. M., and Hillyard, D. R. (1990) *EMBO J. 9*, 1015−1020.
4. Colledge, C. J., Hunsperger, J. P., Imperial, J. S., and Hillyard, D. R. (1992) *Toxicon 30*, 1111−1116.
5. Olivera, B. M. (1998) *Mol. Biol. Cell 8*, 2102−2109.
6. McIntosh, J. M., Olivera, B. M., and Cruz, L. J. (1998) *Methods Enzymol. 294*, 605−624.
7. McIntosh, J. M., Santos, A. D., and Olivera, B. M. (1999) *Annu. Rev. Biochem.* (in press).
8. Galzi, J.-L., Revah, F., Bessis, A., and Changeux, J.-P. (1991) *Annu. Rev. Pharmacol. 31*, 37−72.
9. Arias, H. R. (1997) *Brain Res. Rev. 25*, 133−191.
10. Sargent, P. B. (1993) *Annu. Rev. Neurosci. 16*, 403−443.
11. McGehee, D. S., and Role, L. W. (1995) *Annu. Rev. Physiol. 57*, 521−546.
12. Kobayashi, Y., Ohkubo, T., Kyogoku, Y., Nishiuchi, Y., Sakakibara, S., Braun, W., and Gō, N. (1989) *Biochemistry 28*, 4853−4860.
13. Pardi, A., Galdes, A., Florance, J., and Maniconte, D. (1989) *Biochemistry 28*, 5494−5501.
14. Guddat, L. W., Martin, J. A., Shan, L., Edmundson, A. B., and Gray, W. R. (1996) *Biochemistry 35*, 11329−11335.
15. Gehrmann, J., Alewood, P. F., and Craik, D. J. (1998) *J. Mol. Biol. 278*, 401−415.
16. Maslennikov, I. V., Sobol, A. G., Gladky, K. V., Lugovskoy, A. A., Ostrovsky, A. G., Tsetlin, V. I., Ivanov, V. T., and Arseniev, A. S. (1998) *Eur. J. Biochem. 254*, 238−247.
17. Gouda, H., Yamazaki, K., Hasegawa, J., Kobayashi, Y., Nishiuchi, Y., Sakakibara, S., and Hirono, S. (1997) *Biochim. Biophys. Acta 1343*, 327−334.
18. Hu, S.-H., Gehrmann, J., Guddat, L. W., Alewood, P. F., Craik, D. J., and Martin, J. L. (1996) *Structure 4*, 417−423.
19. Hu, S.-H., Gehrmann, J., Alewood, P. F., Craik, D. J., and Martin, J. L. (1997) *Biochemistry 36*, 11323−11330.
20. Shon, K.-J., Koerber, S. C., Rivier, J. E., Olivera, B. M., and McIntosh, J. M. (1997) *Biochemistry 36*, 15693−15700.
21. Hill, J. M., Oomen, C. J., Miranda, L. P., Bingham, J.-P., Alewood, P. F., and Craik, D. J. (1998) *Biochemistry 37*, 15621−15630.
22. Hu, S.-H., Loughnan, M., Miller, R., Weeks, C. M., Blessing, R. H., Alewood, P. F., Lewis, R. J., and Martin, J. L. (1998) *Biochemistry 37*, 11425−11433.
23. Gotti, C., Fornasari, D., and Clementi, F. (1997) *Prog. Neurobiol. 53*, 199−237.
24. McIntosh, J. M., Yoshikami, D., Mahe, E., Nielsen, D. B., Rivier, J. E., Gray, W. R., and Olivera, B. M. (1994) *J. Biol. Chem. 269*, 16733−16739.
25. Johnson, D. S., Martinez, J., Elgoyhen, A. B., Heinemann, S. F., and McIntosh, J. M. (1995) *Mol. Pharmacol. 48*, 194−199.
26. Pereira, E. F. R., Alkondon, M., McIntosh, J. M., and Albuquerque, E. X. (1996) *J. Pharmacol. Exp. Ther. 278*, 1472−1483.
27. Codignola, A., McIntosh, J. M., Cattaneo, M. G., Vicentini, L. M., Clementi, F., and Sher, E. (1996) *Neurosci. Lett. 206*, 53−56.
28. Kehoe, J., and McIntosh, J. M. (1998) *J. Neurosci. 18*, 8198−8213.
29. Cartier, G. E., Yoshikami, D., Gray, W. R., Luo, S., Olivera, B. M., and McIntosh, J. M. (1996) *J. Biol. Chem. 271*, 7522−7528.
30. Pace, C. N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. (1995) *Protein Sci. 4*, 2411−2423.
31. Marion, D., Ikura, K., Tschudin, R., and Bax, A. (1989) *J. Magn. Reson. 85*, 393−399.
32. Rance, M., Sørensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., and Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 117*, 479−485.
33. Shaka, A. J., Lee, L. J., and Pines, A. (1988) *J. Magn. Reson. 77*, 274−293.
34. Bothner-By, A. A., Stephens, R. L., Lee, J., Warren, C. D., and Jeanloz, R. W. (1984) *J. Am. Chem. Soc. 106*, 811−813.
35. Bax, A., and Davis, D. G. (1985) *J. Magn. Reson. 63*, 207−213.
36. Rucker, S. P., and Shaka, A. J. (1989) *Mol. Phys. 68*, 509−517.
37. Güntert, P., Dötsch, V., Wider, G., and Wüthrich, K. (1992) *J. Biomol. NMR 2*, 619−629.
38. Marion, D., Ikura, K., and Bax, A. (1989) *J. Magn. Reson. 84*, 425−430.
39. Bartels, C., Güntert, P., and Wüthrich, K. (1995) *J. Magn. Reson. A 117*, 330−333.
40. DeMarco, A., and Wüthrich, K. (1976) *J. Magn. Reson. 24*, 201−204.
41. Bartels, C., Xia, T., Billeter, M., Güntert, P., and Wüthrich, K. (1995) *J. Biomol. NMR 5*, 1−10.
42. Bartels, C., Billeter, M., Güntert, P., and Wüthrich, K. (1996) *J. Biomol. NMR 7*, 207−213.
43. Bartels, C., Güntert, P., Billeter, M., and Wüthrich, K. (1997) *J. Comput. Chem. 18*, 139−149.
44. Wüthrich, K. (1986) *NMR of proteins and nucleic acids*, John Wiley & Sons, New York.

45. Güntert, P., Braun, W., and Wüthrich, K. (1991) *J. Mol. Biol. 217*, 517−530.
46. Güntert, P., Braun, W., Billeter, M., and Wüthrich, K. (1989) *J. Am. Chem. Soc. 111*, 3997−4004.
47. Williamson, M. P., Havel, T. F., and Wüthrich, K. (1985) *J. Mol. Biol. 182*, 295−315.
48. Güntert, P., Mumenthaler, C., and Wüthrich, K. (1997) *J. Mol. Biol. 273*, 283−298.
49. Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Jr., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A. (1995) *J. Am. Chem. Soc. 117*, 5179−5197.
50. Luginbühl, P., Güntert, P., Billeter, M., and Wüthrich, K. (1996) *J. Biomol. NMR 8*, 136−146.
51. Widmer, H., Billeter, M., and Wüthrich, K. (1989) *Proteins: Struct., Funct., Genet. 6*, 357−371.
52. Billeter, M., Schaumann, T., Braun, W., and Wüthrich, K. (1990) *Biopolymers 29*, 695−706.
53. Koradi, R., Billeter, M., and Wüthrich, K. (1996) *J. Mol. Graphics 14*, 51−55.
54. McLachlan, A. D. (1979) *J. Mol. Biol. 128*, 49−79.
55. Billeter, M., Kline, A. D., Braun, W., Huber, R., and Wüthrich, K. (1989) *J. Mol. Biol. 206*, 677−687.
56. Billeter, M. (1992) *Q. Rev. Biophys. 25*, 325−377.
57. Levitt, M. (1983) *J. Mol. Biol. 170*, 723−764.
58. Kabsch, W., and Sander, C. (1983) *Biopolymers 22*, 2577−2637.
59. Lee, B., and Richards, F. M. (1971) *J. Mol. Biol. 55*, 379−400.
60. Richmond, T. J. (1984) *J. Mol. Biol. 178*, 68−89.
61. Nicholls, A., and Honig, B. (1990) *J. Comput. Chem. 12*, 435−445.
62. Richards, F. M. (1983) *Annu. Rev. Biophys. Bioeng. 6*, 151−176.
63. Connolly, M. L. (1983) *J. Appl. Crystallogr. 16*, 548−558.
64. Presta, L. G., and Rose, G. D. (1988) *Science 240*, 1632−1641.
65. Richardson, J. S., and Richardson, D. C. (1988) *Science 240*, 1648−1652.
66. Harper, E. T., and Rose, G. D. (1993) *Biochemistry 32*, 7605−7609.
67. Wishart, D. S., Sykes, B. D., and Richards, F. M. (1991) *J. Mol. Biol. 222*, 311−333.
68. Quiram, P. A., and Sine, S. M. (1998) *J. Biol. Chem. 273*, 11007−11011.
69. Luo, S., Nguyen, T. A., Cartier, G. E., Olivera, B. M., Yoshikami, D., and McIntosh, J. M. (1998) *Soc. Neurosci. Abstr. 24*, 87.
70. Kohn, A. J. (1990) *Malacologia 32*, 55−67.
71. Olivera, B. M., Rivier, J., Scott, J. K., Hillyard, D. R., and Cruz, L. J. (1991) *J. Biol. Chem. 266*, 22067−22070.
72. Momany, F. A., McGuire, R. F., Burgess, A. W., and Scheraga, H. A. (1975) *J. Phys. Chem. 79*, 2361−2381.
73. Némethy, G., Pottle, M. S., and Scheraga, H. A. (1983) *J. Phys. Chem. 87*, 1883−1887.
74. Loughnan, M., Bond, T., Atkins, A., Cuevas, J., Adams, D. J., Broxton, N. M., Livett, B. G., Down, J. G., Jones, A., Alewood, P. F., and Lewis, R. J. (1998) *J. Biol. Chem. 273*, 15667−15674.
75. Fainzilber, M., Hasson, A., Oren, R., Burlingame, A. L., Gordon, D., Spira, M. E., and Zlotkin, E. (1994) *Biochemistry 33*, 9523−9529

BI9826254