Solution structure of α-conotoxin SI

Andrew J. Benie[a], David Whitford[a], Balazs Hargittai[b,1], George Barany[b,2], Robert W. Janes[c,*]

[a] *Molecular and Cellular Biology, Division of Biomedical Sciences, St. Bartholomew's and the Royal London School of Medicine and Dentistry, Queen Mary and Westfield College, University of London, Mile End Road, London E1 4NS, UK*
[b] *Department of Chemistry, University of Minnesota, 207 Pleasant St. S.E., Minneapolis, MN 55455, USA*
[c] *School of Biological Sciences, Queen Mary and Westfield College, University of London, Mile End Road, London E1 4NS, UK*

Received 3 May 2000

Edited by Thomas L. James

Abstract The nuclear magnetic resonance solution structure of α-conotoxin SI has been determined at pH 4.2. The 36 lowest energy structures show that α-conotoxin SI exists in a single major solution conformation and is stabilized by six hydrogen bonds. Comparisons are made between the SI solution structure and the solution and crystal structures of α-conotoxin GI. Surprisingly, a high degree of similarity between the backbone conformations of the GI crystal and the SI solution structures is seen in the region of lowest sequence homology, namely residues Gly-8 to Ser-12. This similarity is more surprising when considering that in SI a proline replaces the Arg-9 found in GI. The correspondence in conformation in this region provides the definitive evidence that it is the loss of the arginine basic charge at residue 9 which determines the differences in toxicity between GI and SI, rather than any changes in conformation induced by the cyclic proline residue.

Key words: Conotoxin; Nuclear magnetic resonance; Three-dimensional structure; Nicotinic acetylcholine receptor; Circular dichroism; Disulfide linked peptide

1. Introduction

The nicotinic acetylcholine receptors (nAChRs) are a class of ligand-gated channels mediated by acetylcholine (ACh), which when ligand-bound, allow the passage of potassium and sodium ions across the postsynaptic membrane, hence enabling neurotransmission. The nAChRs can be categorized into two main classes: neuromuscular nAChRs and neuronal nAChRs. The nAChRs of mammalian neuromuscular junctions are formed of five subunits, $\alpha_2\beta\gamma\delta$, arranged in a pseudosymmetric-pentameric structure, as are those from *Torpedo californica* electric organ. Two sites are available for ACh binding located at the α/γ and α/δ subunit interfaces [1]. It is hypothesized that the neuronal form of nAChRs also adopts a pseudosymmetric-pentameric arrangement.

The α-conotoxins, from marine snails of the genus *Conus*, form a family of polypeptide toxins that selectively block nAChRs, and range in size from 12 to 25 residues. Their structures contain two or three disulfide bonds which constrain the conformations that these polypeptides can adopt. The smaller 'classical' α-conotoxins, ranging in size from 12 to 20 amino acids, selectively block the neuromuscular nAChRs. α-Conotoxin GI [2] from *C. geographus*, is 13 amino acids in length with disulfides between residues 2 and 7 and 3 and 13 (Table 1), and is the best characterized member of the family. α-Conotoxin MI from *C. magus* [3] is 14 amino acids long, having an additional N-terminal amino acid, but maintains an identical disulfide pattern (Table 1). α-Conotoxin SI [4] from the species *C. striatus*, the striated cone, is 13 residues long (Table 1), with identical disulfide pairings to the others.

α-Conotoxins GI and MI display a greater degree of toxic activity than SI, and more similar toxicity profiles, even though in terms of sequence homology, GI and MI are more diverse than GI and SI, differing by five in comparison to three residues. α-Conotoxins GI and MI can distinguish between the two different ACh binding sites in the neuromuscular nAChRs [5,6]. Both display a higher affinity in binding to the α/δ site, being some four orders of magnitude greater than the low-affinity α/γ site, in the nAChRs of mammalian muscle [6]. In contrast, MI displays a higher affinity for the α/γ site of *T. californica* electric organ than for its α/δ site [7]. While α-conotoxin SI displays a preferential affinity towards the α/δ site in mammalian muscle nAChRs, it is unique within the family in that it does not distinguish between the two ACh binding sites in *T. californica* electric organ [8]. Additionally, the block of mammalian muscle nAChRs produced by α-conotoxin SI is negligible, in contrast to GI and MI, which are

Table 1
Sequences of selected α-conotoxins, showing the common disulfide pattern used in each

Conotoxin	Sequence*														
SI		Ile	Cys	Cys	Asn	Pro	Ala	Cys	Gly	Pro	Lys	Tyr	Ser	Cys	-NH_2
GI		Glu	Cys	Cys	Asn	Pro	Ala	Cys	Gly	Arg	His	Tyr	Ser	Cys	-NH_2
MI	Gly	Arg	Cys	Cys	His	Pro	Ala	Cys	Gly	Lys	Asn	Tyr	Ser	Cys	-NH_2

*Residues with a black background indicate sequence identities within these α-conotoxins.

*Corresponding author. Fax: (44)-20-8983 0973.
E-mail: r.w.janes@qmw.ac.uk

[1] Present address: Department of Chemistry, University of Arizona, 1306 University Blvd., Tucson, AZ 85721, USA.
[2] Joint appointment in Department of Laboratory Medicine and Pathology.

Abbreviations: ACh, acetylcholine; CD, circular dichroism; LED, longitudinal eddy current delay; nAChR, nicotinic acetylcholine receptor; RAPL, relative amide proton lifetime; RMSD, root mean square deviation

0014-5793/00/$20.00
PII: S0014-5793(00)01724-5

lethal at very low concentrations [4,9]. α-Conotoxin SI does, however, significantly block the *T. californica* electric organ nAChRs [10]. These notable differences in toxicity within the members of the α-conotoxin family ultimately arise from three factors: differences in the toxin sequences, differences in the nAChR sequences at the two binding sites, and subsequently, differences in the interactions of the toxins with the residues that line the receptor ACh binding pockets. It has been demonstrated that the different degrees of toxicity displayed by conotoxin SI result primarily from the presence of Pro in position 9, as opposed to the Arg or Lys present at this position in GI and MI, respectively [8,10]. This substitution places a hydrophobic cyclic residue in the site instead of a polar charged side chain. In addition, proline due to its cyclic nature is limited by steric constraints in the ϕ/ψ angles that it can adopt, and this imposes conformational restraints on the remaining residues of the chain C-terminal to position 9 in SI. The aims of the present investigation were to determine the three-dimensional solution conformation of α-conotoxin SI, to compare and contrast this conformation with the solution [11] and crystal [12] structures of α-conotoxin GI, and to establish the extent to which proline in position 9 alters the conformation of the subsequent residues at the C-terminus of the polypeptide.

2. Materials and methods

2.1. Synthesis of α-conotoxin SI

Most of the materials, solvents, and methods used in the synthesis have been described previously ([13], and references therein). Fmoc-Cys(Xan)-OH was prepared in our laboratory by a previously published procedure [14]. Analytical and preparative HPLC methodologies have also been previously reported [13]. Low-resolution fast atom bombardment mass spectrometry was carried out on a VG Analytical 707E-HF low-resolution double focusing mass spectrometer, to verify the identity of the peptides (intermediate and final product) produced.

2.2. Circular dichroism (CD) spectroscopy

α-Conotoxin SI samples at protein concentrations of ~1 mg/ml (~1 mM in each case), were dissolved in either 90% H_2O/10% ^2H_2O at pH 4.2 (adjusted using dilute HCl, and uncorrected for the deuterium isotope effect), or in a sodium acetate solution buffered to pH 7.0. Measurements for baselines contained only the corresponding pH solutions without peptide. The CD spectra were obtained using an Aviv 62ds spectropolarimeter over a wavelength range from 300 to 180 nm, at an interval of 0.2 nm in a 0.2 cm pathlength cell. For each sample, five spectra and five baselines were collected and averaged [15] and smoothed using a Savitsky–Golay filter [16]. The secondary structures were calculated using a normalized constrained least squares algorithm [17], using two different reference data bases [18,19]. The normalized root mean square deviation (RMSD) parameter was calculated as a measure of the quality of the fit of the calculated secondary structures to the data.

2.3. Nuclear magnetic resonance (NMR) spectroscopy

Polypeptide samples (~5 mM) in 90% H_2O/10% ^2H_2O were adjusted to pH 4.2 using dilute HCl (uncorrected for the deuterium isotope effect). 2D NOESY and TOCSY spectra were acquired on a Varian Unity spectrometer, operating at a nominal proton frequency of 599.92 MHz. DQF-COSY and ^1H-^{13}C HSQC spectra were acquired using a Varian Unity*Plus* spectrometer with a nominal proton frequency of 500.10 MHz. The 2D *J*-resolved spectrum [20] was acquired on a Bruker AMX spectrometer with a nominal proton frequency of 600.12 MHz. All of the experiments were carried out at 283 K using 5 mm triple resonance probes equipped with gradients along the *z*-axis. All spectra, with two exceptions, were performed with 4096 points in the direct dimension and between 460 and 600 points in the indirect dimension. The spectral widths in these experiments were 6 kHz in both dimensions and quadrature detection was performed using the method described by States [21]. The exceptions were the ^1H-^{13}C HSQC and 2D *J*-resolved experiments. The ^1H-^{13}C HSQC was performed using the States method for quadrature detection. The spectral widths were 6 kHz for proton and 17.6 kHz for carbon, the corresponding number of points acquired were 1216 and 860 respectively. The 2D *J*-resolved experiment deviates slightly from that described in the reference [20] in that here the TPPI method was used for quadrature detection. The experiment was performed with a spectral width of 7.2 kHz and 4096 points in the direct dimension and 100 Hz and 256 points in the indirect dimension. The TOCSY spectra were acquired using the 'clean MLEV17' [22] pulse train and a spin lock time of either 40 or 80 ms, and the NOESY spectra were acquired with mixing times of 150, 250, and 500 ms. The delay between the first ^1H 90° pulse and the first ^{13}C 180° pulse in the ^1H-^{13}C HSQC was set to 3.6 ms [$\approx 1/(2J_{CH})$]. All spectra had recycle times of 2.5 s. The WATERGATE procedure [23] was used for water suppression, except in the DQF-COSY and the 2D *J*-resolved spectra where presaturation was used. The spectra were processed using the NMRpipe program [24], and analyzed using the program XEasy [25]. In all cases except the 2D *J*-resolved, where no window functions were used, the processing was the same. The application of a Gibbs filter was followed by a Hamming window, or in the case of the DQF-COSY a cosine squared window, and then zero filling to the next power of 2 prior to Fourier transformation. Where required, additional solvent suppression was performed by the use of low-frequency deconvolution [26]. Spectra used for coupling constant analysis were processed using Bruker's XWINNMR software and analyzed using the DECO module of XWINNMR in the case of the DQF-COSY and the program INFIT [27] for the 2D *J*-resolved spectrum.

2.4. Relative amide proton lifetime determination

The magnetic field gradient dependence of the amide residues was determined by the use of the longitudinal eddy current delay (LED) NMR experiment [28]. The LED experiment performed incorporated bipolar gradients [29] in addition to water suppression by the use of the WATERGATE procedure.

A series of spectra were measured for values of the gradient strength in the range of 0–700 mT m^{-1} in random order steps of 20 mT m^{-1}, using bipolar sine-shaped gradients of base length 2 ms and with a diffusion period of 200 ms. The spectra acquired had spectral widths of 7.2 kHz with a corresponding number of points acquired being 32 768 for a total of 512 transients per experiment.

The intensities of the peaks were determined using Bruker's XWINNMR software. To determine the RAPL values the diffusion of the polypeptide in the solution was obtained. This was achieved by fitting the resolved aliphatic resonances into the equation $I_i = I_{0N} \exp(-K^2\ \Delta\ D_i)$. Here, I_i is the observed intensity of resonance *i*, I_{0N} is the intensity of the non-exchanging component, D_i is the diffusion coefficient for resonance *i*, Δ is the length of the diffusion period in the LED sequence, *K* is equal to $\gamma G\delta$, where γ is the gyromagnetic ratio of the nucleus responsible for resonance *i*, and *G* and δ are the strength and duration of the magnetic field gradient. This polypeptide diffusion coefficient was then normalized using an equivalent fit for the water resonance into the above equation, using the known diffusion coefficient of 2.30×10^{-9} m^2 s^{-1} for water [30]. The relative amide proton lifetimes were determined using the methods according to Liu et al. [31], and fitted to the equation $I_i = I_{0E} \exp(-K^2\ \Delta\ [D_w\ \{1-f\} + fDp]) + I_{0N} \exp(-K^2\ \Delta\ D_p)$. Here, I_{0E} is the intensity of the exchanging components of resonance *i*, in the absence of any gradient, D_w and D_p are the diffusion coefficients for water and polypeptide, respectively, and *f* is the fractional lifetime of the proton on the polypeptide.

2.5. Constraints used in the structure calculations

Approximately 92% of the α-conotoxin SI observed in the NMR spectra is in one major conformation. The solution structures were determined for this form. The sequential assignment was performed using the method described by Wüthrich and co-workers [32]. The NOESY spectra were integrated using the method described by Denk [33] in the program XEasy. These intensities were converted into upper distance limits, with the CALIBA program [34], using the default parameters. Stereospecific assignment of resonances were obtained by the analysis of an initial set of structures with the program GLOMSA [34]. $^3J_{HNH\alpha}$ coupling constants were converted into limits for the backbone angle ϕ, with the appropriate Karplus curve

[35] using the program HABAS [34]. Hydrogen bond restraints were introduced for those residues possessing relative amide proton lifetimes of 200 ms and the correct geometry in all 36 conformers comprising the initial set of structures.

2.6. Structure calculations

A preliminary set of 50 structures was calculated with $T_{HIGH} = 8.0$ and 10 000 steps using only the NOE restraints in the program DYANA [36]. These structures were then used to determine the stereospecific assignments and potential hydrogen bonds (see above). A final set of structures was then calculated incorporating all of the structural restraints. The 36 conformers with the lowest target function energies constitute the final family for the SI structure.

2.7. Structure analyses

Structures were overlaid, and electrostatic surface potentials were calculated using the program MOLMOL [37] by pairwise least-squares analyses of selected residues. PROMOTIF [38] was used to determine secondary and tertiary structural features of the 36 lowest energy structures, and PROCHECK-NMR [39] was used to examine the quality of the structure geometry for these lowest energy conformations. Hydrogen bonds and surface accessibility were determined using RasMol (Version 2.6b2) [40], and MOLMOL, respectively.

3. Results

3.1. α-Conotoxin synthesis

The synthesis of α-conotoxin SI proceeded by an optimized modification [13] of a previously published strategy [41]. In brief, Fmoc solid-phase peptide synthesis was used with a combination of *S*-acetamidomethyl and *S*-xanthenyl cysteine protecting groups to direct disulfide formation in solution, the large loop being formed first. Fast atom bombardment mass spectrometry was used to characterize peptide intermediates and the final product. The final yield of α-conotoxin SI was 21%, after purification by preparative HPLC, to give material that was >99% homogeneous, and which coeluted with an authentic commercially bought standard.

3.2. CD results

Examination of the CD spectra (Fig. 1) for α-conotoxin SI at pH 4.2 and pH 7.0 shows them to be virtually identical. While less accurate than for more 'standard' larger globular protein structures from which the reference data bases have been derived [42], calculations of secondary structure content made using two different data bases [18,19] gave comparable results for both pH values. Together these factors strongly suggest that the structure determined for α-conotoxin SI under the NMR pH 4.2 conditions should be essentially that found at a physiologically more relevant pH 7.0.

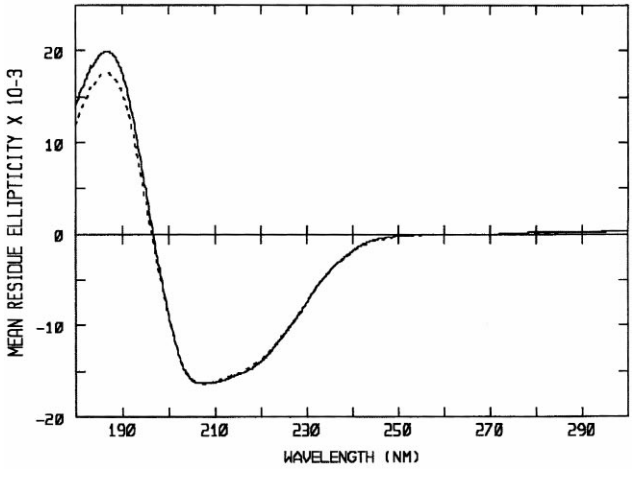

Fig. 1. CD spectra of α-conotoxin SI at pH 4.2 (solid line) and pH 7.0 (dashed line).

3.3. Sequence-specific assignment

The use of the DQF-COSY and TOCSY spectra allowed nearly all of the resonances to be identified (data not shown). The remaining ambiguities, the overlapping resonances of the Pro-5 and Pro-9 H^{β} and H^{γ} protons, were resolved with the use of the ^{1}H-^{13}C HSQC spectrum. The 'fingerprint' region of the DQF-COSY spectrum (Fig. 2A) allowed the identification of 11 of the 12 predicted HN-Hα crosspeaks. The absent crosspeak is for the N-terminal isoleucine residue, which under the experimental conditions underwent rapid proton exchange with the solvent. The DQF-COSY also shows the presence of a set of less intense resonances for a minor conformation. The integration of these peaks suggests this conformation constitutes less than 8% of the total peptide observed. The DQF-COSY spectra allowed the simple and rapid identification of the aromatic resonances of the tyrosine residue. The observation of degenerate H^{δ} and H^{ε} resonances, suggests that the tyrosine ring is free to rotate in α-conotoxin SI, and this is consistent with the solution structure.

It is possible to identify three regions in the NOESY spectra on the basis of connectivities in the amide region (Fig. 2B). In the first, the sequential assignment of the spin systems in the region Cys-2 to Asn-4 was assisted by the presence of a weak Asn-4 $^{4}J_{H\beta H\delta}$ crosspeak in the TOCSY spectrum (with a mixing time of 80 ms). In the second region, Ala-6 to Gly-8, unambiguous assignment was possible due to the presence of three different spin systems. The third region, Lys-10 to Cys-13, has a unique spin system for the lysine. The remaining three residues, Ile-1, Pro-5, and Pro-9, do not give rise to HN-Hα crosspeaks. The locations for the prolines in the sequence were obtained by analyses of NOEs from neighboring residues.

3.4. Coupling constant analysis

The extraction of the $^{3}J_{HNH\alpha}$ coupling constants from both the DQF-COSY and the 2D *J*-resolved spectra allowed the determination of the backbone angle ϕ, for nine out of the 13 residues in α-conotoxin SI. The two different approaches to the elucidation of the $^{3}J_{HNH\alpha}$ coupling constants were found to be in good agreement ($\Delta J > 1.0$ Hz).

3.5. Relative amide proton lifetime determination

The polypeptide diffusion coefficient was determined to be

Table 2
Amide lifetimes during the diffusion period (200 ms), giving indirect evidence of slow exchange with the solution

Residue	Lifetime on the peptide (ms)
Cys-2	20
Cys-3	40
Asn-4	200
Ala-6	200
Cys-7	200
Gly-8	200
Lys-10	200
Tyr-11	200
Ser-12	200
Cys-13	100

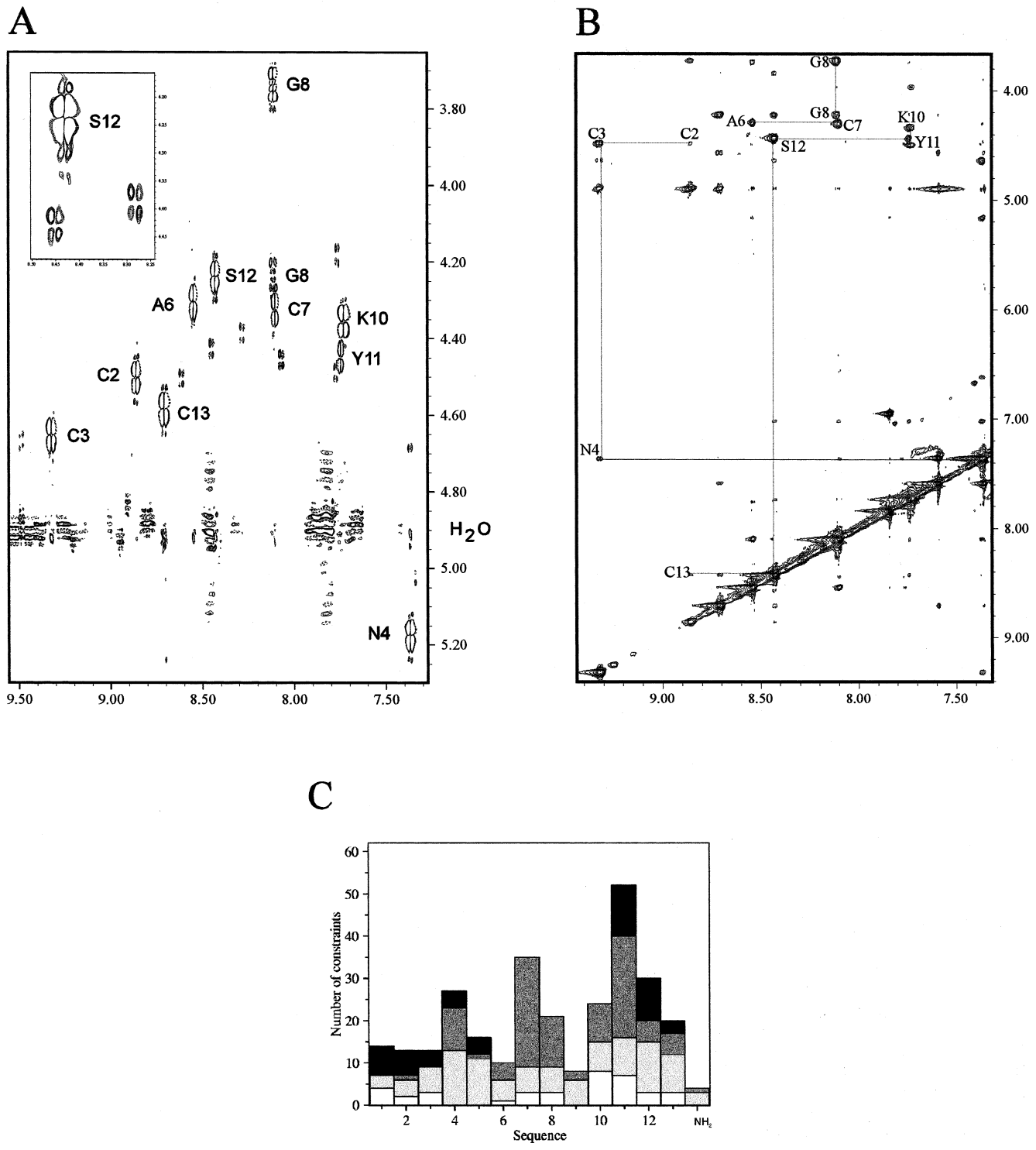

Fig. 2. A: 'Fingerprint' region of the DQF-COSY spectrum showing the major and minor conformation peaks for α-conotoxin SI. The inset box shows a major peak and two adjacent minor conformation peaks. The marking H_2O indicates horizontal features arising from the water suppression technique used (see text). B: Homonuclear NOESY spectrum of α-conotoxin SI showing the NOE connectivities used to assign the regions: Cys-2–Asn-4; Ala-6–Gly-8; Lys-10–Cys13. C: The total number of NOE constraints per residue. White represents intra-residue constraints; light gray are short range ($i,i+1$), dark gray are medium range ($i,i<4$), and black are long range constraints ($i,i\geq 4$).

9.36×10^{-10} m^2 s^{-1}, which was necessary for the subsequent determination of the relative amide proton lifetime (RAPL) values. RAPLs like H/D exchange rates provide indirect evidence of the presence of hydrogen bonds. A long RAPL is an indicator that the proton in question may be involved in a hydrogen bond or is buried from solvent accessibility in some way. From the RAPL values given in Table 2, a number of protons have long lifetimes, indicating that they are poten-

tially in slow exchange with the solution. As indicated below, three proved in suitable geometry after initial structure calculations, to be included as hydrogen-bonded constraints in the later stages of structure calculations.

3.6. Solution structure calculations

A total of 457 NOESY crosspeaks were assigned, integrated and converted into upper distance limits with the program CALIBA, reported in Fig. 2C. This corresponds to a total of 162 meaningful distance restraints (non-meaningful restraints are those that correspond to fixed distances or are so weak that they cannot be violated). A total of 37 intra-residue, 50 short range ($i,i+1$), 50 medium range ($i,i<4$) and 25 long range ($i,i\geq4$) meaningful NOEs were observed. This gives an average of $\sim$12 meaningful NOEs per residue. The relatively large number of long range NOEs is to be expected for a polypeptide like α-conotoxin SI, both small in length and constrained by two disulfides.

The initial structure calculation allowed the stereospecific assignment of five sets of diastereotopic protons, and the addition of constraints for three hydrogen bonds, supported by the data from the RAPL results (Table 2).

3.7. α-Conotoxin SI structure analyses

Fig. 3A shows the superposition of the 36 lowest energy NMR α-conotoxin SI structures. Since Ile-1 and Cys-13 were less well defined positionally because of fewer NOE constraints than the remaining residues, the superpositions were calculated using the backbone atoms of residues Cys-2 through Ser-12. These superpositions gave a pairwise root mean square deviation (RMSD) of 0.19 ± 0.10 Å for the backbone atoms and 0.62 ± 0.16 Å for the non-hydrogen atoms as calculated by MOLMOL [37]. Thus, the ensemble of structures are very similar, probably as a consequence of the constraints imposed by the disulfide bonds between Cys-2 and Cys-7, and Cys-3 and Cys-13. Of the 324 non-terminal, non-glycine, non-proline residues in the 36 lowest energy structures, 61% (199 residues) fall in the most favored region of the Ramachandran plot of ϕ/ψ angles, and the remaining 39% (125 residues) lie in the additionally allowed regions, as calculated by PROCHECK-NMR [39].

The disulfide bond between Cys-2 and Cys-7 was found to be a left-handed spiral in 32 of the 36 structures, as defined by PROMOTIF [38]. The remaining four structures were one right-handed hook, one right-handed spiral, and two with

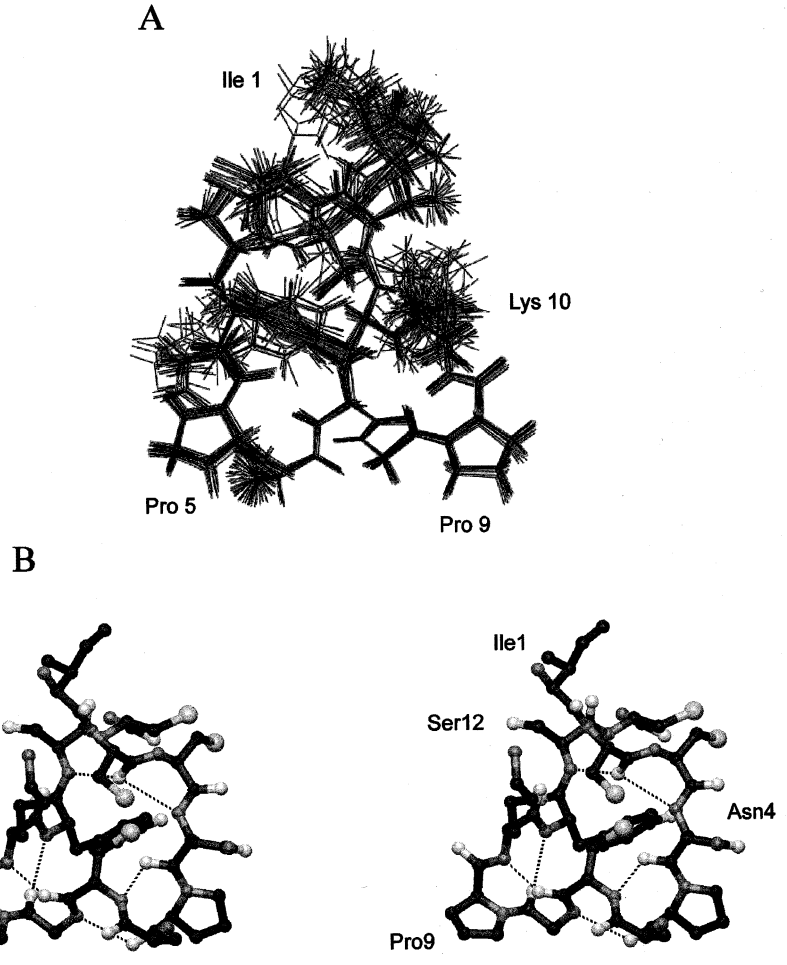

Fig. 3. A: Overlay of the 36 lowest energy α-conotoxin SI NMR solution structures, calculated using the backbone atoms of residues Cys-2 to Ser-12 by a pairwise least-squares procedure. The figure and alignments were produced using the program MOLMOL [37]. B: A stereo representation of the lowest energy α-conotoxin SI NMR solution structures, showing the six hydrogen bonds found in the structure. The carbon atoms are black, the oxygen atoms are white, the nitrogen atoms are gray, and the sulfur atoms are larger radius white atoms. The hydrogen bonds are shown as dashed lines. The figure was produced using the program MOLMOL.

Table 3
Hydrogen bonds found in the lowest energy SI solution structure

D–A[a]	Å	H–A (Å)	D–H–A (°)
N4N–C2O	3.19	2.41	134.4
C7N–N4O	3.27	2.36	150.7
G8N–P5O	2.95	2.04	149.3
K10N–C7O	2.53	1.66	143.1
Y11N–G8O	3.05	2.08	164.4
S12N–C2O	2.70	1.77	152.7

[a]D is the donor atom, A is the acceptor atom in the hydrogen bonding pattern.

an uncategorized motif, with geometries, $c_{\alpha\beta} = \sim 60°$, $c_{\beta s} = \sim 128°$, $c_{ss'} = \sim -162°$, $c_{s'\beta'} = \sim -156°$, $c_{\beta'\alpha'} = \sim -60°$. The disulfide bond between Cys-3 and Cys-13 did not correspond to a categorized disulfide motif. The lowest energy structure had a geometry with $c_{\alpha\beta} = -163.3°$, $c_{\beta s} = 101.1°$, $c_{ss'} = -49.8°$, $c_{s'\beta'} = -167.7°$, $c_{\beta'\alpha'} = 172.4°$, as did 22 of the other 36 lowest energy structures. Eleven of the remaining structures had a disulfide geometry represented by $c_{\alpha\beta} = \sim -150°$, $c_{\beta s} = \sim 109°$, $c_{ss'} = \sim -51°$, $c_{s'\beta'} = \sim 134°$, $c_{\beta'\alpha'} = \sim -53°$, while the other two structures were classified as a short right-handed hook motif. The two proline residues at positions 5 and 9 in α-conotoxin SI are both in the *trans* conformation in all of the 36 lowest energy structures.

The secondary structure of α-conotoxin SI comprises two overlapping 3_{10} helices between residues Asn-4 and Gly-8, and Cys-7 and Tyr-11 in the lowest energy structure, and the majority of the remaining 36 structures, as determined by PROMOTIF. In 12 of the 36 structures, the first helix is replaced by two overlapping β-turns between residues Asn-4 to Cys-7, and Pro-5 to Gly-8. However, as the defined ϕ/ψ angles for a 3_{10} helix and a β-turn are highly comparable, this does not constitute a different conformation. Nine of the 10 lowest energy structures, and half the ensemble as a whole, contain an inverse γ-turn between residues Cys-2 to Asn-4 (an 'inverse B' as defined in PROMOTIF, for all but the two highest energy structures containing a γ-turn, which are classified as 'inverse P'). These structural motifs are supported by a total of six hydrogen bonds (Table 3 and Fig. 3B). All have long RAPLs for their donor atoms (Table 2), but only three, those present in all 36 initial structures calculated, were included for final structure calculations. Four of these hydrogen bonds are involved in the 3_{10} helices and one is involved in the inverse γ-turn, described above. A further hydrogen bond, Ser-12N to Cys-2O, completes the set. The hydrogen bonding is therefore quite extensive. Of interest, the proton on Ala-6N also has a long RAPL but it is not involved with any hydrogen bonding and faces towards the solvent in the structures. Steric hindrance effects from Asn-4, Pro-5 and its own methyl side chain mean that solvent cannot easily reach this proton to enable exchange to take place, nor can any intermediate transition state that would arise during exchange be formed as a result; hence the long RAPL observed for this proton.

Table 4
Accessible surface area for each residue in the lowest energy SI solution structure

Residue	Accessible area (%)[a]
Ile-1	64.6
Cys-2	14.1
Cys-3	38.9
Asn-4	32.2
Pro-5	51.9
Ala-6	49.8
Cys-7	5.6
Gly-8	24.2
Pro-9	65.0
Lys-10	45.9
Tyr-11	41.5
Ser-12	26.0
Cys-13	37.1[b]

[a]Calculated in MOLMOL, based on the method according to [45], with a precision value of 10.
[b]This value includes the amidation.

For a small polypeptide, it is interesting to note that a number of residues are buried in the structure (Table 4). Ten of the 13 residues are less than half solvent exposed. Residues Cys-2 and Cys-7 make a buried hydrophobic core to the polypeptide, although the other disulfide pair, Cys-3 and Cys-13, is more exposed. The most notable residue with regards to the extent of surface exposure is Pro-9 which is well over half exposed to solvent, reflecting its position as being at a 'corner' of the molecule.

4. Discussion

4.1. Relationship to related structures

Structures have been reported for the related family members of the α-conotoxins GI and MI. For GI, solution [11,43,44] and crystal [12] structures are available. Comparisons were made between the current SI structure and the GI structures with coordinates deposited in the Protein Data Bank, a solution [11], and the crystal [12] structure. Table 5

Table 5
Selected RMSDs for the 36 SI structures, and various comparisons between the SI and GI solution structures [11], and the GI crystal structure [12]

Residue range	All 36 SI[a]	SI and GI_X[b]	SI and GI_S[c]	SI and GI_X and GL_S[d]	GI_X and GI_S[e]
2⇒6	0.16 (10)	0.435	0.523	0.48 (4)	0.492
3⇒7	0.11 (7)	0.400	0.588	0.51 (10)	0.555
4⇒8	0.08 (5)	0.284	0.429	0.39 (9)	0.451
5⇒9	0.08 (5)	0.250	0.652	0.55 (26)	0.743
6⇒10	0.05 (2)	0.270	0.667	0.55 (24)	0.711
7⇒11	0.05 (3)	0.259	0.541	0.46 (18)	0.580
8⇒12	0.07 (4)	0.216	0.618	0.47 (22)	0.565

[a]All 36 ensemble members for the SI solution structure, with estimated standard deviations in parentheses.
[b]The lowest energy SI structure and the GI crystal structure.
[c]The lowest energy SI structure and the GI solution structure.
[d]The lowest energy SI structure and the GI crystal and solution structures, with estimated standard deviations in parentheses.
[e]Comparison between the GI crystal and GI solution structures.

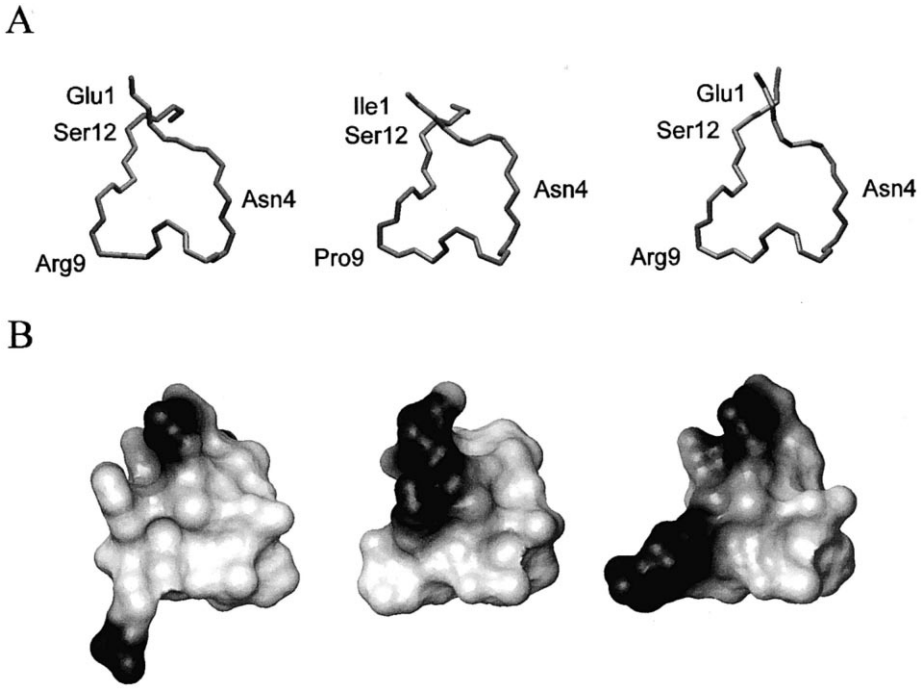

Fig. 4. A: Backbone structures of the most representative GI solution structure (left), the lowest energy solution structure of α-conotoxin SI (middle), and the GI crystal structure (right), aligned using the backbone residues Cys-2 to Ser-12. B: Surface topologies and electrostatic potentials for the structures as laid out, and with the same orientation, as in A. The large protrusion towards the lower left of the GI solution structure is from the Arg-9, while in the GI crystal structure the protrusion here is from Arg-9 and His-10. The electrostatic potentials are represented by black shading for positively charged regions within each of the molecules. The figure and the electrostatic surface charges were produced by MOLMOL.

gives comparison values for pairwise RMSD calculations for selected five-residue spans along the toxins. Fig. 4 shows the solution structures of GI compared with that of SI, and with the crystal structure of GI.

4.2. Comparison of the solution structures of α-conotoxins SI and GI

The secondary structure as calculated by PROMOTIF for the most representative solution structure of α-conotoxin GI comprises a 3_{10} helical segment between residues 5 and 10. The disulfide motifs are a left-handed spiral for Cys-2 and Cys-7, and a right-handed spiral for Cys-3 and Cys-13. No turns of any kind are reported for the structure. This is in broad general agreement with the secondary structure reported here for SI. A least squares pairwise RMSD comparison of the backbone atoms of residues Cys-2 to Ser-12 for these structures gives an overall value of 0.726 Å. Residues Cys-2 to Gly-8 in these polypeptides are identical, and a similar comparison based on this region yields an RMSD value of 0.604 Å. Comparing in a similar manner the 'second half' of the molecule, residues Ala-6 to Ser-12, gives an RMSD value of 0.642 Å. These values indicate that the conformations of the N-termini of the two polypeptides are slightly more similar than the C-termini, although the difference is not very substantial. From the data in Table 5 for the solution structures of SI and GI, it is clear that there is a general agreement between the conformations with no significant features arising along the polypeptide chains.

4.3. Comparison of the solution structure of α-conotoxin SI and the crystal structure of α-conotoxin GI

From PROMOTIF, the crystal structure of α-conotoxin GI comprises two overlapping 3_{10} helices between residues Asn-4 and Gly-8, and Pro-5 and Tyr-11. The disulfides are described as a left-handed spiral for Cys-2 and Cys-7, and a right-handed spiral for Cys-3 and Cys-13. Once again, no turns are reported for the structure. This secondary structure description is more akin to that reported here for SI than is the GI solution structure, although in SI, an inverse γ-turn is found between Cys-2 and Asn-4. A pairwise alignment of the backbone atoms of residues Cys-2 to Ser-12 between the two structures gives a value of 0.503 Å RMSD. Using again subsets of residues, for Cys-2 to Gly-8, the RMSD value is 0.393 Å, while that for Ala-6 to Ser-12 is 0.287 Å. Here, these values show that the C-terminal residues with less sequence homology between the two polypeptide toxins overlay notably better than the N-terminal residues which are identical in both structures. Additionally, referring to Table 5, the RMSD values show that the backbone of the solution structure of SI far more closely resembles the backbone of the crystal structure of GI, than a comparable solution structure of GI, and that the SI structure is closer to the GI crystal structure than is the solution structure of GI. There is also a trend that the C-termini are closer in conformation for the SI solution and the GI crystal structures, although this region presents the lowest sequence homology between the two α-conotoxins.

4.4. Relevance to toxicity

The differences in the sequence of the residues in positions 9 and 10 in α-conotoxins SI, GI, and MI (10 and 11 for MI) have been shown to be significant in terms of the species specificity of conotoxin blocking, and in the selectivity towards one of the two subunit ACh binding sites [8]. However, in this region of the toxins, where it would be expected that

the largest backbone conformational differences would be present, the solution structure of α-conotoxin SI and crystal structure of GI exhibit the highest degree of similarity in their conformations, as reflected by the RMSD value, than in any other regions. In both the solution and crystal structures of GI the two side chains of residues Arg-9 and His-10 are highly exposed and are polar. In the solution structure of SI the hydrophobic Pro-9 side chain is relatively exposed. While the backbone conformations are highly comparable in the C-terminal region, in contrast the surface topology of these conotoxin structures, as presented in Fig. 4, clearly shows there are substantial differences in their shapes and electrostatic potentials. In the GI structures there are two positively charged regions which are well separated from each other, their different appearances arising from differing orientations of the Arg-9 side chains. In the SI structure there is one positively charged region which includes the N-terminus, common to all these structures. Therefore the electrostatic surfaces being presented to the ACh binding sites of the nAChRs are significantly different.

These structural results can now be compared with the activity of synthetic analogues of the α-conotoxins, to help understand their biological role. Such studies support the hypothesis that the nature of the side chains at positions 9 and 10 are essential to the selectivities and binding affinities of nAChRs from both muscle and *T. californica*. For example, analogues of SI have been synthesized in which Lys-10 is replaced by either histidine or asparagine, as found in GI and MI, respectively [8]. Both analogues retain 'SI-like' toxicity towards mammalian nAChRs, but a reduced affinity for one of the two sites in *T. californica* nAChRs. It seems then that this *T. californica* binding site, in particular, is highly sensitive to the surface features and charge of the toxin, but that the muscle nAChRs are less sensitive to these types of alterations in geometry and electrostatics.

Another synthetic analogue has been reported in which Pro-9 in SI has been changed to a lysine [8], removing the conformational restriction imposed by the proline ring and adding a positively charged residue. This produces a polypeptide with affinities and selectivities which are more like those of MI and GI. This suggests that the Lys-9 side chain has a significant interaction with the binding site and that the addition of positive charge at this position may be of more importance in determining selectivity than the conformational restraints imposed by the proline. This interpretation is supported by studies on a GI analogue in which the native Arg-9 has been replaced by an alanine [8]. This substitution, in which the positive charge was eliminated, but the substitution did not include a conformationally restricted amino acid, resulted in a molecule with more SI-like properties.

Thus it appears that surface topology and electrostatic potential, which are both significantly affected by the different residues in positions 9 and 10, are critically important for the affinities and relative selectivities of these α-conotoxins for the various types of nAChR binding sites. In this study we have shown that the solution structure of α-conotoxin SI is indeed highly comparable to the related GI conotoxin in backbone conformation. Surprisingly, the similarity between the two conotoxins is greatest between the SI solution, and the GI crystal structures, particularly where the two toxins differ most in their sequence homology, and where the conformationally restricted proline is found in the SI structure. This study therefore provides the definitive evidence that the inclusion in the SI sequence of proline at position 9 does not result in a distortion away from the 'conotoxin conformation' found in GI. However, comparisons of both the surface topology and electrostatic potential for the conotoxin structures shows them to be markedly different. Such features as these must therefore play an important part in determining the specificity and binding properties of these toxins to nAChR binding sites.

The coordinates for this structure are deposited at the Protein Data Bank (PDB), accession code 1QMW, and NMR data have been deposited at the BioMagResBank (BMRB), accession code 4503.

Acknowledgements: This work was supported by grants from The Royal Society (18967) to R.W.J., and from the National Institutes of Health (GM 43552) to G.B. A.J.B. was a recipient of an EPSRC studentship. We thank the Department of Crystallography at Birkbeck College, University of London, for the use of their circular dichroism spectrometer, and the Biomedical NMR center at National Institute of Medical Research for access to their 11.7T and 14.1T NMR spectrometers. We thank ULIRS for access to their 14.1T NMR spectrometer, and for access to specific computing facilities. We thank S.E.J. Rigby, H. Toms and B.A. Wallace for useful discussions, and F.W. Muskett is acknowledged for advice on aspects of DYANA.

References

[1] Machold, J., Weise, C., Utkin, Y., Tsetlin, V. and Hucho, F. (1995) Eur. J. Biochem. 234, 427–430.
[2] Cruz, L.J., Gray, W.R. and Olivera, B.M. (1978) Arch. Biochem. Biophys. 190, 539–548.
[3] McIntosh, M., Cruz, L.J., Hunkapiller, M.W., Gray, W.R. and Olivera, B.M. (1982) Arch. Biochem. Biophys. 218, 329–334.
[4] Zafaralla, G.C., Ramilo, C., Gray, W.R., Karlstrom, R., Olivera, B.M. and Cruz, L.J. (1988) Biochemistry 27, 7102–7105.
[5] Sine, S.M., Kreienkamp, H.-J., Bren, N., Maeda, R.M. and Taylor, P. (1995) Neuron 15, 205–211.
[6] Groebe, D.R., Dumm, J.M., Levitan, E.S. and Abramson, S.N. (1995) Mol. Pharmacol. 48, 105–111.
[7] Hann, R.M., Pagan, O.R. and Eterovic, V.A. (1994) Biochemistry 33, 14058–14063.
[8] Groebe, D.R., Gray, W.R. and Abramson, S.N. (1997) Biochemistry 36, 6469–6474.
[9] Ramilo, C.A., Zafaralla, G.C., Nadasdi, L., Hammerl, L.G., Yoshikami, D., Gray, W.R., Kristipati, R., Ramachran, J., Miljanich, G., Olivera, B.M. and Cruz, L.J. (1992) Biochemistry 31, 9919–9926.
[10] Hann, R.M., Pagan, O.R., Gregory, L.M., Jacome, T. and Eterovic, V.A. (1997) Biochemistry 36, 9051–9056.
[11] Gehrmann, J., Alewood, P.F. and Craik, D.J. (1998) J. Mol. Biol. 278, 401–415.
[12] Guddat, L.W., Martin, J.A., Shan, L., Edmundson, A.B. and Gray, W.R. (1996) Biochemistry 35, 11329–11335.
[13] Hargittai, B. and Barany, G. (1999) J. Peptide Res. 54, 468–479.
[14] Han, Y.X. and Barany, G. (1997) J. Org. Chem. 62, 3841–3848.
[15] Mielke, D.L. and Wallace, B.A. (1988) J. Biol. Chem. 263, 3177–3182.
[16] Savitsky, A. and Golay, M.J.E. (1964) Anal. Chem. 36, 1627–1639.
[17] Wallace, B.A. and Teeters, C.L. (1987) Biochemistry 26, 65–70.
[18] Chang, C.T., Wu, C.S.C. and Yang, J.T. (1978) Anal. Biochem. 91, 13–31.
[19] Brahms, S. and Brahms, J. (1980) J. Mol. Biol. 138, 149–178.
[20] Williamson, M.P. (1983) J. Magn. Reson. 55, 471–474.
[21] States, D.J., Haberkorn, R.A. and Ruben, D.J. (1982) J. Magn. Reson. 48, 286–292.
[22] Griesinger, C., Otting, G., Wüthrich, K. and Ernst, R.R. (1988) J. Am. Chem. Soc. 110, 7870–7872.
[23] Sklenar, V., Piotto, M., Leppik, R. and Saudej, V. (1993) J. Magn. Reson. 102, 241–245.

[24] Delaglio, F., Grzesiek, S., Vuister, G., Zhu, G., Pfeifer, J. and Bax, A. (1995) J. Biomol. NMR 6, 277–293.
[25] Bartels, C., Xia, T.H., Billeter, M., Güntert, P. and Wüthrich, K. (1995) J. Biomol. NMR 6, 1–10.
[26] Marion, D., Ikura, M. and Bax, A. (1989) J. Magn. Reson. 84, 425–430.
[27] Szyperski, T., Güntert, P., Otting, G. and Wüthrich, K. (1992) J. Magn. Reson. 99, 552–560.
[28] Gibbs, S.J. and Johnson Jr., C.S. (1991) J. Magn. Reson. 93, 395–402.
[29] Wu, D.H., Chen, A.D. and Johnson Jr., C.S. (1995) J. Magn. Reson. A115, 260–264.
[30] Hrovat, M.I. and Wade, C.G. (1981) J. Magn. Reson. 44, 62–75.
[31] Liu, M., Toms, H.C., Hawkes, G.E., Nicholson, J.K. and Lindon, J.C. (1999) J. Biomol. NMR 13, 25–30.
[32] Wüthrich, K., Billeter, M. and Braun, W. (1984) J. Mol. Biol. 180, 715–740.
[33] Denk, W., Baumann, R. and Wagner, G. (1986) J. Magn. Reson. 67, 386–390.
[34] Güntert, P., Braun, W. and Wüthrich, K. (1991) J. Mol. Biol. 217, 517–530.
[35] Vuister, G.W. and Bax, A. (1993) J. Am. Chem. Soc. 115, 7772–7777.
[36] Güntert, P., Mumenthaler, C. and Wüthrich, K. (1997) J. Mol. Biol. 273, 283–298.
[37] Koradi, R., Billeter, M. and Wüthrich, K. (1996) J. Mol. Graphics 14, 51–55.
[38] Hutchinson, E.G. and Thornton, J.M. (1996) Protein Sci. 5, 212–220.
[39] Laskowski, R.A., Rullmann, J.A.C., MacArthur, M.W., Kaptein, R. and Thornton, J.M. (1996) J. Biomol. NMR 8, 477–486.
[40] Sayle, R.A. and Milner-White, E.J. (1995) Trends Biochem. Sci. 20, 374–376.
[41] Munson, M.C. and Barany, G. (1993) J. Am. Chem. Soc. 115, 10203–10210.
[42] Wallace, B.A. and Corder, R. (1997) J. Peptide Res. 49, 331–335.
[43] Pardi, A., Galdes, A., Florance, J. and Maniconte, D. (1989) Biochemistry 28, 5494–5501.
[44] 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.
[45] Lee, B.K. and Richards, F.M. (1971) J. Mol. Biol. 55, 379–400.