Solution Structures of the *cis*- and *trans*-Pro30 Isomers of a Novel 38-Residue Toxin from the Venom of *Hadronyche Infensa sp*. that Contains a Cystine-Knot Motif within Its Four Disulfide Bonds[†,‡]

K. Johan Rosengren,[§] David Wilson,[§] Norelle L. Daly,[§] Paul F. Alewood,[§] and David J. Craik*[,§]

Institute for Molecular Bioscience, ARC Special Research Centre for Functional and Applied Genomics, University of Queensland, Brisbane QLD 4072, Australia

Received October 15, 2001; Revised Manuscript Received December 28, 2001

ABSTRACT: The primary sequence and three-dimensional structure of a novel peptide toxin isolated from the Australian funnel-web spider *Hadronyche infensa sp.* is reported. ACTX-Hi:OB4219 contains 38 amino acids, including eight-cysteine residues that form four disulfide bonds. The connectivities of these disulfide bonds were previously unknown but have been unambiguously determined in this study. Three of these disulfide bonds are arranged in an inhibitor cystine-knot (ICK) motif, which is observed in a range of other disulfide-rich peptide toxins. The motif incorporates an embedded ring in the structure formed by two of the disulfides and their connecting backbone segments penetrated by a third disulfide bond. Using NMR spectroscopy, we determined that despite the isolation of a single native homologous product by RP-HPLC, ACTX-Hi:OB4219 possesses two equally populated conformers in solution. These two conformers were determined to arise from cis/trans isomerization of the bond preceding Pro30. Full assignment of the NMR spectra for both conformers allowed for the calculation of their structures, revealing the presence of a triple-stranded antiparallel β sheet consistent with the inhibitor cystine-knot (ICK) motif.

The venoms of arachnids (spiders and scorpions) have been shown to contain complex mixtures of a number of different types of molecules. In particular, spider venoms are proving to be a rich source of peptidic neurotoxins. Typically, these small, disulfide-rich components are 30−50 amino acids in length and possess a wide range of activities. Some early examples of the toxins discovered in spider venoms are the ω-agatoxins (*1*), μ-agatoxins (*2*−*5*), and curtatoxins (*3*) isolated from the American funnel-web or grass spiders, *Agelenopsis aperta* and *Hololena curta* (Araneomorphae: Agelenidae).

The μ-agatoxins are 36−38 amino acid peptides containing an eight-cysteine framework forming four disulfide bonds. These toxins exhibit an irreversible, insect-selective activity, invoking a spastic excitatory paralysis attributed to action on sodium channels. Determination of the solution structure of one of the μ-agatoxins isolated from the venom of *Agelenopsis aperta* (μ-Aga-I)[1] revealed a tightly constrained triple-stranded antiparallel β-sheet structure similar to that of the ω-agatoxins (ω-Aga-IVA and ω-Aga-IVB) (*6*). These structures all conform to a common structural motif, referred to as the inhibitor cystine-knot (ICK) motif (*7*). The ICK motif is present in a number of different peptides, isolated from a variety of sources and possessing a diverse range of functions. While a large number of these ICK peptides have been isolated from the venoms of cone snails and spiders, they have also been isolated from numerous plant and fungal species. The activities associated with these peptides range from blocking and regulation of the three major voltage-gated ion channel types (Na^+, K^+, and Ca^{2+}) to enzyme inhibition and hemolysis.

Recently, several molecules have been isolated and characterized from the venom of Australian funnel-web spiders (Mygalomorphae: Hexathelidae: Atracinae). In general, these peptides differ in sequence to the agatoxins from the American funnel-web spiders. Four different families of molecules have been reported from Australian funnel-web spiders, each possessing a different cysteine framework. The 42 amino acid δ-atracotoxins (δ-ACTXs) (*8*) possess an eight-cysteine framework, including a unique cysteine triplet at positions 14−16, and are active at site 3 of the voltage-sensitive sodium channel where they slow channel inactivation. The 36−37 amino acid Janus-faced atracotoxins (J-ACTXs) (*9*) contain a different eight-cysteine framework that involves a rare vicinal disulfide bond, and while the exact site of action of these molecules remains unclear, they have insecticidal neurotoxicity. The ω-atracotoxins (ω-ACTXs) (*10*, *11*) contain a six-cysteine frame-

[†] This work was supported by a grant (D.J.C.) from the Australian Research Council.

[‡] Coordinates for both isomers have been deposited in the Protein Data Bank (PDB ID codes cis, 1KQH; trans, 1KQI).

* To whom correspondence should be addressed: Fax: 61-7-3365 2487. Phone: 61-7-3365 4945. E-mail: d.craik@imb.uq.edu.au.

[§] University of Queensland.

[1] Abbreviations: ICK, inhibitor cystine-knot; NMR, nuclear magnetic resonance; RP-HPLC, reversed-phase high-performance liquid chromatography; μ-Aga-I, μ-agatoxin I; ω-Aga-IVA, ω-agatoxin IVA; ω-Aga-IVB, ω-agatoxin IVB; δ-ACTXs, δ-atracotoxins; J-ACTXs, Janus-faced atracotoxins; ω-ACTXs, ω-atracotoxins; TFA, trifluoroacetic acid; DQF-COSY: double quantum filtered-correlation spectroscopy; TOCSY, total correlation spectroscopy; NOESY, nuclear Overhauser enhancement spectroscopy; ESI-MS, electrospray mass spectrometry.

10.1021/bi011932y CCC: $22.00
Published on Web 02/12/2002

work in their 36−37 amino acid sequence. They block voltage-gated calcium channels in insects but not mammals. The fourth family of peptides is the larger 68 amino acid mamba intestinal toxin-like atracotoxins (MIT-like ACTXs) (*12*) that possess a 10-cysteine framework and for which the activity remains unclear. Despite these distinct differences, the structures determined for the components from the Australian funnel-web spiders have conformed to the ICK motif, or a proposed simpler ancestral fold, the disulfide directed β-hairpin motif (DDH), as in the case of the J-ACTXs.

As part of an extensive study on the venom components of the Australian funnel-web spiders and their potential applications, we have discovered a novel 38 amino acid peptide (M_r = 4219 Da) containing an eight-cysteine/four-disulfide bond framework. On the basis of the naming convention introduced by Fletcher et al. (*11*), this molecule from the venom of a female specimen of the Australian funnel-web spider, *H. infensa* Orchid Beach, was termed ACTX-Hi:OB4219.

The sequence of this component has been determined using a combination of Edman degradation N-terminal amino acid sequencing and rapid amplification of cDNA ends (RACE) (*13*) techniques. On the basis of sequence comparison with other toxins, ACTX-Hi:OB4219 revealed an identical cysteine framework and loop sizes to the μ-agatoxins. The cysteine framework is highly conserved, but limited identity of other residues (maximum, six amino acids = 37% identity) is present. In this paper, we report the three-dimensional (3D) solution structure of ACTX-Hi:OB4219 as determined by NMR spectroscopy. In solution, the molecule contains two equally populated conformations resulting from cis/trans isomerization of one of its proline residues. This finding is consistent with the NMR spectra obtained for μ-Aga-IV, which revealed the presence of two significant conformations through cis/trans isomerization of Pro15. In that case, the spectral complexity prevented a full determination of both 3D structures, but for ACTX-Hi:OB4219, it was possible to determine the 3D structure of both conformations.

MATERIALS AND METHODS

Isolation of ACTX-Hi:OB4219. ACTX-Hi:OB4219 was isolated from the venom of the funnel-web *H. infensa* Orchid Beach. Preparative RP-HPLC on a Vydac C18 column with gradients of 0.1% aqueous TFA and 90% acetonitrile/0.09% TFA were employed in the purification. Over 70 individual milkings (50−70 mg) of female and juvenile *H. infensa* Orchid Beach specimens (female and juvenile venom was previously determined to be compositionally similar) provided ~3 mg of homogeneous native sample for NMR analysis. Samples prepared for NMR spectroscopy were dissolved in the following solvent systems: 80% H_2O/10% D_2O/10% CD_3CN at pH 5.0 and 90% D_2O/10% CD_3CN at pH 5.0. pH values were all measured at 298 K, and no correction for deuterium isotope effects was made.

RACE Sequence Determination of ACTX-Hi:OB4219. The venom glands of a live Australian funnel-web spider specimen were dissected and the mRNA extracted using a Quickprep Micro mRNA purification kit (Pharmacia Biotech, Cleveland, OH). Marathon adaptor double-stranded cDNA was obtained using a Marathon cDNA synthesis kit (Clontech Laboratories, Palo Alto, CA). PCR of the cDNA template was performed using an AmpliTaq Gold with GeneAmp kit (PerkinElmer, Shelton, CT), a gene specific primer (5-ttcaacATGAGGAATACTACC-3′), and a universal anchor primer (5′- AACTGGAAGAATTCGCGGCCGCAGGAA T-3′). Thermalcycling followed the cycle protocol; 1 cycle of 95 °C for 5 min, 35 cycles of 95 °C for 30 s, 55 °C for 1 min, 72 °C for 1.5 min, and 1 cycle of 72 °C for 10 min. PCR products were cleaned by agarose gel electrophoresis and then cloned into a pCR 2.1 vector and transformed into host INVαF′ cells using an Invitrogen TA cloning kit (Bresatec, Australia). Clones containing the correct insert size were sequenced by the AGRF (Brisbane, Australia).

NMR Spectroscopy. Spectra were recorded on a Bruker ARX 500 and a Bruker DMX 750 spectrometer with sample temperatures of 293, 298, and 305 K. The carrier frequency was, in all experiments, set at the center of the spectra, on the water resonance frequency, and quadrature detection was used. All spectra were acquired in phase-sensitive mode using TPPI (*14*). Exclusively homonuclear spectra were recorded, including double quantum filtered DQF-COSY (*15*); TOCSY (*16*) with MLEV17 (*17*), isotropic mixing period of 80 ms; ECOSY (*18*); and NOESY (*19*), with mixing times of 100 ms, 200 ms, and 250 ms. The water proton signal was, in the DQF-COSY experiment, suppressed by low power irradiation during the relaxation delay (1.8 s). For the TOCSY and NOESY experiments, water suppression was achieved using a modified WATERGATE (*20*) sequence in which two gradient pulses of 1 ms duration and 6 G cm^{-1} strength were applied on either side of the binomial pulse. All two-dimensional (2D) spectra were collected over 4096 data points in the f2 dimension and 512 or 600 increments in the f1 dimension over a spectral width corresponding to 12 ppm. For identification of slowly exchanging amides, a series of one-dimensional (1D) and TOCSY spectra were run immediately after dissolving the sample in D_2O.

All spectra were processed on a Silicon Graphics workstation using XWINNMR (Bruker). The f1 dimension was generally zero-filled to 2048 real data points with the f1 and f2 dimensions being multiplied by a sine-squared function shifted by 90° prior to Fourier transformation. Spectra were referenced to DSS. Processed spectra were analyzed and assigned using the program XEASY (*21*). Spectra were assigned using the sequential assignment technique (*22*), with particular care required to deal with the two coexisting conformers. The process was facilitated, in part, using the automatic assignment program NOAH, which is a part of the DYANA (*23*) package.

Structure Calculations. Cross-peaks in NOESY spectra recorded in 80% H_2O/10% D_2O/10% CD_3CN with a mixing time of 200 ms were integrated and calibrated in XEASY (*21*), and distance constraints were derived using DYANA (*23*). Corrections for pseudo-atoms were added to distance constraints where needed (*24*). Backbone dihedral angle restraints were derived from $^3J_{HN\alpha}$ coupling constants measured from line shape analysis of antiphase cross-peak splitting in the DQF-COSY spectrum. Angles were restrained to $-120° \pm 40$ for $^3J_{HN\alpha} > 9$ Hz and to $-60° \pm 30$ for $^3J_{HN\alpha} < 5$ Hz. For two residues, Ala6 (both conformations) and Phe27 (cis conformation), the ϕ angle was restrained to $60° \pm 30$, based on a coupling constant of around 6 Hz together with a very strong intraresidual dαN(i,i) NOE. For

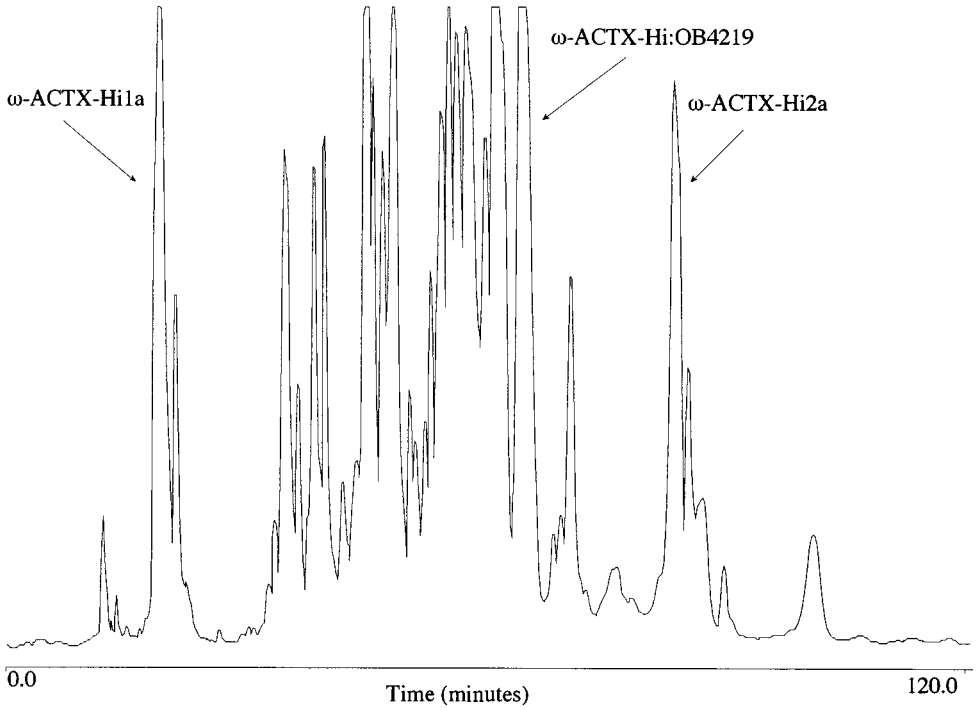

FIGURE 1: Preparative RP-HPLC of crude pooled female and juvenile *H. infensa sp*. venom, illustrating the location of ACTX-Hi:OB4219, *ω*-ACTX-Hi1a (M_r = 3922 Da), and *ω*-ACTX-Hi2a (M_r = 4009 Da) (0.5% gradient, 0−60% MeCN/0.1% TFA in 120 min, UV absorbance 214 nm).

both residues, this was consistent with preliminary structure calculations. Additional ϕ angle restraints of $-100° \pm 80$ were applied where the intraresidue dαN(i,i) NOE was clearly weaker than the sequential dαN(i,i+1) NOE.

Stereospecific assignments of β-methylene protons and χ1 dihedral angles were derived from $^3J_{\alpha\beta}$ coupling constants, measured from ECOSY spectra, in combination with NOE peak intensities (*25*). The χ1 angles were restrained to $60° \pm 30$, $180° \pm 30$, or $-60° \pm 30$ for 15 residues, including six of the eight cysteine residues. Slowly exchanging amide protons identified by D_2O exchange experiments were analyzed in preliminary structures. In cases where hydrogen bonds could be determined and unambiguously assigned, these were introduced as distance restraints for following calculations. Preliminary structures were calculated using a torsion angle simulated annealing protocol within DYANA. Final structures were calculated using simulated annealing and energy minimization protocols within CNS version 1.0 (*26*). The starting structures were generated using random (ϕ,ι) dihedral angles and energy-minimized to produce structures with the correct local geometry. A set of 50 structures was generated by a torsion angle simulated annealing protocol (*27*, *28*). This protocol involves a high-temperature phase comprising 4000 steps of 0.015 ps of torsion angle dynamics, a cooling phase with 4000 steps of 0.015 ps of torsion angle dynamics during which the temperature is lowered to 0 K, and finally an energy minimization phase comprising 5000 steps of Powell minimization. Structures consistent with restraints were subjected to further molecular dynamics and energy minimization in a water shell, as described by Linge and Nilges (*29*). The refinement in explicit water involves the following steps. First, heating to 500 K via steps of 100 K, each comprising 50 steps of 0.005 ps of Cartesian dynamics. Second, 2500 steps of 0.005 ps of Cartesian dynamics at 500 K before a cooling phase where the temperature is lowered in steps of 100 K, each comprising 2500 steps of 0.005 ps of Cartesian dynamics. Finally, the structures were minimized with 2000 steps of Powell minimization. Identical protocols were used for both the cis and the trans conformer. Coordinates for both conformers have been deposited in the PDB (ID code cis, 1KQH; trans, 1KQI).

RESULTS

Isolation and Characterization. The venom of the *H. infensa* Orchid Beach specimen (collected from Orchid Beach, Fraser Island, QLD, Australia) is highly complex, demonstrating the presence of more than 70 components by analytical RP-HPLC (Figure 1) and RP-HPLC/ESI-MS. The peak corresponding to ACTX-Hi:OB4219 was initially identified from RP-HPLC/ESI-MS analysis of crude venom from a female *H. infensa* Orchid Beach specimen, and further analysis revealed it to be a major component in the venom of these spiders. Collection of the peak from analytical RP-HPLC revealed the presence of a relatively homogeneous sample with only a small amount of impurity. The cysteine residues were reduced and alkylated with maleimide and the reduced/alkylated sample purified by RP-HPLC. Automated N-terminal amino acid sequencing of the isolated reduced/alkylated sample revealed a 30 amino acid partial sequence of a novel peptide.

The full amino acid sequence of ACTX-Hi:OB4219 was determined by RACE technology using a gene specific primer designed around the translation start site. The ACTX-Hi:OB4219 gene product was identified by 3′ RACE using a cDNA template library constructed from the venom glands of a female *H. infensa* Orchid Beach specimen. Subsequent cloning of the cDNA into a vector and transformation into host cells allowed for DNA sequencing of the gene,

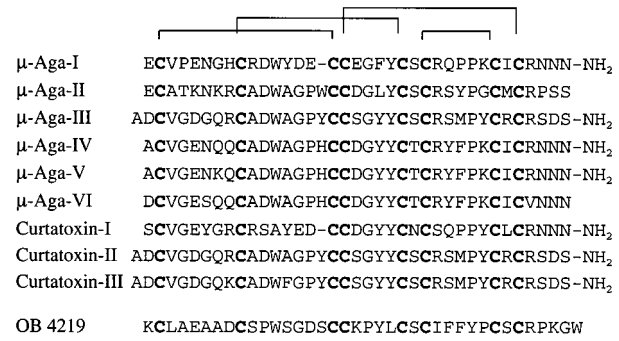

FIGURE 2: Sequence comparison of ACTX-Hi:OB4219 with the μ-agatoxins and curtatoxins. The latter peptides occur in the venom of the American funnel-web spiders, *Agelenopsis aperta* and *Hololena curta*.

determining the translated leader and mature peptide sequences. Sequence comparison with the N-terminal amino acid sequence predicted the cleavage position of the mature peptide from the leader peptide. The experimental mass of the native peptide was consistent with the RACE derived sequence. Figure 2 shows the sequence of ACTX-Hi:OB4219 and, for comparison, those of the μ-agatoxins and curtatoxins.

NMR Assignments. Initial 1D ^{1}H NMR spectra recorded on ACTX-Hi:OB4219 in aqueous solution had poor signal intensity and large line widths, which is most likely the result of sample aggregation. However, spectra recorded in the presence of >10% CD_3CN produced line widths consistent with that of a peptide of the known molecular weight. Therefore, all spectral data, including TOCSY, NOESY, DQF-COSY, and ECOSY, were recorded in either 80% H_2O/10% D_2O/10% CD_3CN or 90%D_2O/10% CD_3CN. Spectra recorded at 298 K were primarily used for assignments, while spectra recorded at other temperatures were used to resolve ambiguities.

Initial inspection and preliminary assignment of the spectral data revealed the presence of far more signals than expected. This suggested that, despite the presence of a single peak in the RP-HPLC chromatogram, ACTX-Hi:OB4219 adopts more than one conformation in solution. Several of the corresponding protons from the different conformations appeared as "brothered" resonances with similar patterns of NOEs but slightly different chemical shifts. Some signals were widely separated, demonstrating that the respective protons clearly experience different environments and NOE interactions. The presence of the two sets of signals resulted in complicated spectra that made the assignment very difficult. Consequently, the initial attempts to assign spectra recorded at 500 MHz were only partially successful. However, after collection and analysis of spectra recorded at 750 MHz, a complete assignment of both conformations was possible. The two conformations were subsequently shown to result from cis/trans isomerization of the peptide bond preceding Pro30. Two of the remaining three prolines (Pro11 and Pro35) display strong $d_{\alpha i \rightarrow \delta i+1}$ cross-peaks in the D_2O NOESY spectra, showing that they are in the trans conformation, while Pro20 displays a strong $d_{\alpha i \rightarrow \alpha i+1}$ cross-peak, consistent with a cis conformation. The ^{1}H chemical shift assignments for both conformers of ACTX-Hi:OB4219 are provided as Supporting Information.

Structure Determination and Analysis. Careful analysis of the 200 ms NOESY spectrum (750 MHz, 298 K) using the program XEASY allowed for the assignment of each spin system to a specific amino acid in either the cis or the trans conformation. All nonintraresidual peaks within each conformation were then subsequently assigned both manually and with the use of the NOAH automatic assignment algorithm within the DYANA program package. Interproton distance restraints were derived from the NOE intensities and were used in structure calculations using a torsion angle simulated annealing protocol within DYANA. An iterative strategy in which preliminary structures were analyzed to resolve spectral ambiguities to allow for the introduction of new restraints was used and resulted in a final set of restraints consisting of 411 and 442 inter-residual distances for the cis and trans conformations, respectively.

These restraints, from which redundancies based on the covalent geometry had been eliminated, included 239 and 252 sequential, 37 and 44 medium-range, and 135 and 146 long-range distances, respectively, for the two conformations. In addition, a set of 28 backbone angle restraints for each conformation and 14 (cis conformation) and 13 (trans conformation) side-chain angle restraints were derived on the basis of coupling constants and short mixing time NOE intensities. Furthermore, 18 restraints for 9 hydrogen bonds were derived from the preliminary structures. In the final round of structure calculations, these restraints were used to calculate a family of 50 structures for each conformer, using a torsion angle simulated annealing protocol (*27*, *28*) within CNS, v.1.0 (*26*). Structures consistent with restraints were subjected to further molecular dynamics and energy minimization in a water shell (*29*).

Description of the 3D Structure. (1) Secondary Structure. A summary of the information on short- and medium-range NOEs, coupling constants, chemical shift index, and slowly exchanging amide protons is presented for both conformations in Figure 3. These data, in conjunction with information about long-range NOEs, clearly show that ACTX-Hi:OB4219 has an absence of helical regions and the presence of regions of extended β-type structure. The principal element of secondary structure is a triple-stranded antiparallel β sheet comprising three short strands involving residues 6−10, 21−25, and 31−35. The arrangement of the three strands, together with identified NOEs and proposed hydrogen bonds, based on slowly exchanging amide protons, is presented in Figure 4.

(2) Disulfide Connectivity. The amino acid sequence of ACTX-Hi:OB4219 includes eight-cysteine residues involved in four disulfide bonds. The connectivities of the disulfide bonds were unknown but were initially assumed to be identical to μ-Aga-I on the basis of sequence framework similarities. This was supported by the experimental data, which included NOEs diagnostic for disulfide bonds between all of the cysteine residues that were proposed to be connected (*30*). Early in the analysis, it was clear that the only possible difference from the assumed disulfide connectivity could be the alternative pairing, 2−17, 18−33, instead of 2−18, 17−33. However, this possibility was excluded by structure calculations that included the alternative pairing; the resultant structures clearly violated the experimental data. The disulfide pairing was therefore unambiguously determined to be 2−18, 9−23, 17−33, 25−31 and is identical to that in μ-Aga-I. The protein is folded in such a way that the disulfide bond between Cys17 and

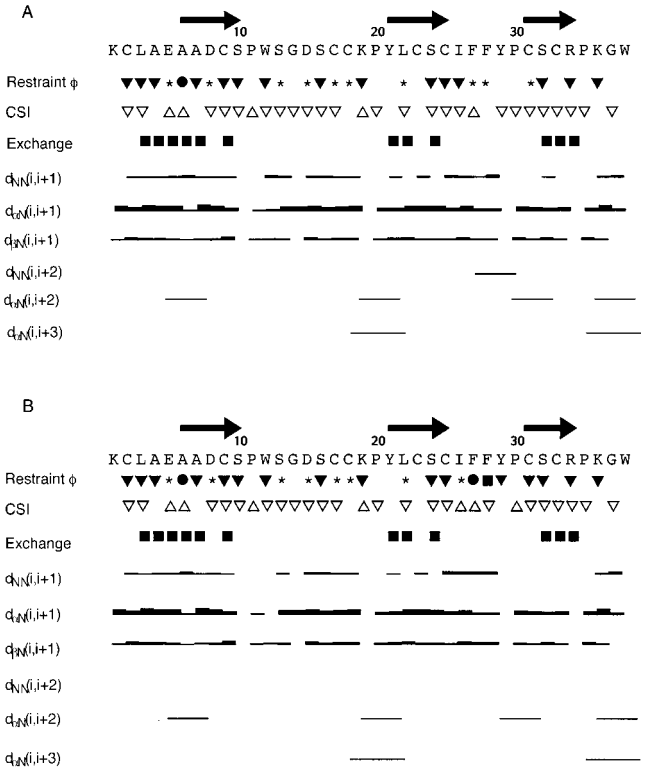

FIGURE 3: Summary of the short- and medium-range NOE interactions, chemical shift index (CSI), slowly exchanging amide protons, and coupling constants for the cis (A) and trans (B) conformations of ACTX-Hi:OB4219. Filled circles (●) indicate a ϕ angle of 60° ± 30°, stars (★) represent a ϕ angle of −100° ± 80°, and filled triangles (▼) indicate a ϕ angle of −120° ± 40°. Empty upside down triangles (▽) represent a CSI of −1, and empty triangles (△) indicate a CSI of +1. Filled squares (■) illustrate amide protons still visible 24 h after exchange into D_2O. Black bars represent NOEs present in the 200 ms NOESY spectrum (293 K, pH 5.0). The thickness of the bars corresponds to the NOE intensity. Arrows indicate β strands in the calculated structures.

Cys33 penetrates a ring formed by the backbone of residues 2−9 together with the disulfide bond between Cys9 and Cys23, the backbone of residues 18−23, and the disulfide bond between Cys2 and Cys18. A similar motif has been identified in many classes of peptides, including μ-Aga-I, and is commonly referred to as a cystine knot (*7*, *31*).

A family of 20 lowest energy structures consistent with the experimental data of each conformer was chosen from the final set of 50 structures to represent the solution structure of each conformer of ACTX-Hi:OB4219. The statistics for the two families are given in Table 1. The structures have no distance violations greater than 0.2 Å and no dihedral violations greater than 3.0°. Further, they have good covalent geometry, indicated by small deviations from ideal bond lengths and bond angles and favorable nonbonded contacts as indicated by low values of Lennard-Jones potential.

The structures are well-defined with the exception of the N-terminus and the C-terminus of the cis conformation. Some flexibility is also seen in the loop including residues 11−15, but that is most likely due to a lack of restraints in this region as a result of overlap. The mean pairwise rmsd for the backbone atoms over the whole molecule of the cis conformation is 0.98 ± 0.25 Å and 1.77 ± 0.41 Å for the heavy atoms. The rmsd for the well-defined regions, residues 6−9, 21−25, and 31−35, is 0.27 ± 0.07 Å for the backbone

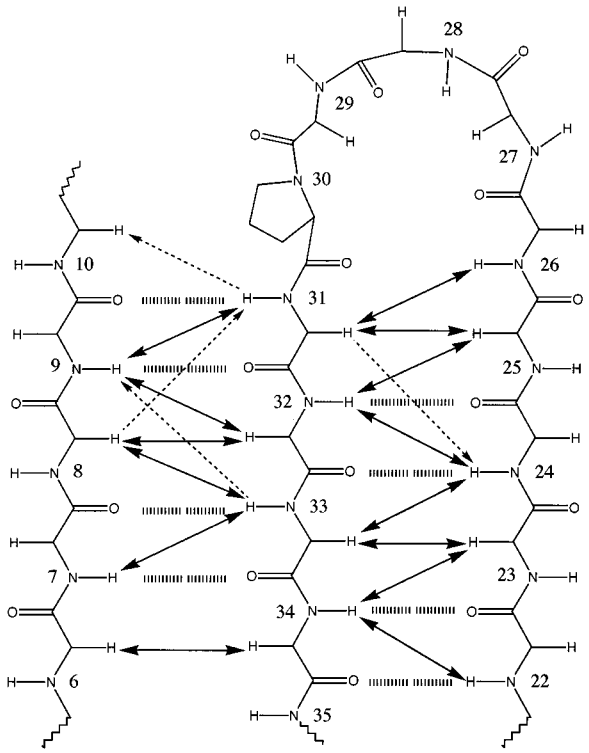

FIGURE 4: Schematic diagram of the triple-stranded antiparallel β sheet and the residue 27−30 β turn present in the ACTX-Hi:OB4219 structure. Arrows represent observed NOEs between protons in the sheet. Dashed arrows represent NOEs that could be present but that are not included in structure calculations because of overlap. Dashed lines correspond to hydrogen bonds predicted by the program MOLMOL (*25*) on the basis of calculated structures.

Table 1: Structural and Energetic Statistics for the Final Family of 20 ACTX-Hi:OB4219 Structures

	cis	trans
Mean RMS Deviations from Experimental Restraints		
NOE (Å)	0.024 ± 0.001	0.024 ± 0.002
dihedral angles (deg)	0.58 ± 0.09	0.57 ± 0.09
Mean RMS Deviations from Idealized Covalent Geometry		
bonds (Å)	0.0040 ± 0.0002	0.0039 ± 0.0002
angles (deg)	0.57 ± 0.04	0.56 ± 0.03
impropers (deg)	0.42 ± 0.04	0.43 ± 0.03
Restraint Violations		
NOE violations > 0.2 Å	0	0
maximum NOE violation (Å)	0.20	0.20
dihedral violations > 3°	0	0
maximum dihedral violation (deg)	3.0	3.0
Mean Energies (kJ mol^{-1})		
*E*noe	13.8 ± 1.7	15.1 ± 2.2
*E*cdih	1.4 ± 0.45	1.25 ± 0.44
*E*vdw	−93.3 ± 5.0	−103 ± 5.0
*E*bond + *E*angle + *E*improper	68.9 ± 11.1	65.8 ± 6.4
*E*elec	−1334 ± 38	−1256 ± 23
*E*total	−1192 ± 28	−1121 ± 27
Pairwise RMSD		
backbone atoms (N, Cα, C; residues 1−38) (Å)	0.98 ± 0.25	1.01 ± 0.30
heavy atoms (Å)	1.77 ± 0.41	1.88 ± 0.41
backbone atoms, β sheet (residues: 6−9,21−25,31−35) (Å)	0.27 ± 0.07	0.35 ± 0.11
heavy atoms (Å)	0.74 ± 0.17	0.86 ± 0.11

atoms and 0.74 ± 0.17 Å for the heavy atoms of the cis conformation. The corresponding values over the entire molecule for the trans conformation are 1.01 ± 0.30 Å for the backbone atoms and 1.88 ± 0.41 Å for the heavy atoms. The corresponding rmsd values for the well-defined regions of the trans conformation are 0.35 ± 0.11 Å for the backbone

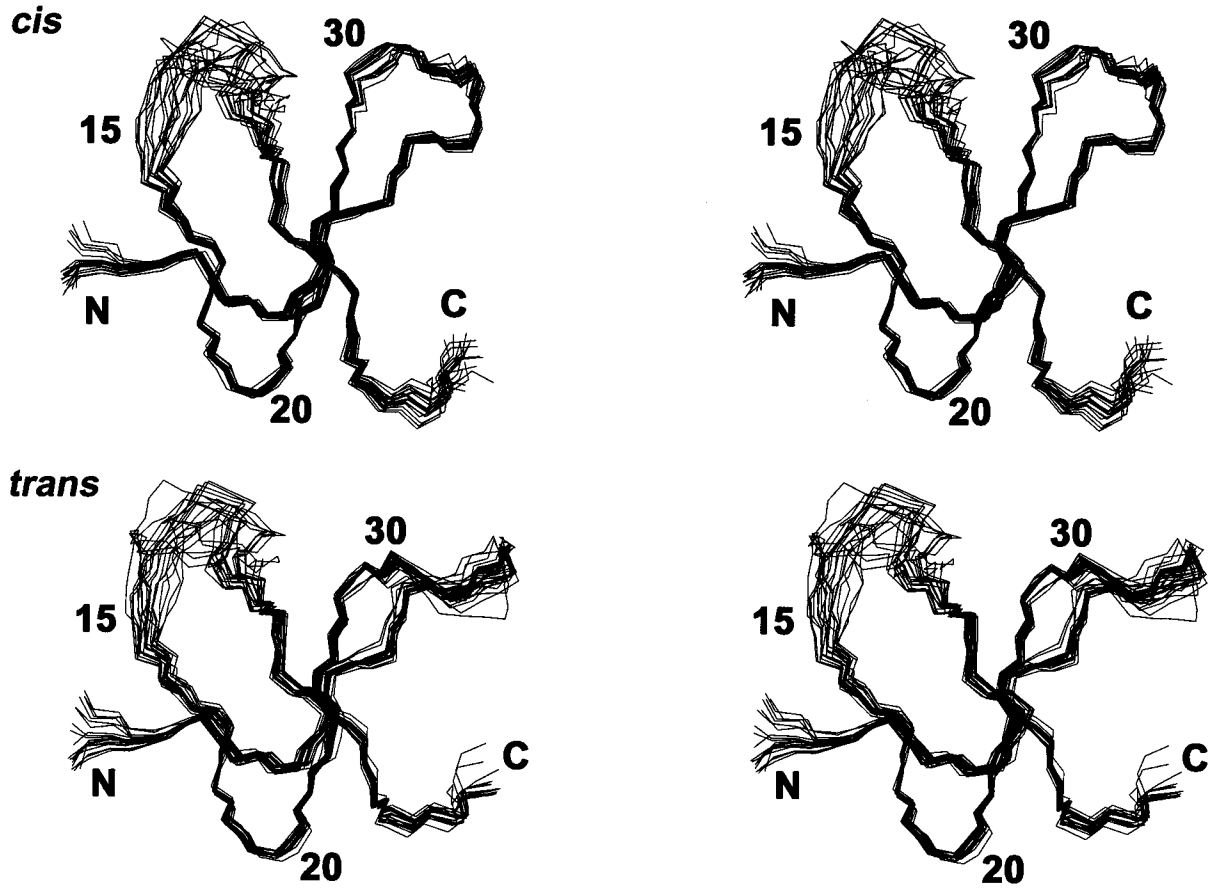

FIGURE 5: Stereoview of backbone superimpositions of the 20 lowest energy structures for the cis (upper) and trans (lower) conformations of ACTX-Hi:OB4219. N- and C-termini and selected residues are labeled.

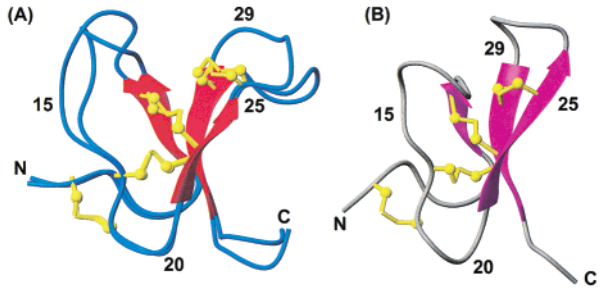

FIGURE 6: Comparison of ACTX-Hi:OB4219 cis and trans conformations (view A) and μAga I (view B). The β sheets are represented as arrows in red and magenta. Disulfide bonds are shown in yellow in ball-and-stick representation. N- and C-termini and selected residues are labeled.

atoms and 0.86 ± 0.11 Å for the heavy atoms. Stereoviews of the two families of structures superimposed over the backbone atoms are shown in Figure 5. A ribbon representation of the structures showing the disulfide bonds is given in Figure 6A. Analysis of the angles shows that ~79% of the residues in both conformations lie in the most favorable regions of the Ramachandran plot and that the remaining 21% all lie in the additionally allowed region.

The solution structure of ACTX-Hi:OB4219 was determined to be very flat, with the core comprising mainly cysteine residues and a fold primarily stabilized by disulfide bonds. The poor definition of the N-terminus is most likely associated with flexibility of this region in solution, as is commonly observed in protein structures. The disulfide bond between residues Cys2 and Cys18, which is in the right-handed hook (RHH) conformation (*32*, *33*), ties the N-terminus to the loop involving residues 11−20. The linker between this loop and the N-terminus is the first β strand, strand 1, involving residues 6−10. This strand is linked to the second β strand, which includes residues 21−25, by the disulfide bond between Cys9 and Cys23, which is in a short right-handed hook (SRH) conformation. The 11−20 loop is further tied to the β sheet by the disulfide bond between Cys17 and Cys33. This disulfide bond adopts a left-handed spiral (LHS) conformation. A turn of no definite classification is formed in both conformations by residues 26−30. This turn includes the Pro30 involved in cis/trans isomerization and is where the major differences between the two conformations occur. The turn serves as a link between strand 2 and the last β strand, strand 3, which includes residues 31−35 and it is stabilized by the disulfide bond between Cys25 and Cys31, which is in the short right-handed hook (SRH) conformation (*32*, *33*). The C-terminus is slightly disordered for the cis conformation, while, in the trans conformations, it appears to be interacting with the turn between residues 26−30. This interaction can be identified by a large number of NOEs between the Trp38 aromatic protons and the Ile26 side chain.

DISCUSSION

In the current study, we have determined the 3D structure of two solution conformers of the Australian funnel-web spider toxin ACTX-Hi:OB4219. The two conformers derive from a single isomer purified by RP-HPLC. In solution, an

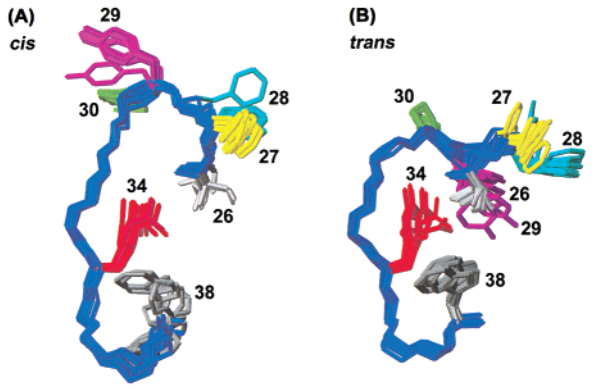

FIGURE 7: Comparison of the 10 lowest energy structures of the cis (A) and the trans (B) C-terminal part of ACTX-Hi:OB4219. The side chains of Ile26, Phe27, Phe28, Tyr29, Pro30, and Trp38 are shown in white, yellow, blue, magenta, green, and gray, respectively.

equal population of each conformer is produced by cis/trans isomerization of Pro30.

Though the presence of multiple conformations is not uncommon among small disulfide-rich peptides, there are few examples where complete spectral assignments have been possible for mixtures of two or more conformers and have led to the successful calculation of the separate structures (*34*). One example of a peptide showing multiple conformations in solution is the ACTX-Hi:OB4219 homologue μ-Aga-IV, isolated from the venom of the unrelated American funnel-web spider, *Agelenopsis aperta*. Attempts to determine the solution structure of μ-Aga-IV were made by Omecinsky et al. but were unsuccessful due to the complexity of the spectral data (*6*), requiring the structures to be modeled instead.

Some regions of the spectra of ACTX-Hi:OB4219 are not affected by the isomerization of Pro30. However, this does not simplify the assignment procedure. Instead, it introduces the significant problem of linking regions of degenerate signals separated by regions of nondegenerate signals. This problem was solved by the identification of a number of key long-range NOE connectivities between the different regions of nondegenerate signals for the two conformations. The amino acids affected by the isomerization are primarily clustered in the β-sheet regions and C-terminal region of the molecule.

The major differences in the structures of the cis and trans isomers are presented in Figure 7, highlighting the important side chains in the C-terminal region of the molecule. It is clear that the major change in the structure is the conformation and orientation of the turn involving Pro30. In the trans case, this turn adopts a conformation that leads to hydrophobic interactions between Ile26 and Trp38. These interactions give rise to a large number of NOEs that are specific for the trans isomer. The turn 27−30 is highly hydrophobic and incorporates residues Phe27, Phe28, Tyr29, and Pro30. The orientation of the side chains of these aromatic residues, especially Tyr29, is affected by cis/trans isomerization and may result in some of the chemical shift differences observed. It is well-known that electron rich aromatic rings often affect the shielding of surrounding protons. The greatest separation in chemical shifts between the two conformers is observed for residues Ile26 and Cys31. For Ile26, the differences in chemical shifts between the two conformations are 0.406 and 0.88 ppm for the Hα and HN protons respectively, while, for Cys31, the differences are 0.041 and 0.499 ppm, respectively.

The solution structure of ACTX-Hi:OB4219, illustrated in Figure 6, shows that the peptide possesses a very similar fold to that of μ-Aga-I. Both are characterized by an inhibitor cystine-knot (ICK) motif that has previously been found in various peptides from phylogenetically different sources, including cone snails, spiders, plants, and fungi (*31*). Although a range of activities have been reported for peptides incorporating the ICK motif, one principal target is the voltage-gated ion channels. The ICK motif incorporates a triple-stranded antiparallel β sheet, which is stabilized by three disulfide bonds. Two of the disulfide bonds are orientated such that they form a ring together with the backbone of the intervening amino acids. The third disulfide bond penetrates through this ring. Members of this family have the consensus sequence $CX_{3-7}CX_{3-6}CX_{0-7}CX_{1-4}CX_{4-13}C$, where X can be any residue (*31*). This framework appears ideal for presenting a number of different functional groups and provides a likely explanation for the diverse range of targets observed for the members of this family. The differences in overall folding and activity of the members are determined by the number of residues in the loops as well as the differences in primary sequence. In ACTX-Hi:OB4219, the sequence can be written as $CX_6CX_7CX_0CX_4CX_1CX_9C$, while μ-Aga-I has only six residues in the second loop. Both ACTX-Hi:OB4219 and μ-Aga-I have an additional disulfide bond within the last loop. This disulfide forms a bridge between strand 2 and strand 3, which defines the end of the sheet and provides further stabilization to the turn involving residues 27−30.

When comparing the two structures of ACTX-Hi:OB4219 and μ-Aga-I, it is clear that the loop between residues 8−16 is the least well-defined. For μ-Aga-I, a broadening of the resonances in this region has been observed (*6*), indicating that there might be some flexibility. However, no such broadening occurs in ACTX-Hi:OB4219, and it has been reported that no broadening can be detected for these resonances in spectra of μ-Aga-IV. These differences in line widths may be caused by the fact that this loop is one residue shorter in μ-Aga-I than in μ-Aga-IV and ACTX-Hi:OB4219, which may cause a change in the rigidity of this loop. On the basis of our data, we assume that the poor definition in this region is more a result of lack of restraints rather than structural flexibility.

No data on the biological activity of ACTX-Hi:OB4219 are available, but the fact that it is one of the major components of the venom from an insectivorous spider suggests an insecticidal activity. This is supported by our structural data, which shows that ACTX-Hi:OB4219 is structurally similar to μ-Aga-I and μ-Aga-IV, both of which have been shown to target insect sodium channels. The structure also shows that ACTX-Hi:OB4219 incorporates an ICK motif, which is a strong indication that it possesses activity on ion channels. However, the sequential differences between ACTX-Hi:OB4219 and the μAga toxins are significant, and because peptides with similar structure have been shown to be active on several different ion channels, it is difficult to speculate which channel ACTX-Hi:OB4219 may target.

ACKNOWLEDGMENT

D.J.C. is an Australian Research Council Professorial Fellow.

SUPPORTING INFORMATION AVAILABLE

Two tables of ^{1}H chemical shifts for the cis and trans conformations of ACTX-Hi:OB4219, a figure showing the 200 ms NOESY spectrum of ACTX-Hi:OB4219, and a schematic diagram demonstrating the regions of ACTX-Hi:OB4219 that show separate resonances for the two conformations. This material is available free of charge via the Internet at http://pubs.acs.org.

NOTE ADDED AFTER ASAP POSTING

This article was released ASAP 02/12/02 before the Supporting Information paragraph had been added. The complete version was posted 03/05/02.

REFERENCES

1. Adams, M. E., Bindokas, V. P., Hasegawa, L., and Venema, V. J. (1990) *J. Biol. Chem. 265*, 861−867.
2. Adams, M. E., Herold, E. E., and Venema, V. J. (1989) *J. Comp. Physiol. A 164*, 333−342.
3. Quistad, G. B., Reuter, C. C., Skinner, W. S., Dennis, P. A., Suwanrumpha, S., and Fu, E. W. (1991) *Toxicon 29*, 329−336.
4. Skinner, W. S., Adams, M. E., Quistad, G. B., Kataoka, H., Cesarin, B. J., Enderlin, F. E., and Schooley, D. A. (1989) *J. Biol. Chem. 264*, 2150−2155.
5. Stapleton, A., Blankenship, D. T., Ackermann, B. L., Chen, T. M., Gorder, G. W., Manley, G. D., Palfreyman, M. G., Coutant, J. E., and Cardin, A. D. (1990) *J. Biol. Chem. 265*, 2054−2059.
6. Omecinsky, D. O., Holub, K. E., Adams, M. E., and Reily, M. D. (1996) *Biochemistry 35*, 2836−2844.
7. Pallaghy, P. K., Nielsen, K. J., Craik, D. J., and Norton, R. S. (1994) *Protein Sci. 3*, 1833−1839.
8. Sheumack, D. D., Claassens, R., Whiteley, N. M., and Howden, M. E. H. (1985) *FEBS Lett. 181*, 154−156.
9. Wang, X., Connor, M., Smith, R., Maciejewski, M. W., Howden, M. E., Nicholson, G. M., Christie, M. J., and King, G. F. (2000) *Nat. Struct. Biol. 7*, 505−513.
10. Atkinson, R. K., Tyler, M. I., and Vonarx, E. J. (1993) Patent No. WO 93/15108.
11. Fletcher, J. I., Smith, R., O'Donoghue, S. I., Nilges, M., Connor, M., Howden, M. E. H., Christie, M. J., and King, G. F. (1997) *Nat. Struct. Biol. 4*, 559−566.
12. Szeto, T. H., Wang, X. H., Smith, R., Connor, M., Christie, M. J., Nicholson, G. M., and King, G. F. (2000) *Toxicon 38*, 429−442.
13. Frohman, M. A., Dush, M. K., and Martin, G. R. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 8998−9002.
14. Marion, D., and Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967−974.
15. Rance, M., Sørensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., and Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 117*, 479−485.
16. Braunschweiler, L., and Ernst, R. R. (1983) *J. Magn. Reson. 53*, 521−528.
17. Bax, A., and Davis, D. G. (1985) *J. Magn. Reson. 65*, 355−360.
18. Griesinger, C., Sørensen, O. W., and Ernst, R. R. (1987) *J. Magn. Reson. 75*, 474−492.
19. Jeener, J., Meier, B. H., Bachmann, P., and Ernst, R. R. (1979) *J. Chem. Phys. 71*, 4546−4553.
20. Piotto, M., Saudek, V., and Sklenar, V. (1992) *J. Biomol. NMR 2*, 661−665.
21. Eccles, C., Güntert, P., Billeter, M., and Wüthrich, K. (1991) *J. Biomol. NMR 1*, 111−130.
22. Wüthrich, K. (1986) *NMR of proteins and nucleic acids*, Wiley-Interscience, New York.
23. Güntert, P., Mumenthaler, C., and Wüthrich, K. (1997) *J. Mol. Biol. 273*, 283−298.
24. Wüthrich, K., Billeter, M., and Braun, W. (1983) *J. Mol. Biol. 169*, 949−961.
25. Koradi, R., Billeter, M., and Wüthrich, K. (1996) *J. Mol. Graphics 14*, 51−55, 29−32.
26. Brünger, A. T., Adams, P. D., and Rice, L. M. (1997) *Structure 5*, 325−336.
27. Rice, L. M., and Brünger, A. T. (1994) *Proteins 19*, 277−290.
28. Stein, E. G., Rice, L. M., and Brünger, A. T. (1997) *J. Magn. Reson. 124*, 154−164.
29. Linge, J. P., and Nilges, M. (1999) *J. Biomol. NMR 13*, 51−59.
30. Klaus, W., Broger, C., Gerber, P., and Senn, H. (1993) *J. Mol. Biol. 232*, 897−906.
31. Craik, D. J., Daly, N. L., and Waine, C. (2001) *Toxicon 39*, 43−60.
32. Hutchinson, E. G., and Thornton, J. M. (1996) *Protein Sci. 5*, 212−220.
33. Richardson, J. S. (1981) *Adv. Protein Chem. 34*, 167−339.
34. Kessler, H., Konat, R. K., and Schmitt, W. (1996) in *Chapter 6: Conformational analysis of peptides: Application to drug design* (Craik, D. J., Ed.), CRC Press, Boca Raton, FL.

BI011932Y