Three-Dimensional Structure in Solution of the Polypeptide Cardiac Stimulant Anthopleurin-A[†,‡]

Paul K. Pallaghy, Martin J. Scanlon, Stephen A. Monks, and Raymond S. Norton*

NMR Laboratory, Biomolecular Research Institute, 381 Royal Parade, Parkville 3052, Australia

Received October 28, 1994; Revised Manuscript Received December 22, 1994[⊗]

ABSTRACT: The three-dimensional structure in aqueous solution of the 49-residue polypeptide anthopleurin-A (AP-A), from the sea anemone *Anthopleura xanthogrammica*, has been determined from ^{1}H NMR data. A restraint set consisting of 411 interproton distance restraints inferred from NOEs and 19 backbone and 13 side chain dihedral angle restraints from spin−spin coupling constants, as well as 15 lower bound restraints based on the absence of NOEs in the spectra, was used as input for distance geometry calculations in DIANA and simulated annealing and restrained energy minimization in X-PLOR. Stereospecific assignments for 12 β-methylene pairs were also included. The final set of 20 structures had mean pairwise rms differences over the whole molecule of 2.04 Å for the backbone heavy atoms (N, C^α, and C) and 2.59 Å for all heavy atoms. For the well-defined region encompassing residues 2−7 and 17−49, the corresponding values were 0.82 and 1.27 Å, respectively. AP-A adopts a compact structure consisting of four short strands of antiparallel β-sheet (residues 2−4, 20−23, 34−37, and 45−48) connected by three loops. The first loop commences with a type I β-turn which includes two important Asp residues; this loop is the least well-defined region of the protein, although a β-turn involving residues 13−16 is observed in nearly half the structures. The loop linking the second and third strands is constrained by the 29−47 disulfide bond and contains two well-defined β-turns, while the third loop contains the Gly40-Pro41 sequence, which has been identifed previously as the site of *cis*−*trans* isomerism. The carboxylate group of Asp7 is close to the ϵ-ammonium group of Lys37, suggesting that they may form a salt bridge. A pH titration monitored by 2D NMR supports this by showing that Asp7 has a low pK_a. It is proposed that this region of the molecule and the nearby residues Asp9 and His39 form part of the molecular surface which interacts with the mammalian cardiac sodium channel.

Anthopleurin-A (AP-A),[1] from the northern Pacific sea anemone *Anthopleura xanthogrammica*, is a polypeptide of 49 residues, cross-linked by three disulfide bonds (Tanaka et al., 1977; Norton, 1981). AP-A is active as a cardiac stimulant at nanomolar concentrations *in vitro*, making it some 200-fold more potent on a molar basis than the cardiac glycoside digoxin. Its positive inotropic activity is not associated with any significant effects on heart rate or blood pressure (Blair et al., 1978), and in conscious dogs its therapeutic index is 7.5, which is about 3-fold higher than that of digoxin (Scriabine et al., 1979). AP-A is active under conditions of stress and hypocalcaemia (Norton, 1981; Kodama et al., 1981), as well as in ischaemic cardiac muscle, where many other positive inotropes give equivocal results (Gross et al., 1985). The mechanism of action of AP-A involves a delayed inactivation of the myocardial voltage-gated Na^+ channel (Norton, 1991).

The closely-related polypeptide *Anemonia sulcata* toxin II (ATX II) has been found to have class III antiarrhythmic activity, indicating its potential in the management of cardiac arrhythmias (Platou et al., 1986). A positive inotropic drug which was also effective as an antiarrhythmic could offer significant advantages therapeutically (Vaughan Williams, 1975; Sasayama, 1992).

Being polypeptides, the sea anemone toxins have limited therapeutic potential in their own right, but they may be useful leads in the development of new positive inotropic agents. In this context, a knowledge of their three-dimensional structures is essential. A low-resolution structure for AP-A in solution has been determined previously from NMR data (Torda et al., 1988), which showed that the molecule consisted of a four-stranded, antiparallel β-sheet linked by three loops, two of which were well-defined by the NMR data and the third poorly defined, possibly due to conformational mobility. This poorly defined loop (from Asp7 to Thr17) contains at least some of the residues considered to be essential for activity (Norton, 1991), including Asp7 and/or Asp9 (Newcomb et al., 1980; Barhanin et al., 1981; Gruen & Norton, 1985). Furthermore, tryptic cleavage of this loop adjacent to Arg14 destroyed the positive inotropic activity of AP-A (Gould et al., 1990), indicating that the integrity of this loop is necessary for activity. One goal of the present study, therefore, was to define the structure of this loop more precisely than in the previously determined structures.

The structures in solution of two related sea anemone polypeptides, ATX Ia (Widmer et al., 1989) and Sh I (Fogh

[†]This work was supported in part by a grant from the Australian Research Council (R.S.N.) and an Australian Postgraduate Research (Industry) Award (S.A.M.).

[‡] Coordinates of all final structures and the NMR restraints used in their determination have been deposited at the Brookhaven Protein Data Bank, Upton, NY 11973 (filename 1AHL).

* To whom correspondence should be addressed. Fax: +61-3-903 9655. Phone: +61-3-903 9650. e-mail: ray@mel.dbe.csiro.au.

[⊗] Abstract published in *Advance ACS Abstracts,* February 15, 1995.

[1] Abbreviations: AP-A(B), anthopleurin-A(B); ATX I−V, *Anemonia sulcata* toxins I−V; Sh I, *Stichodactyla helianthus* neurotoxin I; 1D, one-dimensional; 2D, two-dimensional; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser enhancement; NOESY, 2D NOE spectroscopy; TOCSY, 2D total correlation spectroscopy; DQF-COSY, double-quantum-filtered 2D correlated spectroscopy; E-COSY, exclusive COSY; rms, root mean square.

0006-2960/95/0434-3782$09.00/0

et al., 1990; Wilcox et al., 1993), have also been determined by NMR spectroscopy. Their overall structures were similar to each other and to that of AP-A, but neither is an ideal model for AP-A as they are both potent crustacean neurotoxins with little or no activity on mammalian tissue (Kem, 1988; Norton, 1991). The affinity of AP-A for the cardiac sodium channel no doubt reflects the presence of certain key residues which are lacking in ATX Ia and Sh I, but these have not yet been identified unequivocally (Norton, 1991). As the presence of these residues will affect the local conformation even though the global fold is similar, we have undertaken a high-resolution structure determination of AP-A by ^{1}H NMR spectroscopy. Here we describe this structure and compare it with those of ATX Ia and Sh I. In conjunction with previous data on residues essential for function, the structure provides an indication of which regions of the molecular surface might contribute to the affinity of AP-A for the cardiac sodium channel.

MATERIALS AND METHODS

Materials. Specimens of *Anthopleura xanthogrammica* were obtained frozen from Bodega Bay Marine Laboratories, CA. AP-A was isolated by the method of Schweitz et al. (1981), as modified by Gould and Norton (1995). ^2H_2O, ^{2}HCl, and NaO^{2}H were obtained from Cambridge Isotope Laboratories, Woburn, MA.

NMR Spectroscopy. ^{1}H NMR spectra were recorded on Bruker AMX-500 and -600 spectrometers. All 2D spectra were recorded in phase-sensitive mode using the time-proportional phase incrementation method (Marion & Wüthrich, 1983). Solvent suppression was achieved by selective, low-power irradiation of the water resonance during the relaxation delay, which was typically 2.25 s, and for NOESY spectra also during the mixing time. NMR data used for structure determination were acquired at 600 MHz on a 3 mM sample of AP-A in 90% H_2O/10% ^2H_2O or ^2H_2O at pH 4.8 and a temperature of 300 K. Spectra were also recorded at 285 and 315 K to assist with cross-peak assigments. Homonuclear 2D NOESY spectra (Anil-Kumar et al., 1980; Macura et al., 1981) were recorded with mixing times of 40, 80, 120, 200, and 250 ms in 90% H_2O/10% ^2H_2O and 50 and 200 ms in ^2H_2O. Homonuclear 2D TOCSY spectra (Braunschweiler & Ernst, 1983) were recorded using the DIPSI-2 sequence (Rucker & Shaka, 1989) for spin-lock times of 70−80 ms. Stereospecific assignments and side chain torsion angle constraints, as well as additional distance restraints, were obtained from a DQF-COSY (Rance et al., 1983), complementary E-COSY (Griesinger et al., 1987), and two NOESY spectra recorded on a sample of AP-A in ^2H_2O. Typically, spectra were acquired with 400 t_1 increments, 64−96 scans per increment, and 4096 data points. Slowly-exchanging amide protons were identified by lyophilizing the sample from 90% H_2O/10% ^2H_2O and then recording a series of 1D and TOCSY spectra immediately after dissolution in ^2H_2O.

Spectra were processed on Silicon Graphics IRIS 4D/30 workstations using UXNMR (Bruker) and analyzed using FELIX, version 2.05 (Biosym). For most spectra a phase-shifted (30−90°), sine-squared window function was applied prior to Fourier transformation. For the NOESY spectra used to measure cross-peak volumes for distance restraints, a 90°-shifted sine-squared window function was applied. Final matrix sizes were generally 2048 × 2048 real points although additional zero filling was used for the complementary E-COSY and DQF-COSY spectra.

$^3J_{HN\alpha H}$ coupling constants were measured from a DQF-COSY spectrum of AP-A in 90% H_2O/10% ^2H_2O at 300 K. After processing, the appropriate rows were selected from the 2D matrix and then inverse Fourier transformed, zero filled to 32 K points, and Fourier transformed to give a final digital resolution of 0.25 Hz/point. $^3J_{HN\alpha H}$ coupling constants were then derived from a simulation of the ω_2-cross-sections of the NH−C$^\alpha$H cross-peaks. $^3J_{C\alpha HC\beta H}$ coupling constants were measured from a complementary E-COSY spectrum in ^2H_2O. Due to the significant cross-peak overlap caused by the presence of multiple conformations of AP-A in solution (Scanlon & Norton, 1994), it was not always possible to measure the $^3J_{C\alpha HC\beta H}$ coupling constant from the passive component of the E-COSY cross-peak. In these cases the appropriate row was extracted from the 2D matrix and analyzed using inverse Fourier transformation as described above. The coupling constant was then determined by simulation of the cross-section of the cross-peak. These coupling constants, together with $d_{\alpha\beta}(i,i)$ and $d_{N\beta}(i,i)$ NOEs from short mixing time NOESY spectra in ^2H_2O and H_2O, were used to determine if side chains could be placed in one of the staggered side chain rotamer conformations (χ^1 = 60°, 180°, and −60°) and to make stereospecific assignments (Wagner et al., 1987; Hyberts et al., 1987).

Structural Restraints. NOESY cross-peak volumes were measured from 200 ms mixing time spectra in H_2O and ^2H_2O. Peaks from the upper side of the diagonal were used except where peaks from the lower side were better resolved. The volumes of cross-peaks from degenerate methylene protons and Tyr C(2,6)H and C(3,5)H protons were divided by two, and those of cross-peaks involving methyl groups were divided by three. Peak volumes in each matrix were calibrated using an average value of the volumes of well-resolved geminal C$^\beta$H cross-peaks, measured from both sides of the diagonal and assigned a distance of 1.79 Å. After conversion of the volumes to distances, corrections of 0.5 and 1.0 Å were added to distance restraints involving backbone protons only and at least one side chain proton, respectively. In addition, pseudoatom corrections were added as follows: (i) For DIANA calculations, 0.9, 1.1, and 2.4 Å, respectively, were added for degenerate methylene, methyl, and tyrosine ring center pseudoatoms. Restraints to nondegenerate methylene protons which were not stereospecifically assigned were corrected automatically by DIANA. (ii) For X-PLOR calculations, the bounds output file from DIANA was converted to X-PLOR format, and for the tyrosine ring protons the restraints were reinterpreted for separate (2,6) and (3,5) pseudoatoms (correction of 2.2 Å).

The backbone dihedral angle ϕ was constrained as follows: −120° for $^3J_{HN\alpha H}$ > 8 Hz and −60° for $^3J_{HN\alpha H}$ < 5 Hz (Pardi et al., 1984). Deviations from these angles varied depending on the value of the coupling constant, the minimum deviation being 30°. Backbone dihedral restraints were not applied for $^3J_{HN\alpha H}$ values between 5 and 8 Hz. χ^1 torsion angles were constrained to values of 60°, 180°, or −60° as appropriate, with minimum deviations of 30°.

Calculation of Structures. Initial structures were generated with the distance geometry program DIANA (Güntert et al., 1991) using dihedral angle restraints derived from coupling

constant data and distance constraints derived from NOE cross-peaks assigned unambiguously in both chemical shift dimensions. No hydrogen-bonding restraints were used in the calculations. The Gly40−Pro41 ω-angle was set to 0°, as the configuration has been shown recently to be *cis* in the major species in solution (Scanlon & Norton, 1994). Structures were generated starting with randomized dihedral angles and using the standard search protocol and parameters for DIANA, except that at each level of minimization a maximum of 300 conjugate gradient steps was performed. The redundant dihedral angle constraints (REDAC) strategy (Güntert & Wüthrich, 1991) was employed to reduce the computational time for generation of an acceptable family of structures. At the final stage of minimization, the REDAC dihedral constraints were removed.

Ambiguous NOESY cross-peak assignments were resolved where possible using these initial structures. An assignment was accepted if in all structures the appropriate interproton distance was <5 Å and the distances between alternative proton pairs were >7 Å. In many cases the final assignment was confirmed by analysis of spectra at different temperatures. A small number of lower bound restraints >1.8 Å (Wilcox et al., 1993; Manoleras & Norton, 1994) was also used in the final structure calculations. These were obtained with an in-house program which checked the DIANA structures for distances of <3.5 Å that were not represented in the NOE restraint list. Distances identified in this way were compared with the experimental NOESY spectra to confirm that cross-peaks could have been observed had they been present. Where a cross-peak was clearly absent, a lower bound restraint of 3.5 Å was added to the restraint set; in all other cases the lower bound was 1.78 Å. This analysis was limited to interactions involving pairs of NH, $C^{\alpha}H$, and $C^{\beta}H$ protons not affected by water suppression.

Once the final set of restraints had been obtained, a new family of distance geometry structures was generated using DIANA, and the 50 structures with the lowest penalty functions were refined by simulated annealing in X-PLOR, version 3.1 (Brünger, 1992). Another family of 200 structures was generated in X-PLOR, starting from a linear template structure and different randomized initial velocity distributions, and subjected to the same simulated annealing protocol. The all-hydrogen distance geometry force field was employed, which includes covalent geometry, planarity, and hard sphere van der Waals terms but not Lennard-Jones, electrostatic, or empirical dihedral angle (apart from planarity) terms. Simulated annealing was performed using 20 000 steps at 1000 K and 10 000 steps as the molecule was gradually cooled to 300 K. A time step of 1 fs was employed throughout. The initial van der Waals weighting was very low (0.002) so as to allow atoms to pass through each other early in the simulation. Disulfide bridges were included as distance restraints of 2.02 ± 0.1 Å between the sulfur atoms and 3.10 ± 0.1 Å between $C^{\beta}{}_i$ and $S^{\gamma}{}_j$. The Gly40−Pro41 ω-angle was again set to 0°; 200 steps of energy minimization in the distance geometry force field were performed.

The following modifications were made to the psfgen.inp, temgen.inp, and sa.inp macro files in X-PLOR following the observation that the N-terminal C^{α} improper energy and the Pro ring bond angle energy were dominating the total energy of the molecule: (i) an atom-based improper statement for the chirality of the N-terminal glycine C^{α} was included to overrule a conflicting type-based improper in parmallhdg.pro, which occurs for N-terminal glycines, and (ii) atom-based bond angle statements were included to overrule the $N-C^{\alpha}-C^{\beta}$, $C^{\alpha}-N-C^{\delta}$, and $C^{\gamma}-C^{\delta}-N$ type-based parameters for Pro rings (adjusted to 103.5°, 114.0°, and 104.0°, respectively, from 109.5°, 119.5° and 109.5°).

The 30 best refined DIANA structures and 30 best structures originating from X-PLOR were then subjected to further simulated annealing in which they were gradually cooled from 300 to 0 K in 20 000 steps and then energy minimized using 1000 steps of Powell conjugate gradient minimization. For each structure this procedure was carried out 10 times and the best of these 10 in terms of total energy and NOE energy was selected. Finally, the 20 structures with the lowest total and NOE energies were energy minimized in the empirical CHARMm force field (Brooks et al., 1983), with a distance-dependent dielectric instead of explicit water molecules. The force field was modified to leave the backbone ϕ and ψ angles unconstrained by the empirical dihedral angle term; this was achieved by removing the relevant type-based dihedral terms from the parmallh3x.pro file. These 20 structures constituted the final set used for structural analysis. Compared with the structures that were energy minimized in the X-PLOR distance geometry force field, the final structures had mean pairwise rms differences over the backbone heavy atoms (N, C^{α}, and C) and all heavy atoms of 0.70 and 0.79 Å, respectively.

To investigate the structural consequences of *cis*−*trans* isomerism about the Gly40−Pro41 peptide bond, the final 20 structures were subjected to 20 ps of unrestrained dynamics at 300 K in the CHARMm force field. Two brief simulated annealing runs were then carried out on each structure, one with the Gly40−Pro41 peptide bond in the *cis* conformation and the other with it *trans*, in which the structures were heated to 1000 K and then cooled to 50 K over 1.2 ps. The two resulting families of structures were then subjected to a further 20 ps of simulated annealing, with cooling from 300 to 50 K.

Structural characteristics were analyzed using the program SSTRUC, based on the methods of Kabsch & Sander (1983). Structures were displayed in Insight II (Biosym Technologies, San Diego, CA). Surface areas were calculated in Insight from Connolly surfaces generated with a 1.4 Å probe radius.

RESULTS

Resonance Assignments. It has been known for some time that AP-A exists in solution in two conformations in a ratio of 2:1, as reflected in the splitting of a number of ^{1}H NMR resonances (Gooley et al., 1984, 1988). Recently, we showed that the major source of conformational heterogeneity was *cis*−*trans* isomerism about the Gly40−Pro41 peptide bond, with the major conformer adopting the *cis* conformation (Scanlon & Norton, 1994). Additional minor conformers were also observed, although the source of this heterogeneity has not been identified. In this paper the structure of the major form of AP-A in solution is determined and the consequences of *cis*−*trans* isomerism at Gly40-Pro41 are inferred from molecular dynamics simulations.

The ^{1}H NMR assignments for most spin systems in the major form of AP-A and many in the minor form have been published previously (Mabbutt & Norton, 1990). Our assignments, obtained at 600 MHz, for the major form are in excellent agreement with these published data. Additional

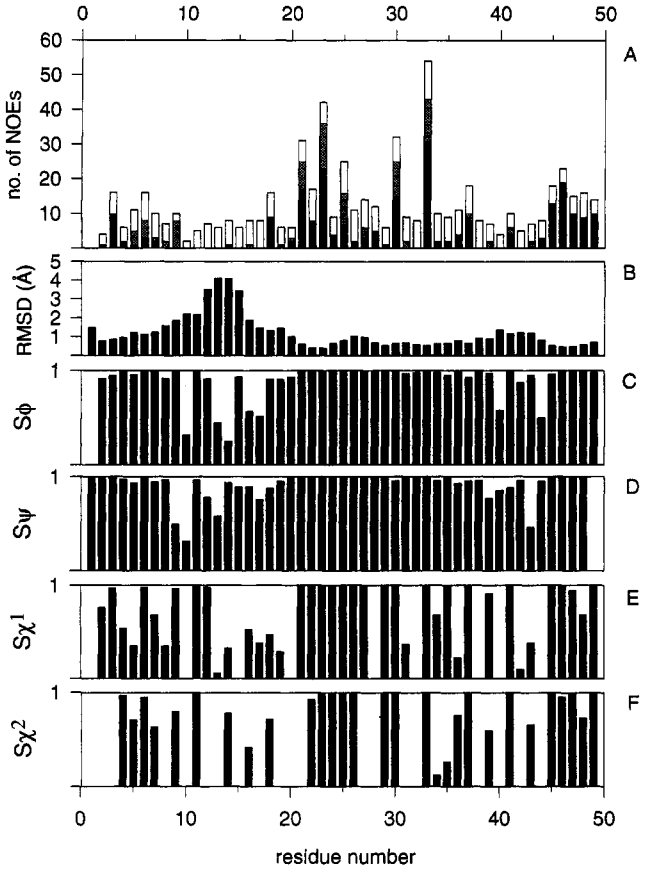

FIGURE 1: Parameters characterizing the final family of structures for AP-A, plotted as a function of residue number. (A) Distribution of upper bound restraints used in the final round of structure calculations. Long range restraints (■), defined as $|i - j| \geq 5$ for a restraint between residues i and j, medium range (light shading), where $1 < |i - j| < 5$, and sequential (□) restraints are shown. Each restraint is counted twice, once for each proton involved. Intraresidue restraints are not shown. (B) rms differences from the mean structure for the backbone heavy atoms (N, C^{α}, and C) when the final 20 structures are superimposed over the entire length of the molecule. (C–F) Angular order parameters (Hyberts et al., 1992; Pallaghy et al., 1993) for the backbone dihedral angles ϕ and ψ and the side chain angles χ^1 and χ^2. The gaps in the χ^1 plot are due to the eight Gly residues and Ala38.

assignments were obtained as follows: Ser8 NH, 8.38; Trp23 N(1)H, 5.92; Pro41 $C^{\alpha}H$, 4.71; $C^{\beta}H_2$, 2.24, 2.28; $C^{\gamma}H_2$, 1.49, 1.53; $C^{\delta}H_2$, 3.35, 3.49; Gln49 $N^{\epsilon}H_2$, 6.50, 6.64 ppm. The partial assignments for the Pro41 spin system reported by Mabbutt and Norton (1990) appear to be for the minor (*trans*) form. Partial or complete assignments for 38 of the 49 spin systems in the minor form have also been obtained in the present work.

Structure Determination and Analysis. The structures were generated using a restraint set consisting of 411 interproton distance restraints inferred from NOEs and 19 backbone and 13 side chain dihedral angle restraints from spin–spin coupling constants, as well as 15 lower bound restraints based on the absence of NOEs in the spectra. Representative NOESY spectra are shown in Scanlon and Norton (1994). The upper bound distance restraint set, from which values redundant with the covalent geometry have been eliminated, consisted of 105 intraresidue, 146 sequential, 45 medium range ($1 < |i - j| < 5$), and 115 long range restraints; their distribution as a function of residue number is summarized in Figure 1A. Initial structures were calculated using either DIANA or X-PLOR (with the distance geometry force field) and then refined by simulated annealing in X-PLOR. Finally, the structures were energy minimized in X-PLOR with the CHARMm force field. The final 20 structures consisted of 10 which originated from DIANA structures and 10 from X-PLOR, as it was found that the the two sets of structures had overlapping energies. A summary of geometric and energetic parameters for these structures is given in Table 1.

Table 1: Structural Statistics for 20 Energy-Minimized Structures of AP-A from X-PLOR[a]

	force field	
	distance geometry	CHARMm
rms deviations from experimental distance restraints Å (411)[b]	0.0415 ± 0.0006	0.0255 ± 0.0008
rms deviations from experimental dihedral restraints (deg) (32)[b]	1.00 ± 0.048	0.543 ± 0.098
rms deviations from idealized geometry		
bonds Å	0.0034 ± 0.00005	0.010 ± 0.0004
angles (deg)	0.54 ± 0.006	3.24 ± 0.040
impropers (deg)[c]	0.38 ± 0.007	0.32 ± 0.014
energies (kcal mol^{-1})		
E_{NOE}[d]	40.5 ± 1.1	15.4 ± 0.9
E_{cdih}[d]	2.0 ± 0.2	0.6 ± 0.2
E_{L-J}[e]	8.0 ± 0.5	−164.6 ± 5.9
$E_{bond} + E_{angle} + E_{improper}$[f]	73.2 ± 1.9	215.7 ± 5.7
E_{elec}[g]		−648.7 ± 18.4

[a] The best 20 structures after energy minimization in the distance geometry force field of X-PLOR were subsequently energy minimized in the CHARMm force field, as described under Materials and Methods. Values in the table are mean ± standard deviation. [b] The number of restraints is shown in parentheses. None of the structures had distance violations >0.3 Å or dihedral angle violations >5°. [c] Improper torsion terms serve to maintain planar geometries and chiralities. [d] Force constants for calculation of the square-well potentials for the NOE and dihedral angle restraints were 50 kcal mol^{-1} $Å^{-1}$ and 200 kcal mol^{-1} rad^{-2}, respectively. [e] The quartic van der Waals repulsion term in the distance geometry force field was calculated with a force constant of 4 kcal mol^{-1} $Å^{-4}$ and the hard sphere van der Waals radius set to 0.75 times the standard value used in the CHARMm empirical energy function. For the structures minimized in CHARMm, the Lennard-Jones–van der Waals energy was calculated with the CHARMm empirical energy function. [f] These terms serve to maintain the covalent geometry. [g] Electrostatic contribution to the overall energy, calculated with a distance-dependent dielectric.

The rms differences for the backbone heavy atoms when the final 20 structures are superimposed over the entire sequence are shown as a function of residue number in Figure 1B. These data indicate that the structure is well-defined except at the N-terminus and around residues 8–16. The backbone angular order parameters (Figure 1C,D) indicate that residues 17 and 39–44 are not as precisely defined as the bulk of the structure (excluding residues 8–16), but for the purpose of calculating rms differences they have been included with the well-defined residues. Thus, over residues 2–7 and 17–49 the mean pairwise rms differences for the backbone heavy atoms (N, C^{α}, and C) and all heavy atoms were 0.82 and 1.27 Å, respectively. Corresponding values for the whole molecule were 2.04 and 2.59 Å, respectively. The β-sheet residues (2–4, 20–23, 34–37, and 45–48) are very well-defined, with a mean pairwise rms difference over the backbone heavy atoms of 0.50 Å.

Stereoviews of the best 20 structures superimposed over the backbone heavy atoms of the well-defined residues are

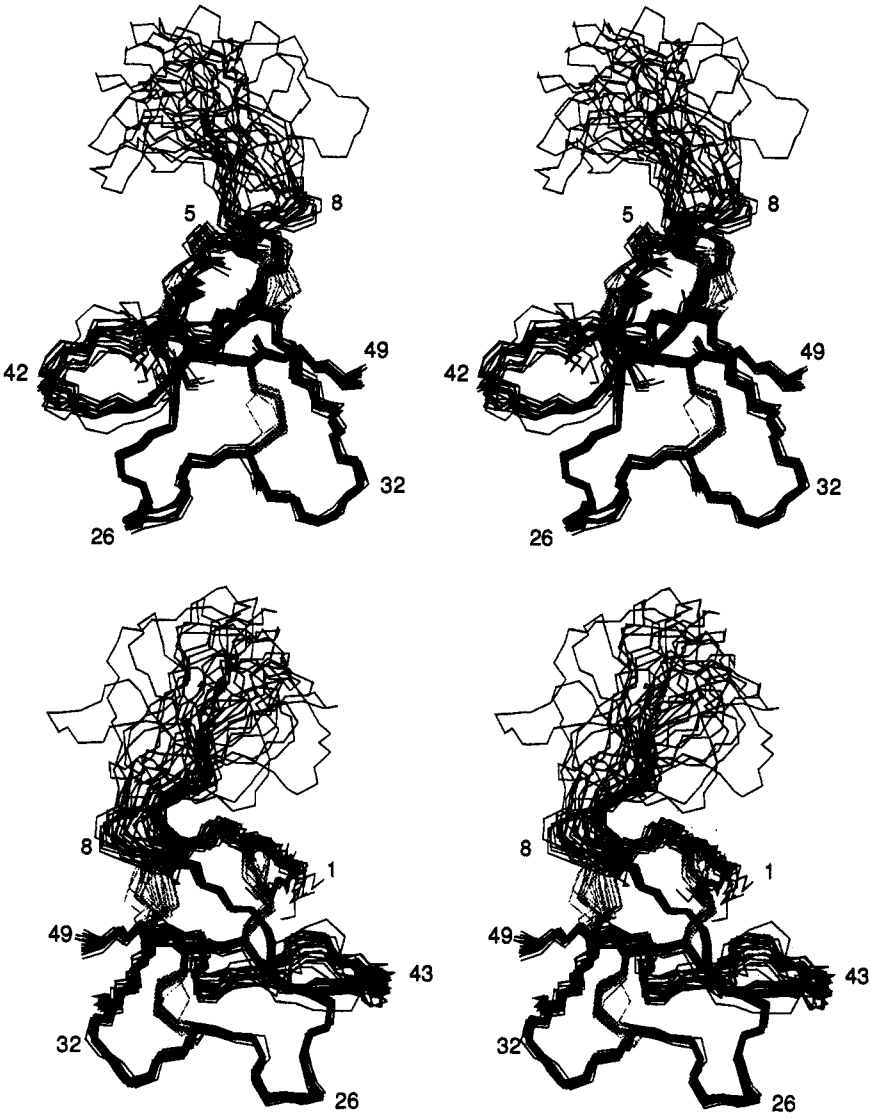

FIGURE 2: Stereoviews of the 20 final structures of AP-A superimposed over the backbone heavy atoms (N, C^{α}, and C) of residues 2−7 and 17−49. The three disulfide bonds are shown in lighter shading. The lower view is related to the upper one by an approximately 180° rotation about the vertical axis.

shown in Figure 2. The poorly-defined region evident from the data presented in Figure 1 is part of a loop linking the first and second strands of the sheet. Its orientation in the overall structure is poorly defined because residues in this loop display no long range and very few medium range NOEs (Figure 1A). The other region that shows slightly higher backbone rms differences and angular order parameters (Figure 1) involves residues 39−44, which forms a loop linking the third and fourth strands of the sheet (Figure 2) and includes the Gly40−Pro41 peptide bond, which undergoes *cis*−*trans* isomerism (Scanlon & Norton, 1994). Once again, there is a dearth of medium and long range NOEs to these residues. The χ^1 angles (Figure 1E) are reasonably well-defined in the region of the sheet. The largest cluster of ordered side chains is found in the second strand of the sheet and the loop joining the second and third strands. The locations of side chains having angular order parameters for χ^1 and χ^2 (Figure 1) of >0.8 are shown in Figure 3.

A Ramachandran plot of the angular average of the final 20 structures is shown in Figure 4. Most residues with well-defined backbone dihedral angles ($S > 0.8$) fall within the allowed regions. Three of the five glycines have well-defined positive ϕ angles.

Description of the Structure. Secondary structure assignment was based on the identification of hydrogen bonds, the locations of which are shown in Figure 5. A MOLSCRIPT representation of the secondary structure and disulfide bonds is shown in Figure 6. The major structural feature is a four-stranded, antiparallel β-sheet consisting of residues 2−4, 20−23, 34−37, and 45−48 and displaying a 1, 2x, −1 topology (Richardson, 1977). A β-bulge occurs in strands III and IV, with the amides of Cys36 and Lys37 in strand III hydrogen bonding to the carbonyl of Cys46 in strand IV. The bulge most closely matches the classic C_+ subclass (Chan et al., 1993) with Cys36 (position 1 in the bulge) having α-helical ϕ and ψ angles and Lys37 (position 2 in the bulge) having β-sheet angles. The loops linking the strands in the sheet contain β-turns at the following locations: type I at residues 6−9 (ϕ_2 −63°, ψ_2 −12°; ϕ_3 −96°, ψ_3 −8°), type II at residues 13−16 (ϕ_2 −79°, ψ_2 154°; ϕ_3 103°, ψ_3 −16°), type I at residues 25−28 (ϕ_2 −65°, ψ_2 −14°; ϕ_3 −121°, ψ_3 12°), and type II at residues 30−33 (ϕ_2 −59°, ψ_2 106°; ϕ_3 134°, ψ_3 −1°). The first and second of these turns are located in the least well-defined part of the molecule and are present in 95% and 40%, respectively, of the structures, whereas the third and fourth turns are present in all 20 structures.

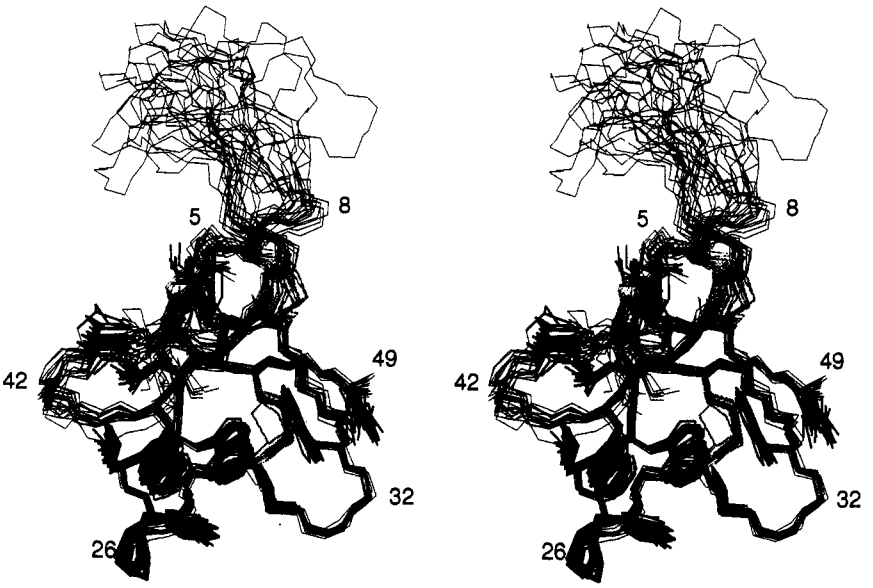

FIGURE 3: Stereoviews of the 20 final structures of AP-A superimposed over the backbone heavy atoms (N, C^{α}, and C) of residues 2−7 and 17−49 and showing in bold the locations of side chains having angular order parameters for χ^1 and χ^2 (Figure 1) > 0.8. Only side chains from the well-defined part of the structure are shown, although Pro11 is the only residue in the poorly-defined loop with both χ^1 and $\chi^2 > 0.8$.

The loop linking the third and fourth strands of the sheet does not contain any well-defined turn. The second and third strands of the sheet are linked by a loop that includes turns at residues 25−28 and 30−33, but this region cannot be regarded as helical.[2]

The exchange rates of backbone amide protons in AP-A have been measured by Torda and Norton (1987). The hydrogen bonds in the β-sheet (Figure 5) account for eight of those identified as slowly exchanging, and it is likely that the resonance at 8.39 ppm observed but not assigned by Torda and Norton (1987) arises from Lys37. The only hydrogen-bonded amides in the sheet that were not identified as slowly exchanging were those of Val2 and Gly20, which may reflect a higher degree of solvent exposure in this region of the molecule. Of the amides hydrogen bonded in β-turns, that of Trp33 was found to exchange slowly with solvent, and resonances that could be assigned to the Asp9 (Gould et al., 1992) and Gly28 amides had intermediate exchange rates, but the amide of Asn16 was not slowly exchanging, probably because of its location in the poorly-defined loop. The slowly-exchanging amides not accounted for by these secondary structure elements are those of Cys6, His39, Cys46, and Lys48. On the basis of the secondary structure of AP-A inferred from assignment data (Gooley & Norton, 1986; Mabbutt & Norton, 1990) and the sheet structure illustrated in Figure 5, it was expected that the Lys48 amide would be hydrogen bonded to the carbonyl of His34. This was not detected as a hydrogen bond in any of our structures because the His34 carbonyl is twisted out of the plane of the sheet. The amide of His39 hydrogen bonds to the carbonyl of Lys37 in one-third of the final structures, but no clear hydrogen bond acceptor is apparent for Cys6. The amide of Cys46 hydrogen bonds to the carbonyl of Gly40 in 12 of the 20 structures and is discussed below in relation to differences between the *cis* and *trans* conformers. An additional hydrogen bond, presumably stabilized by the Cys6−Cys36 disulfide bond, was observed in all structures from Asp7 NH (in turn I) to Cys36 (in strand III), although the amide of Asp7 was not found to be slowly exchanging. The solvent accessibilities of the amide protons of Cys6, Asp7, His39, Cys46, and Lys48 were all near zero.

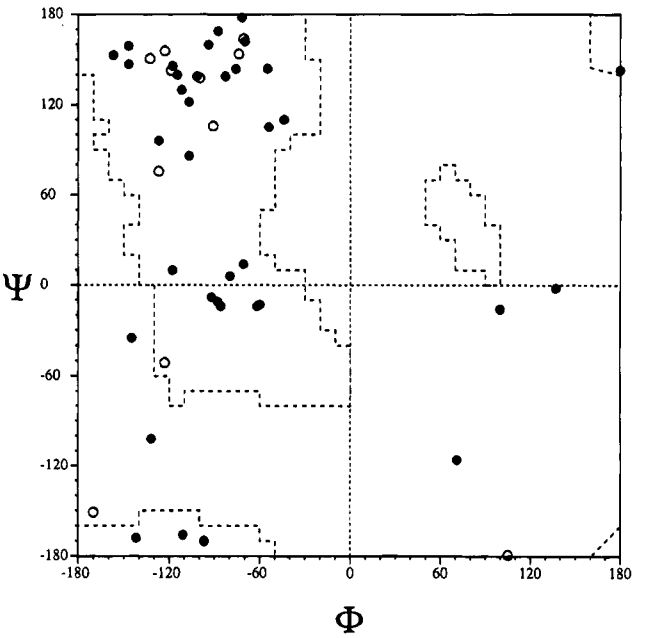

FIGURE 4: Ramachandran plot for the angular average of the 20 final structures of AP-A. Filled circles denote residues with both backbone dihedral angles well-defined (S > 0.8), as documented in Figure 1.

The conformation of the 29−47 disulfide bond is quite well-defined (Figures 1 and 2), and it has a χ_{SS} of −69°, but the combination of side chain dihedral angles does not fit any of the classifications of Srinavasan et al. (1990). This bond is the least solvent exposed of the three. The conformation of the 4−46 bond is less well defined because of low angular order parameters for χ^1 of Cys4 and χ_{SS}, while

[2] Orengo and Thornton (1993) claim that Sh I contains a 3_{10} helix in this region, but we did not detect this in the original or the refined structures (Fogh et al., 1990; Wilcox et al., 1993) nor in AP-A.

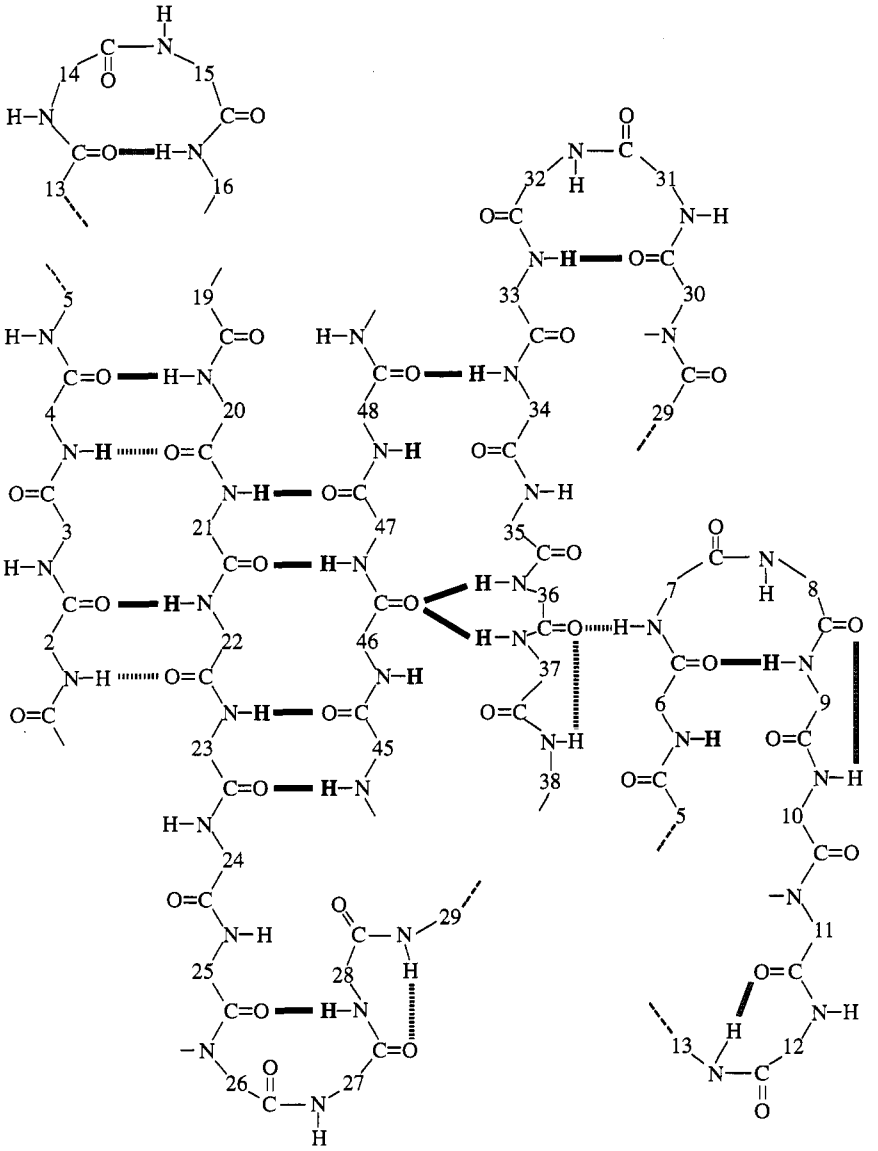

FIGURE 5: Hydrogen bonds in the 20 final structures of AP-A. Hydrogen bonds that were present in 19 or more of the structures are shown as solid black lines and those in 12−18 as hatched lines. Hydrogen bonds in the poorly-defined loop that were present in only eight to nine structures are shown with light shading. Slowly-exchanging backbone amide protons identified by Torda and Norton (1987) are shown in bold. Not shown is a hydrogen bond from Cys46 NH to the carbonyl of Gly40, which is present in 12 structures.

that of the 6−36 bond is the least well-defined because of low angular order parameters for χ_{SS} and the side chain of Cys36. The poorer definition of the conformations of these two disulfides may reflect local structural disorder.

The orientation of the loop connecting strands I and II of the sheet is poorly defined relative to the bulk of the structure. If, however, the structures are superimposed on residues 8−16 rather than on the well-defined residues, it becomes apparent that the structure of the loop is not completely random, as shown in Figure 7. The mean pairwise rms differences over the backbone heavy atoms and all heavy atoms of the loop are 2.25 and 3.51 Å, respectively, clearly higher than the corresponding values obtained when residues 2−7 and 17−49 are superimposed but low enough to suggest that there is some internal order in the loop. The $^3J_{HNC\alpha H}$ coupling constants for residues in the loop are all in the range 6−8 Hz, consistent with a largely-extended structure undergoing some conformational averaging, but the NH line widths are comparable to those for the rest of the molecule, implying that the backbone is not undergoing fast internal motion. These observations raise the question of whether the loop as a whole is oscillating around "hinges" at either end. It is apparent from the angular order parameters for ϕ and ψ (Figure 1) that Gly10 has very low order and could be a potential hinge site at the N-terminal end of the loop, while residue 17 may serve a similar function at its C-terminus. A role such as this for Gly10 may explain why it is conserved throughout all of the long sea anemone polypeptides (Norton, 1991).

[3] Upfield chemical shifts are observed for three of the amides involved in β-turns in AP-A, as follows: Asp9, 0.67; Asn16, 0.65; Gly28, 1.46 ppm. Only the Trp33 NH, which is hydrogen bonded in the 30−33 turn, experiences a downfield shift of 0.35 ppm. These shifts are consistent with previous observations for hydrogen-bonded amides in turns, as summarized by Gould et al. (1992). The unusually large upfield shift for Gly28 appears to be due to ring current shifts from Trp45 (our unpublished results). A similar shift is also seen in ATX Ia (Widmer et al., 1988; Norton, 1991), probably caused by interaction with Tyr45, but not in the type 2 polypeptides (Norton, 1991), where the aromatic residue at position 45 is replaced by polar side chains.

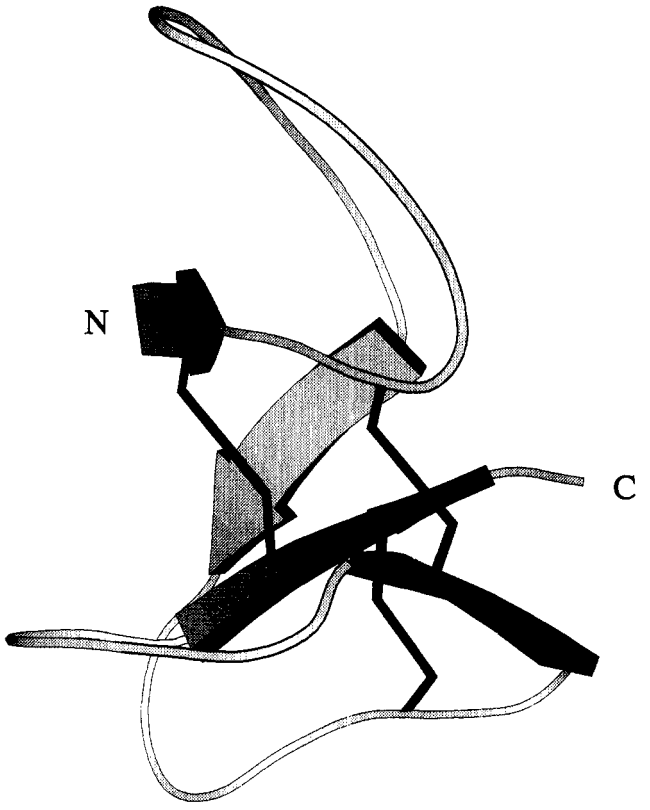

FIGURE 6: Richardson style diagram of the polypeptide backbone of the individual structure of AP-A that is closest to the average over the whole molecule. Disulfide bonds are shown. The diagram was generated using MOLSCRIPT (Kraulis, 1991).

A type II β-turn involving residues 13−16 is present in nearly half of the final structures. This probably accounts for the perturbed chemical shift[3] of the amide of residue 16 in this class of polypeptides (Norton, 1991) and peptide analogues of the loop (Gould et al., 1992). Tryptic cleavage of the polypeptide backbone between Arg14 and Gly15 removes this chemical shift perturbation (Gould et al., 1990).

Interactions among Charged Side Chains. AP-A contains three carboxylate groups, from Asp7, Asp9, and the C-terminus, three ammonium groups from Lys37, Lys48, and the N-terminus, two imidazolium groups from His34 and His39, and one guanidino group from Arg14. The pK_a values of most of these groups have been measured previously (Norton et al., 1982; Gooley et al., 1988). The values for the N- and C-termini and the two imidazolium groups were close to those found in small peptides, while those for the Lys ammonium groups were slightly higher than in model peptides. One of the Asp carboxyl groups had a pK_a of about 3.5, slightly lower than the value of 3.9 in peptides, but the other had a significantly lower pK_a of around 2. Furthermore, protonation of this carboxylate was associated with significant chemical shift changes for a number of residues, implying that it is involved in electrostatic interactions that stabilize the native structure. Establishing which Asp has the low pK_a has been difficult, partly because of their close proximity in the amino acid sequence.

The locations of the Asp side chains and nearby charged residues are shown in Figure 8. The closest charged group to either of the Asp side chains is the ϵ-ammonium group of Lys37, which is near Asp7; over the 20 final structures the mean distance (±SD) between Asp7 C^{γ} and Lys37 N^{ζ} is 4.0 ± 1.3 Å. By contrast, Asp9 is 8.2 ± 1.2 Å from Lys37 and 8.1 ± 2.7 Å from Lys48, and Asp7 is 10.9 ± 2.0 Å from Lys48. These observations suggest that Asp7 is the low-pK_a residue and that its pK_a is low because of a salt bridge to Lys37. Interaction with the imidazolium ring of His39 may also contribute to the low pK_a of Asp7, as its C^{γ} is only 6.3 ± 1.1 Å from the ring C2 of His39 (although the pK_a of His39 is not significantly perturbed).

In order to determine if Asp7 was indeed the low-pK_a residue, we recorded 2D ^{1}H NMR spectra of AP-A in H_2O at several pH values in the range 5−2. The amide resonance of Gln49 had a pK_a of 3.2, slightly lower than observed previously for the C-terminus (Norton et al., 1982; Gooley et al., 1988), and a titration shift ($\delta_B - \delta_A$) of −0.44 ppm. The amide of Asp9 had a pK_a of 3.6 and a titration shift ($\delta_B - \delta_A$) of −0.50 ppm, the latter being larger than observed for Asp NH resonances in linear peptides (Bundi & Wüthrich, 1979), possibly as a result of a local conformational change associated with carboxylate protonation. As the pH was lowered, the amide of Asp7 experienced a small upfield shift with a pK_a of about 3.3 but had not undergone the expected downfield shift associated with sidechain carboxylate protonation by pH 2.6, below which it broadened and became difficult to follow. Accurate pK_a values were not obtained for the $C^{\beta}H$ resonances of either Asp residue because of peak overlap, but it was clear that the Asp9 resonances shifted significantly between pH 4 and 3, whereas those of Asp7 did not begin to move until pH 2.6. While indirect, these observations all point to Asp9 as being the residue with a pK_a around 3.5 and Asp7 having a pK_a of ≤2. This is supported by the behavior of the Lys37 $C^{\epsilon}H$ resonances, one of which shifts by about 0.3 ppm with pK_a ≤2.

Structural Differences between Conformers. Scanlon and Norton (1994) noted that AP-A in solution was a mixture of

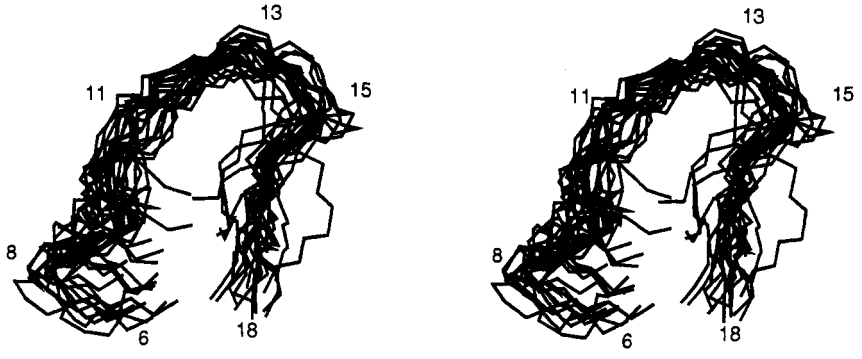

FIGURE 7: Stereoview of residues 6−18 in the 20 final structures of AP-A, superimposed over the backbone heavy atoms (N, C^{α}, and C) of the poorly-defined residues 8−16.

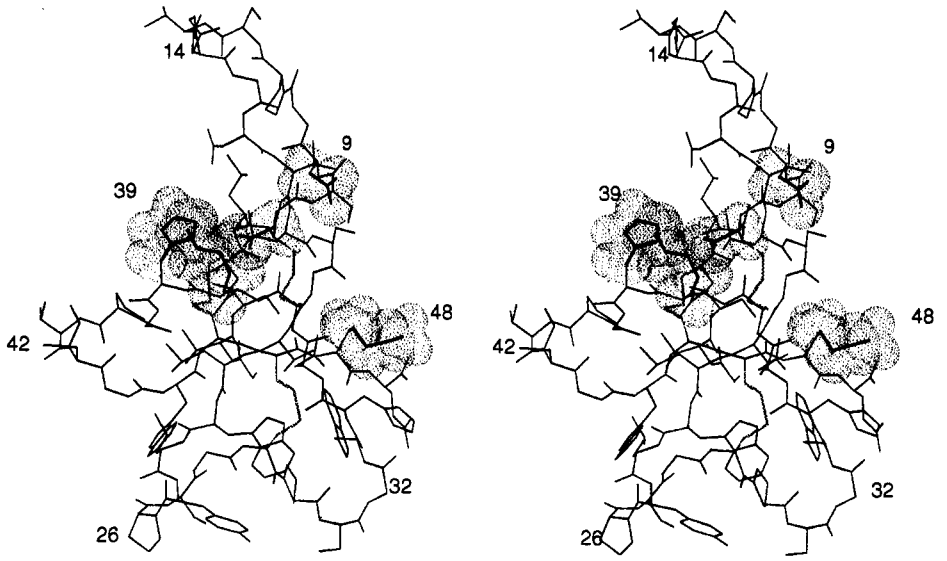

FIGURE 8: Stereoview of the individual structure of AP-A which is closest to the average over the whole molecule showing the backbone and all side chains (in bold). The van der Waals surfaces for the charged residues 7, 9, 37, 39, and 48 are shaded.

conformers in slow exchange on the NMR time scale and with relative populations of approximately 18:9:1:1. The source of the major heterogeneity was *cis*−*trans* isomerism about the Gly40−Pro41 peptide bond, but the source of the minor heterogeneity has not yet been established. In this work the structure of the most abundant conformer of AP-A, with all peptide bonds except the Gly40−Pro41 bond in the *trans* conformation, has been determined. Although an extensive set of resonance assignments was available for the next most abundant minor conformer (with a *trans* Gly40−Pro41 peptide bond), it was difficult to obtain a satisfactory set of NOE-based distance restraints because of peak overlap. In general, the pattern of NOEs in the major conformer was mirrored in the minor conformer, implying that their overall structures are very similar.

In order to assess the structural implications of the altered peptide bond conformation at Gly40−Pro41, we first subjected the 20 final structures to unrestrained molecular dynamics to confirm that they were stable in the absence of experimental restraints. The conformation of the peptide bond was then changed to *trans* by simulated annealing commencing at 1000 K, and the structures were subjected to further dynamics. The family of structures with the *trans* peptide bond coincided remarkably well with the control structures in which the bond remained in the *cis* conformation, although there was greater spread in both sets of structures compared with the starting structures because of the absence of experimental restraints. This shows that the conformation of the Gly40−Pro41 bond can be changed with minimal effect on the rest of the structure.

Further information on the differences between the two conformers comes from chemical shift differences, which are often the most sensitive indicators of local conformational perturbations (Hinds & Norton, 1993). Some 13 backbone amide chemical shifts differ by >0.1 ppm between the two conformers (Scanlon & Norton, 1994), but the largest shifts are found for Cys46 NH, which moves upfield by 0.8 ppm in the *trans* conformer, and His39, which moves downfield by 0.5 ppm. The magnitude and direction of the Cys46 NH shift are consistent with the weakening of a hydrogen bond (Wagner et al., 1983). Furthermore, the Cys46 NH chemical shift is not affected significantly by ring current shifts from Trp45 in the major conformation (unpublished results), and the chemical shifts of the $C^{\alpha}H$ and $C^{\beta}H$ resonances are very similar in the two conformers, emphasizing that specific interactions involving the amide proton are the main difference between the two. The amide of Cys46 hydrogen bonds with the carbonyl of Gly40 in 12 of the final structures. During unrestrained dynamics, this bond was maintained in structures with a *cis* Gly40−Pro41 bond but was not present in any structures with the *trans* Gly40−Pro41 bond because of a change in orientation of the Gly40 carbonyl. In the majority of structures with the *trans* bond, Cys46 NH formed a hydrogen-bond with the carbonyl of Gly44 (which is solvent exposed in structures with the *cis* bond). This may explain why the Cys46 NH resonance in the minor conformation is still at low field (9.9 ppm), consistent with its participation in some hydrogen bonding interactions. Thus, it is likely that stronger hydrogen bonding of Cys46 NH in the *cis* conformer contributes to its greater stability relative to the *trans* conformer. Indeed, the shift in relative population from a *cis*:*trans* ratio of 1:4, which might be observed in the absence of other constraints (Grathwohl & Wüthrich, 1976), to the ratio in AP-A of 2:1 corresponds to a difference in ΔG° of about 1.2 kcal mol^{-1}, which is less than one-half the free energy change associated with hydrogen bond formation. The amide exchange rates of Cys46 NH in the two forms are quite similar (Torda & Norton, 1987; Gould et al., 1990), as are the temperature coefficients of the NH chemical shifts.

DISCUSSION

In this paper we have described a high-resolution structure for AP-A in solution. The previously published structure of AP-A derived from NMR data (Torda et al., 1988) was of limited resolution, sufficient to describe the backbone fold of the molecule but very few of the side chains, and not deposited in the Protein Data Bank (Bernstein et al., 1977). The present structure provides an accurate and precise description of the backbone and most of the side chains in AP-A, including those known to be important for cardiac stimulatory activity.

The main structural feature of AP-A is a four-stranded, antiparallel β-sheet, with side chains emanating from this

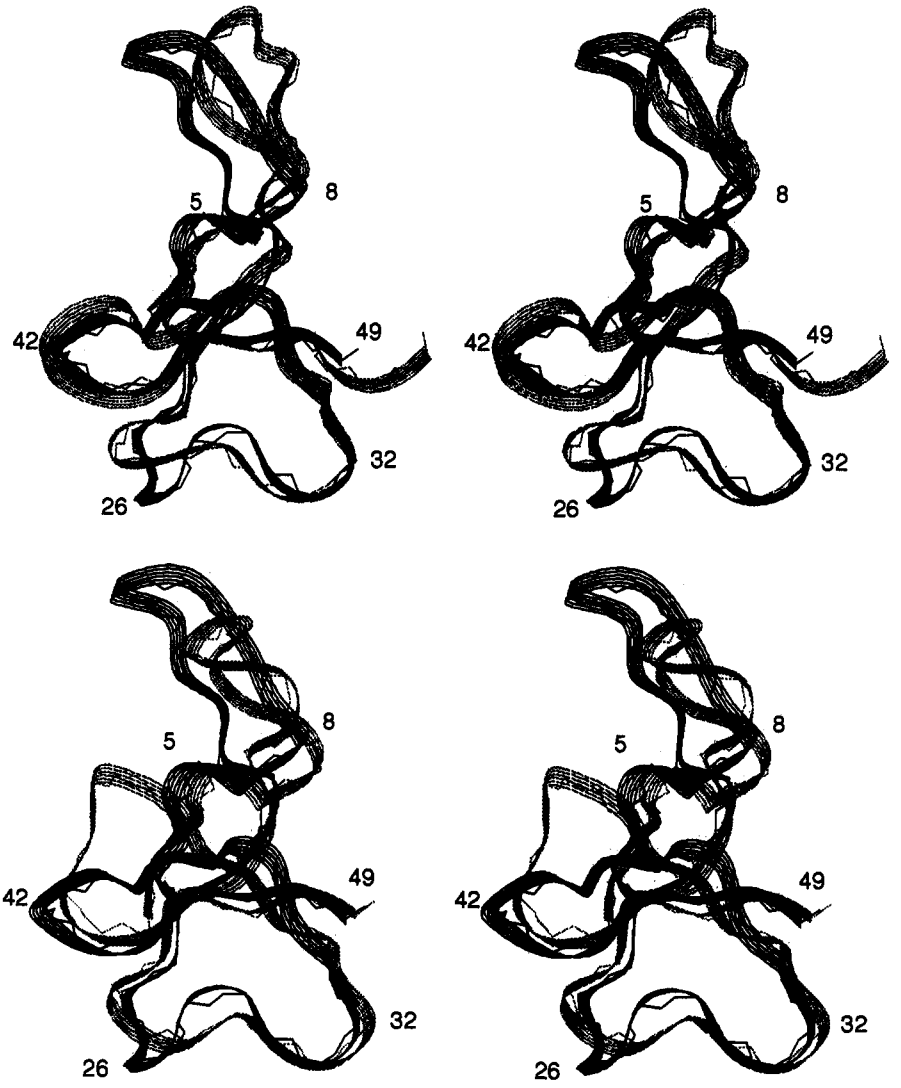

FIGURE 9: Stereoviews of superpositions of the structures of AP-A and Sh I (upper) and AP-A and ATX Ia (lower). In each case AP-A is the darker shaded structure. For AP-A the structure which is closest to the average is shown (cf. Figure 8), while for Sh I the closest to the average of the minimized 'new' structures of Wilcox et al. (1993) is shown and for ATX Ia the lowest energy structure of Widmer et al. (1989). The structures were superimposed over the backbone heavy atoms of residues 2−6, 20−24, 27−36, and 45−49 in AP-A, corresponding to 2−6, 20−34, and 42−46 in ATX Ia and 1−5, 19−33, and 42−46 in Sh I. Residue numbers on the structures are those of AP-A.

sheet making the major contribution to the hydrophobic core of the molecule (Figure 3). The solvent accessibilities of the indole rings of the three Trp residues, which form part of this core, can be compared with previous observations using photochemically-induced dynamic nuclear polarization (Norton et al., 1986), which showed that Trp33 and Trp45 were accessible to the water-soluble flavin dye used in these experiments, whereas Trp23 had limited access. In our structures the accessibilities of the entire indole rings were in the order Trp45 > Trp23 ≥ Trp33, whereas the accessibilities of the indole NH groups were Trp33 > Trp45 > Trp23, with the Trp23 NH being completely inaccessible in all of the structures. This suggests that the accessibility of the indole NH may be of greater importance in the development of polarization in these experiments than the accessibility of the ring as a whole.

The aromatic rings of Trp23 and Trp33 are oriented such that the indole NH of Trp23 lies directly above the six-membered ring of Trp33. This no doubt accounts for the remarkable upfield shift of the Trp23 NH, from its expected position of 10.2 (Wüthrich, 1986) to 5.92 ppm. Burley and Petsko (1986) have shown that amino groups of Lys, Arg, Asn, and Gln and imidazole NH groups of His are often found above and below aromatic rings and close to their center, where they make energetically favorable interactions with the π-electrons. The interaction of the indole NH of Trp23 with the π-orbital of the six-membered ring of Trp33 probably reflects a specific interaction of this type, supplemented by edge-to-face interactions of the two indole rings (Burley & Petsko, 1986). The mean distance of the Trp23 nitrogen from the centre of the six-membered ring of Trp33 is 3.1 Å, close to the optimal distance from an amino nitrogen to a ring centre (Levitt & Perutz, 1988), and the NH group points directly toward the ring.

Comparison with Other Long Sea Anemone Polypeptides. The overall fold of AP-A is very similar to those of the type 1 toxin ATX Ia (Widmer et al., 1989) and the type 2 toxin Sh I (Fogh et al., 1990; Wilcox et al., 1993), as shown in Figure 9. Mean pairwise rms differences between the 20 final structures of AP-A and the eight published structures of ATX Ia and the 12 minimized 'new' structures of Sh I are 1.48 and 1.54 Å, respectively, when superimposed over the backbone heavy atoms of residues 2−6, 20−24, 27−36, and 45−49 in AP-A, which correspond to 2−6, 20−34,

and 42−46 in ATX Ia and 1−5, 19−33, and 42−46 in Sh I. The corresponding value for Sh I vs ATX Ia is 1.65 Å.

The main difference between AP-A and ATX Ia is in the region of residues 36−44, due primarily to a bend in the backbone of AP-A caused by the β-bulge at residues 36−37, which is not present in ATX Ia (Figure 9). When residues corresponding to 37−40 and 42−44 of AP-A are included in the superpositions, rms differences of 2.54 and 1.81 Å are found vs ATX Ia and Sh I, respectively, and 2.76 Å between ATX Ia and Sh I, emphasizing the divergence between ATX Ia and the other two structures in this region. However, as noted previously (Fogh et al., 1990), the more extended structure of ATX Ia may reflect the inclusion of an inappropriate hydrogen bond restraint during one stage of the structure calculations (Widmer et al., 1989). A β-bulge is also found in Sh I at the corresponding position (residues 33−34) (Fogh et al., 1990; Wilcox et al., 1993); as a result, the structures of Sh I and AP-A coincide more closely than those of ATX Ia and AP-A. The pattern of slowly-exchanging amide protons in this region is similar for all three proteins, making it likely that the local structures are similar.

The overall structure of the β-sheet is similar in all three molecules, but there are some local differences. In AP-A the N-terminal strand of the sheet has four hydrogen bonds to the second strand, whereas in ATX Ia the 2→22 bond was not present, and in Sh I neither the 2→22 nor the 20→4 was found. By contrast, the 48→34 bond, which is not found in AP-A because of a slight twist of the backbone around His34, was present in both ATX Ia and Sh I. As mentioned above, AP-A and Sh I have a β-bulge at residues 36−37, which includes 36→46 and 37→46 hydrogen bonds. In ATX Ia the 37→46 bond was not observed but a 46→36 bond was observed in some of the structures, reflecting its inclusion as a restraint during distance geometry calculations (see above). All three molecules have a well-defined β-turn with a 33→30 bond preceding strand III of the sheet, but only AP-A and Sh I have the 7→36 bond linking the start of the first loop with strand III. Sh I differs from AP-A and ATX Ia at the end of strand II where a 25→43 bond is found, effectively lengthening the central strands of the β-sheet, and in the loop joining strands III and IV, where a β-turn with a 44→41 bond occurs. AP-A and ATX Ia also differ from one another in this region because of the Tyr25-Pro26 insert in AP-A, which is associated with a β-turn and a 28→25 bond in all of the structures, and the deletion of Pro41 in ATX Ia, which removes the major source of conformational heterogeneity in AP-A. Thus, there are some genuine differences among the structures in this region. In general, however, AP-A appears to be as similar to Sh I as it is to ATX Ia, emphasizing that there are only local differences between type 1 and type 2 polypeptides and that their global structures are the same. In making these comparisons it should be borne in mind that the AP-A and Sh I structures have been determined with more extensive restraint sets than for ATX Ia.

None of the molecules has well-defined hydrogen bonds in the region around the end of strand I of the sheet and the start of strand II. In all three the amide of Cys6 is slowly exchanging and experiences a significant downfield shift, and from initial analyses of the secondary structure of ATX Ia (Widmer et al., 1988) and Sh I (Fogh et al., 1989), although not AP-A (Mabbutt & Norton, 1990), it appeared that Cys6 was hydrogen bonded to residue 17. The amide resonance of Ser/Thr19 in all three molecules is also shifted downfield significantly, although a hydrogen-bonding partner could not be specified. Subsequent tertiary structure determinations for ATX Ia and Sh I showed that several hydrogen-bonding interactions were possible for these two amides, leading Widmer et al. (1989) to suggest that this region of ATX Ia may be involved in a dynamic equilibrium among several possible conformations. A similar situation appears to exist in AP-A,[4] with neither Cys6 nor Ser19 forming stable hydrogen bonds. Although the backbone angular order parameters for both residues are well-defined (Figure 1), dynamic processes in this region may contribute to the hingelike motion of the 8−16 loop referred to above.

Finally, we note that the hydrophobic core in AP-A (Figure 3) and ATX Ia (Widmer et al., 1989) is depleted in Sh I due to the substitution of Trp23 and Trp/Tyr45 by Asp residues. This may contribute to the reduced thermal stability of Sh I (Norton et al., 1989) relative to AP-A and ATX Ia (Gooley et al., 1988).

Comparison with Other Polypeptide Structures. It has been noted previously that there are some structural similarities between the long sea anemone polypeptides and the 43-residue polypeptide BDS I (Driscoll et al., 1989; Widmer et al., 1989; Llewellyn & Norton, 1991). Compared with the short sea anemone polypeptide ATX III (Norton, 1991), there are no overall structural similarities, but now that high-resolution structures are available for AP-A and ATX III (Manoleras & Norton, 1994), it should be possible to undertake a detailed comparison of their molecular surfaces in an effort to identify common features. Possible similarities with the α-scorpion toxins, which appear to bind to the same site on the voltage-gated sodium channel (Catterall & Beress, 1978), have also been discussed (Norton, 1991).

A less obvious relationship is with a structural motif (designated the inhibitor−cystine knot motif) identified recently in a series of disulfide cross-linked polypeptides represented by ω-conotoxin GVIA and several proteinase inhibitors (Pallaghy et al., 1994). These polypeptides contain a triple-stranded antiparallel β-sheet with a 2x,-1 topology, as well as a disulfide knot. Although the disulfide knot is not present in the sea anemone polypeptides, the topology of strands II−IV of the sheet in AP-A is the same as in the inhibitor cystine knot motif, with the 23→45, 47→21, and 36→46 hydrogen bonds of the sheet in AP-A (Figure 5) aligning with the 8→24, 26→6, and 20→25 bonds in ω-conotoxin GVIA (Sevilla et al., 1993; Davis et al., 1993; Skalicky et al., 1993; Pallaghy et al., 1993, 1994).

Essential Residues in AP-A. Chemical modification studies on AP-A (Newcomb et al., 1980; Gruen & Norton, 1985) and ATX II (Barhanin et al., 1981) indicate that one or both of the carboxylate groups of Asp7 and Asp9 is essential for activity. The significance of these groups may reflect the fact that they interact directly with the sodium channel or that they are essential in maintaining the tertiary structure, or a combination of both. In the structure these residues

[4] Characteristic $d_{NN}(4,20)$ and $d_{\alpha\alpha}(5,18)$ NOEs from this region in ATX Ia (Widmer et al., 1988) and Sh I (Fogh et al., 1989) were not observed in AP-A by Mabbutt and Norton (1990). In the current data the $d_{NN}(4,20)$ NOE was observed but the $d_{\alpha\alpha}(5,18)$ NOE was not because of the near degeneracy of the $C^{\alpha}H$ resonances. A $d_{\alpha\alpha}(20,48)$ NOE observed in ATX Ia but not Sh I was present but weak in current spectra of AP-A.

are part of a type I β-turn at the start of the Arg14-containing loop. In the type 2 polypeptide Sh I, the homologous residues Asp6 and Glu8 are also located in a type I β-turn (Wilcox et al., 1993) and are also essential for neurotoxic activity (Pennington et al., 1990). Although the structure of AP-A is not as well-defined in the region of these Asp residues as in the core of the molecule, and the locations of several side chains on the surface of the molecule are also poorly defined (reflecting, at least in some cases, genuine conformational disorder in solution), it is clear that the carboxyl of Asp7 is close enough to Lys37 to form a salt bridge and that it may also interact with His39. Asp9 is nearby, and Lys48 is on the same face of the protein (Figure 8). The fact that modification of the Asp carboxylate or Lys ammonium groups (Gould & Norton, 1995) destroys cardiac stimulatory activity implies that this region of AP-A forms part of the sodium channel binding surface. It is noteworthy that in both ATX Ia and Sh I, which are potent crustacean neurotoxins but essentially inactive as mammalian cardiac stimulants, Lys37 and both histidines are missing, suggesting that one or more of these side chains may be important in conferring specificity for the mammalian cardiac sodium channel. Arg14, which previously had been thought to be important, appears to have only a minor role in the activity of both AP-A (Gould & Norton, 1995) and AP-B (Khera & Blumenthal, 1994). From the structures presented here it appears that the loss of cardiac stimulatory activity associated with tryptic cleavage adjacent to Arg14 (Gould et al., 1990) may be an indirect effect mediated by conformational changes around the Asp residues near the start of the loop. The structures obtained in this work will provide the basis for interpretation of site-directed mutagenesis and other studies aimed at identifying all of the residues essential for activity.

ACKNOWLEDGMENT

We thank David Smith for the program SSTRUC, David Craik for the program COUPLING, and Richard Ford and Lisa Cowen for assistance with computing.

REFERENCES

Anil-Kumar, Ernst, R. R., & Wüthrich, K. (1980) *Biochem. Biophys. Res. Commun. 95*, 1−6.

Barhanin, J., Hugues, M., Schweitz, H., Vincent, J.-P., & Lazdunski, M. (1981) *J. Biol. Chem. 256*, 5764−5769.

Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., & Tasumi, M. (1977) *J. Mol. Biol. 112*, 535−542.

Blair, R. W., Peterson, D. F., & Bishop, V. S. (1978) *J. Pharmacol. Exp. Ther. 207*, 271−276.

Braunschweiler, L., & Ernst, R. R. (1983) *J. Magn. Reson. 53*, 521−528.

Brooks, B. R, Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., & Karplus, M. (1983) *J. Comput. Chem. 4*, 187−217.

Brünger, A. T. (1992) *X-PLOR Version 3.1. A System for X-ray Crystallography and NMR*, Yale University, New Haven, CT.

Bundi, A., & Wüthrich, K. (1979) *Biopolymers 18*, 299−311.

Burley, S. K., & Petsko, G. A. (1986) *FEBS Lett. 203*, 139−143.

Catterall, W. A., & Beress, L. (1978) *J. Biol. Chem. 253*, 7393−7396.

Chan, A. W. E., Hutchinson, E. G., Harris, D., & Thornton, J. M. (1993) *Protein Sci. 2*, 1574−1590.

Davis, J. H., Bradley, E. K., Miljanich, G. P., Nadasdi, L., Ramachandran, J., & Basus, V. J. (1993) *Biochemistry 32*, 7396−7405.

Driscoll, P. C., Gronenborn, A. M., Beress, L., & Clore, G. M. (1989) *Biochemistry 28*, 2188−2198.

Fogh, R. H., Mabbutt, B. C., Kem, W. R., & Norton, R. S. (1989) *Biochemistry 28*, 1826−1834.

Fogh, R. H., Kem, W. R., & Norton, R. S. (1990) *J. Biol. Chem. 265*, 13016−13028.

Gooley, P. R., & Norton, R. S. (1986) *Biochemistry 25*, 2349−2356.

Gooley, P. R., Blunt, J. W., & Norton, R. S. (1984) *FEBS Lett. 174*, 15−19.

Gooley, P. R., Blunt, J. W., Beress, L., & Norton, R. S. (1988) *Biopolymers 27*, 1143−1157.

Gould, A. R., & Norton, R. S. (1995) *Toxicon* (in press).

Gould, A. R., Mabbutt, B. C., & Norton, R. S. (1990) *Eur. J. Biochem. 189*, 145−153.

Gould, A. R., Mabbutt, B. C., Llewellyn, L. E., Goss, N. H., & Norton, R. S. (1992) *Eur. J. Biochem. 206*, 641−651.

Grathwohl, C., & Wüthrich, K. (1976) *Biopolymers 15*, 2025−2041.

Griesinger, C., Sorensen, O. W., & Ernst, R. R. (1987) *J. Magn. Reson. 75*, 474−492.

Gross, G. J., Warltier, D. C., Hardman, H. F., & Shibata, S. (1985) *Eur. J. Pharmacol. 110*, 271−276.

Gruen, L. C., & Norton, R. S. (1985) *Biochem. Intl. 11*, 69−76.

Güntert, P., & Wüthrich, K. (1991) *J. Biomol. NMR 1*, 447−456.

Güntert, P., Braun, W., & Wüthrich, K. (1991) *J. Mol. Biol. 217*, 517−530.

Hinds, M. G., & Norton, R. S. (1993) *J. Protein Chem. 12*, 371−378.

Hyberts, S. G., Märki, W., & Wagner, G. (1987) *Eur. J. Biochem. 164*, 625−635.

Hyberts, S. G., Goldberg, M. S., Havel, T. F., & Wagner, G. (1992) *Protein Sci. 1*, 736−751.

Kabsch, W., & Sander, C. (1983) *Biopolymers 22*, 2577−2637.

Kem, W. R. (1988) in *The Biology of Nematocysts* (Hessinger, D., & Lenhoff, H., Eds.) pp 375−405, Academic Press, New York.

Khera, P. K., & Blumenthal, K. M. (1994) *J. Biol. Chem. 269*, 921−925.

Kodama, I., Toyama, J., Shibata, S., & Norton, T. R. (1981) *J. Cardiovasc. Pharmacol. 3*, 75−86.

Kraulis, P. (1991) *J. Appl. Crystallogr. 24*, 946−950.

Levitt, M., & Perutz, M. F. (1988) *J. Mol. Biol. 201*, 751−754.

Llewellyn, L. E., & Norton, R. S. (1991) *Biochem. Intl. 24*, 937−946.

Mabbutt, B. C., & Norton, R. S. (1990) *Eur. J. Biochem. 187*, 555−563.

Macura, S., Huang, Y., Suter, D., & Ernst, R. R. (1981) *J. Magn. Reson. 43*, 259−281.

Manoleras, N., & Norton, R. S. (1994) *Biochemistry 33*, 11051−11061.

Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967−974.

Newcomb, R., Yasunobu, K. T., Seriguchi, D., & Norton, T. R. (1980) in *Frontiers in Protein Chemistry* (Liu, D. T., Mamiya, G., & Yasunobu, K. T., Eds.) pp 539−550, Elsevier Press, Amsterdam.

Norton, R. S. (1991) *Toxicon 29*, 1051−1084.

Norton, R. S., Norton, T. R., Sleigh, R. W., & Bishop, D. G. (1982) *Arch. Biochem. Biophys. 213*, 87−97.

Norton, R. S., Beress, L., Stob, S., Boelens, R., & Kaptein, R. (1986) *Eur. J. Biochem. 157*, 343−346.

Norton, R. S., Cossins, A. I., & Kem, W. R. (1989) *Biochemistry 28*, 1820−1826.

Norton, T. R. (1981) *Fed. Proc. 40*, 21−25.

Orengo, C. A., & Thornton, J. M. (1993) *Structure 1*, 105−120.

Pallaghy, P. K., Duggan, B. M., Pennington, M. W., & Norton, R. S. (1993) *J. Mol. Biol. 234*, 405−420.

Pallaghy, P. K., Nielsen, K. J., Craik, D. J., & Norton, R. S. (1994) *Protein Sci. 3*, 1833−1839.

Pardi, A., Billeter, M., & Wüthrich, K. (1984) *J. Mol. Biol. 180*, 741−751.

Pennington, M. W., Kem, W. R., & Dunn, B. M. (1990) *Peptide Res. 3*, 228−232.

Platou, E. S., Refsum, H., & Hotvedt, R. (1986) *J. Cardiovasc. Pharmacol. 8*, 459−465.

Rance, M., Sorensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 117,* 479−485.
Richardson, J. S. (1977) *Nature 268,* 495−500.
Rucker, S. P., & Shaka, A. J. (1989) *Mol. Phys. 68,* 509−517.
Sasayama, S. (1992) *Cardiovasc. Drugs Ther. 6,* 15−18.
Scanlon, M. J., & Norton, R. S. (1994) *Protein Sci. 3,* 1121−1124.
Schweitz, H., Vincent, J.-P., Barhanin, J., Frelin, C., Linden, G., Hugues, M., & Lazdunski, M. (1981) *Biochemistry 20,* 5245−5252.
Scriabine, A., van Arman, C. G., Morgan, G., Morris, A. A., Bennett, C. D., & Bohidar, N. R. (1979) *J. Cardiovasc. Pharmacol. 1,* 571−583.
Sevilla, P., Bruix, M., Santoro, J., Gago, F., Garcia, A. G., & Rico, M. (1993) *Biochem. Biophys. Res. Commun. 192,* 1238−1244.
Skalicky, J. J., Metzler, W. J., Ciesla, D. J., Galdes, A., & Pardi, A. (1993) *Protein Sci. 2,* 1591−1603.
Srinivasan, N., Sowdhamini, R., Ramakrishnan, C., & Balaram, P. (1990) *Int. J. Pept. Protein Res. 36,* 147−155.
Tanaka, M., Haniu, M., Yasunobu, K. T., & Norton, T. R. (1977) *Biochemistry 16,* 204−208.
Torda, A. E., & Norton, R. S. (1987) *Biochem. Intl. 15,* 659−666.
Torda, A. E., Mabbutt, B. C., van Gunsteren, W. F., & Norton, R. S. (1988) *FEBS Lett. 239,* 266−270.
Vaughan Williams, E. M. (1975) *Pharmacol. Ther. B 1,* 115−138.
Wagner, G., Pardi, A., & Wüthrich, K. (1983) *J. Am. Chem. Soc. 105,* 5948−5949.
Wagner, G., Braun, W., Havel, T. F., Schaumann, T., Gō, N., & Wüthrich, K. (1987) *J. Mol. Biol. 196,* 611−639.
Widmer, H., Wagner, G., Schweitz, H., Lazdunski, M., & Wüthrich, K. (1988) *Eur. J. Biochem. 171,* 177−192.
Widmer, H., Billeter, M., & Wüthrich, K. (1989) *Proteins 6,* 357−371.
Wilcox, G. R., Fogh, R. H., & Norton, R. S. (1993) *J. Biol. Chem. 268,* 24707−24719.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids,* Wiley, New York.

BI942524F