NMR determination of the major solution conformation of a peptoid pentamer with chiral side chains

Philippe Armand*[†], Kent Kirshenbaum*[‡], Richard A. Goldsmith*[§], Shauna Farr-Jones[‡], Annelise E. Barron*[‡¶], Kiet T. V. Truong*, Ken A. Dill[‡], Dale F. Mierke[∥]**, Fred E. Cohen*[†,††,‡‡], Ronald N. Zuckermann*[§§], and Erin K. Bradley*[¶¶]

*Chiron Technologies, Chiron Corporation, 4560 Horton Street, Emeryville, CA 94608; Departments of [†]Biomedical Sciences, [‡]Pharmaceutical Chemistry, [††]Medicine, and [‡‡]Cellular and Molecular Pharmacology, University of California, San Francisco, CA 94143; [∥]Gustaf H. Carlson School of Chemistry, Clark University, 950 Main Street, Worcester, MA 01610; and **Department of Pharmacology and Molecular Toxicology, University of Massachusetts, Medical Center, 55 Lake Avenue North, Worcester, MA 01655

ABSTRACT   Polymers of N-substituted glycines ("peptoids") containing chiral centers at the α position of their side chains can form stable structures in solution. We studied a prototypical peptoid, consisting of five *para*-substituted (*S*)-*N*-(1-phenylethyl)glycine residues, by NMR spectroscopy. Multiple configurational isomers were observed, but because of extensive signal overlap, only the major isomer containing all *cis*-amide bonds was examined in detail. The NMR data for this molecule, in conjunction with previous CD spectroscopic results, indicate that the major species in methanol is a right-handed helix with *cis*-amide bonds. The periodicity of the helix is three residues per turn, with a pitch of ≈6 Å. This conformation is similar to that anticipated by computational studies of a chiral peptoid octamer. The helical repeat orients the amide bond chromophores in a manner consistent with the intensity of the CD signal exhibited by this molecule. Many other chiral polypeptoids have similar CD spectra, suggesting that a whole family of peptoids containing chiral side chains is capable of adopting this secondary structure motif. Taken together, our experimental and theoretical studies of the structural properties of chiral peptoids lay the groundwork for the rational design of more complex polypeptoid molecules, with a variety of applications, ranging from nanostructures to nonviral gene delivery systems.

Polymers of N-substituted glycines, termed peptoids, form a new class of biocompatible, synthetically accessible heteropolymers. Their sequence-specific, automated, and highly efficient synthesis has allowed the creation of combinatorial libraries of peptoid oligomers for drug discovery (1–3). By using recent improvements in the efficiency of the coupling chemistry, peptoids up to 50 residues in length have been synthesized. Among those, cationic peptoid 36-mers have been identified that bind DNA, protect the DNA from nuclease digestion, and facilitate gene transfection (4).

We recently showed that many peptoids with chiral centers at the side chain α position have strong CD signals, indicating the presence of a regularly repeating conformation, in both aqueous and organic solvents (5). This structure is remarkably stable, as demonstrated by both CD and differential scanning calorimetry (DSC) measurements (5). A model of this conformation was recently proposed based on the results of molecular mechanics and semi-empirical quantum mechanical calculations (6). Modeling predicted that peptoids containing (*S*)-*N*-(1-phenylethyl)glycine residues would form right-handed helices with a periodicity of approximately three

residues per turn and cis-amide bonds, similar to the polyproline type I conformation. In this conformation, the backbone carbonyls are aligned along the long helical axis, providing an explanation for the intensity of the helical CD signal.

Presented here is the structure determination by NMR spectroscopy of the major solution isomer of compound **1**, a pentapeptoid containing five chiral side chains (Fig. 1*A*). This peptoid displays the characteristic helix-like double minimum CD signal described by Kirshenbaum *et al.* (5). The distance restraints obtained by NMR on **1** are best satisfied by a helical conformation that agrees with our previously published prediction (6). Identification of this structural motif and its CD signal should allow more directed studies of peptoid structural properties, and guide the design of more complex folded polypeptoid molecules for use in a variety of biological, chemical, and material science settings.

MATERIALS AND METHODS

Synthesis, Purification, and CD. The peptoid oligomers were synthesized and purified as described (5). The chiral amines used for the oligomer synthesis were obtained as follows: (*S*)-1-(*p*-nitrophenyl)ethylamine, (*S*)-1-(*p*-chlorophenyl)-ethylamine, and (*S*)-1-(*p*-methoxyphenyl)ethylamine were obtained in >99% enantiomeric excess (e.e.) from Celgene (Warren, NJ); (*S*)-1-(*p*-fluorophenyl)ethylamine was obtained by successive recrystallizations of the tartrate salt of the racemic amine. Briefly, the L-tartrate salt of racemic 1-(*p*-fluorophenyl)ethylamine (Aldrich) was recrystallized three times from 9:1 MeOH:H$_2$O to obtain the desired *S* isomer in >99% e.e. as determined by GC of the Mosher amide (7). The ^{13}C-containing peptoids were synthesized by substituting bromoacetic-1,2-^{13}C$_2$ acid (Cambridge Isotope Laboratories, Andover, MA) for bromoacetic acid at the desired position of the growing oligomer. CD studies were performed as described (5).

NMR Data Acquisition. All molecules were lyophilized from HPLC buffer and then relyophilized from deuterated aceto-

---

This paper was submitted directly (Track II) to the *Proceedings* office.
Abbreviations: 1D, one-dimensional; 2D, two-dimensional; DQ-COSY, double quantum filtered correlation spectroscopy; TROESY, transverse rotating-frame Overhauser effect spectroscopy; HMQC, heteronuclear multiple quantum coherence spectroscopy; HMBC, heteronuclear multiple bond coherence spectroscopy; ROE, rotating-frame Overhauser effect; ADR, absent distance restraint.
[§]Present address: Axys Pharmaceuticals, 180 Kimball Way, South San Francisco, CA 94080.
[¶]Present address: Department of Chemical Engineering, Northwestern University, 2145 Sheridan Road, Evanston, IL 60208.
[§§]To whom reprint requests should be addressed. e-mail: Ron_Zuckermann@cc.chiron.com.
[¶¶]Present address: CombiChem Inc., 1804 Embarcadero Road, Suite 201, Palo Alto, CA 94303.
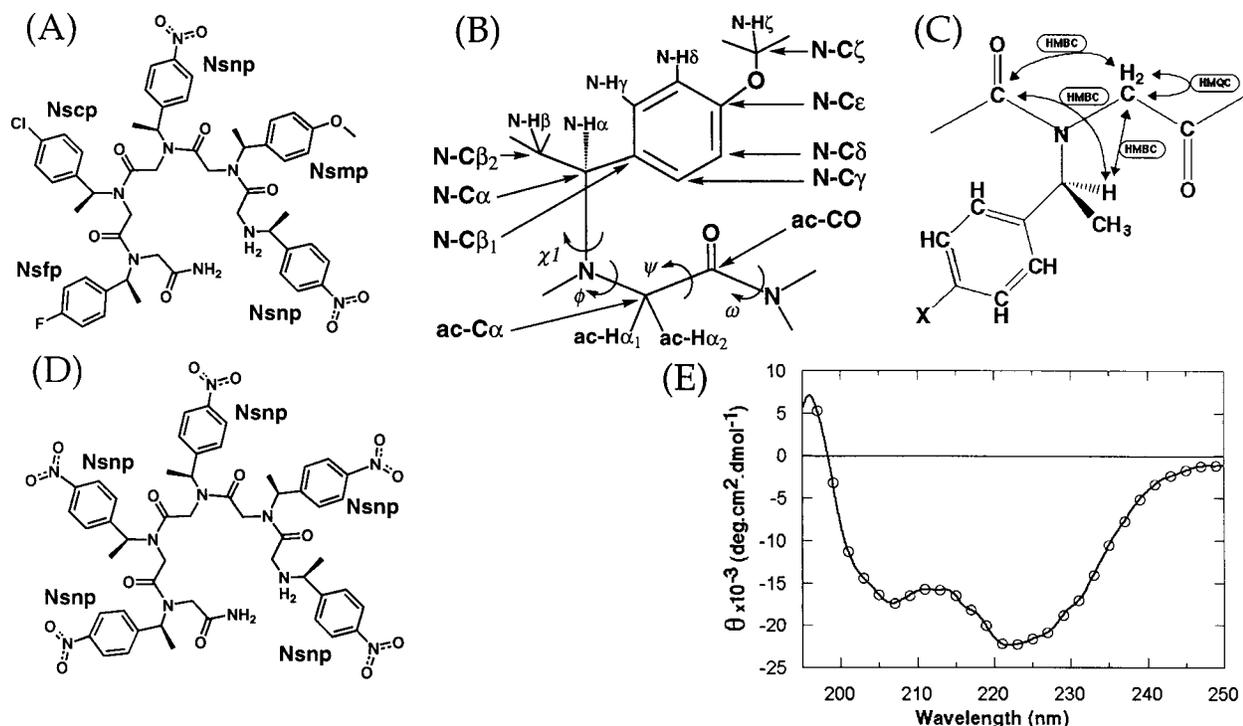
FIG. 1. (*A*) Compound **1** (*N*snp¹-*N*smp²-*N*snp³-*N*scp⁴-*N*sfp⁵). *N*snp = (*S*)-*N*-(1-(*p*-nitrophenylethyl))glycine, *N*smp = (*S*)-*N*-(1-(*p*-methoxyphenylethyl))glycine, *N*scp = (*S*)-*N*-(1-(*p*-chlorophenylethyl))glycine, *N*sfp = (*S*)-*N*-(1-(*p*-fluoro-phenylethyl))glycine. The amide bonds are drawn in their *cis* configurations. (*B*) Atom and dihedral nomenclature. These are shown for *N*smp. The nomenclatures for the other residues are identical except for the absence of *N*smp-C$^\zeta$ and *N*smp-H$^\zeta$ atoms. Atom names follow (15); however, the side chain name subscript has been dropped when not referring to a particular residue (e.g., N-C$^\alpha$ instead of Nxxx-C$^\alpha$) and has been omitted in this figure for clarity. Dihedral angles are named by analogy to those of peptides. (*C*) Simplified schematic representation of the assignment process for part of **1**. Arrows indicate the connections that were used and the spectrum types with which they were established. (*D*) Pentamer of (*S*)-*N*-(1-(*p*-nitrophenylethyl))glycine [(*N*snp)₅]. (*E*) CD spectrum of **1** in 100% methanol. The data were collected at 10°C and at a concentration of 0.1 mM.

nitrile-*d*₃ and then from deuterated methanol-*d*₄ before being dissolved in NMR solvents. NMR was performed at a peptoid concentration of ≈5 mM in methanol-*d*₄ at 10°C. Spectra were acquired on a Varian Unity 300. Standard one-dimensional (1D) proton spectra were collected as 8,192 data points and zero-filled to 16,384. The 1D spectra were also taken at 100 μM, 1 mM, and 10 mM to verify that spectra (chemical shift and populations) were concentration independent (data not shown). Two-dimensional (2D) homonuclear [double quantum filtered correlation spectroscopy (DQCOSY) (8, 9) and transverse rotating-frame Overhauser effect spectroscopy (TROESY) (10)] and heteronuclear [heteronuclear multiple quantum coherence spectroscopy (HMQC) (11) and heteronuclear multiple bond coherence spectroscopy (HMBC) (12)] spectra were collected with the following parameters. The HMQC experiments were performed with *J* = 140 Hz and the HMBC experiments with *J* = 6, 9, and 12 Hz. The heteronuclear experiments were collected with spectral widths of 4,000 Hz in the ¹H dimension, and 16,000 Hz and 18,000 Hz in the ¹³C dimension for the HMQC and the HMBC, respectively. The rotating-frame Overhauser effect (ROE) data were collected by using the TROESY pulse sequence of Hwang and Shaka (10) with a 250- and 300-ms mixing time. The spinlock pulse was 180°(*x*) 180°(−*x*), with a field strength of 8,000 Hz. The numbers of *t1* increments, transients, and *t2* complex data points were, respectively, 512, 256, and 2,048 for the TROESY and natural abundance HMBC spectra; 256, 16, and 2,048 for the HMQC spectra of the labeled compounds; and 512, 64, and 2,048 for the HMBC spectra of those labeled compounds. Spectral widths of 4,000 Hz were used in both dimensions for the TROESY and DQCOSY experiments. Data processing was carried out with the NMR PACK software (13). Heteronuclear spectra were zero-filled to give a final real data matrix

of 1,024 points in ω2 (¹H) and 2,048 points in ω1 (¹³C). The HMQC was processed as phase sensitive in both dimensions, and the HMBC was processed as phase sensitive in ω2 and absolute value in ω1. Visualization of spectral data and resonance assignments were carried out with the NMR PACK (13) and SPARKY 3 (14) software.

Resonance Assignments. A modification of the strategy published by Bradley (15, 16) was employed for assigning ¹³C and ¹H resonances. Each peptoid residue has two spin systems, the side chain branching from the main chain nitrogen, "Nxxx", and the main chain glycine unit, "ac" (Fig. 1*B*). Although it is theoretically possible to assign these spin systems from natural abundance spectra, the dispersion of the ¹³C and ¹H chemical shifts of the main chain ac subunits was too small to permit this. Therefore, five derivatives of compound **1**, each containing a pair of ¹³C-labels (C$^\alpha$ and CO) in a single ac subunit, were used to obtain those assignments.

The ac spin systems for each residue were assigned first by using the ¹³C-HMQC spectra of the five individual molecules, each one labeled with ¹³C at a different ac-C$^\alpha$ position. The main chain subunits were then connected to the side chains with the ¹³C-HMBC spectra of the same individually labeled molecules. The connections used for the assignment procedure are shown schematically in Fig. 1*C*, and the heteronuclear spectra used to assign the main chain and side chain α protons of residue 1 are shown in Fig. 2. In the HMQC of the ¹³C-labeled compounds, the cross peaks involving the labeled ac-C$^\alpha$ positions are much more intense than the peaks for carbons at natural abundance. Thus even with the significant chemical shift degeneracy seen for the main chain carbons and protons, the assignments can be made unambiguously. The same is true for the side chain H$^\alpha$ which is the only proton connected through three bonds to the ¹³C-labeled ac-C$^\alpha$ (Fig. 1*C*).
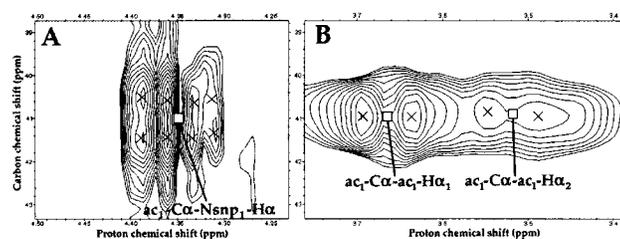
FIG. 2. Detail of the heteronuclear spectra used to assign the resonances of the main chain and side chain α protons of residue **1**. The spectra were collected on a molecule containing ^{13}C labels at the ac^1-C^α and ac^1-CO positions. (*A*) HMBC spectrum showing the ac^1-C^α-Nsnp^1-H^α cross peak for the major isomer. (*B*) HMQC spectrum showing the ac^1-C^α-Nsnp^1-H^α cross peaks for the major isomer.

Although the experiments with the label at the C^α position clearly established the sequential assignments, the main chain connections were further validated by the observation of cross peaks in the HMBC spectra for the three-bond connection between the ^{13}C carbonyl carbons and the following Nxxx-H^α across each amide bond (Fig. 1*C*). The side chains themselves were assigned with a combination of DQCosy and natural abundance HMBC spectra.

Refinement. All unambiguously assigned ROE cross peaks involving the major isomer were used to establish interproton distance bounds. Distance bounds were obtained by assuming an r^{-6} dependence of the cross peak intensities, calibrated on a set of intra-residue side chain cross peaks (H^α-H^β and H^γ-H^δ), and adjusted for the use of pseudoatoms (17). This yielded a set of 15 distance restraints (Table 1), of which 7 were inter-residue. Refinement was attempted with this set of data without success; most of the structures generated with this restraint set had short inter-proton distances, which conflicted with the absence of the corresponding cross peaks in the TROESY spectra. To generate structures that better reflected the experimental data, a set of 51 absent distance restraints (ADRs) was added, in a manner similar to that described in ref. 18; each ADR involved proton pairs that showed no cross peak in the TROESY spectra. Only pairs for which both protons appeared in other ROE cross peaks were included. This ensured that the absence of an ROE reflected a large inter-proton distance and not other factors such as exchange or relaxation processes. The ADRs were entered by setting the lower distance bounds at 3.0 Å. These two sets of restraints were combined and used in metric matrix distance geometry

calculations following the random metrization algorithm of Havel (19). A total of 1,800 conformations that satisfied all distance bounds were generated. Each of these conformations was then subjected to constrained minimization, performed with the DISCOVER program by using the CFF91 force field (Biosym/MSI, San Diego) without cross terms or charges, gradually scaling up the nonbond terms. All conformations that violated the ADRs or that violated the ROE restraints by more than 0.5 Å were then excluded. Structures were visualized by using the INSIGHT II 95.0 molecular modeling system (Biosym/MSI). Ensemble calculations were run as described in ref. 20.

RESULTS

Selection of the Molecule. Compound **1** (Fig. 1*A*) was chosen because it shows the helix-like double minimum signal in its CD spectrum and because its 1D proton NMR spectrum in methanol made it a promising candidate for structure determination. Methanol was used as the solvent because **1** is not soluble in water at the high concentrations required for NMR studies. The atom and dihedral angle nomenclature used in this report is given in Fig. 1*B*. Fig. 1*E* shows the CD spectrum of **1** in methanol.

Data Collection and Resonance Assignments. Both homonuclear and heteronuclear spectra of **1** in methanol-d_4 were collected and used to assign the ^{1}H and ^{13}C chemical shifts of the major solution isomer. The presence of additional minor peaks for each proton was noted in all spectra (HMQC, HMBC, DQCosy, TROESY). This finding points to the existence of multiple species, most likely configurational isomers with different amide bond geometries. There appear to be two such minor isomers that are appreciably populated. Based on the normalized integration in a 1D spectrum of the peak assigned to the major isomer's Nsnp^3-H^γ atom, it is estimated that this major isomer represents between 50% and 60% of the molecules in solution.

The presence of those minor isomers posed the problem of unambiguously assigning all the peaks belonging to the major species. Both proton 1D peak intensities, which reflect relative populations, and TROESY intensities for the fixed distance intra-side chain cross peaks (H^α-H^β, H^α-H^γ, and H^β-H^γ) were used to confirm that the peaks belonged to the major isomer (Table 2). Attempts to identify solvent and temperature conditions that would eliminate the minor isomers were unsuccessful. Attempts were also made to assign the minor isomers with the use of exchange cross peaks from TROESY spectra acquired at higher temperatures, but the data were insufficient for a full assignment of the minor species. Thus, great care was

Table 1. List of distance restraints used in structure refinement

Residue 1	Atom 1	Residue 2	Atom 2	Upper bound, Å
Intra-residue				
ac^2	H^α	Nsmp^2	H^α	4.25
ac^2	H^α	Nsmp^2	H^β	3.90
ac^2	H^α	Nsmp^2	H^γ	4.05
ac^3	H^α	Nsnp^3	H^β	3.50
ac^3	H^α	Nsnp^3	H^γ	4.05
ac^5	H^α	Nsfp^5	H^α	4.05
ac^5	H^α	Nsfp^5	H^β	3.90
ac^5	H^α	Nsfp^5	H^γ	3.90
Short range (i to i+1)				
Nsmp^2	H^γ	ac^3	H^α	3.40
Nsmp^2	H^δ	ac^3	H^α	5.50
Nsmp^2	H^β	Nsnp^3	H^δ	6.50
Medium range (i to i+2, i to i+3)				
Nsmp^2	H^ζ	Nsfp^5	H^β	3.90
Nsmp^2	H^δ	Nsfp^5	H^β	4.25
Nsnp^1	H^δ	Nsnp^3	H^β	5.25
Nsnp^1	H^γ	Nsnp^3	H^β	8.00

Table 2. Chemical shifts (in ppm) for the major isomer of **1**

Residue	H^α	H^{β_2}	H^γ	H^δ	H^ζ
Side chain					
Nsnp^1	4.37	1.47	7.51	8.04	
Nsmp^2	5.57	1.13	6.91	6.51	3.31
Nsnp^3	5.56	1.08	7.22	7.87	
Nscp^4	5.37	1.00	6.78	6.92/7.00	
Nsfp^5	5.29	0.95	6.90	6.71	

	H^α	C^α	CO
Main chain			
ac^1	3.52/3.59	40.9	160.8
ac^2	3.58	39.3	165.0
ac^3	3.68	40.5	163.8
ac^4	3.69	39.1	164.3
ac^5	3.38	38.5	ND

Note that peak overlap prevented us from assigning any cross peak involving the sidechain protons of residue 4. ND, not determined.

taken to use only unambiguously assigned cross peaks of the major isomer in the refinement.

Once the resonances of the ac-H^α and side chain H^α were assigned for the major species in solution, it was possible to determine the geometry about each of the main chain amide bonds from the TROESY spectra. In this work, *cis*-amide bonds will refer to geometries where the main chain C^α atoms are *cis* to each other. For residues with an intervening *cis*-amide bond, ROE cross peaks are expected between the main chain protons of the two residues. Unfortunately, for the pentamer studied here, the main chain protons are almost degenerate and the resulting cross peaks are too close to the diagonal to be distinguishable except for residues 4 and 5, where the expected peak is present. If the amide bond were trans, ROE cross peaks would be present between the main chain protons of the residue preceding the amide bond and one or more of the side chain protons (H^α, H^β, or H^γ) of the following residue. Those cross peaks would occur in the spectra at positions far off the diagonal, where they would be clearly visible. Yet those cross peaks are definitely absent from the spectra for the major isomer of **1**. It can therefore be concluded that the amide bonds of the major isomer of **1** all have a *cis* geometry.

Refinement. The refinement process yielded a set of 39 conformations with all *cis*-amide bonds satisfying all ADRs and all ROEs to within 0.5 Å. Those conformations were analyzed in terms of conformational clusters. The term "cluster" will be used to refer to a group of conformations whose maximal pairwise root mean square (rms) distance for the backbone heavy atoms and the side chain α carbons of the internal residues (residues 2–4) is less than 1.4 Å. The 39 structures can be described as a single cluster, cluster A (Fig. 3*A*). If a less stringent fit to the NMR data is allowed, by increasing the ROE violation tolerance to 1.0 Å, an additional 101 structures are obtained, which have higher penalty function values. Of those 101 structures, 29 belong to cluster A. The other 72 structures fall into 7 clusters, clusters B–H, comprising 22, 21, 16, 7, 3, 2 and 1 structures, respectively.

Fig. 4*A* shows the distribution of the proper backbone dihedral angles of cluster A and clusters B–H. For the internal residues, the contour lines of the potential of an octamer of *N*snp, obtained from molecular mechanics (6), are superimposed on the plots. Points that lie on the upper right-hand side of the maps represent right-handed conformations. The dihedral angles of residues 2 and 3 are more tightly clustered than those involving residues 4 and 5. This is not surprising because there is more contact information on the N-terminal residues than on the C-terminal ones. The uncertainty about the C-terminal conformation (Figs. 3*A* and 4*A*) could therefore be caused by true conformational flexibility or to a lack of ROE restraints.

The information content of our data was assessed in a series of refinements with partial restraint sets. A set consisting of the four medium-range restraints (*N*snp³H$^\beta$-*N*snp¹H$^\delta$, *N*snp³H$^\beta$-*N*snp¹H$^\gamma$, *N*sfp⁵H$^\beta$-*N*smp²H$^\delta$, and *N*sfp⁵H$^\beta$-*N*smp²H$^\zeta$) was sufficient to define the general fold and periodicity of the internal three residues. However, it could define neither the handedness nor the individual backbone dihedral angle values. The addition of the ADRs was required to narrow the distribution of the ϕ/ψ values, whereas the addition of the short-range distance restraints was necessary to detect a preference for the right-handed conformations.

Other Isomers. We have only described the major isomer of **1**. However, it is clear from the NMR spectra that there are multiple species in slow exchange in solution. It is likely that they result from *cis-trans* isomerizations about the amide bonds (and therefore represent configurational isomerization). This is clearly the case for one of the minor isomers, where ROE cross peaks are found between a resonance of *N*sfp$^5_{minor}$-H^α and ac$^4_{minor}$-H^α, as well as between the corresponding *N*sfp$^5_{minor}$-H^β
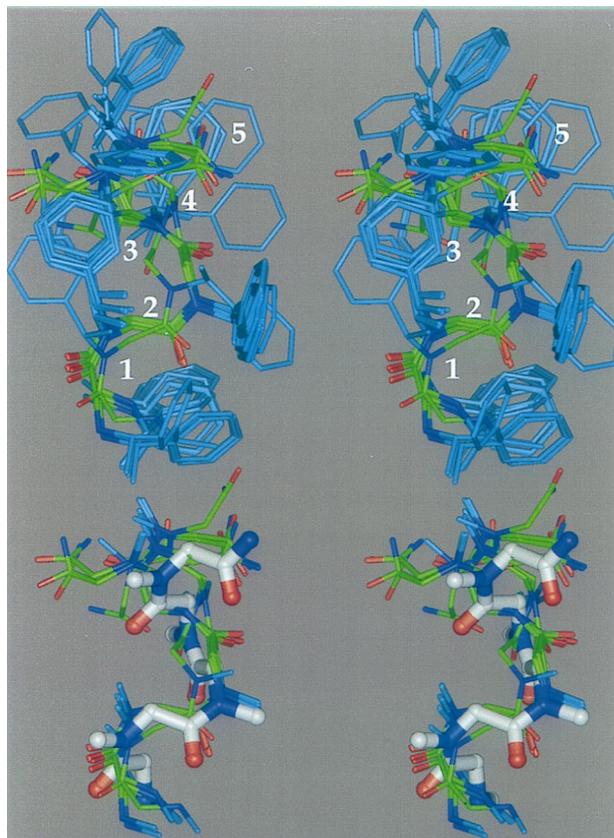

Fig. 3. Stereo representations of the conformations remaining after restraint violation analysis. (*Upper*) Cluster A. This describes the major conformation of **1** in methanol. Backbone carbons are shown in green, side chain carbons in light blue, nitrogen atoms in dark blue, and oxygen atoms in red. The *para* substituents on the aromatic rings and the hydrogen atoms have been omitted for clarity. Residue numbers are shown in white. (*Lower*) Comparison of cluster A conformations with previously published model (6). Only backbone atoms and side chain α carbons are shown. The model is shown as a thicker ball and stick molecule, and its carbon atoms are colored white.

and ac$^4_{minor}$-H^α. Because both the *N*sfp$^5_{minor}$-H^α and the ac$^4_{minor}$-H^α could be unambiguously assigned in the HMBC spectra of the compounds labeled at residues 4 and 3, these two cross peaks must reflect a trans-amide bond at that position.

Ensemble Calculations. The application of the ROE distance restraints to each and every structure of an ensemble is not appropriate for cases in which conformational dynamics are taking place rapidly on the NMR time scale. One way to address this situation is by ensemble dynamics (20), in which the experimental observations are fulfilled by the average over the ensemble: each member of the ensemble must satisfy the constitutional constraints from the structural formula, whereas only the average over the ensemble must fulfill the experimental measurements. This method has been previously used to characterize equilibria between different conformations (20–22).

An ensemble calculation was performed on a set of 780 structures, containing 20 copies of each structure in cluster A. All structures were allowed to independently explore conformational space, with only the requirement that the average pairwise distances must satisfy the distance restraints described above. This yielded a set of conformations compatible with the experimental data under a fast exchange regime. Fig. 4 compares the distribution of values for the ϕ and ψ dihedrals of residues 2–5 before and after the ensemble calculation. It is clear that the regularity in the backbone dihedral angles of the starting conformations is decreased by the ensemble treat-
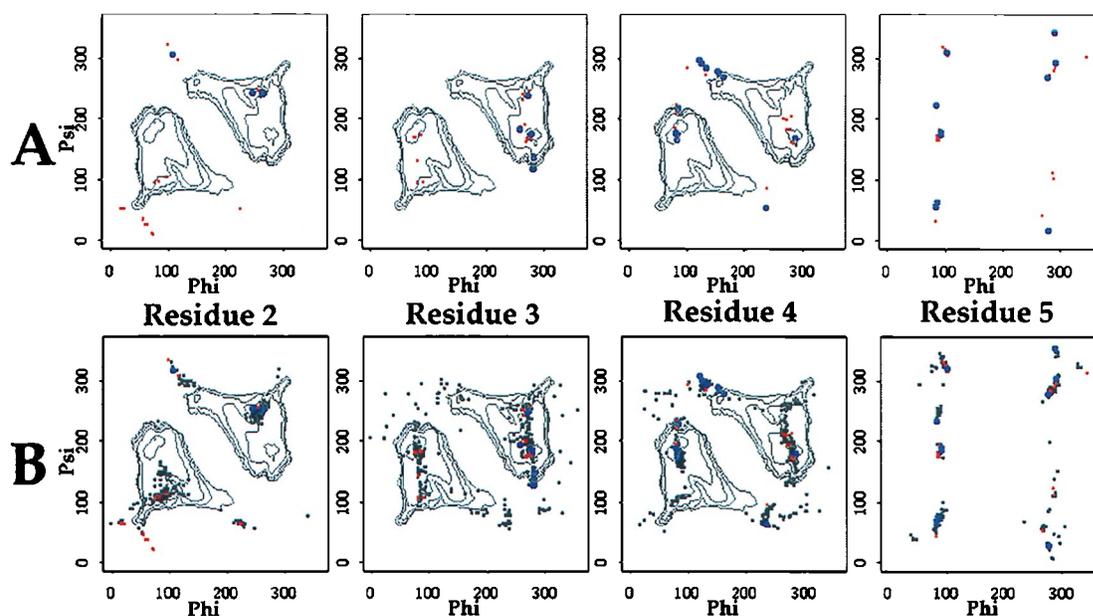
FIG. 4.   Distribution of the values of the main chain dihedral angles (ϕ and ψ) for residues 2 to 5 (residue **1** is not included because there is no proper ϕ dihedral angle). (*A*) Cluster A (large blue circles) and clusters B–H (small red circles) after minimization and before ensemble calculation. (*B*) Same, with cluster A angles after ensemble calculation (small black circles). White contour lines for residues 2–4 indicate the potential surface for *N*snp oligomers, after ref. 6.

ment. Nonetheless, a cluster analysis reveals that 354 of the 780 structures (45%) still belong to cluster A. The remainder of the ensemble falls into 29 distinct clusters. Of these, 2 contain about 10% of the structures each, whereas the rest contain <5% and typically <1% of the structures each. The predominance of cluster A conformations is not the result of limited conformational sampling; indeed, when a control calculation was done without any distance restraints, fewer than 10% of the resulting structures belonged to cluster A. Those results imply that the experimental data are most compatible with a highly populated conformational family (cluster A), validating the results of the standard refinement method.

DISCUSSION

During initial investigations into the biophysical characteristics of peptoids, a wide range of chiral molecules of varying length, charge, and polarity were synthesized. Many of those molecules display a double minimum around 205 and 220 nm in their CD spectrum. This signal was observed in both aqueous and organic solvents, and molecules that could be tested in both types of solvents showed no significant change in their CD spectrum in going from one to the other. Because this signal reflects the presence of a regularly repeating backbone conformation (5), it is likely that all molecules that display this double minimum adopt a similar backbone conformation. To determine this conformation by NMR, molecules with this CD signal were screened for chemical shift dispersion and sharp line widths in their 1D proton NMR spectra. A molecule that satisfied these criteria was a pentamer of (*S*)-*N*-[1-(*p*-nitrophenylethyl)]glycine (*N*snp) (Fig. 1*D*). By analogy to peptides (20), a molecule of such short length could be expected to populate multiple conformations in solution. Nevertheless, this length was sufficient to give rise to a strong double minimum CD signal (5), indicating that the backbone conformation of interest was significantly populated. Moreover, the molecule was small enough to allow the resonance assignments necessary for 2D NMR structure determination. Because initial 2D NMR studies showed overlap in the aromatic region of the spectra, various substitutions were introduced at the *para* position of the side chain aromatic rings to

increase chemical shift dispersion. This ultimately resulted in the selection of compound **1** (Fig. 1*A*) for this work.

Both standard and ensemble refinement methods suggest that the major conformation of **1** is accurately described by a single conformational cluster (cluster A, Fig. 3*A*). This cluster provides the most consistent fit to the experimental restraints. In addition, for most members of cluster A, the backbone dihedral angles of the internal residues (residues 2–4) are closely distributed around 270°, giving rise to regular, right-handed backbone conformations (Figs. 3*A* and 4*A*). This regular repeat is consistent with the ordering of the backbone chromophores demonstrated by the CD data (5). By contrast, most members of clusters B–H have irregular backbones that are discordant with the intense, helix-like CD spectrum. Finally, the ensemble calculations show that cluster A conformations remain by far the most prevalent even if other conformations are allowed to contribute to the experimental fit. This validates the assumption that the ROE cross peaks which were assigned to the major isomer can, to a first approximation, be treated as coming from a single conformation. Indeed, if another conformer were significantly populated in solution and rapidly exchanging with cluster A conformations, the ensemble calculation would have revealed its presence. Because the cluster analysis of the ensemble does not indicate the presence of another important conformer, the ROE data most likely reflect a large population of tightly clustered conformations.

Based on those results, the major conformation of **1** in methanol can be described as having a helical shape, a right-handed twist, a periodicity of three residues per turn, and a pitch of $\approx$6 Å. The amide bonds are all *cis*. The χ_1 angles (defined as the $C_{(i-1)}, N_{(i)}, N\text{-}C^\alpha_{(i)}, N\text{-}C^{\beta_1}_{(i)}$ dihedral angles) of residues 2–5 are all approximately $-120°$. This conformation is very similar to the model previously proposed (6) for an oligomer of (*S*)-*N*-(1-phenylethyl)glycine (Fig. 3*B*). Indeed, the handedness, periodicity, pitch, amide bond geometries, and χ_1 angles were all accurately predicted. The average backbone rms difference between the model and the conformations in cluster A is 1.6 Å, or 1.1 Å if only the internal residues are considered.

The simple molecular mechanics and more detailed quantum mechanical calculations used in the modeling suggested that steric terms figure prominently in the conformational preferences of N^α-chiral peptoids (6). The concordance of the model with the conformation determined by NMR strengthens this hypothesis. Thus, although these results cannot establish the nature of the driving forces in peptoid folding, they do imply that steric interactions play a prominent role in determining the folded conformation of peptoids containing chiral centers at the side chains' α position. Conversely, because CD spectroscopy indicates that this conformation is very similar in aqueous and organic solvents (5), it seems unlikely that the solvent plays a major role in defining the folded conformation. This result can be contrasted to the behavior of polyproline, another polymer without hydrogen bonds, in which solvent seems to have a major impact on the polymer's conformational preferences (23).

Another important result is the presence of a minor isomer of **1** with a *trans*-amide bond. This implies that the energetic difference between the *cis* and *trans* geometries is not large enough to result in the exclusive population of one geometry, and proves that *cis/trans* isomerization is a source of conformational flexibility in this molecule. The population of both *cis*- and *trans*-amide bonds is not surprising, as N-alkylation should lower the energetic difference between *cis* and *trans* forms of peptoids (5, 24), as it does for proline (25). The *cis/trans* isomerizations have been noted in small proline-containing peptides in solution (21) and in previous NMR studies of peptoids (16, 26). It is possible that other amide bonds also adopt *trans* geometries in the minor isomers of **1**. However, modeling suggests that the amide bond of residue 4 could have a particular reason for favoring a *trans* isomer: this geometry could allow the C-terminal amide hydrogens to form a hydrogen bond with the carbonyl oxygen of residue 4, if residues 2–4 adopt the dihedral angle values of the major isomer (data not shown).

Although this isomerization occurs on a slow time scale, the major conformation of **1** is also likely to fluctuate on a fast time scale. The ensemble calculations help to define the flexibility of this isomer. It suggests that although the right-handed conformation is the most populated in solution, other conformations with less regular geometries or opposite handedness may also be populated and exchange with the right-handed helix rapidly relative to the NMR time scale (Fig. 4*B*). Because the penalty function in the ensemble calculation consists only of distance restraints and hard-sphere repulsions, it would be inappropriate to infer the relative energies of the different conformations from their relative populations in the ensemble.

CONCLUSION

The NMR data for the chiral pentapeptoid **1** strongly suggest that the major conformation in methanol consists of a right-handed helix with three residues per turn, a pitch of ≈ 6 Å, and cis-amide bonds. This is consistent with the CD data for this molecule. Because many oligopeptoids containing chiral centers at the side chains' α position have almost identical CD characteristics as **1**, the structure proposed here may also describe the major conformation adopted by those other peptoids. Hence, this represents an important step in the study of peptoid structure. Together with the success of standard modeling methods in predicting the general features of this conformation, this work opens a very promising path to rational peptoid design, which we are currently pursuing. The helical motif described here could eventually form the structural basis for a new class of biomimetic molecules with applications ranging broadly throughout chemistry and biology.

1. Simon, R. J., Kania, R. S., Zuckermann, R. N., Huebner, V. D., Jewell, D. A., *et al.* (1992) *Proc. Natl. Acad. Sci. USA* **89,** 9367–9371.
2. Zuckermann R. N., Martin, E. J., Spellmeyer, D. C., Stauber G. B., Shoemaker K. R., *et al.* (1994) *J. Med. Chem.* **37,** 2678–2685.
3. Zuckermann, R. N., Kerr, J. M., Kent, S. B. H. & Moos, W. H. (1992) *J. Am. Chem. Soc.* **114,** 10646–10647.
4. Murphy, J. E., Uno, T., Hamer, J. D., Cohen, F. E., Dwarki, V. & Zuckermann, R. N. (1997) *Proc. Natl. Acad, Sci. USA* **95,** 1517–1522.
5. Kirshenbaum, K., Barron, A. E., Goldsmith, R. E., Armand, P. A., Bradley, E. K., Truong, K. T. V., Dill, K. A., Cohen, F. E. & Zuckermann, R. N. (1997) *Proc. Natl. Acad Sci. USA* **95,** 4303–4308.
6. Armand, P., Kirshenbaum, K., Falicov, A., Dunbrack, R. L., Jr, Dill, K. A., Zuckermann, R. N. & Cohen, F. E. (1997) *Folding Design* **2,** 369–375.
7. Dale J. A. & Mosher, H. S. (1973) *J. Am. Chem. Soc.* **95,** 512–519.
8. Marion, D. & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun.* **113,** 967–974.
9. Rance, M., Sorensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R. & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun.* **117,** 479–485.
10. Hwang, T.-L. & Shaka, A. J. (1992) *J. Am. Chem. Soc.* **114,** 3157–3159.
11. Mueller, L. (1979) *J. Am. Chem. Soc.* **101,** 4481–4484.
12. Bax, A. & Summers, M. F. (1986) *J. Am. Chem. Soc.* **107,** 2093–2094.
13. Day, M., Kneller, D. & Kuntz, I. D. (1993) *NMR Pack* (Univ. of California, San Francisco).
14. Goddard, T. G. & Kneller D. G. (1977) SPARKY 3: NMR Assignment and Integration Software (Univ. of California, San Francisco).
15. Bradley, E. K. (1996) *J. Magn. Res. Ser. B* **110,** 195–197.
16. Bradley, E. K., Kerr, J. M., Richter, L. S., Figliozzi, G. M., Goff, D. A., Zuckermann, R. N., Spellmeyer, D. C. & Blaney, J. M. (1997) *Mol. Diversity* **3,** 1–15.
17. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids* (Wiley, New York).
18. Blackledge, M. J., Brüschweiler, R., Griesinger, C., Schmidt, J. M., Xu, P. & Ernst, R. R. (1993) *Biochemistry* **32,** 10960–10974.
19. Havel, T. F. (1991) *Prog. Phys. Mol. Biol.* **56,** 43–78.
20. Pellegrini, M., Gobbo, M., Rocchi, R., Peggion, E., Mammi, S. & Mierke, D. F. (1996) *Biopolymers* **40,** 561–569.
21. Mierke, D. F., Kurz, M. & Kessler, H. (1994) *J. Am. Chem. Soc.* **116,** 1042–1049.
22. Pellegrini, M., Mammi, S., Peggion, E. & Mierke, D. F. (1997) *J. Med. Chem.* **40,** 92–98.
23. Harrington, W. F. & Sela, M. (1958) *Biochim. Biophys. Acta* **27,** 24–41.
24. Moehle, K. & Hofmann, H.-J. (1996) *Biopolymers* **38,** 781–790.
25. Creighton, T. E. (1993) *Proteins: Structures and Molecular Properties* (Freeman, New York), 2nd Ed.
26. Melacini, G., Feng, Y. & Goodman, M. (1996) *J. Am. Chem. Soc.* **118,** 10725–10732.