Structures of the Contryphan Family of Cyclic Peptides. Role of Electrostatic Interactions in cis−trans Isomerism[†,‡]

Paul K. Pallaghy,[§] Weilan He,[§] Elsie C. Jimenez,[||,⊥] Baldomero M. Olivera,[||] and Raymond S. Norton*[,§]

Biomolecular Research Institute, 343 Royal Parade, Parkville 3052, Australia, and Department of Biology, University of Utah, Salt Lake City, Utah 84112

Received May 12, 2000

ABSTRACT: The contryphan family of cyclic peptides, isolated recently from various species of cone shell, has the conserved sequence motif NH_3^+-X_1COD-WX_5PWC-NH_2, where X_1 is either Gly or absent, O is 4-*trans*-hydroxyproline, and X_5 is Glu, Asp, or Gln. The solution structures described herein of two new naturally occurring contryphan sequences, contryphan-Sm and des[Gly1]-contryphan-R, are similar to those of contryphan-R, the structure of which has been determined recently [Pallaghy et al. (1999) *Biochemistry 38*, 11553−11559]. The ^{1}H NMR chemical shifts of another naturally occurring peptide, contryphan-P, indicate that it also adopts a similar structure. All of these contryphans exist in solution as a mixture of two conformers due to cis−trans isomerization about the Cys2−Hyp3 peptide bond. The lower cis−trans ratio for contryphan-Sm enabled elucidation of the 3D structure of both its major and its minor forms, for which the patterns of $^3J_{H\alpha HN}$ coupling constants are very different. As with contryphan-R, the structure of the major form of contryphan-Sm (cis Cys2−Hyp3 peptide bond) contains an N-terminal chain reversal and a C-terminal type I β-turn. The minor conformer (trans peptide bond) forms a hairpin structure with sheetlike hydrogen bonds and a type II β-turn, with the D-Trp4 at the 'Gly position' of the turn. The ratio of conformers arising from cis−trans isomerism around the peptide bond preceding Hyp3 is sensitive to both the amino acid sequence and the solution conditions, varying from 2.7:1 to 17:1 across the five sequences. The sequence and structural determinants of the cis−trans isomerism have been elucidated by comparison of the cis−trans ratios for these peptides with those for contryphan-R and an *N*-acetylated derivative thereof. The cis−trans ratio is reduced for peptides in which either the charged N-terminal ammonium or the X_5 side-chain carboxylate is neutralized, implying that an electrostatic interaction between these groups stabilizes the cis conformer relative to the trans. These results on the structures and cis−trans equilibrium of different conformers suggest a paradigm of 'locally determined but globally selected' folding for cyclic peptides and constrained protein loops, where the series of stereochemical centers in the loop dictates the favorable conformations and the equilibrium is determined by a small number of side-chain interactions.

Short disulfide-bridged peptides frequently exhibit well-structured conformations in solution. It is becoming apparent that the conformations of these peptides are primarily dictated by the number of residues in the loop, the series of chiral centers in the backbone, and the positions of constrained residues such as prolines, rather than the noncovalent side-chain interactions responsible for the folding of globular proteins (*1*−*3*). The associated 'cyclic-peptide folding problem' is instructive for the larger protein folding problem, especially for constrained protein loops (*2*), and has consequences for the design of conformationally constrained peptides for applications such as therapeutics, vaccines, tissue targeting agents, and diagnostics (*3*−*6*).

Peptides displaying sequence-dependent cis−trans equilibria are also good models for the study of intramolecular protein interactions since the side-chain to side-chain interaction strength can be estimated from the effect of specific sequence changes on the equilibrium (*7*). The determination of multiple meta-stable 3D[1] structures accessible to individual cyclic peptide sequences is only rarely achievable (e.g., *8*) and can also contribute to our understanding of the role of local interactions in protein folding through identification of the structural determinants of such equilibria. Establishing the sequence determinants of cyclic peptide structure and conformational equilibria is also important in facilitating the design of cyclic peptides for a range of applications.

We have recently described the structure of a novel, cyclic octapeptide sequence, contryphan-R (Figure 1). This peptide

[†] This work was supported in part by NIH Grant GM 48677.

[‡] The structures and NMR-derived restraints have been deposited in the Protein Data Bank under accession numbers 1DGO (des[Gly1]-contryphan-R major conformer), 1DFY (contryphan-Sm major conformer), and 1DFZ (contryphan-Sm minor conformer).

* Correspondence should be addressed to this author. E-mail: ray.norton@bioresi.com.au, FAX: +61 3 9903 9655, Phone: +61 3 9903 9650.

[§] Biomolecular Research Institute.

[||] University of Utah.

[⊥] Present address: Discipline of Chemistry, University of the Philippines College Baguio, Baguio City 2600, Philippines.

[1] Abbreviations: 1D, one-dimensional; 2D, two-dimensional; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser enhancement; NOESY, 2D NOE spectroscopy; TOCSY, 2D total correlation spectroscopy; RMS, root-mean-square; RP-HPLC, reverse-phase high-performance liquid chromatography.

10.1021/bi0010930 CCC: $19.00
Published on Web 09/28/2000

contryphan-R	GCO$\underline{\text{W}}$**E**PWC-NH_2
bromocontryphan	GCO$\underline{\text{W}}$**E**P~~W~~C-NH_2
des[Gly1]-contryphan-R	CO$\underline{\text{W}}$**E**PWC-NH_2
contryphan-P	GCO$\underline{\text{W}}$**D**PWC-NH_2
Leu-contryphan-P	GC**V**$\underline{\text{L}}$**L**PWC-**OH**
contryphan-Sm	GCO$\underline{\text{W}}$**Q**PWC-NH_2

FIGURE 1: Amino acid sequences of the contryphan family, with variations from the GCOWXPWC-NH_2 consensus sequence highlighted in boldface. 4-*trans*-Hydroxyproline is represented by O, D-handed residues are underlined, and W with a double strike-through represents 6-bromotryptophan. Contryphan-R (*9*), des-[Gly1]-contryphan-R (*9*), and bromocontryphan (*11*) are from *Conus radiatus*, contryphan-Sm (*12*) is from *C. stercusmuscarum*, and Leu-contryphan-P (*10*) is from *C. purpurascens*. The contryphan-P sequence has been deduced from cDNA cloned from *C. purpurascens* (*12*).

causes 'stiff-tail' syndrome when injected in mice (*9*) although its receptor is not yet identified. It has a well-ordered structure in solution, with an N-terminal chain reversal centered on Hyp3−D-Trp4 (with the peptide bond preceding Hyp3 in the cis conformation) and a C-terminal type I β-turn (*2*). It is evident from RP-HPLC (*10*) and NMR (*2*) that a minor conformation (with a trans peptide bond preceding Hyp3) is present in a ratio of 8:1. The effective concentration of this conformer was too low to allow structure determination, but it was clear from examination of its chemical shifts that the structure must be quite different from that of the major form.

Here we present an NMR study of the structures of three new naturally occurring contryphans, contryphan-Sm, contryphan-P, and des[Gly1]-contryphan-R (Figure 1) (*9, 10, 12*), and the factors that govern the ratio of cis and trans conformers in these peptides. Because of the lower cis−trans isomer ratio for contryphan-Sm, we have been able to determine the backbone fold of the minor form as well as the complete structure of the major form. We also present the structure of des[Gly1]-contryphan-R as well as sequential assignments and a reconciliation of HPLC and NMR determinations of the cis−trans isomer ratio for contryphan-P. An N-terminally acetylated analogue of contryphan-R was synthesized to enable the effect of the N-terminal charge on the cis−trans ratio to be determined. The structures and cis−trans ratios of the five sequences, including those of contryphan-R, are compared and discussed in terms of the consequences for peptide and protein folding and design.

EXPERIMENTAL PROCEDURES

Most samples consisted of 1−5 mM synthetic peptide in 90% H_2O/10% ^2H_2O and pH 2.9−3.2 (*9, 10, 12*) (Table 1). A small sample of acetylated contryphan-R was provided by Dr. Jonathan Baell for the purpose of measuring the cis−trans ratio in the absence of the N-terminal charge. With contryphan-R, we had confirmed previously that the same NOE set was present at pH 5.2, where Glu5, the only titratable carboxyl group in the molecule, would be deprotonated and able to participate in ionic interactions (*2*).

The ratios of cis to trans forms of the contryphans were estimated from the integrated intensities of nonoverlapped amide resonances. The effects of pH, amino acid sequence, and solvent were examined. In particular, the solvent conditions of contryphan-P were adjusted to match the HPLC conditions [35% buffer B60 (*10*), i.e., 21% acetonitrile by volume and pH 2.0] in an attempt to reconcile the cis−trans ratios measured by HPLC and NMR.

Structures were determined from 2D ^{1}H NMR spectra recorded on a Bruker DRX-600 spectrometer at 5 °C. The methods of spectral acquisition and analysis, extraction of constraints, and structure calculation were as described previously (*2, 3*). The TOCSY and NOESY spin-lock and mixing times were 70 and 300 ms, respectively. A series of 1D experiments was recorded for each peptide over the temperature range 5−25 °C to determine the amide proton temperature coefficients. Amide exchange half-lives were calculated from exponential decay fits to plots of nonoverlapped peak volumes measured from a series of short TOCSY experiments versus time. Structures were based on NOE and ϕ and χ^1 dihedral angle constraints. After simulated annealing of a random template structure in the X-PLOR (*13*) distance geometry force-field, the structures were placed in a cube of water (side length 35 Å) and energy-minimized with periodic boundary conditions in the CHARMM-19 force-field (*14*) in X-PLOR. The 20 best structures and NMR-derived constraints for the major forms of des[Gly1]-contryphan-R and contryphan-Sm and the minor form of contryphan-Sm have been deposited in the Protein Databank (*15*) (accession numbers 1DGO, 1DFY, and 1DFZ, respectively) and the chemical shift values in the BioMagResBank (*16*) (accession numbers BMRB-4469, BMRB-4464, and BMRB-4468, respectively). Chemical shift calculations were made using Shiftcalc (*17*), a program embedded in Molmol (*18*).

RESULTS

NMR Spectroscopy. Published RP-HPLC data for contryphan-Sm, contryphan-P, and des[Gly1]-contryphan-R show that separate peaks, eluting several minutes apart and having a population ratio in the range 3:1 to 17:1, were due to interconverting forms of the pure peptides (*10*). At least two separate sets of resonances were visible in the ^{1}H NMR spectra of all of these peptides, with contryphan-Sm demonstrating the largest proportion of the minor form (3.3:1) in 90% H_2O/10% ^2H_2O (Table 1). For contryphan-P, the cis−trans ratio in aqueous solution measured by NMR differed markedly from that measured by HPLC (*10*) (8.1:1 and 4.5:1, respectively). After adjusting the pH and adding acetonitrile in an effort to better match the HPLC solution conditions, the cis−trans ratios measured by NMR and HPLC became closer (3.0:1 and 4.5:1, respectively). For contryphan-R, the same modification in solution conditions did not lead to a significant change in ratio as measured by NMR (7.9:1 in aqueous solution compared to 8.0:1 with acetonitrile at pH 2.0), which is as expected since both results were consistent with the HPLC estimate (approximately 10:1) (*10*). The ratio for contryphan-R in aqueous solution could be raised from 7.9:1 to 12.1:1 by increasing the pH from 2.9 to 5.2. The cis−trans ratio for the acetylated form of contryphan-R was 3:0 at pH 3.0 and independent of pH in the range pH 3−5 (Table 1).

Complete resonance assignments for the major conformers and most assignments for the minor conformers were made for contryphan-P, contryphan-Sm, and des[Gly1]-contryphan-R. The chemical shifts are generally very similar to those of contryphan-R, and, as noted for contryphan-R, the chemical

Table 1: NMR and HPLC Data for Contryphan Peptides

	des[Gly1]-contryphan-R	major (cis) form contryphan-Sm	minor (trans) form contryphan-Sm	contryphan-P	contryphan-R	Ac-contryphan-R
concentration (mM)	5.0	4.3[a]	1.2[a]	ca. 1.0	4.0	very low
pH	2.9	2.9	2.9	3.2	3.0	3.0
cis−trans ratio estimates at 5 °C						
H_2O, pH ca. 3.0 (NMR)	16.8:1	3.3:1	−	8.1:1	7.9:1	3.0:1
H_2O, pH ca. 5.1 (NMR)	−	−	−	−	12.1:1	−
21% CH_3CN, pH ca. 2.0 (NMR)	−	−	−	3.0:1	8.0:1	−
RP-HPLC[b]	−	ca. 6:1	−	4.5:1	ca. 10.0:1	−
N-term NH_3^+−Xxx5 $O^{\delta/\epsilon}$ distance	7.0 ± 0.9 Å	6.3 ± 1.2 Å	9.8 ± 1.1 Å	−	8.4 ± 0.7 Å	−
no. medium- & long-range NOEs	27	15	3	−	22	−
ϕ-angle dihedral series[c]	−↑↑↑↓	↑↑↑↑↓	↑↓↑↓↑	−	↑↑↑↑↓	−
amide exchange at pH	3.3	2.9	2.9	−	2.7	−
amide exchange $t_{1/2}$ at 5 °C (h)[d]						
Xxx5	21	44	27	−	38	−
Trp7	9.6	7.2	8.9	−	6.6	−
Cys8	8.8	8.9	<1	−	8.7	−

[a] Effective concentration for the major and minor forms of contryphan-Sm. [b] The values listed are an estimate based on RP-HPLC peak heights (*10*), except in the case of contryphan-P for which the ratio was determined accurately through measurement of the kinetics (*10*). [c] From Cys2, D-Trp4, Xxx5, Trp7, and Cys8 $^3J_{H\alpha HN}$ coupling constants, respectively. ↑ represents a β-sheet-type ϕ angle (ca. −120° except +120° for D-handed residues), ↓ represents an α-helical-type ϕ angle (ca. −60°), and − indicates that the angle was not constrainable. [d] $t_{1/2}$ < 1 h for Cys2 and D-Trp4.

shifts of the major and minor forms are very different from one another (*2*). The major and minor conformers are characterized by the conformation at the Cys2−Hyp3 peptide bond being cis and trans, respectively, for all peptides, as identified by the expected sequential NOEs ($d_{\alpha\alpha}$ and $d_{\alpha\delta}$ for cis and trans, respectively). The Xxx5−Pro6 peptide bond is trans in both conformers for all peptides. A third conformer was apparent, being most populated in contryphan-Sm with a ratio of ca. 10:1 to the major form and ca. 3:1 to the minor form; this is most likely the form where both the Cys2−Hyp3 and Gln5−Pro6 peptide bonds are cis. The amide exchange rates and coupling constants are tabulated in Table 1 and are all similar to those observed for contryphan-R (*2*).

Structure Calculations. The NOE cross-peaks were strong enough to enable well-defined structures of the major forms of des[Gly1]-contryphan-R and contryphan-Sm to be determined. A reasonably well-defined structure, at least for the backbone, was obtained for the minor form of contryphan-Sm, somewhat surprisingly given that there were only three medium-range NOEs, four ϕ-angle constraints, and two χ^1-angle constraints. In all three cases, structural constraints were extracted from a 300 ms NOESY spectrum and input to simulated annealing and energy minimization calculations in explicit solvent. The families of structures are shown in Figure 2. The mean pairwise RMS differences calculated over the backbone heavy atoms and all heavy atoms, respectively, of the whole molecule, as well as other structural characteristics are summarized in Figure 3 and Table 2. The concentrations of the contryphan-P and acetylated contryphan-R samples were too low to allow a reliable structure determination. However, the large deviations from random-coil chemical shifts evident in the spectrum of the major form of contryphan-R were observed, suggesting that their structures are all very similar.

Structures of the Major Forms. The major conformers of des[Gly1]-contryphan-R and contryphan-Sm exhibit the 'contryphan fold' (*2*), consisting of a seven-residue loop stabilized by the Cys2−Cys8 disulfide bridge (Figure 2). There are two chain reversals: a non-hydrogen-bonded reversal between residues 1 and 5 and a type I β-turn encompassing residues 5−8 and stabilized by a hydrogen bond from the slowly exchanging Cys8 NH to Glx5 carbonyl, as in contryphan-R (*2*). The cis Cys2−Hyp3 peptide bond and D-Trp4 presumably play a role in generating the first chain reversal.

In contryphan-R, the distance between the two potential salt bridge partners, the Glu5 carboxyl and the N-terminal ammonium groups, was 8.4 ± 0.7 Å, suggesting that a salt bridge did not occur. This was confirmed by the observations that (i) the same NOE set was present at pH 5.2 (where the Glu5 carboxyl group would be ionized) and (ii) the Glu5 carboxyl pK_a was higher than the value in unstructured small peptides (the opposite would be expected if a salt bridge were present) (*2*). In des[Gly1]-contryphan-R and contryphan-Sm, where the N-terminal charge is shifted and Glu is substituted with Gln respectively, relative to contryphan-R, the corresponding distances are 7.0 ± 0.9 and 6.3 ± 1.2 Å, respectively, demonstrating that a salt bridge is not present in these molecules either (Table 1).

Comparison of the structures of des[Gly1]-contryphan-R and contryphan-Sm with that of contryphan-R (Figure 2) shows that both differ from contryphan-R in having a better defined Cys8 HN to Glx5 O hydrogen bond, a more arched backbone structure around Glx5, and a shift of the terminus of the Glx5 side chain toward the N-terminus (as reflected in the distances quoted above). The backbone RMS deviations of the mean structures from each other are, however, all approximately the same (0.89, 0.92, and 0.91 Å for des-[Gly1]-contryphan-R vs contryphan-R, contryphan-Sm vs contryphan-R, and des[Gly1]-contryphan-R vs contryphan-Sm, respectively).

A feature of the contryphan ^{1}H NMR spectra is the significant perturbation from random-coil chemical shifts for a number of resonances (as much as −1.60 and +0.73 for the Gln5 and D-Trp4 amide protons of contryphan-Sm, respectively). To evaluate the role of the Trp residues in these perturbations, the ring current contributions (*17*) to all chemical shifts were calculated from the structures and compared with the experimental values. Qualitatively, these calculations confirm that ring current effects from the Trp residues could cause significant shifts from random-coil values. Quantitatively, however, the calculated chemical

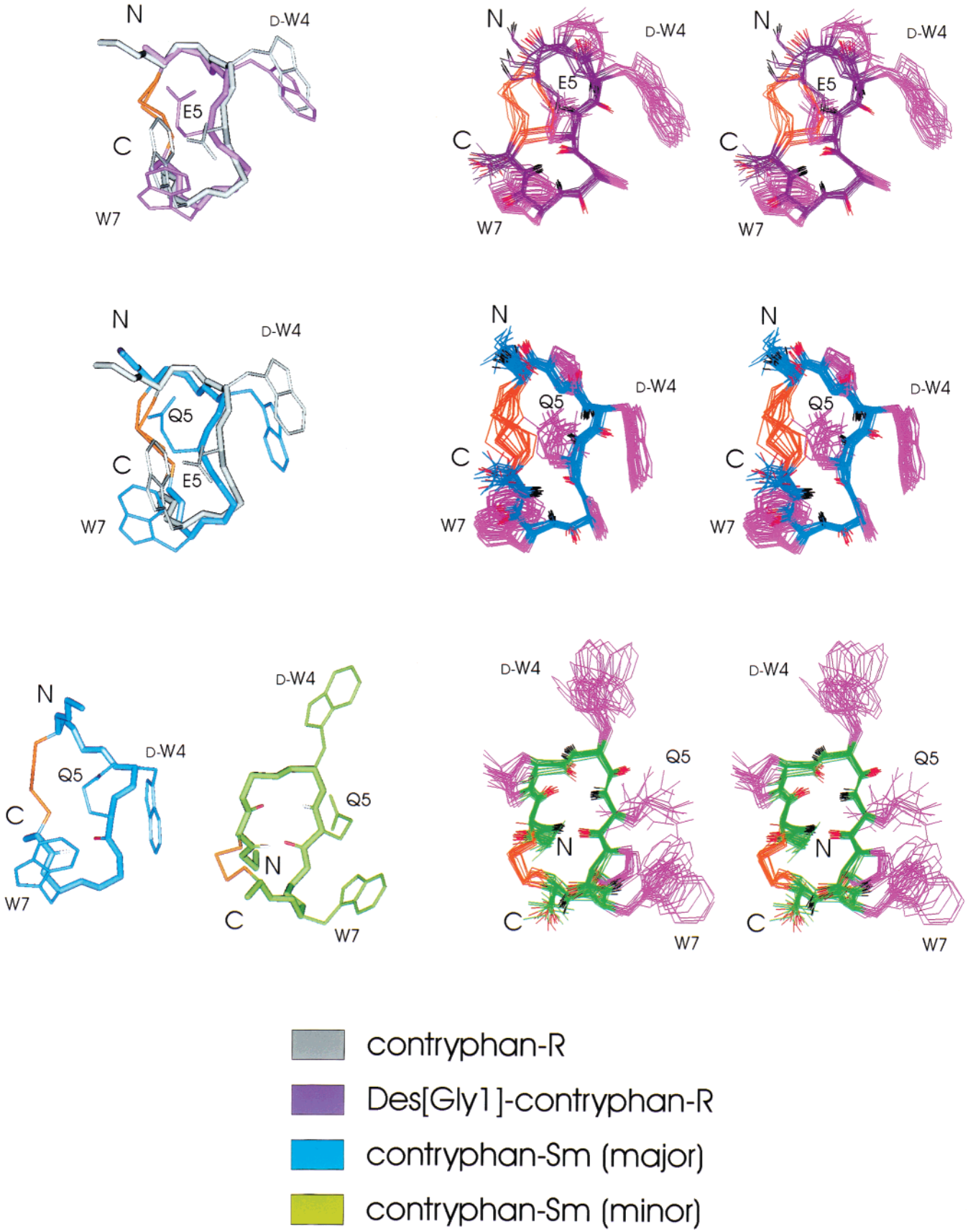

FIGURE 2: Structural comparisons (left) and stereoviews of the families (right) for des[Gly1]-contryphan-R (purple) and the major (blue) and minor (green) forms of contryphan-Sm. Contryphan-R (*2*) is shown in gray. The disulfide bridges are shown in orange, the carbonyl oxygen atoms in red, and amide protons in black.

shifts for the six largest deviations from random-coil shifts (four of which were for amide protons) were not in good agreement with the experimental values. Adjustments of between 40° and 60° to the χ^2 angle for both Trp residues

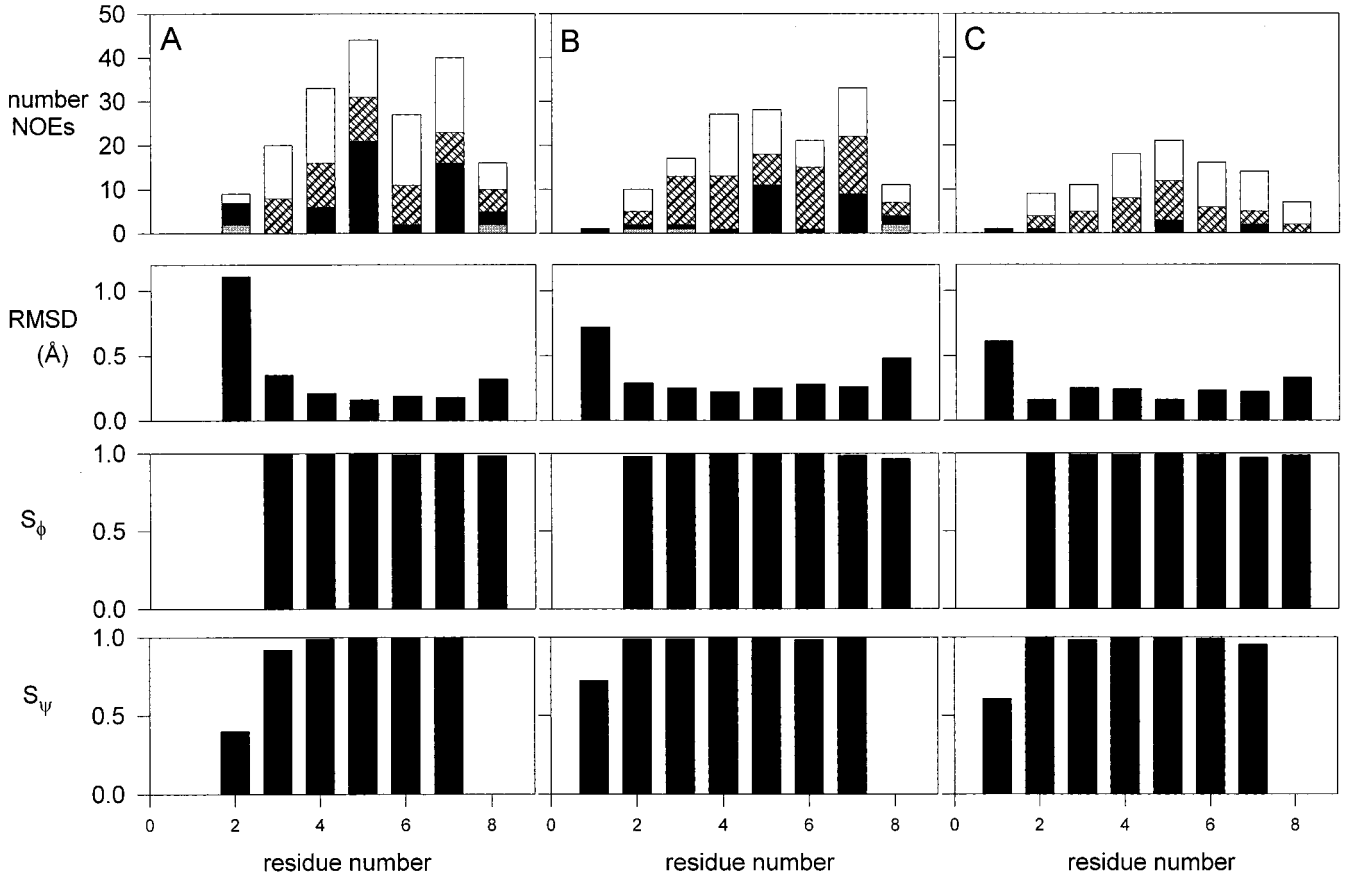

FIGURE 3: Number of NOEs, average RMS deviation from the mean structure, and ϕ and ψ angular order parameters (*19, 20*) vs residue for des[Gly1]-contryphan-R (A) and the major (B) and minor (C) forms of contryphan-Sm. The numbers of NOEs are indicated for the long-range ($i - j \geq 5$, black), medium-range ($1 < i - j < 5$, hatched), sequential (shaded), and intraresidue (unfilled) connectivities.

of contryphan-Sm (D-Trp4 $\chi^2 = -40° \rightarrow -100°$ and Trp7 $\chi^2 = -85° \rightarrow -45°$) enabled the sign of the deviation from random-coil shifts to be reconciled with the data and the magnitude to be accounted for with an average of 36% error. Following these side-chain rotations, the respective Trp side chains still occupied the same canonical rotamers as before the adjustments, but some NOEs were now violated (by up to 0.7 Å). Given that chemical shift calculations for amide protons are less reliable than those for CH protons because of potential hydrogen-bonding interactions, and bearing in mind that these calculations are for solvent-exposed Trp side chains that are undoubtedly somewhat mobile in solution, we believe that the side-chain positions shown in Figure 2 are a reasonable representation of the average structures in solution, with some allowance for side-chain mobility. The side-chain adjustments described above are not incorporated in the structural families presented in Figure 2.

Structure of the Minor Form of Contryphan-Sm. As suggested by the pattern of chemical shift differences from random-coil values and the $^3J_{H\alpha HN}$ coupling constants, the structure of the minor form of contryphan-Sm is very different from that of the major form and is characterized by the Cys2−Hyp3 peptide bond adopting the trans conformation. The peptide has a β-hairpin/sheetlike structure, with two interstrand hydrogen bonds and a type II β-turn at the hairpin (Figure 2). D-Trp4, in a positive ϕ-angle conformation, occupies position 3 in the type II β-turn, where a glycine would normally be found. One of the interstrand hydrogen bonds is very skewed, but this may be due to the low number of constraints used in the generation of these structures.

This structure is significantly less compact than that of the major form, with D-Trp4, Gln5, and Trp7 exposed to the solvent rather than being packed against the backbone. While this could be due to the dearth of constraints (due to the low effective concentration), the side chains appear sufficiently well-defined up to the γ heavy atom to rule this out. The distance between Gln5 and the N-terminus was 9.8 ± 1.1 Å, the largest for any analogue or conformer (Table 1).

Amide Exchange Rates for Major and Minor Forms. The amide exchange data are reasonably consistent between the major forms of contryphan-R, des[Gly1]-contryphan-R, and contryphan-Sm, with the backbone amide signals of Glx5, Trp7, and Cys8 remaining visible in the 2D TOCSY experiment (Table 1). The half-lives for these amides are 21−44 h, 6.6−8.9 h, and 4.1−8.8 h, respectively, the major form of contryphan-Sm having the slowest exchanging amide proton (Gln5, 44 h).

Cys8 is involved in the only unambiguous hydrogen bond detected in the major form in all three contryphans whose structures are known, stabilizing the type I β-turn, and yet it is not the most slowly exchanging amide. This is due at least partly to the interconversion between the two conformers. In contryphan-P, for example, the major form converts to the minor form with a half-life, $t_{1/2}$, of 86 min at 25 °C, and the reverse takes place with a $t_{1/2}$ of 19 min (*10*). The Cys8 amide proton of the minor form has a very short $t_{1/2}$ (<1 h) (Table 1). Since the rates of isomerism are faster than the slow amide exchange rates, an amide proton protected in one conformer but unprotected in the other will display

Table 2: Structural Characteristics for the 20 Best Structures of Each Peptide[a]

	des[Gly1]-contryphan-R	contryphan-Sm major (cis) form	contryphan-Sm minor (trans) form
RMS deviations from exptl distance restraints (Å)	0.027 ± 0.003	0.022 ± 0.003	0.028 ± 0.003
no. of exptl distance restraints	135	114	76
RMS deviations from exptl dihedral restraints (deg)	0.56 ± 0.37	0.09 ± 0.17	0.45 ± 0.28
no. of exptl dihedral restraints	7	8	6
RMS deviations from idealized geometry			
bonds (Å)	0.013 ± 0.002	0.012 ± 0.001	0.014 ± 0.003
angles (deg)	4.0 ± 0.1	3.60 ± 0.06	3.67 ± 0.07
impropers (deg)	0.4 ± 0.1	0.4 ± 0.1	0.6 ± 0.4
energies (kcal mol^{-1})[b]			
E_{NOE}	4.9 ± 0.8	2.9 ± 0.8	3.0 ± 0.6
E_{cdih}	0.25 ± 0.25	0.02 ± 0.05	0.13 ± 0.16
$E_{bond}+E_{angle}+E_{improper}$	61 ± 5	54 ± 2	59 ± 10
E_{L-J}	−9 ± 2	−18 ± 2	−17 ± 3
E_{elec}	−155 ± 4	−191 ± 5	−179 ± 12
pairwise RMS differences			
backbone (N, C^a, C) (Å)	0.66 ± 0.31	0.56 ± 0.17	0.46 ± 0.15
all heavy (Å)	0.93 ± 0.25	0.76 ± 0.15	1.46 ± 0.30

[a] The best 20 structures after solvated energy minimization in X-PLOR using the CHARMM-19 force field. Values are mean ± standard deviation. 100% of the residues were found to be in the favorable or most favorable regions of the Ramachandran plot by PROCHECK−NMR (*21*). [b] 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.

different corresponding amide exchange rates, and the apparent $t_{1/2}$ of the amide proton in the more protected conformer will be underestimated. It is likely that this is what is reflected in the exchange data for the amide proton of Cys8, which is involved in hydrogen bonding in the major conformer but not in the minor conformer.

The $t_{1/2}$ values for the amide protons of Gln5 and Trp7 of contryphan-Sm are similar in the major and minor conformers (Table 1). These protons have potential nearby acceptor atoms in the major conformer. The type I β-turn centered on Pro6−Trp7 is rather skewed, and the Cys8 HN−Gln5 O hydrogen bond may be bifurcated with a Trp7 HN−Gln5 O hydrogen bond, and the Glx5 amide proton appears to be interacting with the sulfur of Cys8 in many of the structures of the families. In the minor conformer of contryphan-Sm, the Gln5 amide proton is involved in one of the bridging hydrogen bonds, whereas the Trp7 amide proton may be buried between the Pro6 and Trp7 side chains. The Cys2 amide proton is involved in the second bridging hydrogen bond and is probably absent in the exchange experiments due to the fast exchange of that proton in the major conformer.

DISCUSSION

In this section we discuss the sequence and structural determinants of cis−trans equilibria in this series of contryphan peptides. Our data also allow some more general conclusions to be drawn concerning the factors that determine the folding of cyclic peptides and protein loops.

cis−trans Ratio. All of the data presented imply that an *i*/*i*+4 electrostatic interaction between the N-terminal ammonium group and the charged side chain at position 5 is a major determinant of the cis−trans equilibrium. Most of the sequence variation occurs at position 5, three residues from the Cys2−Hyp3 peptide bond responsible for isomerization. Neutralization of this residue from Glu or Asp to Gln reduces the cis−trans ratio from 8:1 to 3:1 at pH 3 in aqueous solution. Variation from Glu to Asp has no effect on the cis−trans ratio under these conditions. Neutralizing the N-terminal charge through acetylation similarly reduces the ratio from 8:1 to 2.7:1. Thus, neutralization of either charge in the peptide yields a cis−trans ratio of approximately 3:1. Moving the only formal positive charge in the molecule (the N-terminal ammonium) one residue closer in the sequence to the negatively charged Glu5 in des[Gly1]-contryphan-R pushes the ratio in the other direction to ca. 17:1.

The cis−trans ratio reflects the relative stability of the major and minor forms. Taken together, the data above indicate that the electrostatic interaction stabilizes the major conformation and that this interaction is weakened in contryphan-Sm (Gln5) and Ac-contryphan-R, where the electrostatic interaction is neutralized, and strengthened in des[Gly1]-contryphan-R, where the distance between the charged groups is shortened. This is supported by the observation that the N-terminal NH_3^+−Xxx5 $O^{\delta/\epsilon}$ distance is shorter in des[Gly1]-contryphan-R than in contryphan-R (Table 1). In addition, the N-terminal NH_3^+−Xxx5 $O^{\delta/\epsilon}$ distance is larger in the minor form of contryphan-Sm than the major form.

In aqueous solution, the cis−trans ratio was pH dependent for contryphan-R, implying a role for the electrostatic interaction. In the presence of acetonitrile, and at pH 2.0 (to mimic the RP-HPLC conditions), the ratios for contryphan-P (Xxx5 = Asp) measured by NMR and HPLC were reconciled. The HPLC conditions did not alter the ratio for contryphan-R (Xxx5 = Glu) but did for contryphan-P (Xxx5 = Asp). The insensitivity of the ratio in contryphan-R may reflect a stronger electrostatic interaction with the longer Glu5 side chain, which is in turn less affected by solvent conditions.

Our results on the cis−trans ratios in the contryphans have some similarities to those reported for a series of X−Pro dipeptides (*7*), where X was varied to include Gly, L- and D-Ala, Leu, and Phe, and the effects of blocking the N- and C-termini and varying the pH were also examined. The cis forms were stabilized by the larger L-handed hydrophobic side chains. For most sequences, the cis forms were destabilized at acidic pH and when the C-terminus was blocked, both conditions under which the C-terminus would be neutralized. This is consistent with the present results, where the cis−trans ratio is lower for contryphan-Sm, with Gln5, than in contryphan-R or -P, with a negatively charged side chain at position 5, even though there is a five-residue gap between the interacting partners and the potential negative charge is located on a side chain in the contryphans. The pH dependence of the cis−trans ratio for contryphan-R (Table 1) supports these considerations, with the cis−trans ratio increasing with pH over the range 3.0−5.1. It is possible that cis peptide bonds are stabilized by long-range bridging interactions (such as NH_3^+−Xxx5 $O^{\delta/\epsilon}$) because of the U-turn-like nature of the cis bond.

Major Form Structures. The structures of the major forms of des[Gly1]-contryphan-R and contryphan-Sm are similar to that of contryphan-R, as expected. More generally, these results, and the similar chemical shifts for contryphan-P, demonstrate that the underlying sequence motif (CP$_D$-XXPXC) defines a scaffold that is structurally robust with respect to side-chain substitutions, as further demonstrated in our recent work with engineered contryphan sequences designed to mimic an unrelated toxin surface (*3*).

Minor Form Structure. The structure of the minor form of contryphan-Sm is very different from that of the major form, having a β-hairpin-like structure rather than the double-turn structure of the major form. The extent of the difference was evident very early from the series of measurable $^3J_{H\alpha HN}$ coupling constants (for Cys2, D-Trp4, Gln5, Trp7, Cys8), as follows: ↑↑↑↑↓ (major form) and ↑↓↑↓↑ (minor form), where ↑ and ↓ represent >8 Hz and <6 Hz coupling constants indicative of β-sheet- and α-helix-like ϕ-angle conformations, respectively (Table 1).

For small cyclic peptides, it is intuitive that cis−trans isomerization should cause a significant global structural change due to the constraints imposed by cycle closure. The large difference between the 3D structures of the two forms of contryphan-Sm explains partially why they were separable by RP-HPLC (*10*, *12*). Specifically, the less compact minor form, having a larger hydrophobic surface area, would be expected to interact more strongly with the nonpolar stationary phase, and thus to elute more slowly on RP-HPLC (*10*, *12*). In addition, the structural constraints imposed by cyclization would increase the activation barrier for transition between multiple forms due to the significant structural rearrangement necessary, thereby reducing the conformational exchange rate and allowing separation on the chromatographic time-scale. Similar chromatographic separation of multiple isomers of larger polypeptides and proteins with proline-containing constrained loops (e.g. ref *22*) is not generally observed because of the relatively small change to the overall protein structure, and hence to the protein's chromatographic properties, caused by the cis−trans isomerism, even though in some cases the rate of interchange between the two isomers might be similar to that in the contryphans or even slower. In the case of linear peptides, cis−trans isomerism would cause more localized structural changes, and the conformational exchange rates would be faster than the chromatographic time-scale. In addition, most linear peptides are inherently flexible, and the chromatographic properties represent the average over a large ensemble of rapidly interconverting conformations. The combination of more rapid interchange between the cis and trans isomers, smaller structural differences between the isomers, and conformational averaging within each isomer makes chromatographic separation of isomers much less likely for linear peptides containing proline residues.

The cis form is over an order of magnitude more predominant in active (*9*) des[Gly1]-contryphan-R, suggesting that it is this form that is biologically relevant, although this remains to be proved.

Implications for Peptide and Protein Folding and Design. Elucidation of the sequence determinants of cis−trans equilibria and determination of the alternative structures of contryphan-Sm have several corollaries for the local folding of peptides and protein loops. It is significant that all potentially measurable ϕ angles [the N-terminal residue and (hydroxy-)proline ϕ angles cannot be measured due to the absence of $^3J_{H\alpha HN}$ couplings] were in non-random-coil-like conformations for both conformers of contryphan-Sm, suggesting that the backbone structures of both are well ordered, with all residues in the constrained loop adopting favorable Ramachandran configurations. This result was not necessarily expected and suggests that arbitrary constrained sequences might intrinsically adopt a small number of non-random-coil-like Ramachandran configurations. This is the case even in the absence of evolutionary pressure acting on a conformation to maintain high potency via a rigid structure (as presumably does occur for one of the isomers of the contryphans), and this suggests that small cyclic peptides might, in general, adopt well-defined structures with backbone dihedral angles in the favored regions of a Ramachandran plot.

The consequences for structure prediction of cyclic peptides and protein loops are encouraging. It may be possible to use simple geometric force-fields to search for the small number of conformations with favorable Ramachandran configurations consistent with the loop size, and submit these to energy minimization and ranking in more sophisticated force-fields, possibly with explicit solvent models (*2*).

The notions outlined above, together with the sensitivity of the cis−trans ratio in the contryphan peptides to sequence variation and pH, support a paradigm of 'locally determined, globally selected' folding for cyclic peptides and constrained protein loops. In this framework, which is either simply a conceptual aid or a model of what actually occurs in cyclic peptide dynamics (or both), a small number of favored conformations is determined by the local stereochemistry, and the equilibrium between these states is selected by the global favorability of a small number of side-chain interactions. The similarity of the contryphan-Sm and -R major form structures, while having different cis−trans ratios, lends particular support to this view of the folding of cyclic peptides and constrained protein loops. The nature of the side chains has little or no bearing on the structure, which is dictated by the series of backbone chiralities, and instead determines the equilibrium populations.

CONCLUSIONS

The cis−trans ratio of the contryphan family peptides is dictated by the strength of the interaction between the N-terminus and the residue 5 side chain. In particular, the major cis form is stabilized in sequences where an attractive electrostatic interaction can occur between the N-terminal ammonium group and the carboxylate moiety of the Glu5 or Asp5 side-chain. Our determination of the 3D structure of the minor form of contryphan-Sm demonstrates that more than one structure consistent with the loop size is energetically favorable and that these two structures are very different. The very different series of coupling constants with non-random-coil values for both conformers suggests that the contryphans, and probably cyclic peptides and constrained protein loops in general, populate Ramachandran-allowed configurations consistent with the loop constraint. We have also demonstrated that the contryphan family of peptides exhibits a canonical fold that is robust with respect to sequence changes at position 5 and the N-terminus.

ACKNOWLEDGMENT

We are grateful to Jonathan Baell for acetylating the sample of contryphan-R.

REFERENCES

1. Wermuth, J., Goodman, S. L., Jonczyk, A., and Kessler, H. (1997) *J. Am. Chem. Soc. 119*, 1328−1335.
2. Pallaghy, P. K., Melnikova, A. P., Jimenez, E. C., Olivera, B. M., and Norton, R. S. (1999) *Biochemistry 38*, 11553−11559.
3. Pallaghy, P. K., and Norton, R. S. (2000) *Biopolymers 54*, 173−179.
4. Porcelli, M., Casu, M., Lai, A., Saba, G., Pinori, M., Cappelletti, S., and Mascagni, P. (1999) *Biopolymers 50*, 211−219.
5. Satoh, T., Aramini, J. M., Li, S., Friedman, T. M., Gao, J., Edling, A. E., Townsend, R., Koch, U., Choksi, S., Germann, M. W., Korngold, R., and Huang, Z. (1997) *J. Biol. Chem. 272*, 12175−12180.
6. Dreesman, G. R., Sparrow J. T., Frenchick, P. J., and Kennedy, R. C. (1985) *Adv. Exp. Med. Biol. 185*, 129−137.
7. Grathwohl, C., and Wüthrich, K. (1976) *Biopolymers 15*, 2025−2041.
8. . Maslennikov, I. V., Sobol, A. G., Gladky, K. V., Lugovskoy, A. A., Ostrovsky, A. G., Tsetlin, V. I., Ivanov, V. T., and Arseniev, A. S. (1998) *Eur. J. Biochem. 254*, 238−247.
9. Jimenez, E. C., Olivera, B. M., Gray, W. R., and Cruz, L. J. (1996) *J. Biol. Chem. 271*, 28002−28005.
10. Jacobsen, R. B., Jimenez, E. C., De la Cruz, R. G. C., Gray, W. R., Cruz, L. J., and Olivera, B. M. (1999) *J. Pept. Res. 54*, 93−99.
11. Jimenez, E. C., Craig, A. G., Watkins, M., Hillyard, D. R., Gray, W. R., Gulyas, J., Rivier, J. E., Cruz, L. J., and Olivera, B. M. (1997) *Biochemistry 36*, 989−994.
12. Jacobsen, R., Jimenez, E. C., Grilley, M., Watkins, M., Hillyard, D., Cruz, L. J., and Olivera, B. M. (1998) *J. Pept. Res. 51*, 173−179.
13. Brünger, A. T. (1992) *X-PLOR Version 3.1. A system for X-ray crystallography and NMR*, Yale University, New Haven, CT.
14. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. (1983) *J. Comput. Chem. 4*, 187−217.
15. Bernstein, F. C., Koetzle, T. F., Williams, G. J., Meyer, E. E., Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., and Tasumi, M. (1977) *J. Mol. Biol. 112*, 535−542.
16. Seavey, B. R., Farr, E. A., Westler, W. M., and Markley, J. L. (1991) *J. Biomol. NMR 1*, 217−236.
17. Case, D. A. (1995) *J. Biomol. NMR 6*, 341−346.
18. Koradi, R., Billeter, M., and Wüthrich, K. (1996) *J. Mol. Graphics 14*, 51−55.
19. Hyberts, S. G., Goldberg, M. S., Havel, T. F., and Wagner, G. (1992) *Protein Sci. 1*, 736−751.
20. Pallaghy, P. K., Duggan, B. M., Pennington, M. W., and Norton, R. S. (1993) *J. Mol. Biol. 234*, 405−420.
21. Laskowski, R. A., Rullmann, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) *J. Biomol. NMR 8*, 477−486.
22. Scanlon, M. J., and Norton, R. S. (1994) *Protein Sci. 3*, 1121−1124.

BI0010930