Cyclic Peptides. 15. Lanthanide-Assisted ^{13}C and ^{1}H NMR Analysis of Preferred Side-Chain Rotamers in Proline-Containing Cyclic Dipeptides[1]

Paul E. Young, Vincent Madison, and Elkan R. Blout*

Contribution from the Department of Biological Chemistry, Harvard Medical School, Boston, Massachusetts 02115. *Received October 15, 1975*

Abstract: For a series of cyclic dipeptides containing proline, ^{1}H and ^{13}C NMR spectra have been obtained in the presence of varying amounts of $Yb(fod)_3$ and $Eu(fod)_3$. The resulting chemical shift and coupling constant data have been used to ascertain the preferred rotamer for the nonprolyl side chain. To determine whether more than one rotamer is populated, a rotamer averaging method has been applied. The data, taken in conjunction with minimum energy calculations, imply that aliphatic side chains adopt rotamers determined primarily by steric considerations, whereas aromatic side chains have additional stabilizing interactions.

Of the numerous factors affecting the conformations adopted and maintained by proteins in solution, the role played by amino acid side chains has been little examined. Many investigations have centered on determining backbone conformation of main chains and relating secondary structure to amino acid primary sequence.[2] More recently, interest has turned toward determining whether different amino acid side chains may actually prefer different rotamers about the C_α–C_β bond.[3] Though side chains in proteins, linear peptides, or cyclic peptides may not always adopt a preferred rotamer, there is experimental evidence (vide infra) that they do so in certain instances. It was our objective, therefore, to examine the hypothesis that all nonpolar side chains, aliphatic as well as aromatic, prefer a specific rotamer in solution.

Evidence for preferred side-chain rotamers has been deduced predominantly from circular dichroism (CD),[4] infrared (ir),[5] and nuclear magnetic resonance (NMR)[6] data on small peptides. The majority of such studies have been restricted to compounds with aromatic[6a] or polar side chains,[6d] since these groups cause perturbations which can be readily detected experimentally. It has been deduced that in the cyclic dipeptides, *cyclo*(Gly-L-Phe) and *cyclo*(Gly-L-Tyr), the aromatic rings of phenylalanine and tyrosine adopt a preferred conformation in which they are folded over the diketopiperazine (DKP) ring (see Figure 1 for definition of rotamer nomenclature). This rotamer is also found in the cyclic dipeptide, *cyclo*(L-Pro-D-Phe) (Figure 2, b).[6c] In *cyclo*(L-Pro-L-Phe), on the other hand, Blaha and co-workers have found from a combination of ir[5] and NMR[6c] data that the rotamer with the aromatic ring extended toward nitrogen (Figure 2, c) is preferred. In *cyclo*(L-Tyr-Gly-Gly)$_2$, however, Kopple et al. have concluded that the aromatic ring is extended toward oxygen (Figure 2, e).[7] In a few cases proton NMR coupling constant data have been used to infer relative rotamer populations.[6b] The interpretation of such coupling constant data is ambiguous, however, due to (a) uncertainty of the limiting values of trans and gauche H_α–H_β vicinal coupling constants, (b) equality of the coupling constants for the two extended conformations (Figure 1), and (c) nearly identical coupling constants for a folded side chain and one freely rotating about the C_α–C_β bond. In cyclic peptides without aromatic rings or polar side chains, most reported work has been limited to theoretical energy calculations[3a-c] and x-ray crystallography.[8]

In a previous communication[9] we reported the application of the lanthanide-assisted NMR method to *cyclo*(L-Pro-L-Pro) and *cyclo*(L-Pro-D-Pro). In this paper we emphasize the use of the lanthanide method for the determination of the *amino* acid side-chain conformation in cyclic dipeptides containing proline and an amino acid with a nonpolar side chain. The inclusion of proline in the cyclic dipeptides increases the solubility of the peptides in chloroform, a solvent which is compatible with the shift reagent. In addition, the pyrrolidine ring decreases the conformational mobility of the whole molecule so that the set of side-chain rotamers for the nonprolyl residue is well defined. The nearly invariant positions of the atoms of the bicyclic backbone (the DKP and pyrrolidine rings) can be used to define the lanthanide binding sites.

In addition to relatively unambiguous determination of the preferred dihedral angle[10] for the nonprolyl side chains about the C_α–C_β bond (χ_1) and, in some cases, of that about the C_β–C_γ bond (χ_2), the lanthanide method is of considerable further value, since (a) the proton spectrum is expanded allowing elucidation of otherwise indeterminable coupling constants; (b) conformational information concerning the DKP backbone can be obtained; and (c) most carbon and proton signals can be distinguished and accurately assigned.

Experimental Section

NMR Spectra. Proton spectra with ytterbium and europium were obtained on a Varian HA-100 NMR spectrometer equipped with Fourier transform capability. ^{13}C spectra and homonuclear proton spin decoupled spectra were obtained on a Varian XL-100 spectrometer. The same samples, ca. 0.04 g in 0.5 ml of solution in a 5-mm tube, were used to obtain both proton and carbon spectra, thus minimizing error due to weighing or pipetting. Additions of ytterbium were made by microliter pipette from freshly prepared solutions. Mole ratios up to 1:10 (lanthanide:peptide) were used to obtain all necessary data for carbon atoms. Most proton data could be obtained with ratios of 1:5 or less, but separation of many H_γ protons required additions as high as 1:2. At these high ratios some curvature in a plot of chemical shift vs. lanthanide concentration is detectable, and, therefore, the shift values for these protons, as well as the initial chemical shifts obtained by extrapolation, are probably less accurate.

Shift Reagents. $Yb(fod)_3$ and $Eu(fod)_3$-d_{27} were obtained from Alfa Chemical Co. and used without further purification. Fresh solutions of these complexes in chloroform-*d*, predried by treatment with activated molecular sieves, were prepared for each use.

Cyclic Dipeptides. *tert*-Butyloxycarbonylamino acids and amino acid methyl esters were obtained from Fox and Schwartz-Mann Chemical Co. Boc-D-Ile-OH was prepared from D-Ile and *tert*-butoxycarbonyl azide.[11] All linear dipeptides were prepared by the mixed anhydride method. All cyclic dipeptides were prepared by the method of Nitecki et al.[12] The cyclic dipeptides are characterized in Table I. They were readily crystallized from ether, acetone, or ether–acetone mixtures.

Methods. As in previous work,[9] experimental data yield the least-squares slopes for the chemical shifts of each ^{1}H and ^{13}C atom plotted against the lanthanide–peptide molar ratio. Figure 3 illustrates that these plots are linear, as required for subsequent analysis. As ytterbium shift reagents act almost exclusively by a pseudocontact mechanism,[13a] these compounds were used for determination of slopes in the chemical

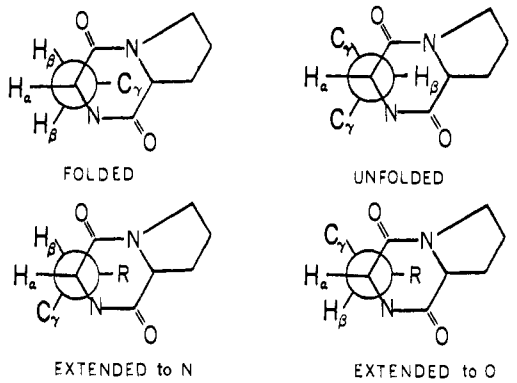

Figure 1. Low-energy side-chain rotamers available to X residues in Pro-X cyclic dipeptides. For X residues with branching at the β carbon, R = C_γ; for residues without branching, R = H_β.

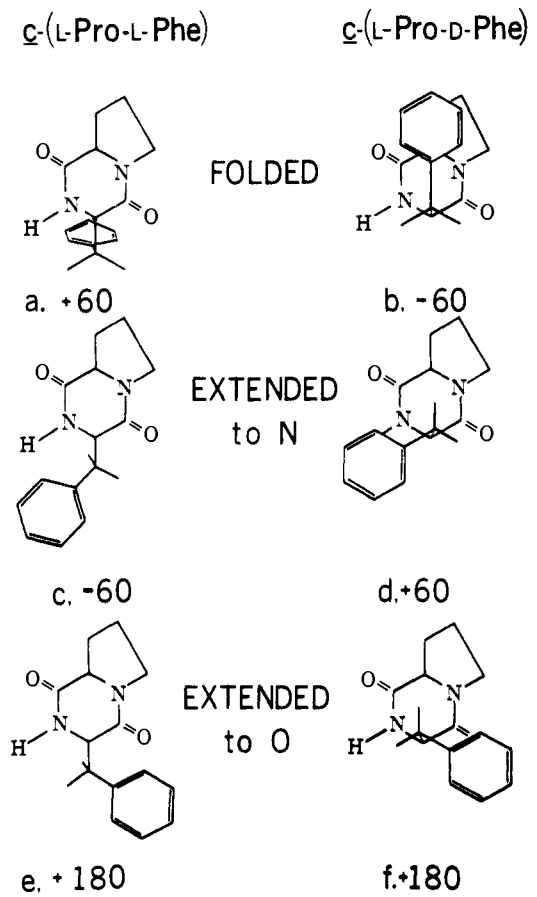

Figure 2. Comparison of the three available rotamers for the two phenylalanine-containing dipeptides. The numbers below each conformer are the χ_1 values according to ref 10. χ_1 describes the dihedral angle about the C_α–C_β bond and χ_2 describes that about the C_β–C_γ bond. The zero angle for χ_1 is that at which the C_α–N bond is eclipsed with the C_β–C_γ bond. Zero for χ_2 is analogously defined with the C_α–C_β and C_γ–C_δ bonds. When there is more than one C_γ or C_δ atom, the normal IUPAC rules for dominance apply. Clockwise rotation of the atoms attached to C_α or C_β is positive for χ_1 and χ_2, respectively, when the molecule is viewed so that these atoms are held closer to the observer.

shift plots. Other lanthanides, such as europium, are less suitable for this purpose, as they may have a significant contact component in the induced chemical shift (especially in carbon spectra).[13b]

The general approach requires minimization of the difference between observed and calculated ytterbium chemical shifts ($\delta\nu/\nu$) for a set of atoms to give "the best match" using the equation[14]

$$\frac{\delta\nu}{\nu} = K\left(\frac{3\cos^2\lambda - 1}{r^3}\right)$$

where K is a constant, and the geometric parameters, λ and r, are

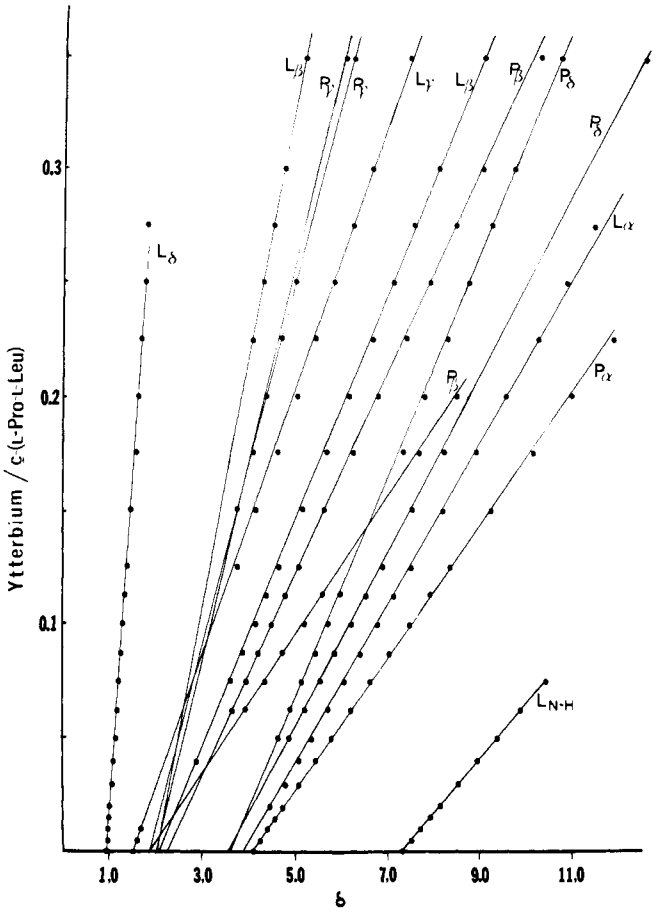

Figure 3. Plots of chemical shifts of protons of *cyclo*(L-Pro-L-Leu) at indicated mole ratios of Yb(fod)$_3$-*cyclo*(L-Pro-L-Leu).

defined in Figure 4.

For all the dipeptides, the bicyclic backbone composed of the DKP and pyrrolidine rings was assumed to have identical coordinates[15]—obtained from calculations[1] on *cyclo*(L-Pro-Gly). This assumption is supported by measurement of coupling constants of proline protons, by computations of conformational energies, and by crystallographic results.[16] A computer program, SHIFT, locates ytterbium in two binding sites relative to the fixed peptide backbone to give the best match of the lanthanide data for the 17 backbone hydrogen and carbon atoms. With the lanthanide positions determined, the shifts of the 9–13 side-chain atoms could be predicted for various possible rotamers and compared to the observed shifts to determine the best match—the one with the lowest R factor.[17]

To locate the best rotamers quickly, the matches for all staggered rotamers were calculated and compared to the observed shifts. The conformational space about the best match was then explored more thoroughly. In a second approach a value for χ_1 was determined by matching experimental shifts with those calculated for the atoms directly bonded to C_β. That is, R factors were determined at 10° rotational intervals about the C_α–C_β bond. For the best values (smallest R factor) of χ_1 so determined, χ_2 was then rotated at 10° intervals. In the region of the best match χ_1 and χ_2 were then both varied in steps of 5°. The two approaches produce the same preferred rotamers, reinforcing our belief that the rotamer which gives the absolute minimum R factor has been located. The ratio of R factors for other conformers to the minimum R factor gives the confidence level that the conformer of minimum R factor is more highly populated. We will generally say that a particular conformer is the "preferred rotamer" if the confidence level, when compared to all other examined rotamers, is greater than, or equal to, 90% (see ref 18).

To further test the hypothesis that a particular side-chain conformer is dominant, the fractional population was determined for each rotamer deemed populated from energy calculations.[1] A computer program, MOLE, determines the mole fraction of each rotamer which will give the best match between predicted and observed lanthanide shifts. The rotamers considered for each compound included the

Table I. Cyclic Dipeptide Characterizations

Compd	Mp, °C	Precursor	Yield,[a] % (crystn solvent)	Mol ion, found (calcd)	C, found (calcd)	H, found (calcd)	N, found (calcd)
c(L-Pro-L-Phe)	125–127	Boc-L-Phe-L-Pro-OMe	44 (ether)	244.123 (244.121)	68.45 (68.85)	6.49 (6.56)	11.31 (11.48)
c(L-Pro-L-Val)	166–168	Boc-L-Val-L-Pro-OMe	27 (ether)	196.122 (196.121)	61.28 (61.22)	8.25 (8.16)	14.21 (14.29)
c(L-Pro-L-Leu)	151–153	Boc-L-Pro-L-Leu-OMe	39 (ether)	210.137 (210.137)	62.27 (62.86)	8.49 (8.57)	13.08 (13.33)
c(L-Pro-D-Phe)	136–138	Boc-L-Pro-D-Phe-OMe	46 (ether–acetone)	244.119 (244.121)	68.83 (68.85)	6.50 (6.56)	11.39 (11.48)
c(L-Pro-D-Val)	139–141	Boc-L-Pro-D-Val-OMe	14 (ether)	196.121 (196.121)	61.23 (61.22)	8.11 (8.16)	14.16 (14.29)
c(L-Pro-D-Leu)	143–145	Boc-L-Pro-D-Leu-OMe	14 (acetone)	210.135 (210.137)	62.91 (62.86)	8.67 (8.57)	13.24 (13.33)
c(L-Pro-D-*allo*-Ile)	107–108	Boc-L-Pro-D-*allo*-Ile-OMe	9 (hexane–ether–acetone)	210.139 (210.137)	63.02 (62.86)	8.47 (8.57)	13.38 (13.33)
c(L-Pro-D-Ile)	161–163	Boc-D-Ile-L-Pro-OMe	41 (acetone)	210.138 (210.137)	62.78 (62.86)	8.57 (8.57)	13.31 (13.33)
c(L-Pro-D-Pro)	179–181	Boc-D-Pro-L-Pro-OMe	68 (acetone)	194.102 (194.106)	61.67 (61.86)	7.05 (7.22)	14.18 (14.43)
c(L-Pro-L-Tyr)	152–154	Boc-L-Tyr-L-Pro-OMe	15 (chloroform)	260[c] (260)	64.57 (64.60)	6.16 (6.20)	10.72 (10.76)
c(D-Pro-L-Tyr)	218–220	Boc-D-Pro-L-Tyr-OEt	11 (ether–acetone)	260[c] (260)	64.68 (64.60)	6.15 (6.20)	10.75 (10.76)
c(L-Pro-L-Ala)	165–168	Boc-L-Ala-L-Pro-OMe	60 (ethanol–ether)	168[c] (168)	56.67 (57.14)	7.01 (7.19)	16.42 (16.65)
c(L-Pro-D-Ala)	139–142	Cbz-D-Ala-L-Pro-OMe[b]	50 (ethanol)	168[c] (168)	56.76 (57.14)	7.06 (7.19)	16.59 (16.65)

[a] Yield is of pure, recrystallized solid. [b] Prepared by hydrogenation of dipeptide precursor. [c] Low-resolution mass spectra.

preferred rotamer from the *R* factor analysis and up to four additional rotamers whose computed intramolecular potential energies were within 3 kcal/mol of the global minimum.[1] In no case were conformers considered which were more than 3 kcal/mol above the calculated global minimum. Since for *cyclo*(L-Pro-D-Ile) no preferred rotamer was predicted, only minimum energy rotamers were utilized.

Relative side-chain rotamer populations were also calculated from vicinal H_α–H_β coupling constants. Addition of Eu(fod)$_3$-d_{27} separated the chemical shifts of the protons without loss of fine structure so that the coupling constants could be measured.[19] For compounds with two β protons, mole fractions (P_a and P_b) of the two extended rotamers (one H_β gauche, the second H_β trans to H_α) and the mole fraction (P_c) of the one folded rotamer (both H_β's gauche to H_α) were calculated from the observed coupling constants (J_{o_1} and J_{o_2}) by the method of Pachler.[20a] Trans (J_t) and gauche (J_g) proton coupling constants of 13.6 and 2.6 Hz, respectively, were assumed.[20b]

$$P_a = (J_{o_1} - J_g)/(J_t - J_g)$$

$$P_b = (J_{o_2} - J_g)/(J_t - J_g)$$

$$P_c = 1 - (P_a + P_b)$$

For the valine- and isoleucine-containing compounds there is only one observed vicinal coupling constant and, thus, only the population of the unfolded rotamer (H_β and H_α trans) can be estimated.

Chemically equivalent atoms (e.g., the three protons of a methyl group or the two ortho carbons in the phenylalanine ring) were not resolved in NMR spectra even in the presence of lanthanide ions. This indicated magnetic equivalence of the atoms could be achieved by free rotation about the preceding bond or, alternatively, by a mechanism in which each of the atoms would have the same occupancy time in a given spatial site. For example, in the phenylalanine ring the two ortho carbons will be magnetically equivalent either if there is free rotation about the C_α–C_β bond or if there is rapid conversion between two geometrically equivalent values of χ_2, such as 90° and −90°. We adopted a model corresponding to the latter possibility in the calculation of lanthanide shifts by averaging the values for the three methyl protons, two ortho carbons, etc., with the side chain fixed in one of the geometrically equivalent conformers.

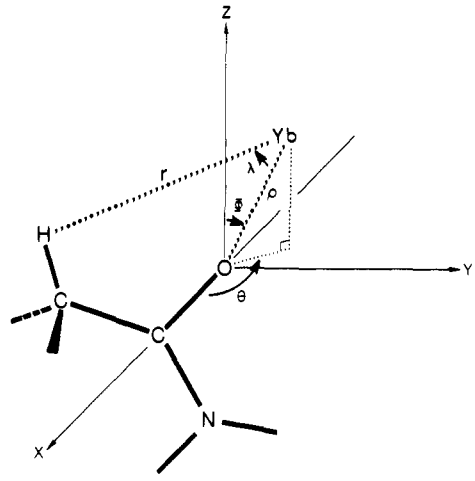

Figure 4. Coordinate system used to locate lanthanide ions. r is the lanthanide–proton (or carbon) distance. ρ is the lanthanide–oxygen distance given in Table I. λ is the proton- (or carbon) lanthanide–oxygen angle. θ and ϕ are the angular spherical coordinates of the lanthanide with the origin of oxygen.

Results

Ytterbium Positions. Binding of ytterbium at both carbonyl oxygen atoms in each dipeptide has been assumed. The ytterbium positions (see Figure 4 and Table II) determined by the analyses are reasonably constant with $\rho = 2.4 \pm 0.2$ and the Yb–O–C angle varying from 137 to 163° for $X_{C=O}$ and from 161 to 173° for $Pro_{C=O}$. At the nonprolyl carbonyl binding site, the lanthanide is located on the face of the carbonyl away from nitrogen (toward C_α) and on the same side of the DKP ring as the D side chain. The proline carbonyl binds the lanthanide more nearly in the amide plane but still on the face away from the nitrogen. The latter lanthanide binding site is on the side of the DKP ring away from the D side chain.

Table II. Ytterbium Positions

Compd	ρ, Å	Angle Yb-O-C	K	R
c(L-Pro-L-Pro)	2.5	169.8	1333	0.045
c(L-Pro-L-Phe)				
L-Pro	2.5	163.7	1305	0.030
L-Phe	2.5	149.5	1904	
c(L-Pro-L-Val)				
L-Pro	2.6	172.7	1232	0.049
L-Val	2.5	163.1	2225	
c(L-Pro-L-Leu)				
L-Pro	2.6	160.7	2092	0.062
L-Leu	2.5	155.2	1273	
c(L-Pro-D-Phe)				
L-Pro	2.5	163.2	2239	0.043
D-Phe	2.6	147.5	1319	
c(L-Pro-D-Val)				
L-Pro	2.5	171.2	2312	0.051
D-Val	2.4	151.7	1183	
c(L-Pro-D-Leu)				
L-Pro	2.5	167.5	2526	0.066
D-Leu	2.5	137.2	715.3	
c(L-Pro-D-Ile)				
L-Pro	2.4	164.9	2262	0.040
D-Ile	2.2	157.4	782.4	
c(L-Pro-D-*allo*-Ile)				
L-Pro	2.4	167.6	2145	0.042
D-*allo*-Ile	2.5	149.5	1178	

^{13}C Chemical Shifts. The ^{13}C chemical shifts of the various cyclic dipeptides are listed in Table III. Assignment of all resonances could be made with the aid of shift reagents, proton-coupled spectra, and previous data on amino acids.[21]

In all cases but one, the proline α carbon was readily identified as the resonance between 133.6 and 134.6 ppm (upfield of CS_2). The lanthanide data suggested that in *cyclo*(L-Pro-D-Phe) the peak at 133.95 be assigned to the phenylalanine α carbon and the peak at 135.05 to the proline α carbon.[22]

Backbone Conformation. The use of europium to separate resonances led to the observation that the proline protons in all the cyclic dipeptides exhibit very similar coupling constants. In Figure 5 the upfield region of the proton spectra of *cyclo*(L-Pro-L-Pro), *cyclo*(L-Pro-L-Leu), and *cyclo*(L-Pro-D-Val) are compared. Once the assignments for the proline protons of one compound are known, those for all three compounds can be obtained simply by comparing coupling patterns. Since we have simulated[9] the spectrum of *cyclo*(L-Pro-L-Pro), assignments of the proline protons can be made with confidence and are readily confirmed by spin decoupling. Perhaps the most significant observation is that the proline ring in a bicyclic DKP retains the same conformation whether the second side chain is L or D. These results are not surprising in view of our earlier finding[9] that the pyrrolidine rings were similar in *cyclo*(L-Pro-L-Pro) and *cyclo*(L-Pro-D-Pro), DKP's which may represent the two extremes of backbone conformation within the *cyclo*(L-Pro-X) series.

Insofar as the proline ring geometry is sensitive to changes in the DKP backbone conformation, these observations also strengthen our initial assumptions that the same set of backbone coordinates may be used for both LL and LD dipeptides.

Side-Chain Rotamers. Madison et al.[1] have calculated the rotamers that would be expected for the side chain of each dipeptide if steric factors alone were responsible for the preference. For each compound the experimentally preferred rotamer from the lanthanide analysis is compared with that computed to have lowest energy (Table IV). The nomenclature used to define these rotamers is described in Figure 1. Recall that, although β-branched amino acids may have one C_γ folded and a second C_γ extended, we have chosen to classify them according to the extended C_γ to avoid ambiguity.

The results of the side-chain rotamer analysis can be summarized as follows.

(1) For seven of the eight compounds examined, the lanthanide method predicts a preferred rotamer. [The one exception is *cyclo*(L-Pro-D-Ile).] The lanthanide analysis and energy calculations predict the same preferred rotamer in all cases except for *cyclo*(L-Pro-L-Phe) which has a solvent-dependent rotamer distribution.

(2) In chloroform solution five of the compounds that exhibit a preferred rotamer are in the conformation we have designated "extended to nitrogen". For the remaining compounds the preferred rotamer is folded for *cyclo*(L-Pro-D-Phe) and unfolded for *cyclo*(L-Pro-D-Val).

(3) In the seven cases in which a preferred rotamer is found with a confidence level greater than 90%, application of the rotamer averaging method does not indicate a different preference, although in two cases inclusion of more than one rotamer improves the *R* factor by 20% or more. In the one case where no preference for a single rotamer was indicated, *cyclo*(L-Pro-D-Ile), the averaging method indicates a significant percentage of each of the three χ_1 rotamers.

(4) Examination of Table III in conjunction with Table IV suggests that the ^{13}C chemical shifts for the γ-methyl carbon

Table III. Carbon Chemical Shifts of Cyclic Dipeptides[a]

Compd	$X_{C=O}$	X_α	X_β	X_{γ_1}	X_{γ_2}	X_{δ_1}	X_{δ_2}	$\chi_{o,m,p}$	$Pro_{C=O}$	Pro_α	Pro_β	Pro_γ	Pro_δ
c(L-Pro-Gly)	29.15	146.01							22.73	134.13	164.22	170.27	147.40
c(L-Pro-L-Pro)									26.27	132.13	165.00	169.32	147.46
c(L-Pro-L-Phe)	27.56	136.32	155.85	56.49				63.30 (o), 63.68 (m), 65.35 (p)	23.15	133.60	164.36	170.23	147.35
c(L-Pro-L-Val)	27.53	132.13	164.16	176.62	173.73				22.11	133.89	164.16	170.31	147.59
c(L-Pro-L-Leu)	26.24	139.15	154.12	168.14		169.46	171.27		21.96	133.71	164.64	169.91	147.26
c(L-Pro-D-Pro)									28.36	131.56	162.78	170.88	147.59
c(L-Pro-D-Phe)	27.63	133.95	152.34	57.19				62.63 (o), 64.16 (m), 65.33 (p)	23.01	135.05	163.81	171.09	147.71
c(L-Pro-D-Val)	27.17	129.29	159.64	174.93	173.73				22.76	134.34	163.31	170.72	147.12
c(L-Pro-D-Leu)	26.05	136.42	149.94	168.17		169.65	171.16		22.79	134.60	163.67	170.43	147.10
c(L-Pro-D-Ile)	27.27	129.93	152.92	167.99	177.38	181.38			22.85	134.29	163.28	170.70	147.12
c(L-Pro-D-*allo*-Ile)	26.98	130.91	152.90	166.87	178.38	181.19			23.00	134.31	163.20	170.80	147.14

[a] In chloroform-*d*; parts per million upfield from CS_2.

Table IV. Rotamer Populations of Cyclic Dipeptides

Compd	$J_{\alpha\beta}$, Hz	Nature of rotamer[a]	χ_1[b]	χ_2[b]	Fraction of population		R factor[e]	Confidence level, %
c(L-Pro-L-Phe)	10.0	Folded	+60	+90	0.33[c]	(0.24)[d]	0.082, 0.063	90[f]
	3.6	Extended to N	−50		0.67	(0.67)		
		Extended to O				(0.09)		
c(L-Pro-L-Val)	2.7	Extended to N	−60		0.90	(0.99)	0.057, 0.054	99
		Unfolded	−170		0.10	(0.01)		
c(L-Pro-L-Leu)	9.0	Extended to N	−70	+180			0.077, 0.072	97.5
	4.0	Extended to N	−65[g]	+175[g]	0.51	(0.58)		
		Extended to N	−80	+70	0.36			
		Folded	+70	+170	0.13	(0.29)		
		Extended to O				(0.13)		
c(L-Pro-D-Phe)	4.3	Folded	−40	+100	0.85	(0.59)	0.059, 0.051	99
	5.5	Extended to N	+65	+75	0.15[h]	(0.26)[h]		
		Extended to O				(0.15)[h]		
c(L-Pro-D-Val)	6.6	Unfolded	+60		1.00	(0.36)	0.081, 0.081	90
		Unfolded	+65			(0.64)		
		Extended						
c(L-Pro-D-Leu)	9.0[i]	Extended to N	+70	+70	0.67	(0.58)	0.092, 0.088	95
	4.0[i]	Extended to N	+70	+40	0.33			
		Extended to N	+80	+160				
		Folded				(0.29)		
		Extended to O				(0.13)		
c(L-Pro-D-*allo*-Ile)	4.0[i]	Extended to N	+60	+170			0.079, 0.051	90
		Extended to N	+60	+160	0.58	(0.87)		
		Extended to O	−70	+180	0.17			
		Unfolded	+180	−170	0.25	(0.13)		
c(L-Pro-D-Ile)	5.5[i]	Extended to O	+170	−170	0.25	(0.74)	0.086, 0.042	50
		Extended to O	+170	−60	0.01			
		Extended to N	−60	+180	0.40			
		Unfolded	+60	+180	0.34	(0.26)		
		Unfolded	+70[j]	+160[j]				

[a] For definitions, see Figure 1. [b] The first rotamer listed for each compound is the preferred rotamer from minimum energy calculations. [c] Populations from lanthanide rotamer averaging method. [d] Populations from coupling constant analysis. These fractional populations were assigned to individual rotamers in accord with the lanthanide results. [e] The first R factor is for preferred rotamer before averaging. The second R factor is for the mixture of all rotamers included in the averaging method. [f] This confidence level was calculated for χ_1 only. The data were insufficient for assessment of χ_2. [g] The x-ray values are[8a] $\chi_1 = -72.3°$, $\chi_2 = +177.7°$. [h] This assignment is uncertain due to ambiguity in the assignment of Phe β protons. [i] Lanthanide broadening contributes to increased uncertainty in the coupling constant. [j] Preferred rotamer from the lanthanide method before averaging.

atoms of the valine and isoleucine compounds are at substantially higher field when the methyl group is folded over the DKP backbone. Thus, the folded C_γ of L-Val is at 176.62 ppm, while the unfolded C_γ's of D-Val are 173.73 and 174.93 ppm. The folded C_γ of D-*allo*-Ile is at 178.38 vs. 177.38 ppm for D-Ile (unfolded).

(5) Although the fractional populations calculated by the lanthanide and coupling constant methods are not identical, the predictions of major and minor rotamers are similar in all cases but one. For *cyclo*(L-Pro-D-Val), $J_{\alpha\beta} = 6.6$ Hz, yielding an estimate that $\sim\frac{3}{8}$ of the molecules are unfolded, while the lanthanide data suggest a much stronger preference. It seems likely that the valine trans and gauche coupling constants are lower than those assumed by Pachler, since it has been noted that increasing methyl substitution lowers the observed vicinal coupling constants,[24] and $J_{\alpha\beta}$ also underpredicts the population of the unfolded rotamer for *cyclo*(L-Pro-L-Val).

Discussion

In Table IV, the calculated low-energy rotamers of the cyclic dipeptides are compared with the lanthanide results for the ytterbium–cyclic dipeptide complexes. In all cases the calculated and experimental rotamers are the same (about both χ_1 and χ_2), except for *cyclo*(L-Pro-L-Phe) and *cyclo*(L-Pro-D-Ile). Since several *cyclo*(L-Pro-D-Ile) rotamers appear to be populated, it may be concluded that in all cases where a preferred rotamer exists for an aliphatic side chain, calculations and experiment agree. In fact, the experimental preponderance of one conformer is often greater than would be expected from the differences among calculated energies for the rotamers. Nevertheless, it seems likely that the experimentally determined rotamers for the aliphatic side chains primarily reflect steric interactions. The good agreement between calculation and experiment holds despite the fact that the lanthanide data are extracted from the ytterbium–dipeptide complex, which suggests that the geometries of the dipeptides are not significantly altered by the presence of ytterbium.

In the case of the LL dipeptides, the experimental data indicate that the β substituents largely determine which rotamer is populated. In *cyclo*(L-Pro-L-Val) the two methyl groups occupy the folded and extended toward nitrogen positions. For both *cyclo*(L-Pro-L-Leu) and *cyclo*(L-Pro-L-Phe), the C_γ is extended toward nitrogen in the preferred rotamer. Therefore, in the LL series it is tempting to fix the positions in order of decreasing ability to accommodate steric bulk as (a) extended toward nitrogen, (b) folded, and (c) extended toward oxygen. For alkyl side chains computations are consistent with this ordering, but the steric requirements of the planar aromatic ring differ from those of tetrahedral alkyl chains, so that the folded conformer of *cyclo*(L-Pro-L-Phe) has the lower computed van der Waals energy.[1] The likelihood that nonsteric factors influence the preference of *cyclo*(L-Pro-L-Phe) for the extended rotamer in chloroform will be discussed below.

For the LD diastereomers the bulkiness of the substituent on the γ-carbon atom seems to play a major role in determining the conformational preference. *Cyclo*(L-Pro-D-Val) adopts the unfolded conformation with one methyl group extended toward nitrogen, the other toward oxygen. *Cyclo*(L-Pro-D-Ile) may be derived from this compound by replacing the methyl group extended toward nitrogen with an ethyl group. This

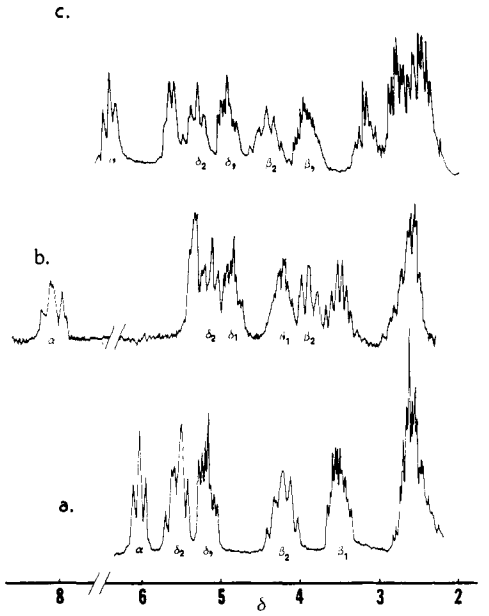

Figure 5. Comparison of the 100-MHz proton spectra (upfield region only) of (a) *cyclo*(L-Pro-L-Pro), (b) *cyclo*(L-Pro-D-Val), and (c) *cyclo*(L-Pro-L-Leu) with sufficient Eu(fod)$_3$-d_{27} to separate all proline α, β, and δ resonances. Europium to substrate molar ratios are, respectively, 0.21, 0.27, and 0.18. The subscripts 1 and 2 signify proline ring protons which are syn and anti, respectively, to the proline α proton.

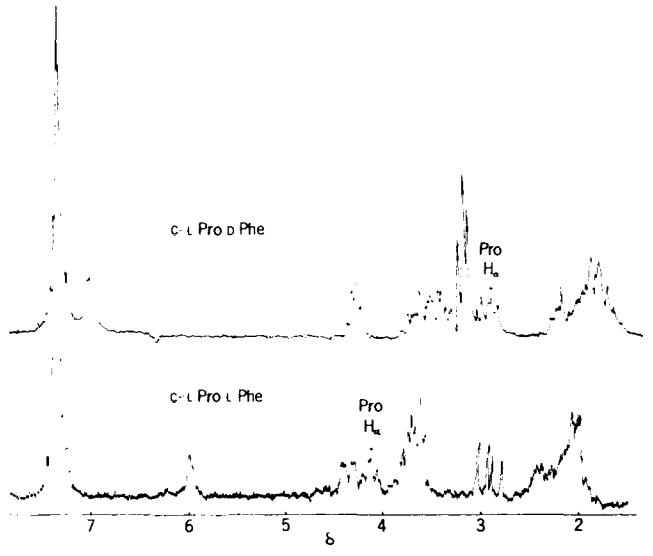

Figure 6. Comparison of the proton spectra of *cyclo*(L-Pro-L-Phe) and *cyclo*(L-Pro-D-Phe) in $CDCl_3$ at 100 MHz.

seemingly minor change removes all preference for the unfolded conformation, and the two extended conformers (Figure 1) are now significantly populated. The configuration at the β carbon of *allo*-Ile is the reverse of that in Ile. *Cyclo*(L-Pro-D-*allo*-Ile) could be derived from the preferred rotamer of *cyclo*(L-Pro-D-Val) by replacing the methyl group extended toward oxygen with an ethyl group. This substitution also destabilizes the unfolded rotamer, presumably due to steric repulsion between the bulky ethyl substituent and the carbonyl group. Apparently, in *cyclo*(L-Pro-D-*allo*-Ile) the ethyl group is best accommodated near the nitrogen which places the methyl group in the folded position over the DKP ring (Figure 1). As in the case of *cyclo*(L-Pro-D-Ile), the other two χ_1 rotamers are also significantly populated. Leucine has only one β substituent, and its side chain in *cyclo*(L-Pro-D-Leu) adopts the conformation extended toward nitrogen. However, this conformational preference results primarily from the steric bulk of the two methyl groups attached to the γ carbon.

In *cyclo*(L-Pro-L-Phe) and *cyclo*(L-Pro-D-Phe) it is likely that nonsteric forces affect the conformational distribution. Calculations indicate that the folded conformer has the lowest energy for both compounds by 0.8 and 0.2 kcal/mol, respectively.[1] The preferred rotamer observed for *cyclo*(L-Pro-L-Phe) in chloroform is extended toward nitrogen and may be stabilized by induced polarization of the electronegative aromatic ring by the proximate electropositive amide hydrogen. Such dipole–induced dipole interactions were not included in our computations. Blaha and co-workers[5] also suggested that an N–H to phenyl ring interaction in *cyclo*(L-Pro-L-Phe) might stabilize the rotamer extended toward nitrogen. Evidence for this type of interaction is found in that under identical experimental conditions the amide proton in *cyclo*(L-Pro-L-Phe) is shifted upfield more than 1 ppm compared to that in *cyclo*(L-Pro-D-Phe) (Figure 6)—consistent with an aromatic ring current shift of the L-Phe amide proton. The fact that in polar solvents the folded conformer predominates over the one extended toward nitrogen is consistent with electrostatic stabilization of the latter. In chloroform solution *cyclo*(L-Pro-D-Phe) adopts the folded conformer which is slightly favored sterically. The fact that the folded conformer predominates over a range of solvents and temperatures suggests that it possesses additional stability perhaps from interactions between the aromatic and peptide π systems.

Conclusions

We have shown that lanthanide-assisted analysis of NMR spectra can identify preferred rotamers of amino acid side chains. In fact, seven of the eight cyclic dipeptides examined have a preferred rotamer about the C_α–C_β bond.

The findings reported herein indicate that steric factors predominate in determining the conformational preferences of amino acid side chains. For the aromatic groups particular rotamers (such as the folded ones) may be stabilized by interactions between the peptide groups and the polarizable aromatic electron cloud. Initial experiments[25] indicate that the lanthanide method will find useful application with larger peptides.

Acknowledgments. This work was supported, in part, by U.S. Public Health Service Grants AM07300 and AM10794. We are grateful to Dr. Eric T. Fossel for his assistance in the taking of numerous NMR spectra, and we thank the Harvard University Chemistry Department for the use of their Varian XL-100 NMR spectrometer. We are also grateful to Dr. K. Biemann and Dr. C. Hignite of Massachusetts Institute of Technology (under National Institutes of Health Grant No. RR00317) for the mass spectra. This work also employed a CFT-20 NMR spectrometer, provided through National Science Foundation Grant GB-41535. We also thank Dr. C. H. Niu for samples of *cyclo*(L-Ala-L-Pro) and *cyclo*(D-Ala-L-Pro). V.M. and P.E.Y. held National Institutes of Health Postdoctoral Fellowships.

References and Notes

(1) For the previous paper in this series, see V. Madison, P. E. Young, and E. R. Blout, *J. Am. Chem. Soc.*, preceding paper in this issue.
(2) See, for example, R. D. B. Fraser and T. MacRae, "Conformation in Fibrous Proteins and Related Synthetic Polypeptides", Academic Press, New York, N.Y., 1973, p 276, and references cited therein.
(3) (a) V. Sasiekharan and P. H. Ponnuswamy, *Biopolymers*, **10**, 583 (1972); (b) R. Chandreskaran, A. V. Lakshminarayanan, P. Monhanakrishnan, and G. N. Ramachandran, *ibid.*, **12**, 1421 (1973); (c) J. Caillet, B. Pullman, and B. Maigret, *ibid.*, **10**, 221 (1971); (d) G. N. Ramachandran and A. V. Lakshminarayanan, *ibid.*, **4**, 495 (1966); (e) C. M. Deber and H. Joshua, *ibid.*, **11**, 2493 (1972).
(4) (a) C. A. Bush and D. E. Gibbs, *Biopolymers*, **11**, 2421 (1972); (b) S. M. Ziegler and C. A. Bush, *Biochemistry*, **10**, 1330 (1971).
(5) J. Vicar, J. Smolikova, and K. Blaha, *Collect. Czech. Chem. Commun.*, **38**, 1957 (1973).
(6) (a) K. D. Kopple and D. H. Marr, *J. Am. Chem. Soc.*, **89**, 6193 (1967); (b) R. Q. Newmark and M. A. Miller, *J. Phys. Chem.*, **75**, 505 (1971), and ref-

erences cited therein; (c) J. Vicar, M. Budesinsky, and K. Blaha, *Collect. Czech. Chem. Commun.*, **38**, 1940 (1973); (d) Ziauddin, K. D. Kopple, and C. A. Bush, *Tetrahedron Lett.*, 483 (1972); (e) K. D. Kopple and M. Ohnishi, *J. Am. Chem. Soc.*, **91**, 962 (1969), and references cited therein; (f) G. Gawne, G. W. Kenner, N. H. Rogers, R. C. Sheppard, and K. Titlestad, "Peptides 1968", E. Bricas, Ed., Wiley, New York, N.Y., 1969, p 28; (g) R. Deslauerlers and I. C. P. Smith, *Biochem. Biophys. Res. Commun.*, **40**, 179 (1970).
(7) (a) K. D. Kopple, M. Ohnishi, and A. Go, *Biochemistry*, **8**, 4087 (1969); (b) *J. Am. Chem. Soc.*, **91**, 4264 (1969).
(8) (a) I. L. Karle, *J. Am. Chem. Soc.*, **94**, 81 (1972); (b) L. E. Webb and C.-F. Lin, *ibid.*, **93**, 3818 (1971).
(9) P. E. Young, V. Madison, and E. R. Blout, *J. Am. Chem. Soc.*, **95**, 6142 (1973).
(10) The conventions followed in this paper are given by IUPAC–IUB Commission on Nomenclature, *Biochemistry*, **9**, 3471 (1970).
(11) E. Schnabel, *Justus Liebigs Ann. Chem.*, **702**, 188 (1967).
(12) D. E. Nitecki, B. Halpern, and J. W. Westley, *J. Org. Chem.*, **33**, 864 (1968).
(13) (a) J. Reuben, *J. Magn. Reson.*, **11**, 103 (1973); (b) M. Hirayama, E. Edagawa, and Y. Hamyu, *Chem. Commun.*, 1343 (1972); S. R. Johns, R. A. Smith, G. E. Hawkes, and J. D. Roberts, *Proc. Natl. Acad. Sci. U.S.A.*, **70**, 939 (1973); B. F. G. Johnson, J. Lewis, P. Mardle, and J. R. Norton, *Chem. Commun.*, 535 (1972).
(14) (a) For general reviews, see R. von Ammon and R. D. Fischer, *Angew. Chem., Int. Ed. Engl.*, **11**, 675 (1972); (b) J. K. M. Sanders and D. H. Williams, *J. Am. Chem. Soc.*, **93**, 641 (1971); (c) R. E. Sievers, Ed., "Nuclear Magnetic Resonance Shift Reagents", Academic Press, New York, N.Y., 1973; (d) B. A. Levine and R. J. P. Williams, *Proc. R. Soc., London, Ser. A*, **345**, 5 (1975).
(15) The analysis was also completed for *cyclo*(L-Pro-L-Leu) using the x-ray crystal coordinates. There was no significant difference between the two sets of results.
(16) See Table IV in I. L. Karle, H. C. J. Ottenheym, and B. Witkop, *J. Am. Chem. Soc.*, **96**, 539 (1974).
(17) (a) M. R. Wilcott, III, R. E. Lenkinski, and R. E. Davis, *J. Am. Chem. Soc.*, **94**, 1742 (1972); (b) R. E. Davis and M. R. Wilcott, III, *ibid.*, **94**, 1744 (1972).
(18) Recently, K. L. Servis and D. J. Bowler, *J. Am. Chem. Soc.*, **97**, 80 (1975), have applied a similar method, without the use of energy calculations, to the analysis of side-chain rotamers in alkylcyclohexanones.
(19) Although the measured coupling constants are necessarily an average of those of the complexed and uncomplexed substrate, several considerations make their use for the estimation of rotamer populations of the uncomplexed substrate seem reasonable. First, $J_{\alpha\beta}$ could almost always be extracted for spectra containing less than 0.05 mol of europium per mole of substrate—most of the peptide is *not* complexed. Second, the coupling constants did not change significantly over the whole range of additions. Indeed, even those coupling constants that could be measured before europium addition remained nearly constant as the europium mole fraction increased. Third, in those cases where preferred rotamers have been determined by other methods [as for *cyclo*(L-Pro-D-Phe)[6c] and *cyclo*(L-Pro-L-Leu)[8a]], agreement with the present results is good.
(20) (a) K. G. R. Pachler, *Spectrochim. Acta*, **20**, 581 (1964); (b) R. J. Abraham and K. A. McLachlan, *Mol. Phys.*, **5**, 513 (1962).
(21) W. Horsley, H. Sternlicht, and J. S. Cohen, *J. Am. Chem. Soc.*, **92**, 680 (1970).
(22) As a check on these assignments, *cyclo*(L-Pro-D-Phe), partially deuterated in the phenylalanine α position, was synthesized. Its ^{13}C NMR spectrum confirmed that the more shielded α carbon is that of proline. The proline α carbon is thus 0.66 ppm higher field than the average of the four proline α carbons of the other LD compounds and 0.92 ppm higher field than that of *cyclo*(L-Pro-Gly). The proline α proton of *cyclo*(L-Pro-D-Phe) is also shifted upfield. As shown in Figure 6, in *cyclo*(L-Pro-D-Phe) the proline H_α appears at δ 2.85, as compared to δ 4.05 in *cyclo*(L-Pro-L-Phe). This upfield shift in *cyclo*(L-Pro-D-Phe) is due to shielding by the aromatic ring in the folded conformation (Figure 2, b). The upfield shift of the proline α carbon is about one-half that of the proline α proton and may be at least partially due to thru-space anisotropic shielding effects of the aromatic ring.[23]
(23) (a) R. H. Levin and J. D. Roberts, *Tetrahedron Lett.*, 135 (1973); (b) R. Deslauriers, Z. Gzonka, K. Schaumberg, T. Shiba, and R. Walter, *J. Am. Chem. Soc.*, **97**, 5093 (1975), and references cited therein.
(24) G. N. Ramachandran, R. Chandrasekaran, and K. D. Kopple, *Biopolymers*, **10**, 2113 (1971).
(25) (a) K. L. Servis and D. J. Patel, *Tetrahedron*, **31**, 1359 (1975); (b) P. E. Young, V. Madison, C.-H. Niu, and E. R. Blout, "Peptides: Chemistry, Structure, and Biology", R. Walter and J. Meienhofer, Ed., Ann Arbor Science Publishers, Ann Arbor, Mich., 1975, p 187.

Spatial Configuration of Ordered Polynucleotide Chains. 3. Polycyclonucleotides[1]

Wilma K. Olson* and Rama D. Dasika

Contribution from the Department of Chemistry, Douglass College, Rutgers, the State University, New Brunswick, New Jersey 08903. *Received October 31, 1975*

Abstract: Approximate details of the spatial configuration of the ordered 8-2′-purine polycyclonucleotide chain in dilute solution are reported from a combined theoretical analysis of chain flexibility and base stacking. The limited experimental evidence detailing the rotational preferences in this molecule is here supplemented by semiempirical energy estimates of conformation. Only a fraction of the wide variety of regular polycyclonucleotide helices generated on this basis is found to accommodate the array of stacked bases that characterizes the ordered form of the molecule. The bases comprising these stacked helices are arranged, as has been observed previously, almost exclusively in left-handed stacking patterns. In contrast to earlier suggestions, however, the backbone structures to which the stacked polycyclonucleotide bases attach are observed to be right-handed helices. In fact, the sugar–phosphate units of these polycyclonucleotide helices are identical with backbone conformations deduced in x-ray fiber diffraction analyses of ordered double-stranded polynucleotides. Unlike the bases aligned in planes approximately perpendicular to the long axes of ordered polynucleotide chains, the polycyclonucleotide bases attached to the same backbone frameworks are found to stack in planes that approximately parallel the helix axis. This unusual parallel alignment permits the bases of the polycyclonucleotide simultaneously both to exhibit left-handed stacking and to conform to right-handed helical organization.

It has been well established that the introduction of covalent chemical linkages between the sugar and base moieties in a so-called polycyclonucleotide chain will alter the physical and biological properties of the nucleic acid system.[2-10] This effect is especially dramatic in the case of 8-2′-purine and 6-2′-pyrimidine polycyclonucleotide chains where the covalent attachment between the base and 2′-carbon fixes the glycosyl rotation χ (describing the mutual orientation of the two moieties) in the unusual so-called[11] "high anti" conformation. This conformation is an approximately intermediate arrangement between the commonly occurring anti and syn glycosyl conformations. As evident from Figure 1 [where a unit of the 8-2′-anhydro-8-X-9-(β-D-arabinofuranosyl)adenyl or poly(cyclo-A) chain is depicted], the base plane in the high anti orientation approximately parallels the sense of direction of the sugar–phosphate backbone repeating unit. The rotation angle defined by atoms C(2′)–C(1′)–N(9)–C(4) of this unit is found in a trans arrangement. This unusual conformation also characterizes the crystal structures of a number of naturally occurring azanucleosides including formycin,[11] 8-azaadenosine,[12] and 6-azacytidine.[13,14] In the unmodified polynucleotide chain, on the other hand, the anti or syn bases are oriented approximately perpendicular to the direction of the chain repeat unit. In the former conformation the rotation