The Link between Sequence and Conformation in Protein Structures Appears To Be Stereochemically Established

ARTICLE *in* THE JOURNAL OF PHYSICAL CHEMISTRY B · JUNE 2006

Impact Factor: 3.3 · DOI: 10.1021/jp056417e · Source: PubMed

4 AUTHORS, INCLUDING:

Vibin Ramakrishnan
Indian Institute of Technology Guwahati
13 PUBLICATIONS **72** CITATIONS
SEE PROFILE

Ranjit Ranbhor
Sun Pharmaceutical Industries
6 PUBLICATIONS **68** CITATIONS
SEE PROFILE

Anil Kumar
University of Toronto
13 PUBLICATIONS **104** CITATIONS
SEE PROFILE

The Link between Sequence and Conformation in Protein Structures Appears To Be Stereochemically Established

Vibin Ramakrishnan,[†,§] Ranjit Ranbhor,[†] Anil Kumar,[‡] and Susheel Durani*[,‡]

School of Biosciences and Bioengineering and Department of Chemistry, Indian Institute of Technology Bombay, Mumbai 400076, India

Received: November 7, 2005; In Final Form: February 24, 2006

In search of the link between sequence and conformation in protein structures, we perform molecular dynamics analysis of the effect of stereochemical mutation in end-protected octa-alanine Ac-Ala_8-NHMe from poly-L to an alternating-L,D structure. The mutation has a dramatic effect, transforming the peptide from a condition of extreme sensitivity to one of extreme insensitivity to solvent. Examining the molecular folds of poly-L and alternating-L,D structure in atomistic detail, we find them to differ in the relationship between peptide dipolar interactions at the local and nonlocal levels, either conflicting or harmonious depending upon the chain stereochemistry. The stereochemical transformation of interpeptide electrostatics from a condition of conflict to one of harmony explains the long-standing puzzle of why poly-L and alternating-L,D peptides strongly differ in properties such as "stiffness" and solvent sensitivity. Furthermore, it is possible that poly-L stereochemistry is also the fulcrum of protein sensitivity to the effects of amino acid side-chain structures via dielectric arbitrations in interpeptide electrostatics. Indeed the evidence is accumulating that the amino acid side chains differing in α-helix and *β*-sheet propensities also differ in their desolvating effects in the adjacent and nearest-neighbor peptides and thus possibly in the solvent screening of peptide dipolar interactions.

Introduction

Encoding information chemically, nature achieves its physical materialization in the process of protein folding. The underlying mechanisms remain poorly understood and characterize the protein-folding problem. The problem is less about predicting conformation—as knowledge-based prediction algorithms seem feasible[1]—but more about knowing the chemical specificity of the conformation for underlying causes. The choices in conformation are few, but the variation in chemical sequential structures is astronomical given the diversity in side-chain sizes, shapes, polarities, types of polar groups, and their location with respect to the main chain, etc. The effects will manifest conformationally, but the causes—steric, entropic, hydrophobic, electrostatic, direct or solvent-mediated, etc.—are complex and have been difficult to resolve. For instance leucine and isolucine, identical in structure except for the location of the branch point, differ in their effects; the entropic penalty for the side chain,[2,3] interactions between main-chain and side-chain atoms,[4] and differential masking of the main chain from the solvent[5,6] were suggested reasons in different studies.

Protein structures not only are sensitive to the chemical effects of their own sequences but also of their environment, such as solvent. The effect is captured in simple oligoalanines even though lacking in sequence differentiation like a protein.[7,8] The models will emulate the different conformational states of the protein structure in a length-dependent manner. The shorter versions will emulate the protein denatured states being predominantly PPII in conformation,[8] while the versions long enough to be stabilized cooperatively are α-helices that are proteinlike in their sensitivity to solvents, denaturants, and even side-chain effects.[7] Cooperative ordering to conformations that are sensitive to solvents, denaturants, and amino acid side chains is thus a generic of the main-chain structure in oligoalanines and therefore proteins.

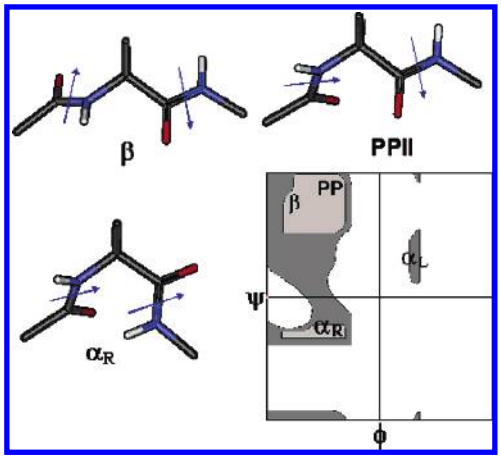


Figure 1. *β* ($\phi = -120°$, $\psi = 120°$), PPII ($\phi = -75°$, $\psi = 145°$), and α ($\phi = -57°$, $\psi = -47°$) conformers of L-alanine dipeptide, showing the placements in the Ramachandran diagram and the mutual alignment of peptide dipole vectors, which are antiparallel in the *β*-conformer, parallel in the α-conformer, and more or less perpendicular in the PPII conformer.

The fundamental conformational alternatives of the protein structure are the α-helix and *β*-sheet types of patterning of its peptide dipoles (Figures 1 and 2).[9,10] The alternatives are of differing sensitivities to the amino acid side-chain effects but for reasons unknown. Varied explanations are invoked, and electrostatic interaction between peptide dipoles is one of them.[5,11] Differing in mutual orientation of the dipoles, as noted in Figures 1 and 2, and therefore in peptide dipolar interactions,

* Author to whom correspondence should be addressed. Phone: +91-22-25767164. Fax: +91-22-25767152. E-mail: sdurani@iitb.ac.in.
† School of Biosciences and Bioengineering.
‡ Department of Chemistry.
§ Present address: Department of Biology, Rensselaer Polytechnic Institute, Troy, NY 12180.

10.1021/jp056417e CCC: $33.50
Published on Web 04/14/2006

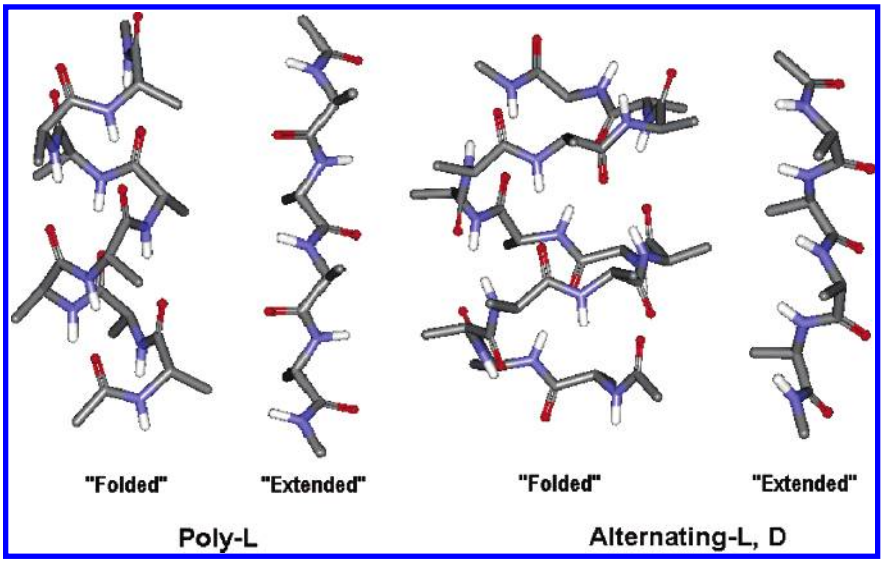


Figure 2. "Folded" and "extended" conformers of poly-L and alternating-L,D alanines, highlighting the differing dipole orientations. The poly-L alanine is folded in the α-helical conformation (dipoles parallel) and extended in the β-conformation (dipoles antiparallel), while alternating-L,D alanine is folded in the β-conformation (dipoles antiparallel) and extended in the α-conformation (dipoles parallel). The ϕ,ψ values in the modeled structures are defined in the Materials and Methods section.

the α-helix and β-sheet structures could differ in their sensitivities to the dielectric constant of the medium. Thus the argument is made that the side chains, controlling accessibility in their adjacent and nearest-neighbor peptide groups to the solvent, could control the solvent screening of peptide dipolar interactions and thus the local conformation in the polypeptide main chain. According to the model then, the differing conformational effects of isoleucine and leucine are the consequence of their differential desolvation of the local peptide groups, due to the differences in the branch point in their side chains.[5,6,11]

Are the local interactions between peptide dipoles really the fundamental arbiters of the conformation and its sequence dependence in protein structures? Addressing the question, we examine polypeptide structure here for the effects of stereochemical mutation from poly-L to alternating-L,D structure. Protein structures are always poly-L, and this is the defining consideration for the dipolar interactions characterizing the α-helix and β-sheet structures.[9,10] The dipolar interactions can be reversed if the polypeptide structure is mutated from poly-L to alternating-L,D. Illustrated in Figure 2, the canonical "folded" and "extended" conformers of poly-L and alternating-L,D structures could differ in their solvent sensitivities and perhaps in sequence sensitivities as well, should the solvent screening of peptide dipolar interactions really comprise the presiding mechanism of folding. That this indeed might be the case is suggested in the following three notable observations: (1) Alternating-L,D helices, exemplified by the gramicidin-A β-helix, are ordered under the conditions of solvent that are clearly denaturing for the helices of a proteinlike poly-L structure,[12−15] (2) not only are poly-L peptides sensitive to solvents, they are also sensitive to structure variations such as variation in the chain length,[7] and (3) modified from poly-L to alternating-L,D structure, polypeptides diminish 10-fold in "stiffness", from ∼9 to ∼0.9, in a length-dependent manner and to a value even lower than the theoretical minimum for the statistical coil (the characteristic ratio 1) "unperturbed" by the steric effects in molecular structure.[16−18] Flory made this observation in 1965 and it puzzled him.[16] Offering a partial explanation, he speculated that the exceptional inflated characteristic ratio of ∼9 in "random-coil" peptides of poly-L structure, compared, for instance, with the characteristic ratio of ∼3 of poly-L lactates, could be due to the unfavorable peptide dipolar interactions in the α-helical conformation.[16] Thus in his invocation of a strong restraint in the conformational access for the poly-L peptide structure, Flory contradicted his own statistical-coil model for proteins in the 1969 book *Statistical Mechanics of Chain Molecules*.[16] Having escaped attention, the anomaly has been profound in its implications for protein-folding theories. Over decades, "statistical coil" occupied the status of the presiding model for the "unfolded" protein, only to be superseded with the view, consistent ironically with what Flory implied in 1969,[16] that the protein and peptide random coils really are quite structured.[7]

In search of the link between sequence and conformation in protein structures, we address here the Flory puzzle regarding the link between "stiffness" and stereochemistry and interpeptide electrostatics in the polypeptide structure. To this end we perform a molecular dynamics (MD) analysis of the effect of stereochemical mutation in end-protected octa-alanine Ac-Ala_8-NHMe from poly-L to an alternating-L,D structure. Elsewhere we have shown that the equilibrium ensemble of the poly-L version of the peptide in water, generated with MD, was in excellent agreement with the results of experimental measurements reported with similar models.[19] Here we use water as the reference solvent and methanol as a typical denaturant. Different in dielectric strength, the solvents could differ in their effects on poly-L and alternating-L,D peptides, should peptide dipolar interaction really be the presiding mechanism for protein folding. The results of the study establish that the stereochemical mutation has a dramatic effect, transforming the peptide from a condition of extreme sensitivity to one of extreme insensitivity to solvent. Searching for explanations, we find that the peptide molecular folds differ in the peptide dipolar interactions at the local and nonlocal levels, either harmonious or antagonistic, depending upon the chain stereochemistry. The strong contrasts between poly-L and alternating-L,D peptides in both chain stiffness and solvent sensitivity are thus founded on a unifying principle, harmony or conflict in peptide dipolar interaction at the local and nonlocal levels. Explaining the Flory puzzle, the opposing forces can also explain the protein sensitivity to not just the chemical environment but perhaps also sequences. Indeed, the evidence has been accumulating that the amino acid

SCHEME 1: Interpeptide Hydrogen Bonds Defining the Turns of Type γ (*i* to *i* − 2), β (*i* to *i* − 3), α (*i* to *i* − 4), and Gramicidin (*i* to *i* + 5 and *i* to *i* − 7)

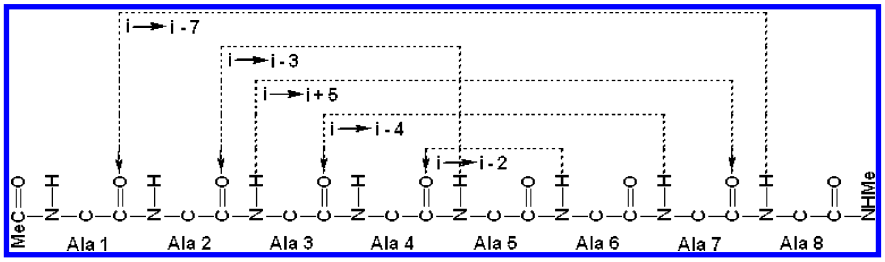


TABLE 1: Percent Occupancies of Specific ϕ,ψ Basins in Equilibrium Ensembles of Alanine Dipeptide and Octa-alanine[a]

peptide, stereochemistry	solvent (time and temperature)[b]	% PPII[c]	% β[d]	% α[e]
Ac-Ala$_1$-NHMe	water (25 ns, 298 K)	40	38	10
Ac-Ala$_1$-NHMe	methanol (20 ns, 298 K)	41	51	0
Ac-Ala$_8$-NHMe, poly-L	water (10 ns, 373 K)	28	24	27
Ac-Ala$_8$-NHMe, poly-L	water (40 ns, 298 K)	35	22	26
Ac-Ala$_8$-NHMe, poly-L	water (105 ns, 298 K)	36	27	16
Ac-Ala$_8$-NHMe, poly-L	methanol (10 ns, 323 K)	48	33	12
Ac-Ala$_8$-NHMe, poly-L	methanol (26 ns, 323 K)	45	40	8
Ac-Ala$_8$-NHMe, alternating-L,D	water (10 ns, 373 K)	38	36	19
Ac-Ala$_8$-NHMe, alternating-L,D	water (25 ns, 298 K)	38	37	15
Ac-Ala$_8$-NHMe, alternating-L,D	methanol (10 ns, 323 K)	43	37	10

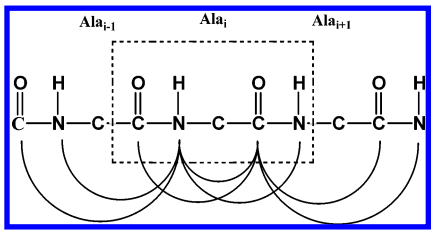
[a] The basins are of a combined L and D enantiomeric structure, as defined in the footnotes. [b] Length, in nanoseconds, and temperature, in Kelvin, of the MD trajectory. [c] PPII: $\phi^L = -30°$ to $-90°$, $\psi^L = 80°$ to $170°$, $\phi^D = 30°$ to $90°$, $\psi^D = -80°$ to $-170°$. [d] β: $\phi^L = -90°$ to $-170°$, $\psi^L = 80°$ to $180°$, $\phi^D = 90°$ to $170°$, $\psi^D = -80°$ to $-180°$. [e] α: $\phi^L = -20°$ to $-100°$, $\psi^L = -20°$ to $-80°$, $\phi^D = 20°$ to $100°$ $\psi^D = 20°$ to $80°$.

side chains differing in α-helix and β-sheet propensities also differ in their local desolvating effects in the protein main chain, and thus presumably in the solvent screening of peptide dipolar interactions. Given this background, we develop the argument that the poly-L structure is the fulcrum of chemical expressions in protein conformation; placing peptide dipolar interactions at the local and nonlocal levels in mutual conflict creates an apparatus of extreme sensitivity to dielectric. Presumably, conflict resolution between folding and unfolding underlying the expression in protein structures of the instructions in their chemical sequences is fundamentally dielectric in mechanism.

Materials and Methods

Gromos96 43A1 was implemented in GROMACS[20,21] on an Intel Pentium IV, stand-alone or clustered, in Linux 7.3. Molecular dynamics was performed in explicit solvent under NVT conditions (constant in number of particles, volume, and temperature) with a periodic boundary. It was our aim to generate thermodynamic ensembles of alanine dipeptide (Ac-L-Ala-NHMe) and octa-alanine (Ac-Ala$_8$-NHMe; poly-L and alternating-L,D; Scheme 1) at ambient temperature (298 K); however, the higher temperature studies at 323 and 373 K were prompted to assess the conditions at equilibrium or to hasten its attainment. Experimental conditions in protein folding are typically NPT (constant in pressure rather than volume), but NVT, 1 atm at 298 K if the density is 1 g/cm^3, is equivalent under the isothermal condition. Pressure changes occurred under NVT with temperature (data not shown) but were significant, as reported,[22] at only 400 K or higher. The equilibrium ensembles prepared under the extremes of temperature in this study (298 and 373 K) were in good agreement with both macroscopic and microscopic observables.

SCHEME 2: Interacting NH and C=O Dipoles within a Residue and Its Nearest Neighbors Defining E_{local}



Conformational ensembles were generated with the octa-alanine isomers in the initial "folded" and "extended" conformations illustrated in Figure 2. In poly-L octa-alanine, the folded conformation was modeled to $\phi = -57°$, $\psi = -47°$, the standard α-helix values, while the extended conformation was modeled to $\phi = -120°$, $\psi = 120°$. In alternating-L,D octa-alanine, the extended conformation was modeled to $\phi = -57°$, $\psi = -47°$ in L residues and to $\phi = 57°$, $\psi = 47°$ in D residues. The folded conformation in this case was modeled to the β-helix of gramicidin-A, assigning the ϕ,ψ values corresponding to positions 3−11 of the structure from PDB code 1GRM.[23]

Water, the SPC model, or methanol, as in GROMACS, was added to a periodic cubic box of edge length 4.5 nm for octa-alanine and 2.9 nm for the alanine dipeptide, containing the energy-minimized solute, to the required density.[24] Molecular dynamics was initialized after rounds of minimization, first with and then without position restraints on the solute atoms. The initial 3 ns was exempted for equilibration, and the trajectories were sampled every 5 ps thereafter. The nonbonded cutoff was 1.4 nm with a shift at 0.8 nm. The integration step was 2 fs. Initial velocities were drawn from Maxwellian distributions. The temperature was coupled to an external bath with a relaxation time constant of 0.1 ps. Bond lengths were constrained with the Shake algorithm to the geometric accuracy of 10^{-4}.

Hydrogen bonds were enumerated with a 0.35 nm cutoff in distance and a 60° cutoff in angle. The ϕ,ψ dihedral angles extracted with GROMACS were plotted to $10° \times 10°$ pixels on $360° \times 360°$ maps with MATLAB. Quantitative comparisons between diastereomers were made on the basis of occupancy statistics of the combined enantiomeric ϕ,ψ basins "α", "β", and "PPII" as defined in the footnotes of Table 1.

Conformational microstates were enumerated in GROMACS with the clustering procedure of Daura et al.[25,26] Briefly, a least-squares translational−rotational fit for backbone atoms (N, C$_\beta$, C$_\alpha$, C) between every pair of conformers gave the number of neighbors for every conformer in an ensemble, with the root-mean-square deviation (RMSD) ≤ 0.15 nm defining the neighbors. The conformer with the largest number of neighbors was defined as the central member of the first cluster or the most populous microstate. All members of this microstate were removed from the ensemble, and the procedure was iterated until all the remaining conformers in the ensemble were assigned to specific microstates, diminishing in population.

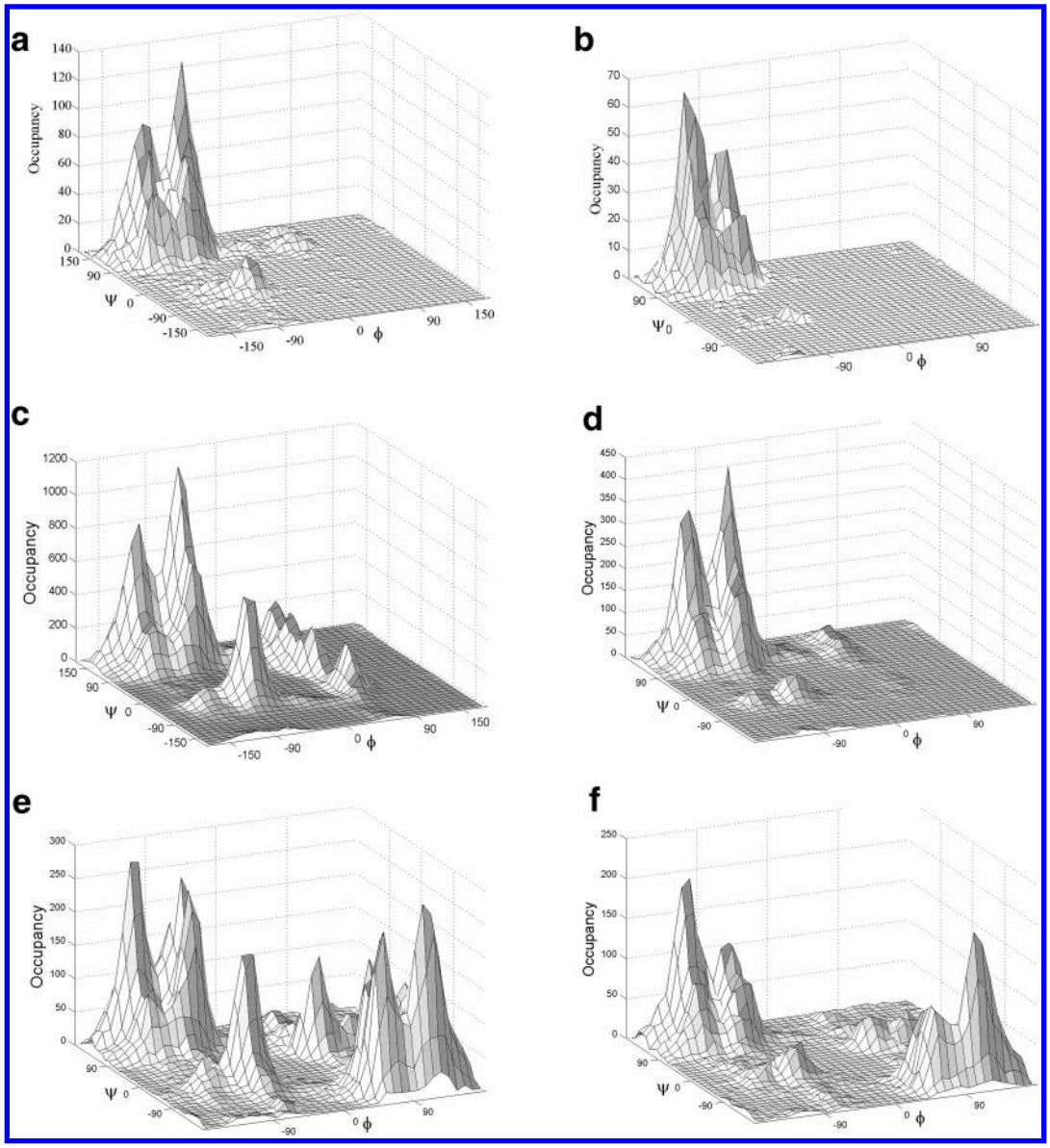


Figure 3. Solvent sensitivity of alanine dipeptide and of poly-L and alternating-L,D octa-alanines, as reflected in the occupancy of ϕ,ψ basins in equilibrium ensembles of the peptides in water and methanol. The ensembles were prepared with MD under the conditions of time and temperature specified in Table 2: Ac-Ala-NHMe in (a) water and (b) methanol; poly-L Ac-Ala$_8$-NHMe in (c) water and (d) methanol; alternating-L,D Ac-Ala$_8$-NHMe in (e) water and (f) methanol.

Electrostatic calculations were performed with the peptides in the extended and folded conformations defined in Figure 2. Atomic partial charge parameters of GROMOS96 were used, and the Coulomb equation was solved with the dielectric constant equal to 1. Two calculations were performed on each structure, E_{total} and E_{local}, corresponding, respectively, to the total Coulomb potential over the entire polypeptide and the Coulomb potential between the local peptide groups, within a residue, and its immediate neighbors spanning tripeptide segments as illustrated in Scheme 2, summed over an entire peptide. The difference between the values calculated for a structure gave its E_{nonlocal} according to the relation $E_{\text{total}} = E_{\text{local}} + E_{\text{nonlocal}}$. The changes in Coulomb potential corresponding to the change in conformation from extended to folded, being ΔE_{total}, ΔE_{local}, and $\Delta E_{\text{nonlocal}}$, were thus calculated for the end-protected poly-alanine Ac-Ala$_N$-NHMe varying in N from 4 to 20.

Results

With MD and using alanine dipeptide Ac-Ala-NHMe as the benchmark, we examine end-protected octa-alanine Ac-Ala$_8$-NHMe for the effect of stereochemical mutation on its solvent sensitivity. Two peptide groups are joined in a single stereocenter in alanine dipeptide, which is of L configuration in this study. Nine peptides are joined in eight stereocenters in octa-alanine, which we model in poly-L and alternating-L,D configurations. We perform the MD in three steps: defining the conditions for equilibrium, preparing the equilibrium ensembles of interest, and analyzing the role of stereochemistry in the solvent sensitivity of octa-alanine.

Evaluating the Conditions for Equilibrium. MD was initiated with octa-alanines in the folded or extended conformations shown in Figure 2, and the trajectories were examined for the time evolution in (i) hydrogen bonds, (ii) distribution in ϕ,ψ angles, (iii) population in ϕ,ψ basins, and (iv) the number of microstates. A convergence between these observables defined the requirement in time or temperature to the attainment of equilibrium. The results of the fixed time (5 ns) and variable temperature experiments are in Figures S1−S3 of the Supporting Information. With water as the solvent, the distributions in hydrogen bonds (Figure S1) and ϕ,ψ angles (Figure S2) and

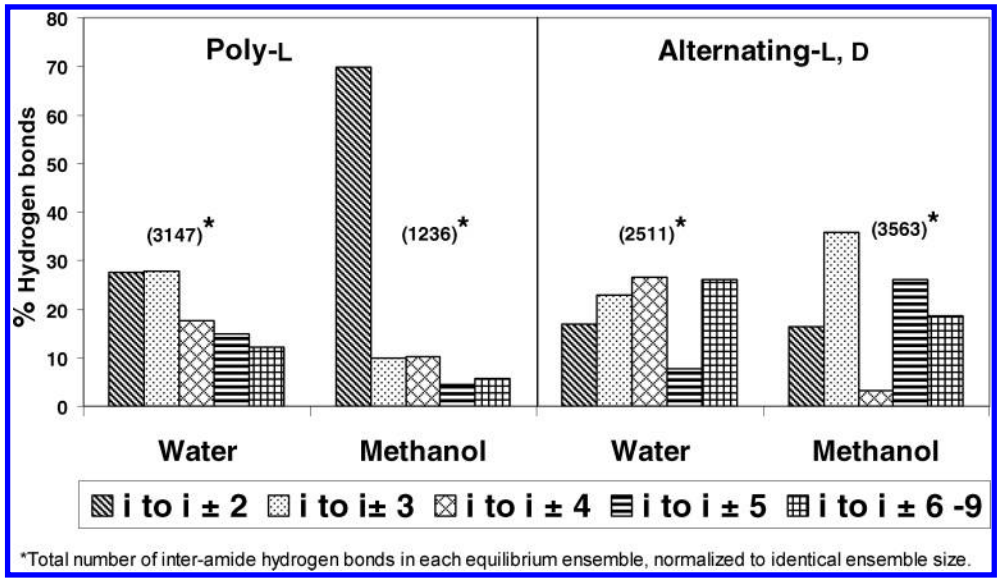


Figure 4. Solvent sensitivities of poly-L and alternating-L,D octa-alanines, as reflected in the number of interpeptide hydrogen bonds and the frequency of specific types of hydrogen bonds in equilibrium ensembles of the peptides in water and methanol. The ensembles were prepared with MD under the conditions of time and temperature specified in Table 2.

the time evolution in microstates (Figure S3) are dependent on the initial conformation in both the octa-alanines at 298 and 323 K but not at 373 K. Apparently, an equilibrium state is approximated in the systems over 5 ns at 373 K but not at 298 or 323 K. Likewise, with methanol as the solvent, the distributions in hydrogen bonds (Figure S1) and ϕ,ψ angles (Figure S2) and the time evolutions in microstates (Figure S3) are dependent on the initial conformations in octa-alanines at 298 K but not at 323 K. Apparently, a state of equilibrium is approximated in these systems over 5 ns at 323 K but not at 298 K. The values of the specific observables over 5 ns at 373 K in water and at 323 K in methanol are thus the benchmarks to judge the state of equilibrium in the systems of interest.

Preparing Equilibrium Ensembles. The 298 K MD trajectories in water initiated with the octa-alanines in extended conformations were continued. A state of equilibrium could be approximated at 298 K over 40 ns in poly-L octa-alanine and over 25 ns in alternating-L, D octa-alanine in water, as judged with (i) ϕ,ψ basin populations (Table 1) the distributions in (ii) ϕ,ψ angles (Figures 3 and S2) and (iii) hydrogen bonds (Figures 4 and S1), and (iv) time evolution of microstates (Figure 5). The 298 K MD trajectory of poly-L octa-alanine in water was continued to 105 ns. The data for basin occupancy statistics over this longer trajectory are shown separately in Table 1. Clearly from these data, even the shorter 40 ns trajectory is an acceptable approximation of the equilibrium state in this system. Thus only the shorter trajectory (40 ns) was examined in detail for its microstates and their properties.

With methanol as the solvent, the 323 K trajectories initiated with octa-alanines in folded and extended conformations were already approximately in equilibrium. The trajectory of poly-L octa-alanine was extended to 26 ns, and again from the basin occupancy statistics in Table 1, even the shorter trajectory is a fair approximation for an ensemble in equilibrium. The shorter 5 ns trajectories were merged, nanosecond by nanosecond, and an enumeration of the microstates was performed both before and after the merger. The data on time evolution of microstates are in Figure 5, and clearly the population saturation in microstates was better approximated in the combined 10 ns trajectories. Thus these trajectories were used for the analysis of the microstates and their properties in detail. Since in water the equilibrium ensembles were prepared at 298 K, the 298 K

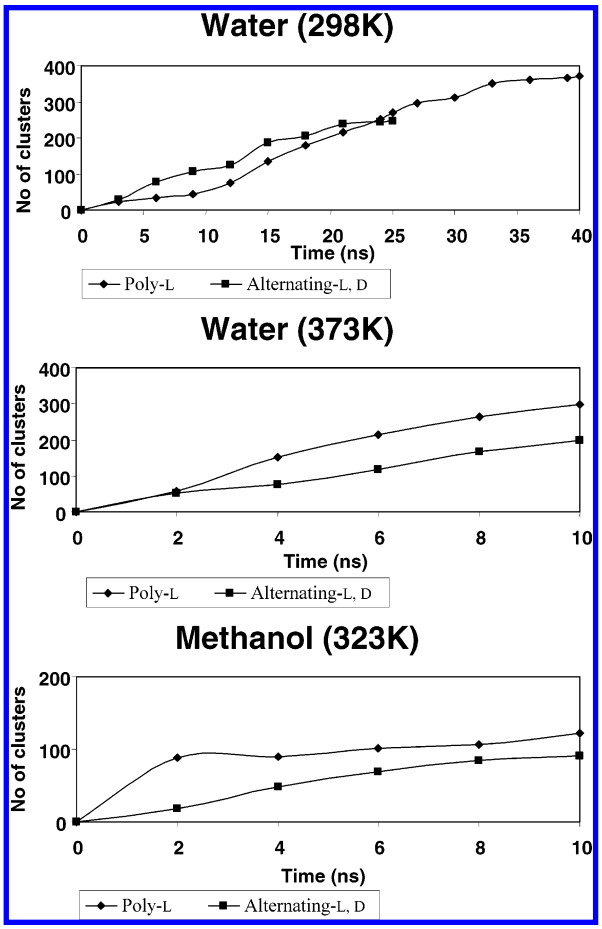


Figure 5. Time evolution of microstates (conformational clusters at 0.15 nm RMSD cutoff) in the MD trajectories of octa-alanine under the conditions of solvent, temperature, stereochemistry, and the initial peptide conformation specified.

trajectories of isomeric octa-alanines in methanol were also merged, nanosecond by nanosecond. There was no evidence that the microstates are saturated even in the combined trajectories in this case (data not shown); however, the qualitative mix of the folds (Figures S4c and S4d) is broadly the same

TABLE 2: Number of Conformers in the Eight Largest Clusters, or Microstates, Populating Equilibrium Ensembles of Poly-L and Alternating-L,D Octa-alanine Prepared with MD[a]

	poly-L		alternating-L,D	
	water (40 ns, 298 K)	methanol (10 ns, 323 K)	water (25 ns, 298 K)	methanol (10 ns, 323 K)
1	1422 (4)	260 (0)	401 (3)	301 (1)
2	491 (3)	148 (0)	337 (3)	185 (3)
3	247 (3)	109 (0)	291 (2)	169 (2)
4	151 (0)	90 (1)	228 (1)	74 (1)
5	149 (4)	79 (0)	162 (3)	73 (1)
6	130 (3)	64 (4)	160 (3)	57 (3)
7	120 (0)	60 (0)	147 (3)	53 (1)
8	106 (2)	58 (0)	142 (0)	52 (3)

[a] The figure in parentheses is the number of interpeptide hydrogen bonds in the central member of each cluster.

as in 323 K 10 ns ensembles, as noted by comparing the Figures S4a and S4b and 6.

Conformational ensembles were also prepared for alanine dipeptide (Ac-Ala-NHMe) in water and methanol at 298 K over, respectively, 25 and 20 ns. Given the simplicity of the system compared with octa-alanine, a state of equilibrium is presumably well approximated in both the ensembles. The distribution statistics of ϕ,ψ angles in the ensembles are in Figure 3 and Table 1.

Evaluating the State of Folding from Basin Occupancies. The inversion symmetric α_L and α_R Ramachandran basins (Figure 1) are fundamental diagnostics of the "folds" of protein structure; the occupancy of at least one of the basins is mandatory for the β- and α-helical turns; only the γ-turn, the C_2 ribbon, the so-called collagen PPII and PPI helices, and the β-stranded structures will fall outside the α_L or α_R basins. The ϕ,ψ distributions and the occupancy statistics of the ϕ,ψ basins in Figure 3 and Table 1 imply that while alanine dipeptide may be partially folded in water it is completely unfolded in methanol; the α-conformational basin is practically empty in this solvent. On similar reasoning, it can be inferred from the results in Figure 3 and Table 1 that poly-L octa-alanine is folded to a much smaller extent in methanol than water. Even alternating-L,D octa-alanine is noted to display a solvent dependence in both ϕ,ψ distribution and occupancy statistics of ϕ,ψ basins (Figure 3 and Table 1)—the α basin is 1.5 times more populous in water than in methanol—but no judgment about the extent of folding can be made in this case since even PPII/β-type ϕ,ψ angles can comprise folds in this system. Indeed, gramicidin-A adopts a variety of folds[12,14,15] that are β/PPII in ϕ,ψ angles; in fact repetitive α-type ϕ,ψ angles in this peptide are extended[27] as noted in Figure 2. Clearly, no judgment about the extent of folding can be made in the alternating-L,D octa-alanine on the basis of the ϕ,ψ distribution or the occupancy statistics of different ϕ,ψ basins.

Evaluating the State of Folding from Hydrogen Bonds. Unlike ϕ,ψ angles, hydrogen bonds are a direct arbiter of folds independent of chain stereochemistry. From the results in the left panel of Figure 4 it is noted that overall there are nearly 2.5 times more hydrogen bonds in poly-L octa-alanine in water than in methanol. Further, an appreciable proportion of the hydrogen bonds in water are the β-turn (i to $i \pm 3$) and α-helix (i to $i \pm 4$) types (Scheme 1), while a large proportion of those in methanol are the γ-turn (i to $i \pm 2$) type, for which the ϕ,ψ angles occur in the Ramachandran β-region in close proximity of PPII sub-basin.[28] Clearly, while poly-L octa-alanine is extensively folded in water it is more or less completely unfolded in methanol. The results in the right panel of Figure 4 imply that the peptide sensitivity to solvent is stereospecific. Alternating-L,D octa-alanine can adopt gramicidin-A-type folds.[12,14,15] Modeled in Figure 2, the gramicidin β-helix is with its peptide dipoles in a mutual antiparallel arrangement and in alternately i to $i + 5$ and i to $i - 7$ hydrogen bonds (Scheme 1). From the results in the right panel of Figure 4, alternating-L,D octa-alanine is noted to feature a significantly lesser number of hydrogen bonds in water than in methanol. The mix of the hydrogen bonds is generally the same except for a slight preponderance of α-type hydrogen bonds (i to $i \pm 4$) in water and of β-turn (i to $i \pm 3$) and gramicidin-type hydrogen bonds (i to $i \pm 5$) in methanol. Clearly, whether slightly more folded in methanol or with folds that are somewhat higher in frequency of hydrogen bonds, alternating-L,D octa-alanine is by and large comparable in its degree of folding in the two solvents.

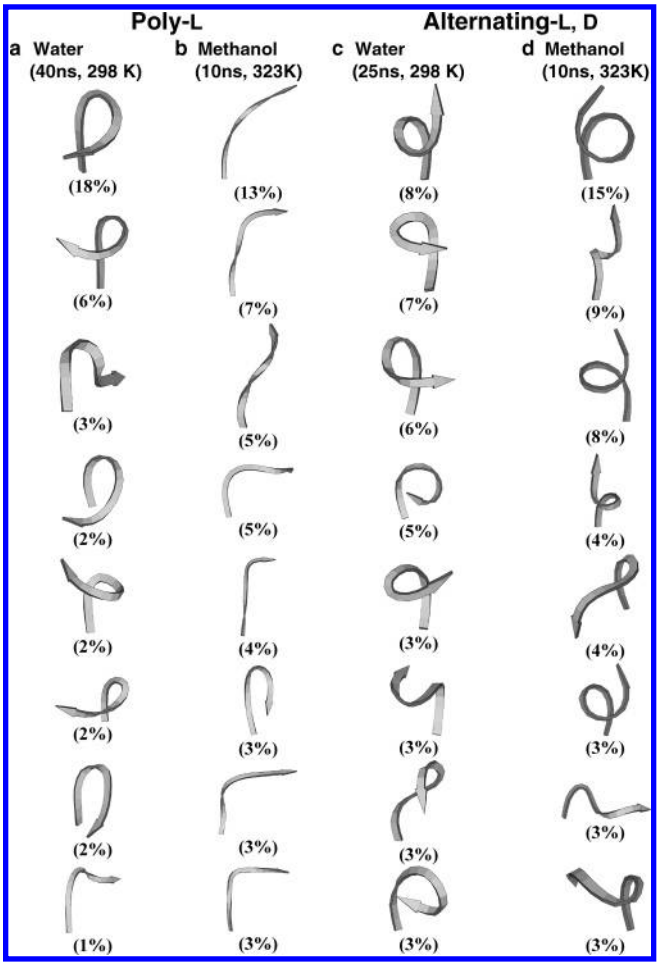


Figure 6. Solvent-specific nature of the microstates of equilibrium ensembles of poly-L and alternating-L,D octa-alanine. The ensembles were prepared with MD under the conditions of time and temperature specified in Table 2. The eight most populous microstates of each ensemble are shown in ribbon representations of the central members, with the percent population of the microstate shown in parentheses.

Evaluating the State of Folding from Microstates. The microstates of an ensemble are a direct portrayal of its molecular "folds" in atomistic detail. The cluster statistics of the eight largest microstates in each equilibrium ensemble are in Table 2. The residue-by-residue ϕ,ψ angles in the central member of each microstate are in Table S1 and Figure S6, and their hydrogen bond registries are shown in Figure S5. The most populous microstates in ribbon representations of their central members are in Figure 6, and the percent population of each microstate is shown in parentheses. The stereoviews of wire frames of some of the largest clusters are in Figures S4a and

S4b. The most populous microstates in 298 K nonequilibrium ensembles in methanol are in ribbon representations of their central members in Figure S4c and in stereoviews of their wire frames in Figure S4d.

The overall conclusion is that whether in real population equilibrium or not, the microstates of isomeric octa-alanines are in a state of differing folding. In the case of poly-L octa-alanine, almost all microstates are folded in water but unfolded in methanol. In the case of alternating-L,D octa-alanine, however, nearly all of the microstates are folded in both water and methanol. Clearly, the isomeric octa-alanines are in strong contrast to their solvent sensitivities; the alternating-L,D isomer is practically inert while the poly-L isomer is extremely sensitive.

Interpeptide Electrostatics in Extended and Folded Conformations. To examine if interpeptide electrostatics might be involved in the differing sensitivities of the isomeric octa-alanines to solvent, a Coulomb calculation was performed of the change in electrostatic potential on the folding of a peptide from the canonical extended to helical conformations, as defined in the Material and Methods section and illustrated in Figure 2. The net change in Coulomb potential in the process, ΔE_{total}, was evaluated for its local and nonlocal components, respectively, ΔE_{local} and $\Delta E_{nonlocal}$, according to the relation $\Delta E_{total} = \Delta E_{local} + \Delta E_{nonlocal}$. E_{local} in a peptide is defined, as illustrated in Scheme 2, as the interaction between its dipoles within a residue and its immediate neighbors in a tripeptide segment, summed over the entire polypeptide, while $E_{nonlocal}$ defines every other possible peptide dipolar interaction in a peptide, including the interactions that are definable as its hydrogen bonds. The calculations were performed with the peptides increasing in chain length and with a dielectric constant of 1. The results are summarized in Figure 7 and Table S2. The peptides with isomeric stereochemical structures are noted to differ in the manner in which ΔE_{total} partitions into ΔE_{local} and $\Delta E_{nonlocal}$ and in the length dependence of ΔE_{local} and $\Delta E_{nonlocal}$ and therefore of ΔE_{total}. In poly-L peptides, ΔE_{local} and $\Delta E_{nonlocal}$ are more or less comparable in magnitude but opposite in sign and therefore subtractive in the contributions in ΔE_{total}. In contrast, the terms are identical in sign and therefore additive in the contributions in ΔE_{total} in alternating-L,D peptides. As a consequence, there is a large disparity in both the magnitude of ΔE_{total} and its length dependence in the two stereochemical series. Specifically, ΔE_{local} and $\Delta E_{nonlocal}$ are in conflict in poly-L peptides and vanish on mutating to an alternating-L,D structure; thus the conflict has its roots in stereochemistry.

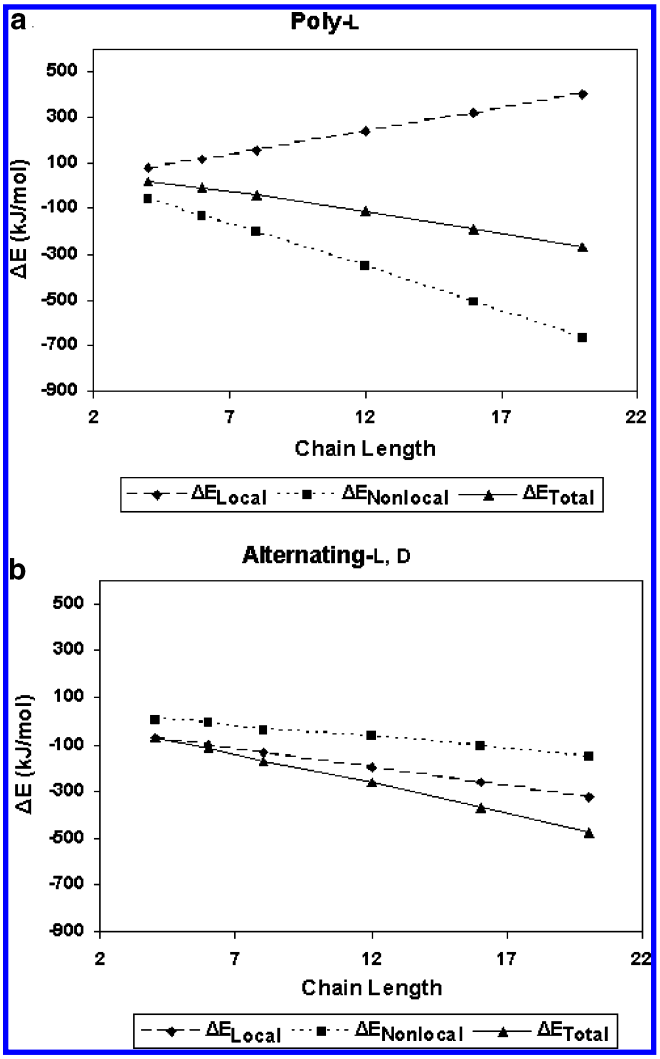


Figure 7. Plot of chain length (the number of residues) and the difference Coulomb potential (ΔE) in poly-L and alternating-L,D alanines between the extended and helical conformers defined in the Materials and Methods section and illustrated in Figure 2.

Discussion

Ramachandran[29] described α_R and β basins in the ϕ,ψ torsional space of the alanine dipeptide molecular structure as the fundamental conformational alternatives for protein structure (Figure 1). Clearly valid for the native protein structures, the generalization was presumed to be also valid for proteins in denatured states. Diverse results in a variety of systems have falsified the notion,[30,31] affirming the pioneering observation of Tiffany and Krimm,[32] made a year before the publication of Flory's book,[16] that there could be "order" in "disordered" peptides along with specificity in conformation, in contradiction to Flory's random-coil hypothesis, according to which the α and β Ramachandran basins are statistically sampled.[16] The determinants of the order and specificity in the so-called random-coil states in polypeptide structures are now established to reside within the elements of alanine dipeptide;[33,34] consequently poly-alanines have been relevant models in probing the principles.[30,31] Poly-alanines also harbor the elements of sensitivity of polypeptide structure to the effects of both solvents and amino acid side chains. Stereochemistry in the main chain and the interactions between peptide dipoles could be involved, having been implicated by Flory as the determinants of unusual stiffness of polypeptide structures,[16] though in a manner that puzzled him.

Noting that poly-L peptides are much stiffer than expected on the basis of atomic excluded volume effects, Flory invoked repulsive interaction between peptide dipoles in α-conformations as the reason.[16] In accord with his speculation and in agreement with more recent insights,[30,31,33,34] we note here that there is a clear tendency in alanine dipeptide to avoid the α-conformation, presumably for electrostatic reasons. The molecule equilibrates largely within the broader β-Ramachandran basin and in a bimodal distribution (Figures 3a and 3b and Table 1, rows 1 and 2), displaying a clear preference for the "PPII" sub-basin in water but for the "extended-β" sub-basin in methanol. Recently much attention has focused on the reasons why alanine dipeptide will prefer PPII rather than the extended-β conformation in water. Sterics or solvation was invoked as the reason in different studies.[35−37] From the results in this study we find that solvent has a role, but whether for reasons of dielectrics or hydrogen bonding is not our immediate concern here; rather, we are interested in the reason or reasons for the apparent aversion of alanine dipeptide for the α-type ϕ,ψ angles. Stronger

TABLE 3: Change in Coulomb Potential (ΔE in kJ/mol) on Changing a Peptide from the "Extended" to "Folded" Conformation as Defined in the Materials and Methods Section and Illustrated in Figure 2

	poly-L			alternating-L,D		
chain length	ΔE_{local}	$\Delta E_{nonlocal}$	ΔE_{total}[a]	ΔE_{local}	$\Delta E_{nonlocal}$	ΔE_{total}[a]
4	79	−60	19	−72	2	−70
6	119	−128	−9	−104	−8	−112
8	158	−199	−41	−134	−38	−172
12	240	−353	−113	−197	−64	−261
16	322	−510	−188	−262	−109	−371
20	402	−667	−265	−324	−154	−479

[a] $E_{total} = \Delta E_{local} + \Delta E_{nonlocal}$.

in methanol (Figure 3 and Table 1), the solvent of lower dielectric strength, the aversion could be electrostatic, in conformity with Flory's speculation.[16] ΔE_{local} is ∼13.5 kJ/mol (∼3.2 kcal/mol) as alanine dipeptide unfolds from "α" to "β/PPII" conformation, involving an approximately ±180° rotation in ψ, in concordance with the value of ∼3.5 kcal/mol that Avbelj and Moult calculated for the process.[5]

While ΔE_{local} is unfavorable, $\Delta E_{nonlocal}$ contributes to folding in poly-L alanines and in a length-dependent manner, growing more strongly with length than ΔE_{local}, outstripping it when the length is six residues, as noted in Figure 7 and Table 3. Participating in $\Delta E_{nonlocal}$, hydrogen bonds can contribute in folding, and the possible choice of donor and acceptor peptides is stereochemically dertermined.[9,10] When poly-L, either the i to $i-3$ type hydrogen bonds defining a β-turn[38] or the i to $i-4$ type hydrogen bonds defining α-helical turn[10] (Scheme 1) are needed for folding. Both are α-type in the ϕ,ψ angles, in at least one of the residues in a β-turn.[38] Unfavorable in ΔE_{local}, the folds are dependent on the assistance of solvent; as the stronger dielectric medium, water promotes folding in poly-L octa-alanine, but methanol does not. In alternating-L,D octa-alanine, the gramicidin-type i to $i+5$ and i to $i-7$ hydrogen bonds (Scheme 1 and Figure 2) are possible, and since α-type ϕ,ψ angles are avoided, the folding can proceed with harmony between ΔE_{local} and $\Delta E_{nonlocal}$; thus ΔE_{total} is negative in this series even in the shortest possible peptide, while it diminishes more strongly with length than in the poly-L series (Figure 7 and Table 3). Thus alternating-L,D octa-alanine is fully folded in both water and methanol; solvent has no explicit role in folding.

Does the harmony or conflict between the local and the nonlocal peptide dipolar interactions, that is, between ΔE_{local} and $\Delta E_{nonlocal}$, explain the catastrophic loss of stiffness in poly-L peptides when modified to alternating-L,D structures, as noted by Flory? Stiffness in a polymer is its persistence length or its characteristic ratio,[16] being the ratio between its experimentally determined end-to-end distance or radius of gyration and the corresponding value theoretically calculated in the statistical-coil model unperturbed by the exclusion volume of atoms. The value is ∼9 for poly-L peptides but only ∼3 for poly-L lactates.[16] Given this 3-fold difference in systems of comparable atomic exclusion volumes, Flory invoked repulsive interaction between peptide dipoles in the α-conformation as a possible reason.[16−18] Against this backdrop, Flory noted that when modified to alternating-L,D structure, the characteristic ratio in poly-L peptides will diminish from the unusual ∼9.0 to the surprising ∼0.9, in a length-dependent manner, and to a value lower than the theoretical minimum for the unperturbed statistical-coil.[16] While Flory speculated that the gramicidin-type folds might be involved, he was puzzled by the reversal of the characteristic ratio with length (ref 16, p 286), diminishing in alternating-L,D peptides but increasing sharply in poly-L peptides.[16] The puzzle is now resolved with the demonstration that the mutual relationship between ΔE_{local} and $\Delta E_{nonlocal}$ is stereochemically determined. When the terms are in harmony, ΔE_{total} is strong in length dependence and can cause the characteristic ratio to diminish with length when the peptide structure is alternating-L,D. Also, given favorable ΔE_{local}, the gramicidin-A β-helix should be insensitive to solvent, which it is,[12,14,15] being ordered in apolar solvents and in lipid-type environments in detergents, the conditions that are typically denaturing for not just the poly-L helix but also for a typical water-soluble globular proteins in general. With ΔE_{local} and $\Delta E_{nonlocal}$ in conflict, ΔE_{total} is comparatively small in both magnitude and length dependence in poly-L peptides. The unfavorable ΔE_{local} should render poly-L peptides sensitive to the dielectric arbitration of solvent, which indeed is the case.[7] Further, considering the disparities in length dependence of ΔE_{local} and $\Delta E_{nonlocal}$ (Figure 7 and Table 3), the poly-L peptides should be sensitive to variations in chain length, which indeed is the case; extremely short poly-alanines tend to be in an extended PPII conformation[8,30] as the effect of ΔE_{local} prevails, and poly-alanines with approximately 20 or more residues are α-helical[7] as the effect of $\Delta E_{nonlocal}$ prevails, while the intermediate length poly-alanines could be in folded PPII conformations[19] as ΔE_{local} and $\Delta E_{nonlocal}$ become more evenly poised. As a corollary, the all-or-none type of cooperativity of the α-helical fold might be a consequence of the opposing effects of ΔE_{local} and $\Delta E_{nonlocal}$, just as the apparent solvent sterility of gramicidin β-helix might be a consequence of the harmony between ΔE_{local} and $\Delta E_{nonlocal}$.

The protein molecular structure, it would appear, is a dielectric sensor par-excellence by the very nature of its chemical constitution. The question remains; does this property also comprise the heart of the expression mechanism in proteins of the instructions they harbor in their sequences? Indeed, there is an evolving thread in the literature in support of the idea. Highlighting the planarity and polarity of a peptide group,[39] Pauling implicated hydrogen bonding interactions between peptide dipoles as the defining principle for secondary structure.[9,10] Differing in dipole orientations, the α-helix and β-sheet secondary structures differ in E_{local}; the effect is invoked in both kinetics and thermodynamics of protein folding; helix folding according to Zimm- and Bragg-type theories is large and unfavorable in the initiation constant;[40,41] adverse ΔE_{local} is forwarded as a reason; being a macro-dipole, the α-helix is sensitive to the effects of charged side chains,[7] the interactions with E_{local} being a reason. The peptide group is large in hydration free energy and almost entirely due to the enthalpy of association with water.[42] Variable in E_{local}, the different conformers of alanine dipeptide structure might be variable in their enthalpy of association with water, and replacements of side chains may influence the process by blocking the access to water. Indeed, arbitrating solvent accessibility of adjoining and nearest-neighbor peptides, side chains can modulate the effects of E_{local}.[5] Side chains are shown to define the electrostatic solvation free energy of the local peptide dipoles—being a measure of their E_{local} plus accessibility to solvent—and the effect is shown to correlate with the side-chain preferences for β-conformation,[6] for helical conformation,[5] and for a capping position in a helical motif.[43] The α-helix and β-sheet motifs differ in NMR chemical shifts;[44] the effect is shown to correlate with both electrostatic solvation free energy in the main chain and accessibility of the main chain to solvent.[45] Nearest-neighbor effects appear in $^3J_{HNC_\alpha H}$ coupling constants and therefore in ϕ values in short, unordered pep-

tides;[46] the effect is shown to correlate with the class of the amino acid side chain, S-class (small) or L-class (large), and the effects of solvent accessibility of the main chain.[47] The net peptide dipole vectors in globular protein structures are correlated with the electrostatic effects of the charged side chains;[48] the interactions of E_{local} with electrostatic fields of the folded protein structures could be the reason. Clearly, controlling solvent accessibilities of their adjoining and nearest-neighbor peptide dipoles, amino acid side chains can control the solvent screening of E_{local} and therefore the local conformation in the main chain. Further, the electrostatic cum dielectric effects of the protein covalent structures can modulate the effects of E_{local} and thus the local conformation in the main chain.

The protein structures are far too complex to expect that any singular effect will in entirety explain the observed relationship of sequence with conformation in protein structures. Surely, sterics, hydrophobics, conformational entropy, interactions among side chains, and interactions of side chains with the main chain can all contribute; however, being a generic of the main-chain structure, the interactions between peptide dipoles could be in the presiding role. Knowing that side chains can arbitrate in the solvent screening of the interaction between peptide dipoles, it is our central conclusion in this study that the opposing effects of ΔE_{local} and $\Delta E_{nonlocal}$, riding on the fulcrum of poly-L stereochemistry, are pivotal in the response mechanism of proteins to the effects in sequences.

Conclusion

The principle of correspondence between sequence and conformation in native protein structures has been one of the most enduring scientific problems of recent decades. Here, we have illuminated the problem from a fresh angle, demonstrating that the correspondence may be stereochemically established. It has been increasingly clear over the recent years that the peptide dipoles of the polypeptide structure are in a conflicting type of interaction at the local and nonlocal levels. We have demonstrated in the present study that the conflict is stereochemical, operative in poly-L peptides but not in alternating-L,D peptides. We have further established that the conflict is specifically responsible for the extreme sensitivity of polypeptide structure to solvent, via the dielectric mechanism. We have further argued that the effect may also contribute to the expression of side-chain effects in the main chain conformation, with dielectrics as the presiding mechanism. Thus in conclusion we assert that poly-L stereochemistry may be the fulcrum of expression in proteins of the instructions that they encode chemically in their sequential structures.

Acknowledgment. This work was supported by grants from the Council of Scientific and Industrial Research and the Board of Research in Nuclear Sciences, Department of Atomic Energy, Government of India.

Supporting Information Available: Residue-by-residue ϕ,ψ angles in central members of the eight most populous microstates of the equilibrium ensembles of poly-L and alternating-L,D octa-alanine in water and methanol, chain-length-dependent local Coulomb energies (E_{local}) of poly-L and alternating-L,D alanines in extended and folded conformations and the change in local Coulomb energy on folding (ΔE_{local}), the statistics of interpeptide hydrogen bonds in the conformational folds populating 3−8 ns segments of MD trajectories of extended or folded octa-alanines of poly-L and alternating-L,D structures, occupancy of ϕ,ψ basins in 3−8 ns segments of MD trajectories of extended or folded octa-alanines of poly-L structure, time evolution of microstates in the MD trajectories of octa-alanine, stereo-images of wire frames of the most populous conformational clusters of the equilibrium ensembles of poly-L and alternating-L,D octa-alanine in water and methanol, ribbon representations of the central members of the eight most populous microstates of the nonequilibrium ensembles of poly-L and alternating-L,D octa-alanine in water and methanol, stereo-images of wire frames of the most populous conformational clusters of the nonequilibrium ensembles of poly-L and alternating-L,D octa-alanines in methanol, and hydrogen bond registries in the central members of the eight most populous microstates of the equilibrium ensembles of poly-L and alternating-L,D octa-alanine in water and methanol. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Schonbrun, J.; Wedemeyer, W. J.; Baker, D. *Curr. Opin. Struct. Biol.* **2002**, *12*, 348−354.
(2) Creamer, T. P. *Proteins: Struct., Funct., Genet.* **2000**, *40*, 443−450.
(3) Doig, A. J.; Sternberg, M. J. *Protein Sci.* **1995**, *4*, 2247−2251.
(4) Street, A. G.; Mayo, S. L. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 9074−9076.
(5) Avbelj, F.; Moult, J. *Biochemistry* **1995**, *34*, 755−764.
(6) Avbelj, F.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 1309−1313.
(7) Chakrabartty, A.; Baldwin, R. L. *Adv. Protein Chem.* **1995**, *46*, 141−176.
(8) Shi, Z.; Olson, C. A.; Rose, G. D.; Baldwin, R. L.; Kallenbach, N. R. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 9190−9195.
(9) Pauling, L.; Corey, R. B. *Nature* **1951**, *168*, 550−551.
(10) Pauling, L.; Corey, R. B.; Branson, H. R. *Proc. Natl. Acad. Sci. U.S.A.* **1951**, *37*, 205−211.
(11) Avbelj, F.; Fele, L. *J. Mol. Biol.* **1998**, *279*, 665−684.
(12) Wallace, B. A. *J. Struct. Biol.* **1998**, *121*, 123−141.
(13) Wallace, B. A.; Ravikumar, K. *Science* **1988**, *241*, 182−187.
(14) Veatch, W. R.; Fossel, E. T.; Blout, E. R. *Biochemistry* **1974**, *13*, 5249−5256.
(15) Urry, D. W. *Proc. Natl. Acad. Sci. U.S.A.* **1971**, *68*, 672−676.
(16) Flory, P. J. *Statistical Mechanics of Chain Molecules*; InterScience Publishers: New York, 1969.
(17) Brant, D. A.; Flory, P. J. *J. Am. Chem. Soc.* **1965**, *87*, 2788−2791.
(18) Brant, D. A.; Flory, P. J. *J. Am. Chem. Soc.* **1965**, *87*, 2791−2800.
(19) Ramakrishnan, V.; Ranbhor, R.; Durani, S. *J. Am. Chem. Soc.* **2004**, *126*, 16332−16333.
(20) Lindahl, E.; Hess, B.; van der Spoel, D. *J. Mol. Model.* **2001**, *7*, 306−317.
(21) van Gunsteren, W. F.; Billeter, S. R.; Eising, A. A.; Hünenberger, P. H.; Krüger, P.; Mark, A. E.; Scott, W. R. P.; Tironi, I. G. *Biomolecular Simulation: The GROMOS96 Manual and User Guide*; Hochschulverlag AG an der ETH Zürich: Zürich, Switzerland, 1996.
(22) Walser, R.; Mark, A. E.; van Gunsteren, W. F. *Biophys. J.* **2000**, *78*, 2752−2760.
(23) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235−242.
(24) Kell, G. S. *J. Chem. Eng. Data* **1975**, *20*, 97−105.
(25) Daura, X.; Gademann, K.; Jaun, B.; Seebach, D.; van Gunsteren, W. F.; Mark, A. E. *Angew. Chem., Int. Ed.* **1999**, *38*, 236−240.
(26) Daura, X.; van Gunsteren, W. F.; Mark, A. E. *Proteins: Struct., Funct., Genet.* **1999**, *34*, 269−280.
(27) Watson, J. D.; Milner-White, E. J. *J. Mol. Biol.* **2002**, *315*, 183−191.
(28) Milner-White, E. J. *J. Mol. Biol.* **1990**, *216*, 386−397.
(29) Ramachandran, G. N.; Sasisekharan, V. *Adv. Protein Chem.* **1968**, *23*, 283−438.
(30) Shi, Z.; Woody, R. W.; Kallenbach, N. R. *Adv. Protein Chem.* **2002**, *62*, 163−240.
(31) Baldwin, R. L. *Adv. Protein Chem.* **2002**, *62*, 361−367.
(32) Tiffany, M. L.; Krimm, S. *Biopolymers* **1968**, *6*, 1379−1382.
(33) Poon, C.; Samulski, E. T.; Weise, C. F.; Weisshaar, J. C. *J. Am. Chem. Soc.* **2000**, *122*, 5642−5643.
(34) Han, W.-G.; Jalkanen, K. J.; Elstner, M.; Suhai, S. *J. Phys. Chem. B* **1998**, *102*, 2587−2602.
(35) Pappu, R. V.; Rose, G. D. *Protein Sci.* **2002**, *11*, 2437−2455.
(36) Mezei, M.; Fleming, P. J.; Srinivasan, R.; Rose, G. D. *Proteins* **2004**, *55*, 502−507.

(37) Drozdov, A. N.; Grossfield, A.; Pappu, R. V. *J. Am. Chem. Soc.* **2004**, *126*, 2574−2581.
(38) Rose, G. D.; Gierasch, L. M.; Smith, J. A. *Adv. Protein Chem.* **1985**, *37*, 1−109.
(39) Pauling, L. *Nature of the Chemical Bond*, 3rd ed.; Cornell University Press: New York, 1960.
(40) Zimm, B. H.; Bragg, J. K. *J. Chem. Phys.* **1959**, *31*, 526−535.
(41) Scheraga, H. A.; Vila, J. A.; Ripoll, D. R. *Biophys. Chem.* **2002**, *101*−*102*, 255−265.
(42) Wolfenden, R. *Biochemistry* **1978**, *17*, 201−204.
(43) Thomas, S. T.; Loladze, V. V.; Makhatadze, G. I. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 10670−10675.
(44) Wishart, D. S.; Sykes, B. D.; Richards, F. M. *Biochemistry* **1992**, *31*, 1647−1651.
(45) Avbelj, F.; Kocjan, D.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 17394−17397.
(46) Penkett, C. J.; Redfield, C.; Dodd, I.; Hubbard, J.; McBay, D. L.; Mossakowska, D. E.; Smith, R. A.; Dobson, C. M.; Smith, L. J. *J. Mol. Biol.* **1997**, *274*, 152−159.
(47) Avbelj, F.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 10967−10972.
(48) Ripoll, D. R.; Vila, J. A.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 7559−7564.