Conformations and free energy landscapes of polyproline peptides

Mahmoud Moradi[a], Volodymyr Babin[a], Christopher Roland[a], Thomas A. Darden[b], and Celeste Sagui[a,1]

[a]Center for High Performance Simulations and Department of Physics, North Carolina State University, Raleigh, NC 27695; and [b]OpenEye Scientific Software, 9 Bisbee Court, Suite D, Santa Fe, NM 87508

Edited by Harold A. Scheraga, Cornell University, Ithaca, NY, and approved September 10, 2009 (received for review June 11, 2009)

The structure of the proline amino acid allows folded polyproline peptides to exist as both left- (PPII) and right-handed (PPI) helices. We have characterized the free energy landscapes of hexamer, nanomer, and tridecamer polyproline peptides in gas phase and implicit water as well as explicit hexane and 1-propanol for the nanomer. To enhance the sampling provided by regular molecular dynamics, we used the recently developed adaptively biased molecular dynamics method, which describes Landau free energy maps in terms of relevant collective variables. These maps, as a function of the collective variables of handedness, radius of gyration, and three others based on the peptide torsion angle ω, were used to determine the relative stability of the different structures, along with an estimate of the transition pathways connecting the different minima. Results show the existence of several metastable isomers and therefore provide a complementary view to experimental conclusions based on photo-induced electron transfer experiments with regard to the existence of stable heterogeneous subpopulations in PPII polyproline.

cis-trans isomerization | left-handed helix | molecular dynamics | PPI | PPII

The concept of molecular chirality is used to describe molecular structures that are not superposable on their mirror images. Chiral molecules are quite prevalent in biological systems, which are primarily homochiral systems. For example, most proteins contain only L-amino acids, while DNA is made up primarily of D-deoxyribose. The relationship between chirality and helical polypeptide structure was first mentioned by Pauling (1) in reference to the α-helix. The naturally occurring L-amino acids predominantly form right-handed helices, whereas their stereoisomers D-amino acids favor left-handed helices. Indeed, right-handed helices are prevalent in biology, not only in peptides (with structural motifs such as α-helix , 3_{10} helix, and π helix) but also in the double-helical structure of B- and A-DNA. Although much less common, left-handed helices with the same chiral units as right-handed helices also exist, such as those found in PPII and in Z-DNA.

In this paper, we investigate the free energy landscape of several short polyproline peptides. Proline is unique among the natural amino acids in that its side chain is cyclized to the backbone, restricting its backbone dihedral angle to $\phi = -75°$, giving proline an exceptional rigidity and a considerably restricted conformational space. Polyproline is known to form helical structures with two well-characterized conformations: a left-handed polyproline helix (PPII) is formed when the sequential residues all adopt backbone dihedral angles (ϕ, ψ) of $(-75°, 146°)$, with all prolyl bonds in the *trans*-isomer conformation (i.e. backbone dihedral angle $\omega = 180°$) with 3 residues per turn; and a more compact right-handed polyproline helix (PPI) is formed with all sequential residues adopting dihedral angles of roughly $(-75°, 160°)$ and all prolyl bonds assume a *cis*-isomer conformation (i.e. backbone dihedral angle $\omega = 0°$) with 3.3 residues per turn. Of the 20 natural amino acids, only proline is "comfortable" in the *cis*-isomer conformation, and proline seems to be quite effective in stabilizing left-handed helices. The probability distribution for *cis-trans* prolyl bonds is affected by neighboring amino acids (2, 3), pH and ionic strength (4), solvent (5–11) and chain length (12). It has been noted experimentally that the PPII structure is favored in water, benzyl alcohol, and most of the other solvents, while the PPI structure is favored in the presence of aliphatic alcohols like propanol (12–15). The two forms can reversibly interconvert by means of changes in solvent composition (5, 7, 10, 16, 17), as analyzed via circular dichronism (CD) spectroscopy experiments (12, 18–20). It is also known that the five-membered pyrrolidine ring may adopt distinct up- and down-puckered conformations (21), which were deemed to be almost equally probable based on X-ray analyses of peptides (22) and proteins (23). Although these studies also suggest a correlation between puckering and the distribution of *cis-trans* bonds, it is now known that the prolyl ring can flip between its two states irrespective of the *cis-trans* nature of the peptide bond (24). The puckering preference of both the PPII and PPI structures has also been investigated in refs. 9 and 25–27.

The structural role of proline depends on the position it occupies in a protein. When in the middle of α helices and β sheets, proline acts as a structural disruptor, but it is generally found at the beginning of α helices and edge of β sheets, as well as in turns. It has been noted that *cis-trans* isomerization of X-Pro peptide groups is one of the rate-determining steps for folding and unfolding of various proteins (28–32) and several studies have investigated the *cis-trans* isomerization of proline-containing peptides (33–37). The role that PPII conformations play in the nonstructured states of polypeptides has also been much discussed (38–41), and experimental and theoretical evidence shows that the presence of nonproline residues decreases the PPII helix content in a pro-rich environment (39, 40). Proline oligomers also have considerable interest in their own right. Traditionally, the relatively rigid structure of PPII has been used as a "molecular ruler" in structural molecular biology, especially for the validation of spectroscopic rulers in Föster resonance energy transfer (FRET) experiments (42). However, both FRET and photoinduced-electron transfer (PET) studies show deviations of experimentally observed end-to-end distances of polyproline from theoretical predictions. Recently, Doose et al. used PET techniques to probe the structure and dynamics of PPII with 1–10 residues in aqueous solution (43). The authors showed that "polyproline samples exhibit static structural heterogeneity with subpopulations of distinct end-to-end distances that do not interconvert on the scales from nano- to milliseconds". This heterogeneity was attributed to interspersed *cis* isomers that disrupt the otherwise ideal all *trans* PPII structure, in agreement with theoretical studies based on conformational energy calculations (44, 45) and previous simulation results (40). The authors concluded that the stability of these heterogeneous subpopulations, caused by prolyl *cis-trans* isomerization, requires characterization of the *cis* isomers in order for polyproline to be

Author contributions: M.M., V.B., C.R., T.A.D., and C.S. designed research; M.M. and V.B. performed research; M.M. analyzed data; and M.M., V.B., C.R., and C.S. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. E-mail: sagui@ncsu.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0906500106/DCSupplemental.

used as a molecular ruler, for instance to calibrate FRET efficiency measurements. From a computational point of view, such a characterization is rather difficult because *cis-trans* isomerization in polyproline is much slower than the formation and rupture of hydrogen bonds in an α-helix-coil transition. Indeed, characteristic time scales for prolyl isomerization range from tens to hundreds of seconds at room temperature (2, 46), implying solvent-dependent barriers of the order of 10 to 20 kcal/mol (24, 27, 47–49).

Given these high barriers, traditional molecular dynamics (MD) simulations cannot explore the relevant conformational space for the wide range of *cis-trans* transitions. Rather, based on the recently developed adaptively biased molecular dynamics (ABMD) method (50), accurate free energy landscapes of short polyproline peptides in a few selected environments, as a function of several relevant-order parameters or collective variables, have been calculated. The ABMD method belongs to the general category of umbrella sampling methods with a time-dependent potential and has recently been implemented in the AMBER-10 modeling package (51). Different sets of collective variables have been used to study not only the two helical conformations of the polyproline peptide but also the transition pathways and the many conformations that are neither PPI nor PPII. The results provide a complementary view to the conclusions of Doose et al. (43) about the existence of stable heterogeneous subpopulations. Finally, all relevant simulation details are provided in the *SI Text* with clarifying Figs. S1–S3.

Free Energy Methods and Collective Variables

To calculate accurate free energy maps, we used the ABMD method (50), with supplementary equilibrium umbrella sampling runs (52). In addition, steered molecular dynamics (SMD) runs were used to investigate specific pathways on these free energy maps (53). The ABMD method belongs to the general category of umbrella sampling methods with a time-dependent potential and provides for an elegant way of computing the potential of mean force or free energy as a function of a collective variable $\sigma(\mathbf{r_1}, \ldots, \mathbf{r_N})$:

$$f(\xi) = -k_B T \ln p(\xi), \qquad [1]$$

where $p(\xi) =< \delta[\xi - \sigma(\mathbf{r}_1, \ldots, \mathbf{r}_N)] >$ is the probability density of the collective variable (the angular brackets denote an ensemble average; k_B, the Boltzmann constant; T, temperature). The method estimates the Landau free energy of a reaction coordinate from an evolving ensemble of realizations and uses that estimate to bias the system dynamics to flatten an effective free energy surface (which may not be exactly the same as the total conformational energy). Finally, all relevant simulation details are provided in the *SI Text*.

The ABMD method was used to calculate the free energy landscapes of short polyproline peptides [Ace-$(Pro)_n$-Nme, n = 6, 9, 13], as a function of several sets of collective variables. These collective variables are carefully chosen as to reflect the "slow modes" of the system, which are ultimately responsible for the large-scale molecular structure—i.e., the changes associated with the *cis-trans* isomerization. The "fast modes" of the system, which take place on a relatively short time scale, are properly sampled and de facto integrated out when calculating the Landau free energy. An important example of a fast mode is provided by the puckering of the pyrrolidine ring. The solvated transition barrier between the puckered up and down states of the system is of the order of a few k_BT (54), so that at T = 300 K, the ring oscillates rapidly between the two puckering conformations (see *SI Text* and Fig. S2), which are then integrated in the ensemble average of the slow modes. As a first set, we considered the collective variables of handedness (H) and the radius of gyration (R_g), which are natural because polyproline peptides are characterized by not only PPII and PPI but also by more compact "globular" conformations. The

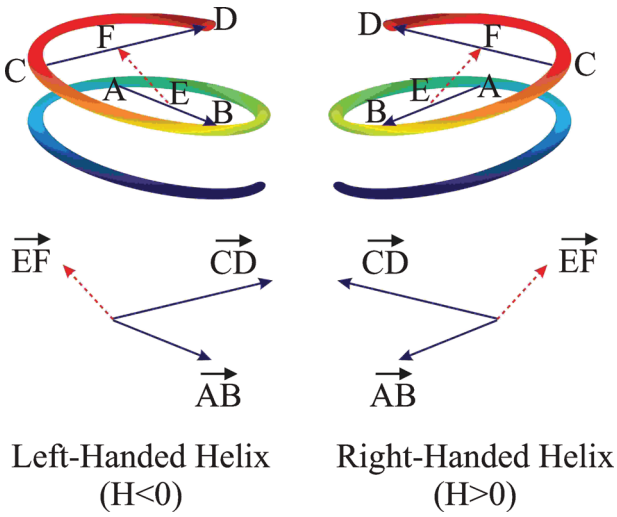

Fig. 1. Geometrical construction involved in the definition of handedness: $H < 0$ for left-handed and $H > 0$ for right-handed structures.

definition of H is illustrated in Fig. 1. Moving through the helical turns, a sequence of four points $A, \ldots, D$ is defined. These points determine the vectors $\overrightarrow{AB}$ and $\overrightarrow{CD}$. The midpoints of these vectors (E and F) form the vector $\overrightarrow{EF}$. The handedness of these points is then defined as

$$H_{ABCD} = \frac{\overrightarrow{EF} \cdot \overrightarrow{CD} \times \overrightarrow{AB}}{|\overrightarrow{EF}|.|\overrightarrow{CD}|.|\overrightarrow{AB}|}. \qquad [2]$$

Such a definition discriminates between PPI ($H_{ABCD} > 0$) helices and PPII ($H_{ABCD} < 0$) helices. To complete the definition, the four-point contributions are then summed along the backbone, thereby defining the collective variable H for the entire peptide: $H = \sum_{i=1}^{n-3} H_{i,i+1,i+2,i+3}$. Roughly speaking, the magnitude of H is a measure of the number of turns associated with the helices. For polyproline, the position of the nitrogen atoms was found to be a good choice for the location of the $A, \ldots, D$ points.

Even though the handedness is enough to differentiate between PPI and PPII, many compact structures are characterized by the same value of H. To remove this degeneracy, we considered the free energies associated with the radius of gyration R_g of the heavy atoms: $R_g = \sqrt{\sum_a \frac{m_a}{m_\Sigma}(\mathbf{r}_a - \mathbf{R}_\Sigma)^2}$. Here, $\mathbf{R}_\Sigma = \sum_a (m_a/m_\Sigma)\mathbf{r}_a$ is the center of mass, $m_\Sigma = \sum_a m_a$, and the sums run over all atoms except hydrogen. Note that R_g is not the only quantity that can be used to characterize the spatial extent of a peptide's conformation. Experimentally, results are often discussed in terms of the peptide's "end-to-end" (d_{end}) distance. Free energy maps based on either of these two variables capture essentially the same physical content, as demonstrated in the *SI Text* where several maps based on d_{end} are shown in Figs. S4–S6.

We have also defined other collective variables specifically associated with the *cis-trans* isomerization. The torsion angle that describes this transition has value $\omega = 0°$ for the *cis* (C) isomer, and $\omega = 180°$ for the *trans* (T) isomer. Thus, we define Ω as the sum of the cosines of the ω angles, which for an n-mer peptide is

$$\Omega_{n-mer} = \sum_{i=1}^{n-1} \cos \omega_i. \qquad [3]$$

For a *cis-trans* prolyl bond, $\cos \omega_i$ is +1 (−1) and, therefore, Ω_{n-mer} can take any of the following values: $-n+1$, $-n+3$, …, $n-3$, or $n-1$. The collective variable Ω, like H, is degenerate. For a sequence with $n_b = n-1$ bonds, the number of conformations with

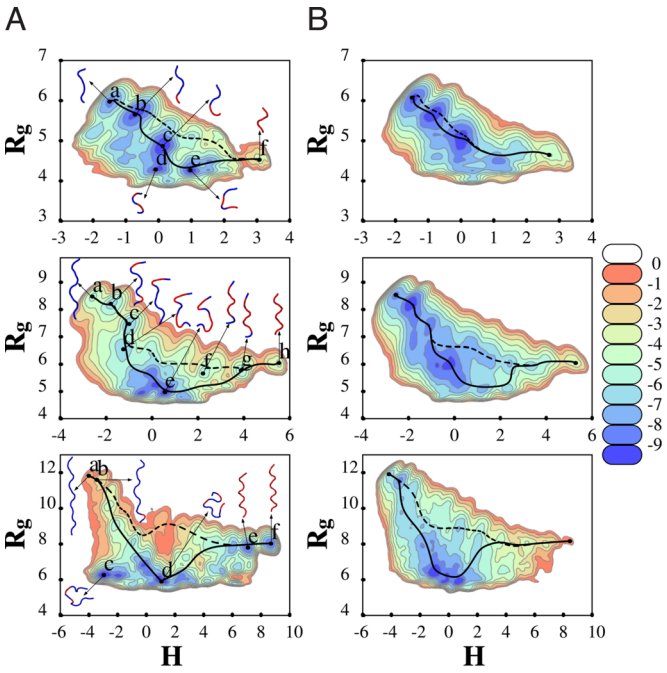

Fig. 2. Free energy landscapes (kcal/mol) as a function of collective variables (H, R_g). Landscapes are shown for a 6-mer (*Top*), 9-mer (*Middle*), and 13-mer (*Bottom*) polyproline peptide: in vacuo (*A*) and in implicit water (*B*). Examples of transition paths between PPI and PPII are shown: The solid line passes through the global minimum whereas the dashed line avoids it. A ribbon representation is used for some of the structures associated with some of the major minima, with *cis-trans* prolyl bonds highlighted in red (blue), respectively. In the 13-mer case, all the prolyl bonds of both *E* and *F* structures are *cis*, but the prolyl amide bond in the amidated terminal of the peptide (not shown) is *trans* in *E* and *cis* in *F*.

n_C *cis* bonds and n_T *trans* bonds (with $n_b = n_C + n_T$) is $n_b!/(n_C!n_T!)$. Naturally, perfect PPI is characterized by $\Omega_{n-\mathrm{mer}} = n - 1$ and perfect PPII by $\Omega_{n-mer} = -(n-1)$.

Another possibility is to consider the "interface" between bonds and define the collective variable Λ as

$$\Lambda_{n-\mathrm{mer}} = \sum_{i=1}^{n-2} \cos(\omega_i + \omega_{i+1}). \qquad [4]$$

For an n-mer, there are $n-1$ prolyl bonds and $n-2$ interfaces. If two neighboring bonds have the same dihedral angle ω, then their interface contributes $\cos(\omega_i + \omega_{i+1}) = +1$; otherwise it gives -1. Correspondingly, $\Lambda_{n-\mathrm{mer}}$ can take on any of the values $-n+2, -n+4, \ldots, n-4, n-2$. $\Lambda_{n-\mathrm{mer}}$ removes some of the degeneracy associated with Ω. For $|\Omega| = n-1$, there is only one value of $\Lambda(\Lambda = n-2)$; for $\Omega = 0$ (which exists only for odd values of n), there are $(n-2)$ values of Λ; for $|\Omega| = m$ (with $m \neq n-1$ and $m \neq 0$), there are $n-1-m$ values of Ω.

Finally, given the binary nature of this torsion angle and considering that for an n-mer there are 2^{n-1} possible configurations of the bonds, it is possible to define a collective variable that avoids the degeneracy problem entirely. First, one maps the *cis-trans* bond conformations into binary numbers by using $b_i = (\cos\omega_i + 1)/2$ and then converts the sequences $\{b_i\}$ into a decimal number:

$$\Gamma_{n-mer}(\{b_i\}) = \sum_{i=1}^{n-1} b_i 2^{i-1}. \qquad [5]$$

Therefore Γ is a number ranging from 0 (PPII) to $2^{n-1}-1$ (PPI). This new collective variable acts like a label for each possible conformation. It can be related to Ω via $\Omega_{n-mer} = \sum_{i=1}^{n-1}(2b_i - 1)$. Notice that the decimal number thus assigned does not have any physical meaning: It is just a label for the different isomers. The collective variable Γ is only convenient for the investigation of relatively short polypeptides because the number of minima associated with Γ increases exponentially with peptide length.

Results and Discussion

We now discuss the free energy landscapes of Ace-(Pro)$_n$-Nme peptides ($n = 6, 9, 13$) both in vacuo and for selected solvated environments (implicit water, hexane, and 1-propanol).

Fig. 2*A* shows the (H, R_g) free energy map for the peptides in vacuo. For discussion purposes, we focus primarily on the nanomer free energy map. It displays three sets of minima: Three minima (f, g, h) with $H > 2$ correspond to PPI; a broad basin with several minima (d, e) between $-1.5 < H < 2$ with $R_g < 7$ correspond to twisted and globular structures; and three minima (a, b, c) with $H < -0.95$ and $R_g > 7$, correspond to PPII. A ribbon representation of the typical structures in these minima is also shown in Fig. 2*A*. A quantitative description of the minima (location, depth) and characteristic isomer is given in Table S1. The isomer associated with minimum (a) corresponds most closely to the ideal PPII structure ($RMSD = 0.4$ Å), whereas that associated with (h) is closest to PPI ($RMSD = 1.0$ Å). There are qualitative changes in the free energy maps as the length of polyproline is increased. Although the same three classes of minima are found on all three maps, their relative importance changes. For short 6-mers, the free energy favors the formation of PPII structures, as opposed to PPI. For the longer 13-mer, the trend is reversed, and PPI is favored. The 9-mer appears to correspond to an intermediate case, such that PPI and PPII structures are approximately balanced. Such a change with peptide length has also been noted experimentally in alcohol environments. Free energy maps for polyproline in implicit water (Fig. 2*B*) are qualitatively similar, except that PPII is favored over PPI for all chain lengths, in agreement with experimental results (12). In implicit water, the free energy minima associated with PPII are of the same magnitude as those associated with the compact structures, and they form a "broad valley" with relatively low barriers.

The (Ω, R_g) and (Ω, Λ) free energy maps for n-mers in implicit water are shown in Fig. 3 (gas phase results are qualitatively

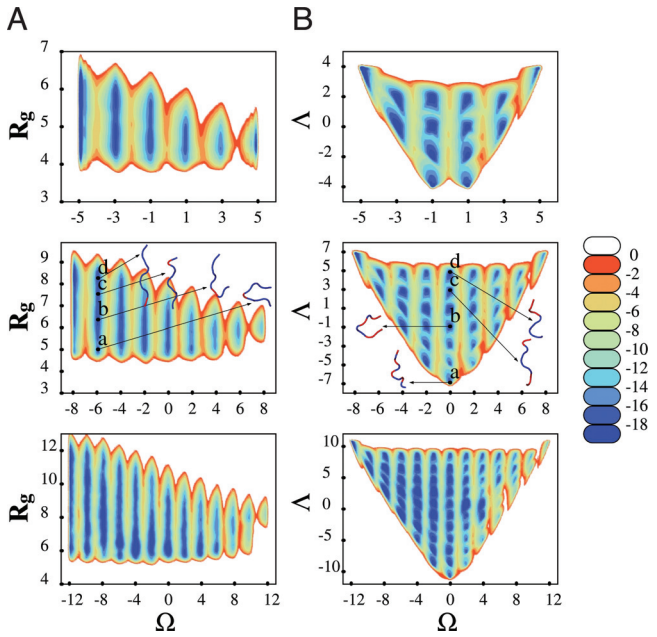

Fig. 3. Shown are (R_g, Ω) and (Λ, Ω) free energy landscapes (kcal/mol) for a 6-mer (*Top*), 9-mer (*Middle*), and 13-mer (*Bottom*) polyproline in implicit water.

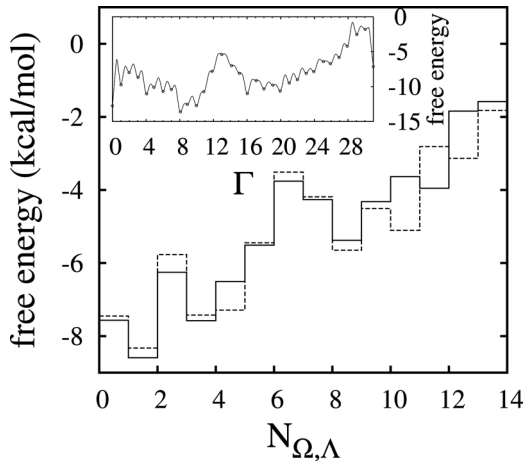

Fig. 4. Free energy in the $N_{\Omega,\Lambda}$ representation. The solid line is obtained from the (Ω, Λ) map by enumerating the minima and plotting the corresponding free energy values; the dotted line is obtained by using the Γ results combined with Eq. **6**. (*Inset*). Free energy profile (kcal/mol) of a hexamer polyproline in vacuo, as a function of the single collective variable Γ.

similar and shown in the *SI Text* and Fig. S7). Because for an n-mer polypeptide there are $n-1$ peptide bonds, Ω runs between ± 5, ± 8 and ± 12 for the different length peptides. Allowed values for Ω are centered on even values for the 9,13-mers, and about odd values for the 6-mer. The free energy minima associated with the Ω values are also elongated in the R_g direction, forming narrow valleys, which is a reflection of the underlying structures seen in the ribbon representations of Fig. 3*A*. For example, the flipping of a single bond within an all-*trans* helix does not greatly modify the helical structure if the bond is near the end, but it produces a sharp turn when it is in the center. For the nanomer, the deepest minima correspond to globular and PPII structures. Because the PPII-like helices are more stretched, they have more flexibility in terms of R_g, unlike PPI-like structures which are more rigid. This flexibility translates into a net decrease of the elongation of the free energy contours as one moves from PPII to PPI. Finally, because for $n-1$ bonds there are $n-2$ possible "interfaces" between them, the possible values of Λ vary between ± 4, ± 7, and ± 11 for the 6−, 9−, and 13-mer, respectively. For the nanomer, $\Lambda = 7$ corresponds to either a complete PPI or PPII helix because both of these structures have no *cis-trans* interfaces. In contrast, $\Omega = 0$ and $\Lambda = -7$ correspond to a situation of all alternating *cis-trans* bonds.

The 1D free energy profile as a function of Γ for a hexamer in vacuo is shown as an inset in Fig. 4. The label Γ runs between 0 (PPII) and 31 (PPI) because there are 32 possible isomers involved. To make the connection with this collective variable and other ABMD variables, we have calculated the corresponding values of Ω and Λ for each of the possible combination of b_i's. Because of the inherent degeneracy, multiple Γ's may have the same value for Ω and Λ. Suppose that $\Gamma_1, \Gamma_2, \ldots, \Gamma_m$ all have the same values of Ω and Λ. In such a case, the probability of finding a structure with a given value of Ω and Λ is the sum of the probabilities related to $\Gamma_1, \Gamma_2, \ldots, \Gamma_m$. We can therefore define a new phase space—labeled with $N_{\Omega,\Lambda}$—and relate the free energies in this space to those in Γ space via

$$\exp\left(\frac{-F_{N_{\Omega,\Lambda}}}{k_B T}\right) = \sum_{i=1}^{m} \exp\left(\frac{-F_{\Gamma_i}}{k_B T}\right). \qquad [6]$$

To illustrate this explicitly, consider the (Ω, Λ) free energy plot for a 6-mer shown in Fig. 3. This map is characterized by 14 distinct minima, which we map onto the single variable $N_{\Omega,\Lambda}$, which now runs from 1 to 14. Moving from left to right, and then bottom to top, we label each minima in (Ω, Λ) space by a number $N_{\Omega,\Lambda}$,

Table 1. The values of the order parameters Γ, Ω, Λ, and $N_{\Omega,\Lambda}$ for different configurations of a 6-mer polyproline peptide

$N_{\Omega,\Lambda}$	Ω	Λ	Γ	Sequence
1	−5	4	0	TTTTT
2	−3	0	2, 4, 8	TCTTT, TTCTT, TTTCT
3	−3	2	1, 16	CTTTT, TTTTC
4	−1	−4	10	TCTCT
5	−1	−2	5, 9, 18, 20	CTCTT, CTTCT, TCTTC, TTCTC
6	−1	0	6, 12, 17	TCCTT, TTCCT, CTTTC
7	−1	2	3, 24	CCTTT, TTTCC
8	1	−4	21	CTCTC
9	1	−2	11, 13, 22, 26	CCTCT, CTCCT, TCCTC, TCTCC
10	1	0	14, 19, 25	TCCCT, CCTTC, CTTCC
11	1	2	7, 28	CCCTT, TTCCC
12	3	0	23, 27, 29	CCCTC, CCTCC, CTCCC
13	3	2	15, 30	CCCCT, TCCCC
14	5	4	31	CCCCC

All $32 = 2^5$ possible conformations of the prolyl bonds are presented. The acetylated end is at the left and the amidated end at the right.

i.e., $N_{-5,4} = 1$ (PPII), $N_{-3,0} = 2$, $N_{-3,2} = 3, \ldots, N_{5,4} = 14$ (PPI). Table 1 gives the Γ, Ω, Λ, and $N_{\Omega,\Lambda}$ values for all 32 possible configurations of a 6-mer polyproline, and Fig. 4 plots the free energy as computed from the variable Γ for the 14 different states $N_{\Omega,\Lambda}$ and compares it with the ABMD runs with (Ω, Λ). The correspondence between them is quite good, with only small, expected free energy differences (see also Fig. S8).

Now we turn to explicit solvent results. We investigated a nanomer in explicit nonpolar hexane and in 1-propanol, with free energy maps shown in Fig. 5. In hexane, the phase space associated with the PPI helices is reduced, as compared with the gas-phase and implicit water results: The only minima identified correspond to PPII and globular structures. By contrast, 9-mer polyproline in 1-propanol shows minima for the compact and PPI structures.

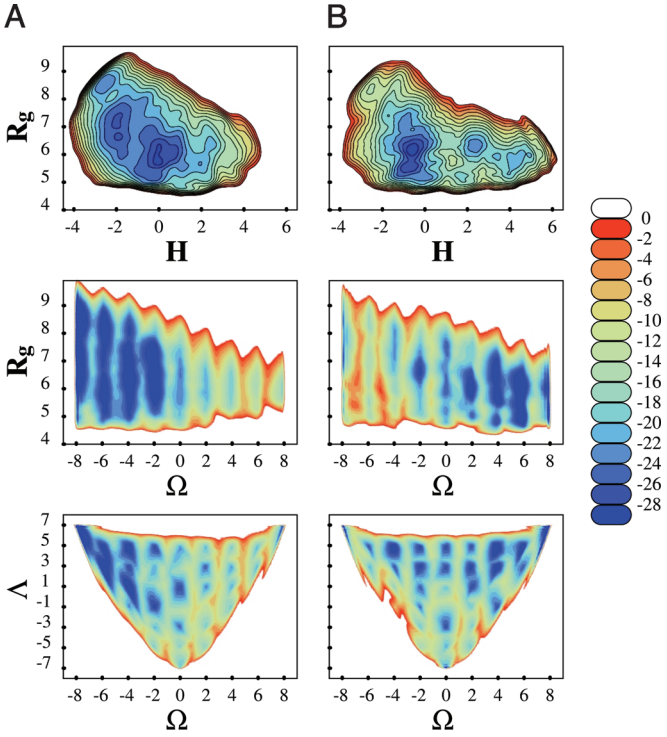

Fig. 5. Free energy landscapes (kcal/mol) of a 9-mer polyproline in hexane (*A*) and 1-propanol (*B*), as obtained from ABMD runs followed up with umbrella corrections.

Table 2. Free energy difference (in kcal = mol) of all *cis* PPI and all *trans* PPII [*f*(PPI)-*f*(PPII)], for 6-, 9-, and 13-mer Ace-$(Pro)_n$-Nme, in vacuo, in implicit water, and in explicit solvents hexane and propanol (for 9-mer), obtained from ABMD simulations with different sets of collective variables

Environment	(H,R_g)	(Ω,R_g)	(Ω,Λ)	Γ
		6-mer		
In vacuo	5.12	5.22	6.02	5.63
Water	4.88	5.20	5.33	5.15
		9-mer		
In vacuo	1.05	1.42	1.24	1.52
Water	4.53	5.32	6.08	6.43
Hexane	18.13	19.68	20.12	–
Propanol	−6.98	−6.13	−8.19	–
		13-mer		
In vacuo	−7.21	−6.22	−6.65	–
Water	6.47	9.60	8.05	–

The fields marked by blank lines were not computed.

Again, these results, agree with trends observed in experiments. What is the origin of this differing behavior of polyproline in the two different solvents? First, the characteristic ring of polyproline precludes the nitrogen atom of the prolyl bond from engaging in hydrogen bonding. This feature not only affects the structure of the helices but also the interaction with solvents. The *cis-trans* conformations are linked to the orientation of the carboxyl groups C = O: In PPI, these groups are almost parallel to the axis of the compact helix and are shielded from the solvent by proline rings; in PPII, the C = O groups are mainly perpendicular to the axis of the helix, and because the helix is more elongated, the carboxyl oxygen is more exposed to solvents. Indeed, results for different radial distributions linking the carboxyl oxygen and different atoms in the solvent showed the following: (*i*) the distances between oxygen in C = O and hexane atoms are shorter in PPII than in PPI (therefore hexane favors PPII); (*ii*) the distances between O and propanol atoms are shorter in PPI than PPII (therefore propanol favors PPI); and (*iii*) the distances related to the hexane atoms in both cases are shorter than the distances related to propanol (therefore the effects of hexane are stronger than the effects of propanol).

As a short summary, Table 2 gives the calculated free energy differences between the PPI and PPII structures, as obtained from the different runs with different collective variables. There is good agreement between all the different numerical values.

To further characterize the ABMD-based free energy maps, we performed SMD runs that use the topology of these landscapes to steer the polyproline peptide between PPI and PPII. The aim here is not to compute accurate free energy differences but rather to compare pathways associated with different mechanisms in a qualitative way. Here, we choose selected trajectories, such as those given by the lowest free energy path (LFEP) method (55) in the (H,R_g) plane, paths on the (Ω,Λ) plane, paths where Ω is the only variable, etc.

Two LFEP paths identified between PPI and PPII are shown in Fig. 2. One such path passes through the global minimum (solid line), but a second path avoids it (dotted line). This second path is characterized by considerably lower free energy barriers, as shown in Fig. 6*A*. It is plausible that the two different paths are associated with different transition mechanisms. Specifically, for the path avoiding the global minimum, sampling of intermediate conformations along the path is mostly consistent with a zipper-like mechanism, in which changes take place via the successive switching of neighboring prolyl bonds (i.e., for 6-mer undergoing a PPII→PPI transition, this would take the form TTTTT→CTTTT→CCTTT… →CCCCC). Nucleation begins primarily at the ends or in their near vicinity, with a rate that appears to be length dependent. In contrast, for the path through the global minima, bond flipping takes place in a less-ordered fashion. The helix seems to undergo melting before reassembling again. This view is supported by examining the structures along the path. The corresponding work done by carrying out SMD along the two paths from PPI to PPII is shown in Fig. 6*B*. There is less work done along the path that avoids the global minimum. There are many degeneracies associated with the H and R_g collective variables, and as a result the work function is just monotonically increasing. It is possible to overcome some of the ambiguities related to the degeneracy in the minima by steering the system from a pure PPI structure to a pure PPII structure by using different collective variables. We have tried this both by using the 1D variable Ω and also by choosing a path in the (Ω, Λ) phase space by keeping Λ at its highest possible value (Fig. 6*C*). In both cases, the total work is less than in the other two trajectories, indicating that it is much more natural for the system to move along such a path. The work as a function of time (Fig. 6*D*) shows eight periodic looking barriers that correspond to the flipping of each of the bonds and that can be interpreted as a more clear signature of a zipper-like mechanism. We have studied these pathways and others not shown here. The SMD runs can give a qualitative description of different pathways as well as possible accompanying mechanisms, but unless all trajectories are examined, it is difficult to make quantitative assertions. The zipper-like mechanism is a highly probable mechanism for helix conversion but not necessarily the most probable mechanism. As another example, the work for the zipper-like mechanism is compared in Fig. 6*D* with that corresponding to an extreme case where all the bonds (except one) flip simultaneously. The cosines of each of the prolyl bonds for both mechanisms are shown in Fig. 6*E* and 6*F*. The huge barrier linked to the "en masse" mechanism would make this transition rather improbable.

Summary

This work presents a comprehensive study of polyproline as a paradigm of molecules whose conformational phase space contains

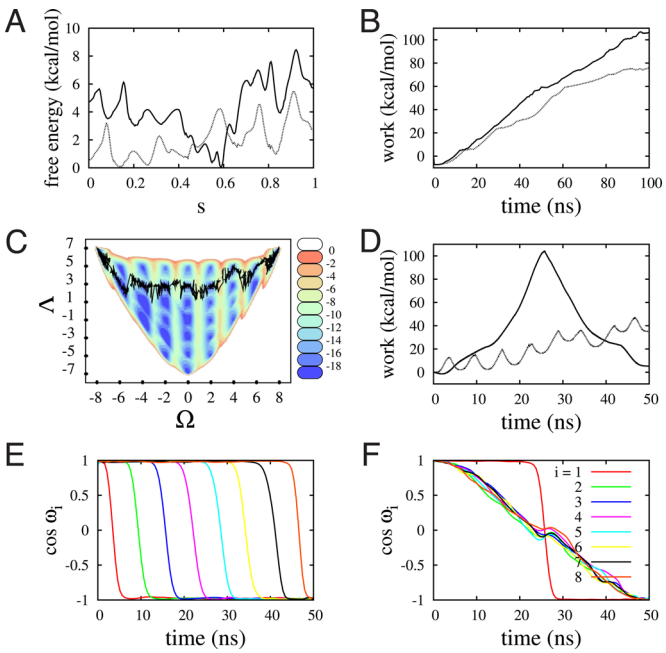

Fig. 6. Paths and work done via SMD for a nanomer in vacuo. (*A*) Free energy profiles for the LFEP paths shown in Fig. 2*A* as function of normalized path length (*s*) in (H,R_g) space, (solid line passes through the global minimum and dotted line avoids it). (*B*) Work done via SMD for the paths shown in *A*. (*C*) A trajectory associated with a zipper-like mechanism in the (Ω, Λ) space. (*D*) Work done via SMD for the path shown in *C* (dashed line) and for another trajectory associated with an "en masse" flipping of the *cis-trans* bonds (solid line). (*E* and *F*) Time evolution of the cosines for each individual prolyl bond. Here, (*E*) is associated with the zipper-like transition and (*F*) with the "en masse" flipping mechanism (except for one bond).

left-handed and/or right-handed helices. The free energies were calculated as a function of a variety of relevant collective variables: handedness H, radius of gyration R_g, and the variables Ω, Λ, and Γ, associated specifically with the *cis-trans* isomerization of the peptide bond. The theoretical framework developed here may readily be used for a quantative description of other helical molecules involving bistable changes in the relevant molecular conformation (e.g., the *anti\syn* flipping of nucleotides in B- to Z-DNA transistions). Specifically, the free energy landscapes of polyproline n-mers ($n = 6, 9, 13$) in vacuo and in implicit water were investigated, along with the nanometer in explicit hexane and 1-propanol. We also considered minimum free energy pathways through these landscapes, which were further investigated via SMD runs. General trends are as follows: The stability of PPI is enhanced by increasing the length of the peptide chain, which means that cooperativity or interresidue interaction is more favorable in PPI than PPII. Pure PPI and PPII are just one of the minima (when present) in the phase space, with many other minima corresponding to other stable subpopulations whose representative structures, in terms of *cis-trans* bonds, are readily obtainable. Favored structures as a result of interactions with different solvents agree qualitatively with experiments. The position of the carboxyl oxygen (parallel to the helical axis and hidden by proline rings in compact PPI; perpendicular to the axis and more exposed to solvent in the more-open PPII) plays an important role in the interaction with solvents. Different mechanisms of transitions are associated with different pathways in the phase space, some extreme examples being the case where all bonds rotate together, a melting and reordering mechanism, and a more probable zipper-like mechanism. Our results provide a complementary view to recent PET experiments regarding the existence of stable heterogeneous subpopulations of polyproline peptide conformations.

ACKNOWLEDGMENTS. We thank Dr. Jung Goo Lee for many useful discussions. This research was supported by the National Science Foundation Career Grants DMR-0348039 and FRG-0804549. In addition, we thank the North Carolina State University High Performance Computing (HPC) center for extensive computational support.

1. Pauling L, Corey RB, Branson HR (1951) The structure of proteins: Two hydrogen-bonded helical configurations of the polypeptide chain. *Proc Natl Acad Sci USA* 37:205–211.
2. Reimer U, et al. (1998) Side-chain effects on peptidyl-prolyl *cis/trans* isomerisation. *J Mol Biol* 279:449–460.
3. Grathwohl C, Wuthrich K (1976) The x-pro peptide bond as an nmr probe for conformational studies of flexible linear peptides. *Biopolymers* 15:2025–2041.
4. Grathwohl C, Wuthrich K (1976) Nmr studies of the molecular conformations in the linear oligopeptides h-(l-ala)*n* l-pro-oh. *Biopolymers* 15:2043–2057.
5. Steinberg IZ, Harrington WF, Berger A, Sela M, Katchalski E (1960) The configurational changes of poly-l-proline in solution. *J Am Chem Soc* 82:5263–5279.
6. Gornick F, Mandelkern L, Diorio AF, Roberts DE (1964) Evidence for a cooperative intramolecular *trans*ition in poly-l-proline. *J Am Chem Soc* 86:2549–2555.
7. Mandelkern L (1967) Poly-L-proline. In *Poly- -Amino Acids*, ed Fasman GD (Marcel Dekker, New York), pp 675–724.
8. Strassmair H, Engel J, Zundel G (1969) Binding of alcohols to the peptide co-group of poly-l-proline in the i and ii conformation. i. Demonstration of the binding by infrared spectroscopy and optical rotatory dispersion. *Biopolymers* 8:237–246.
9. Tanaka S, Scheraga HA (1975) Theory of the cooperative *trans*ition between two ordered conformations of poly(l-proline). ii. Molecular theory in the absence of solvent. *Macromolecules* 8:504–516.
10. Tanaka S, Scheraga HA (1975) Theory of the cooperative *trans*ition between two ordered conformations of poly(l-proline). iii. Molecular theory in the presence of solvent. *Macromolecules* 8:516–521.
11. Kofron JL, Kuzmic P, Kishore V, Colon–Bonilla E, Rich DH (1991) Determination of kinetic constants for peptidyl prolyl *cis-trans* isomerases by an improved spectrophotometric assay. *Biochemistry* 30:6127–6134.
12. Kakinoki S, Hirano Y, Oka M (2005) On the stability of polyproline-i and ii structures of proline oligopeptides. *Poly Bul* 53:109–115.
13. Traub W, Shmueli U (1963) Structure of poly-l-proline i. *Nature* 198:1165–1166.
14. Cowan PM, McGavin S (1955) Structure of poly-l-proline. *Nature* 176:501–503.
15. Mutter M, Wohr T, Gioria S, Keller M (1999) Pseudo-prolines: Induction of *cis/trans*-conformational interconversion by decreased *trans*ition state barriers. *Biopolymers* 51:121–128.
16. Kurtz J, Berger A, Katchalski E (1956) Mutarotation of poly-l-proline. *Nature* 178:1066–1067.
17. Steiberg IZ, Berger A, Katchalski E (1958) Reverse mutarotation of poly-l-proline. *Biochim Biophys Acta* 28:647–648.
18. Okabayashi H, Isemura T, Sakakibara S (1968) Steric structure of l-proline oligopeptides. ii. Far-ultraviolet absorption spectra and optical rotations of l-proline oligopeptides. *Biopolymers* 6:323–330.
19. Helbecque N, Loucheux–Lefebvre MH (1982) Critical chain length for polyproline-ii structure formation in h-gly-(pro)n-oh. *Int J Pept Protein Res* 19:94–101.
20. Crespo L, et al. (2002) Peptide dendrimers based on polyproline helices. *J Am Chem Soc* 124:8876–8883.
21. Momany FA, McGulre RF, Burgess AW, Scheraga HA (1975) Energy parameters in polypeptides. vii. Geometric parameters, partial atomic charges, nonbonded interactions, hydrogen bond interactions, and intrinsic torsional potentials for the naturally occurring amino acids. *J Phys Chem* 79:2361–2381.
22. Madison V (1977) Flexibility of the pyrrolidine ring in proline peptides. *Biopolymers* 16:2671–2692.
23. Vitagliano L, Berisio R, Mastrangelo A, Mazzarella L, Zagari A (2001) Preferred proline puckerings in *cis* and*trans* peptide groups: Implications for collagen stability. *Protein Sci* 10:2627–2632.
24. Kang YK, Choi HY (2004) *Cistrans* isomerization and puckering of proline residue. *Biophys Chem* 111:135–142.
25. Tanaka S, Scheraga HA (1974) Calculation of conformational properties of oligomers of l-proline. *Macromolecules* 7:698–705.
26. Zhong H, Carlson HA (2006) Conformational studies of polyprolines. *J Chem Theory Comput* 2:342–353.
27. Kang YK, Jhon JS, Park HS (2006) Conformational preferences of proline oligopeptides. *J Phys Chem B* 110:17645–17655.
28. Brandts JF, Halvorson HR, Brennan M (1975) Consideration of the possibility that the slow step in protein denaturation reactions is due to *cis-trans* isomerism of proline residues. *Biochemistry* 14:4953–4963.
29. Tanaka S, Scheraga HA (1977) Hypothesis about the mechanism of protein folding. *Macromolecules* 10:291–304.
30. Schmid FX, Mayr LM, Mücke M, Schönbrunner ER (1993) Prolyl isomerases: Role in protein folding. *Adv Protein Chem* 44:25–66.
31. Houry WA, Scheraga HA (1996) Nature of the unfolded state of ribonuclease a: Effect of *cis-trans* x-pro peptide bond isomerization. *Biochemistry* 35:11719–11733.
32. Wedemeyer WJ, Welker E, Scheraga HA (2002) Proline *cis-trans* isomerization and protein folding. *Biochemistry* 41:14637–14644.
33. Zimmerman SS, Pottle MS, Némethy G, Scheraga HA (1977) Conformational analysis of the 20 naturally occurring amino acid residues using ecepp. *Macromolecules* 10:1–9.
34. Vásquez M, Némethy G, Scheraga HA (1983) Computed conformational states of the 20 naturally occurring amino acid residues and of the prototype residue a-aminobutyric acid. *Macromolecules* 16:1043–1049.
35. Némethy G, et al. (1992) Energy parameters in polypeptides: 10. Improved geometrical parameters and nonbonded interactions for use in the ecepp/3 algorithm, with application to proline-containing peptides. *J Phys Chem* 96:6472–6484.
36. Stein RL (1993) Mechanism of enzymatic and nonenzymatic prolyl *cis-trans* isomerization. *Adv Protein Chem* 44:1–24.
37. Kang YK, Jhon JS, Han SJ (1999) Conformational study of ac-xaa-pro-nhme dipeptides: Proline puckering and *trans/cis* imide bond. *J Pept Res* 53:30–40.
38. Woody RW (1992) Circular dichroism and conformation of unordered polypeptides. *Adv Biophys Chem* 2:37–39.
39. Kelly MA, et al. (2001) Host–guest study of left-handed polyproline ii helix formation. *Biochemistry* 40:14376–14383.
40. Vila JA, Baldoni HA, Ripoll DR, Ghosh A, Scheraga HA (2004) Polyproline ii helix conformation in a proline-rich environment: A theoritical study. *Biophys J* 86:731–742.
41. Makowska J, et al. (2006) Polyproline ii conformation is one of many local conformational states and is not overall conformation of unfolded peptides and proteins. *Proc Natl Acad Sci USA* 103:1744–1749.
42. Stryer L, Haugland RP (1967) Probing polyproline structure and dynamics by photoinduced electron *trans*fer provides evidence for deviations from a regular polyproline type ii helix. *Proc Natl Acad Sci USA* 58:719–726.
43. Doose S, Neuweiler H, Barsch H, Sauer M (2007) Probing polyproline structure and dynamics by photoinduced electron *trans*fer provides evidence for deviations from a regular polyproline type ii helix. *Proc Natl Acad Sci USA* 104:17400–17405.
44. Tanaka S, Scheraga HA (1975) Calculation of the characteristic ratio of randomly coiled poly(l-proline). *Macromolecules* 8:623–631.
45. Zimmerman SS, Scheraga HA (1976) Stability of *cis*, *trans*, and nonplanar peptide groups. *Macromolecules* 9:408–416.
46. Grathwohl C, Wuthrich K (1981) Nmr studies of the rates of proline *cis-trans* isomerization in oligopeptides. *Biopolymers* 20:2623–2633.
47. Venkatachalam CM, Price BJ, Krimm S (2004) A theoretical estimate of the energy barriers between stable conformations of the proline dimer. *Biopolymers* 14:1121–1132.
48. Kang YK (2006) Conformational preferences of non-prolyl and prolyl residues. *J Phys Chem B* 110:21338–21348.
49. Jhon JS, Kang YK (1999) Imide *cis-trans* isomerization of n-acetyl-n[prime]-methylprolineamide and solvent effects. *J Phys Chem A* 103:5436–5439.
50. Babin V, Roland C, Sagui C (2008) Adaptively biased molecular dynamics for free energy calculations. *J Chem Phys* 128:134101.
51. Case DA, et al. (2008) AMBER 10 (University of California, San Fran*cis*co).
52. Babin V, Roland C, Darden TA, Sagui C (2006) The free energy landscape of small peptides as obtained from metadynamics with umbrella sampling corrections. *J Chem Phys* 125:2049096.
53. Izrailev S, et al. (1998) *Steered Molecular Dynamics. Computational Molecular Dynamics: Challenges, Methods, Ideas* (Springer, Berlin), pp 39–65.
54. Kang YK (2007) Puckering *trans*ition of proline residue in water. *J Phys Chem B* 111:10550–10556.
55. Ensing B, Laio A, Parrinello M, Klein M (2005) A recipe in the computation of the free energy barrier and the lowest free energy path of concerted reactions. *J Phys Chem B* 109:6676–6687.