Polypeptide Motions Are Dominated by Peptide Group Oscillations Resulting from Dihedral Angle Correlations between Nearest Neighbors†

James E. Fitzgerald,[‡,§] Abhishek K. Jha,[||,⊥] Tobin R. Sosnick,*[,§,⊥] and Karl F. Freed*[,||]

Department of Physics, Department of Mathematics, Department of Biochemistry and Molecular Biology, Department of Chemistry, Institute for Biophysical Dynamics, and The James Franck Institute, The University of Chicago, Chicago, Illinois 60637

Received August 3, 2006; Revised Manuscript Received October 27, 2006

ABSTRACT: To identify basic local backbone motions in unfolded chains, simulations are performed for a variety of peptide systems using three popular force fields and for implicit and explicit solvent models. A dominant "crankshaft-like" motion is found that involves only a localized oscillation of the plane of the peptide group. This motion results in a strong anticorrelated motion of the Φ angle of the *i*th residue (Φ_i) and the Ψ angle of the residue $i - 1$ (Ψ_{i-1}) on the 0.1 ps time scale. Only a slight correlation is found between the motions of the two backbone dihedral angles of the same residue. Aside from the special cases of glycine and proline, no correlations are found between backbone dihedral angles that are separated by more than one torsion angle. These short time, correlated motions are found both in equilibrium fluctuations and during the transit process between Ramachandran basins, e.g., from the β to the α region. A residue's complete transit from one Ramachandran basin to another, however, occurs in a manner independent of its neighbors' conformational transitions. These properties appear to be intrinsic because they are robust across different force fields, solvent models, nonbonded interaction routines, and most amino acids.

All biological processes are perforce dynamical events. Local and global conformational changes are crucial to many of the biological functions performed by proteins. Since protein motions occur on a wide range of time scales (from subpicoseconds to seconds), they have been probed using many different experimental and theoretical techniques. However, relatively little attention has been devoted to elucidating the nature of the local changes (*1*−*8*), in spite of the fact such changes significantly contribute to most conformational transformations.

The geometry of a protein backbone is largely described by a pair of backbone dihedral angles, Φ and Ψ. A third torsion angle ω exists between the amide nitrogen and carbonyl carbon, but this angle is essentially fixed at 180° due to the partial double bond character of the bond between the nitrogen and oxygen. If the protein backbone were rigid, a change in either of the Φ or Ψ angles would pivot a large portion of the protein, moving each atom by an amount proportional to its distance from the axis of rotation. In a protein consisting of hundreds of amino acids, this rotation would involve a huge displacement of the protein and would be strongly prohibited due to inertial and viscous drag. Hence, such large-scale motions do not occur.

To alter the dihedral angles while avoiding a huge unfavorable displacement, several folding models have utilized crankshaft motions, cooperative motions involving three neighboring torsion angles in which all other torsion angles remain relatively unperturbed (*1*, *7*, *9*−*12*). Through this crankshaft motion, neighboring residues cooperate to eliminate large-scale motions, long-range steric clashes, and excessive viscous or inertial drag caused by moving large portions of the protein through the solvent. However, whether cooperative crankshaft motions that eliminate huge displacements exist remains unclear. Simulations of synthetic polymers demonstrate the lack of a link between dynamical chain cooperativity and crankshaft motions (*3*, *13*, *14*). Instead, the simulations find that some macromolecules achieve the chain cooperativity necessary to avoid excessive displacements without crankshaft motions (*3*, *13*, *14*). The conditions under which crankshaft motions might persist remain unanswered for proteins.

Local moves are essential to any motion where long-range displacements must be avoided. Theoretical investigations have considered the types of movements required to leave long-range structure unperturbed (*15*, *16*). These local movements have been used to model the dynamics of protein backbones in an effort to make Monte Carlo conformational searches more efficient (*16*, *17*). In addition to strictly local crankshaft moves, these studies also utilize longer-range

† This work is supported by grants from the National Institutes of Health (GM55694 to T.R.S.) and the National Science Foundation (CHE-0416017 to K.F.F.) and Grant 0353989. J.E.F. is a University of Chicago Beckman Scholar and acknowledges the Arnold and Mabel Beckman Foundation. A.K.J. acknowledges the support of Burroughs Wellcome Fund Interfaces 1001774.

* To whom correspondence should be addressed. T.R.S.: phone, (773) 834-0657; fax, (773) 702-0439; e-mail, trsosnic@uchicago.edu. K.F.F.: phone, (773) 702-7202; fax, (773) 702-5863; e-mail, freed@uchicago.edu.

‡ Department of Physics and Department of Mathematics.

§ Department of Biochemistry and Molecular Biology.

|| Department of Chemistry and The James Franck Institute.

⊥ Institute for Biophysical Dynamics.

10.1021/bi061575x CCC: $37.00
Published on Web 12/22/2006

motions involving several residues to avoid large, unphysical displacements of the chain by rotating multiple interior residues' dihedral angles in such a manner that the ends of the loop are constrained to their initial positions.

Evidence of the presence of anticorrelated motions of the backbone torsion angles ($\Delta\Phi_i$ and $\Delta\Psi_{i-1}$) has appeared in normal mode analyses for alanine and glycine in helical structures (*1*). More recently, these correlations have been observed in several simulations of proteins, which find that this anticorrelation is part of a crankshaft motion in which more distant torsion angles are uncorrelated (*2*, *4*−*7*, *18*−*20*). However, these computational studies are performed using molecular dynamics simulations that are too short to ensure sufficient equilibration of the trajectories (presumably due to computational limitations). In addition, recent simulations demonstrate that different force fields yield strikingly different results for the backbone dynamics of small peptides (*21*−*25*). It is, therefore, important to reexamine these types of motions in a more general and systematic context.

Very recently, structure refinement using heteronuclear NMR relaxation measurements has detected evidence of local crankshaft motions that involve anticorrelations between the motions of torsion angles Φ_i and Ψ_{i-1} (*26*). Simulations using experimental constraints find that to a large degree, crankshaft motions exist across the entire protein backbone, with only a few regions exhibiting a lack of this correlated mode of motion.

Here we delineate the conditions under which crankshaft motions are present by performing simulations with modern force fields and various solvent models to test the degree to which these motions are ubiquitous in protein backbone dynamics. Our simulations consider three popular force fields, a variety of nonbonded interaction routines, and both implicit and explicit solvent models. Simulations are performed using a variety of amino acid sequences, thereby enabling a comparison of how the different force fields and solvent models describe the correlated motions and other dynamic and equilibrium properties. Finally, we discuss the implications of our findings.

METHODS

Simulation Protocol. Computer simulations are run for a variety of sequences, force fields, solvation models, and nonbonded interaction routines. The sequences are chosen to test whether there is a dependence on the number and type of residues. Simulations of an alanine$_7$ (Ala_7) and an alanine$_{20}$ (Ala_{20}) chain are used to test the influence of peptide length. A heteropolymer (EDEVARLKKLLW) is investigated because it both has sequence heterogeneity and is short in length while possessing significant helical structure according to circular dichroism measurements (*27*). This helical structure is important for observing numerous transitions involving the α-basin, which are found to be much rarer than $\beta \leftrightarrow$ Polyproline II (PPII) transitions in the Ala_7 and Ala_{20} oligomers. Because proline and glycine behave rather differently compared to other residues, simulations have also been run for a proline$_7$ (Pro_7) and a glycine$_7$ (Gly_7) chain. All peptides are amino-acetylated and carboxy-amidated to eliminate charges at the termini.

Langevin dynamics (LD) simulations of these sequences have been probed using the implicit solvent model developed by Shen and Freed (*28*, *29*) which has been tested by comparison with explicit solvent simulations. The all-atom LD simulations proceed by solving the Langevin equation

$$m_i \frac{\partial \vec{u}_i}{\partial t} = -\vec{\nabla}_i U - \varsigma_i \vec{u}_i + \vec{A}_i(t)$$

where $\vec{u}_i$ is the velocity of atom i, ζ_i is its friction coefficient, m_i is its mass, $\vec{A}_i(t)$ is the random force acting on atom i, and U is the energy function. These simulations are performed with several different energy functions as described below. In general, the energy function is a combination of an all-atom force field that represents the solute−solute interactions and implicit solvation terms which attempt to reproduce the influence of the solvent. These energy functions are given by

$$U_{\text{solute}} = U_{\text{bond}} + U_{\text{bend}} + U_{\text{torsion}} + U_{\text{improper}} + U_{\text{VDW}}$$

$$U_{\text{solvent}} = U_{\text{charge}}(\epsilon) + U_{\text{solvation}}(\sigma)$$

such that the total potential is

$$U = U_{\text{solute}} + U_{\text{solvent}}$$

The energy function, U_{solute}, is henceforth termed the force field. We use several commonly used force fields, namely, CHARMM27 (*30*), all-atom OPLS (*31*), and Garcia and Sanbonmatsu's GS-AMBER94 (*24*). Differences between these force fields result in different equilibrium Φ and Ψ distributions, as well as in the prediction of different dynamical properties (*21*). All three force fields utilize individual energy terms of the form

$$U_{\text{bond}} = \sum_{\text{bonds}} {}^1/_2 K_b (b - b_o)^2$$

$$U_{\text{bend}} = \sum_{\text{bond angles}} {}^1/_2 K_\theta (\theta - \theta_o)^2$$

$$U_{\text{torsion}} = \sum_{\text{torsion angles}} K_\phi [1 + \cos(n\phi - \delta)]$$

$$U_{\text{improper}} = \sum_{\text{improper torsion angles}} {}^1/_2 K_\chi (\chi - \chi_o)^2$$

$$U_{\text{VDW}} = \sum_{\text{nonbonded pairs}} 4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r}\right)^{12} - \left(\frac{\sigma_{ij}}{r}\right)^{6}\right]$$

where b is the bond length, θ is the bond angle, χ is the improper torsion angle, ϕ is the torsion angle, the K values are the respective force constants, and the subscript o indicates an equilibrium value. The parameter n in the torsion potential is the multiplicity, and δ is the phase. The van der Waals potential contains ϵ_{ij} (the well depth), σ_{ij} (the Lennard-Jones diameter), and r (the nonbonded distance). The difference between force fields enters through the values used for the parameters in these energy terms. In particular, the coefficients used in the torsion potential differ strongly in these force fields. Previous work catalogues these differences and discusses them in detail (*21*).

The solvation potential $U_{solvent}$ defines the implicit solvent model. The first term of this energy function is a screened Coulomb potential of the form

$$U_{charge}(\epsilon) = \sum_{i<j} \frac{q_i q_j}{\epsilon(r_{ij}) r_{ij}}$$

where $\epsilon(r_{ij})$ is a Ramstein−Lavery style distance-dependent dielectric "constant" (*32*) and the q_i values are atomic partial charges. This distance-dependent dielectric constant is given by

$$\epsilon(r_{ij}) = D_\infty - \frac{D_\infty - D_o}{2}(S^2 r_{ij}^{\,2} + 2S r_{ij} + 2)e^{-S r_{ij}}$$

where D_∞ is the bulk dielectric constant of water, D_o is the limit at small distances, and S is a sigmoidal parameter which can be adjusted to optimize agreement with experiment or explicit solvent simulations. We use the following parameters: $D_\infty = 78.5$, $D_o = 1$, and $S = 0.3$.

The second term of the solvation model uses the Ooi−Scheraga solvent-accessible surface area (SASA) potential (*33*), which is of the form

$$U_{solvation}(\sigma) = \sum_{i=1}^{N} g_i \sigma_i$$

where σ_i is the accessible surface area of a hypersurface bisecting the first solvent shell surrounding protein atom i and g_i is an empirical (free energy) parameter dependent on atom type.

All implicit solvent simulations fix the temperature at 300 K using the random force, and the simulation is performed using a modified, speed-enhanced (*34*) TINKER molecular mechanics package (http://dasher.wustl.edu/tinker/). Integration proceeds using the Verlet algorithm (*35*) with an integration step of 2 fs and coordinates retained every 0.1 ps. The lengths of all bonds involving hydrogen are frozen using the RATTLE algorithm (*36*), and nonbonded forces are computed using the FAST-LD routine (*37*). The frictional forces and random forces are computed using the Pastor−Karplus accessible surface area model (*38*) with the experimental solvent viscosity of 0.89 cP. Finally, accessible surface areas, atomic friction coefficients, and solvation potentials are updated every 0.2 ps. Unless otherwise noted, all implicit solvent simulations are 30 ns in duration.

A 30 ns explicit solvent simulation for Ala_7 uses conventional nonbonded interactions, the TIP3P model for water (*39*), and the CHARMM32 (*40*) package with the CHARMM27 force field (*30*). The protein is solvated with a box of TIP3P molecules with a length of 40 Å per side and with periodic boundary conditions. A nonbonded cutoff of 12 Å is imposed with the force shift technique beginning at 8 Å (*41*). The protein is maintained at the center of the solvation box by fixing the coordinate of the α-carbon on the fourth amino acid, and the lengths of all bonds involving hydrogen are held constant using the SHAKE algorithm (*42*). Integration proceeds with the Verlet algorithm. The integration step is 1 fs, and coordinates are saved every 0.1 ps. The temperature is maintained at 300 K by rescaling velocities every 100 fs.

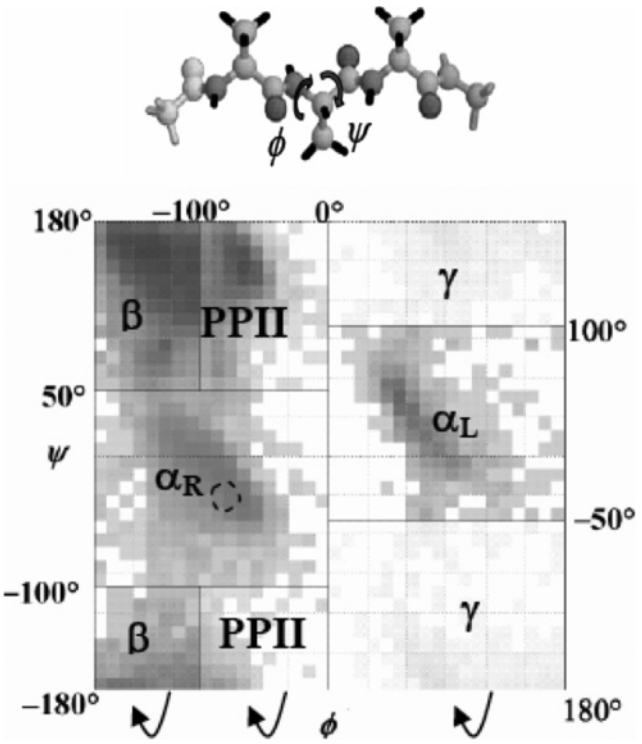

FIGURE 1: Conventional Ramachandran map.

A 30 ns gas phase MD simulation of Ala_7 with all nonbonded interactions retained is performed using a modified version of the TINKER package by removing all solvent effects, i.e., by setting the dielectric constant to 1 and by dropping the SASA term as well as the viscous drag and random forces. The temperature is again fixed at 300 K using the Berendsen thermostat with velocities rescaled every 100 fs. The integration step is 2 fs, and coordinates are retained every 0.1 ps.

Implicit solvent simulations have also been run using several different choices for the treatment of nonbonded interactions to probe the range over which correlated backbone motions may persist. The notation Nm indicates that nonbonded interactions of the ith residue are computed only using residues in the range from $i - m$ to $i + m$. Simulations are performed using N1−N4 and N∞ nonbonded interaction schemes for the Ala_7 system, where the notation N∞ corresponds to the full retention of nonbonded interactions. Simulations of the heteropolymer (EDEVARLKKLLW) are performed using both the N2 and N∞ nonbonded interaction schemes. The simulations of homopolymers Pro_7 and Gly_7 are simulated only using the N2 nonbonded interaction scheme.

The longest simulation of 200 ns for Ala_7 uses the GS-AMBER94 force field, our implicit solvent model, and the N2 nonbonded interaction scheme. This simulation is used in the discussion unless otherwise specified because it provides the best statistics, especially for basin hopping transitions. The other simulations provide extensive checks that demonstrate the consistency of the results deduced from this primary simulation.

Basin Populations and Transitions. The backbone dihedral angles Φ and Ψ preferentially adopt certain ranges of values generically termed Ramachandran basins. Often, the Ramachandran basins are defined such that five basins completely fill the plane (Figure 1). However, these basin areas are larger

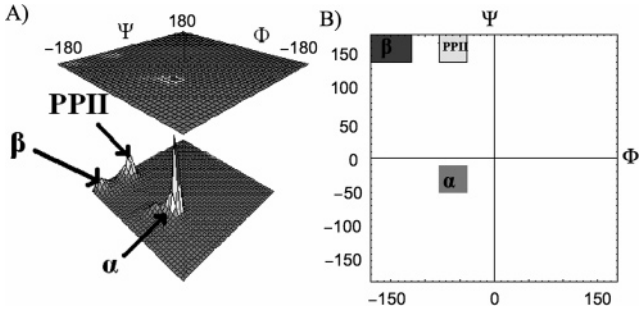

FIGURE 2: (A) Ramachandran populations obtained from a 2 ns trajectory from an explicit solvent simulation of Ala_7 using the CHARMM27 force field and TIP3P explicit solvent. This illustrative trajectory is chosen because all three basins are well sampled, although the populations from this portion do not correspond to equilibrium basin populations. The minima obtained from this distribution are used to specify the core of each basin. The populations are averaged over all residues. (B) Basin definitions used to discuss transition rates. The cores of the β, PPII, and α basins are labeled.

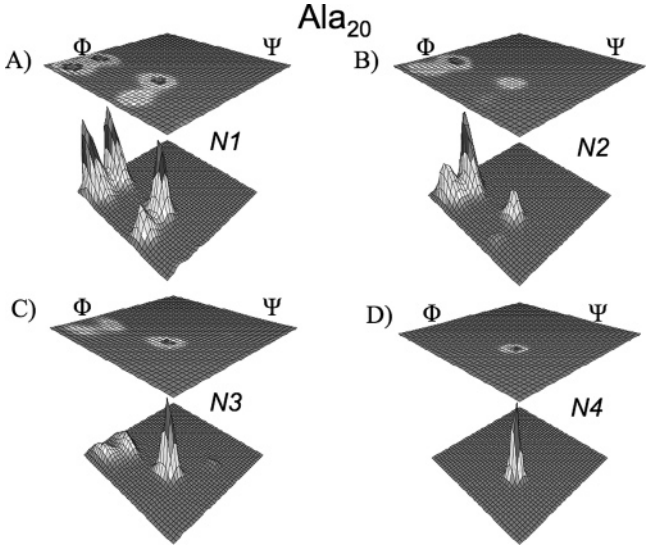

FIGURE 3: Probability distribution of the Ramachandran map for Ala_{20} obtained using the implicit solvent model with the GS-AMBER94 force field: (A) N1, (B) N2, (C) N3, and (D) N4 nonbonded interaction schemes.

than those we use when considering basin transitions. In the latter case, only the core of each basin is used to specify the basin to ensure the examination of authentic basin-to-basin transitions instead of fluctuations near the boundary between basins. To identify these core regions, Figure 2A displays the Ramachandran basin populations observed in a 2 ns segment of the explicit solvent trajectory. This explicit solvent simulation (which includes all nonbonded interactions) indicates the presence of three well-defined and isolated minima that correspond to the α, β, and PPII basins. The core regions of the basins are defined by choosing boundaries so that at least 90% of the population is included in one of the basin cores, while the separation between basin cores is large compared to fluctuations within a basin core (Figure 2B). The basin core definitions are obtained using a short segment of the trajectory because after a sufficiently long time, Ala_7 becomes helical, and the α-basin is dominantly populated. Hence, using the entire trajectory obscures the minima of the β and PPII basins. Consequently, the particular 2 ns segment is chosen as one in which all three basins are well-populated.

We now define the basin transitions. If a residue is in basin A at time t_0, enters basin B at time t_1, and does not traverse any other basin between these two times, the residue is said to have undergone a transition from A to B at time t_1.

RESULTS AND DISCUSSION

Influence of Nonbonded Interactions on Ramachandran Populations. Nonbonded interactions strongly influence peptide stability and kinetics. This paper is concerned with local correlations that occur due to constraints imposed by the local protein geometry. Conclusions based on simulations that include all nonbonded interactions pose uncertainty about whether the observed correlations arise from local or nonlocal effects. Therefore, separate simulations are performed using a full treatment of nonbonded interactions and using a reduced treatment that is designed to focus on backbone motions in the absence of long-range interactions. The reduced treatment uses a modified potential that incorporates only the sequence-local portion of the nonbonded interactions. This type of reduced treatment is well-known in polymer physics to correspond to describing the ideal θ state in which the polymer−solvent interactions effectively offset the nonbonded interactions (*43*−*45*) so that only the short-range interactions along the chain persist. The θ state is thus useful for focusing on local structure and dynamics. Hence, our use of the reduced treatments is in the spirit of studying properties under θ conditions.

The notation N*m* signifies that the reduced description ignores nonbonded interactions for residues that are more distant than $\pm m$ residues in the primary sequence. Simulations using N1, N2, or N3 interactions inhibit secondary structure formation. Consequently, stable long-range structure in the peptide is absent, thereby enabling the peptide to undergo more frequent conformational transitions. In the simulations for Ala_7, for example, the peptide rapidly adopts a helical conformation when all nonbonded interactions are included, but it remains a statistical coil for the N1−N3 models. Figure 3 displays the populations in the form of Ramachandran maps from implicit solvent simulations for Ala_{20} using the N1−N4 interaction schemes with the GS-AMBER94 force field.

Despite differences in Ramachandran populations, the same correlated local motions are found for the N1−N∞ simulations (i.e., neighbor-only through all residue nonbonded interaction models). Because we find the same correlated motions independent of the manner in which nonbonded interactions are treated, of the amino acids, of the solvent model, and of the force field, this section illustrates the general behavior of the correlated backbone motions from N2 simulations for which the longest trajectories and hence the best statistics are available. This N2 nonbonded interaction routine maximally reduces the influence of the tertiary nonbonded interactions because the choice allows the largest degree of conformational freedom while still maintaining the three strongly populated basins. Thus, all results from the N2 simulations are a consequence of locally determined transitions and not of interference from the long-range interactions present in secondary structures. This choice is consistent with our focus on fundamental

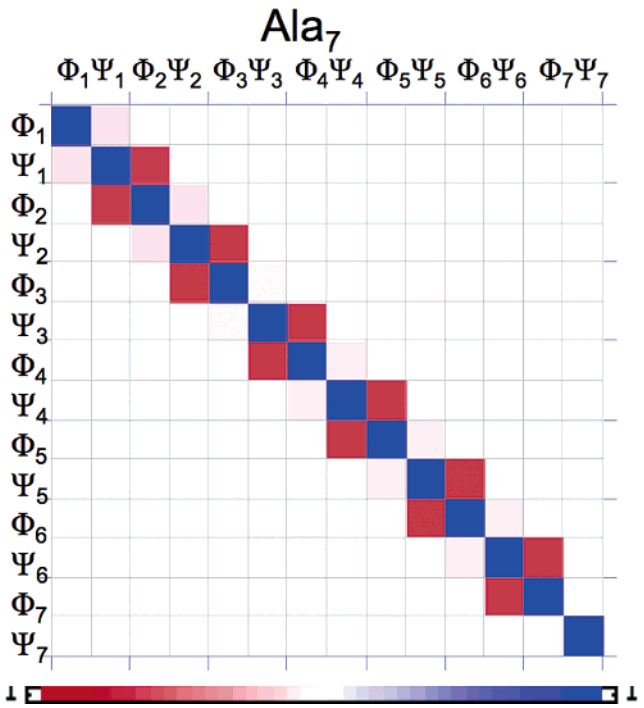

FIGURE 4: Graphical representation of the correlation coefficients for movements of the torsion angles of Ala_7 using the implicit solvent model, the GS-AMBER94 force field, and N2 interactions. The motion of a torsion angle is quantified as the change the torsion angle undergoes in 0.1 ps. The *x*- and *y*-axes represent torsion angles (i.e., 1 for Φ_1, 2 for Ψ_1, 3 for Φ_2, etc.). Red indicates anticorrelation, and blue designates positive correlation. All statistically insignificant correlation coefficients have been set to zero.

backbone motions. The reduced nonbonded interaction routines are compared with a full treatment of nonbonded interactions in a figure in the Appendix, where the results for this full treatment are summarized Table 4. This demonstrates that the conclusions concerning correlated motions are the same independent of nonbonded interaction model. The Appendix also contains a comparison of the results obtained from different force fields and solvent models.

Description of Observed Motions in Ala7. Correlated motions between backbone torsion angles are analyzed by computing the correlation coefficient $\sigma(\Delta\alpha,\Delta\beta)$

$$\sigma(\Delta\alpha,\Delta\beta) = \frac{\langle(\Delta\alpha)(\Delta\beta)\rangle}{\sqrt{\langle(\Delta\alpha)^2\rangle\langle(\Delta\beta)^2\rangle}}$$

where α and β are dihedral angles (of the same or different residues) and $\Delta\alpha$ denotes the change in angle α between consecutive frames in the trajectory. Unless otherwise noted, the time between consecutive frames in the trajectory is 0.1 ps. Some calculations consider the change in torsion angles over the finer time scale of 1 fs to probe the degree to which the crankshaft motion is concerted. Figure 4 displays the correlation coefficients from the simulation of Ala_7 with implicit solvent, the GS-AMBER94 force field, and N2 interactions for each pair of torsion angles. All correlations not on the supradiagonal (matrix elements $A_{i,i-1}$), diagonal (matrix elements $A_{i,i}$), or superdiagonal (matrix elements $A_{i,i+1}$) are statistically insignificant. Hence, long-range correlations are absent. Because the only relevant correlations are motions between neighboring torsion angles for residues

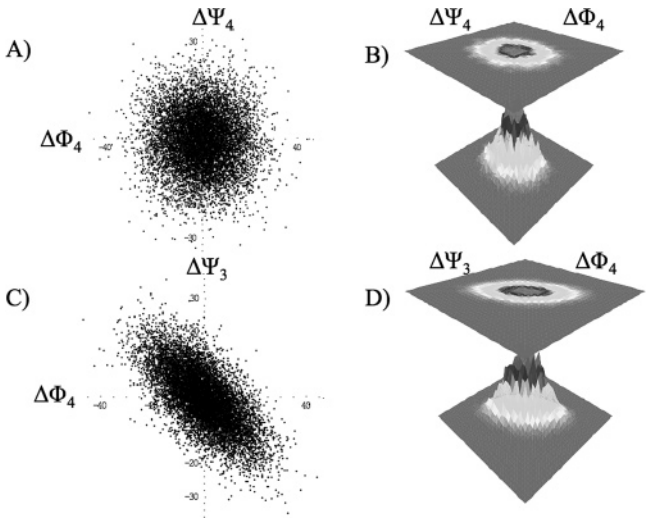

FIGURE 5: Change in the Φ_4 and Ψ_4 angles of the fourth residue of Ala_7 shown through (A) a scatter plot and (B) a probability distribution. Panels C and D are the same, respectively, except for the change in Φ_4 and Ψ_3. Changes in this figure are calculated over 0.1 ps.

j and *j* − 1, these correlations lie on the superdiagonal of the matrix. The diagonal of the matrices always presents a perfect correlation since the two angles examined are identical. Also, note that the matrix is symmetric by construction.

We begin by examining the correlations between the Φ and Ψ angles of the same residue. The intraresidue correlations and inter-residue correlations occur in an alternating pattern along the superdiagonal of this matrix. The intraresidue correlations are seen to be quite weak and in many cases are vanishingly small. More explicitly, the correlation coefficients for neighboring torsion angles of the same residue lie in the range of −0.11 to −0.05. For the fourth residue of Ala_7, the correlation coefficient between $\Delta\Phi_4$ and $\Delta\Psi_4$ is −0.10 (Figure 5A,B). There is thus a very slight anticorrelation of motions between $\Delta\Phi_4$ and $\Delta\Psi_4$ which can be discerned by a slight clustering of the data around a negative diagonal. However, this correlation is very small and is essentially zero for some residues. An error estimate for the correlation coefficient $\sigma(\Delta\Phi_i,\Delta\Psi_i)$ is obtained by decomposing the last 5 ns of the trajectory into 1 ns windows. The intraresidue correlation coefficient is calculated for each window and ranges from −0.05 to 0.06. Therefore, those correlations with absolute magnitudes of <0.05 are deemed to be statistically insignificant.

We next examine the relationship between the Φ angle of the *i*th residue and the Ψ angle of residue $i - 1$, which exhibits a well-defined anticorrelation. This neighbor correlation barely varies along the peptide chain, with a range of only −0.68 to −0.67. Panels C and D of Figure 5 depict this anticorrelation by plotting $\Delta\Phi_4$ versus $\Delta\Psi_3$. The dominance of this anticorrelation is shown below to be quite general in peptide dynamics. An error analysis demonstrates that the neighbor correlations from separate windows vary only between −0.68 and −0.67, and consequently, the error is negligible compared to the magnitude of the correlation.

The observed correlations involve localized cooperative motions that eliminate long-range chain displacements and that represent a type of crankshaft motion. These correlations are readily understood physically as oscillations of the plane

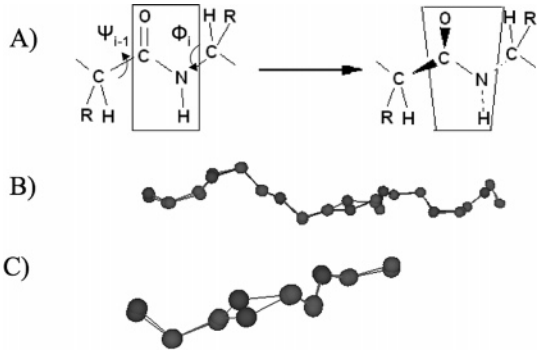

FIGURE 6: (A) Conformational transitions are accomplished by a strong anticorrelation of the motions of Φ_i and Ψ_{i-1} that corresponds to an oscillation of the peptide plane and that largely leaves all other atoms unchanged in space. Superimposed structures of Ala_7 before and after a large change (64.4°) in the Φ angle of the fourth residue are depicted for (B) the entire chain or (C) residues 3−5. Notice that the chains nearly perfectly superimpose, except for the atoms involved in the rotation of Ψ_3, Φ_4, or Ψ_4.

of the peptide group (Figure 6A). Because the peptide group acts as a rigid structure and because the displacement of both Φ_i and Ψ_{i-1} involves atoms C_{i-1} and N_i, which are elements of the peptide plane, the motions of this rigid object necessitate an anticorrelation between $\Delta\Phi_i$ and $\Delta\Psi_{i-1}$. To further illustrate this mode of motion, panels B and C of Figure 6 present two snapshots from consecutive frames between which the largest change in Φ_4 is observed (64.4°). The two structures are superimposed to minimize the rmsd. The minimized rmsd is 0.07 Å, indicative of an exceedingly small difference between the two structures that is appreciable only near Φ_4 and Ψ_3 and to a lesser extent Ψ_4. The rest of the chain is altered only by small equilibrium fluctuations.

To further quantify these torsion angle correlations, we introduce a measure called threshold plots. These plots are created by culling all frames from the trajectory that satisfy the condition $|\Delta\beta| \geq x$, where β is a torsion angle and x is a threshold value. For each threshold value, the average absolute change in a second torsion angle α is computed and is defined as the dependent variable y. Threshold plots have the advantage of displaying the scale of the changes, unlike the normalized correlation coefficient which describes only the extent to which two angles move in unison. A threshold plot is of the form

$$y(x) = \langle \{ |\Delta\alpha|; |\Delta\beta| \geq x \} \rangle$$

where α and β are torsion angles of the protein backbone and the average is over the culled trajectory. Figure 7 presents threshold plots for several choices of angle α, where β is fixed to be one of the torsion angles of the fourth residue of Ala_7. Data are taken from the simulation of Ala_7 using the implicit solvent model with the GS-AMBER94 force field and N2 interactions. In graphs of this form, a vanishing slope indicates that motions in α and β are independent since when the change in β is large, the change in α does not depart from its average fluctuations (the y-intercept of threshold plots is exactly the equilibrium average of the dynamic variable). On the other hand, a positive slope is indicative of either a correlation or an anticorrelation because when the magnitude of the change in β is large, the positive slope indicates that the change in α likewise tends to be large. This positive slope, in turn, means that the α and β torsion angles move cooperatively, and when one deviates from equilibrium fluctuations, the other tends to deviate from equilibrium fluctuations as well. A negative slope implies that when one torsion angle undergoes large motions, the other becomes more rigid and moves less than it would during equilibrium fluctuations within a basin. This unphysical behavior is not observed.

Figure 7 depicts the presence of a strong cooperativity between changes in Φ_4 and Ψ_3 and a much weaker cooperativity between Φ_4 and Ψ_4. Furthermore, there is an absence of appreciable correlation at longer distances, since all the other curves relating Φ_4 and Ψ_4 to the other dihedral angles along the chain have a vanishing slope. The threshold plots also provide a quantitative understanding of the degree of movement expected from a torsion angle when its neighbor moves by a particular amount. Threshold plots have been computed and examined for all torsion angles in the peptide, but the other cases are not displayed because they all behave identically.

While the change in torsion angles between two consecutive frames exhibits correlations only for nearest neighbors, there still remains the possibility that a finite propagation time exists such that there may be a correlation between $\Delta\alpha(t)$ and $\Delta\beta(t + dt)$, where dt is some propagation time. When dt is set to be the time step between consecutive frames (0.1 ps), again no statistically significant correlation is evident between distant torsion angles, and the correlated motions between neighboring torsion angles have already relaxed in the sense that the strongest correlations at $t + dt$ are weaker than the correlation between Φ_i and Ψ_i at time t. Thus, the correlation time of these motions must be less than 0.1 ps.

Furthermore, after a single time step of 0.1 ps, the single-angle correlations $\sigma[\Delta\Phi_i(t),\Delta\Phi_i(t + dt)]$ and $\sigma[\Delta\Psi_i(t), \Delta\Psi_i(t + dt)]$ are small, with their correlations being -0.27 and -0.22, respectively, while after two time steps, the single-angle correlations $\sigma[\Delta\Phi_i(t),\Delta\Phi_i(t + 2dt)]$ and $\sigma[\Delta\Psi_i(t), \Delta\Psi_i(t + 2dt)]$ are very close to being statistically insignificant, with both correlations being -0.06. Recall that by definition a stochastic process $x(t)$ is said to be Markovian if

$$P[x(t_n)|x(t_0), ..., x(t_{n-1})] = P[x(t_n)|x(t_{n-1})]$$

This equation simply states that the process at time t_n is independent of the process at times shorter than t_{n-1}. Considering the dynamics of dihedral angles as stochastic processes, the motion of the protein backbone is essentially a Markovian process on a 0.1 ps time scale. Since correlations $\sigma[\Delta\Phi_i(t),\Delta\Phi_i(t + dt)]$ and $\sigma[\Delta\Psi_i(t),\Delta\Psi_i(t + dt)]$ are also small in the sense that any torsion angle is more correlated with its neighbor (at the given time) than with itself (in the previous frame), to a good approximation, the motion of a protein backbone is Markovian on a time scale of less than 0.1 ps. The time scales of most studies are long compared to 0.1 ps, and the Markovian assumption for torsion dynamics is justified in these studies. The correlation coefficients reported here are obtained from the implicit solvent trajectory. While the statistics are poorer, the explicit solvent simulation, which uses a conventional nonbonded cutoff of 12 Å, produces slightly higher correlations, and a few tenths of a picosecond may be necessary for an explicit solvent trajectory to appear Markovian. But again this is

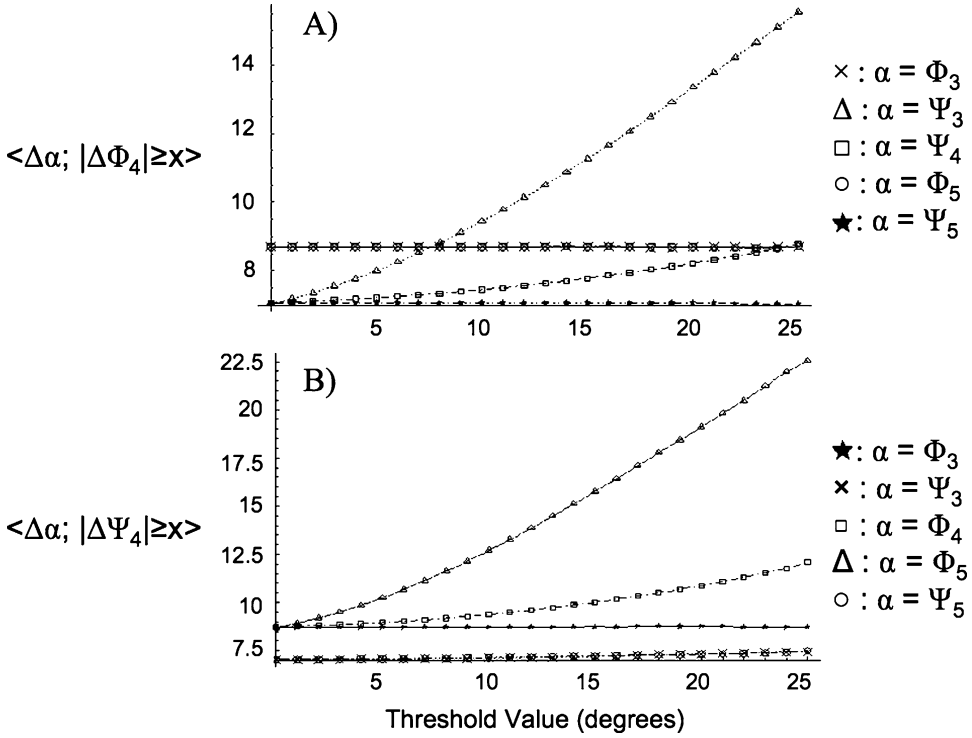

FIGURE 7: Threshold plots for several different torsion angles of Ala_7 as a function of the motions of the fourth residue of Ala_7. Data are taken from the implicit solvent simulation using the GS-AMBER94 force field and N2 interactions. Changes are recorded over 0.1 ps. (A) Variance in the Φ angle of the fourth residue. (B) Variance in the Ψ angle of the fourth residue.

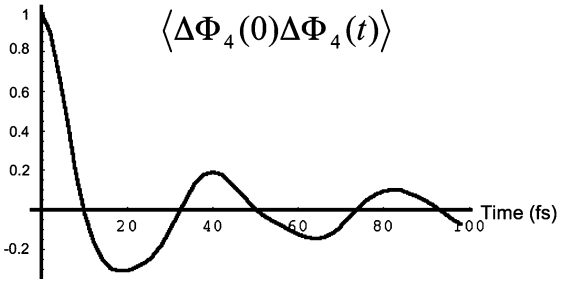

FIGURE 8: Finely binned normalized autocorrelation function, $\langle \Delta\Phi_4(0)\Delta\Phi_4(t)\rangle$, obtained from a 1 ns simulation using the GS-AMBER94 force field, our implicit solvent model, and N2 interactions. Data are binned with a bin size of 2 fs, and the change in a torsion angle is defined as the change observed over 2 fs windows.

shorter than time scales of general interest. This finding is of great practical interest for methods that attempt to determine protein dynamics with kinetic Monte Carlo methods. Also, the observed correlations provide a physically realistic move set for use in these methods.

To better ascertain the Markovian nature of these trajectories, a 1 ns trajectory is performed (according the protocol used above) during which coordinates are retained every 2 fs. Autocorrelation functions are computed for the changes in torsion angles over 2 fs windows. The autocorrelation functions decay to zero within 10 fs and subsequently exhibit damped oscillation, indicating negative and positive correlations (Figure 8). Cross-correlation functions for correlated torsion angles are also found to exhibit a similar oscillatory character. All correlations are zero by 0.1 ps, consistent with the results determined using the coarser binning.

To summarize this section, we have established that the dynamics of a torsion angle affect only the nearest torsion angle on each side (excluding ω which is essentially constant). This correlated motion represents a rocking of the rigid plane of the peptide group in which flanking residues are fixed, as their movement is unaltered from normal fluctuations. Furthermore, the motions are Markovian on time scales of interest (>0.1 ps), and the change in any torsion angle is more dependent on the motion of its neighboring torsion angle than on its own previous motions.

Basin Transitions. So far, we have focused on equilibrium averages obtained over the entire trajectory. These averages are dominated by the populations near the bottom of the Ramachandran basins. It is interesting to determine if the correlated motions persist during the transition from one Ramachandran basin to another. We therefore compare the dynamics of equilibrium fluctuations to those of basin transitions. The analysis of this section is performed using the 200 ns simulation of Ala_7 which uses the implicit solvent, the N2 nonbonded interactions, and the GS-AMBER94 force field because transitions involving the α basin are quite rare and the other simulations are too short for properly sampling transitions involving the α basin. Furthermore, the N2 interactions of this simulation prevent helix formation and consequently more readily allow transitions out of the very stable α basin. Because the fundamental motions present in this simulation are shown to be robust to choice of force field, solvent models, and nonbonded interaction routines (see the Appendix and Tables 4 and 5), no qualitative differences are expected between these results and those which could be obtained from other simulations.

It is instructive to begin by providing a brief description of the manner in which basin transitions occur. As described in Methods, if residue n is in basin A at time t_0 and the residue enters basin B at time t_1 (without traversing any other basin), the residue is said to have undergone a basin transition from A to B at time t_1. Two characteristic time distributions

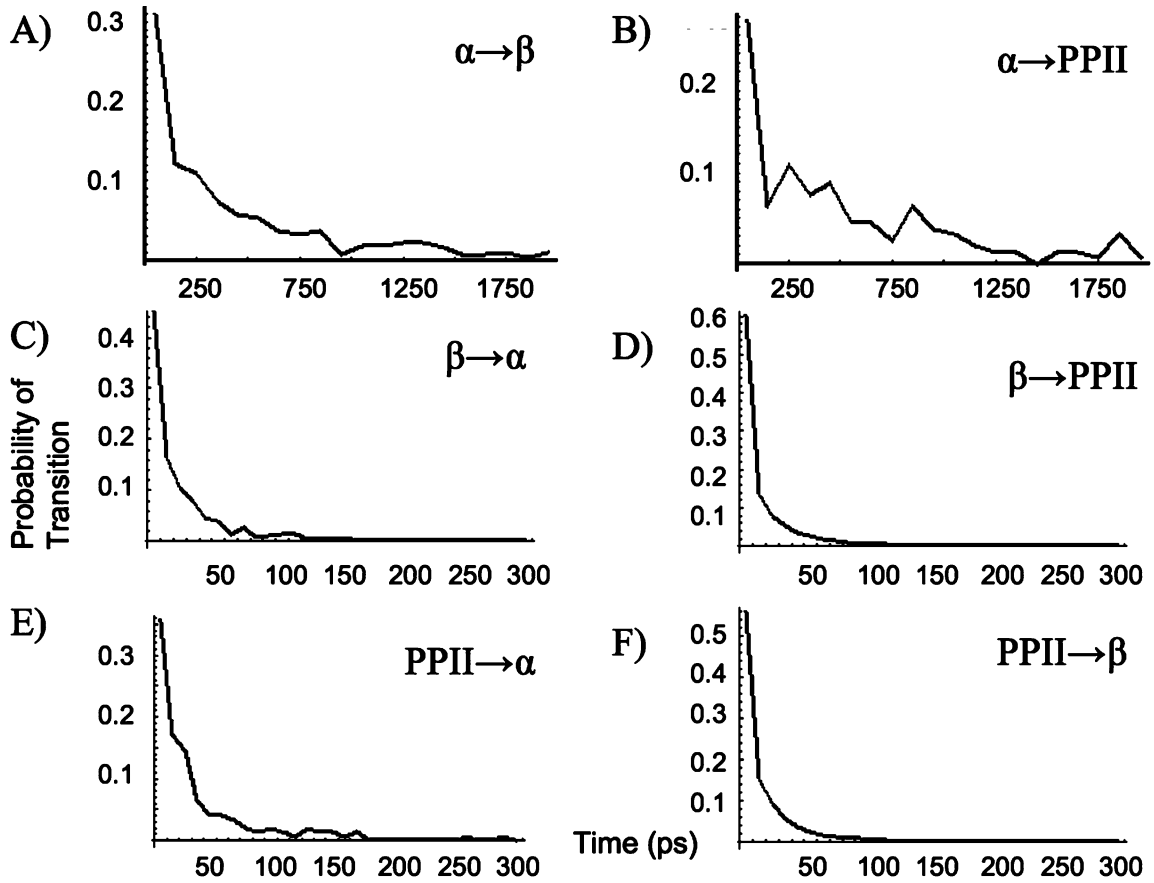

FIGURE 9: Waiting time distributions obtained from the 200 ns simulation of Ala_7 using the implicit solvent model, GS-AMBER94 force field, and N2 interaction scheme: (A) α → *β* (380 transitions), (B) α → PPII (158 transitions), (C) *β* → α (389 transitions), (D) *β* → PPII (25 320 transitions), (E) PPII → α (152 transitions), and (F) PPII → *β* (25 325 transitions).

are used to describe the time scales associated with basin transitions.

We first consider the "waiting time" that an amino acid stays in a particular well before transitioning to another well. Figure 9 presents the waiting time distributions for all six types of transitions (α ↔ *β*, α ↔ PPII, and *β* ↔ PPII). Notice that the transitions originating in the α basin wait significantly longer before they eventually undergo a transition. Notice also that transitions involving the α basin are remarkably rare compared to the *β* ↔ PPII transitions. Both these facts are reflections of the stability of the α basin in Ala_7 and the height of the barriers separating the basins. This difference explains the increased noise in the waiting time distributions associated with α basin transitions.

Second, we consider the time that elapses between the time when the residue leaves the first well and the time when the residue arrives in the second well, termed the passage time. Figure 10 displays the passage time distributions for all six types of transitions (α ↔ *β*, α ↔ PPII, and *β* ↔ PPII). Transitions involving the α basin tend to have one characteristic passage time distribution, whereas *β* ↔ PPII transitions have shorter characteristic passage times, with the vast majority being <5 ps and essentially no transitions having a passage time of >15 ps. Contrastingly, α transitions have a much broader distribution of passage times, and transitions with passage times of greater that 50 ps constitute roughly 15% of the transitions. This result is somewhat surprising because exponential basin hopping rates are normally associated with a scenario in which the peptide moves through the high-energy region of the Ramachandran map in a rapid, barrier crossing mode.

The physical origin of the long passage times associated with α transitions is linked to the fact that during basin transitions, the motions of the peptide are the same as the motions within a Ramachandran basin. Figure 11 presents the correlation coefficient matrices computed over subsets of the trajectory during which a particular basin transition is underway. The first thing to note is that once again the motions are highly local with the dominant anticorrelation being in the form $\sigma(\Delta\Phi_i,\Delta\Psi_{i-1})$. During some transitions in some amino acids, there is a very small degree of cooperativity at longer ranges that cannot be dismissed as being statistically insignificant. However, in all cases, these correlations are still exceedingly small ($|\sigma| \sim 0.05-0.06$). Basin transitions can, and most often do, occur without any correlations beyond an adjacent torsion angle. The correlations during basin transitions are essentially the same as those for motions near the bottom of the wells. The average change of a torsion angle is also the same during basin transitions as it is in equilibrium fluctuations.

We next examine the extent to which basin transitions are Markovian. Recall that during equilibrium fluctuations, the same angle correlation coefficients $\sigma[\Delta\Phi_i(t),\Delta\Phi_i(t + dt)]$ and $\sigma[\Delta\Psi_i(t),\Delta\Psi_i(t + dt)]$ are found to be −0.27 and −0.22, respectively, when $dt = 0.1$ ps. When these correlations are recomputed by restricting the averaging to the subset of the trajectory consisting of basin transitions, only very minor differences emerge from the equilibrium case (Table 1).

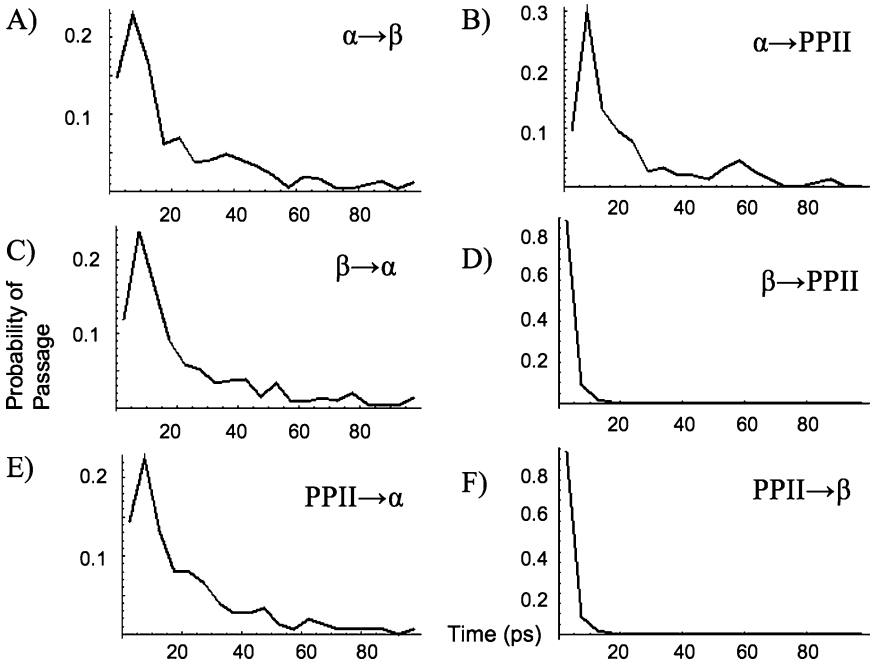

FIGURE 10: Passage time distributions obtained from the 200 ns simulation of Ala_7 using the implicit solvent model, the GS-AMBER94 force field, and the N2 interaction scheme: (A) α → *β* (380 transitions), (B) α → PPII (158 transitions), (C) *β* → α (389 transitions), (D) *β* → PPII (25 320 transitions), (E) PPII → α (152 transitions), and (F) PPII → *β* (25 325 transitions).

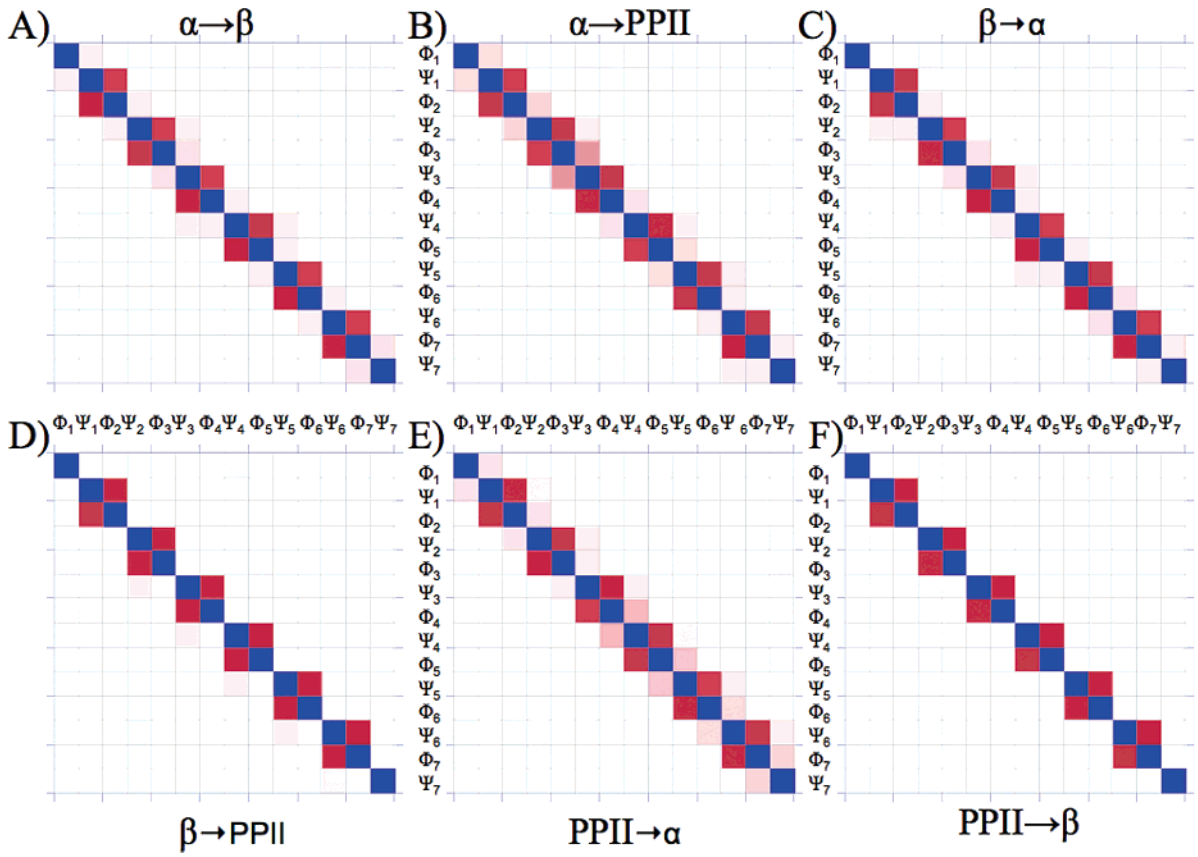

FIGURE 11: Graphical representation of the correlation coefficients for movements of the torsion angles (during basin transitions) of Ala_7 using the implicit solvent model, the GS-AMBER94 force field, and N2 interactions. Red indicates anticorrelation, and blue designates correlation. The *x* and *y*-axes represent torsion angles (i.e., 1 for Φ_1, 2 for Ψ_1, 3 for Φ_2, etc.). The ticks on the axes separate one residue from another, and the grid on the plot separates torsion angles: (A) α → *β* transitions, (B) α → PPII transitions, (C) *β* → α transitions, (D) *β* → PPII transitions, (E) PPII → α transitions, and (F) PPII → *β* transitions. Statistically insignificant correlations have been set equal to zero.

The previous two paragraphs provide an interesting characterization of the dynamics of basin transitions in peptides. Since during basin transitions, the motion of any torsion angle is shown to be largely independent of its previous dynamics, we find that basin transitions do not proceed through a few large motions, but instead through a

Table 1: Autocorrelation Coefficients for Ala_7[a]

ensemble	$\sigma[\Delta\Phi_4(t),\Delta\Phi_4(t+dt)]$	$\sigma[\Delta\Psi_4(t),\Delta\Psi_4(t+dt)]$
entire trajectory	−0.27	−0.22
α → β transitions	−0.22	−0.19
α → PPII transitions	−0.22	−0.20
β → α transitions	−0.21	−0.20
β → PPII transitions	−0.19	−0.20
PPII → α transitions	−0.30	−0.21
PPII → β transitions	−0.21	−0.21

[a] This simulation uses the GS-AMBER94 force field, our implicit solvent model, and the N2 interaction routine. The simulation is 200 ns in duration, and in this table, $dt = 0.1$ ps.

series of many small uncorrelated changes. Also, there is no long-range cooperativity or inertial effects. The peptide crosses the high-energy barrier via small, diffusive-like rocking motions, analogous to the barrier crossings envisioned by Kramers' theory (*46*). In particular, because of this diffusive barrier crossing, the distance between the basins in torsion space correlates with barrier height and largely characterizes the passage time, while the well depth and breadth largely determine the waiting time distribution. Furthermore, it is impossible to tell which residues in a peptide are in energetically unfavorable conformations by examining the peptide dynamics alone. With this context, we can understand the long passage times associated with α basin transitions.

Importantly, we find no evidence for coupling of basin transitions between neighboring residues. From our previous discussion, the only expected couplings of basin transitions are those that involve a change in the Φ angle of the *i*th residue and an opposing change in the Ψ angle of residue $i - 1$. This anticorrelation implies coupled transitions would be observed only when either the *i*th residue undergoes a transition from the β to the PPII basin while residue $i - 1$ undergoes a transition from the β or PPII basin to the α basin or the *i*th residue undergoes a transition from the PPII to β basin while residue $i - 1$ undergoes a transition from the α basin to either the β or PPII basin. However, by examining Figures 9 and 10, we see that the time scales associated with these types of transitions are vastly different. By the time an α-related transition has been completed, a multitude of β ↔ PPII transitions will have occurred. Effectively, basin transitions on one residue are independent of transitions on other residues, a result that is pertinent to devising kinetic Monte Carlo schemes for protein dynamics.

This lack of coupling between basin transitions also is indicated by considering their time scales. The motions of the proteins relax on the order of 0.1 ps, while basin transitions can take tens of picoseconds to occur. Hence, a coupling of basin transitions is not necessary. A neighboring residue has adequate time to relax while its neighbor is passing from one basin to another. To further investigate this matter, we have examined the probability of a residue undergoing a transition as a function of how long before a given neighbor transitioned. In most cases, statistics are insufficient to draw rigorous deductions, but in all cases, there is no suggestion of coupled transitions (i.e., many transitions are observed that do not couple with any other transition). Therefore, we conclude that the backbone dynamics during basin transitions are very similar to those of fluctuations within a Ramachandran basin. All anticorrelations are on the same order of magnitude, and the processes are Markovian to the same extent as the equilibrium fluctuations. No evidence appears for coupled basin transitions of neighboring residues.

Table 2: Correlation Coefficients for the Heteropeptide EDEVARLKKLLW[a]

residue	$\sigma(\Delta\Phi_X,\Delta\Psi_X)$	$\sigma(\Delta\Phi_X,\Delta\Psi_{X-1})$
Glu	−0.01	not applicable
Asp	−0.25	−0.66
Glu	−0.06	−0.70
Val	0.02	−0.73
Ala	−0.14	−0.72
Arg	−0.08	−0.70
Leu	−0.03	−0.71
Lys	−0.13	−0.72
Lys	−0.03	−0.71
Leu	−0.04	−0.72
Leu	−0.09	−0.73
Trp	−0.09	−0.71

[a] This simulation uses the GS-AMBER94 force field, our implicit solvent model, and the N2 interaction routine. The simulation is 50 ns in duration.

Correlated Motions in Different Force Fields, Solvent Models, and Systems. The dynamic behavior is further investigated by varying the identity of the amino acid, the length of the peptide, the force field, the solvent model, and the nonbonded interaction scheme to ensure that the observed motions are fundamental to peptides rather than an artifact created by the conditions of a particular simulation. The simulations with various different force fields, nonbonded interaction routines, and solvent models all agree well with the results already described. Thus, the detailed discussion of the influence of force field and solvent model is included as the Appendix. No comparison can be made concerning basin transitions since the other simulations are not long enough for observation of sufficient sampling of transitions involving the α basin.

Length Dependence. Because the correlations between Ψ_{i-1} and Φ_i are found with N2 local interactions, no difference is expected due to the length of the peptide. Simulations of Ala_7 and Ala_{20} for 200 and 50 ns, respectively, enable the testing of this expectation. Both simulations use an implicit solvent model with the GS-AMBER94 force field and the N2 nonbonded interaction scheme. When the middle residues in each system are compared, the Ala_7 simulation yields $\sigma(\Phi_4,\Psi_4) = -0.10$ and $\sigma(\Phi_4,\Psi_3) = -0.68$, whereas the Ala_{20} simulation produces the statistically equivalent $\sigma(\Phi_{10},\Psi_{10}) = -0.08$ and $\sigma(\Phi_{10},\Psi_9) = -0.68$, thereby confirming the expectation that the correlations are independent of the length of the peptide.

Sequence Dependence. We next examine if the correlated motions vary with the amino acid side chain by examining simulations of heteropolymer EDEVARLKKLLW (described in Methods) and of peptides Gly_7 and Pro_7. The latter two are chosen for comparison as the most atypical examples to serve as a test of generic behavior. Table 2 presents the correlation coefficients for the heteropeptide, and Figure 12 depicts the correlation coefficients as a function of residue number. As with the Ala_7 peptide, the dominant anticorrelated motions are between Φ_i and Ψ_{i-1}. In addition, the correlation coefficient for these motions is close to constant for each residue in the chain. The smaller correlation coefficient for

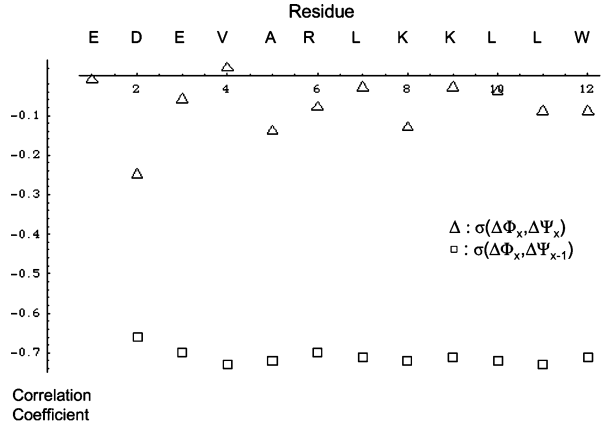

FIGURE 12: Variation of the correlation coefficients of the backbone dihedral angles along the EDEVARLKKLLW sequence. Despite the peptide's sequence heterogeneity, the correlation between Φ_i and Ψ_{i-1} motions always dominates and is reasonably constant across the backbone. The correlations between Φ_i and Ψ_i motions are more variable.

the Φ_i and Ψ_i motions is slightly more variable, but all intraresidue correlations or anticorrelations are very weak compared to the Φ_i and Ψ_{i-1} anticorrelation. The matrix plot of the correlation coefficients (Figure 1A of the Supporting Information) is identical to the one presented for Ala_7 in Figure 2 for all off-diagonal elements, while the diagonal elements are summarized in Figure 12. When this peptide is simulated with N∞ interactions, the diagonal elements are very similar with only a small amount of variation, but in some cases, there is a very small anticorrelation for which $|\sigma(\Delta\Psi_i,\Delta\Psi_{i\pm1})| \approx 0.06$. The statistical significance of this correlation is very minor. No other correlations exist (Figure 1B of the Supporting Information).

Proline and glycine often exhibit unusual behavior, and so polyglycine and polyproline are studied because they present the most likely exceptions to the previously discussed results. Table 3 contains the correlation coefficients obtained from the Gly_7 and Pro_7 simulations, and Figure 13 displays matrix plots of the correlation coefficients. In both Gly_7 and Pro_7, the Φ_i,Ψ_{i-1} anticorrelation ([−0.55,−0.56] and [−0.60,−0.61], respectively) is stronger than the Φ_i,Ψ_i anticorrelation ([−0.11,−0.15] and [−0.17,−0.22], respectively). However, the correlation coefficients for Φ_i and Ψ_{i-1} motions are ∼15% smaller than those for other residues. On the other hand, there is an up to 3-fold stronger anticorrelation between the motions of Φ_i and Ψ_i for the Gly_7 and Pro_7 peptides (compare Tables 2 and 3). Interestingly, these two homopolymers both exhibit weak correlations at longer distances. The extra statistically significant correlation for Gly_7 is of the form $(\Delta\Psi_i,\Delta\Psi_{i\pm1})$ and takes values in the range of −0.11 to −0.05, while for Pro_7, the extra correlation is of the form $(\Delta\Psi_i,\Delta\Psi_{i\pm2})$ with values in the range of −0.06 to −0.05. These correlations are statistically significant, but they have such small magnitudes that their origin and importance are unclear.

Table 3: Correlation Coefficients for Gly_7 and Pro_7[a]

	Gly_7		Pro_7	
residue	$\sigma(\Delta\Phi_X,\Delta\Psi_X)$	$\sigma(\Delta\Phi_X,\Delta\Psi_{X-1})$	$\sigma(\Delta\Phi_X,\Delta\Psi_X)$	$\sigma(\Delta\Phi_X,\Delta\Psi_{X-1})$
1	−0.14	not applicable	−0.19	not applicable
2	−0.15	−0.55	−0.17	−0.60
3	−0.14	−0.55	−0.18	−0.60
4	−0.11	−0.55	−0.18	−0.60
5	−0.15	−0.55	−0.20	−0.60
6	−0.13	−0.55	−0.19	−0.60
7	−0.11	−0.56	−0.22	−0.61

[a] This simulation uses the GS-AMBER94 force field, our implicit solvent model, and the N2 interaction routine. The simulation is 30 ns in duration.

The oscillatory motion of the peptide group observed in the heterogeneous, polyglycine, and polyproline peptides exhibits slight variations with amino acid type, but the dominant trends persist. Hence, the existence of these oscillations is largely independent of the amino acid identity, but the exact nature of the motion varies slightly with residue identity.

Implications for Computer Simulations. Knowledge of the fundamental local backbone motions has applications for

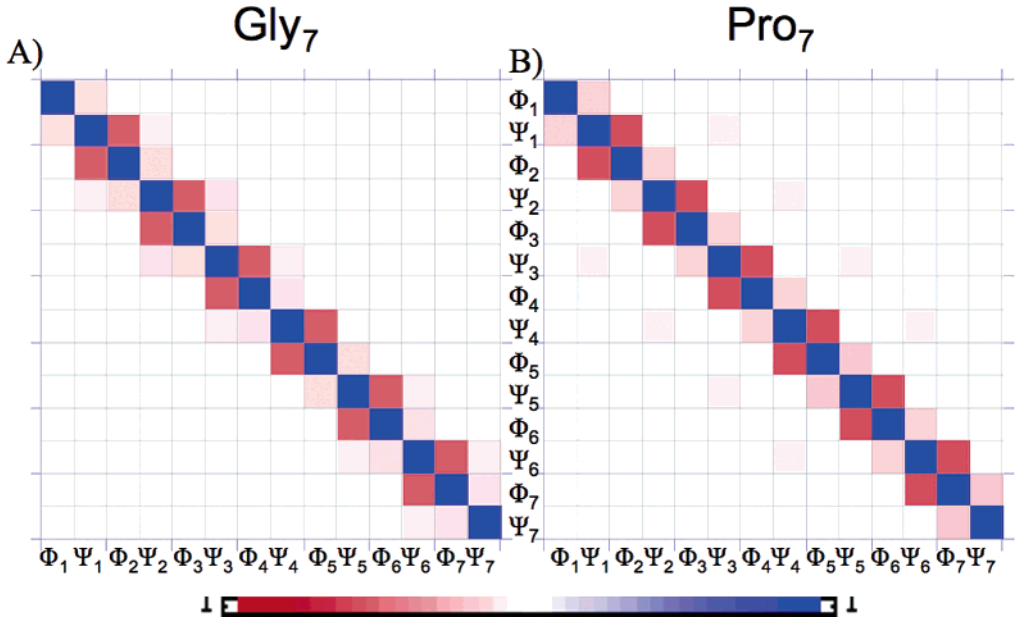

FIGURE 13: Correlation between dihedral angles for (A) Gly_7 and (B) Pro_7. Statistically insignificant correlations have been set equal to zero.

Table 4: Dependence of Correlation Coefficients on Force Field[a]

	GS-AMBER94		CHARMM27		OPLS	
residue	$\sigma(\Delta\Phi_X,\Delta\Psi_X)$	$\sigma(\Delta\Phi_X,\Delta\Psi_{X-1})$	$\sigma(\Delta\Phi_X,\Delta\Psi_X)$	$\sigma(\Delta\Phi_X,\Delta\Psi_{X-1})$	$\sigma(\Delta\Phi_X,\Delta\Psi_X)$	$\sigma(\Delta\Phi_X,\Delta\Psi_{X-1})$
1	−0.11	not applicable	0.04	not applicable	−0.05	not applicable
2	−0.11	−0.67	0.05	−0.72	0.00	−0.67
3	−0.05	−0.68	0.03	−0.71	0.01	−0.68
4	−0.10	−0.68	0.08	−0.71	0.02	−0.67
5	−0.06	−0.68	0.09	−0.72	−0.03	−0.66
6	−0.09	−0.68	0.07	−0.72	−0.05	−0.67
7	−0.05	−0.68	0.08	−0.71	0.09	−0.68

[a] All simulations use our implicit solvent model and the N2 interaction routine. Each simulation is 30 ns in duration.

Table 5: Dependence of Correlation Coefficients on Solvent Model and Nonbonded Interaction Routine for Ala_7[a]

	implicit: N2		explicit: N∞		vacuum: N∞	
residue	$\sigma(\Delta\Phi_X,\Delta\Psi_X)$	$\sigma(\Delta\Phi_X,\Delta\Psi_{X-1})$	$\sigma(\Delta\Phi_X,\Delta\Psi_X)$	$\sigma(\Delta\Phi_X,\Delta\Psi_{X-1})$	$\sigma(\Delta\Phi_X,\Delta\Psi_X)$	$\sigma(\Delta\Phi_X,\Delta\Psi_{X-1})$
1	0.04	not applicable	−0.20	not applicable	0.02	not applicable
2	0.05	−0.72	−0.17	−0.67	0.01	−0.73
3	0.03	−0.71	−0.20	−0.60	−0.01	−0.75
4	0.08	−0.71	−0.25	−0.56	−0.15	−0.73
5	0.09	−0.72	−0.36	−0.62	−0.13	−0.72
6	0.07	−0.72	−0.27	−0.58	−0.01	−0.72
7	0.08	−0.71	−0.23	−0.59	0.01	−0.74

[a] All simulations use the CHARMM27 force field. The implicit solvent simulations used the N2 interaction routine, while the explicit solvent and vacuum simulations used the N∞ interaction routine. Each simulation is 30 ns in duration.

understanding the dynamics of proteins and for predicting the native structure of proteins. In general, Monte Carlo (MC) simulations designed to predict protein structure need not involve physically accessible moves. Consequently, no information concerning the folding mechanism can be rigorously extracted from such simulations. On the other hand, kinetic MC simulations employing physical accessible moves and realistic hopping rates may provide mechanistic information. Thus, the correlated motions observed in this study should be included in the set of physically realistic moves to be used in MC simulations.

Understanding the fundamental backbone motions of proteins is also important for determining reaction coordinates for macromolecular processes. Several studies of alanine dipeptide show that the Φ and Ψ torsion angles alone are insufficient for completely describing backbone conformational transitions (*47*, *48*). Since displacements of the Φ angle are more correlated with those of the preceding torsion angle than of Ψ from the same residue, in searching for reaction coordinates in larger systems, we conclude detailed information about the nature of correlated motions along the protein backbone should be useful.

CONCLUSION

We observe correlated backbone motions involving a strong anticorrelation in the motions of Ψ_{i-1} and Φ_i. This motion corresponds to the oscillation of the plane of the peptide group and is robust with respect to the identity of the amino acid, the force field, the degree to which nonbonded interactions are retained or neglected, and the solvent model. The exact correlation coefficients vary slightly with different simulation conditions, but the peptide group oscillation is always present and represents the largest correlated motion. The oscillation enables local backbone motions without large-scale displacements of the remainder of the chain through space. The motion largely arises from inertial effects and the rigidity of the peptide bond, as observed in the vacuum simulations as well. The correlated motion can be viewed as a rocking of the plane of the peptide group, and this motion is found both for fluctuations within each Ramachandran basin and during transit between different basins. For neighboring residues, however, the transitions from one Ramachandran basin to another, which are required for the backbone to sample different secondary structure types, are independent of each other.

ACKNOWLEDGMENT

J.E.F. thanks Professors David Angulo and Gregor von Laszewski for useful discussion. We thank Professor A. Dinner for assistance with the explicit solvent simulations, access to his computer cluster, and a critical review of the manuscript.

APPENDIX

Force Field Dependence. Simulations of Ala_7 with implicit solvent and N2 interactions are compared for all-atom force fields GS-AMBER94, CHARMM, and OPLS (Table 4). The dominance of the anticorrelation of Φ_i and Ψ_{i-1} motions persists for all force fields, and the precise correlation coefficient depends only very slightly on the force field, with CHARMM producing only slightly more anticorrelated motions than the other two force fields. The two-dimensional matrix plots of the correlation coefficients are not shown since they are nearly identical to Figure 2, and any differences can be discerned from Table 4. However, the Φ_i and Ψ_i motions exhibit a stronger dependence on the force field. The GS-AMBER94 force field produces a slight anticorrelation between the motions of Φ_i and Ψ_i in the range from −0.11 to −0.05; the CHARMM force field yields a less significant correlation in the range of 0.05 to 0.09, and the OPLS force field exhibits statistically insignificant correlations except in one case where the correlation coefficient is 0.09 and in two where there is a anticorrelation of −0.05.

Differences between Solvent Models and Nonbonded Interaction Routines. Implicit solvent, explicit solvent, and no solvent (vacuum) simulations for Ala_7 with the CHARMM force field are compared to assess the importance of the solvent model and the importance of inertia relative to solvent viscosity for the chain motions. The implicit solvent simulation is performed with the N2 nonbonded interaction scheme, while the explicit solvent simulation uses the TIP3P water model, a nonbonded cutoff of 12 Å, and the force shift technique starting at 8 Å. The explicit solvent and the vacuum simulations include nonbonded interactions between all residues. The correlation coefficients obtained from implicit solvent and vacuum simulations are very similar (Table 5). The explicit solvent simulations retain the inequality $\sigma(\Delta\Phi_i,\Delta\Psi_{i-1}) > \sigma(\Delta\Phi_i,\Delta\Psi_i)$, but the inequality is not quite as strong as for the implicit solvent and vacuum simulations. More specifically, the explicit solvent calculations yield a slightly stronger correlation between torsion angles on the same amino acid [explicit solvent gives $\sigma(\Delta\Phi_i,\Delta\Psi_i)$ in the range $[-0.36,-0.17]$, while implicit solvent gives $\sigma(\Delta\Phi_i,\Delta\Psi_i)$ in the range [0.03, 0.09]] and a slightly weaker correlation between $\Delta\Phi_i$ and $\Delta\Psi_{i-1}$ [explicit solvent gives $\sigma(\Delta\Phi_i,\Delta\Psi_i)$ in the range $[-0.67,-0.56]$, while implicit solvent gives $\sigma(\Delta\Phi_i,\Delta\Psi_i)$ in the range $[-0.72,-0.71]$]. In all cases, all correlations die out at longer distances, leaving the peptide group's oscillatory motion as the dominant (anti)-correlated motion. Because the dominance of the anticorrelation of $(\Delta\Phi_i,\Delta\Psi_{i-1})$ also persists in the vacuum simulations and is independent of force field, the prevalence of this anticorrelated motion must arise from structural properties of the peptide, namely, from inertial and/or rigidity effects, rather than viscous drag.

Simulations of the heteropolymer EDEVARLKKLLW have also been run using N2 and N∞ interactions. Figure 1 of the Supporting Information demonstrates that the pattern of correlations is almost identical for both of these trajectories. Further discussion of these simulations is contained in the text.

SUPPORTING INFORMATION AVAILABLE

Graphical representation of the correlation coefficients for movements of the torsion angles of the heteropolymer EDEVARLKKLLW using the implicit solvent model, the Garcia force field, and N2 interactions and N∞ interactions. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Go, M., and Go, N. (1976) Fluctuations of an α-Helix, *Biopolymers 15*, 1119−1127.
2. Mccammon, J. A., Gelin, B. R., and Karplus, M. (1977) Dynamics of Folded Proteins, *Nature 267*, 585−590.
3. Helfand, E., Wasserman, Z. R., and Weber, T. A. (1979) Brownian Dynamics Study of Polymer Conformational Transitions, *J. Chem. Phys. 70*, 2016−2017.
4. Levitt, M. (1983) Molecular-Dynamics of Native Protein. 2. Analysis and Nature of Motion, *J. Mol. Biol. 168*, 621−657.
5. Dellwo, M. J., and Wand, A. J. (1989) Model-Independent and Model-Dependent Analysis of the Global and Internal Dynamics of Cyclosporine-A, *J. Am. Chem. Soc. 111*, 4571−4578.
6. Garcia, A. E. (1992) Large-Amplitude Nonlinear Motions in Proteins, *Phys. Rev. Lett. 68*, 2696−2699.
7. Fadel, A. R., Jin, D. Q., Montelione, G. T., and Levy, R. M. (1995) Crankshaft Motions of the Polypeptide Backbone in Molecular-Dynamics Simulations of Human Type-α Transforming Growth-Factor, *J. Biomol. NMR 6*, 221−226.
8. Tama, F., and Sanejouand, Y. H. (2001) Conformational change of proteins arising from normal mode calculations, *Protein Eng. 14*, 1−6.
9. Hoang, T. X., and Cieplak, M. (1998) Protein folding and models of dynamics on the lattice, *J. Chem. Phys. 109*, 9192−9196.
10. Demene, H., and Sugar, I. P. (1999) Protein conformation and dynamics. Effects of crankschaft motions on N-15 relaxation data and H-1 NMR NOESY spectra, *Biophys. J. 76*, A116.
11. Demene, H., and Sugar, I. P. (1999) Protein conformation and dynamics. Effects of crankshaft motions on H-1 NMR cross-relaxation effects, *J. Phys. Chem. A 103*, 4664−4672.
12. Hilhorst, H. J., and Deutch, J. M. (1975) Analysis of Monte-Carlo Results on Kinetics of Lattice Polymer-Chains with Excluded Volume, *J. Chem. Phys. 63*, 5153−5161.
13. Helfand, E. (1971) Theory of Kinetics of Conformational Transitions in Polymers, *J. Chem. Phys. 54*, 4651.
14. Helfand, E. (1971) Kinetics of Conformational Transitions, *Bull. Am. Phys. Soc. 16*, 388.
15. Go, N., and Scheraga, H. A. (1973) Ring-Closure in Chain Molecules with C_n, I, or S_{2n} Symmetry, *Macromolecules 6*, 273−281.
16. Betancourt, M. R. (2005) Efficient Monte Carlo trial moves for polypeptide simulations, *J. Chem. Phys. 123*, 174905.
17. Elofsson, A., Legrand, S. M., and Eisenberg, D. (1995) Local Moves: An Efficient Algorithm for Simulation of Protein Folding, *Proteins 23*, 73−82.
18. Chandrasekhar, I., Clore, G. M., Szabo, A., Gronenborn, A. M., and Brooks, B. R. (1992) A 500-Ps Molecular-Dynamics Simulation Study of Interleukin-1-β in Water: Correlation with Nuclear-Magnetic-Resonance Spectroscopy and Crystallography, *J. Mol. Biol. 226*, 239−250.
19. Palmer, A. G., and Case, D. A. (1992) Molecular-Dynamics Analysis of Nmr Relaxation in a Zinc-Finger Peptide, *J. Am. Chem. Soc 114*, 9059−9067.
20. Brunne, R. M., Berndt, K. D., Guntert, P., Wuthrich, K., and Vangunsteren, W. F. (1995) Structure and Internal Dynamics of the Bovine Pancreatic Trypsin-Inhibitor in Aqueous-Solution from Long-Time Molecular-Dynamics Simulations, *Proteins 23*, 49−62.
21. Zaman, M. H., Shen, M. Y., Berry, R. S., Freed, K. F., and Sosnick, T. R. (2003) Investigations into sequence and conformational dependence of backbone entropy, inter-basin dynamics and the flory isolated-pair hypothesis for peptides, *J. Mol. Biol. 331*, 693−711.
22. Zaman, M. H., Shen, M. Y., Berry, R. S., and Freed, K. F. (2003) Computer simulation of met-enkephalin using explicit atom and united atom potentials: Similarities, differences, and suggestions for improvement, *J. Phys. Chem. B 107*, 1685−1691.
23. Mu, Y. G., Kosov, D. S., and Stock, G. (2003) Conformational dynamics of trialanine in water. 2. Comparison of AMBER, CHARMM, GROMOS, and OPLS force fields to NMR and infrared experiments, *J. Phys. Chem. B 107*, 5064−5073.
24. Garcia, A. E., and Sanbonmatsu, K. Y. (2002) α-Helical stabilization by side chain shielding of backbone hydrogen bonds, *Proc. Natl. Acad. Sci. U.S.A. 99*, 2782−2787.
25. Hu, H., Elstner, M., and Hermans, J. (2003) Comparison of a QM/MM force field and molecular mechanics force fields in simulations of alanine and glycine "dipeptides" (Ace-Ala-Nme and Ace-Gly-Nme) in water in relation to the problem of modeling the unfolded peptide backbone in solution, *Proteins 50*, 451−463.
26. Clore, G. M., and Schwieters, C. D. (2004) Amplitudes of protein backbone dynamics and correlated motions in a small α/β protein: Correspondence of dipolar coupling and heteronuclear relaxation measurements, *Biochemistry 43*, 10678−10691.
27. Jha, A. (2002) Unpublished work, University of Chicago, Chicago.
28. Shen, M. Y., and Freed, K. F. (2002) Long time dynamics of met-enkephalin: Comparison of explicit and implicit solvent models, *Biophys. J. 82*, 1791−1808.
29. Shen, M. Y., and Freed, K. F. (2002) All-atom fast protein folding simulations: The villin headpiece, *Proteins 49*, 439−445.
30. MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, W. E., Roux, B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Yin, D.,

and Karplus, M. (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins, *J. Phys. Chem. B 102*, 3586−3616.
31. Jorgensen, W. L., and Tiradorives, J. (1988) The Opls Potential Functions for Proteins: Energy Minimizations for Crystals of Cyclic-Peptides and Crambin, *J. Am. Chem. Soc. 110*, 1657−1666.
32. Ramstein, J., and Lavery, R. (1988) Energetic Coupling between DNA Bending and Base Pair Opening, *Proc. Natl. Acad. Sci. U.S.A. 85*, 7231−7235.
33. Ooi, T., Oobatake, M., Nemethy, G., and Scheraga, H. A. (1987) Accessible Surface-Areas as a Measure of the Thermodynamic Parameters of Hydration of Peptides, *Proc. Natl. Acad. Sci. U.S.A. 84*, 3086−3090.
34. Shen, M. Y., and Freed, K. F. (2005) A simple method for faster nonbonded force evaluations, *J. Comput. Chem. 26*, 691−698.
35. Allen, M. P., and Tildesley, D. J. (1989) *Computer simulation of liquids*, Oxford University Press, Oxford, England.
36. Andersen, H. C. (1983) Rattle: A Velocity Version of the Shake Algorithm for Molecular-Dynamics Calculations, *J. Comput. Phys. 52*, 24−34.
37. Shen, M. Y. (2002) in *Chemistry*, pp 198, University of Chicago, Chicago.
38. Pastor, R. W., and Karplus, M. (1988) Parametrization of the Friction Constant for Stochastic Simulations of Polymers, *J. Phys. Chem. 92*, 2636−2641.
39. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of Simple Potential Functions for Simulating Liquid Water, *J. Chem. Phys. 79*, 926−935.
40. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. (1983) Charmm: A Program for Macromolecular Energy, Minimization, and Dynamics Calculations, *J. Comput. Chem. 4*, 187−217.
41. Steinbach, P. J., and Brooks, B. R. (1994) New Spherical-Cutoff Methods for Long-Range Forces in Macromolecular Simulation, *J. Comput. Chem. 15*, 667−683.
42. Ryckaert, J. P., Ciccotti, G., and Berendsen, H. J. C. (1977) Numerical-Integration of Cartesian Equations of Motion of a System with Constraints: Molecular-Dynamics of N-Alkanes, *J. Comput. Phys. 23*, 327−341.
43. Flory, P. J. (1953) *Principles of polymer chemistry*, Cornell University Press, Ithaca, NY.
44. Yamakawa, H. (1971) *Modern theory of polymer solutions*, Harper & Row, New York,.
45. Freed, K. F. (1987) *Renormalization group theory of macromolecules*, J. Wiley, New York.
46. Kramers, H. A. (1940) Brownian motion in a field of force and the diffusion model of chemical reactions, *Physica 7*, 284−304.
47. Ma, A., and Dinner, A. R. (2005) Automatic method for identifying reaction coordinates in complex systems, *J. Phys. Chem. B 109*, 6769−6779.
48. Bolhuis, P. G., Dellago, C., and Chandler, D. (2000) Reaction coordinates of biomolecular isomerization, *Proc. Natl. Acad. Sci. U.S.A. 97*, 5877−5882.

BI061575X