PROTEINS
STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

Equilibrium transitions between side-chain conformations in leucine and isoleucine

Diego Caballero,[1,2] W. Wendell Smith,[1] Corey S. O'Hern,[1,2,3,4] and Lynne Regan[2,5,6]*

[1] Department of Physics, Yale University, New Haven, Connecticut
[2] Integrated Graduate Program in Physical and Engineering Biology, Yale University, New Haven, Connecticut
[3] Department of Mechanical Engineering & Materials Science, Yale University, New Haven, Connecticut
[4] Department of Applied Physics, Yale University, New Haven, Connecticut
[5] Department of Molecular Biophysics & Biochemistry, Yale University, New Haven, Connecticut
[6] Department of Chemistry, Yale University, New Haven, Connecticut

ABSTRACT

Despite recent improvements in computational methods for protein design, we still lack a quantitative, predictive understanding of the intrinsic probabilities for amino acids to adopt particular side-chain conformations. Surprisingly, this question has remained unsettled for many years, in part because of inconsistent results from different experimental approaches. To explicitly determine the relative populations of different side-chain dihedral angles, we performed all-atom hard-sphere Langevin Dynamics simulations of leucine (Leu) and isoleucine (Ile) dipeptide mimetics with stereo-chemical constraints and repulsive-only steric interactions between non-bonded atoms. We determine the relative populations of the different χ_1 and χ_2 dihedral angle combinations as a function of the backbone dihedral angles ϕ and ψ. We also propose, and test, a mechanism for inter-conversion between the different side-chain conformations. Specifically, we discover that some of the transitions between side-chain dihedral angle combinations are very frequent, whereas others are orders of magnitude less frequent, because they require rare coordinated motions to avoid steric clashes. For example, to transition between different values of χ_2, the Leu side-chain bond angles κ_1 and κ_2 must increase, whereas to transition in χ_1, the Ile bond angles λ_1 and λ_2 must increase. These results emphasize the importance of computational approaches in stimulating further experimental studies of the conformations of side-chains in proteins. Moreover, our studies emphasize the power of simple steric models to inform our understanding of protein structure, dynamics, and design.

Proteins 2015; 83:1488–1499.

Key words: Markov chains; hydrophobic amino acids; side-chain conformations; side-chain dihedral angles; rotamer prediction; langevin dynamics of proteins; protein structure prediction; NMR studies of proteins; protein–protein interactions; protein folding.

INTRODUCTION

For most amino acid types, there are a small number of side-chain conformations that are highly populated,[1] both in protein crystal structures and in solution. Any particular residue in a crystal structure adopts only one of these conformations, and it is presumed that the conformation that is adopted is significantly affected by local interactions.[2–4] For unfolded proteins and short peptides in solution, however, side-chains sample different conformations as the system thermally fluctuates.[5–13] Nevertheless, studies have suggested that the different side-chain conformations sampled in the unfolded state mirror the set that is observed in a large ensemble of protein crystal structures.

Because of their role in protein stability,[14–17] secondary structure determination,[3,18] and protein–protein interactions,[19,20] it is important to develop the ability to understand and predict the probability distributions of amino acid side-chain dihedral angles. Although in

Additional Supporting Information may be found in the online version of this article.

Grant sponsor: National Science Foundation (NSF); Grant number: PHY-1019147; Grant sponsor: Raymond and Beverly Sackler Institute for Biological, Physical and Engineering Sciences; Grant sponsor: Kavli Institute for Theoretical Physics; Grant number: PHY11-25915; Grant sponsor: NSF; Grant number: CNS 08-21132.
*Correspondence to: Lynne Regan, Integrated Graduate Program in Physical and Engineering Biology, Yale University, New Haven, Connecticut. E-mail: lynne.regan@yale.edu
Received 11 February 2015; Revised 30 April 2015; Accepted 20 May 2015
Published online 28 May 2015 in Wiley Online Library (wileyonlinelibrary.com).
DOI: 10.1002/prot.24837

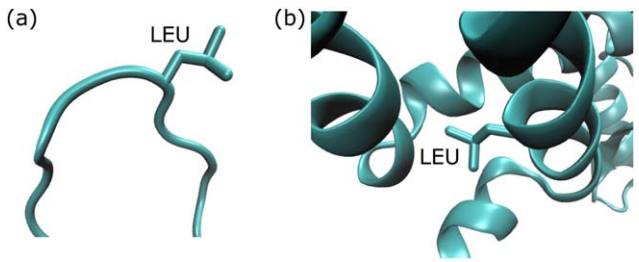

Figure 1

(**a**) X-ray crystal structure in ribbon representation of a coil region of the protein T4 Lysozyme (PDB ID: 1L63) with Leu 15 highlighted. (**b**) Close-up of Leu 99, which occurs on an α-helical segment in the core of T4 Lysozyme. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

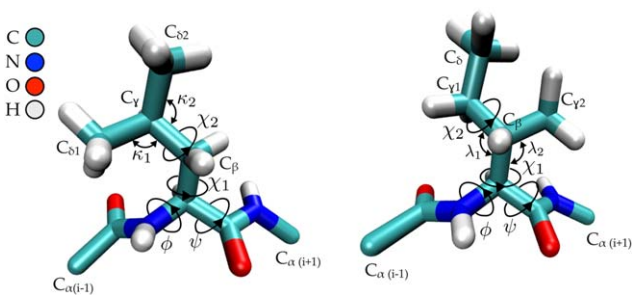

Figure 2

Stick representations of leucine (left) and isoleucine (right) dipeptide mimetics, N-acetyl-L-Leu-methylamide and N-acetyl-L-Ile-methylamide, respectively. The carbon, nitrogen, oxygen, and hydrogen atoms are shaded green, blue, red, and white. The backbone dihedral angles ϕ and ψ, side-chain dihedral angles χ_1 and χ_2, and bond angles κ_1 and κ_2 on Leu (λ_1 and λ_2 on Ile) are indicated. The residues before and after the ith central residue are labeled $i-1$ and $i+1$. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

the unfolded state there may be instances of residual structure, the dominant determinant of side-chain conformations in the fully denatured state is local steric constraints.[21,22] Thus, it is reasonable to model the equilibrium probability distributions for the side-chain dihedral angles for an amino acid in a denatured protein using molecular simulations of an isolated dipeptide [Fig. 1(a)], where the backbone and side-chain conformations can transition among different states.

As a protein folds to its native state, the conformational space available to a given side-chain becomes increasingly more restricted, until in the folded protein, the side-chain is essentially fixed [Fig. 1(b)] and cannot substantially change its conformation without unfolding the protein to some degree.[4,23] For example, Vajpai *et al.*[10] measured the side-chain dihedral angle χ_1 for each residue in the urea denatured states of ubiquitin and the B1 domain of IgG-binding protein G. They found that the predominant rotamer observed in the folded protein was not necessarily the rotamer observed in the urea denatured state. About half of the residues in each protein populated 20% or less the side-chain conformation that occurs in the native state. Similarly, the side-chain χ_1 values for leucine (Leu) and isoleucine (Ile) in an infrequently populated intermediate state and the native state have been shown to be quite different.[8,9]

One can thus envision two different types of probabilities. An *out-of-equilibrium* (observed) side-chain dihedral angle probability distribution $P^o(\chi_1, \chi_2)$ obtained by averaging the side-chain dihedral angle conformations for a given amino acid type over each fixed crystal structure and an *equilibrium* side-chain dihedral angle probability distribution $P^{eq}(\chi_1, \chi_2)$ obtained by averaging the side-chain conformations of a single residue (or dipeptide) over time as it undergoes thermal fluctuations. This manuscript will focus on characterizing the similarities between these two side-chain dihedral angle distributions for Leu and Ile (Fig. 2). We chose the amino acids Leu and Ile because they are abundant in proteins and well studied.[24,25] In addition, Leu and Ile are hydrophobic and not highly charged so that a simple hard-sphere model with stereochemical constraints can accurately reproduce their side-chain dihedral angle distributions.[26,27]

We perform Langevin Dynamics simulations of a hard-sphere model for Leu and Ile dipeptide mimetics with imposed stereochemical constraints on the bond lengths and angles and calculate the equilibrium probability distributions for the side-chain dihedral angles $P^{eq}(\chi_1, \chi_2)$. We compare the results of the Langevin Dynamics simulations to the side-chain dihedral angle distributions $P^o(\chi_1, \chi_2)$ obtained from a high-resolution non-redundant subset of coil structures ("Wu Coil-3 library") from the protein data bank (PDB).[28] We find significant similarities between the predicted and observed distributions, but also point out several key differences.

From the hard-sphere Langevin Dynamics simulations, we also show that transitions from one frequently sampled side-chain conformation to another occur when the $\kappa_{1,2}$ ($\lambda_{1,2}$) bond angles among the atoms C_β-C_γ-$C_{\delta 1,2}$ (C_α-C_β-$C_{\gamma 1,2}$) are larger than the values obtained from averaging the $\kappa_{1,2}$ ($\lambda_{1,2}$) bond angles over all Leu (Ile) residues in high-resolution protein crystal structures. Further, the probability distributions $P_t^o(\kappa_{1,2})$ ($P_t^o(\lambda_{1,2})$) for the κ (λ) bond angles in the transition regions of χ_1-χ_2 space are shifted to larger values than the full bond angle distributions from all observed high-resolution crystal structures. These results suggest that larger than average values of the bond angles $\kappa_{1,2}$ and $\lambda_{1,2}$ enable transitions among particular side-chain dihedral angle conformations.

MATERIALS AND METHODS

We studied all-atom hard-sphere representations of leucine (N-acetyl-L-Leu-N′-methylamide) and isoleucine

(N-acetyl-L-Ile-N'-methylamide) dipeptide mimetics, as shown in Figure 2. The Leu and Ile dipeptide mimetics possess 24 bonds between pairs of atoms and 42 bond angles (including bonds that involve hydrogen atoms) between atom triples. The averages and standard deviations of the bond lengths (l_{ij}^0 and Δl_{ij}), bond angles (θ_{ijk}^0 and $\Delta\theta_{ijk}$), and ω backbone dihedral angles ($\omega_{ijkl}^0 \approx 0$ and $\Delta\omega_{ijkl}$) were obtained for Leu and Ile from a database of high-resolution protein crystal structures, that is, the "culled Dunbrack database", and are listed in the Supporting Information. This database is a subset of protein ctystal structures with resolution ≤ 1.0 Å and R factor ≤ 0.2 from the PDB provided by Dr. Roland Dunbrack, Jr.[29,30] Note that this data set is not a subset of the set presented in Shapovalov and Dunbrack,[31] even though a similar methodology was used to obtain it. The selected subset of protein crystal structures includes 2204 Leu and 1555 Ile residues. This subset of structures is just one of several high-resolution protein databases that could have been used.[32,33]

The atomic diameters σ_i, C(sp^3): 1.5 Å, C(sp^2): 1.4 Å, N: 1.4 Å, O: 1.35 Å, and H: 1.05 Å, are similar to values employed in previous studies,[26,27,34–36] except the oxygen diameter was decreased from 1.4 Å to 1.35 Å to improve sampling in ϕ-ψ space.[36] Hydrogen atoms were added to the dipeptide using the REDUCE software program.[37] The simulations of the dipeptide mimetics include the following four interaction potentials between spherical atoms i and j: (1) a purely repulsive Lennard-Jones potential,[38]

$$V_{\mathrm{LJ}} = \epsilon\left(1-\left(\frac{\sigma_{ij}}{r_{ij}}\right)^6\right)^2 \Theta(\sigma_{ij}-r_{ij}), \tag{1}$$

where ϵ is the characteristic energy scale of the interaction, r_{ij} is the separation between non-bonded atoms i and j, $\sigma_{ij} = (\sigma_i+\sigma_j)/2$, and $\Theta(x)$ is the Heaviside step function that prevents interactions between atoms when they are not in contact; (2) a harmonic potential to constrain the bond lengths,

$$V_{\mathrm{BL}} = \frac{K^{l_{ij}}}{2}\left(r_{ij}-l_{ij}^0\right)^2, \tag{2}$$

where $K^{l_{ij}} = T/(\Delta l_{ij})^2$ and T is the temperature in units of the Boltzmann constant; (3) a harmonic potential to constrain the bond angles,

$$V_{\mathrm{BA}} = \frac{K^{\theta_{ijk}}}{2}\left(\theta_{ijk}-\theta_{ijk}^0\right)^2, \tag{3}$$

where $K^{\theta_{ijk}} = T/(\Delta\theta_{ijk})^2$; and (4) a harmonic potential to constrain the two ω_{ijkl} dihedral angles (defined by the groups of four atoms $C_{\alpha,i-1}-C_{i-1}-N-C_\alpha$ and $C_\alpha-C-N_{i+1}-C_{\alpha,i+1}$) to be planar,

$$V_\omega = \frac{K^{\omega_{ijkl}}}{2}\omega_{ijkl}^2, \tag{4}$$

where $K^{\omega_{ijkl}} = T/(\Delta\omega_{ijkl})^2$. Note that the spring constants $K^{l_{ij}}$, $K^{\theta_{ijk}}$, and $K^{\omega_{ijkl}}$ are chosen so that the standard deviations at temperature T of the bond lengths, bond angles, and ω dihedral angles match those for Leu and Ile residues from the culled Dunbrack database of high-resolution protein crystal structures. The total potential energy V of a dipeptide is obtained by summing the interactions in Eqs. (1)–(4) over all non-bonded pairs of atoms, bonds, bond angles, and the two backbone dihedral angles ω that occur in each dipeptide:

$$V = \sum_{i>j} V_{LJ}(r_{ij}) + \sum_{\text{bonds}} V_{BL} + \sum_{\text{bond angles}} V_{BA} + \sum_{\omega \text{ dihedral angles}} V_\omega. \tag{5}$$

We performed implicit-solvent Langevin dynamics simulations of Leu and Ile dipeptide mimetics by numerically integrating

$$m_i\frac{d^2\vec{r}_i}{dt^2} = -\xi\frac{d\vec{r}_i}{dt} + \vec{\Gamma}_i - \frac{\partial V}{\partial\vec{r}_i} \tag{6}$$

for the atomic positions $\vec{r}_i$, where m_i is the mass of atom i, the Gaussian-distributed, δ-function correlated random forces $\vec{\Gamma}_i$ on atom i obey $\langle\vec{\Gamma}_i(t)\cdot\vec{\Gamma}_j(t')\rangle = 2\xi T\delta_{ij}\delta(t-t')$, and $\delta(x)$ (δ_{ij}) is the Dirac (Kronecker) δ-function. We implemented a modified Velocity Verlet algorithm[39] to integrate Eq. (6) with a time step $\Delta t = 10^{-4}\ t_0$, where $t_0 = \sigma_H\sqrt{m_H/\epsilon}$, and damping parameter $\xi = 5\epsilon t_0/\sigma_H^2$.

The initial atomic velocities were drawn from a Maxwell-Boltzmann distribution at temperature T^*, where $T^* = T/\epsilon \approx 10^{-2}$. The ratio T/ϵ determines the average amount of overlap (that is, pair separations that satisfy $r_{ij} < \sigma_{ij}$) between non-bonded atoms that occurs in the simulations. In the $(T/\epsilon) \to 0$ limit, the system explores only sterically allowed conformations. We showed in previous studies[36] that the average number of overlaps between pairs of non-bonded atoms becomes nonzero above the characteristic temperature T^*, which is the temperature of the simulations, and thus our simulations are carried out in the limit of hard-sphere interactions.

We find that for both Leu and Ile dipeptides the system is able to rapidly sample the sterically allowed backbone dihedral angle (ϕ-ψ) space. In contrast, sampling of the side-chain dihedral angle (χ_1-χ_2) space is slow due to the high energy barriers that separate different sterically allowed χ_1-χ_2 combinations. To efficiently sample χ_1-χ_2 space, we first performed $N_s = 64$ simulations of the dipeptides with random initial velocities and side-chain conformations in each of $N_b = 9$ 120°×120° boxes

in χ_1-χ_2 space for a sufficient amount of time to equilibrate the ϕ-ψ degrees of freedom.[36] During these simulations, we saved $N_t = 10^5$ static snapshots of the system (that is, the bond lengths, bond angles, ω dihedral angles, ϕ and ψ) equally spaced in time. For each snapshot s and initial condition c, we evaluated the Boltzmann weight at each dihedral angle combination χ_1 and χ_2 and temperature T^*:

$$P_{sc}^{eq}(\chi_1, \chi_2) \propto \exp\left(\frac{-V(\chi_1, \chi_2)}{T^*}\right) \tag{7}$$

with χ_1 and χ_2 sampled in 5° increments. We then averaged the Boltzmann weight over snapshots and initial conditions,

$$P^{eq}(\chi_1, \chi_2) \sim \frac{1}{N_s N_t} \sum_{s=1}^{N_t} \sum_{c=1}^{N_s} P_{sc}^{eq}(\chi_1, \chi_2) \tag{8}$$

and normalized the distribution in χ_1-χ_2 space such that $\int d\chi_1 d\chi_2 P^{eq}(\chi_1, \chi_2) = 1$.

We also coarse-grained the equilibrium probabilities $P^{eq}(\chi_1, \chi_2)$ over each of the N_b 120°×120° regions in χ_1-χ_2 space to yield the discrete probabilities P_j^{eq}, with $j = 1, \ldots, N_b$. P_j^{eq} is normalized such that $N_b^{-1} \sum_{j=1}^{N_b} P_j^{eq} = 1$ as shown in Figure 3. The discrete equilibrium probabilities were compared with the results from a Markov chain analysis that yields the steady-state probabilities probabilities in each χ_1-χ_2 region given trajectories of $\chi_1(t)$ and $\chi_2(t)$ for each initial condition over long times.[40,41] Details of the Markov chain analysis are provided in the Supporting Information.

To study the influence of the backbone bond angles κ (Leu) and λ (Ile) on the transition probabilities between different side-chain conformations, we also performed Langevin Dynamics simulations with elevated average values $\langle\kappa\rangle$ and $\langle\lambda\rangle$ compared to the average values obtained from protein crystal structures. We considered three cases: (1) $\langle\kappa_1\rangle = \kappa_0 > \bar{\kappa}_1$ and $\langle\kappa_2\rangle = \kappa_0 > \bar{\kappa}_2$ ($\langle\lambda_1\rangle = \lambda_0 > \bar{\lambda}_1$ and $\langle\lambda_2\rangle = \lambda_0 > \bar{\lambda}_2$); (2) only $\langle\kappa_1\rangle = \kappa_0 > \bar{\kappa}_1$ (only $\langle\lambda_1\rangle = \lambda_0 > \bar{\lambda}_1$); and (3) only $\langle\kappa_2\rangle = \kappa_0 > \bar{\kappa}_2$ (only $\langle\lambda_2\rangle = \lambda_0 > \bar{\lambda}_2$), where $\kappa_0 = 114.5° \pm 1°$, $\bar{\kappa}_1 = 110.6° \pm 1.8°$, and $\bar{\kappa}_2 = 111.1° \pm 1.8°$; and $\lambda_0 = 114.5° \pm 1°$, $\bar{\lambda}_1 = 111.2° \pm 1.6°$, and $\bar{\lambda}_2 = 110.7° \pm 1.4°$. Note that the standard deviations of the bond angle distributions in each of the three constrained simulations were smaller than those obtained from protein crystal structures.

We compare the results for the equilibrium side-chain dihedral angle distributions $P^{eq}(\chi_1, \chi_2)$ with the out-of-equilibrium distributions $P^o(\chi_1, \chi_2)$ obtained from Leu and Ile residues in protein crystal structures in the 'Coil-3' library.[28] This library is derived from protein crystal structures with resolution ≤ 2.0Å, R-factor ≤ 0.2, and a 50% sequence identity cutoff, but only includes residues that do not occur in α-helical or β-sheet secondary structures, are not pre-proline residues, or located in turns. The 'Wu Coil-3' library was developed to remove the oversampling of α-helix backbone conformations in the PDB. We obtain similar results using other coil libraries.[42,43]

The Coil-3 library contains $N_L = 56,009$ and $N_I = 34,098$ Leu and Ile residues, respectively, including both buried and surface residues. The observed probability distributions are calculated using

$$P_{L,I}^o(\chi_1, \chi_2) = \frac{N(\chi_1 - \frac{\Delta\chi_1}{2} < \chi_1 < \chi_1 + \frac{\Delta\chi_1}{2} \text{ and } \chi_2 - \frac{\Delta\chi_2}{2} < \chi_2 < \chi_2 + \frac{\Delta\chi_2}{2})}{N_{L,I}\Delta\chi_1\Delta\chi_2}, \tag{9}$$

where $N_{L,I}(\chi_1 - \frac{\Delta\chi_1}{2} < \chi_1 < \chi_1 + \frac{\Delta\chi_1}{2}$ and $\chi_2 - \frac{\Delta\chi_2}{2} < \chi_2 < \chi_2 + \frac{\Delta\chi_2}{2})$ is the number of Leu or Ile residues in the Coil-3 library that have side-chain dihedral angle combinations in a bin with area $\Delta\chi_1\Delta\chi_2$ centered on (χ_1, χ_2).

RESULTS

We performed Langevin Dynamics simulations of hard-sphere models with stereochemical constraints for Leu and Ile dipeptide mimetics (Fig. 2). With these simulations, we are able to rapidly sample all sterically allowed backbone dihedral angle combinations (ϕ, ψ) and transition frequently between the α-helical and β-sheet backbone conformations through the bridge region [Fig. 3(a)].[36] By contrast, although we observe rapid transitions between some side-chain dihedral angle combinations (χ_1, χ_2), transitions between others are orders of magnitude less frequent. Thus, instead of using long-time averages to measure the equilibrium probability distributions $P^{eq}(\chi_1, \chi_2)$, we averaged the Boltzmann factor at each (χ_1, χ_2) over all backbone dihedral angle, bond length, and bond angle combinations sampled during the Langevin Dynamics simulations (right column of Fig. 3). The resulting equilibrium distributions were validated against the results of Markov chain analyses. (See Supporting Information.)

Leucine

The equilibrium side-chain dihedral angle distributions $P^{eq}(\chi_1, \chi_2)$ (as well as the coarse-grained probabilities P_j^{eq}) for Leu are shown in Figure 3. We display the side-chain dihedral angle distributions associated with the complete ϕ-ψ map, in addition to the α, β' (β-sheet and PP_{II}), and bridge regions separately. Similar results for the side-chain dihedral angle distributions for Leu were shown previously[26,27]. The most probable side-chain conformations occur in boxes 5, 6, 8, and 9 (where 6 and 8 are more frequent than 5 and 9), whereas boxes 2, 3, 4, and 7 are rarely populated. Box 1 is never populated.

The Leu dipeptide can freely interconvert between side-chain dihedral angle combinations in boxes 5, 6, 8 and 9; whereas side-chain conformations in boxes 2 and

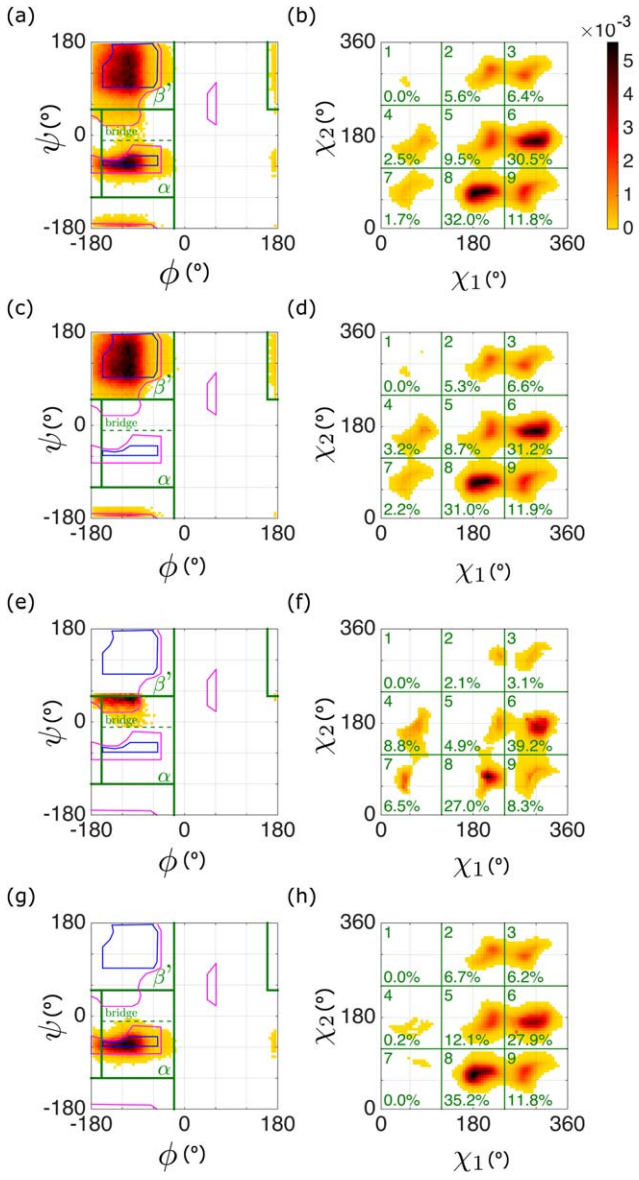

Figure 3

Equilibrium probability distributions $P^{eq}(\chi_1, \chi_2)$ [(**b**), (**d**), (**f**), and (**h**)] for the side-chain dihedral angles in Leu obtained by averaging the Boltzmann weight (Eq. 7) over (**a**) all sampled backbone conformations ϕ and ψ or (**c**) β' ($180° < \phi < 20°$ and $50° < \psi < 180°$, $180° < \phi < 200°$ and $180° < \psi < 120°$, $160° < \phi < 180°$ and $50° < \psi < 180°$), (**e**) bridge ($160° < \phi < 20°$ and $10° < \psi < 50°$), and (**g**) α ($160° < \phi < 20°$ and $120° < \psi < -10°$) regions (delineated by the dashed and solid green lines). Values of the probability decrease from dark to light as indicated on the color scale. The backbone and side-chain dihedral angle distributions are normalized in each panel according to $\int P(\phi, \psi) d\phi d\psi = 1$ and $\int P(\chi_1, \chi_2) d\chi_1 d\chi_2 = 1$. Coarse-grained probabilities P_j^{eq} for each of the $j = 1, \ldots, 9$ 120°×120° boxes are also shown with $\sum_j P_j^{eq} = 1$. Figure 7 shows the analogous data for Ile.

3, or 4 and 7 are isolated from other boxes. Overall, 'horizontal' transitions (that is, changes in χ_1) are significantly more frequent than transitions in χ_2. This asymmetry in the dynamics can be explained by considering the structure of the Leu side-chain, which branches at the C_γ atom (Fig. 2). For Leu, potential clashes can arise between two C atoms ($C_{\delta 1}$ and $C_{\delta 2}$) as χ_2 varies.

We also identify several features of $P^{eq}(\chi_1, \chi_2)$ that depend on the backbone conformation. Correlations between backbone and side-chain conformations have been observed before,[44] and are expected since the angles ϕ, ψ, and χ_1 involve a common bond (N-C_α)[21,45] and side-chain atoms can clash with main chain atoms[46] for particular values of χ_1, χ_2, ϕ, and ψ. For example, in Figure 3(h) we find that χ_1-χ_2 combinations in boxes 4 and 7 are much more infrequent when ϕ-ψ occurs in the α region as others have shown.[18,47] In addition, when the backbone of the Leu dipeptide occupies the bridge region, the χ_1-χ_2 combinations that populate boxes 5 and 8 are shifted to larger values of χ_1 [compare Fig. 3(b,f)].

In previous studies of Alanine dipeptides,[36] we identified a strong correlation between the main chain bond angle τ and transitions from α-helical to β-sheet backbone conformations and vice versa. In particular, when τ is large, the probability for transitions between α-helical and β-sheet backbone conformations increases. For Leu and Ile, we investigate whether there are similar correlations between the transition rates from one side-chain conformation to another and the side-chain bond angles κ_1 and κ_2 (Leu) or λ_1 and λ_2 (Ile) (Fig. 2).

In Figure 4(a), we decomposed χ_1-χ_2 space for Leu into three frequently sampled and three transition regions along χ_2. We then measured the probability distributions $P(\kappa_1)$ and $P(\kappa_2)$ in each χ_2 region. We find that the distributions $P(\kappa_1)$ and $P(\kappa_2)$ in the frequently sampled regions match the distributions expected from Boltzmann weighting V_{BA} [Eq. (3)] for κ_1 and κ_2. In contrast, $P(\kappa_1)$ and $P(\kappa_2)$ for conformations in the transition regions are shifted to larger values compared to the mean values $\bar{\kappa}_1 = 110.6° \pm 1.8°$ and $\bar{\kappa}_2 = 111.1° \pm 1.8°$ obtained from Boltzmann weighting V_{BA} [Fig. 4(b)]. In particular, we find that side-chain conformations in transition region (i) are associated with large κ_2, side-chain conformations in transition region (ii) are associated with large κ_1 and κ_2, and side-chain conformations in region (iii) are associated with large κ_1. As shown in Figure 4(c), elevated values of κ_1 and κ_2 are also found in the transition regions for crystal structures in the Coil-3 library,[28] except in transition region (ii), where there is little to no shift of the observed distributions $P(\kappa_1)$ or $P(\kappa_2)$.

From the Markov chain analysis, described in the Supporting Information, we can determine the 64 rates (inverse lifetimes) $1/\langle t_{ij} \rangle$ for transitioning from a side-chain conformation in box i to one in box j. We exclude transitions into and out of box 1 for Leu because we never observe any. Transition rates are pictured as arrows with their weight proportional to the natural logarithm of the rate in Figure 5(a). As shown earlier in Figure

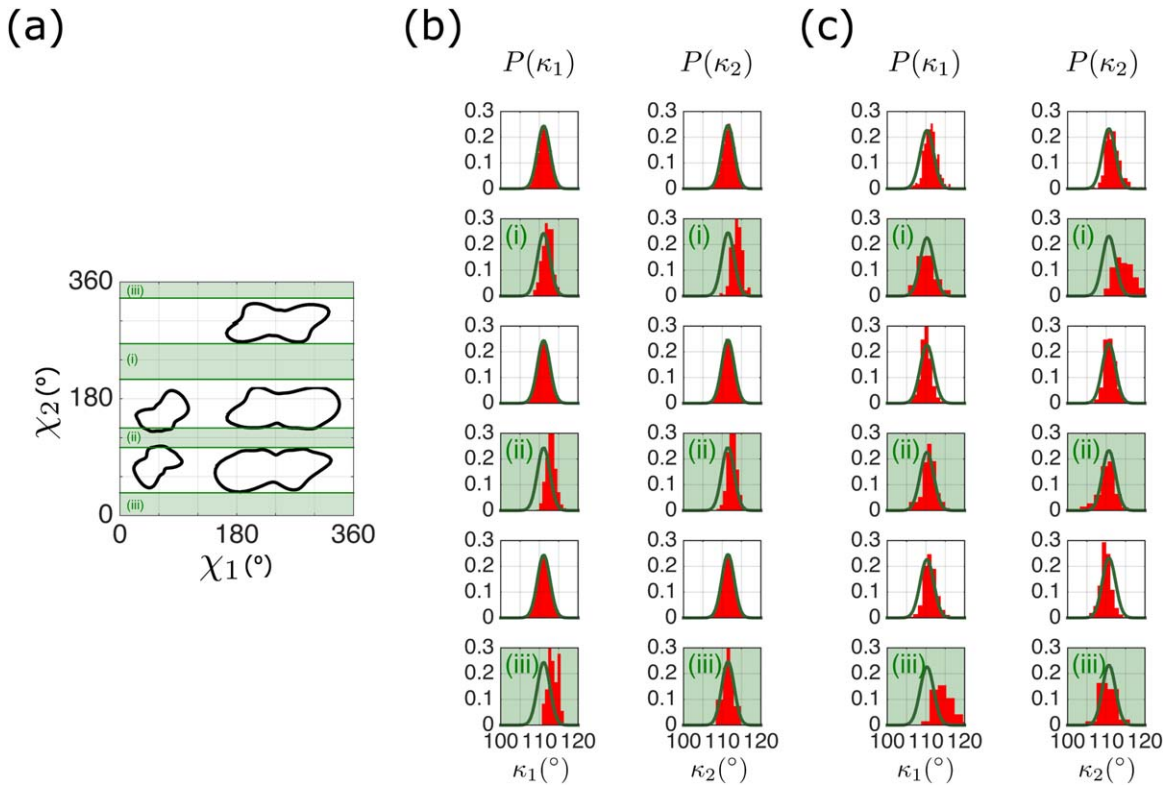

Figure 4

(**a**) Definitions of the transition (green shaded regions i, ii, and iii) and frequently sampled (unshaded) horizontal bands of the side-chain dihedral angle χ_1-χ_2 space for Leu. The black solid lines provide an outline of the area in χ_1-χ_2 space that includes 90% of the conformations sampled during the simulations, which occur mostly in the frequently sampled unshaded bands. The horizontal lines that divide the transition and frequently sampled bands are at $\chi_2 = 35°$, 105°, 140°, 210°, 270°, and 335°. (**b**)–(**c**) Probability distributions $P(\kappa_1)$ (left column) and $P(\kappa_2)$ (right column) for the side-chain bond angles κ_1 and κ_2 (red histograms) for Leu for the three transition and three frequently sampled regions defined in (a). The solid green lines show (b) the bond angle distributions from Boltzmann weighting the bond-angle potential $V_{BA}(\kappa_{1,2})$ or (c) the bond angle distributions from the full Coil-3 database. The distributions $P(\kappa_1)$ and $P(\kappa_2)$ are obtained from (b) Langevin Dynamics simulations of the hard-sphere model for the Leu dipeptide mimetic and (c) Leu residues extracted from the Wu Coil-3 database of protein crystal structures.[28] Each $P(\kappa_1)$ and $P(\kappa_2)$ is normalized such that $\int P(\kappa_{1,2})d\kappa_{1,2} = 1$. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

3(b), transitions in χ_1 are in general much more frequent than those in χ_2.

The fastest transition rates occur between side-chain conformations in boxes 5, 6, 8, and 9 and between conformations in boxes 2 and 3. In contrast, some transition rates are orders of magnitude smaller, for example, all transitions into and out of boxes 4 and 7. The transition rates between boxes 2 and 8, boxes 4 and 6, and several others indicated by dashed arrows fall below our measurement threshold $t_0/\langle t_{\min}\rangle = 10^{-7}$ [Fig. 5(a)]. For these transitions we therefore set $\langle t_{ij}\rangle = R\langle t_{\min}\rangle$, where $R = 3.5$. We chose this value of R because it minimizes the root-mean-square deviation of the equilibrium probabilites from Boltzmann sampling and the Markov chain analysis. The outcome of this minimization is shown in Figure 5(b), where we show excellent agreement for the equilibrium probabilities of side-chain dihedral angle combinations in boxes 2–9 for Leu from Boltzmann sampling (green) and the Markov chain analysis (blue).

We also performed Langevin Dynamics simulations with the side-chain bond angles κ_1 and κ_2 constrained so that their average values are larger than their original average values $\bar{\kappa}_1$ and $\bar{\kappa}_2$. Such simulations allow us to determine the influence of these changes on the transition rates between side-chain conformations. In particular, these simulations allow us to determine which of the 64 transitions between boxes for Leu are affected by increasing the average bond angles $\langle\kappa_1\rangle$ and $\langle\kappa_2\rangle$ (Fig. 6). We consider the three cases: (1) $\langle\kappa_1\rangle = \kappa_0 > \bar{\kappa}_1$ and $\langle\kappa_2\rangle = \kappa_0 > \bar{\kappa}_2$ [panel (b)], (2) only $\langle\kappa_1\rangle = \kappa_0 > \bar{\kappa}_1$ [panel (c)], and (3) only $\langle\kappa_2\rangle = \kappa_0 > \bar{\kappa}_2$ [panel (d)], where $\kappa_0 = 114.5° \pm 1°$, $\bar{\kappa}_1 = 110.6° \pm 1.8°$, and $\bar{\kappa}_2 = 111.1° \pm 1.8°$. We find that larger values of $\langle\kappa_1\rangle$ increase the transition rates between boxes 4 and 7 and boxes 3 and 9, that is, primarily transitions between different values of χ_2. Larger values of $\langle\kappa_2\rangle$ also increase transition rates between boxes 4 and 7, as well as transitions between boxes 2 and 5 and boxes 3 and 6. The increases

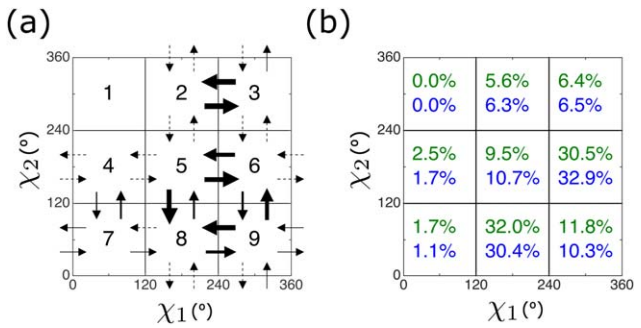

Figure 5

(**a**) The transition rates (inverse lifetimes) for Leu between side-chain conformations in different χ_1-χ_2 boxes obtained from Markov chain analyses are pictured as arrows with widths that are given by the natural logarithm of the transition rate. The transition rate scale ranges from a maximum of 10^{-4} to a minimum of 10^{-7} (in units of t_0^{-1}). The dotted arrows indicate transitions with rates that are below the measurement threshold of the simulations. (**b**) Comparison of the equilibrium probabilities P_j^{eq} for side-chain conformations in boxes $j = 1, \ldots, 9$ for Leu obtained from the Markov chain analysis of the rates in (a) (blue, dark) and Boltzman weighting at each χ_1-χ_2 combination (green, light) [as given Fig. 3(b)]. The probabilities agree to within 3%. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

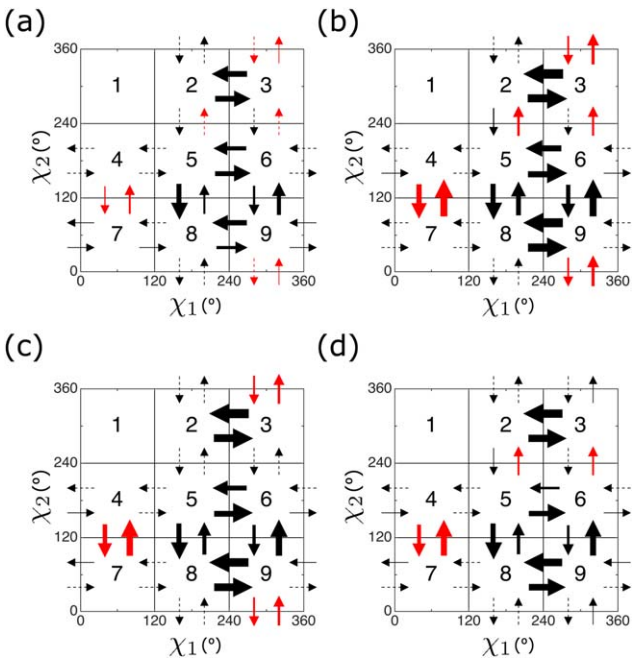

Figure 6

Transition rates (inverse lifetimes) for Leu between side-chain conformations in different χ_1-χ_2 boxes obtained from Markov chain analyses of Langevin Dynamics simulations for several different constrained side-chain backbone angle distributions $P(\kappa)$: (**a**) unconstrained simulations with averages $\langle\kappa_1\rangle = \bar{\kappa}_1$ and $\langle\kappa_2\rangle = \bar{\kappa}_2$; constrained simulations with (**b**) $\langle\kappa_1\rangle = \kappa_0 > \bar{\kappa}_1$ and $\langle\kappa_2\rangle = \bar{\kappa}_2$, (**c**) $\langle\kappa_1\rangle = \bar{\kappa}_1$ and $\langle\kappa_2\rangle = \kappa_0 > \bar{\kappa}_2$, and (**d**) $\langle\kappa_1\rangle = \kappa_0 > \bar{\kappa}_1$ and $\langle\kappa_2\rangle = \kappa_0 > \bar{\kappa}_2$, where $\kappa_0 = 114.5°$. The rates are pictured as arrows with widths that are given by the natural logarithm of the transition rate with maximum 10^{-4} and minimum of 10^{-7} (in units of t_0^{-1}). The dotted arrows indicate transitions with rates that are below the measurement threshold of the simulations. The red (light) arrows in (b)–(d) indicate several of the important transition rates from the constrained simulations that are different from those in (a). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

in the transition rates that result from larger values of κ_1 and κ_2 are consistent with the results shown in Figure 4(b). (See Supporting Information for visualizations of Leu dipeptides undergoing transitions from one side-chain rotameric state to another along χ_1 or χ_2.)

Isoleucine

We also performed the same Langevin Dynamics simulations of Ile dipeptides as we did for Leu. (Fig. 2). We display the side-chain dihedral angle distributions associated with the complete ϕ-ψ map, in addition to the α, β' (β-sheet and PP_{II}), and bridge regions separately (Fig. 7). When we average over all backbone conformations, we find that the most probable χ_1-χ_2 combinations occur in boxes 3, 5, and 6 with $P^{eq} \approx 19\%$, 33%, and 15%, respectively.[26,27] Boxes 4, 8, and 9 are less probable with $P^{eq} \approx 13\%$, 10%, and 7%, respectively. Boxes 1, 2, and 7 are very rarely sampled. We also find that the regions of high probability in the frequently sampled boxes are elongated in χ_2 direction, rather than the χ_1 direction as found for Leu, which indicates that transitions between frequently sampled χ_1-χ_2 combinations are impeded along χ_1. In contrast to Leu, the side-chain of Ile branches at C_β (Fig. 2), and thus transitions along χ_1 can be hindered by possible clashes involving two heavy atoms ($C_{\gamma 1}$ and $C_{\gamma 2}$) rather than only one for χ_2.

The equilibrium distributions $P^{eq}(\chi_1, \chi_2)$ for Ile are qualitatively similar for α and β' backbone conformations. However, the distribution is very different for backbone conformations in the bridge region, where only side-chain conformations in boxes 4, 5, and 7 are sampled with probabilities above 5%. This result emphasizes that the backbone conformation strongly influences the sampling of χ_1-χ_2 combinations in Ile.

We also investigated correlations between the side-chain bond angles λ_1 and λ_2 (Fig. 2) and transitions between χ_1-χ_2 combinations for Ile. In Figure 8(a), we divide χ_1-χ_2 space into 6 regions (vertical columns) along χ_1: 3 transition and 3 frequently sampled regions. The equilibrium bond angle distributions $P(\lambda_1)$ and $P(\lambda_2)$ for each region are shown in Figure 8(b). For all three transition regions, we find elevated values for either λ_1 or λ_2 compared to the distribution obtained from Boltzmann weighting V_{BA}. In both transition regions 1 and 2, we find $\langle\lambda_1\rangle > \bar{\lambda}_1 = 111.2° \pm 1.6°$ and $\langle\lambda_2\rangle > \bar{\lambda}_2 = 110.7° \pm 1.4°$, whereas in transition region (iii), $\langle\lambda_1\rangle > \bar{\lambda}_1$, but $\langle\lambda_2\rangle \approx \bar{\lambda}_2$. Similar results are found for

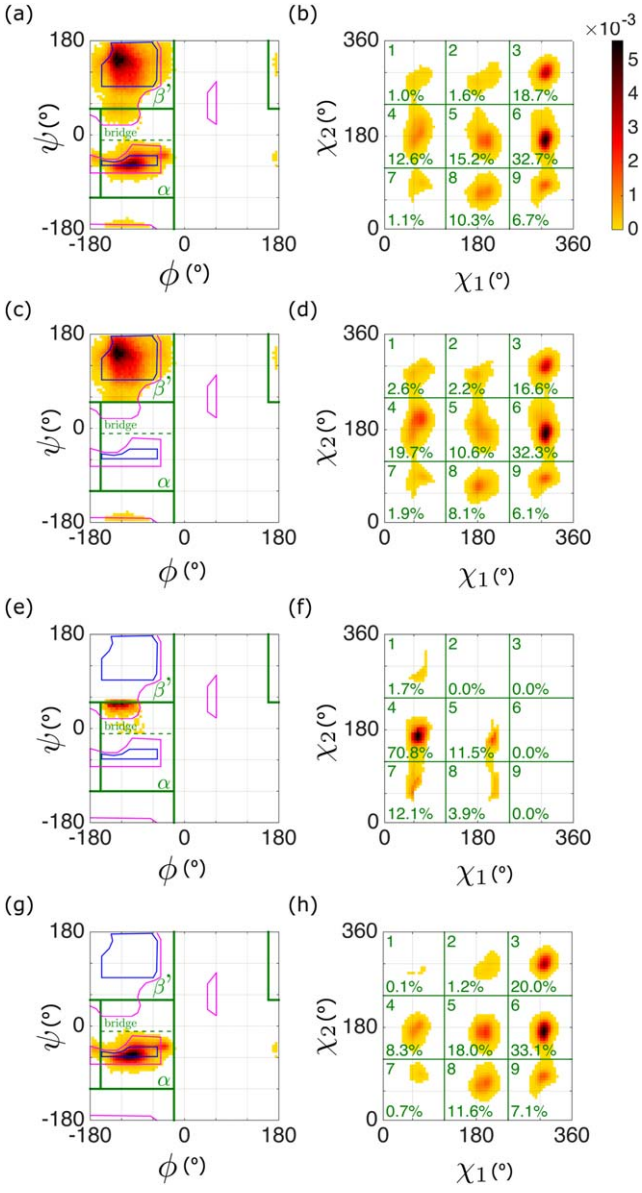

Figure 7

Equilibrium probability distributions $P^{eq}(\chi_1, \chi_2)$ [(**b**), (**d**), (**f**), and (**h**)] for the side-chain dihedral angles in Ile obtained by averaging the Boltzmann weight (Eq. 7) over (**a**) all sampled backbone conformations ϕ and ψ or (**c**) β′ ($180° < \phi < 20°$ and $50° < \psi < 180°$, $180° < \phi < 200°$ and $180° < \psi < 120°$, $160° < \phi < 180°$ and $50° < \psi < 180°$), (**e**) bridge ($160° < \phi < 20°$ and $10° < \psi < 50°$), and (g) α ($160° < \phi < 20°$ and $120° < \psi < -10°$) regions (delineated by the dashed and solid green lines). Values of the probability decrease from dark to light as indicated on the color scale. The backbone and side-chain dihedral angle distributions are normalized in each panel according to $\int P(\phi, \psi) d\phi d\psi = 1$ and $\int P(\chi_1, \chi_2) d\chi_1 d\chi_2 = 1$. Coarse-grained probabilities P_j^{eq} for each of the $j = 1, \ldots, 9$ 120°×120° boxes are also shown with $\sum_j P_j^{eq} = 1$. Figure 3 shows the analogous data for Leu. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

protein crystal structures in the Coil-3 library, except for λ_1 in transition region (i).

The transition rates (defined as inverse lifetimes, obtained via Markov chain analyses) between boxes of χ_1-χ_2 combinations are shown as arrows in Figure 9(a). As mentioned above, transitions along χ_2 are much more frequent than along χ_1. The most frequent transitions occur between boxes 5 and 8, 6 and 9, 4 and 7, and 1 and 7. A consistency check on the two methods (that is, Boltzmann weighting and Markov chains) for calculating the equilibrium probabilities is shown in Figure 9(b). The agreement is within $\lesssim$5%, which provides an error estimate for the calculated equilibrium probabilities.

We have shown that transitions along χ_1 between frequently sampled χ_1-χ_2 combinations are correlated with elevated values of λ_1 and λ_2, but does the converse also hold? To address this question we constrained the sampling of λ_1 and λ_2 so that either $\langle\lambda_1\rangle > \bar{\lambda}_1$ or $\langle\lambda_2\rangle > \bar{\lambda}_2$, or both. As shown in Figure 10, many of the transitions along χ_1 are increased with larger values of $\langle\lambda_1\rangle$ and $\langle\lambda_2\rangle$. In particular, transitions between boxes 5 and 6 increase with elevated $\langle\lambda_1\rangle$ and between boxes 2 and 3 with elevated $\langle\lambda_2\rangle$. These results are consistent with the transition region $\lambda_{1,2}$ distributions seen in Figure 7.

DISCUSSION

We present the results of a computational study whose goals were to determine the equilibrium distribution of the side-chain dihedral angles χ_1 and χ_2 for the amino acids Leu and Ile and also to identify the mechanism by which the inter-conversions between side-chain dihedral angle conformations occur. It is, of course, desirable to compare our computational results with experimental data. However, the data on side-chain dihedral angle distributions is limited and by no means definitive. Experimental data on the mechanism of inter-conversion between different side-chain conformations, to our knowledge, does not exist.

There are two main types of data with which we can compare the results of our calculations: data from high-resolution crystal structures of proteins and data from NMR studies of proteins in solution. Various researchers have curated "coil libraries" from the PDB in an effort to remove the influences of stretches of secondary structure and 'average out' local environmental differences. Although different groups use similar methods to curate such coil libraries, the results can be quite different. This is especially true for estimates of the relative populations of different side-chain conformations - the data we are most interested in. As an illustration of this, in the 'Wu coil-3' library,[28] analysis of the relative proportions of different side-chain conformations of Leu residues with α-helical backbone ϕ-ψ values, gives the populations of

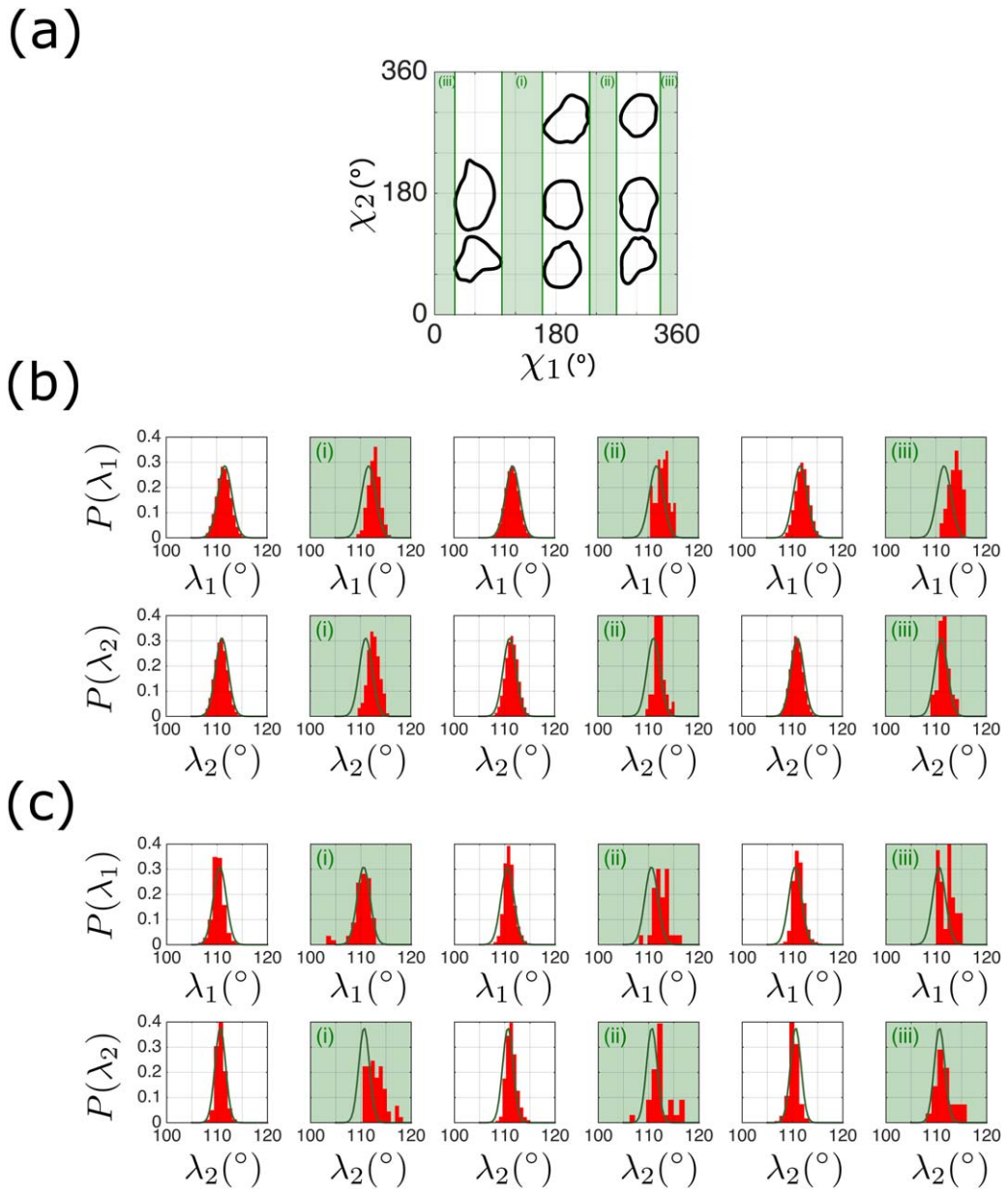

Figure 8

(**a**) Definitions of the transition (green shaded regions i, ii, and iii) and frequently sampled (unshaded) vertical bands of the side-chain dihedral angle χ_1-χ_2) space for Ile. The black solid lines provide an outline of the area in χ_1-χ_2 space that includes 90% of the conformations sampled during the simulations, which occur mostly in the frequently sampled unshaded bands. The vertical lines that divide the transition and frequently sampled bands are at $\chi_1 = 30°$, 95°, 165°, 230°, 270°, and 330°. (b)-(c) Probability distributions $P(\lambda_1)$ (top row) and $P(\lambda_2)$ (bottom row) for the side-chain bond angles λ_1 and λ_2 (red histograms) for Ile for the three transition and three frequently sampled bands defined in (a). The solid green lines show the (b) bond angle distributions from Boltzmann weighting the bond-angle potential $V_{BA}(\lambda_{1,2})$ or (c) the bond angle distributions from the full Coil-3 database. The distributions $P(\lambda_1)$ and $P(\lambda_2)$ are obtained from (**b**) Langevin Dynamics simulations of the hard-sphere model for the Ile dipeptide mimetic and (**c**) Ile residues extracted from the Wu Coil-3 database of protein crystal structures.[28] Each $P(\lambda_1)$ and $P(\lambda_2)$ is normalized such that $\int P(\lambda_{1,2})d\lambda_{1,2} = 1$. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

χ_1 as 24% in *t* and 75% in g-/m (Table I); whereas in the coil library of Swindells *et al.*[43] for Leu residues with α-helical backbone ϕ-ψ values, the same analysis gives the populations of χ_1 as 48% in *t* and 51% in g-/m (Table I). Our prediction for Leu in an α-helical conformation is much closer to the result from the Swindells coil library, as we predict 54% in *t* and 46% in g-/m (Table I). In contrast, our prediction for the relative proportions of different side-chain conformations of Ile residues with α-helical backbone values is much closer to

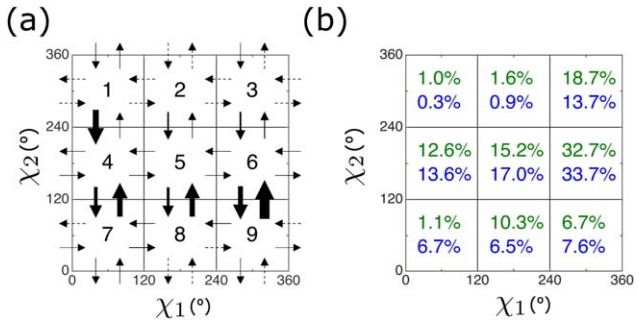

Figure 9

(**a**) The transition rates (inverse lifetimes) for Ile between side-chain conformations in different χ_1-χ_2 boxes obtained from Markov chain analyses are pictured as arrows with widths that are given by the natural logarithm of the transition rate. The transition rate scale ranges from a maximum of 10^{-4} to a minimum of 10^{-7} (in units of t_0^{-1}). The dotted arrows indicate transitions with rates that are below the measurement threshold of the simulations. (**b**) Comparison of the equilibrium probabilities P_j^{eq} for side-chain conformations in boxes $j = 1, \ldots, 9$ for Leu obtained from the Markov chain analysis of the rates in (a) (blue, dark) and Boltzman weighting at each χ_1-χ_2 combination (green, green) (as given Fig. 7(b)). The probabilities agree to within 5%. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

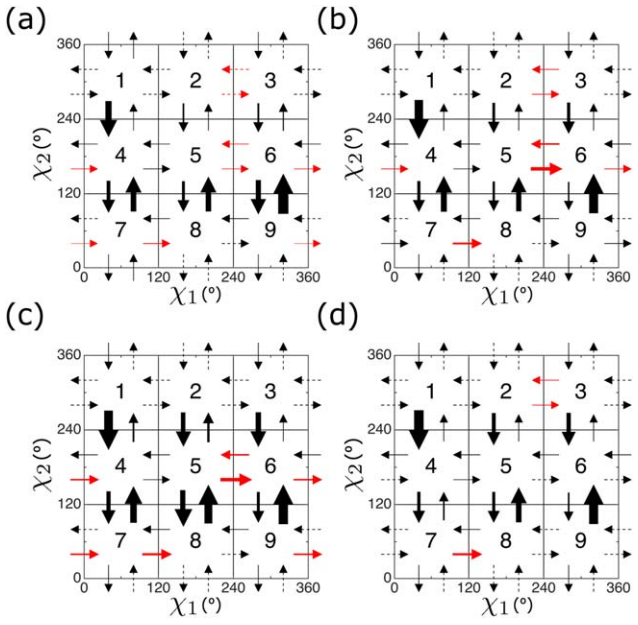

Figure 10

Transition rates (inverse lifetimes) for Ile between side-chain conformations in different χ_1-χ_2 boxes obtained from Markov chain analyses of Langevin Dynamics simulations for several different constrained side-chain backbone angle distributions $P(\lambda)$: (**a**) unconstrained simulations with averages $\langle\lambda_1\rangle = \bar{\lambda}_1$ and $\langle\lambda_2\rangle = \bar{\lambda}_2$; constrained simulations with (**b**) $\langle\lambda_1\rangle = \lambda_0 > \bar{\lambda}_1$ and $\langle\lambda_2\rangle = \bar{\lambda}_2$, (**c**) $\langle\lambda_1\rangle = \bar{\lambda}_1$ and $\langle\lambda_2\rangle = \lambda_0 > \bar{\lambda}_2$, and (**d**) $\langle\lambda_1\rangle = \lambda_0 > \bar{\lambda}_1$ and $\langle\lambda_2\rangle = \lambda_0 > \bar{\lambda}_2$, where $\lambda_0 = 114.5°$. The rates are pictured as arrows with widths that are given by the natural logarithm of the transition rate with maximum 10^{-4} and minimum of 10^{-7} (in units of t_0^{-1}). The dotted arrows indicate transitions with rates that are below the measurement threshold of the simulations. The red (light) arrows in (b)-(d) indicate several of the important transition rates from the constrained simulations that are different from those in (a). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

the 'Wu coil-3' library, as we predict χ_1 populations of 31% for *t* and 60% for g-/m (Table I); and 'Wu coil-3' predicts 16% and 64% for *t* and g-/m respectively (Table I). In this case, Swindells predicts 47% and 14% for *t* and g-/m (Table I). See Table I and Supporting Information Figures S10 and S11 for a more detailed comparison between our predictions and the coil library results.

Alternatively, researchers seek to measure side-chain dihedral angle conformational preferences using NMR methods, with the advantages that the proteins are in solution and it may be possible to collect data on both the folded and denatured protein. Doubly ^{15}N and ^{13}C labeled proteins allow intra-residue homo- and heteronuclear coupling constants $^3J(C', C^\gamma)$ and $^3J(N, C^\gamma)$ to be measured, from which side-chain dihedral angles can be calculated.[5] Alternatively researchers make use of the relationship between ^{13}C chemical shifts and the $^3J(C^\delta, C^\alpha)$ coupling constant to estimate the side-chain dihedral angle from ^{13}C chemical shifts.[7–9]

In these experiments, side-chain dihedral angle conformations for proteins in solution are usually determined from these or other J-couplings by using the Karplus equation.[48] Chou *et al.*[48] have pioneered the use of these backbone and side-chain dipolar couplings in small proteins to derive a consistent set of Karplus parameters. Side-chain dihedral angle distributions measured in this fashion are the experimental data that would be needed to validate or falsify our predicted distributions. However, to date, the NMR methods to determine side-chain dihedral angles have limitations and the results are not sufficiently clear-cut to allow a meaningful comparison. For instance, the correlation between the Leu C^δ ^{13}C chemical shifts and the coupling constant shows strong correlation for some proteins (glycophorin),[11] but a weak correlation for others (staphylococcal nuclease).[13] Another limitation is the extremely small amount of NMR data on side-chain conformational preferences, as seen for example in an early widely cited article that attempted to use coupling constants to determine the χ_1 populations in the denatured state of Hen Egg White Lysozyme (^{13}C, ^{15}N labeled protein in 8M urea pH 2).[5] This protein contains 12 Leu residues of which only six could be assigned. The handful of other proteins whose side-chain dihedral angle distributions have been studied in this fashion contain similarly small numbers of Leu residues. Rapid inter-conversion between conformers is assumed, and measured coupling constants are converted into fractional occupancies as population weighted averages. Thus, we are still lacking a sufficiently reliable experimental grasp on measuring transition rates between different side-chain conformations. It would be

Table I
Comparison between observed and predicted χ_1 conformations, in percentages, as extracted from hard-sphere langevin dynamics (LD) simulations and two different coil libraries: Thornton and colleagues (a/coil and b/coil)[43] and Wu and colleagues (Coil-3).[28]

	g+/p	*t*	g−/m
Leu (α) LD	0.1	54.0	45.9
Leu (α) a/coil	0.7	47.9	51.4
Leu (α) Coil-3	1.3	23.7	75
Leu (β) LD	5.3	45	49.7
Leu (β) b/coil	2.5	33.7	63.8
Leu (β) Coil-3	0.5	24.2	75.3
Ile (α) LD	9.0	30.8	60.2
Ile (α) a/coil	38.8	46.7	13.5
Ile (α) Coil-3	19.6	15.5	63.9
Ile (β) LD	24.2	20.9	54.9
Ile (β) b/coil	39.4	10.1	50.5
Ile (β) Coil-3	19.9	9.6	70.5

χ_1 probabilities are presented as a function of different backbone conformations in coil regions of proteins. We label the different χ_1 states based on two different nomenclatures, g+/*t*/g-[50] and p/*t*/m[32]; where g+/p = 0° < χ_1 < 120°, *t* = 120° < χ_1 < 240°, and g-/m = 240° < χ_1 < 360°.

very interesting to relate the transition probabilities between rotameric states observed in our simulations to some experimental order parameters, like spin relaxation measurements, but this is still a challenge.[49] Another example is glycophorin,[11] where the abundance in the *t* state of each of its eight Leu residues in the trans conformation of χ_1 is measured. The measured abundances range from 29% to 68%. Similarly, for the five Leu residues of Fyn SH3 domain,[9] the fraction of each residue in the *t* conformation in the unfolded state varies from 62% to 68%, and in the folded state from 58% to 85%.

In summary, the current lack of experimental data on Leu and Ile side-chain conformational preferences prevents a definitive test of the predicted distributions $P(\chi_1, \chi_2)$ of the side-chain dihedral angles for Leu and Ile from the hard-sphere dipeptide model. For our studies of the main-chain dihedral angles and inter-conversions between them,[36] there is much more experimental data and our results from the hard-sphere dipeptide model match the data well. Thus, we encourage additional experimental studies of the side-chain dihedral angle distributions for Leu and Ile so that our predictions from the hard-sphere dipeptide model can be quantitatively tested.

CONCLUSION

A fundamental understanding of the backbone and side-chain conformational preferences of amino acids would dramatically improve protein structure prediction and design. To address this issue, we performed Langevin Dynamics simulations of leucine and isoleucine dipeptide mimetics using a minimal model that includes only stereochemical constraints and repulsive hard-sphere Lennard-Jones interactions between non-bonded atoms. We have previously shown that such a model predicts backbone conformation probabilities that are consistent with both random coil libraries and experimental data on short peptides. In this study, we build upon this validated method to predict relative populations of side-chain dihedral angle combinations, for which there is not yet definitive experimental data. In addition, our studies reveal the mechanism of conversion between side-chain conformations.

ACKNOWLEDGMENTS

This work was supported in part by the facilities and staff of the Yale University Faculty of Arts and Sciences High Performance Computing Center.

REFERENCES

1. Kidera A, Konishi Y, Oka M, Ooi T, Scheraga. HA. Statistical analysis of the physical properties of the 20 naturally occurring amino acids. J Protein Chem 1985;4:23–55.
2. Ramakrishnan C, Ramachandran. GN. Stereochemical criteria for polypeptide and protein chain conformations. Biophys J 1965;5: 909–933.
3. Janin J, Wodak S, Levitt M, Maigret. B. Conformation of amino acid side-chains in proteins. J Mol Biol 1978;125:357–386.
4. Scheraga. HA. On the dominance of short-range interactions in polypeptides and proteins. Pure Appl Chem 1973;36:1–8.
5. Hennig M, Bermel W, Spencer A, Dobson CM, Smith LJ, Schwalbe. H. Side-chain conformations in an unfolded protein: χ_1 distributions in denatured hen lysozyme determined by heteronuclear ^{13}C, ^{15}N NMR spectroscopy. J Mol Biol 1999;288: 705–723.
6. Butterfoss GL, Richardson JS, Hermans J. Protein imperfections: separating intrinsic from extrinsic variation of torsion angles. Acta Cryst 2005;D61:88–98.
7. London RE, Wingad BD, Mueller. GA. Dependence of amino acid side chain $^{1}3C$ shifts on dihedral angle: application to conformational analysis. J Am Chem Soc 2008;130:11097–11105.
8. Hansen DF, Neudecker P, Kay. LE. Determination of Isoleucine side-chain conformations in ground and excited states of proteins from chemical shifts. J Am Chem Soc 2010;132:7589–7591.
9. Hansen DF, Neudecker P, Vallurupalli P, Mulder FAA, Kay. LE. Determination of Leu side-chain conformations in excited protein states by NMR relaxation dispersion. J Am Chem Soc 2010;132:42–43.
10. Vajpai N, Gentner M, Huang JR, Blackledge M, Grzesiek. S. Side-chain χ_1 conformations in urea-denatured ubiquitin and protein G from ^{3}J coupling constants and residual dipolar couplings. J Am Chem Soc 2010;132:3196–3203.
11. MacKenzie KR, Prestegard JH, Engelman. DM. Leucine side-chain rotamers in a glycophorin A transmembrane peptide as revealed by three-bond carbon-carbon couplings and $^{1}3C$ chemical shifts. J Biomol NMR, 1996;7:256–260.
12. Kjaergaard M, Iesmantavicius V, Poulsen. FM. The interplay between transient α-helix formation and side chain rotamer distributions in disordered proteins probed by methyl chemical shifts. Prot Sci 2011;20:2023–2034.
13. Nicholson LK, Kay LE, Baldisseri DM, Arango J, Young PE, Bax A, Torchia DA, Dynamics of methyl groups in proteins as studied by proton-detected ^{13}C NMR spectroscopy. application to the Leucine

residues of staphylococcal nuclease. Biochemistry 1992;31:5253–5263.
14. Dill. KA. Dominant forces in protein folding. Biochemistry 1990;29: 7133–7155.
15. Lee C, Levitt. M. Accurate prediction of the stability and activity effects of site-directed mutagenesis on a protein core. Nature 1991; 352:448–451.
16. Eriksson AE, Baase WA, Zhang XJ, Heinz DW, Blaber M, Baldwin EP, Matthews. BW. Response of a protein structure to cavity-creating mutations and its relation to the hydrophobic effect. Science 1992;255:178–184.
17. Petrella RJ, Karplus. M. The energetics of off-rotamer protein side-chain conformations. J Mol Biol 2001;312:1161–1175.
18. McGregor MJ, Islam SA, Sternberg. MJ. Analysis of the relationship between side-chain conformation and secondary structure in globular proteins. J Mol Biol 1987;198:295–310.
19. Fleishman SJ, Whitehead TA, Ekiert DC, Dreyfus C, Corn JE, Strauch E, Wilson IA, Baker. D. Computational design of proteins targeting the conserved stem region of influenza hemagglutinin. Science 2011;332:816–821.
20. Cohen M, Reichmann D, Neuvirth H, Schreiber. G. Similar chemistry, but different bond preferences in inter versus intra-protein interactions. Proteins 2008;72:741–753.
21. Dunbrack RL, Karplus M. Conformational analysis of the backbone-dependent rotamer preferences of protein sidechains. Nat Struct Mol Biol 1994;1:334–340.
22. Ponder JW, Richards. FM. Tertiary templates for proteins: use of packing criteria in the enumeration of allowed sequences for different structural classes. J Mol Biol 1987;193:775–791.
23. Dill KA, Ozkan SB, Shell MS, Weikl. TR. The protein folding problem. Annu Rev Biophys 2008;37:289–316.
24. Kobea B, Kajava. AV. The leucine-rich repeat as a protein recognition motif. Curr Op Struct Biol 2001;11:725–732.
25. Berntsen KRM, Vriend. G. Anomalies in the refinement of isoleucine. Acta Cryst 2014;D70:1037–1049.
26. Zhou AQ, Caballero D, O'Hern CS, Regan. L. New insights into the interdependence between amino acid stereochemistry and protein structure. Biophys J 2013;105:2403–2411.
27. Zhou AQ, O'Hern CS, Regan. L. Predicting the side-chain dihedral angle distributions of non-polar, aromatic, and polar amino acids using hard-sphere models. Proteins 2014;82:2574–2584.
28. Jiang F, Han W, Wu YD. The intrinsic conformational features of amino acids from a protein coil library and their applications in force field development. Phys Chem Chem Phys 2013;15:3413–3428.
29. Wang G, Dunbrack RL, Jr. PISCES: a protein sequence culling server. Bioinformatics 2003;19:1589–1591.
30. Wang G, Dunbrack RL, Jr. PISCES: recent improvements to a PDB sequence culling server. Nucleic Acids Res 2005;33(Web Server issue):W94–W98.
31. Shapovalov MS, Dunbrack RL. A smoothed backbone-dependent rotamer library for proteins derived from adaptive kernel density estimates and regressions. Structure 2011;19:844–858.
32. Lovell SC, Word JM, Richardson JS, Richardson DC. The penultimate rotamer library. Proteins 2000;40:389–408.
33. Lovell SC, Davis IW, Arendall WB, III, de Bakker PIW, Word JM, Prisant MG, Richardson JS, Richardson DC. Structure validation by cα geometry: ϕ, ψ, and C_β deviation. Proteins 2003;50:437–450.
34. Zhou AQ, O'Hern CS, Regan L. Revisiting the Ramachandran plot from a new angle. Prot Sci 2011;20:1166–1171.
35. Zhou AQ, O'Hern CS, Regan L. The power of hard-sphere models: explaining side-chain dihedral angle distributions of Thr and Val. Biophys J 2012;102:2345–2352.
36. Caballero D, Määttä J, Zhou AQ, Sammalkorpi M, Regan L, O'Hern. CS. Intrinsic α-helical and β-sheet conformational preferences: a computational case study of alanine. Prot Sci 2014;23: 970–980.
37. Word JM, Lovell SC, Richardson JS, Richardson. DC. Asparagine and Glutamine: using hydrogen atom contacts in the choice of sidechain amide orientation. J Mol Biol 1999;285:1735–1747.
38. Weeks JD, Chandler D, Andersen. HC. Role of repulsive forces in determining the equilibrium structure of simple liquids. J Chem Phys 1971;54:5237–5247.
39. Allen MP, Tildesley DJ. Computer Simulations of Liquids. New York: Oxford University Press; 1987.
40. Swope WC, Pitera. JW. Describing protein folding kinetics by molecular dynamics simulations. 1. Theory. J Phys Chem B 2004; 108:6571–6581.
41. Swope WC, Pitera JW. Describing protein folding kinetics by molecular dynamics simulations. 2. Example. applications to alanine dipeptide and a β-hairpin peptide. J Phys Chem B 2004;108:6582–6594.
42. Fitzkee NC, Fleming PJ, Rose GD. The protein coil library: a structural database of nonhelix, nonstrand fragments derived from the PDB. Proteins 2005;58:852–854.
43. Swindell MB, MacArthur MW, Thornton JM. Intrinsic ϕ,ψ propensities of amino acids, derived from the coil regions of known structures. Nat Struct Biol 1995;2:596–603.
44. Cote Y, Maisuradze GG, Delarue P, Scheraga HA, Senet P. New insights into protein (un)folding dynamics. J Phys Chem Lett 2015; 6:1082–1086.
45. Chakrabarti P, Pal D. Main-chain conformational features at different conformations ofthe side-chains in proteins. Method Enzymol 1998;11:631–647.
46. Chakrabarti P, Pal D. The interrelationships of side-chain and main-chain conformations in proteins. Prog Biophys Mol Biol 2001;76:1–102.
47. Jiang F, Han W, Wu YD. Influence of side-chain conformations on local conformational features of amino acids and implication for force field development. J Phys Chem B 2010;114:5840–5850.
48. Chou JJ, Case DA, Bax A. Insights into the mobility of methyl-bearing side chains in proteins from 3jcc and 3jcn couplings. J Am Chem Soc 2003;125:8959–8966.
49. Hu H, Hermans J, Lee AL. Relating side chain mobility in proteins to rotameric transitions: insights from molecular dynamics simulations and nmr. J Biomol NMR 2005;32:151–162.
50. Markley JL, Bax A, Arata Y, Hilbers CW, Kaptein R, Sykes BD, Wright PE, Wüthrich. K. Recommendations for the presentation of NMR structures of proteins and nucleic acids. J Mol Biol 1998;280: 933–952.