Revisiting the Ramachandran plot from a new angle

Alice Qinhua Zhou,[1] Corey S. O'Hern,[2,3] and Lynne Regan[1,4]*

[1]Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT, USA
[2]Department of Mechanical Engineering and Materials Science, Yale University, New Haven, CT, USA
[3]Department of Physics, Yale University, New Haven, CT, USA
[4]Department of Chemistry, Yale University, New Haven, CT, USA

Received 11 March 2011; Revised 16 April 2011; Accepted 18 April 2011
DOI: 10.1002/pro.644
Published online 27 April 2011 proteinscience.org

Abstract: The pioneering work of Ramachandran and colleagues emphasized the dominance of steric constraints in specifying the structure of polypeptides. The ubiquitous Ramachandran plot of backbone dihedral angles (ϕ and ψ) defined the allowed regions of conformational space. These predictions were subsequently confirmed in proteins of known structure. Ramachandran and colleagues also investigated the influence of the backbone angle τ on the distribution of allowed ϕ/ψ combinations. The "bridge region" ($\phi \leq 0°$ and $-20° \leq \psi \leq 40°$) was predicted to be particularly sensitive to the value of τ. Here we present an analysis of the distribution of ϕ/ψ angles in 850 non-homologous proteins whose structures are known to a resolution of 1.7 Å or less and sidechain B-factor less than 30 Å^2. We show that the distribution of ϕ/ψ angles for all 87,000 residues in these proteins shows the same dependence on τ as predicted by Ramachandran and colleagues. Our results are important because they make clear that steric constraints alone are sufficient to explain the backbone dihedral angle distributions observed in proteins. Contrary to recent suggestions, no additional energetic contributions, such as hydrogen bonding, need be invoked.

Keywords: Ramachandran plot; backbone conformation prediction; steric constraint; hydrogen bonding

Introduction

The "Ramachandran plot" is an iconic image of modern biochemistry. In the late 1950s and early 1960s, Ramachandran and colleagues investigated the inter-atomic separations between nonbonded atoms in crystal structures of amino acids and related compounds.[1,2] For different types of atom pairs, for example between C and C, C and O, and so on, they specified two sets of allowed inter-atomic separations, the "normally allowed" and a smaller, "outer limit." Subsequently, they assessed all possible combinations of backbone ϕ, ψ angles for an alanyl dipeptide mimetic (*N*-acetyl-L alanine-methylester) (Fig. 1), and identified those ϕ/ψ combinations that are consistent with allowed inter-atomic separations (where ϕ is the dihedral angle defined by rotation around the N-C_α bond of the backbone atoms C′-N-C_α-C′, and ψ is the dihedral angle defined by rotation about the C_α-C′ bond involving the backbone atoms N-C_α-C′-N). Plotting the allowed ϕ/ψ combinations yields Ramachandran plots, which are typically made for both the normal and outer limits.

In this article, we will investigate how the main chain angle τ, which is defined by the backbone bond angle C′-C_α-N, affects the conformation of a

Abbreviations: Ala dipeptide, N-acetyl-L-alanine-methylester; C′ carboxyl carbon

Grant sponsor: National Science Foundation; Grant numbers: DMR-1006537, PHY-1019147; Grant sponsor: The Raymond and Beverly Sackler Institute for Biological, Physical and Engineering Sciences.

*Correspondence to: Lynne Regan, 266 Whitney Avenue, Bass Center, Room 322, New Haven, CT 06511.
E-mail: lynne.regan@yale.edu

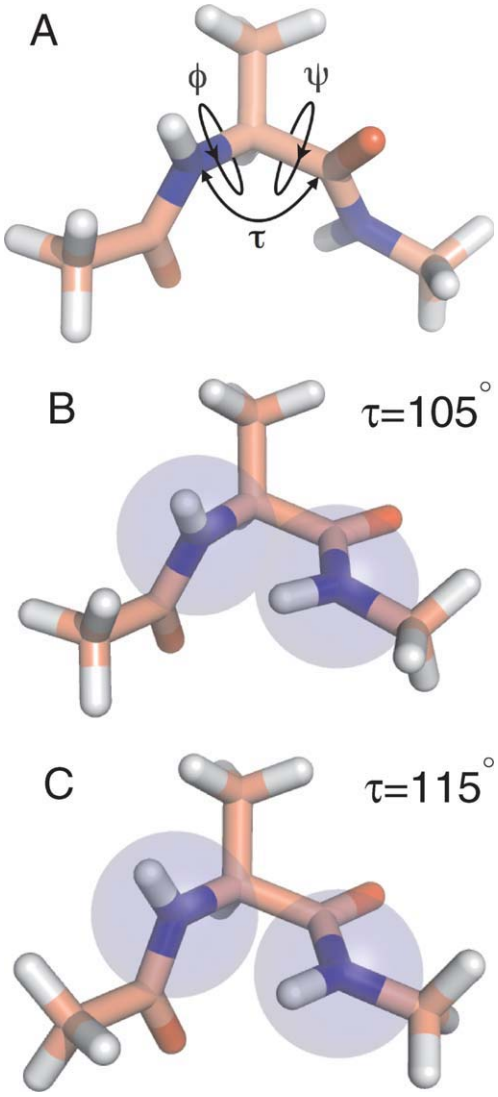

Figure 1. Stick representation of alanyl dipeptide mimetics. Atom types are color-coded: carbon = pink, nitrogen = blue, oxygen = red, hydrogen = white. A: The backbone dihedral angles ϕ and ψ and the bond angle τ are indicated. B: $\tau = 105°$, $\phi = -90°$, $\psi = 0°$ (i.e., bridge region values of ϕ and ψ). Blue-shaded spheres indicate steric overlap between main-chain nitrogens for this value of τ. C: $\tau = 115°$, $\phi = -90°$, $\psi = 0°$ (i.e., bridge region values of ϕ and ψ). Blue-shaded spheres indicate no steric overlap between main-chain nitrogens for this value of τ. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

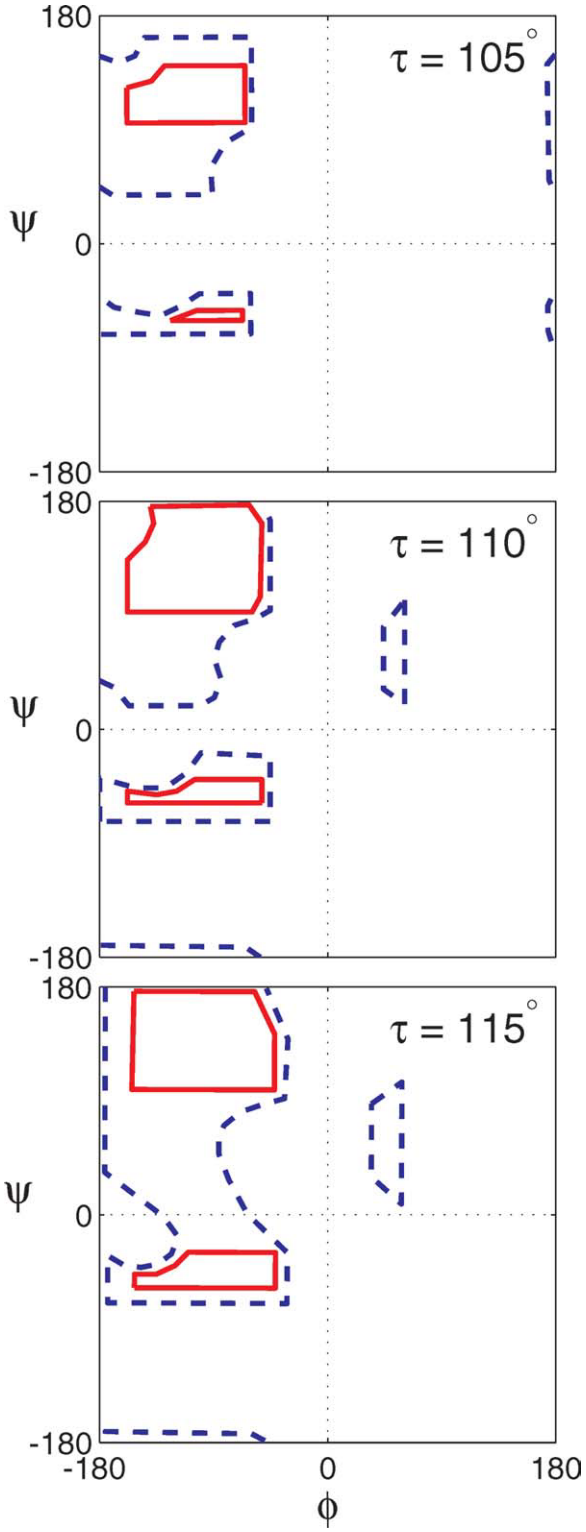

Figure 2. Calculated Ramachandran Plots. Ramachandran plots of allowed ϕ/ψ combinations for three values of τ.[2] The solid red lines enclose the "normally allowed" ϕ/ψ combinations and the dashed blue line indicates the "outer limit". [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

peptide backbone. For an ideal tetrahedral sp^3 carbon, $\tau = 109.5°$ (Fig. 1). Ramachandran and colleagues realized that the allowed combinations of ϕ and ψ angles in a peptide backbone are influenced by the value of τ, and indeed they published plots showing this dependence for the Ala dipeptide.[1,2] Thus, in fact, there are many Ramachandran plots because the allowed regions of ϕ and ψ depend on the value of τ for which the map is calculated (Fig. 2). The crystal structures of proteins confirmed that the ϕ/ψ combinations predicted by Ramachandran, for an "average" value of $\tau = 110°$, were indeed those populated by amino acids within proteins.[2]

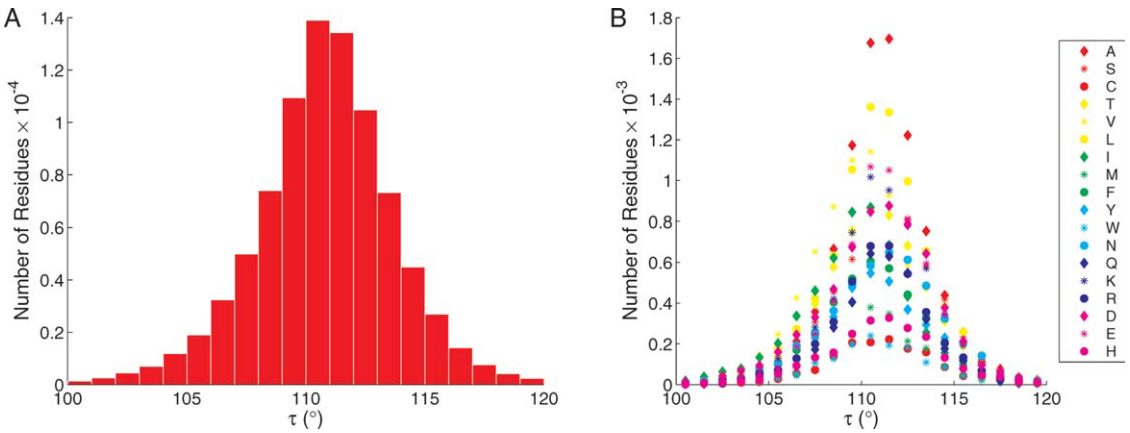

Figure 3. Distribution of the bond angle τ. A: Distribution of τ for all 86,299 residues in the Dunbrack data base (excluding Gly and Pro). Number of residues plotted against the indicated τ ranges. B: Distribution of τ for each type of residue in the Dunbrack data base (excluding Gly and Pro). The residue types are identified using the single letter code. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Nowadays, Ramachandran plots are the "gold standard" against which new crystal structures are evaluated.[3] The remarkable finding of Ramachandran et al. is that they were able to predict the ϕ and ψ dihedral angles of known protein structures without considering electrostatic, solvent-mediated or any other interactions.

Results and Discussion

With the large number of high resolution crystal structures of proteins now available, it is appropriate to revisit the Ramachandran plot, to examine the relationship between allowed ϕ and ψ angles and the backbone bond angle τ. Evidently, this angle can be widened or contracted significantly from the tetrahedral geometry to accommodate various other strains in the structure.[4–6] Figure 3(A) shows a histogram of the values of τ for 86,299 residues in 850 nonhomologous proteins, which we will refer to as the Dunbrack database.[7] The distribution is centered on $\tau = 110.8^\circ$, with a range between 100° and 120° (which includes more than 99% of the data points). Figure 3(B) shows a similar plot, but for each residue individually. It is evident that the distribution of τ is similar for each amino acid. There is no systematic dependence of either the mean value or standard deviation of τ with respect to amino acid type.

For all amino acids in the Dunbrack database, we constructed ϕ/ψ plots for different ranges of τ: 100°–104°, 104°–108°, 108°–112°, 112°–116°, and 116°–120° (Fig. 4). For ease of viewing in Figure 4, we show the scatter plots of ϕ/ψ angles from amino acids in the Dunbrack database overlaid on an average Ramachandran plot.[8] Of particular note are the residues with ϕ/ψ values in the so-called "bridge region" ($\phi \leq 0^\circ$ and $-20^\circ \leq \psi \leq 40^\circ$).[9] It is clear that the fraction of residues with ϕ/ψ angles in the bridge region (F_{bridge}) increases as a function of τ [(Fig. 5(A)]. This increase in F_{bridge} (roughly by a factor of 3) from $\tau = 105^\circ$ to 115° is consistent with the increase in area of the allowed part of the bridge region relative to the area of the total allowed region of the ϕ/ψ map predicted by Ramachandran and colleagues from their hard-sphere models of dipeptides. Figure 5(B) shows a similar plot, but for each residue individually. It is evident that the increase in the fraction of residues with ϕ/ψ values in the bridge region as a function of τ is similar for each amino acid type.

Porter and Rose[9] recently suggested that it might be advantageous to "re-draw the conventional Ramachandran plot by applying a hydrogen-bonding (H-bonding) requirement as an additional energetic criterion." They argued that the ϕ/ψ combinations in the bridge region prevent water from H-bonding with the backbone nitrogen of the neighboring residue. They thus speculated that even though the ϕ/ψ combinations in the bridge region are sterically allowed, the penalty for nitrogen not H-bonding with water excludes residues from occupying the bridge region unless they form intra-peptide H-bonds in the folded protein.

They also noted, however, that in proteins of known structure, many residues are found with ϕ/ψ angles in the bridge region (Fig. 4). They rationalized this apparent contradiction by suggesting that "almost all the 30,924 residues in the disfavored bridge could be classified readily into one of three local hydrogen bonded motifs." In other words, they suggested that the reason that ϕ/ψ angles corresponding to the bridge region are adopted by amino acids in folded proteins is because they are always associated with H-bonding to polar groups other

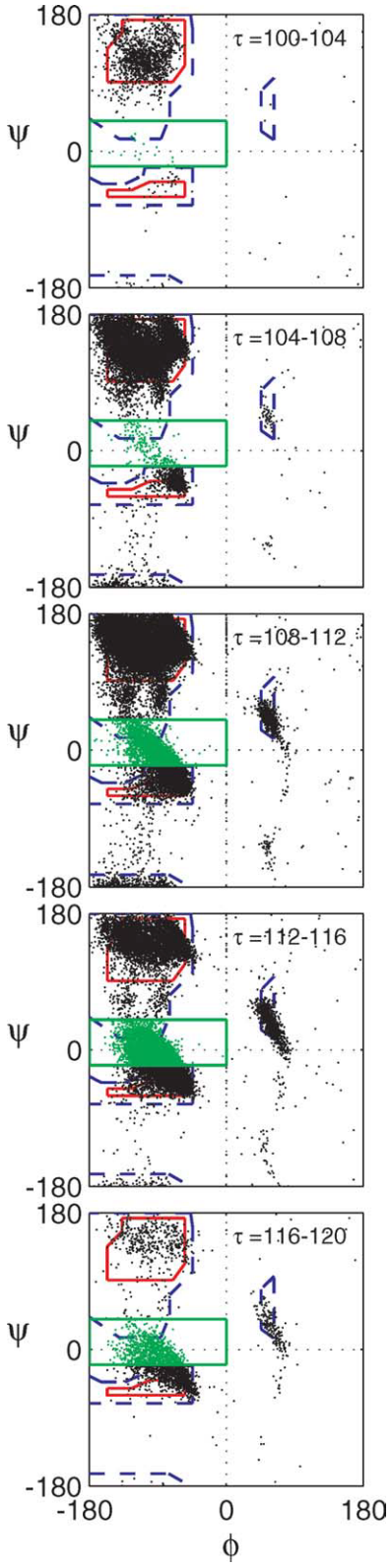

Figure 4. Observed ϕ/ψ combinations for all residues. The observed ϕ/ψ distribution as a function of the indicated ranges of τ for residues in the Dunbrack database (excluding Gly and Pro). The data are overlaid on an average Ramachandran plot.[8] The solid red lines enclose the "normally allowed" ϕ/ψ combinations and the dashed blue line indicates the "outer limit". Residues within the bridge region are colored in green. The bridge region is defined by the area within the solid green lines. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

than water, for example, the carboxyl group on a nearby residue.

In light of our findings concerning the τ dependence of the ϕ/ψ distribution, we chose two different amino acid types: Serine, which is capable of intrapeptide H-bonding, and Leucine, which is not, and tracked the distribution of allowed ϕ/ψ angles as a function of τ for both residue types. These results, shown in Figure 6, make clear that the same trend—increasing τ correlates with increasing percentage of residues with ϕ/ψ angles in the bridge region—applies to both Serine and Leucine equally.

In summary, we have shown that the distribution of backbone dihedral angles observed in proteins of known structure is well explained by Ramachandran and coworker's original analysis of an alanyl dipeptide, where only repulsive hard-sphere interactions together with bond length and angle constraints determine the allowed ϕ/ψ angles. Particularly, the original analysis showed an increase in the region of allowed ϕ/ψ dihedral angles (predominantly in the bridge region) as τ increases. For ϕ/ψ dihedral angles in the bridge region, larger τ relieves the clashes between N and N_{i+1} and N_i and HN_{i+1} [Figs. 1B,C]. Our analysis shows that in proteins of known structure the relationship between the regions of allowed ϕ/ψ dihedral angles and the bond angle τ is predicted by the original calculations of Ramachandran and coworkers. We find no need to invoke additional interactions to explain the backbone conformations of proteins.

Materials and Methods

Protein database

850 high-resolution non-homologous protein structures solved by X-ray crystallography (resolution $\leq$1.7Å, B-factor of sidechains $<30Å^2$, R-factor $\leq$ 0.25, sequence identity $<$50%) were obtained from the Protein Data Bank (PDB) and prepared by R. L. Dunbrack, Jr. as follows: Hydrogen atoms were added to the structures using the REDUCE program.[10] Side chains with atom-atom clashes were either flipped to satisfy hydrogen-bonding requirements or removed by the PROBE program.[11] The placement of the hydrogen atoms does not affect the backbone conformation. The list of PDB chains is available at http://dunbrack.fccc.edu/bbdep/bbdepformat.php (May 2002 version).[7]

Calculations and nomenclature

The ϕ dihedral angle was defined by the clockwise rotation around the N-C bond (viewed from N to C) of the backbone atoms C′-N-C_α-C′. The ψ dihedral angle was defined by the clockwise rotation about the C_α-C′ bond (viewed from C to C′) involving the backbone atoms N-C_α-C′-N. Bridge residues were defined as those with $\phi \leq 0^\circ$ and $-20^\circ \leq \psi \leq 40^\circ$. The main chain angle τ was defined as the bond

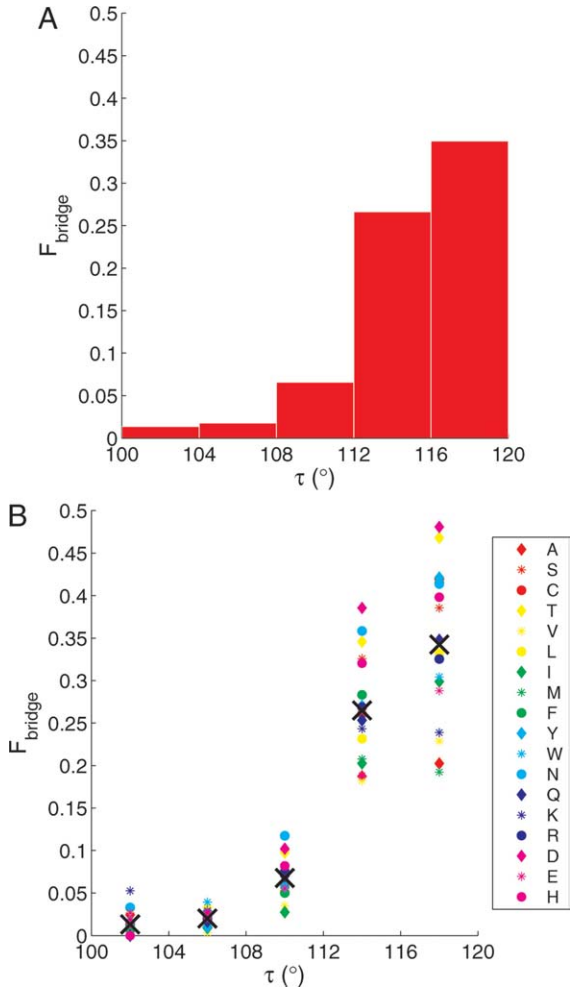

Figure 5. Fraction of residues with ϕ/ψ angles in bridge region. A: Fraction of residues, F_{bridge}, with ϕ/ψ angles in the bridge region as a function of the indicated τ ranges. B: The fraction, F_{bridge}, of each residue type with ϕ/ψ angles in the bridge region as a function of the indicated τ ranges. The residue types are identified using the single letter code. The crosses indicate the average value for each range of τ.

angle between N-C_α-C′. The "average Ramachandran plot" shown in Figures 5 and 6 was taken from the X-PLOR user manual.[8] F_{bridge}, the fraction of residues with ϕ/ψ in the bridge region, is defined by

$$F_{\text{bridge}} = \frac{N_{\text{bridge}}(\tau)}{N(\tau)} \qquad (1)$$

where $N_{\text{bridge}}(\tau)$ is the number of residues with ϕ/ψ angles in the bridge region for a given τ range and $N(\tau)$ is the total number of residues for a given τ range.

We exclude Glycine from all calculations because its lack of a side chain makes the distribution of ϕ/ψ angles significantly different from that of all other amino acid types. We also exclude Proline from all calculations because the pyrrolidine ring essentially fixes ϕ and thus significantly limits the distribution of ϕ/ψ relative to that of all other amino acids.

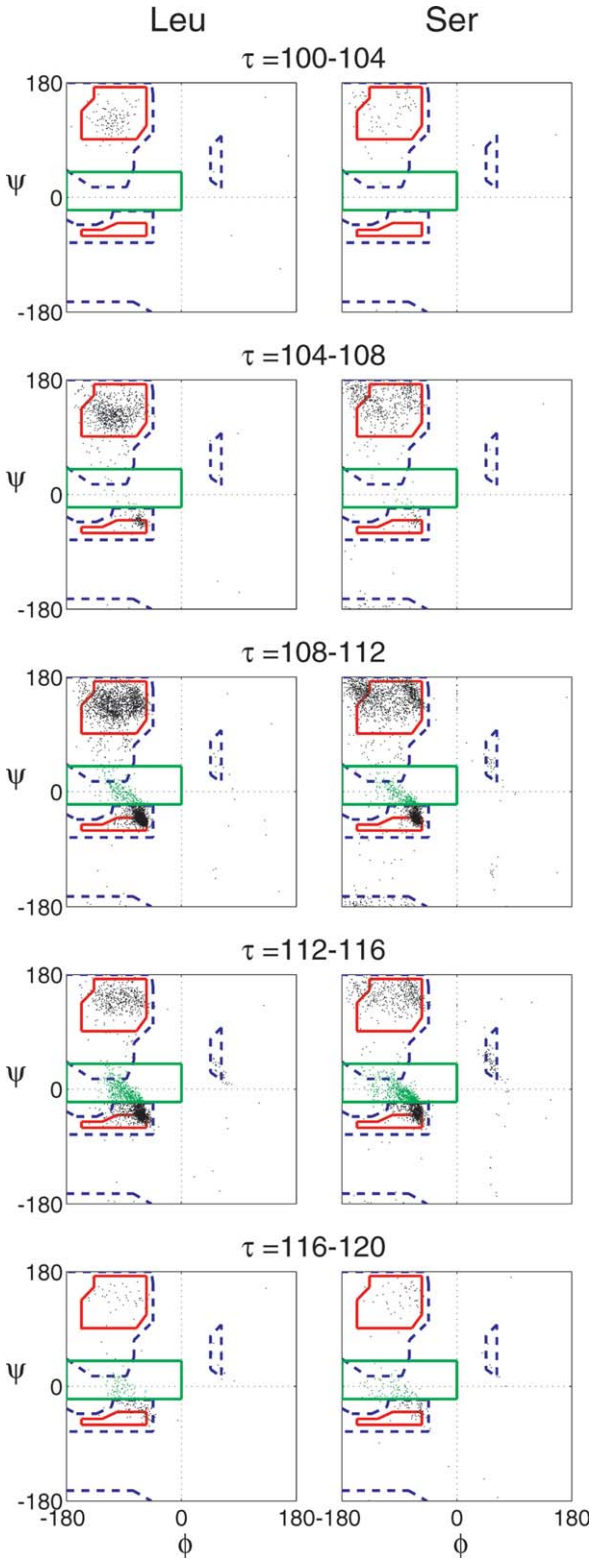

Figure 6. Observed ϕ/ψ combinations for Serine and Leucine. The observed ϕ/ψ distribution as a function of the indicated ranges of τ for all Ser (left column) and all Leu (right column) residues in the Dunbrack database. The data are overlaid on an average Ramachandran plot. The solid red lines enclose the "normally allowed" ϕ/ψ combinations and the dashed blue line indicates the "outer limit." Residues within the bridge region are colored in green. The bridge region is defined by the area within the solid green lines.

Acknowledgments

The authors thank R. L. Dunbrack, Jr. for providing all PDB structures used in this study. They also acknowledge N. Clarke, R. Collins, D. Engelman, T. Grove, R. Ilagan, A. Miranker, S. Mochrie, D. Blaho Noble, E. Spetz, and Y. Xiong for their critical reading of the article.

References

1. Ramakrishnan GN, Ramakrishnan C, Sasisekharan V (1963) Stereochemistry of polypeptide chain configurations. J Mol Biol 7:95–99.
2. Ramakrishnan C, Ramachandran GN (1965) Stereochemical criteria for polypeptide and protein chain conformations. II. Allowed conformations for a pair of peptide units. Biophys J 55:909–933.
3. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Cryst 26: 283–291.
4. Malathy Sony SM, Saraboji K, Sukumar N, Ponnuswamy MN (2006) Role of amino acid properties to determine backbone tau (N-Calpha-C′) stretching angle in peptides and proteins. Biophys Chem 120:24–31.
5. Momany FA, McGuire RF, Burgess AW, Scheraga HA (1975) Energy parameters in polypeptides: VII. Geometric parameters, partial atomic charges, nonbonded interactions, hydrogen bond interactions, and intrinsic torsional potentials for the naturally occurring amino acids. J Phys Chem 79:2361–2381.
6. Jiang X, Cao M, Teppen B, Newton SQ, Schafer LJ (1995) Predictions of protein backbone structural parameters from first principles: systematic comparisons of calculated N–C(α)–C′ angles with high-resolution protein crystallographic results. J Phys Chem 99:10521–10525.
7. Dunbrack RL, Cohen FE (1997) Bayesian statistical analysis of protein side-chain rotamer preferences. Protein Sci 6:1661–1681.
8. Schwieters CD, Kuszewski JJ, Tjandra N, Clore GM (2003) The Xplor-NIH NMR Molecular Structure Determination Package. J Magn Res 160:66–74.
9. Porter LL, Rose GN (2010) Redrawing the Ramachandran plot after inclusion of hydrogen-bonding constraints. PNAS 108:109–113.
10. Word JM, Lovell SC, Richardson JS, Richardson DC (1999) Asparagine and glutamine: using hydrogen atom contacts in the choice of sidechain amide orientation. J Mol Biol 285:1735–1747.
11. Word JM, Lovell SC, LaBean TH, Taylor HC, Zalis ME, Presley BK, Richardson JS, Richardson DC (1999) Visualizing and quantifying molecular goodness-of-fit: small-probe contact dots with explicit hydrogens. J Mol Biol 285:1711–1733.