© 1994 Nature Publishing Group http://www.nature.com/nsmb



# Conformational analysis of the backbone-dependent rotamer preferences of protein sidechains

Roland L. Dunbrack, Jr[1] and Martin Karplus

**Amino acids have sidechain rotamer preferences dependent on the backbone dihedral angles ϕ and ψ. These preferences provide a method for rapid structure prediction which is a significant improvement over backbone-independent rotamer libraries. We demonstrate here that simple arguments based on conformational analysis can account for many of the features of the observed backbone dependence of the sidechain rotamers. Steric repulsions corresponding to the 'butane' and '*syn*-pentane' effects make certain conformers rare, as has been observed experimentally.**

Department of Chemistry, 12 Oxford Street, Harvard University, Cambridge, Massachusetts 02138, USA

[1]Present address: Department of Pharmaceutical Chemistry, University of California, San Francisco, California 94143-0446, USA

Most analyses of protein sidechain conformations have been based on the distribution of sidechain dihedral angles, χ, in proteins of known structure, although calculations with empirical energy functions have also been used for this purpose. From early protein structures it was evident that sidechain dihedral angles tend to cluster around particular values of χ; for example, for tetrahedral carbons, $\chi_1 \cong 60°$, 180°, and -60° (ref. 1). There are a number of rotamer libraries that describe the probable values of sidechain χ angles[2–6]. They however do not consider the possibility of a relation between sidechain χ values and the backbone conformation. The larger data base of accurate protein structures now available, has made possible the determination of a relationship between sidechain and backbone conformation[7–9]. McGregor *et al.*[7] found statistically significant rotamer-preferences for all sidechain types for α-helices, β-sheets, and non-α, non-β conformations. Schrauber *et al.*[9] used circular regions of the ϕ,ψ map centred about positions corresponding to secondary structural elements in a rotamer library. By use of a database of 132 protein structures, we have determined a backbone-dependent rotamer library calculated in 20° × 20° square blocks of the ϕ,ψ map for the purpose of sidechain conformation prediction[8]. In an extended version of the same library[10], making use of 182 proteins, it was shown that $\chi_1$ rotamers could be predicted with 70% accuracy (compared to 55% for a backbone independent library). Prediction using such a library requires a few seconds for a protein of any size in comparison with other methods[11–13] which may take hours on comparable computers.

In this paper we present simple arguments based on the conformational analysis of hydrocarbons[14,15] to explain the backbone-dependent and backbone-independent rotamer preferences of protein sidechains. We show that the steric factors embodied in conformational analysis are sufficient to describe most of the observed trends. The present approach for sidechain rotamers is analogous to the classic analysis of backbone preferences by Ramachandran *et al.*[16]. As in the latter, more detailed calculations based on empirical energy functions can be used to refine the results[2,11,17–20].

## Hydrocarbon conformational analysis

In butane, there is one heavy-atom dihedral angle and the two gauche conformers ($g^+$ and $g^-$, where $g^+ \sim +60°$ and $g^- \sim -60°$) are about 0.8 kcal higher in energy than the trans conformer (t = 180°)[21]; this is the 'butane' effect[22]. In pentane, there are two heavy-atom dihedral angles and six unique conformers; four of these are shown in Fig. 1. The energy of the {t,$g^+$} or {t,$g^-$} conformers of pentane relative to the global energy minimum {t,t} is approximately 0.8–0.9 kcal $mol^{-1}$, equivalent to the butane gauche interaction. If the dihedral angles are both gauche and of the same sign, {$g^+$,$g^+$} and {$g^-$,$g^-$}, the energy is 1.3 kcal $mol^{-1}$ in *ab initio* calculations, relative to {t,t}, due to the presence of two gauche interactions[23]. If both dihedral angles are gauche and of opposite sign, {$g^+$,$g^-$} or {$g^-$,$g^+$}, the energy is approximately 3.3 kcal $mol^{-1}$ above {t,t}[23–25]; this is the '*syn*-pentane' effect[22]. The terminal carbons of pentane are very close together in the exact {+60°,-60°} and {-60°,+60°} conformations, so that the actual dihedral angles of these conformers are expected to deviate significantly from

© 1994 Nature Publishing Group http://www.nature.com/nsmb



these values and are closer to (±70°, ±90°) in *ab initio* calculations[23]. The high energy of the *syn*-pentane conformation thus involves both steric and dihedral angle contributions[27], and extends to a range of ±30° or more about the canonical values of ±60°.

## Sidechain-backbone interactions ($\chi_1$)

The Newman projection[28] of a sidechain with a single γ-carbon in a dipeptide unit (Fig. 2*a*) shows the positions of the three $\chi_1$ rotamers relative to the protein backbone. The $\chi_1$ rotamer populations of residue i are affected by repulsive interactions with the backbone atoms $N_i$ and $C_i$ (butane effect) which are independent of the mainchain dihedral angles, and by repulsive interactions with the backbone atoms $C_{i-1}$, $O_i$, and $N_{i+1}$, as well as the hydrogen bond acceptor of $HN_i$, whose positions depend on φ and ψ (*syn*-pentane effect). The hydrogen acceptor is almost always an oxygen (99.95%) and most of these (70%) are from backbone carbonyls (we use the symbol O---$HN_i$ for the hydrogen bond acceptor). The dihedral angles φ and $\chi_1$ determine the interaction of each Xγ with $C_{i-1}$ and O---$HN_i$, and the angles ψ and $\chi_1$ determine the interaction of Xγ with $O_i$ and $N_{i+1}$ (Fig. 2*a*).

Because of the large unfavourable energy associated with the *syn*-pentane repulsion, certain sidechain rotamers are essentially forbidden over a ±30° to ±50° range about particular values of φ and ψ. Sidechains with a single Xγ heavy atom exhibit such *syn*-pentane interactions with $C_{i-1}$ when $\{\phi,\chi_1\} = \{180°,-60°\}$ and $\{60°,60°\}$, with O---$HN_i$ when $\{\psi,\chi_1\} = \{0°,-60°\}$ and $\{-120°,60°\}$, and with backbone atoms $N_{i+1}$ and $O_i$ when $\{\psi, \chi_1\} = \{0°,180°\}$, and $\{120°,60°\}$ (See Methods). Valine, threonine, and isoleucine are more restricted because of the presence of two γ heavy atoms in each of these sidechains; for example, Val has *syn*-pentane interactions when $\{\phi,\chi_1\} = \{180°,180°\}$, $\{180°,-60°\}$, $\{60°,+60°\}$, $\{60°,-60°\}$,

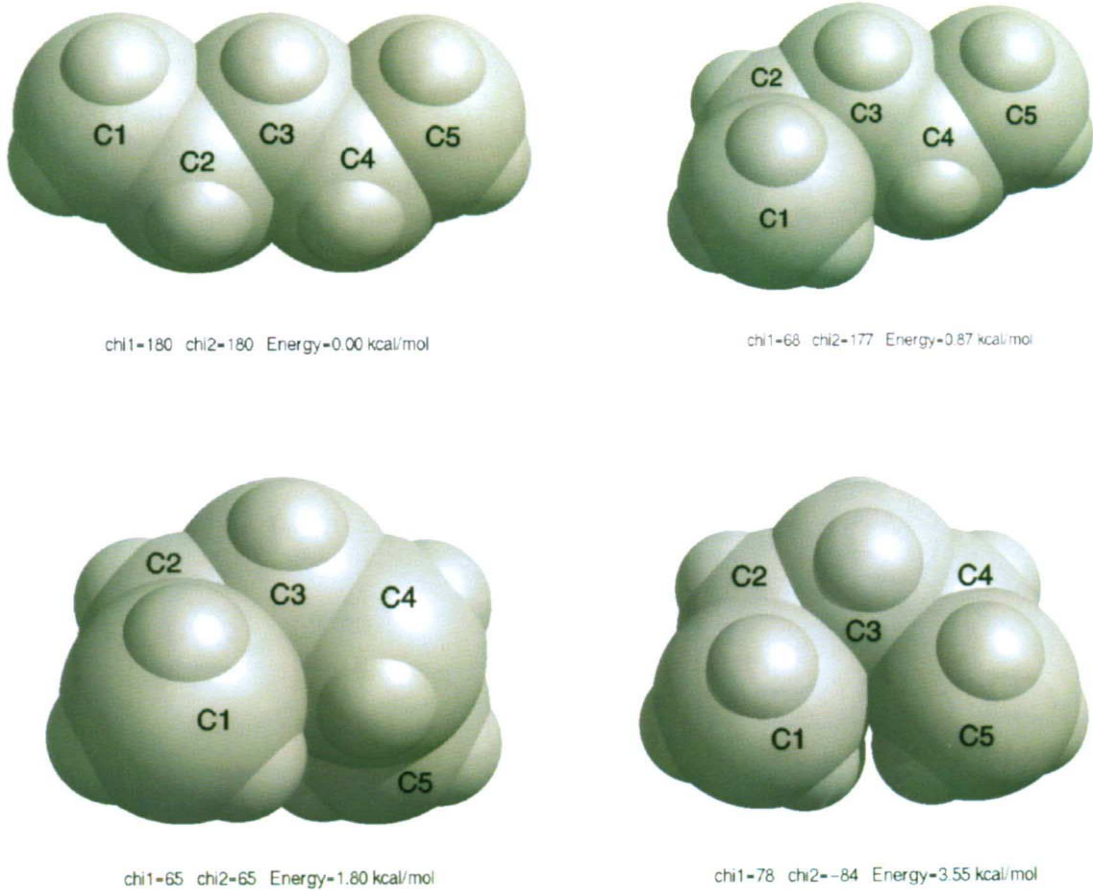


**Fig. 1** Conformations of pentane: *a*, t,t; *b*, $g^+$,t; *c*, $g^+$,$g^+$; and *d*, $g^+$,$g^-$. These structures were derived by energy minimization with the program CHARMM[30], using the recently derived all-hydrogen-atom force field (A. D. MacKerell *et al.*, unpublished results). Values of the heavy-atom dihedral angles ($\chi_1$=C1-C2-C3-C4 and $\chi_2$=C2-C3-C4-C5) and energies of the local minima relative to the global energy minimum structure ({t,t}) from the CHARMM minimizations are shown beneath each structure. The structures were drawn with the MidasPlus[32] software system from the Computer Graphics Laboratory of the University of California, San Francisco.

© 1994 Nature Publishing Group http://www.nature.com/nsmb



{0°, 180°}, {0°, -60°}, {-120°, +60°} and {-120°, -60°} and when {ψ,$\chi_1$} = {180°,+60°}, {180°,180°}, {120°,+60°}, {120°,-60°}, {0°,+60°}, {0°,180°}, {-60°,+60°}, and {-60°,-60°}.

## Maps of $\chi_1$ versus {ϕ,ψ}

The above analysis suggests that certain $\chi_1$ rotamers are unlikely in specific regions of the Ramachandran map. In most cases, one rotamer is excluded by the *syn*-pentane effect and in some cases only one is allowed. Where there are two allowed rotamers, one may predominate because of the butane effect. Maps of the allowed $\chi_1$ values as a function of ϕ and ψ for sidechains with one Xγ (where X is C, O, or S), including aromatics (Fig. 3*a*) and with two Xγ (Val, Ile, and Thr) (Fig. 3*b*) show good overall agreement with results from experimental data for rotamer distributions in proteins (Fig. 4).

The heavy vertical and horizontal lines in the figures separate the regions with different behaviour. For ϕ between -180° and -150°, for example, the *syn*-pentane effect eliminates $g^-$ rotamers due to repulsion between Cγ and $C_{i-1}$. For ϕ between -180° and -150° and ψ near 180°, t rotamers are also excluded due to a *syn*-pentane repulsion between Cγ and $N_{i+1}$, and so $g^+$ predominates. Generally, the probability of $g^-$ is greater than that of $g^+$ when both are allowed, because $g^-$ has only one gauche interaction (with the backbone $N_i$), while $g^+$ is gauche to both $N_i$ and $C_i$ (see Figure 2*b*), regardless of the backbone conformation. When $g^-$ and t are both allowed (for example, {-140° < ϕ < -60°, 90° < ψ < 150°}), they both tend to be well populated.

For -50° < ϕ < 0°, $g^-$ rotamers can have steric interactions with O---$HN_i$ (Fig. 2*a*). There is a *syn*-pentane effect when $\chi_1$ is near -60°, since the dihedral angle involving the oxygen donor, $N_i$, Cα, and Cβ is ~+60° when ϕ=0°, if a linear hydrogen bond is assumed. Thus, there should be fewer $g^-$ rotamers when -50° < ϕ < 0°. This is exemplified by α-helices, where the steric interaction between Cγ and the oxygen hydrogen bonding to $HN_i$ (that is, $O_{i-4}$) has been noted previously[29].

## Comparison with structural data

In almost all cases, the most common rotamer agrees with that predicted by conformational analysis. The only disagreements between Figs 3*a* and 4*a* for the most common rotamer are in the regions {ϕ,ψ} = {-60° to -30°, 150° to 180°}, {-60° to -30°, -30° to +30°}, and {-150° to -120°, -180° to -150°} and between Figs 3*a* and 4*b* in the region {ϕ,ψ} = {-60° to -30°, 150° to 180°}. The conformational analysis predicts $g^+$ rotamers to predominate in the regions where -60° < ϕ < -30° (Fig. 3*a*), while $g^-$ rotamers are found experimentally (Fig. 4*a*, *b*). The data in these regions are sparse, so that the experimental distributions are not well determined. When 10° increments in ϕ and ψ are used, there is a change to $g^+$ rotamers in the blocks where ϕ≥-50°, which are less heavily populated than the ϕ=-60° block. Also, some of the predicted preferences are due to repulsion with the hydrogen bond, which is not expected to be as strong as the interactions with the backbone atoms of residues i-1, i, and i+1.

The relative probabilities of the three rotamers are a result of interactions of each rotamer with the backbone, as a function of ϕ and ψ. These interactions vary continuously with angle and so the sharp categories used in the figures are merely convenient approximations. In Fig. 3*a*,*b* we have drawn vertical lines at -140° and -50° to correspond roughly to the data in a backbone-dependent rotamer library calculated every 10° in ϕ and ψ[10]. In Fig. 4*a*,*b* the data are gathered into 30° increments in ϕ and ψ, and vertical lines are therefore drawn at -150° and -60°. The continuous variation of the distributions is clear from Figs 4*a*,*b*; for example, there is an increase in the proportion of t rotamers and a decrease in the proportion of $g^+$ rotamers as ψ moves from 180° to 120°; a

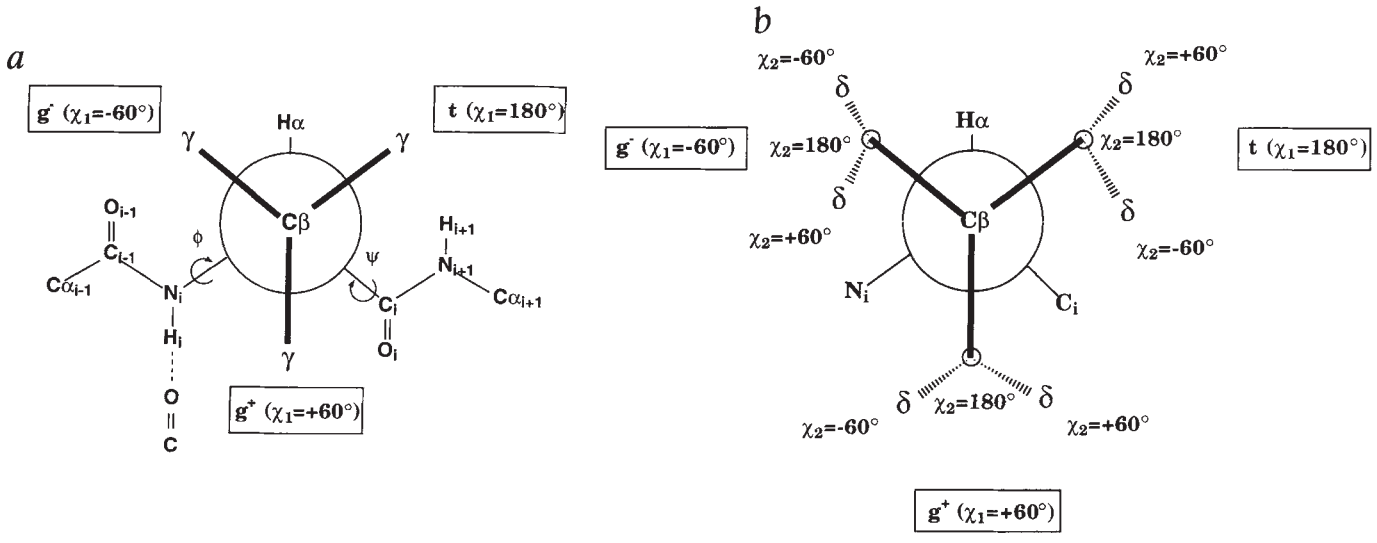


**Fig. 2** Newman projections of *a*, the backbone conformation dependent interactions of protein sidechains with the backbone and *b*, the backbone conformation independent interactions of protein sidechains with the backbone.

© 1994 Nature Publishing Group http://www.nature.com/nsmb



decrease in t rotamers and an increase in $g^+$ rotamers as $\psi$ approaches 0°; and an increase in t rotamers and a decrease in $g^+$ rotamers as $\psi$ approaches -60°. A similar variation occurs for $g^-$ rotamers as a function of $\phi$; that is, there is an increase in $g^-$ as $\phi$ moves from -180° to -90° followed by a decrease in $g^-$ as $\phi$ goes from -90° to 0°. We have calculated the average $\chi_1$ angles in the present database as a function of $\phi$ and $\psi$ for all possible rotamers of each sidechain type (data not shown). The small number of sidechains that have conformations which involve *syn*-pentane interactions between Xγ and the backbone have $\chi_1$ dihedrals that deviate by 20–30° from the canonical values of +60°, 180°, or -60°.

Aromatic sidechains (Fig. 4*b*) behave much like the straight-chain sidechains (Fig. 4*a*), with some small differences in the preferred orientations when two sidechain rotamers are allowed. The remaining single Xγ sidechains (Asp, Asn, Ser and Cys) follow essentially the same scheme (data not shown). Forbidden rotamers (as defined in Fig. 3*a*) are as rare in protein structures for these sidechain types as they are for the straight chain (Fig. 4*a*) and aromatic sidechains (Fig. 4*b*). Preferences among the remaining allowed rotamers for Asp, Asn, Ser, and Cys show somewhat different distributions, because of hydrogen bonding and other electrostatic interactions. This is especially true for serine which shows a marked shift to $g^+$ rotamers, which can form hydrogen bonds to the backbone at certain values of $\phi$ and $\psi$[29].

When there are two Xγ atoms (Fig. 3*b*), only one rotamer is allowed at most values of $\psi$. For example, near $\psi=\pm180°$ and 0° only $g^-$ rotamers for Val ($g^+$ for Ile, Thr) have no strong repulsive interactions with the backbone and are most common (Fig. 4*c*). Near $\psi$=120° and -60° for Val, only t rotamers ($g^-$ for Ile, Thr) are allowed by the steric analysis. Since these regions are also the most populated in {$\phi$,$\psi$} space, Val is most commonly found to be in a t-conformation in backbone-independent rotamer libraries. When $\phi$ is near -180° and $\psi$ is near -180°, -60°, 0°, 120°, or 180°, all three rotamer positions are sterically hindered by *syn*-pentane interactions with the backbone. There are relatively few Val in these regions of {$\phi$,$\psi$} space. They are mostly $g^+$, since both $g^-$ and t have *syn*-pentane interaction with $C_{i-1}$, while $g^+$ has *syn*-pentane interactions with the smaller $O_i$ and $N_{i+1}$. Fig. 4*c* shows excellent agreement with Fig. 3*b*. The results for Thr are quite similar, with small differences due to hydrogen bonding effects[29]. Because these sidechains have only one allowed rotamer in most regions of the

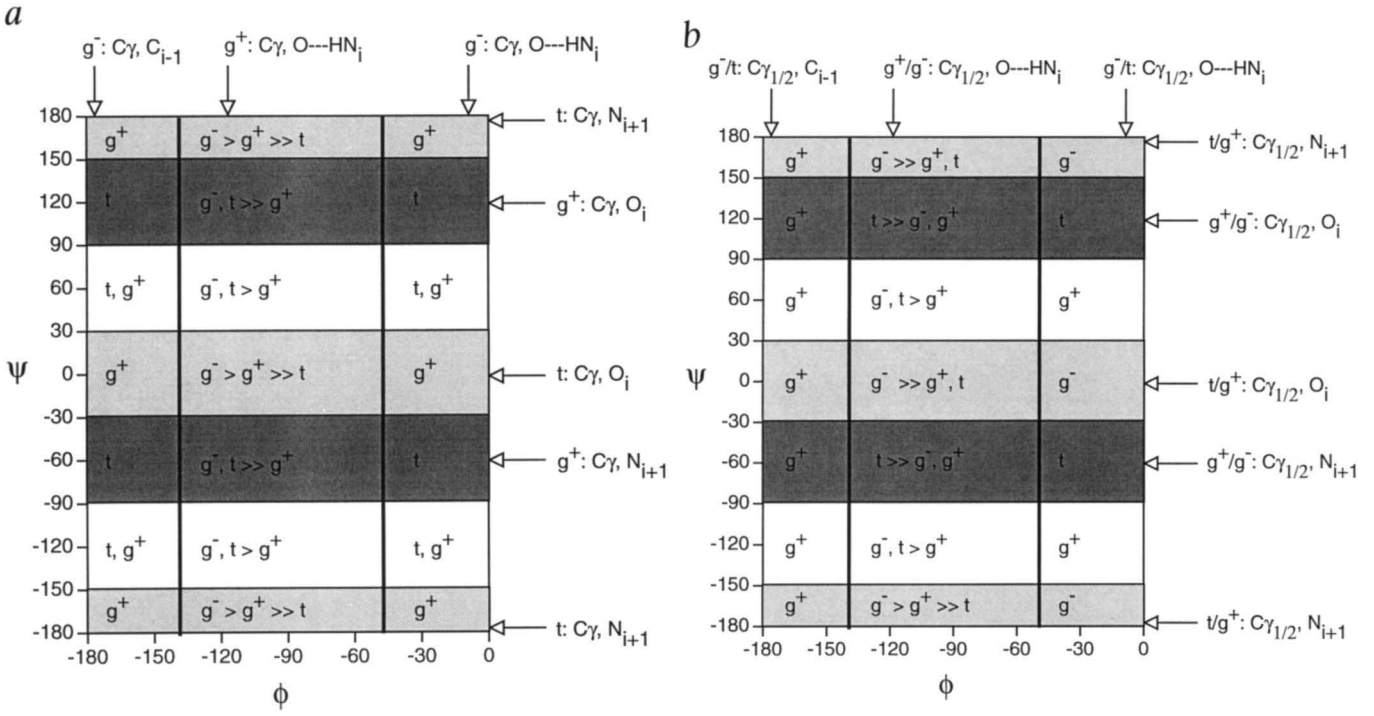


**Fig. 3** Conformational analysis of the backbone-dependence of protein rotamer preferences for *a*, sidechains with a single γ heavy atom (Arg, Lys, Met, Glu, Gln, Leu, Ser, Cys, Asp, Asn, Phe, Tyr, Trp, His) and *b*, sidechains with two γ heavy atoms at $\chi_1$ and $\chi_1$+120° (that is Val; for Ile and Thr with heavy atoms at $\chi_1$ and $\chi_1$-120°, the dihedral angles of $g^+$, t, $g^-$ of Val correspond to t, $g^-$, $g^+$ of Ile and Thr (see Methods)). The *syn*-pentane interactions between $C_{i-1}$, O---$HN_i$, $O_i$, and $N_{i+1}$ and sidechain $\chi_1$ rotamers are marked by arrows on the Ramachandran plots at the central $\phi$ and $\psi$ values at which they are expected to occur; that is, when the dihedral pairs connecting sidechain and backbone heavy atoms are equal to {+60°,-60°} or {-60°,+60°}. These interactions occur in a range of $\phi$ and $\psi$ of ±30° to ±50° about the values marked. It is expected that rotamers involved in *syn*-pentane interactions with backbone atoms will be rare in the $\phi$ and $\psi$ ranges in which the interactions occur. The remaining rotamers will therefore predominate, and these are listed in each block of the Ramachandran maps in the figure. Because of backbone-conformation independent interactions between Cγ and backbone $N_i$ and $C_i$, we expect that $g^-$ rotamers will predominate over $g^+$ rotamers when neither rotamer type forms *syn*-pentane interactions with the backbone (for example, -140° < $\phi$ < -50°, 150° < $\psi$ < 180°). For Val in *b*, when *syn*-pentane interactions occur for all three rotamers (for example, $\phi$ < -140° and $\phi$ > -50°), interactions with $C_{i-1}$ were considered to be larger (more unfavorable) than interactions with $N_{i+1}$ and $O_i$, which in turn were treated as larger than interactions with the hydrogen bond acceptor, O---$HN_i$.

© 1994 Nature Publishing Group http://www.nature.com/nsmb



map, they are the most easily predicted from the values of $\phi$ and $\psi$: 81% of Val, 85% of Ile, and 80% of Thr conformations can be predicted correctly from $\phi$ and $\psi$ and the backbone-dependent rotamer library.

### Interactions determining $\{\chi_1,\chi_2\}$

The *syn*-pentane effect also has a strong influence on which $\{\chi_1,\chi_2\}$ rotamer combinations are allowed and which are sterically hindered. These effects can occur between both backbone $N_i$ and $C_i$ and any Cδ atom of a sidechain, and are independent of the backbone dihedral angles. In Fig. 2*b*, conformations with *syn*-pentane interactions are those where the δ atoms come closest to the backbone atoms $N_i$ and $C_i$. For linear sidechains and Ile, the forbidden rotamers occur when $\{\chi_1,\chi_2\}$ = {60°,60°}, {60°,-60°}, {180°,-60°}, and {-60°,60°} (See Methods). These four rotamers are rare in the database (totaling 6% of Met, 13% of Glu, 9% of Gln, 5% of Arg, 7% of Lys, and 5% of Ile sidechains). For sidechain conformations with *syn*-pentane interactions, one or both of the experimental $\chi$ values are pushed away from canonical values by 10-30° as in pentane itself (Fig. 1*d*). For example, in the present database the $\{g^-,g^+\}$ rotamer of Lys has average $\{\chi_1,\chi_2\}$ values of {-87°,76°} compared to the $\{g^-,g^-\}$ rotamer averages of {-62°,-71°}. For the other *syn*-pentane rotamers, $\{g^+,g^+\}$, $\{g^+,g^-\}$, and $\{t,g^-\}$, the average dihedrals for Lys are {60°,78°}, {50°,-81°}, and {-157°,-81°} respectively. These deviations from the canonical values caused by *syn*-pentane interactions explain some of the observations of 'new' non-ideal rotamers found by Schrauber *et al.*[9].

Leucine is even more hindered than the single Cδ sidechains, because *syn*-pentane interactions of Cδ1 and Cδ2 (at $\chi_2$+120°) with backbone $N_i$ and $C_i$ eliminate conformers with $\{\chi_1,\chi_2\}$ = {60°,60°}, {60°,180°}, {60°,-60°},

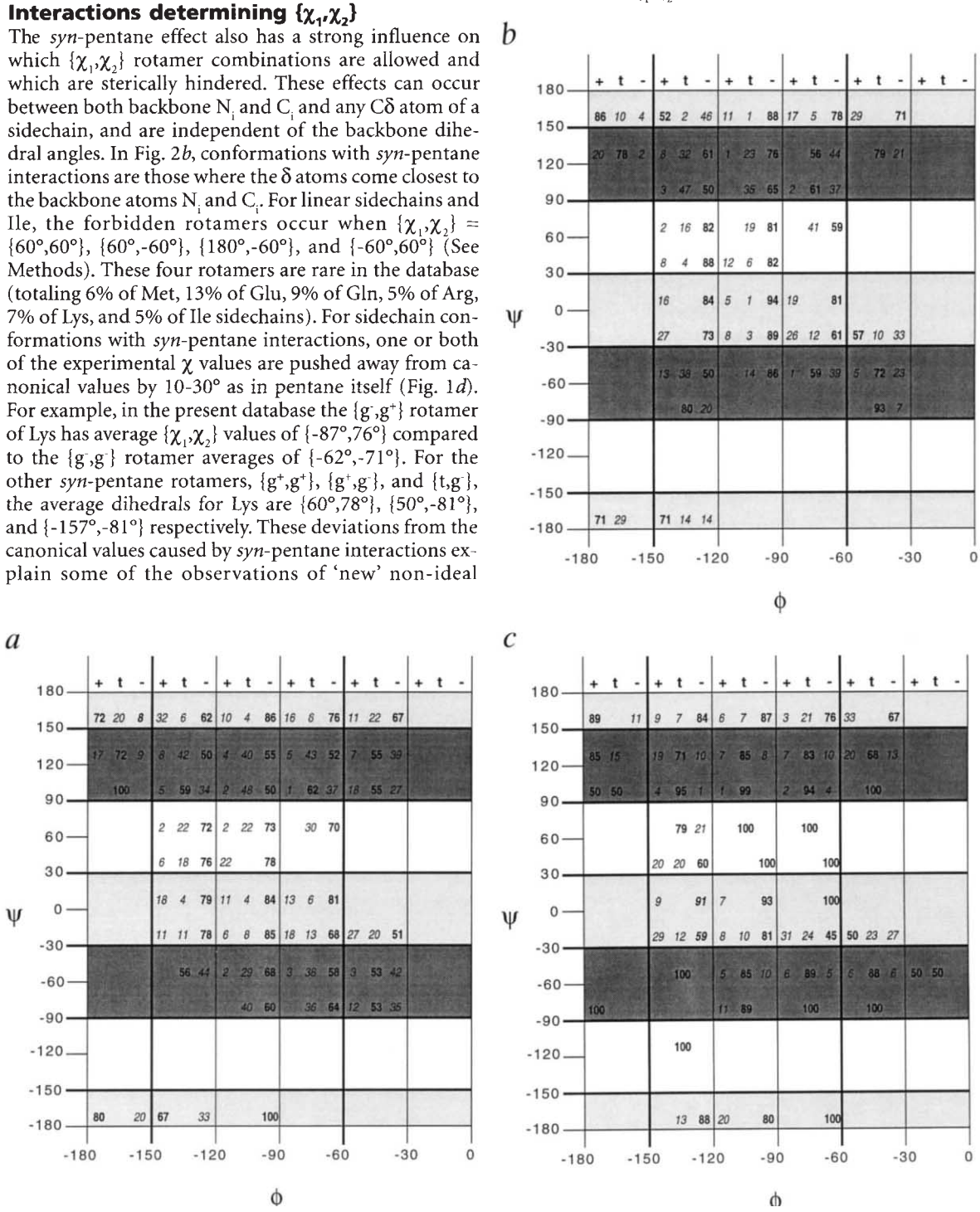


**Fig. 4** Backbone-dependent $\chi_1$ rotamer library for *a*, straight side chains Lys, Arg, Met, Glu, and Gln; *b*, aromatic sidechains Phe, Tyr, His, and Trp; *c*, Val and Ile. Each group of three numbers in a cell represent the percentages of $g^+$, t, and $g^-$ rotamers for $\chi_1$ within the range of $\{\phi,\psi\}$ values marked on the axes. Because of the definition of $\chi_1$ of Ile relative to Val, the values of $\chi_1$-120° were used for Ile (that is, for Ile, '+' in this figure refers to $\chi_1$=180°±60°; t refers to $\chi_1$=-60°±60°, and '-' refers to $\chi_1$=+60°±60°).

© 1994 Nature Publishing Group http://www.nature.com/nsmb



{180°,180°}, {180°,-60°}, {-60°,60°}, and {-60°,-60°}, independent of the backbone conformation. The remaining two rotamer combinations, {180°,60°} and {-60°,180°} represent 27% and 55% of leucines in the present database. *Syn*-pentane interactions effectively eliminate $g^+\chi_1$ rotamers for leucine; less than 2% of leucines have $\chi_1$ with a value of +60°±60°.

For aromatics, the *syn*-pentane effect of the Cδ1 and Cδ2 atoms with the backbone alters the average $\chi_2$ value depending on $\chi_1$. When $\chi_1$ is near -60°, there is a *syn*-pentane interaction for $\chi_2$ near +60°. Consequently, $\chi_2$ is skewed toward an average of 99° for $g^-$ rotamers of Phe in the present database, rather than the value of 90° expected from butane-type interactions. When $\chi_1$ is 180°, a *syn*-pentane interaction between Cδ2 and $C_i$ occurs when $\chi_2$ is 120° (or Cδ1 when $\chi_2$ is -60°). Hence, $\chi_2$ is skewed in the opposite direction, and the average $\chi_2$ for trans $\chi_1$ rotamers is 78°. When $\chi_1$ is 60°, *syn*-pentane interactions can occur when $\chi_2$ is 60° or 120°, and the average $\chi_2$ is near 90° in the backbone-independent rotamer library. The other aromatics, Tyr, His, and Trp, show similar $\chi_2$ deviations, dependent on $\chi_1$.

The above results make clear that the dominant features of sidechain rotamer populations can be understood by simple arguments based on conformational analysis. In individual cases, particularly for highly polar or charged sidechains, additional interactions (for example, electrostatic effects) may perturb the sidechain orientation; for such cases and more quantitative results, detailed calculations with empirical energy functions[30] are required. Steric interactions involving secondary and tertiary structural elements (that is, packing) can also play a role in selecting one rotamer conformation if more than one is allowed as a result of the present analysis[11,31].

## Methods

**Conformational analysis of backbone-dependent rotamer preferences.** The analysis of backbone-dependent rotamers was performed by compiling a list of *syn*-pentane interactions of sidechain γ heavy atoms of residue i with backbone atoms $C_{i-1}$, $O_i$, and $N_{i+1}$, as well as the hydrogen bond acceptor to $HN_i$ ('O---$HN_i$'). *Syn*-pentane interactions occur when the two dihedrals connecting heavy atoms separated by 4 chemical bonds have values of approximately {-60°,+60°} or {+60°,-60°}. The positions of backbone atoms $C_{i-1}$, $O_i$, O---$HN_i$, and $N_{i+1}$ relative to the *i*th sidechain are dependent on the backbone dihedral angles φ and ψ and the sidechain dihedral angle $\chi_1$. The dihedral angles connecting $C_{i-1}$ and Xγ are $C_{i-1}$-$N_i$-Cα-Cβ (~φ–120°)and $N_i$-Cα-Cβ-Xγ ($\chi_1$). The dihedral angles connecting O---$HN_i$ and Xγ are $H_i$-$N_i$-Cα-Cβ (~φ+60°) and $\chi_1$ (assuming a linear hydrogen bond). The dihedral angles connecting backbone $O_i$ and Xγ of the same residue are $O_i$-$C_i$-Cα-Cβ (~ψ-60°) and $C_i$-Cα-Cβ-Xγ (~$\chi_1$-120°). The dihedral angles connecting backbone $N_{i+1}$ and Xγ are $N_{i+1}$-$C_i$-Cα-Cβ (~ψ+120°) and $C_i$-Cα-Cβ-Xγ (~$\chi_1$-120°). In the cases of Thr, Ile, and Val, the interactions of these backbone atoms with the Cγ2 atoms (at $\chi_1$+120° for Val and at $\chi_1$-120° for Ile and Thr) must also be considered.

**Conformational analysis of backbone-independent rotamer preferences.** For the analysis of backbone conformation independent interactions, the relevant dihedrals are $N_i$-Cα-Cβ-Cγ ($\chi_1$) and $C_i$-Cα-Cβ-Cγ ($\chi_1$-120°), each in combination with Cα-Cβ-Cγ-Xδ ($\chi_2$). That is, when {$\chi_1$,$\chi_2$} or {$\chi_1$-120°, $\chi_2$} is equal to {60°,-60°} or {-60°,60°} we expect a strong steric interaction to occur, and a particular rotamer to be rare. For straight chain hydrocarbons, this occurs for four different rotamer combinations, leaving only five {$\chi_1$,$\chi_2$} combinations unaffected. Leucine has Cδ2 at $\chi_2$+120°, and so in addition to {$\chi_1$,$\chi_2$} and {$\chi_1$-120°,$\chi_2$}, the dihedral pairs {$\chi_1$,$\chi_2$+120°} and {$\chi_1$-120°,$\chi_2$+120°} produce steric interactions when equal to {60°,-60°} or {-60°,60°}. This is a total of 8 *syn*-pentane interactions, two of which occur for {$\chi_1$,$\chi_2$}={60°,-60°}, leaving only 2 out of 9 conformations without significant clashes with the backbone. For aromatics, Cδ2 occurs at $\chi_2$+180°, and the combinations of dihedrals that must be checked can be calculated as above.

**Rotamer library.** The rotamer library was determined from the structures of 182 protein chains from 173 entries in the Brookhaven Protein Database at a resolution better than or equal to 2.0 Å. The library is somewhat larger than that used in our previous work[8]. $\chi_1$ rotamers were defined according to the following limits: $g^+$ ($0° \le \chi_1 < 120°$); t ($120° \le \chi_1$ 0°); and $g^-$(-120° $\le \chi_1 < 0°$). Note that some authors use reversed definitions of $g^+$ and $g^-$. The library data in Fig. 4 are shown in 30° ranges of φ and ψ. For the purposes of sidechain conformation prediction, the library was calculated in 20° ranges, every 10° for each sidechain type individually. In the present library, there are 2021 Lys, 1254 Arg, 621 Met, 1863 Glu, 1221 Gln, 1320 Phe, 1287 Tyr, 708 His, 445 Trp, 1798 Ile, and 2449 Val sidechains.

*Note:* The backbone-dependent rotamer library and a program for its use in sidechain conformation prediction is available by e-mail from dunbrack@cgl.ucsf.edu.

Received 10 January; accepted 29 March 1994.

© 1994 Nature Publishing Group http://www.nature.com/nsmb



**Acknowledgements**
We wish to thank S.L. Schreiber for discussions on conformational analysis and J.M. Thornton for suggesting the importance of hydrogen bonding effects. This work was supported in part by a grant from the National Institutes of Health and a gift from Molecular Simulations, Inc.

**1.** Chandrasekaran, R. & Ramachandran, G.N. Studies on the conformation of amino acids. XI. Analysis of the observed side group conformations in proteins. *Int. J. prot. Res.* **2**, 223–233 (1970).
**2.** Janin, J., Wodak, S., Levitt, M. & Maigret, B. Conformations of amino acid sidechains in proteins. *J. molec. Biol.* **125**, 357–386 (1978).
**3.** Ponder, J.W. & Richards, F.M. Tertiary templates for proteins: Use of packing criteria in the enumeration of allowed sequences for different structural classes. *J. molec. Biol.* **193**, 775–792 (1987).
**4.** Bhat, T.N., Sasisekharan, V. & Vijayan, M. An analysis of side-chain conformation in proteins. *Int. J. Peptide Protein Res.* **13**, 170–184 (1979).
**5.** Benedetti, E., Morelli, G., Nemethy, G. & Scheraga, H.A. Statistical and energetic analysis of side-chain conformations in oligopeptides. *Int. J. Peptide Protein Res.* **22**, 1–15 (1983).
**6.** Tuffery, P., Etchebest, C., Hazout, S. & Lavery, R.A new approach to the rapid determination of protein side chain conformations. *J. biomolec. Struct. Dyn.* **8**, 1267–1289 (1991).
**7.** McGregor, M.J., Islam, S.A. & Sternberg, M.J.E. Analysis of the relationship between sidechain conformation and secondary structure in globular proteins. *J. molec. Biol.* **198**, 295–310 (1987).
**8.** Dunbrack, R.L., Jr & Karplus, M. Backbone-dependent rotamer library for proteins: Application to sidechain prediction. *J. molec. Biol.* **230**, 543–571 (1993).
**9.** Schrauber, H., Eisenhaber, F. & Argos, P. Rotamers: To be or not to be? An analysis of amino acid sidechain conformations in globular proteins. *J. molec. Biol.* **230**, 592–612 (1993).
**10.** Dunbrack, R.L., Jr *Conformational analysis of protein sidechains: Empirical energy parameters for proline and development of a backbone-dependent rotamer library* (Ph.D. Thesis, Harvard University, 1993).
**11.** Eisenmenger, F., Argos, P. & Abagyan, R. A method to configure protein sidechains from the mainchain trace in homology modeling. *J. molec. Biol.* **231**, 849–860 (1993).
**12.** Lee, C. & Subbiah, S. Prediction of protein sidechain conformation by packing optimization. *J. molec. Biol.* **217**, 373–388 (1991).
**13.** Desmet, J., DeMaeyer, M., Hazes, B. & Lasters, I. The dead-end elimination theorem and its use in protein sidechain positioning. *Nature* **356**, 539–542 (1992).
**14.** Becker, F. Theoretische Behandlung des Einflusses sterischer Effekte auf die Reaktivität aliphatischer Verbindungen I. *Z. Naturforsch.* **14a**, 547–556 (1959).
**15.** Eliel, E.L., Allinger, N.A., Angyal, S.J. & Morrison, G.A. *Conformational Analysis* (Interscience, New York, 1965).
**16.** Ramachandran, G.N., Ramakrishnan, C., and Sasisekharan, V.J. Stereochemistry of polypeptide chain configurations, *J. molec. Biol.* **7**, 95–99 (1963).
**17.** Gelin, B.R. & Karplus, M. Sidechain torsional potentials: Effect of dipeptide, protein, and solvent environment. *Biochemistry* **18**, 1256–1268 (1979).
**18.** Ponnuswamy, P.K. & Sasisekharan, V. Studies on the conformation of amino acids. IX. Conformations of butyl, seryl, threonyl, cysteinyl, and valyl residues in a dipeptide unit. *Biopolymers* **10**, 565–582 (1971).
**19.** Lewis, P.N., Momany, F.A. & Scheraga, H.A. Energy parameters in polypeptides. VI. Conformational energy analysis of the N-Acetyl N'-methyl amides of the twenty naturally occurring amino acids. *Israel. J. Chem.* **11**, 121–152 (1973).
**20.** Zimmerman, S.S., Pottle, M.S., Nemethy, G. & Scheraga, H.A. Conformational analysis of the 20 naturally occurring amino acid residues using ECEPP. *Macromolecules* **10**, 1–9 (1977).
**21.** Compton, D.A.C., Montero, S. & Murphy, W.F. Low-frequency Raman spectrum and asymmetric potential function for internal rotation of gaseous *n*-butane. *J. phys. Chem.* **84**, 3587–3591 (1980).
**22.** Pitzer, K.S. The vibration frequencies and thermodynamic functions of long chain hydrocarbons. *J. chem. Phys.* **8**, 711–720 (1940).
**23.** Wiberg, K.B. & Murcko, M.A. Rotational barriers. 2. Energies of alkane rotamers. An examination of gauche interactions. *J. Am. chem. Soc.* **110**, 8029–8038 (1988).
**24.** Hoeve, C.A.J. Unperturbed mean-square end-to-end distance in polyethylene. *J. chem. Phys.* **35**, 1266–1267 (1961).
**25.** Abe, A., Jernigan, R.L. & Flory, P.J. Conformational energies of *n*-alkanes and the random configuration of higher homologs including polymethylene. *J. Am. chem. Soc.* **88**, 631–639 (1966).
**26.** Pitzer, K.S. Chemical equilibria, free energies, and heat contents for gaseous hydrocarbons. *Chem. Rev.* **27**, 39–57 (1940).
**27.** Pitzer, R.M. The barrier to internal rotation in ethane. *Accts. chem. Res.* **16**, 201–210 (1983).
**28.** Newman, M.S. A notation for the study of certain stereochemical problems. *J. chem. Ed.* **32**, 344–347 (1955).
**29.** Gray, T.M. & Matthews, B.W. Intrahelical hydrogen bonding of serine, threonine, and cysteine residues with α-helices and its relevance to membrane-bound proteins. *J. molec. Biol.* **175**, 75–81 (1984).
**30.** Brooks, B.R. *et al.* CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. comp. Chem.* **4**, 187–217 (1983).
**31.** Richards, F.M. & Lim, W.A. An analysis of packing in the protein folding problem. *Q. Rev. Biophys.* (in the press).
**32.** Ferrin, T.E. *et al.* The MIDAS display system. *J. molec. Graphics* **6**, 13–27 (1988).