e. Virtual bond model provides an accurate description of the conformational preferences of the backbone

In low resolution approaches, it is important to preserve as much as possible of the unique, diverse characteristics of different residues, while removing the atomic details. The *virtual bond approximation* yields an almost unequivocal description of the backbone conformation, and a reasonable account of residue specificities. As a matter of fact, the distribution of bond angles and torsional angles in the virtual bond model are highly correlated with the local secondary structure, and thus reflect the secondary structure propensities of individual amino acids. See § VII.x.

A total of 2n variables define the backbone geometry for the virtual bond model: the dihedral angle (φ_i) of each virtual bond, and the angle (ϑ_i) between successive dihedral angles. The virtual bond lengths are almost fixed at 3.81 Å, except for *cis* peptide bonds. The angles ϑ_i and φ_i follow interdependent probability distributions specific to each type of residue. *Figure II.1.11* displays the joint distribution of φ_i and φ_{i+1}. Two regions are highly populated, located near $(\varphi_i, \varphi_{i+1}) \approx (50°, 50°)$ and $\approx (200°, 200°)$, characteristic of α-helices, and β-sheet structures, respectively

The distributions $N_A(\varphi_i, \varphi_{i+1})$ for particular amino are illustrated in *Figure II.1.12*. The strong preference of Glu, for example, for α-helices, the versatility of Gly to adopt various torsional states, the aptitude of Val for participating in β-sheets, or the unique rotational preferences of Pro are distinguishable. The conformational energies extracted from such distributions will be shown in § VII to be useful for estimating the secondary structure propensities of particular amino acids.

Likewise, the bond angle ϑ_i is strongly correlated with the torsions φ_i and φ_{i+1} of the adjoining bonds, as illustrated in *Figure II.1.13*. Bond angles of about 90° and 120° are favored in α-helices and β-sheets, respectively. The distributions $N_A(\vartheta_i, \varphi_{i+1})$ and $N_A(\vartheta_i, \varphi)$ for specific residues (not shown) also exhibit residue-specific characteristics {Bahar, Kaplan, et al. 1997 ID: 92}.

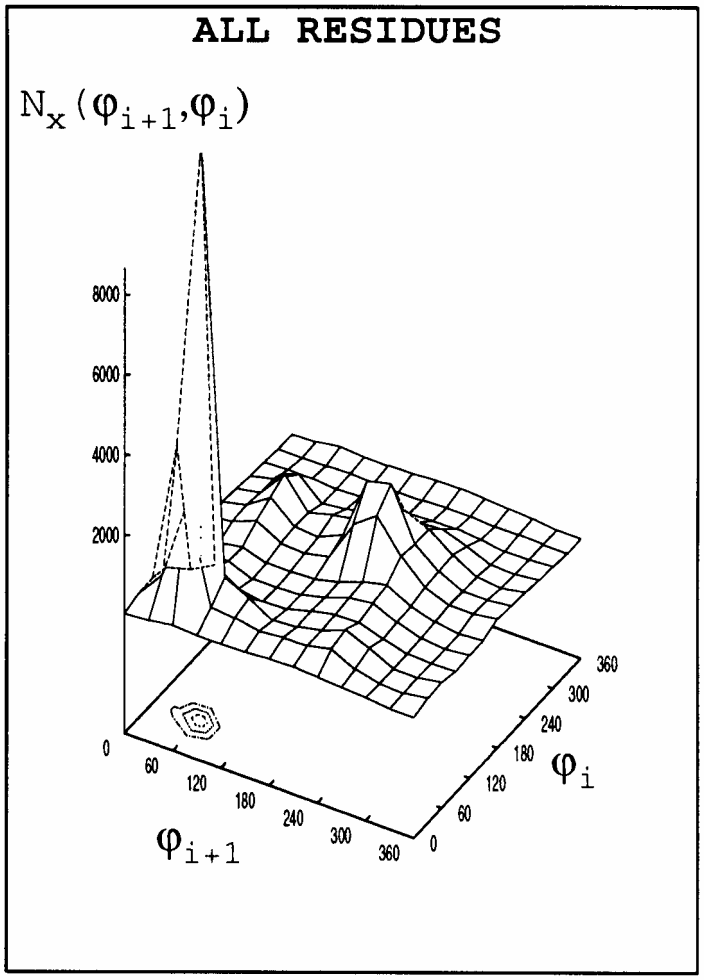

Figure II.1.11. *Distribution of adjacent virtual bonds' dihedral angles $N_X(\varphi_i, \varphi_{i+1})$. The horizontal axes are the torsion angles φ_i and φ_{i+1}, and the surface represents the number of occurrence of each region of size ($\varphi_i \pm 15°$, $\varphi_{i+1} \pm 15°$) in 150 PDB structures. The projection of the surface is shown on the lower plane. The most populated dihedral angle pairs are enclosed by the innermost contours. The sharp peak near (φ_i, φ_{i+1}) = (50°, 50°) corresponds to α-helices, and the broader peak near (φ_i, (200°, 200°) is associated with β-sheets. (from {Bahar, Kaplan, et al. 1997 ID: 92})*

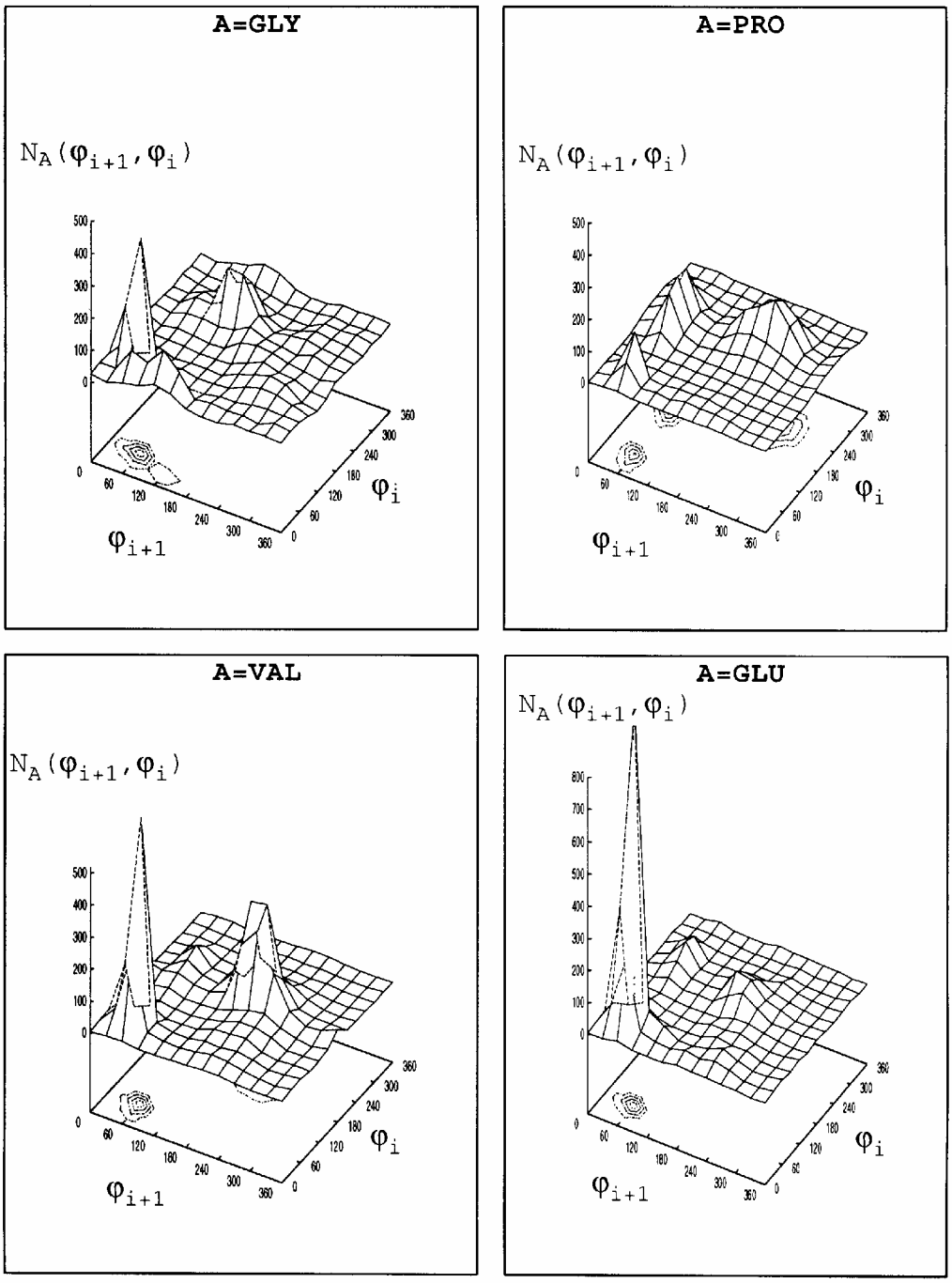

Figure II.1.12. *Number distribution of virtual bond torsions $N_A(\varphi_i, \varphi_{i+1})$ for A = Gly, Glu, Val and Pro. (taken from {Bahar, Kaplan, et al. 1997 ID: 92})*

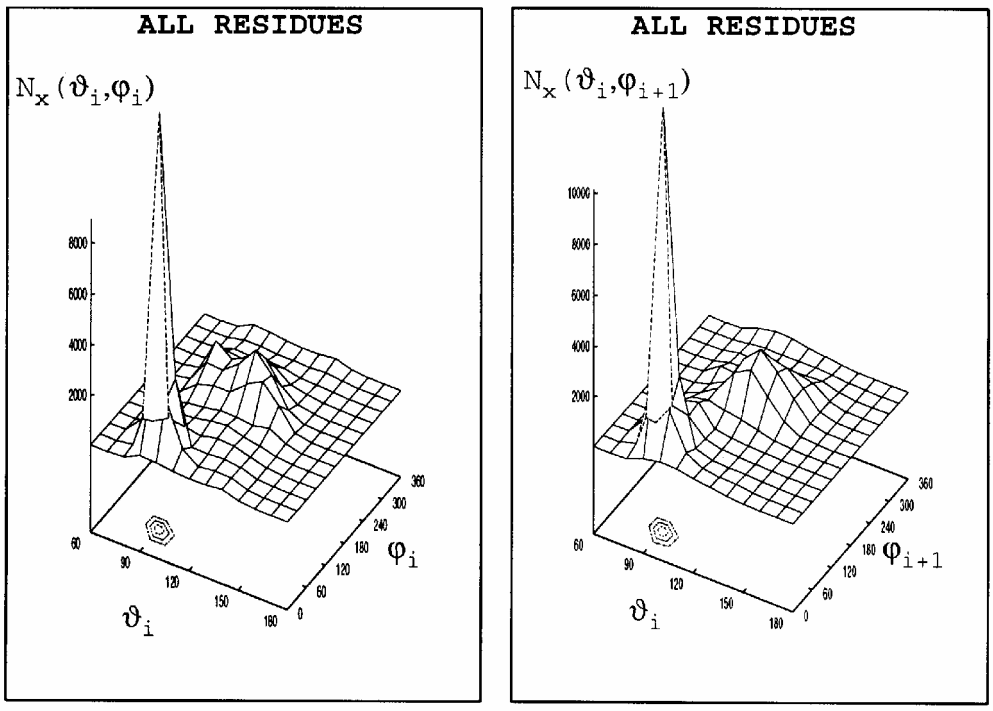

Figure II.1.13. *Coupling between virtual bond torsion and bending angles. The surfaces represent the number distributions from 150 PDB structures, irrespective of residue type.* ***(a)*** *Distribution* $N_X(\vartheta_i, \varphi_{i+1})$ *for any ith bond angle* ϑ_i *and the succeeding bond torsion* φ_{i+1} ***(b)*** *Distribution* $N_X(\vartheta_i, \varphi_i)$ *for any ith bond angle* ϑ_i *and the preceding bond rotation* φ_i. *Bond angles outside the displayed range* $60° < \vartheta_i < 180°$ are not displayed. *Grids of size 10° were taken in analyzing the bond angles (taken from {Bahar, Kaplan, et al. 1997 ID: 92})*

f. Amino acid sidechains prefer angles near their ideal rotational isomeric states

In addition to ϕ and ψ angles, proteins also have freedom in the side chain rotational angles. Moving along the side chain away from the backbone defines carbons identified as C^{β}, C^{γ}, etc., and rotational angles as χ_1, χ_2, etc.; see *Figures II.1.14 and 15*.

Hydrocarbon chains *[(-CH_2-)$_n$]* in the gas phase tend to populate the three rotational isomeric states *trans*, *gauche*$^+$ and *gauche*$^-$ (see *Figures II.1.5* and *6*). *Figures II.1.16* and *17* show that the side chains in proteins also tend to populate the same rotational angles, at least for the χ_1 angles. *Figure II.1.18* shows that this correspondence between side chain angles observed in proteins with their ideal values grows stronger as the structural quality of examined proteins increases. This is usually interpreted to mean that when proteins are known to very high resolution, side chain angles in globular proteins should coincide closely with the angles intrinsically favored by those bond types. The less optimistic interpretation is that protein side chain angles become ideal because computerized structural refinement methods are based on assuming such ideality, and protein structure refinements do in part reflect the artifacts of the refinement process.

But if it is true, observations of ϕ ψ, and χ angles in proteins indicate an important principle that holds at least to first approximation: the folding forces acting on proteins do not perturb the backbone and side chain angles very much relative to the intrinsic values that amino acids and dipeptides would have in the absence of the folding forces. Local factors - a side chain interacting with its own backbone, or a methylene group interacting with a neighboring methylene in a side chain - dictate the bond angle options that are available to a protein. A folding protein, chooses from among those options. See *Figures II.1.16* and *17*. Folding forces rarely distort or strain the angles the backbone and side chain bonds intrinsically prefer, but simply act to select among the intrinsically favorable forms.

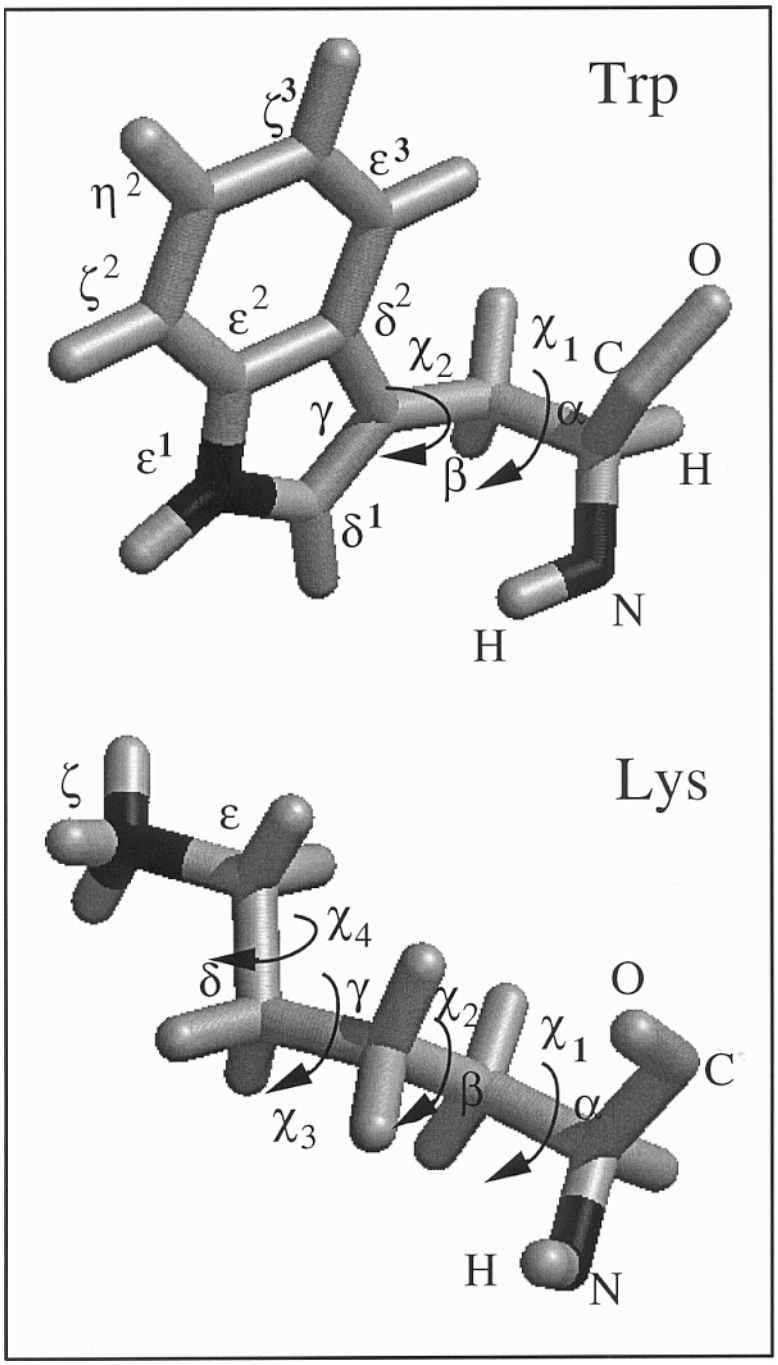

Figure II.1.14. *Two examples, Trp and Lys, illustrating the definition of the torsion angles χ_1, χ_2, etc. of amino acid side chains in proteins, along with the labels α, β, γ, δ, etc. assigned to successive atoms for distinguishing their position along the sidechain. Nitrogen atoms are shown in black. Backbone atoms are on the right end in each case.*

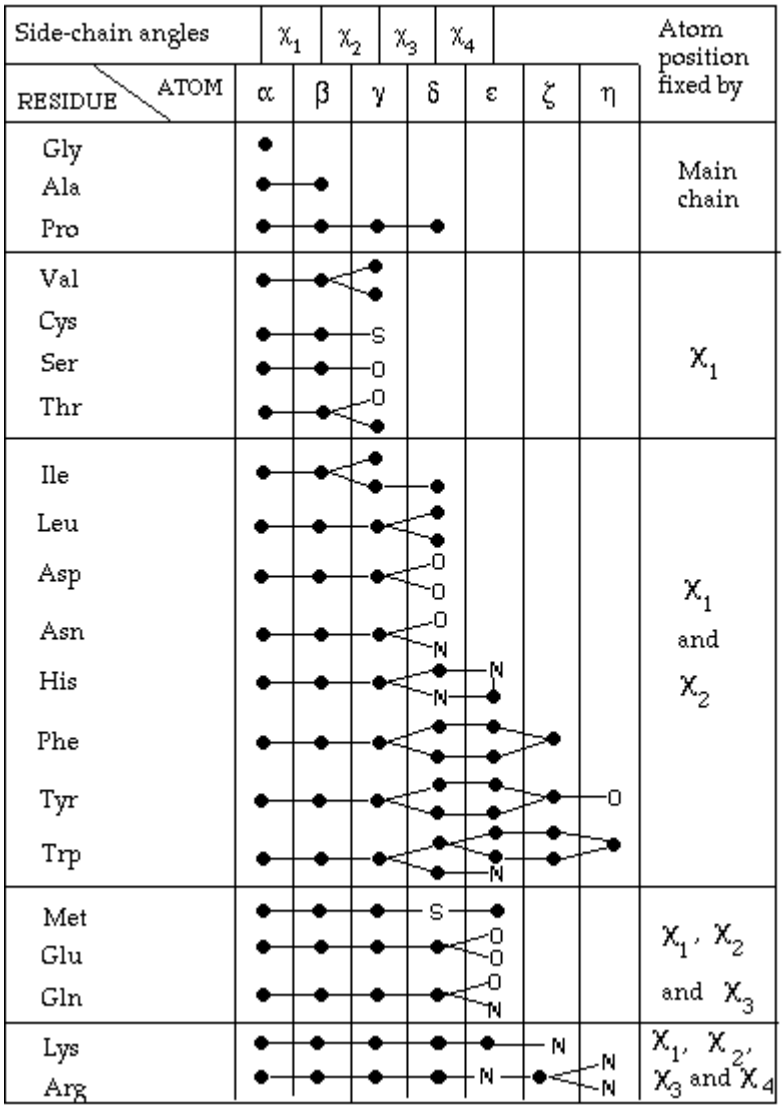

Figure II.1.15. *Definition of sidechain atoms' identification labels and corresponding torsional angles, displayed for all types of amino acid. The last column shows the torsional angles required to be defined in order to fix the positions of all sidechain atoms. See the diagram of {Ponder 2000 ID: 498}.*

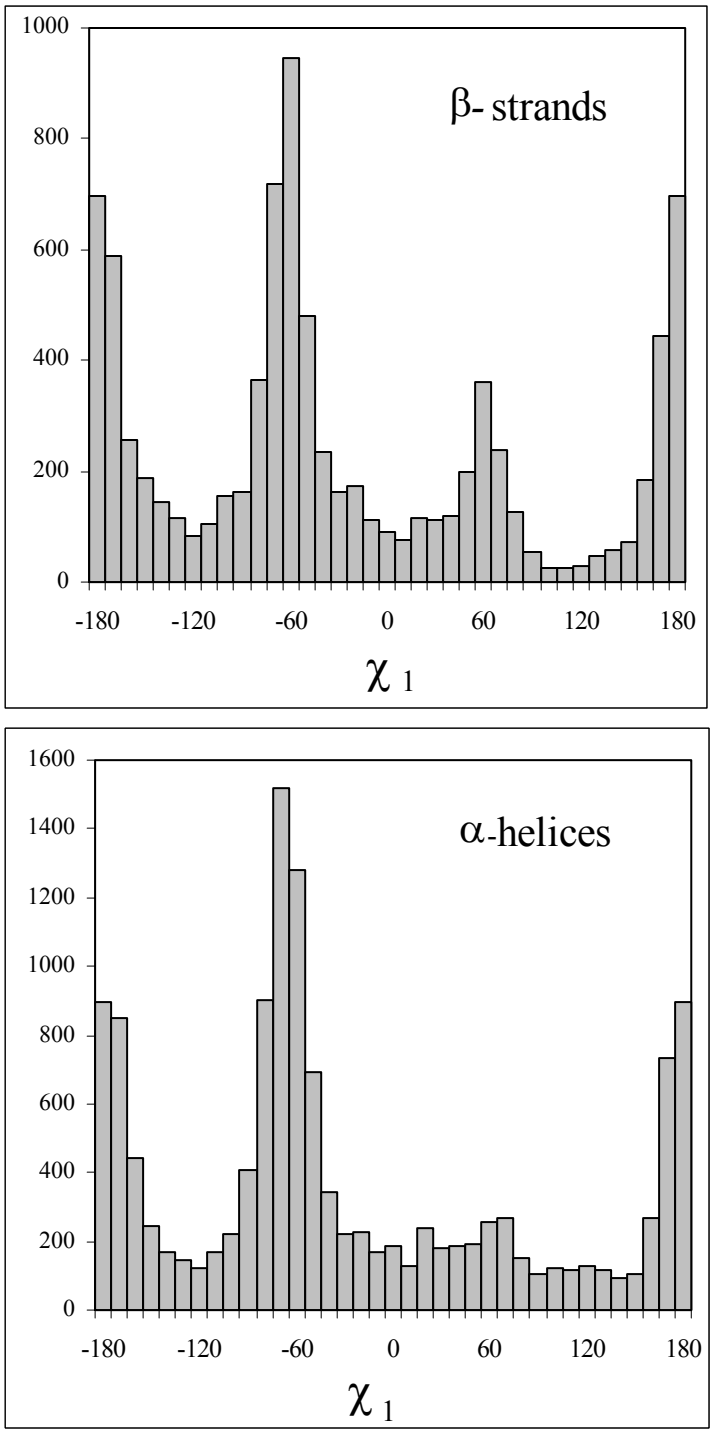

Figure II.1.16. *Distribution of the χ_1 sidechain torsion angles for amino acids in α-helical and β-sheet regions compiled from 150 high resolution protein structures. The ordinate represents the total number of observations for each interval of size $\Delta\chi_1 = 10^0$. Note the relatively low probability of the gauche$^-$ state ($\chi_1 = 60^o$) for side chains on α-helices.*

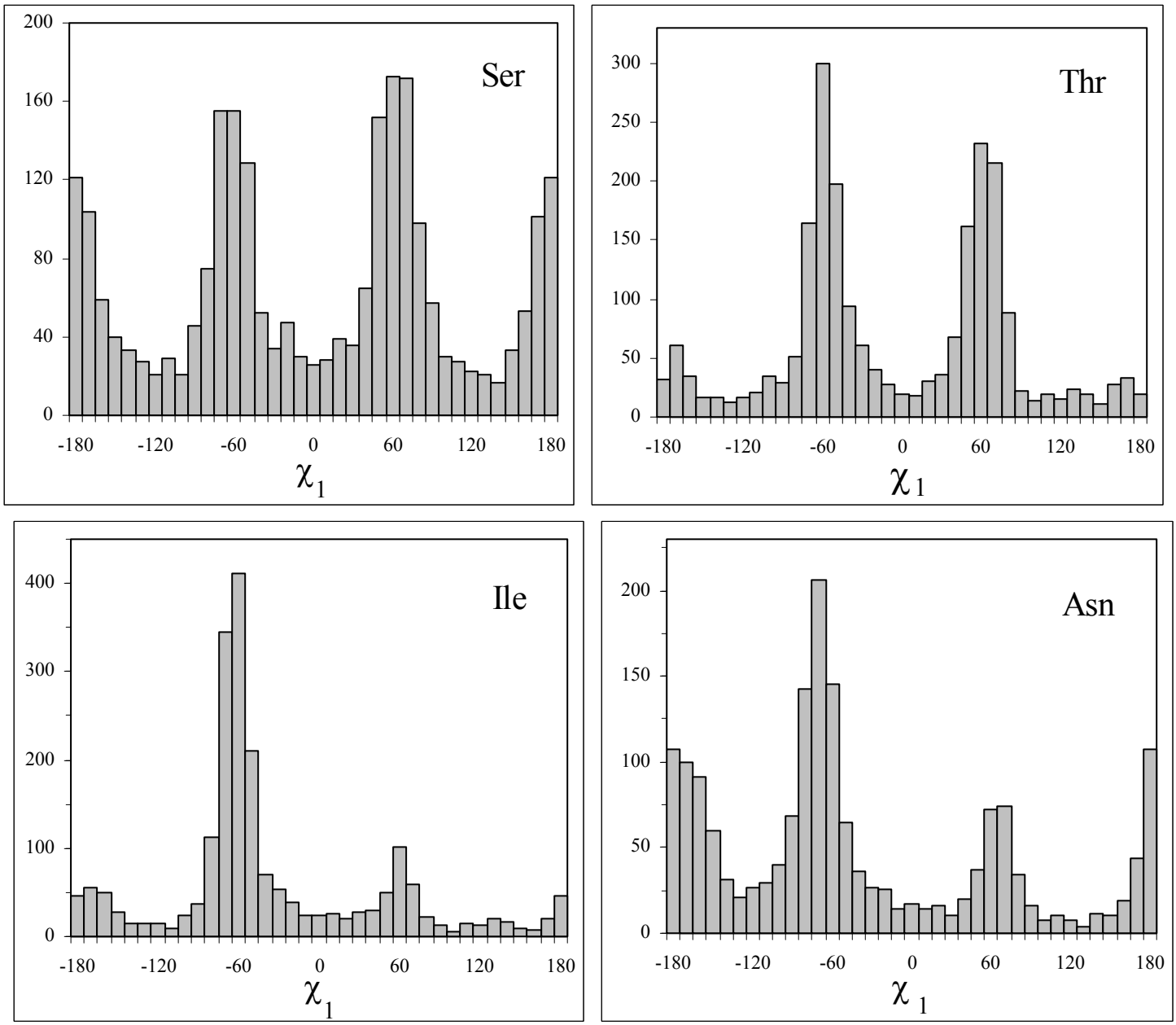

Figure II.1.17. *Comparison of the χ_1 sidechain torsion angles for four different types of amino acids, Ser, Thr, Ile, and Asn, compiled from 150 high resolution databank structures. The figure illustrates that the side chains choose from among the same options (about trans, gauche$^+$ and gauche$^-$ states) although the relative populations of these states differ depending on the type of amino acid.*

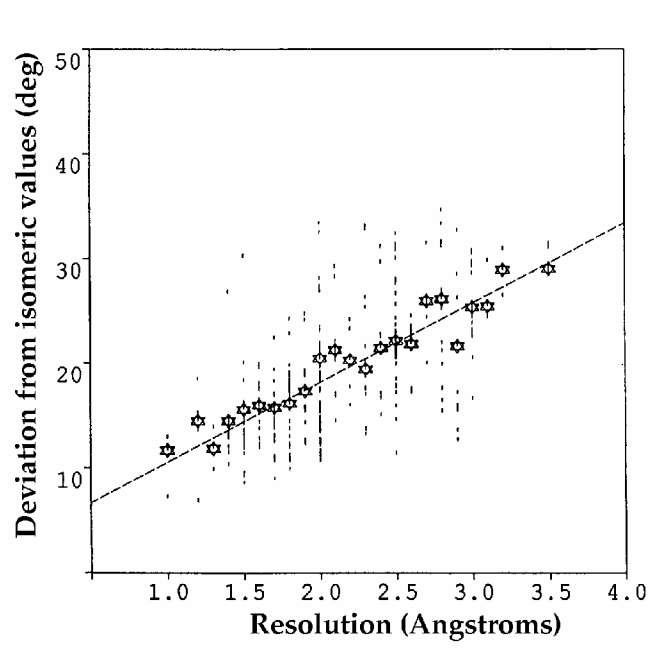

Figure II.1.18. *Deviations of χ_1 angles from the standard gauche$^-$, trans and gauche$^+$ rotamers plotted against the resolution of the experimentally determined protein native structure. Points are for individual proteins, stars are averages for a given resolution. (adapted from {Thornton 1992 ID: 496})*